Mutate & Conjugate: A Method for Rapid In-Cell Validation of Epigenetic Targets



Adam Matthew Thomas

Somerville College, University of Oxford

Thesis submitted for the degree of:

Doctor of Philosophy in Chemical Biology

Hilary Term 2022

Under the supervision of:

Professor Stuart J. Conway (Oxford)

Dr Jacob T. Bush (GlaxoSmithKline)

Department of Chemistry

DECLARATION OF AUTHORSHIP

Name: ADAM THOMAS **Candidate number: 1027374**

College: SOMERVILLE COLLEGE **Supervisor: PROF. STUART CONWAY**

Title of thesis: MUTATE & CONJUGATE: A METHOD FOR RAPID IN-CELL VALIDATION OF EPIGENETIC TARGETS

Please tick to confirm the following:

I have read and understood the University's disciplinary regulations concerning conduct in examinations and, in particular, the regulations on plagiarism (*The University Student Handbook* Section 8.7; available at https://www.ox.ac.uk/students/academic/student-handbook). ☑

I have read and understood the Education Committee's information and guidance on academic good practice and plagiarism at

https://www.ox.ac.uk/students/academic/guidance/skills?wssl=1. ☑

The [thesis/dissertation/extended essay/assignment/project/other submitted work] I am submitting is entirely my own work except where otherwise indicated. ☑

It has not been submitted, either partially or in full, either for this Honour School or qualification or for another Honour School or qualification of this University (except where the Special Regulations for the subject permit this), or for a qualification at any other institution. ☑

I have clearly indicated the presence of all material I have quoted from other sources, including any diagrams, charts, tables or graphs. ☑

I have clearly indicated the presence of all paraphrased material with appropriate references. ☑

I have acknowledged appropriately any assistance I have received in addition to that provided by my [tutor/supervisor/adviser]. ☑

I have not copied from the work of any other candidate. ☑

I have not used the services of any agency providing specimen, model or ghostwritten work in the preparation of this thesis/dissertation/extended essay/assignment/project/other submitted work. (See also section 2.4 of Statute XI on University Discipline under which members of the University are prohibited from providing material of this nature for candidates in examinations at this University or elsewhere: http://www.admin.ox.ac.uk/statutes/352-051a.shtml). ☑

I agree to retain an electronic copy of this work until the publication of my final examination result, except where submission in hand-written format is permitted. ☑

I agree to make any such electronic copy available to the examiners should it be necessary to confirm my word count or to check for plagiarism. ☑

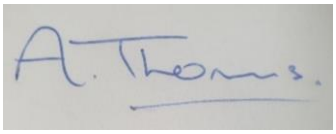
Candidate's signature:  Date: 22nd April 2022

Table of Contents

Abstract ..... **i**

Acknowledgements ..... **ii**

Abbreviations ..... **v**

Chapter 1: Introduction **1**

1.1 Target validation **2**

1.1.1 Target validation in drug discovery **2**

1.1.2 A biological approach to target validation **5**

1.1.3 A chemical approach to target validation **6**

1.2 Covalent ligands and probes **8**

1.2.1 Targeting cysteine residues **10**

1.2.1.1 Disulfides *10*

1.2.1.2 Acrylamides *12*

1.2.1.3 Haloacetamides *16*

1.2.1.4 α,β-Unsaturated esters *18*

1.2.2 Targeting lysine residues **19**

1.2.2.1 Sulfonyl fluorides *20*

1.2.2.2 Activated esters *21*

1.2.3 Targeting other residues **23**

1.2.3.1 Isoxazolium salts for targeting glutamic acid *23*

1.2.3.2 Nitrile oxides for targeting selenocysteine *24*

1.2.4 Clickable probes for target engagement **26**

1.2.4.1 Target identification *27*

1.2.4.2 Bioorthogonal Click Chemistry *28*

1.2.4.3 Photoaffinity labelling *29*

1.2.4.4 Recent Examples *30*

1.3 Epigenetics **32**

1.4 Bromodomains **33**

1.4.1 Bromodomain-containing protein 4 (BRD4) **34**

1.4.1.1 Structure and function *34*

1.4.1.2 Small molecule ligands for BRD4 ..... *34*

1.5 Project Aims **41**

Chapter 2: Design, expression, and structural validation of a cysteine mutant bromodomain **42**

2.1 Introduction and aims 43
2.2 Mutant protein design 44
2.2.1 PyMOL mutagenesis 46
2.2.2 Molecular dynamics simulations 49
2.3 Mutant protein expression and purification 50
2.4 Structural validation 52
2.4.1 Biophysical assays 52
2.4.2 Crystallographic studies 54
2.5 Conclusions 56
Chapter 3: Identification of reactive fragments that covalently bind to BRD4(1)^{L94C} 58
3.1 Introduction and aims 59
3.2 Structure-guided covalent fragment design 59
3.3 Structure-guided covalent fragment synthesis 62
3.3.1 Ketone-linked fragments 62
3.3.2 Amide-linked fragments 69
3.4 Evaluation of purified BRD4(1)^{L94C} labelling using protein mass spectrometry 73
3.4.1 In-house screen 73
3.4.2 Diverse reactive fragment screening platform 78
3.5 Conclusions 82
Chapter 4: Optimisation of reactive fragments targeting mutant BRD4(1)^{L94C} 84
4.1 Introduction and aims 85
4.2 Design of optimised covalent fragments 86
4.3 Optimised covalent fragment synthesis 88
4.3.1 Ketone-linked fragment 88
4.3.2 Amide-linked fragments 89
4.3.2.1 Synthesis of the amine-containing scaffold 89
4.3.2.2 Installation of electrophilic groups 95
4.4 Biological evaluation 102
4.4.1 Evaluation of purified BRD4(1)^{L94C} labelling using protein mass spectrometry 102
4.4.2 AlphaScreen™ assay 104
4.5 Conclusions 106
Chapter 5: Clickable probes to evaluate in-cell selectivity of reactive fragments targeting BRD4(1)^{L94C} 107
5.1 Introduction and aims 108
5.2 Click probe design 108
5.2.1 Isoxazole scaffold 109

5.2.1 γ-Lactam scaffold 110
5.3 Click probe synthesis 112
5.3.1 Isoxazole scaffold 112
*5.3.1.1 Synthesis of **5.1** and **5.4** 112*
*5.3.1.2 Synthesis of **5.2**, **5.3**, **5.5**, and **5.6** 115*
5.3.2 γ-Lactam scaffold 120
5.4 Click probe validation 122
5.4.1 Click efficiency with TAMRA-PEG3-azide 122
5.4.2 Fluorescent labelling of purified BRD4(1)^{L94C} 125
5.4.3 Evaluation of purified BRD4(1)^{L94C} labelling using protein mass spectrometry 127
5.5 In-cell target engagement 129
5.5.1 Plasmid design 129
5.5.2 Validation of protein expression 130
5.5.3 In-cell protein labelling 131
5.5.3.1 In-cell protein labelling by isoxazole-based click probes 132
5.5.3.2 In-cell protein labelling by γ-lactam-based click probes 135
5.6 Conclusions 137
Chapter 6: Investigating the applicability of 'mutate & conjugate' to the development of protein degraders 138
6.1 Introduction and aims 139
6.2 Design of covalent HyTag probes 141
6.3 Synthesis of covalent HyTag probes 142
6.3.1 Synthesis of linker and hydrophobic moiety 142
6.3.2 Synthesis of unsubstituted acrylamide 143
6.3.3 Synthesis of substituted acrylamide 145
6.4 Evaluation of purified BRD4(1)^{L94C} labelling using protein mass spectrometry 146
6.5 In-cell engagement and protein degradation by flow cytometry 148
6.6 Current work 151
6.7 Conclusions 153
Chapter 7: Conclusions and future work 154
7.1 Conclusions 155
7.2 Future work 157
Chapter 8: Experimental 159
8.1 General chemical methods 160
8.2 Synthetic procedures and compound data 163
8.3 General biology methods 245

8.3.1 Chemicals and cell culture reagents **245**
8.3.2 Bacterial cell culture **246**
8.3.3 Mammalian cell culture **247**
8.4 Protein expression and purification **247**
8.4.1 Cloning **247**
8.4.1.1 DNA amplification - PCR *247*
8.4.1.2 DNA purification - Agarose gel electrophoresis *248*
8.4.1.3 Gibson Assembly *248*
8.4.1.4 DNA Sequencing *249*
8.4.2 Competent cells **249**
8.4.3 Fermentation conditions for human bromodomain containing proteins **249**
8.4.4 Purification conditions for human bromodomain containing proteins **250**
8.4.5 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) **252**
8.5 Mammalian protein expression **253**
8.5.1 DNA transfection **253**
8.5.2 Cell harvesting and lysis **254**
8.5.3 Determination of total protein concentration **254**
8.5.4 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) **254**
8.5.5 Western blotting **255**
8.6 Assay procedures **255**
8.6.1 Differential scanning fluorimetry (DSF) **255**
8.6.2 Isothermal titration calorimetry (ITC) **256**
8.6.3 Protein LC-MS assays **256**
8.6.3.1 In-house *256*
8.6.3.2 GSK *257*
8.6.4 Dose response AlphaScreen™ **258**
8.6.5 Click probe efficiency **259**
8.6.6 Fluorescent labelling of purified BRD4(1)^{L94C} **260**
8.6.7 In-cell target engagement **260**
8.6.7.1 Click probe selectivity *260*
8.6.7.2 OXFBD02 Competition assay *261*
8.6.7.3 Immunoprecipitation *261*
8.6.8 Protein degradation **262**
8.7 Protein X-ray crystallography **262**
8.8 Computational Studies **263**
8.8.1 Molecular dynamics (MD) simulations **263**
8.8.2 Ligand Docking Studies **264**
References **265**

Appendices 278

Appendix A: Chapter 2 279

Appendix B: Chapter 3 294

Appendix C: Chapter 4 308

Appendix D: Chapter 5 315

Appendix E: Chapter 6 321

Appendix F: NMR spectra of novel compounds 328

Chapter 3 spectra 328

Chapter 4 spectra 372

Chapter 5 spectra 392

Chapter 6 spectra 432

Mutate & Conjugate:
A method for rapid in-cell validation of epigenetic targets

A bottleneck in the drug discovery process is target validation. Target validation is the process of demonstrating that engagement with a cellular target is linked to the desired therapeutic effect, confirming that the target has a key role in the disease phenotype. Incomplete or incorrect target validation can lead to costly failures at the later stages of drug development. Consequently, techniques that enable rapid and robust target validation are of great value to those seeking to discover new medicines. The advantage of using small molecule ligands over genetic knockdowns, in target validation, is the temporal control over the "knockdown". In this approach, reactive covalent fragments have been combined with a cysteine mutant of the target protein. It was hypothesised that covalent inhibition would circumvent the need to optimise ligand affinity, accelerating probe development, while also offering increased residency times that could potentially act as a functional knockdown, thus providing information on the function of the target.

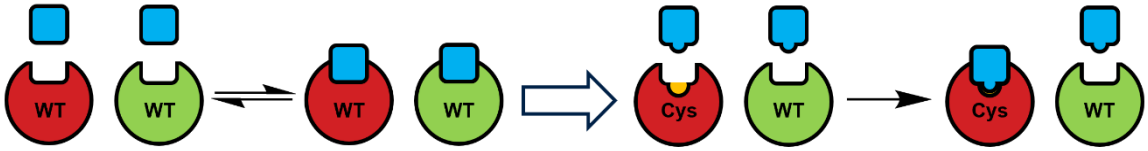


Figure 0.1. Schematic illustration depicting how a 'mutate & conjugate' methodology works.

The BET bromodomain-containing protein, BRD4, was chosen for proof-of-concept studies. Bromodomains are epigenetic 'reader' modules that recognise and bind to acetylated lysines on histone tails, and signal for the recruitment of protein complexes responsible for gene regulation. Owing to the conserved nature of the BET bromodomains, the development of single-domain selective ligands has proved challenging. Consequently, development of ligands that enable the probing of single-domain function are highly sought after. Herein, the application of a 'mutate & conjugate' approach to enable selective biological probing of BRD4(1) is reported. A structurally stable and functioning mutant, BRD4(1)^{L94C}, was designed, expressed, and validated. Using a structure-guided, and a fragment-based approach, ~270 reactive fragments were screened for C94 labelling. Subsequent fragments were developed and exhibited selective inhibition of BRD4(1)^{L94C} over BRD4(1)^{WT} in an AlphaScreen™ assay. Clickable analogues were used to evaluate proteome-wide selectivity, leading to the discovery of a covalent fragment with exquisite in-cell selectivity for BRD4(1)^{L94C}. Finally, a 'mutate & conjugate' strategy was applied to the development of selective protein degraders. This work shows that the combination of reactive fragments and site-directed-mutagenesis could be applied to rapidly develop chemical probes for target validation studies.

Acknowledgements

Firstly, I would like to thank my supervisor Stuart– I have truly enjoyed working with you over the past 4.5 years on this project. Your guidance, energy, and gambling of half pints has been paramount to this project, and I have learnt so much, although, I think you still owe me half a pint! To my industrial supervisor Jake, thank you very much for your advice and thought-provoking conversations during my time in Oxford, as well as the Champagne you bought us in Varsity Club, circa 2018. I would also like to thank Tom Milne for welcoming me into his wonderful lab and for giving me the opportunity to try my hand at cell work, I very much enjoyed my time at the WIMM.

Without a doubt, the Conway group has been the perfect place for me to undertake my DPhil. I've had the pleasure of working with countless numbers of amazing, caring, funny, and inspiring people, and I will hold these memories close.

I would like to thank all the Postdocs who have helped me over the years. Maze, thank you for your biological wisdom, interesting scientific discussions, and for introducing Bodrums to me. Ewen, it was great to work with you. Your love of chemistry, beer, and coffee was a combination that I have adopted! Pietro, your dance moves were formidable, and being in G12 with you was a lot of fun. I hope that someday you can retire and go back to chasing pigeons in a park on a warm summer's day. Debs, you're a legend– you're a brilliant Postdoc (and now PI) and you kept everything running smoothly. Thanks for all the Uni Club pints, ridiculous conversations, 'Lovin' You' duets, and minion memes over the past years, however I'm still waiting on that P53 PROTAC! Maria, thanks very much for the good vibes and the gossip while you were in Oxford. Pascal, thanks for all your help, you always have a great energy and enthusiasm that brightens up any lab. Thanks also for the many rides we did, I will be forever indebted to you for the water you gifted me with the first time we rode up Aston hill on the hottest day of the year! All the best with starting your own research group! Antoine, I've really enjoyed working and sharing a desk with you, you're a great addition to the group. Thank you for all the scientific, music, and general life chats we've shared. Marta, it's been wonderful to work with you, and I'm excited to see where the project goes with your wisdom! Glen, thank you for your insight into chemistry and craft beer, it's been fun working with you.

Mus, the column king, and the world record holder for the longest 'yeah boi', you were awesome to work with in G12, and an all-round top guy– cheers for the twirl bar when my ITC didn't work. I hope your new lab at GSK don't mind you playing 'Cat's in the cradle' and 'Shut up' on repeat! Catherine (Kefrine), thanks for being so welcoming and chatty when I joined, your laugh is infectious, and coupled with my laugh I think we successfully annoyed all the other groups in the

office area! Amy, thanks for putting up with me on our old desk, I think the North-South divide sticker is still there, you were great to have a laugh with and on occasion chat about chemistry. Coco, thanks for being so helpful, and for making that cardboard cut-out of me, unfortunately someone has stolen it! Joe, thank you for your help with docking and doing the MD, your sense of humour never failed to make me laugh. Jess, the meme queen, it was so fun working with you, you're always so positive and have a solid selection of memes. I hope a student in years to come finds your Q drive meme folder. Holly, thank you for your support, country music, and your dreadful dad jokes.

Darius, not gonna lie, I'm still salty about the magic mix, but apart from that it's been awesome working with you– we shared a lot of laughs, and I hope one day you get to meet your favourite band, The 1975. Also, did you know this is a Marvin Gaye sample? Katrina, you kept things in order in G12 while providing the most delicious baked goods. It's been a pleasure working with you, and I wish you all the best with the Western blots! Jenny, thanks for the bike rides and giving me a wheel to sit on so I didn't get dropped! Patrick, thank you for all the great music suggestions and exciting science chat, I hope you make that covalent fragment! Sophie and Harriet, you two were a hilarious combination and gave #Bae4 good vibes, although to be fair Antoine and I weren't short of those. Sophie, congratulations on making your probes and I wish you the best of luck in Oncology. Harriet, it was great to have someone to chat rowing with, even though I am now a shrivelled has-been, good luck with it all, and just remember to 'SEND IT!'. Bernie, thanks very much for amazing cakes, computational skills, and checking that the THF was dry! Alistair, the humour and energy you bring to the lab is great and I wish you all the best with your macrocycles. Oli and Michael, it's great that you both returned to the group for your DPhils! I wish you all the best for the future. Max, your choice of tunes were always left-field but I appreciated the wonky techno vibes! To the new DPhils and rotation students; Keefe, Ada, Ceri, and Katherine, I wish you the best of luck with your DPhil journeys!

Thanks to all the Part II students who I've worked with. A big shout out to James (Tilden) and Lily– it was great to be your fumehood neighbour. James, your tunes and vibes were impeccable, and I will never forget the Mr Men book we put together for you (future Conway group members can find this in my Q drive folder, titled 'Mr Wrong')! Lily, I'm glad you didn't finish your part II looking like that squashed fly! Was great to rep Somerville with you. Additionally, a special mention to Ed who helped with this project in the early stages.

A big thank you to the Milne lab who took me in and taught me the biology ways! It was a great opportunity and I really enjoyed working with you all. Vassi, thank you so much for taking the time

to help me get started, teaching me the techniques, and introducing me to Winesday and the “Cross Keys” at St Peter’s! Nick, thank you for your patience and your amazing troubleshooting skills. You’re going to make a great PI! Nicole, thanks for the gossip and the New Ground Donuts; your photography skills are unmatched, and I hope my snake photo lives on in the Milne lab!

A huge thank you to my collaborators and all the NMR, MS, admin, and facilities staff that have continued to support me throughout and made my DPhil journey as smooth as possible.

Outside of the CRL and the WIMM I would also like to thank friends and the Somerville MCR community who have encouraged and guided me and made my time in Oxford unforgettable. Raph and Joel, it was amazing to live with you guys during the DPhil. Casa Amor will always be remembered for the good food (not Joel’s though), good vibes, cleaning rotas, poor Wi-Fi, and ridiculous moments of complete insanity and weirdness. It’s a shame we never got that hot tub though. I will always remember the Friday nights we spent hanging out in the MCR with Connor and Jon, or the late-night walks around Oxford, inevitably ending at Kebab King.

To the ‘Succulent’ boys, thanks for the support over the past 12+ years, as well as the pints and the raves– I’m sure there will be more to come! To ‘Pharmafam 2.0’, your friendship and support has meant a lot to me since meeting you all in 2013, so thank you.

To Naomi, thank you for being the good egg that you are, for your constant support, encouragement, and gifts of chocolate and coffee, that has helped to power me through the last year or so of my DPhil.

Finally, to my family; Mum, Dad, Luke, and Hannah, who have been unwavering in their support, despite not fully understanding what I do! Thank you for everything you have done for me over the years, I will be forever grateful.

Abbreviations

Å	Angstrom
ABP	Activity-based probe
AlphaScreen™	Amplified luminescent proximity homogenous assay screen
ATP	Adenosine 5’-triphosphate
BCP	Bromodomain containing protein
BET	Bromodomain and extra terminal domain
Boc	*tert*-Butyloxycarbonyl
BRD	Bromodomain
BRD4	Bromodomain-containing protein 4
BSA	Bovine serum albumin
CDI	1,1’-Carbonyldiimidazole
CETSA	Cellular Thermal Shift Assay
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate
COSY	Homonuclear correlation spectroscopy
CREBBP	Cyclic AMP response element binding protein (CREB) binding protein
CuAAC	Copper-catalysed azide-alkyne cycloaddition
CV	Column volume
Da	Daltons
DIPEA	*N,N*-Diisopropylethylamine
DMAP	4-Dimethylaminopyridine
DMF	*N,N*-Dimethylformamide
DMI	3,5-Dimethylisoxazole
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dppf	1,1’-Bis(diphenylphosphino)ferrocene
DSF	Differential scanning fluorimetry
EP300	E1A Binding protein p300
eq.	Equivalents
ESI	Electrospray ionisation
Et_3N	Triethylamine
Et_2O	Diethyl ether
EtOAc	Ethyl acetate

EtOH	Ethanol
FAC	Final assay concentration
FP	Fluorescence polarisation
g	grams
GSH	Glutathione
h	Hour(s)
HAT	Histone acetyltransferase
HATU	1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate
HDAC	Histone deacetylase
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
HMBC	Heteronuclear multiple bond correlation
HPLC	High performance liquid chromatography
HRMS	High resolution mass spectrometry
HSQC	Heteronuclear single quantum correlation
HWE	Horner-Wadsworth-Emmons
Hz	Hertz
IC_{50}	Half maximal inhibitory concentration
IMAC	Immobilised metal affinity chromatography
IP	Immunoprecipitation
IPA	Isopropanol
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
J	Joules
KAc	ε-*N*-Acetylated lysine
K_d	Dissociation constant
K_i	Inhibitory constant
LC-MS	Liquid chromatography-mass spectrometry
LRMS	Low resolution mass spectrometry
M	Molar
MD	Molecular dynamics
MeCN	Acetonitrile
MeI	Iodomethane
MeOH	Methanol

MHz	Megahertz
min	Minute(s)
mL	Millilitres
mol	Moles
mp	Melting point
MS	Mass spectrometry
NBS	*N*-Bromosuccinimide
NCS	*N*-Chlorosuccinimide
NMP	*N*-Methyl-2-pyrrolidone
NMR	Nuclear magnetic resonance
PAL	Photoaffinity labelling
PARP	Poly(ADP-ribose) polymerase
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PDB	Protein data bank
PEG	Polyethylene glycol
PPI	Protein-protein interaction
ppm	Parts per million
PROTAC	Proteolysis-targeting chimeras
PTM	Post-translational modification
PyBOP	Benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate
R_f	Retention factor
RMSD	Root-mean-square deviation
RMSF	Root-mean-square fluctuation
RNAi	Ribonucleic acid (RNA) interference
rt	Room temperature
SAR	Structure-activity relationship
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SFI	Solubility forecast index
SGC	Structural Genomics Consortium
TBAF	Tetra-n-butylammonium fluoride
TBAI	Tetra-n-butylammonium iodide
TFA	Trifluoroacetic acid

THF	Tetrahydrofuran
TMS	Trimethylsilyl
TLC	Thin layer chromatography
UPLC	Ultra performance liquid chromatography
UV	Ultraviolet
WT	Wildtype

Chapter 1

Introduction

1.1 Target validation

1.1.1 Target validation in drug discovery

Over the past two decades the pharmaceutical industry has seen a general decline in productivity, where productivity is defined as the relationship between the number of late-stage (Phase II or Phase III) assets and the cost to develop these assets. Publications as early as 2010 and 2011, from Paul *et al*. and Bunnage,[1,2] highlighted the alarming trend that total research and development (R&D) costs were rising while the numbers of new molecular entities entering the market each year had stagnated, and in some years declined. Unfortunately, this trend continued for years to come.[3,4] Only in 2021 was there a tangible change in the direction of productivity.[5] Data collected by Deloitte, analysing the returns of 15 leading pharmaceutical companies (including GSK, AstraZeneca, Pfizer, Roche, and Novartis) between 2013–2021, shows that 2021 saw the first increase in average late-stage assets above the 2013 level. These data also show that 2021 saw a 16% decrease in the average R&D costs, from $2,376 million per asset in 2020 to $2,006 million per asset in 2021 (**Figure 1.1**). Whilst a large proportion of this increase in productivity was driven by emergency approvals for drugs and vaccines targeting COVID-19, this did not account for all of it; excluding COVID-19 assets shows there was still an increase from 2020.



Figure 1.1. Data collected and analysed by Deloitte LLP, from 15 major pharmaceutical companies between 2013–2021, shows (**A**) the average number of late-stage assets per company and (**B**) the average R&D cost to develop a therapeutic from discovery to launch. Reprinted with permission from Deloitte LLP.[5]

Whilst these data are encouraging and hopefully indicate a change in the productivity of the pharmaceutical industry, the overall failure rate for potential therapies entering pre-clinical studies remains around 95%, with over 90% of potential therapies failing in clinical trials.[4,6–9] With failure rates so high, companies opt for safer options, such as the derivatisation of existing drugs with proven mechanisms of action. This strategy stunts innovation and leaves undrugged diseases poorly treated. Furthermore, failings in the clinic are costly and can lead to the closure of departments, and job losses.[6] Strikingly, it is estimated that 79% of clinical failures are due to lack of efficacy or poor safety profiles,[8,10] often caused by unwanted off-target effects. Worryingly, a large proportion of these failures could have been avoided earlier in the drug discovery process before entering clinical trials, with some of these failures caused, in part, by poor early-stage target validation; however, these are complex problems to tackle.[1,11,12]

Target validation is the process of demonstrating that engagement with a cellular target is linked to the desired therapeutic effect and confirms that the target plays a key role in causing the disease phenotype. Estimates suggest that greater focus on target validation could lead to a ~25% decrease in Phase II attrition rates, which would in turn decrease the average cost of drug development by ~30%.[1,13] Whilst these figures are only predictions, a change this dramatic would no doubt lead to companies venturing into untapped and innovative areas of drug discovery.

A case study demonstrating how increased focus on early-stage target validation can improve productivity is that of AstraZeneca (AZ).[14,15] In 2011, after R&D productivity from 2005–2010 were below industry averages, AZ embarked on changing its R&D culture with the aim of increasing productivity. AZ implemented a new framework, termed the '5Rs': the right target, right tissue, right safety, right patient, and right commercial potential. Increased attention to target validation, or invalidation, and a deeper understanding of the biology, meant that failures were realised earlier in projects before more costly clinical studies were undertaken. This shift in strategy increased the success rate of projects progressing to lead optimisation, from 23% (2005–2010) to 48%

(2012– 2016), while project failures during 2012–2016 were driven more by target invalidation or insufficient target validation than during 2005–2010 (**Figure 1.2 A & B**).[15] Overall, this strategy led to an increase in candidate drugs progressing to Phase III trials from 4% (2005–2010) to 19% (2012– 2016) (**Figure 1.2 C**). These results are inspiring and demonstrate the need for robust and reproducible target validation that interrogates the complex biology and ensures that the target is key to the disease.

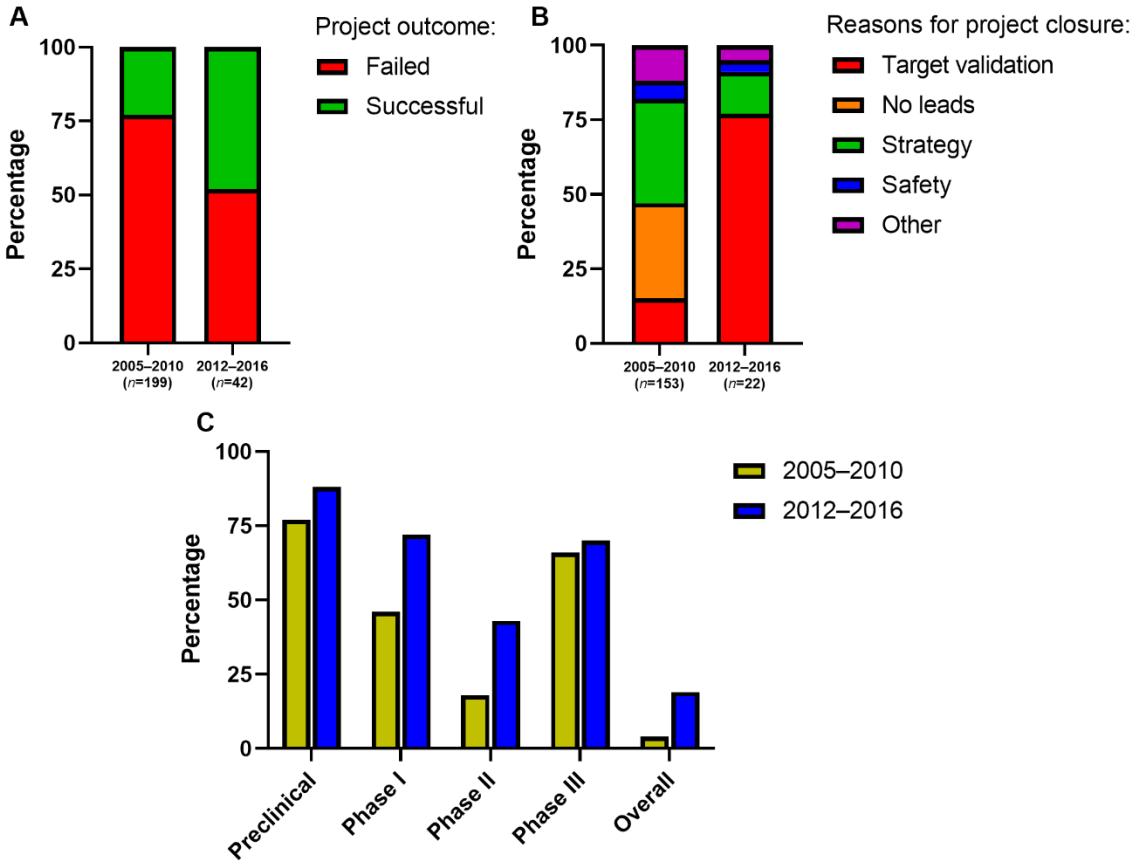


Figure 1.2. Data from AZ, pre- (2005–2010) and post- (2012–2016) introduction of the 5R framework. **A**) Percentage of preclinical projects reaching lead optimisation. **B**) Reasons for early-stage discovery project failure. **C**) Overall project success rates for AZ.[15]

1.1.2 A biological approach to target validation

The use of genetic techniques such as knockdowns, knockouts, and overexpression, in target validation has been invaluable, and with the rise of CRISPR-Cas9 screens this will continue to grow.[9,16] The ability to manipulate targets at the genomic level, in cells or whole organisms, has helped validate loss-of-function phenotypes, and how mutations within the target affects its activity.[17] Nevertheless, these techniques possess some drawbacks that can lead to incorrect interpretation of the target's function (**Figure 1.3**).

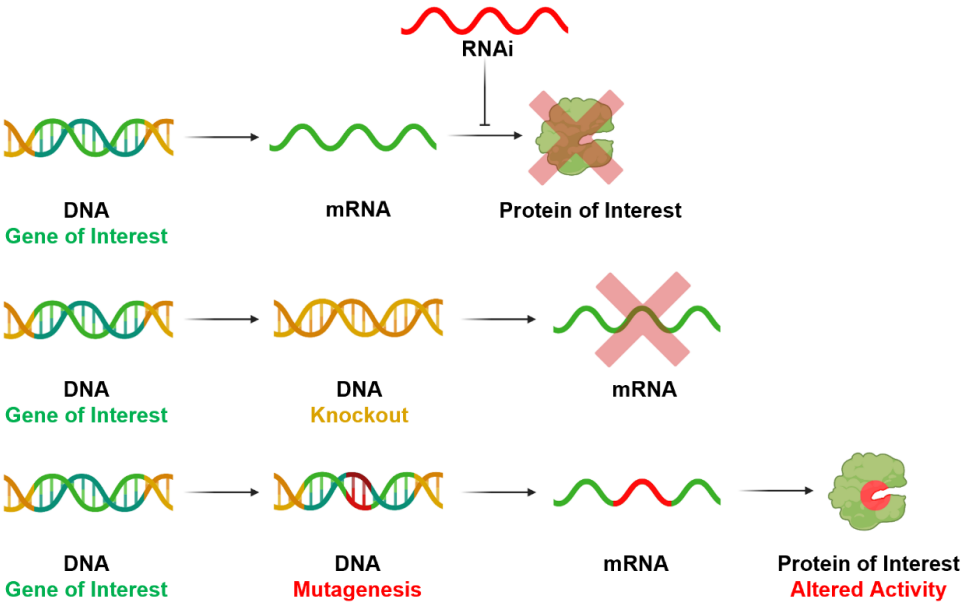


Figure 1.3. Biological methods used to aid target validation through manipulation of the target at the genomic level.[18]

The method of RNA interference (RNAi) is powerful in interrogating a protein's function and allows silencing of genes through inhibiting the translation of the target gene. The development of RNAi reagents for knockdown studies are usually affordable and quick to obtain, making it a standard technique within target validation.[17] However, this approach of removing the majority of the protein from the cell fails to consider four key points: (i) for multidomain proteins, what are the roles of the other domains within the protein and how do they influence the phenotype, (ii) does the target protein's structure play scaffolding roles, (iii) does it partake in protein-protein interactions (PPI) that may influence the phenotype, and (iv) does this change in protein

stoichiometry alter protein complexes. In this way, RNAi validation results may vary from the conclusions drawn from small molecule inhibitors.[17,19,20] Absence of the protein can also lead to compensatory biochemical pathways that might induce non-physiologically relevant conditions,[21] further reducing the validity of the phenotypic results. Additionally, the off-target effects of RNAi can be further-reaching than those associated with small molecules.[17] Off-target effects of small molecules are usually limited to effects mediated by proteins of similar structure, whereas for RNAi, short stretches of homologous sequences might exist with structurally unrelated proteins.[17,22]

1.1.3 A chemical approach to target validation

Chemical probes are biologically active small molecules that can modulate a target protein's function and can help ask mechanistic and phenotypic questions about the target protein.[20] Chemical probes are vital for reliable target validation, and this is where chemical biology and drug discovery work in synergy.[20,23–25] Although some compounds can be used as both chemical probes and drugs, there are usually some distinct differences. Therapeutic drugs are designed to be efficacious at targeting a disease: they might display polypharmacology or have tolerable off-target effects.[20,23] Chemical probes, however, should ideally exhibit exquisite selectivity for their target, this ensures that robust conclusions can be drawn from cellular screens, to elucidate protein function and validate a target, for example, the Structural Genomics Consortium (SGC) requires probes to exhibit >30-fold selectivity over related proteins.[20] Therapeutic small molecules are rigorously assessed for bioavailability, pharmacokinetics and pharmacodynamics, however, chemical probes do not require the same level of optimisation of these properties, and instead selectivity, affinity, solubility, and cell permeability are most important.[20,23–25]

Whilst biological approaches to target validation generally work by inhibiting the proteins biosynthesis, chemical probes instead inhibit protein function once the protein has been biosynthesised (**Figure 1.4**).[17] The advantage of a chemical probe approach is that the protein is

present within the cell and can participate in normal physiological processes— leading to more meaningful phenotypic data. The use of chemical probes therefore allows for temporal control over modulation of the proteins function, i.e., the effect is induced upon addition of the probe. A chemical probe approach also allows for dose-dependent studies to be carried out, which is particularly useful if complete inhibition or RNAi removal of the target protein results in cell death.[19,26] An additional benefit of using chemical probes over genetic techniques is the targeting of specific domains within a protein, whereas techniques such as RNAi remove the entire protein and all functions related to it. Therefore, the phenotypic results acquired using chemical probes are more akin to the effects observed for small molecule therapeutics.

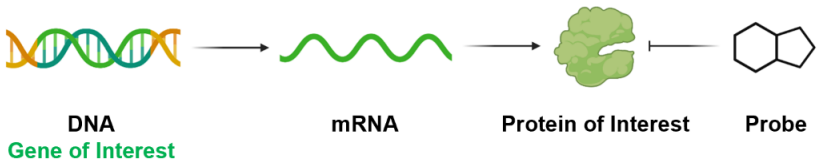


Figure 1.4. Chemical probes provide a powerful tool for understanding and modulating a target protein's functions.[18]

Chemical probes can also be used to demonstrate cellular engagement with the target protein, as well as off-target interactions, through the use of modified analogues bearing tagging groups for protein enrichment, or fluorophores for visualisation.[27–29]

Overall, chemical probes offer an orthogonal approach to validating potential drug targets before an expensive full drug discovery project is commenced. However, the main downside of chemical probes for target validation is the amount of time and resources it can take to develop a probe with desirable affinity, selectivity, and cell permeability; especially if the protein of interest is an emerging target, or poorly understood. Methods that aim to increase the speed of selective probe development are crucial, as rapid target validation leads to faster and more efficient drug discovery.

1.2 Covalent ligands and probes

Most ligands or probes bind to their target protein reversibly, through non-covalent interactions, such as hydrogen bonding, hydrophobic interactions, π-π stacking or cation-π interactions (**Figure 1.5 A**).[30] Covalent ligands and probes, instead, work through a two-step mechanism; first is the non-covalent binding of the ligand to the protein of interest *via* the aforementioned non-covalent interactions. The second step is a chemical reaction between a reactive group on the ligand and a specific amino acid residue within the protein, forming a covalent bond, which can be reversible or irreversible depending on the reactive group (**Figure 1.5 B**). Selectivity for a given target is achieved through the first, non-covalent, step. To allow for covalent tethering to the desired residue, a reactive group must be incorporated into the molecule to allow the reaction with the target amino acid. Whilst the covalent bond may be reversible, tethering the ligand to the protein typically prolongs the residence time of the ligand and its effects, often until the protein is degraded and biosynthesised again.

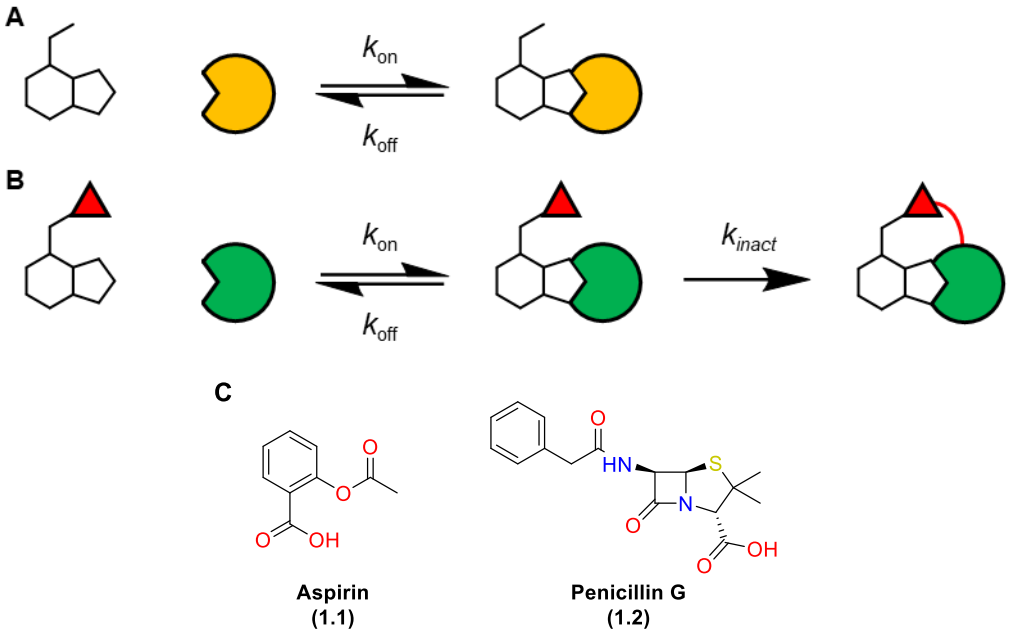


Figure 1.5. Ligands can bind reversibly or irreversibly to target proteins. **A**) Non-covalent ligands bind reversibly to target proteins through interactions such as hydrogen bonding and lipophilic interactions. **B**) Covalent ligands possess reactive groups (red triangle) and can bind irreversibly to target proteins. They have affinity to initially bind non-covalently to the target protein, holding the ligand in position for the covalent bond formation to occur. **C**) Chemical structures of the irreversible inhibitors aspirin and penicillin G.

Traditionally, drug discovery programmes have avoided the development of covalent ligands due to the concerns surrounding off-target binding to other biomolecules, long residency times and potential idiosyncratic toxicity.[31] Although unknown until long after their use became widespread, both aspirin (**1.1**) and penicillin G (**1.2**) exert their action through irreversible covalent inhibition of their targets (**Figure 1.5 C**).[32,33] Recently there has been a renewed interest in the development of covalent ligands, in both medicinal chemistry and chemical biology,[34–37] with much focus on the use of covalent fragment screening for the discovery of ligands and probes for emerging targets, or targets with no known reversible ligands.[38–41]

The development of covalent ligands offers many benefits over their traditional non-covalent counterparts. Selectivity between proteins that share high structural similarity can be achieved by covalently targeting a non-conserved nucleophilic residue within the target. This approach has been vital for the selective targeting of kinases and has led to the discovery of covalent kinase inhibitors targeting non-conserved cysteine residues.[30,42] Similarly, covalent ligands can overcome the reversible competition posed by abundant and high affinity substrates, such as ATP in the case of kinases.[30,42,43]

Covalent ligands also offer the opportunity for targets previously described as 'undruggable'.[34,35,44] 'Undruggable' targets include proteins with shallow or poorly defined binding pockets which are conventionally hard to target through reversible ligands— owing to the lack of available interactions. So long as the target residue is proximal, covalent ligands can potentially overcome this issue as they can require fewer non-covalent interactions,[35] and once bound will remain bound for longer than non-covalent ligands.

The slow reversibility of the covalent bond formation offers a prolonged mechanism of action and possibly complete inactivation of the target's functions. From a chemical biology and target validation perspective these opportunities are desirable and allow for detailed evaluation of the

target's function. However, for medicinal chemistry, these may not always be desired, especially for proteins that are functionally vital and have a slow turnover rate.[42]

1.2.1 Targeting cysteine residues

Cysteine residues are the most frequently targeted residue for covalent labelling and offer a unique set of properties that make them an attractive target. The cysteine thiol is the most nucleophilic side chain of the canonical amino acids and deprotonation yields a highly nucleophilic thiolate, capable of reacting with a diverse range of electrophiles.[45,46] Additionally, the cysteine thiol possesses a pK_a of ~8.6, and therefore ~6% exists as the deprotonated thiolate at pH 7.4.[47,48] Moreover, the pK_a of a cysteine thiol can shift dramatically depending on the microenvironment, and pK_a values as low as 2.5 have been documented.[49] Of the 20 canonical amino acids, cysteine is one of the least abundant, representing only ~2% of the total residues within the proteome (the proteins expressed in cells).[50,51] This low abundance offers a potential strategy to selectively target proteins within a given family, by targeting non-conserved cysteine residues that are present within the target protein. Due to the strong interest in targeting cysteine, there are many reactive groups used for covalent labelling of cysteine,[46] however, (i) disulfides, (ii) acrylamides, (iii) haloacetamides, and (iv) α,β-unsaturated esters, are the main approaches.

1.2.1.1 Disulfides

One of the earliest approaches to actively seek ligands that were able to covalently engage with proteins was disulfide tethering.[52] Libraries of fragments bearing disulfide bonds can be screened against potential targets, and fragments able to bind to the protein surface near the cysteine can undergo disulfide bond exchange, thus forming a covalent bond with the protein (**Figure 1.6 A**).[39,46,52,53] Because of the reversibility of the disulfide bond formation, only fragments with

sufficient affinity for the protein will remain tethered to the protein, while fragments with non-specific or weaker affinity are out competed and cleaved to reveal the thiol. Protein mass spectrometry can then be used to identify any disulfide fragments capable of labelling the protein, and the identified hit fragments can then act as a starting point for ligand development. One of the most notable discoveries from the application of disulfide tethering was the identification of **6H05** (**1.3**) as an allosteric covalent ligand for the oncoprotein K-Ras^{G12C} (**Figure 1.6 B**).[44] Fragment **1.3** was then developed through SAR into covalent inhibitor **1.4**, bearing a reactive acrylamide group.

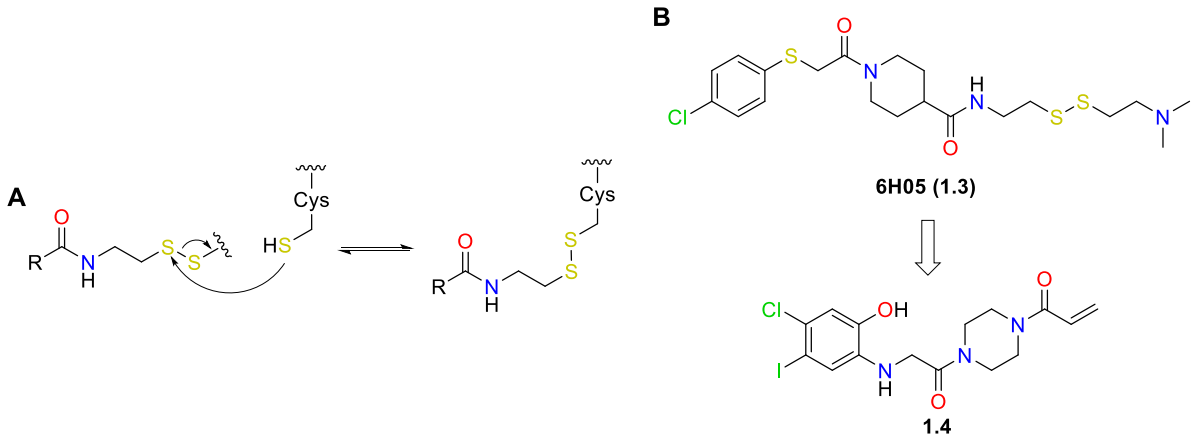


Figure 1.6. Disulfide tethering has led to the discovery of covalent ligands and probes. **A**) Disulfide tethering mechanism. **B**) Fragment **1.3** was discovered by disulfide tethering, which was then optimised to **1.4**.[44]

Site-directed-mutagenesis and disulfide tethering have also been combined to discover ligands for hard to target proteins.[52,54–56] In this technique, a cysteine mutant is engineered into a target protein near a binding site. Screening of disulfide containing fragments, using protein mass spectrometry or fluorescent polarisation (FP) allows for the discovery of weakly binding fragments that might not appear as hits through other screening methods. The hit fragments can then be optimised through SAR to develop non-covalent ligands for the wildtype (WT) protein.

The main challenges facing the disulfide tethering approach is that: (i) generally the disulfide group must be replaced with a more physiologically stable reactive group during ligand development (as seen for compound **1.4**), (ii) the poor accessibility to disulfide screening libraries, and (iii) the synthesis requires the use of thiols that are easily oxidised and have an unpleasant smell.[53] Whilst the use of reactive disulfides has played an important role in the development of covalent ligands,

the focus has now shifted towards the use of carbon-based electrophiles that are typically easier to handle and more accessible.

1.2.1.2 Acrylamides

The use of acrylamide groups for covalent labelling of proteins has become popular in both medicinal chemistry and chemical biology.[45,57] The α,β-unsaturated carbonyl of acrylamides acts as a Michael acceptor that can undergo conjugate addition with cysteine, allowing for covalent bond formation (**Figure 1.7**). Acrylamides are generally relatively weak electrophiles, and so the reacting nucleophile must be close and orientated optimally for covalent bond formation, whereas their enone counterparts exhibit higher electrophilicity and reactivity, and thus are not as popular within the field.[45]

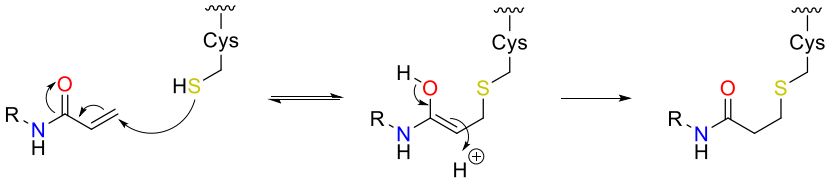
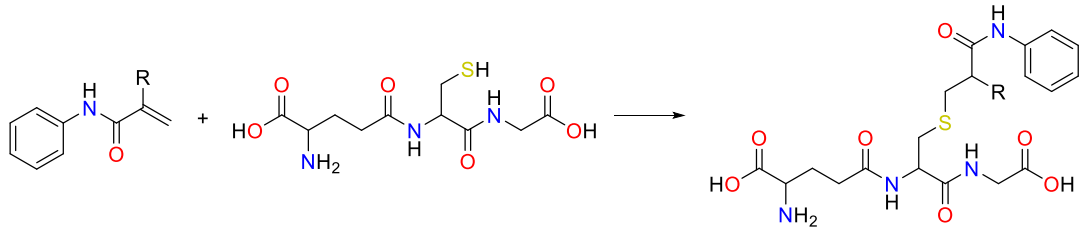


Figure 1.7. Acrylamides react with cysteine thiols *via* a 1,4-conjugate addition.

Due to the popularity of the acrylamide group, much work has gone into investigating how acrylamide modification impacts reactivity and selectivity, to allow for fine-tuning of reactive groups. A comprehensive study by Birkholz *et al*. examined the effects of α- and β-substitution on glutathione (GSH) reactivity with an *N*-phenylacrylamide model using an LC-MS assay and computational studies.[58] Substitution at the α-position by methyl- (**1.8**) or mildly electron-withdrawing groups led to significantly decreased reactivity, while addition of a strongly electron-withdrawing trifluoromethyl group (**1.9**) at this position resulted in over a 10-fold increase in reactivity (**Table 1.1**). Substitution at the α-position with aminomethyl groups (**1.10**) increased reactivity relative to the hydrogen.

Table 1.1. A selection of data demonstrating how α-acrylamide substitution alters reactivity with GSH.[58]

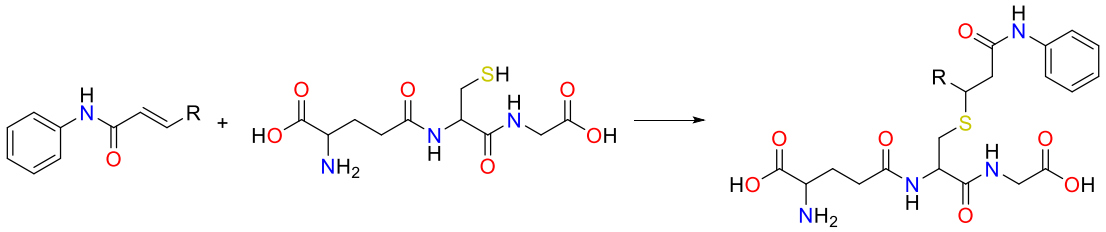


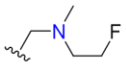
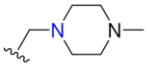
Compound	R	GSH $t_{1/2}$ (min)
1.5	H	179
1.6	F	>512
1.7	Cl	3
1.8	CH_3	>512
1.9	CF_3	15
1.10	CH_2NMe_2	<2

A variety of β-substituted acrylamides were next investigated (**Table 1.2**). A simple methyl substitution (**1.11**) was found to exhibit slower reaction kinetics compared to the unsubstituted acrylamide (**1.5**), likely due to a decrease in acrylamide electrophilicity as a result of electron donation from the methyl group. Substitution with a β-trifluoromethyl group (**1.12**) enhanced reactivity to levels observed for the α-trifluoromethyl substitution. A commonly employed acrylamide modification is the β-dimethylaminomethyl (DMAM) group (**1.13**), was shown to moderately increase the rate of GSH-adduct formation, which is in-line with some previous observations.[59] However, this result is contradictory to the work conducted by Martin *et al*. who observed faster reaction rates for the unsubstituted acrylamide compared to the β-DMAM acrylamide.[60] Interestingly, it is of note that both studies used the same model acrylamides, however, differed in the cysteine model, where Birkholz *et al*. utilised GSH and Martin *et al*. used a Boc- and methyl ester- protected cysteine. This difference in the model substrate could account for the differences in their results. Overall, Birkholz *et al*. concluded that the effects of substitution on reaction rates could generally be understood in terms of electron-donating or withdrawing ability

of the substituent, and that substitution of the β-position with aminomethyls possessing pK_{aH} values of >7 (**1.13** and **1.14**) enhanced reactivity, while pK_{aH} values of <7 (**1.15** and **1.16**) decreased reactivity towards GSH.

Table 1.2. A selection of data demonstrating how β-acrylamide substitution alters reactivity with GSH.[58]

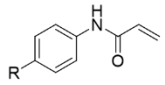


Compound	R	GSH $t_{1/2}$ (min)	Calc. pK_{aH}
1.5	H	179	N/A
1.11	CH_3	>512	N/A
1.12	CF_3	16	N/A
1.13	CH_2NMe_2	116	8.2
1.14	CH_2NEt_2	170	9.2
1.15		326	6.7
1.16		>512	2.8

Although most studies have focused on how modification of the acrylamide affects reactivity, it is important to note that acrylamide reactivity can also be influenced by the nature of the amine component. This is an important factor when considering the design of covalent ligands, where it is likely that the acrylamides may be appended to a variety of amine scaffolds, such as heterocycles or aliphatic groups. Multiple studies have investigated how different phenyl substituents affect acrylamide reactivity.[40,61,62] Generally, it was found that substitution at the *ortho*- and *para*-positions were found to have the most influence on reactivity, and that substitution with strongly electron-withdrawing groups, such as -NO_2, enhanced reactivity compared to *N*-phenylacrylamide. Kathmen *et al*. observed up to a ~2044 fold difference in GSH reactivity between an

electron-withdrawing and an electron-donating aryl substituent (**Table 1.3**).[40] Additionally, Lonsdale *et al*. probed a diverse range of acrylamides formed from a variety of heterocyclic-, phenyl-, and aliphatic-amines. It was found that heterocyclic amine scaffolds were generally the most reactive, while aliphatic amines were the least reactive.[62] In addition, covalent fragment screening efforts by Nonoo *et al*. and Kathman *et al*. uncovered highly reactive fragments with acrylamides directly attached to electron-poor heterocycles.[40,53,63]

Table 1.3. Pseudo-first order reaction kinetics for acrylamides bearing substituted phenyl groups.[40]

Scaffold	R	Compound	Pseudo-first order rate ($k_1 s^{-1}$)	Variation in reactivity
	H	**1.15**	1.645×10^{-4}	~2044 fold
	OMe	**1.17**	3.914×10^{-4}	
	NO_2	**1.18**	8.003×10^{-2}	

A proteomic study of clickable probes based on known kinase inhibitors bearing acrylamide groups found that modifications at the β-positions not only altered reactivity but, interestingly, selectivity.[64] Methyl substitution of the Bruton's tyrosine kinase (BTK) inhibitor, **Ibrutinib** (**1.19**) to develop **1.20**, led to a marked increase in selectivity for BTK, despite a decrease in IC_{50} (80 nM compared to 1 nM), however, this trade-off in affinity for selectivity may be advantageous for chemical probes where minimal off-target interactions are highly desirable (**Figure 1.8 A**). DMAM substitution of **1.19** did not significantly alter cellular reactivity, however, DMAM substitution of the epidermal growth factor receptor (EGFR) inhibitor, **PF-6274484** (**1.21**), led to greatly increased proteome-wide cross-reactivity for **1.22** (**Figure 1.8 B**).

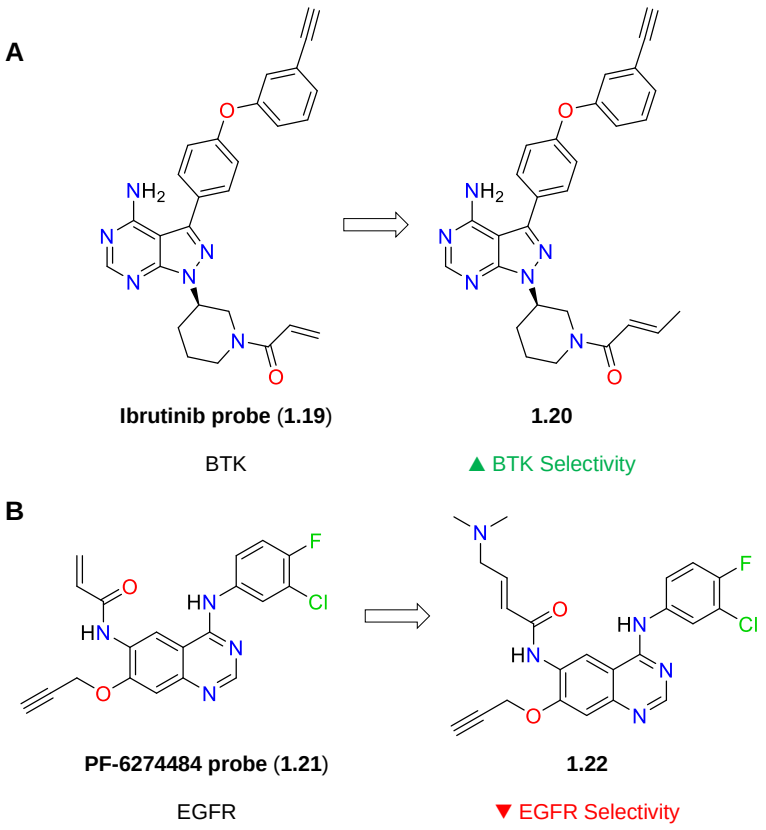


Figure 1.8. Acrylamide modifications altered in-cell reactivity and selectivity. **A**) Methyl substitution increased selectivity for **1.20**. **B**) DMAM substitution of **1.21** to develop **1.22** drastically increased reactivity and off-target binding.[64]

1.2.1.3 Haloacetamides

Haloacetamides can undergo an S_N2 reaction with cysteine (**Figure 1.9**). These reactions proceed rapidly due to the π* orbital of the carbonyl and the σ* orbital of the carbon-halogen bond combining to form a new, lower energy, LUMO that is easily attacked by the nucleophilic cysteine.[65] The most commonly used halides are chloro- and iodo-derivatives.

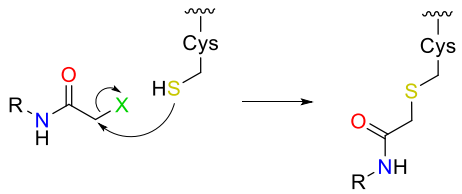


Figure 1.9. Haloacetamides react with cysteine thiols *via* an S_N2 displacement.

In pioneering work that aimed to identify potentially ligandable cysteine residues in the proteome, Cravatt and co-workers developed the isoTOP-ABBP (isotopic tandem orthogonal proteolysis-

activity-based protein profiling) method.[66,67] Briefly, lysates or live-cells were pre-treated with DMSO or a reactive fragment, before incubation with a broad-spectrum cysteine reactive clickable probe (IA-alkyne) comprised of an iodoacetamide electrophile (**Figure 1.10 A**). Samples were then subjected to a CuAAC reaction with isotopically labelled (^{13}C and ^{15}N) biotin-azide tag, where DMSO samples were tagged with the heavy isotope, and the reactive fragment samples with the light isotope. IA-alkyne labelled proteins clicked to isotopically labelled biotin were then enriched by streptavidin, and the IA-alkyne labelled proteins from the DMSO samples and the reactive fragment samples were compared using MS-based proteomics. Therefore, proteins harbouring cysteines targeted by the reactive fragments would be suppressed in the reactive fragment samples compared to the DMSO samples. This method was then applied to discover and develop a covalent chloroacetamide probe **1.24** that was used to study the poorly understood functions of CASP8 and CASP10 in extrinsic apoptosis (**Figure 1.10 B**).[68]

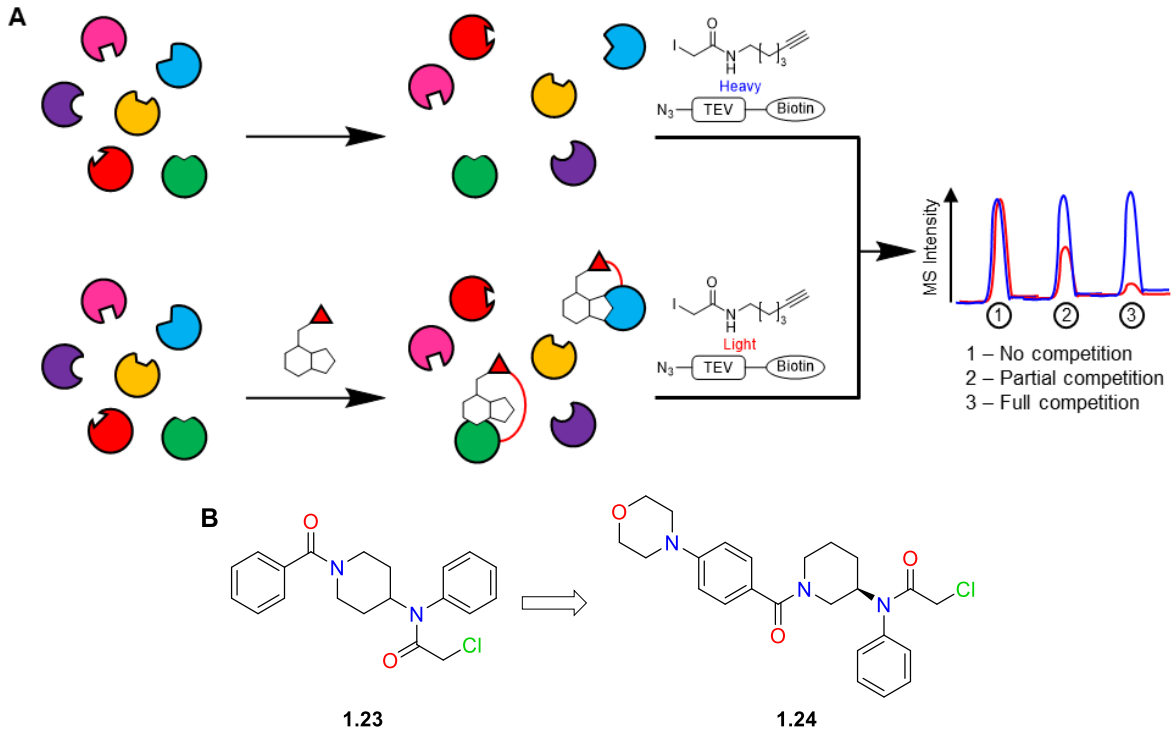


Figure 1.10. Proteome-wide screening for ligandable cysteines. **A**) Schematic representation of the isoTOP-ABBP method. MS competition ratios were used to identify proteins with ligandable cysteines. **B**) Fragment **1.23** was discovered using the isoTOP-ABBP method, and was subsequently developed into **1.24** to probe CASP8 and CASP10 function.[68]

Along with acrylamides, chloroacetamides have become common place in reactive covalent fragment screening.[69] Work by Resnick *et al*. screened approximately 1000 electrophilic fragments comprising a chloroacetamide or an acrylamide reactivity group against 10 proteins of interest.[38] Of the 10 proteins, 7 were labelled with potent and unique fragments. Using the chloroacetamide hits, probes for the deubiquitinase, OTBU2, and the pyrophosphatase, NUDT7, were discovered, for which there were previously no known inhibitors (**Figure 1.11**).

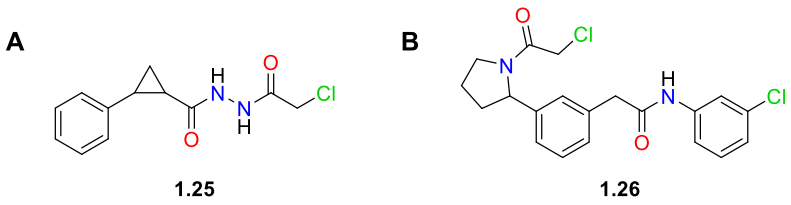


Figure 1.11. Reactive fragment screening was used to identify covalent ligands for proteins with no known inhibitors. **A**) Compound **1.25** is a covalent inhibitor of OTBU2. **B**) Compound **1.26** is a covalent inhibitor of NUDT7.[38]

1.2.1.4 α,β-Unsaturated esters

α,β-Unsaturated esters are Michael acceptors that can form a covalent bond with cysteine *via* a 1,4-conjugate addition reaction and have recently been used to develop covalent ligands (**Figure 1.12**).[45,70,71]

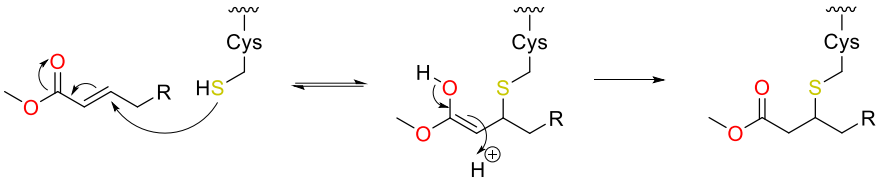
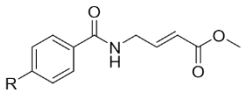


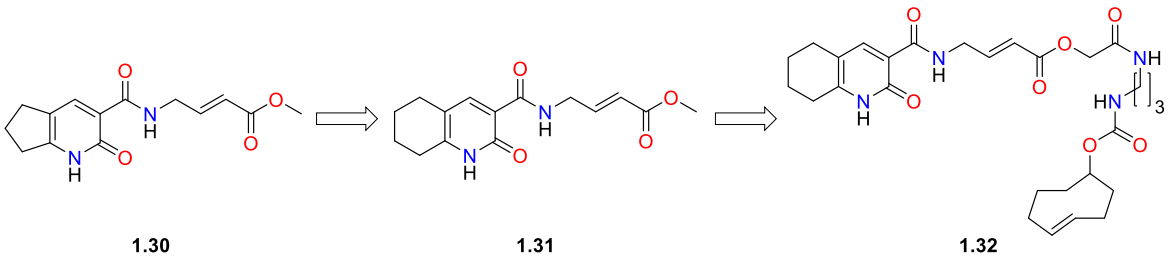
Figure 1.12. α,β-Unsaturated esters react with cysteine thiols *via* a 1,4-conjugate addition.

Kinetic studies by Kathman *et al*. found that the α,β-unsaturated ester group exhibited the least fluctuation in reactivity as a result of changes to a phenyl scaffold when compared to the corresponding acrylamides, vinyl sulfonamides, and sulfones (**Table 1.4**).[40] Using this preferred reactive group and fragment screening, three covalent inhibitors of the cysteine protease, papain, were identified.[40]

Table 1.4. Pseudo-first order reaction kinetics for α,β-Unsaturated esters bearing substituted phenyl groups.[40]

Scaffold	R	Compound	Pseudo-first order rate ($k_1 s^{-1}$)	Variation in reactivity
	H	**1.27**	3.726×10^{-4}	~1.6 fold
	OMe	**1.28**	3.724×10^{-4}	
	NO_2	**1.29**	5.914×10^{-4}	

Inspired by these findings Johansson *et al*. screened over 100 α,β-unsaturated ester fragments to discover and develop covalent probes for the catalytic HOIP subunit of the linear ubiquitin chain assembly complex (LUBAC), an understudied E3 ligase.[41] Optimisation of hit fragment **1.30** led to the discovery of the covalent HOIP inhibitor **1.31**, which also displayed promising proteome-wide selectivity through the use of a clickable analogue **1.32** bearing a *trans*-cyclooctene tag (**Figure 1.13**).



1.30 **1.31** **1.32**

Figure 1.13. The development of HOIP inhibitor **1.31** and clickable analogue **1.32** from fragment **1.30**.[41]

1.2.2 Targeting lysine residues

Lysine residues are more ubiquitous than cysteines and therefore offer an alternative approach to selective covalent labelling of proteins where cysteine residues are unavailable.[34,35,72] Recent proteome-wide studies have revealed a large number of potential lysine residues that can be modified, many of which were found in proteins with no available probes.[73,74] The major drawback of targeting lysine residues is that at physiological pH the vast majority of lysine residues (pK_{aH} of 10.4) will reside in their protonated form, and are therefore non-nucleophilic. This does, however, offer the opportunity for selective targeting of lysines that exist in the neutral and nucleophilic

form, due to their microenvironment. Two commonly used reactive groups to selectively label lysine residues are sulfonyl fluorides and activated esters.

1.2.2.1 Sulfonyl fluorides

Sulfonyl fluorides have previously been described as a privileged reactive group with applications in targeting a variety of different amino acids, such as tyrosine, histidine, and serine. [75,76] Despite their broad range of reactivity, their use in selective lysine labelling has been widespread and has led to the development of several covalent ligands (**Figure 1.14**).

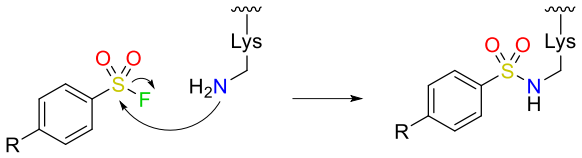


Figure 1.14. Mechanism for the reaction of sulfonyl fluorides with lysine amines.

Pettinger *et al*. optimised the lysine targeting HSP72 inhibitor, **1.33**, by replacing the reactive acrylate group with an aryl sulfonyl fluoride group, **1.34** (**Figure 1.15 A**).[77] This modification enhanced covalent efficiency by over 100-fold, enabling further understanding of HSP72 chaperone protein function. The sulfonyl fluoride reactive group has also played a role in probing kinases. 5'-Fluorosulfonylbenzoyl adenosine (**FSBA**, **1.35**) is a pan-kinase covalent inhibitor and binds into the ATP binding pocket as an ATP analogue (**Figure 1.15 B**). An activity-based probe (ABP) FSAB derivative **1.36** engages with the Src-family kinases in live-cells by covalently labelling the conserved catalytic L295.[78]

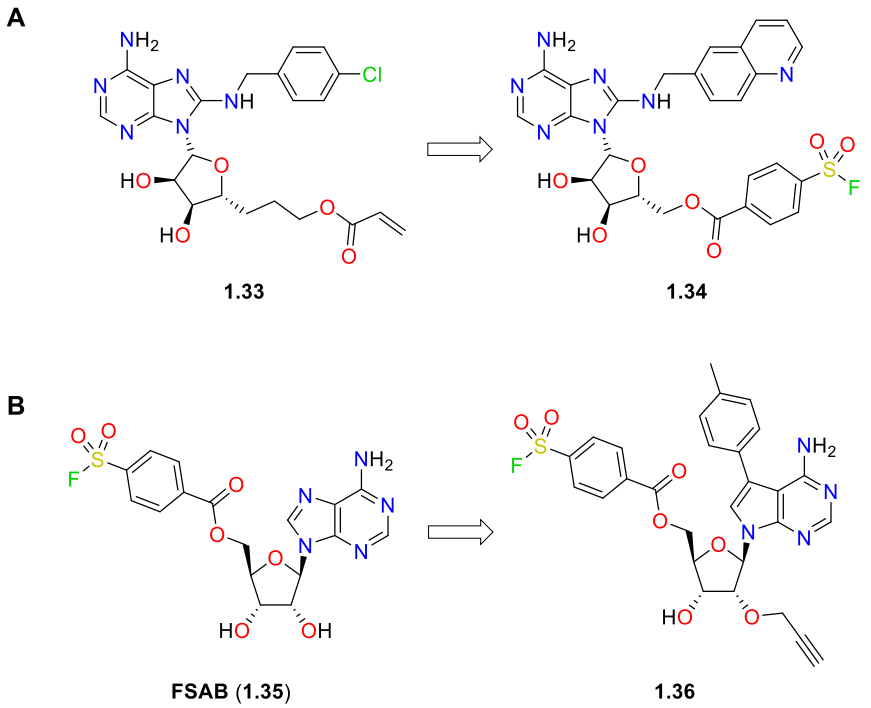


Figure 1.15. Development of lysine targeting covalent probes using a sulfonyl fluoride reactive group. **A**) HSP72 inhibitor **1.33** was kinetically optimised by replacing the acrylate with a sulfonyl fluoride.[77] **B**) **1.35** was developed into an alkyne bearing ABP **1.36** to evaluate in-cell target engagement.[78]

1.2.2.2 Activated esters

Activated esters covalently label lysine residues through the formation of an amide bond (**Figure 1.16**). These reactive groups are promising functionalities in protein kinase labelling, where covalently targeting conserved catalytic residues may be an attractive method for overcoming drug resistance.[79–81]

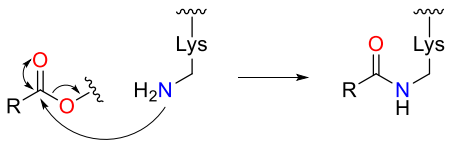


Figure 1.16. Activated esters react with lysine amines *via* an addition-elimination mechanism.

In 2018 Dalton *et al*. published their work on the development of a covalent inhibitor targeting the catalytic L779 of phosphoinositide 3-kinase delta (PI3Kδ).[80] Using the reversible clinical candidate **1.37** as a starting point, the sulfonamide interacting with L779 was changed to a range of activated esters, and cellular assays were used to determine the optimal SAR (**Figure 1.17 A**). From the

selection of esters, the 4-fluorophenol analogue **1.38** was shown to exhibit the greatest selectivity for PI3Kδ over the isoforms PI3Kα, PI3Kβ, and PI3Kγ. Protein mass spectrometry and crystallography further confirmed covalent bond formation, and pulldown studies using a clickable azide analogue revealed target engagement with PI3Kδ and selective labelling of PI3Kδ. Targeting the same protein, and the same L779 residue, Fournier *et al*. took inspiration from nature's post-translational modifications (PTMs) to design a PI3Kδ ligand (**1.39**) that was able to effectively acetylate L779 (KAc) and render PI3Kδ inactive (**Figure 1.17 B**).[81] The authors hypothesise that labelling and inhibiting targets with naturally occurring modifications, such as KAc, may decrease the potential risks of immune-mediated side effects.[81]

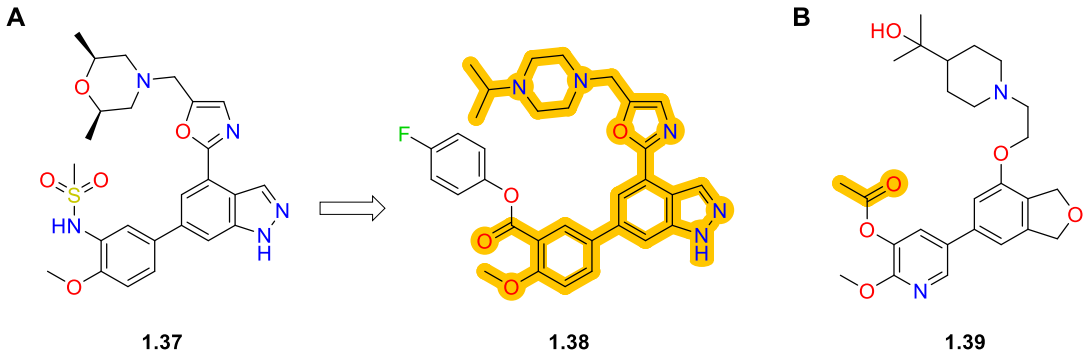


Figure 1.17. Activated esters covalently targeting L779 of PI3Kδ. The lysine labelling groups are highlighted in orange. **A**) Non-covalent ligand **1.37** was used as a template to develop covalent probe **1.38**.[80] **B**) Covalent probe **1.39** was developed to acetylated L779 and inhibit PI3Kδ function.[81]

1.2.3 Targeting other residues

Whilst covalent labelling of cysteines and lysines has demonstrated to be of great utility in selectively targeting closely related or hard to target proteins, it would be highly advantageous to expand this covalent toolbox to target a wider range of residues. Although many of these emerging methods are only in their infancy there is already great promise for new methods to cover a spectrum of different residues.

1.2.3.1 Isoxazolium salts for targeting glutamic acid

Glutamic acids could serve as a desirable covalent target owing to their availability within protein binding sites (~5% compared to ~2% for cysteine).[84] Originally developed as an amide coupling reagent, the isoxazolium salt, Woodward's Reagent K, has been shown to covalently modify a variety of amino acids.[82] Work conducted by Martín-Gago *et al*. aimed to investigate their use for the selective labelling of E88 of the lipoprotein chaperone PDE6δ.[83]

The isoxazolium salts undergo a base-mediated ring-opening by appropriate functional groups within the protein, thus forming the ketenimine which is subsequently attacked by, in this case, glutamic acid to form a stable enol ether adduct (**Figure 1.18 A**). Having identified E88 as a potential target residue, Martín-Gago *et al*. incorporated an isoxazolium salt into a previously published reversible PDE6δ ligand **1.40** that displays low-nanomolar affinity (K_d= 8 nM) (**Figure 1.18 B**).[83] Optimisation led to compound **1.41** which engaged and stabilised PDE6δ in a cellular thermal shift assay (CETSA), while only minimal off-targets were observed— a promising result given the abundance of carboxylic acids in proteins.

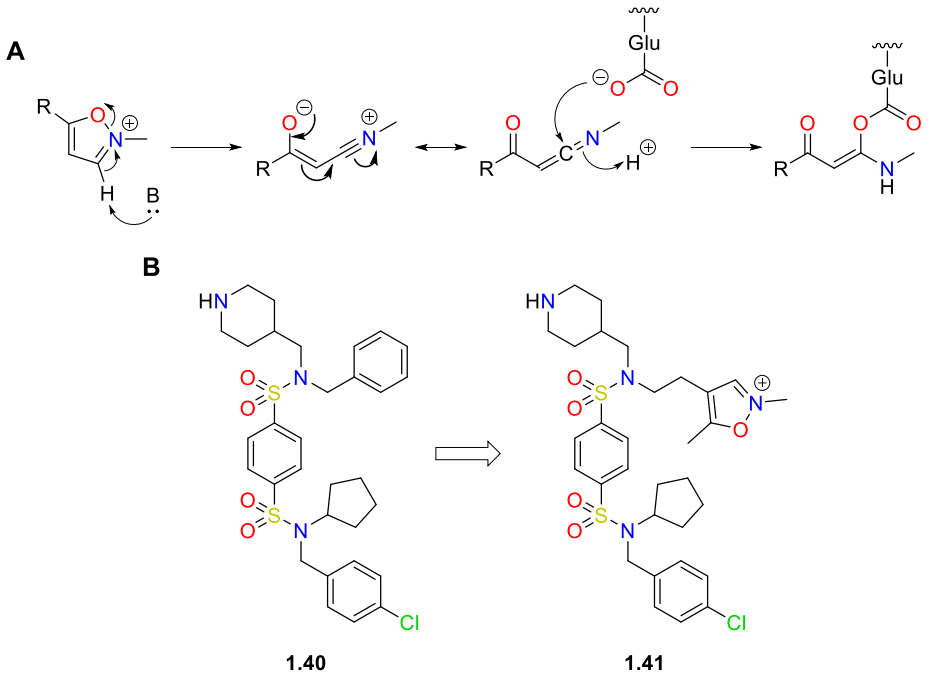


Figure 1.18. Isoxazolium salts can be used to covalently label glutamic acid residues. **A**) Proposed mechanism for the reaction between an isoxazolium salt and glutamic acid. It is hypothesised that basic functional groups within the protein are able to mediate the ring-opening. **B**) Non-covalent PDE6δ ligand **1.40** was developed into a covalent ligand **1.41** by replacing the phenyl ring with a reactive isoxazolium salt.[83]

1.2.3.2 Nitrile oxides for targeting selenocysteine

Selenocysteine is a cysteine analogue in which selenium replaces the sulfur. Although selenocysteine is a non-canonical amino acid, and is not coded for directly, it is present in many proteins, for example, the active sites of redox regulating proteins.[85] Selenocysteines have a lower pK_a value than cysteine, ~5.4 compared to ~8.6, respectively. Because of this, selenocysteines exist mostly in their deprotonated form at physiological pH, rendering them highly nucleophilic, and potential targets for covalent labelling.

Recent work by Eaton *et al*. uncovered the mode of action for a glutathione peroxidase 4 (GPX4) inhibitor **1.42**,[86] previously discovered in a phenotypic screen.[87] GPX4 is a phospholipid hydroperoxidase and a selenoprotein. Rounds of SAR determined that the nitroisoxazole was vital for GPX4 inhibition. Although **1.42** lacked no obvious reactive group, protein LC-MS studies in live cells revealed covalent labelling of GPX4, however, with an adduct corresponding to −41 Da less than expected, while covalent labelling in purified protein was not observed. Furthermore, CETSA

studies showed that **1.42** stabilised GPX4, but did not stabilise the selenocysteine-to-alanine GPX4 mutant ($GPX4^{U46A}$), indicating that covalent labelling was occurring at the selenocysteine. Eaton *et al*. next hypothesised that the vital nitroisoxazole was acting as a prodrug and may be undergoing a base-mediated ring opening to form an α-nitroketoxime **1.44** (**Figure 1.19**). The proposed intermediate **1.44** was synthesised and shown to induce the expected phenotype and label GPX4 in live cells. However, **1.44** was still unable to label purified GPX4, further indicating another metabolic transformation was required to form the active compound. Additional studies led to the discovery that **1.44** was metabolised to the nitrile-oxide **1.46**, which was confirmed to be the active species able to label selenocysteine and inhibit GPX4.

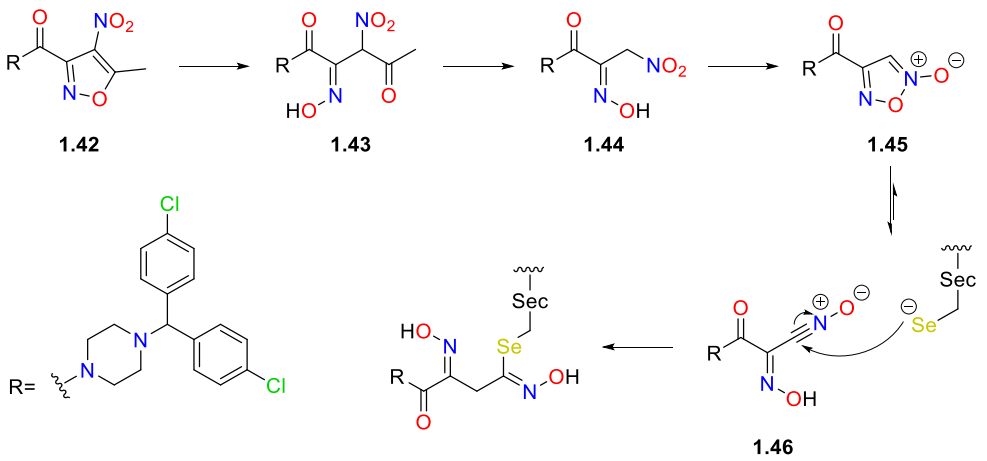


Figure 1.19. The proposed mechanism of action for covalent GPX4 inhibitor **1.42**. Compound **1.42** is activated by metabolic transformations that yield the reactive nitrile-oxide capable of inhibiting GPX4.[86]

1.2.4 Clickable probes for target engagement

Covalent probes can be modified with 'tagging' moieties that allow for proteome-wide selectivity studies and confirmation of cellular target engagement, which is important for target validation.[20,23,88] These tagging groups allow for the identification of interactions between the probe and proteins within the proteome (**Figure 1.20**).[27,28,88] When designing these probes, care must be taken to ensure that these modifications do not perturb the ligand binding to the protein so that accurate conclusions about target engagement can be made.

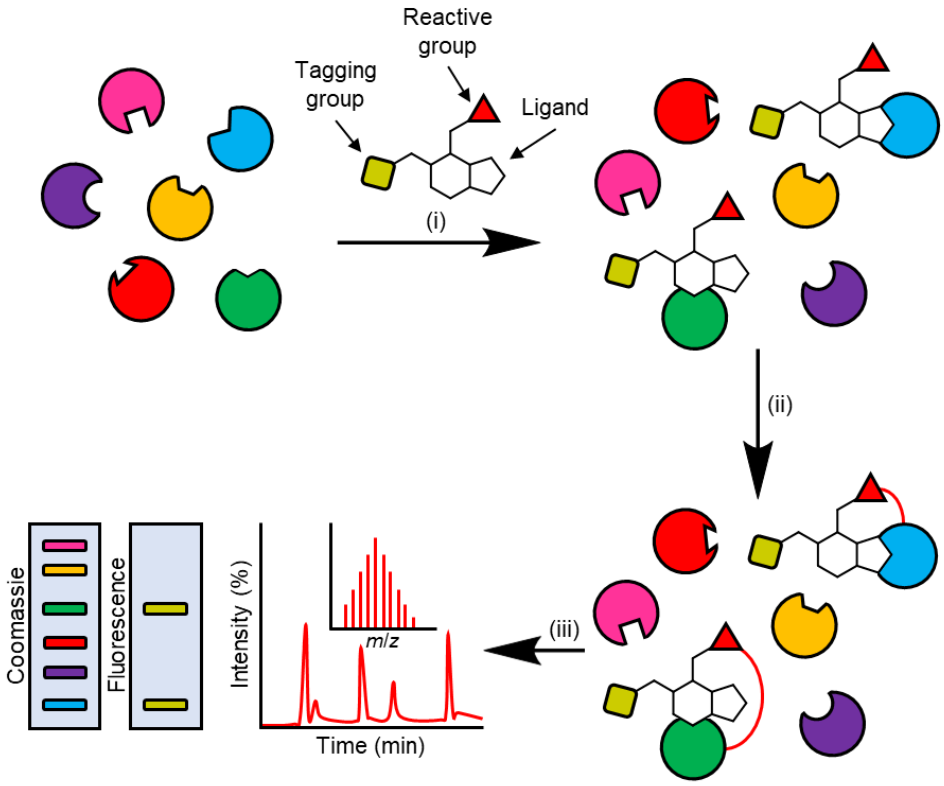


Figure 1.20. Schematic representation of how covalent clickable probes are used to validate targets and confirm target engagement: (i) Probe initially binds to interacting proteins through non-covalent interactions; (ii) A covalent bond forming reaction between the probe and interacting proteins occurs (iii) Tagging groups are then used to identify interacting proteins, either by LC-MS/MS or SDS-PAGE analysis.

1.2.4.1 Target identification

With the interacting proteins covalently bound to the probe, the tagging group allows for the identification of the bound proteins. Two main tagging approaches can be used (i) direct attachment of a fluorescent dye or an affinity tag, such as biotin, or (ii) the use of a silent handle, such as an alkyne or an azide (**Figure 1.21**).[20,27–29] Probes directly attached to fluorescent dyes allow for direct visualisation of labelled proteins through SDS-PAGE separation and fluorescent imaging, while probes attached to affinity labels can be directly enriched and analysed by protein LC-MS/MS. In contrast, probes bearing silent handles can be used to click on fluorescent dyes or affinity tags with temporal control and analysed as described above. The term 'silent handle' refers to the discrete nature of these moieties, which is of great importance when attempting to minimise the disruption to binding or significantly alter the physico-chemical properties of the probe, and hence is generally the preferred method.[23,29] Furthermore, the direct attachment of bulky, lipophilic fluorophores can impact cell permeability and cellular localisation.

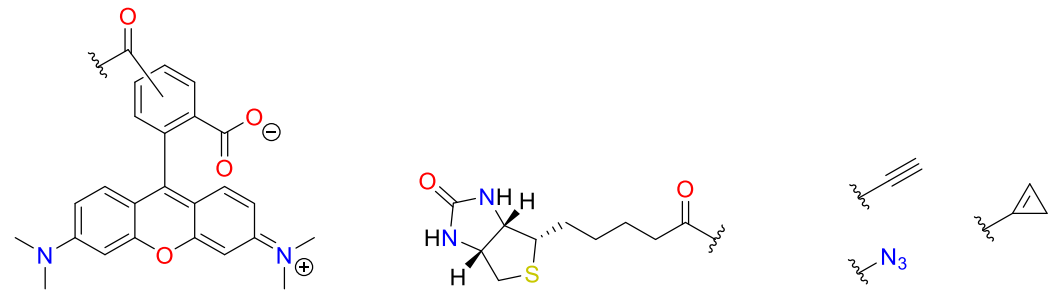


Figure 1.21. Commonly employed tagging groups used for protein identification. Some probes have fluorescent dyes or affinity tags directly ligated to the probe, however the use of silent handles for downstream tagging is most popular, due to the discreteness of the groups.

1.2.4.2 Bioorthogonal Click Chemistry

There are a number of criteria that are important when considering reactions in a biological setting (i) chemical and biological inertness, (ii) selectivity, (iii) kinetics, (iv) aqueous solubility, and (v) efficiency under physiological conditions.[89–91] Bioorthogonal click chemistry methods that meet these criteria have proved pivotal to chemical biology— allowing for the efficient ligation of fluorescent dyes or affinity tags to silent handles, *in vivo*, *in vitro*, and in cell lysate, through a variety of methods (**Table 1.5**).

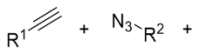
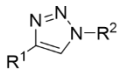
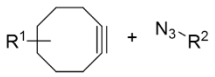
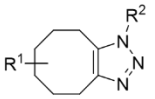
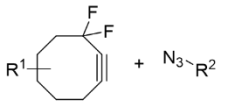
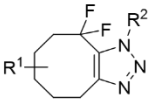
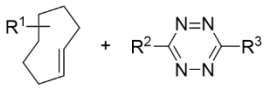
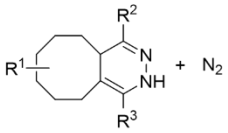
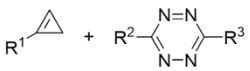
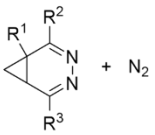
The most common click chemistry approach is the copper-catalysed alkyne azide cycloaddition (CuAAC) originally developed by Sharpless,[92] which formally proceeds *via* a 1,3-dipolar cycloaddition, and is popular due to its functional group tolerance and fast reaction kinetics. The major drawback of this method is the need for a Cu(I) catalyst, which makes it incompatible with live cell work, owing to toxicity.[93]

Copper-free methods that use strained alkynes have become a widely used approach for tagging probes and biomolecules in live cells. In 2004 the Bertozzi group published the discovery of the copper-free strain-promoted [3+2] azide−alkyne cycloaddition (SPAAC) between azides and strained cyclooctynes for selective modification of biomolecules, with no observable cytotoxicity, however, reaction kinetics were slower than those of CuAAC .[94] In 2007 the group published a follow up paper that utilised optimised cyclooctynes furnished with an electron withdrawing difluoro group.[95] This method increased sensitivity and gave faster reaction kinetics compared to the original strain promoted cyclooctynes. Interestingly, until recently these reactions were accepted to be primarily driven by the strain release of cyclooctyne, however, recent work from Sterling *et al*. suggests that delocalisation plays a more critical role in the reactivity of these cyclooctynes.[96]

Another metal-free method is the inverse electron demand Diels−Alder cycloaddition (iEDDA) of tetrazine with alkenes, most commonly *trans*-cyclooctene.[97–100] These reactions are compatible with biological systems and possess high reaction rates, however, the *trans*-cyclooctene groups are

bulky and can alter properties including solubility and cell permeability, as well as disrupting binding to the target. Methods utilising small and metabolically stable cyclopropenes have also been developed.[101,102]

Table 1.5. Examples of commonly employed bioorthogonal click chemistry reactions for functionalisation of silent handles.

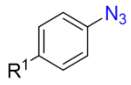
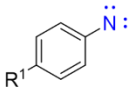
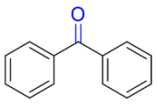
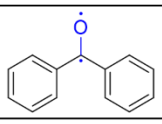
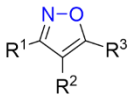
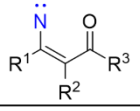
Reaction	Components	Product
Cu(I)-catalysed azide-alkyne cycloaddition (CuAAC)		
Strain-promoted azide–alkyne cycloaddition (SPAAC)		
		
Inverse electron demand Diels–Alder cycloaddition (iEDDA)		
		

1.2.4.3 Photoaffinity labelling

A similar approach for target engagement, known as photoaffinity labelling (PAL), can also be applied to probes that bind non-covalently to target proteins. Termed 'photoaffinity probes', these are non-covalent probes functionalised with a tagging moiety and a photoactivable group that, when irradiated with UV light, forms reactive species that can react with interacting proteins, covalently tethering the probe to the protein (**Table 1.6**). Due to the reactive and unstable nature of these species, only proteins directly interacting with the photoaffinity probe are labelled, enabling detailed proteome fingerprinting of interacting proteins.[27,103] Commonly employed

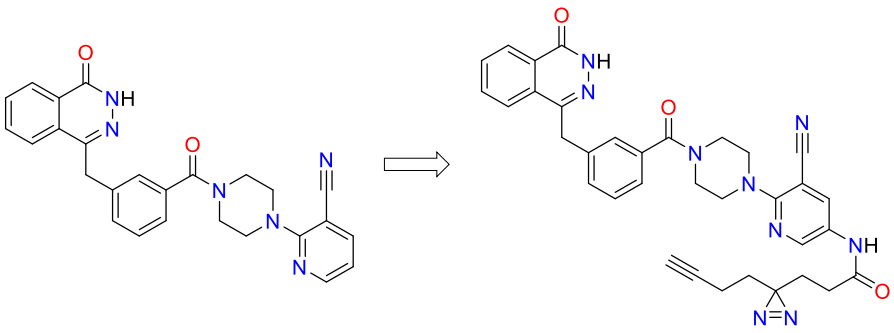
photoactivatable groups are aryl azides, benzophenones, and diazirenes (**Table 1.6**).[104–106] More recently, the isoxazole scaffold has also been added to the photo-crosslinking toolbox.[107] This is of interest as the isoxazole is a pharmacophore with a rich selection of targets and will aid in target engagement studies.[108]

Table 1.6. Examples of commonly used photoactivatable groups, their structure, and irradiation wavelength (nm) to form the reactive species.

Photoactivatable group	Structure	Irradiation (nm)	Reactive species
Aryl azide		254	
Benzophenone		365	
Diazirene		365	
		365	
Isoxazole		254	

1.2.4.4 Recent Examples

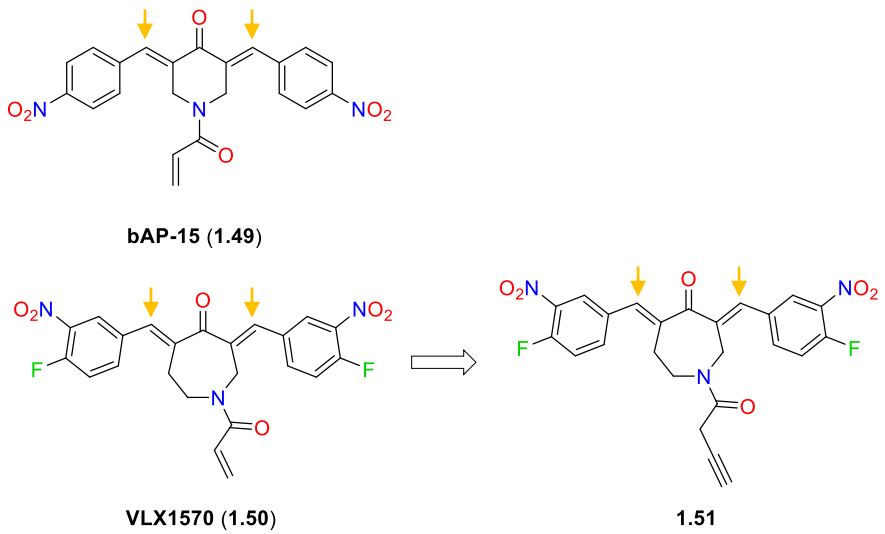
Both clickable covalent probes and photoaffinity probes have proven to be an important method in target validation and target engagement. A recent publication by Howard *et al.*,[109] showed the design and evaluation of a photoaffinity probe **PARPYnD** (**1.48**) based on the previously published triple poly(ADP-ribose) polymerase (PARP) 1/2/6 inhibitor **AZ9428** (**1.47**) (**Figure 1.22**).[110,111] The photoaffinity probe engaged with both PARP1 and PARP2 in live cells. However, whilst the photoaffinity probe and the parent inhibitor were able to engage and inhibit recombinantly expressed PARP6 protein, engagement with endogenously expressed PARP6 was not observed, confirming that **1.47** was not an inhibitor of cellular PARP6. This work demonstrates the importance of in-cell target engagement studies in the wider concept of target validation.



AZ9482 (**1.47**) **PARPYnD** (**1.48**)

Figure 1.22. The photoaffinity probe, **PARPYnD** (**1.48**), was used to study target engagement of **AZ9482** (**1.47**), a previously published triple PARP1/2/6 inhibitor. **1.48** exhibited target engagement with PARP1/2 but not endogenous PARP6.

Compounds **b-AP15** (**1.49**) and **VLX1570** (**1.50**) are structurally analogous and were developed as covalent inhibitors of the deubiquitinating enzymes (DUBs); USP14 and UCH-37 (**Figure 1.23**). In 2017, a Phase I/II clinical trial investigating **1.50** efficacy against multiple myeloma was put on hold, owing to dose-limiting toxicity. A comprehensive proteomic study by Ward *et al*. used a clickable analogue, **1.51**, to reveal that the **1.50** scaffold possesses intrinsic reactivity and was inducing non-specific protein aggregation— likely causing the observed toxicity (**Figure 1.23**).[112] This work therefore invalidated **1.49** and **1.50** as chemical probes, and that the phenotypes observed upon their use were heavily perturbed and not necessarily due to USP14 and UCH-37 inhibition.



bAP-15 (**1.49**)

VLX1570 (**1.50**) **1.51**

Figure 1.23. Chemical structures of **1.49**, **1.50**, and **1.51**. The orange arrows show sites prone to unintentional reactivity with nucleophiles.

1.3 Epigenetics

The term ‘epigenetics’ refers to heritable changes in phenotype that are not linked to a change in genotype.[113,114] These changes in phenotype are instead due to differences in gene expression, as a result of changes in chromatin structure, caused, in part, by post-translational modifications (PTMs) of histone proteins. PTMs refer to the reversible and dynamic modification of proteins that occur after biosynthesis. Epigenetically relevant PTMs include acetylation, methylation, and phosphorylation of amino acid residues of histones.[115] These epigenetic modifications are regulated enzymatically by (i) ‘writers’ which install modifications, (ii) ‘erasers’ which remove modifications, and (iii) ‘readers’ which recognise modification and signal for downstream processes.[116,117]

One of the most common epigenetic modifications is the acetylation of lysines (KAc) on histone tails— the proteins around which DNA is wrapped and packed (**Figure 1.24**).[115] Acetylation of histone tail lysine residues leads to the neutralisation of the charged amine side chain, and ultimately reduces the electrostatic interactions between the DNA and histones.[116] This leads to a more relaxed chromatin structure that is more accessible to transcriptional machinery. Additionally, KAc signals for the binding of chromatin-associated factors.[117] For lysine acetylation, the required ‘writers’ and ‘erasers’ are histone acetyltransferases and histone deacetylases, respectively. The ‘readers’ of lysine acetylation are a family of proteins known as the bromodomains.[117]

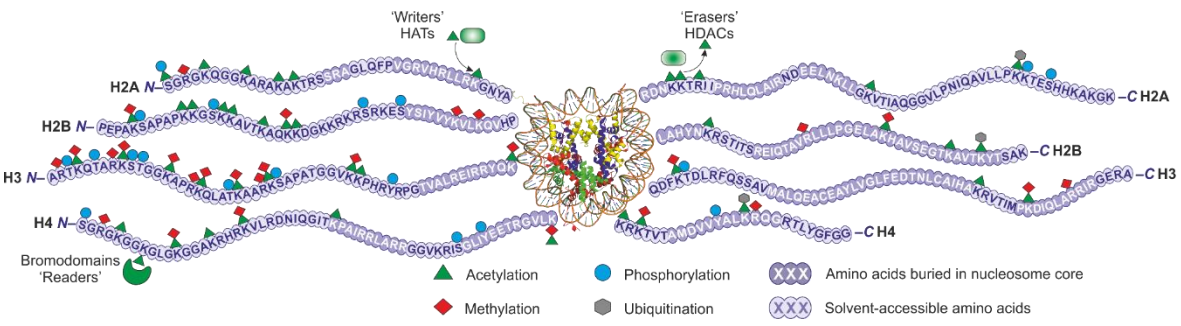


Figure 1.24. DNA is packaged and wrapped around histones. Protruding histone tails can be modified with PTMs, such as methylation, phosphorylation, and most commonly, acetylation. PTMs are introduced, removed, and read by ‘writers’, ‘erasers’, and ‘readers’, respectively. PTMs can alter the accessibility of DNA to transcriptional machinery, thus playing a role in gene regulation. Reprinted from Jennings *et al.*[118]

1.4 Bromodomains

First reported in 1992, and named after the *Drosophila* gene *brahma*,[119] the bromodomains are epigenetic reader proteins that recognise and bind to KAc residues on histone tails. The human genome codes for 61 distinct bromodomains that are distributed across 46 separate bromodomain-containing proteins (BCPs).[120]

Bromodomains are structurally conserved protein modules, of ~110 amino acid residues, that reside in larger protein structures, many of which are important in regulating gene transcription. The tertiary structure of the bromodomains is a characteristic bundle of four α-helices (αZ, αA, αB, and αC) connected by three flexible loops (ZA, AB, and BC loops) (**Figure 1.25**).

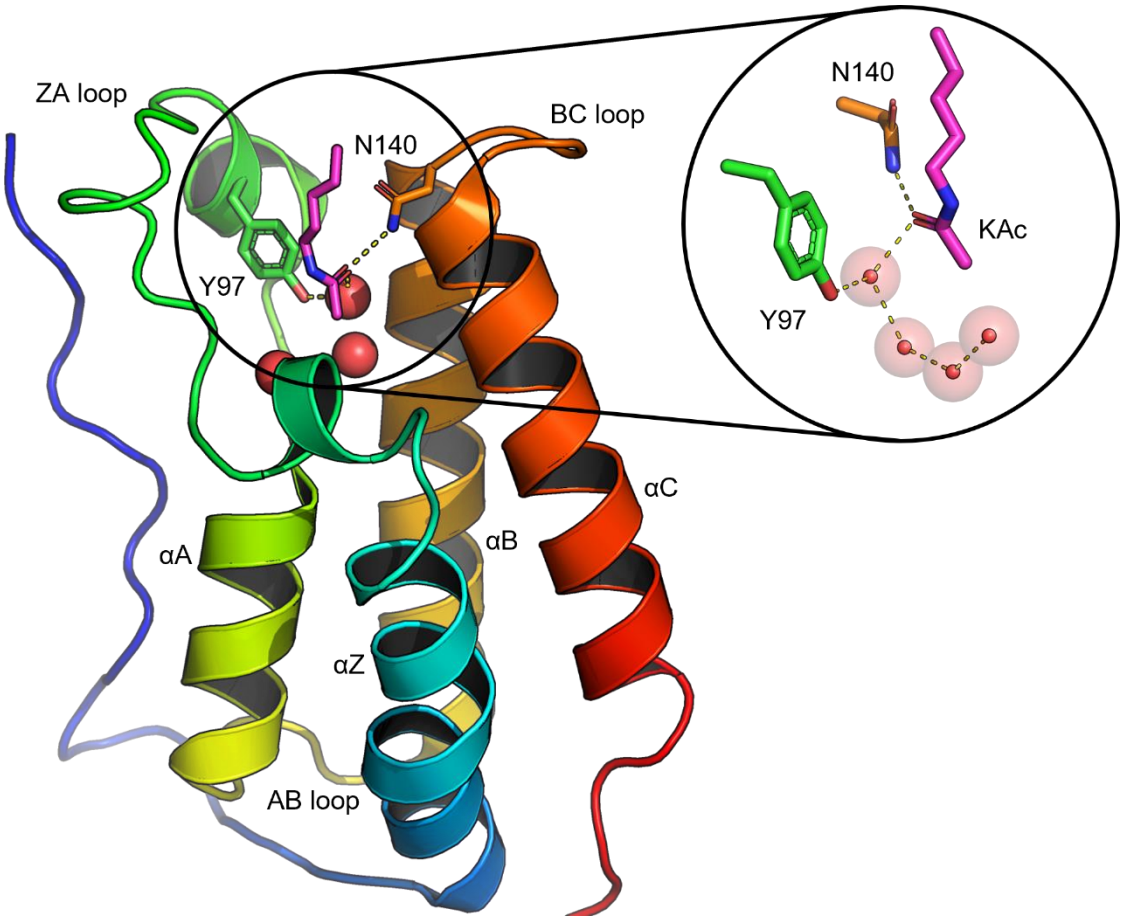


Figure 1.25. X-Ray crystal structure of KAc bound to BRD4(1) (PDB ID: 3UVX).[120] The protein structure is formed of four α-helices (αZ, αA, αB, and αC), and three flexible loop regions (ZA, AB, and BC loop). The KAc group (magenta) hydrogen bonds to N140 and Y97.[120]

The KAc binding pocket is located between the ZA and BC loops and is largely hydrophobic in nature, however, there are typically four conserved water molecules observed at the bottom of the pocket by X-ray crystallography (**Figure 1.25**). KAc binding to bromodomains occurs *via* two main interactions, (i) a direct hydrogen bond between a highly conserved αB asparagine amide and the KAc carbonyl, and (ii) a water mediated hydrogen bond between a conserved ZA loop tyrosine and the KAc carbonyl.[121]

1.4.1 Bromodomain-containing protein 4 (BRD4)

1.4.1.1 Structure and function

Bromodomain-containing protein 4 (BRD4) is part of the bromodomain and extra C-terminal domain (BET) family of BCPs, which also comprises BRD2, BRD3, and BRDT.[120] BET proteins are formed of two separate but conserved N-terminal bromodomains. Whilst the two bromodomains within a BET protein are structurally similar, X-ray crystallography and bioinformatics suggests that the most resemblance is found between the respective first or second sets of BET bromodomains, i.e BRD4(1) bears more resemblance to BRD2(1) than it does to BRD4(2).

There has been much interest in BRD4, as localisation of BRD4 to histones *via* KAc has been found to recruit gene regulating complexes such as the positive transcription elongation factor (P-TEFb) kinase.[122,123] Furthermore, it has been found that BRD4 is the only BET protein that is pan-essential in mammalian cells.[124]

1.4.1.2 Small molecule ligands for BRD4

Owing to their role as epigenetic readers and possible wider-cellular functions, the BET bromodomains are an exciting target. However, they are not yet fully understood. For example, it

is not known why these domains share such a high degree of homology, and why BRD2–4 are so widely expressed.[125]

As bromodomains are components within larger proteins, the use of techniques that remove the whole BCP cannot be used for the complete interpretation and validation of bromodomain function.[121] Therefore, there is a need for selective chemical probes that allow for the interrogation of single domain function. Over the past decade there has been much success in the development of small molecule ligands that inhibit bromodomain function, especially for the BET bromodomain family.

In 2010, Filippakopoulos *et al*. and Nicodeme *et al*. reported the first BET chemical probes.[126,127] These utilised a triazolobenzodiazepine scaffold, **(+)-JQ1** (**1.52**) and **I-BET762** (**1.53**), where the triazole group acted as a KAc mimic (**Figure 1.26**). Whilst these compounds exhibited potent inhibition of BRD4 they were unable to selectively distinguish between the BET bromodomains.[128] Subsequently, a photoaffinity analogue (**1.54**) of **1.52** and **1.53**, identified several hundred potential off-target proteins for these probes, including ubiquitin-specific-processing protease 7 (USP7) and DNA damage-binding protein 1 (DDB1).[129]

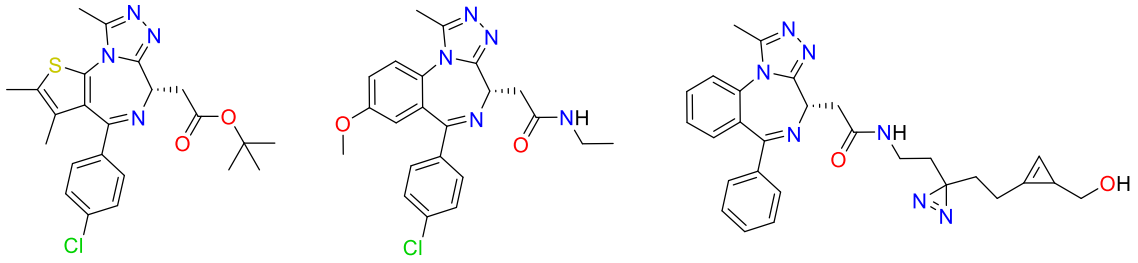


Figure 1.26. (+)-JQ1 (**1.52**) and **I-BET762** (**1.53**) were the first disclosed BET bromodomain inhibitors. To determine the selectivity profiles for the triazolobenzodiazepine-based BET inhibitors, a photoaffinity probe **1.54** was tested. It was found that the triazolobenzodiazepine scaffold interacts with many proteins.

The discovery that the 3,5-dimethylisoxazole (DMI) scaffold can act as a KAc mimic in 2011 by Hewings *et al*. has led to the development of a number of BET bromodomain ligands.[130] The origin of the DMI scaffold began from a fragment screen where the commonly used solvent, *N*-methyl pyrrolidinone (NMP, **1.55**), was found to interact with a range of bromodomain KAc binding pockets

(**Figure 1.27**). Subsequently, a variety of methyl-substituted heterocycles were investigated, and a dihydroquinazolinone-containing DMI **1.56** was found to exhibit low micromolar IC_{50} values for BRD2(1) and BRD4(1). Interestingly, this fragment bears two potential KAc mimics (the dihydroquinazolinone and the DMI), and so X-ray crystallography was employed to decipher which one was occupying the BRD4(1) KAc binding site. The obtained structure (PDB ID: 3SVF) confirmed that the DMI scaffold was binding to the KAc binding pocket by retaining the key interactions seen for KAc—the direct hydrogen bond to N140 and the water mediated hydrogen bond to Y97 (**Figure 1.27**).

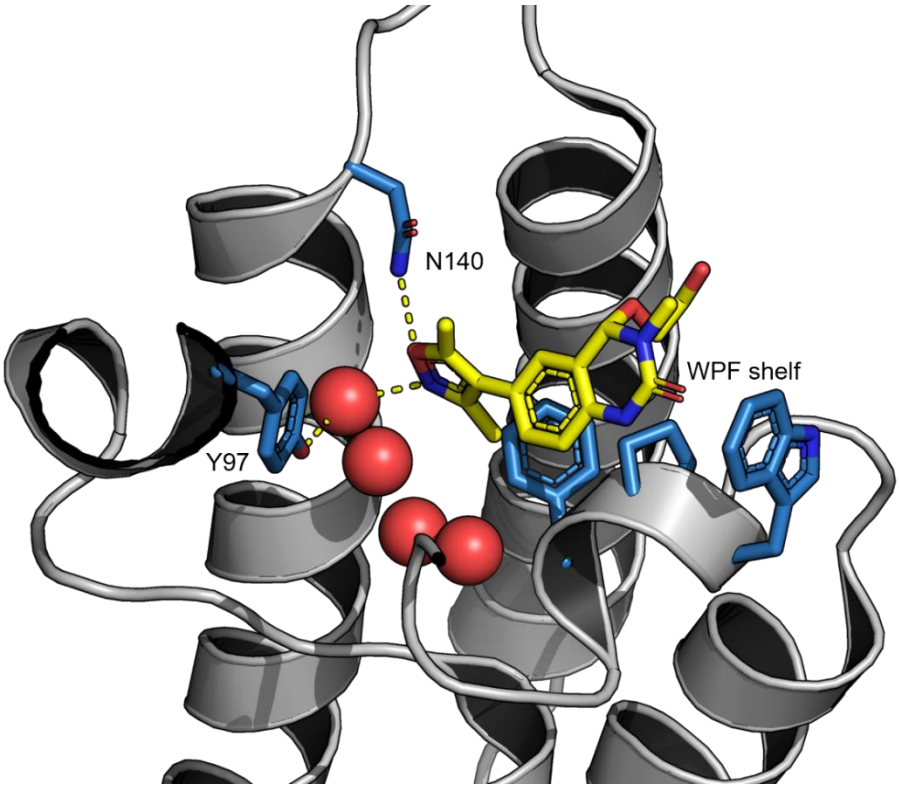


Figure 1.27. X-Ray crystal structure of DMI fragment **1.56** bound to BRD4(1) (PDB ID:3SVF). The DMI acts as a KAc mimic by hydrogen bonding directly to N140, and indirectly to Y97. The lipophilic WPF shelf is shown.[130]

SAR studies on DMI-based ligands led to the exploration and exploitation of the hydrophobic WPF shelf in order to gain affinity, yielding **OXFBD02** (**1.57**) and **OXFBD04** (**1.58**) as high affinity BRD4(1) ligands (**Figure 1.28**).[131,132] BROMOscan screening revealed that **OXFBD02** (**1.57**) and **OXFBD04** (**1.58**) were in fact high affinity pan-BET bromodomain inhibitors that also possessed moderate affinity for the non-BET bromodomains of CREBBP and EP300.[132]

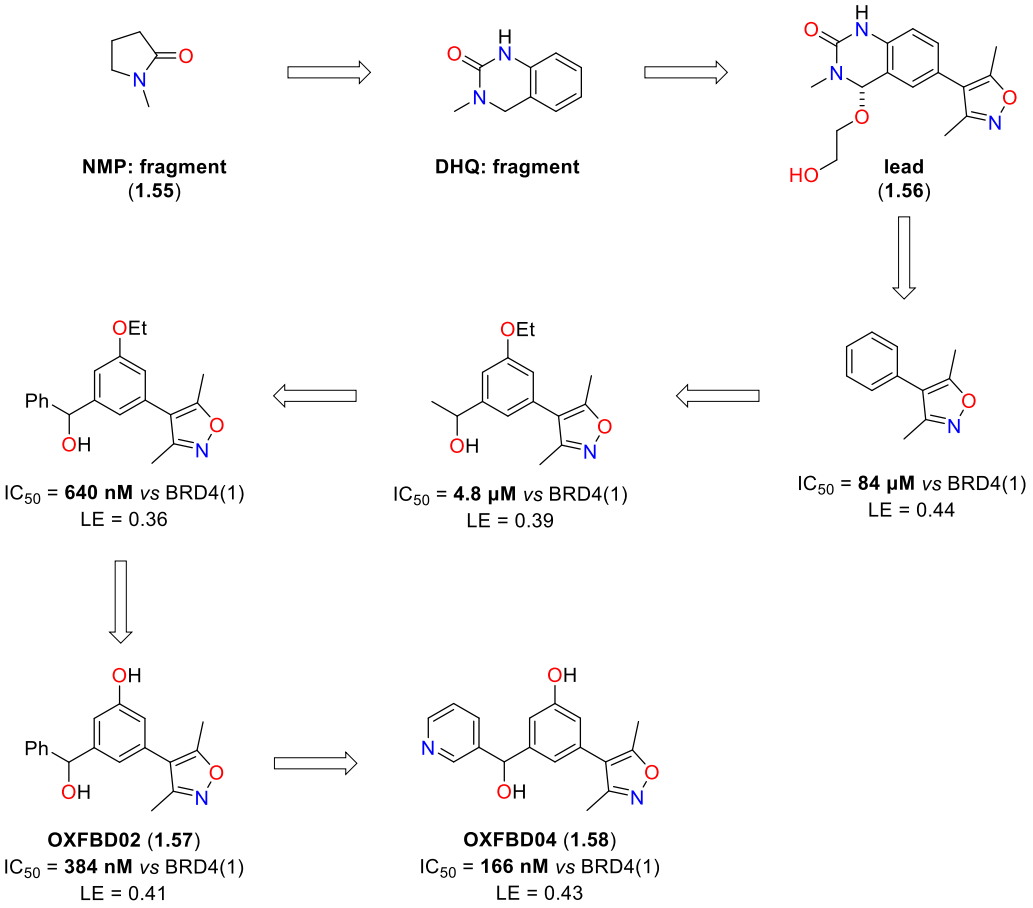


Figure 1.28. The development of **OXFBD02** (**1.57**) and **OXFBD04** (**1.58**). **1.55** was identified in fragment screen, which led to the screening of methyl-substituted heterocycles. An isoxazole-bearing fragment **1.56** was discovered, and binding confirmed by crystallography. Ligand development focused on increasing lipophilic interactions with the WPF shelf of BRD4(1), leading to the development of **1.57** and **1.58**.[130–132]

Although many of these first-generation ligands were developed to target BRD4, the reality is that most are pan-BET ligands and bind indiscriminately to all BET bromodomains with similar affinity.[125] Whilst some pan-BET inhibitors have shown promise in clinical trials for the treatment of cancers, others have failed due to safety profiles.[133] It is now thought that single domain BET inhibitors may lead to improved safety profiles and decreased off-target effects for therapeutic BET inhibitors.[133] From a target validation perspective the use of pan-BET inhibitors is not ideal for the interpretation of single bromodomain function. Much work has gone into the development of methods that allow for selective tuning of these pan-BET ligands to enable targeting of a single BET bromodomain.

In 2014 Baud *et al*. utilised Shokat's pioneering 'bump & hole' approach to selectively probe BET bromodomain functions (**Figure 1.29**).[128] Analysis of the pan-BET ligands, **(+)-JQ1** (**1.52**) and **I-BET762** (**1.53**), bound to BRD4(1) revealed that mutation of the conserved L94 residue (BRD4

numbering) to a smaller alanine residue may generate a 'hole' that could accommodate a bulky 'bump' modification to **1.52** and **1.53**, such as a methyl or ethyl group. The 'bumped' ligands would sterically clash with and therefore not bind to, the wildtype proteins, while binding to the mutant protein. This approach led to high levels of selectivity for the mutant proteins over the wildtype proteins, and the discovery that selective inhibition of BRD4(1) is sufficient for displacement of BRD4 from chromatin.

In a follow-up publication, Runcie *et al*. optimised the 'bump & hole' approach for the BET bromodomains, by utilising a more functional mutant.[134] These experiments confirmed that the first BET bromodomains are more important for chromatin binding than the second BET bromodomains. However, Runcie *et al*. suggested that while BRD4(2) does not play a significant role in chromatin binding it is still vital for gene expression, possibly through the binding of non-histone proteins. Whilst the 'bump & hole' method has provided a useful insight into the biology of bromodomains, it has been shown that compounds such as **1.52** and **1.53** interact with many other proteins, and so, it is not unreasonable to postulate that some of the phenotypic effects observed may stem from interactions with other proteins.

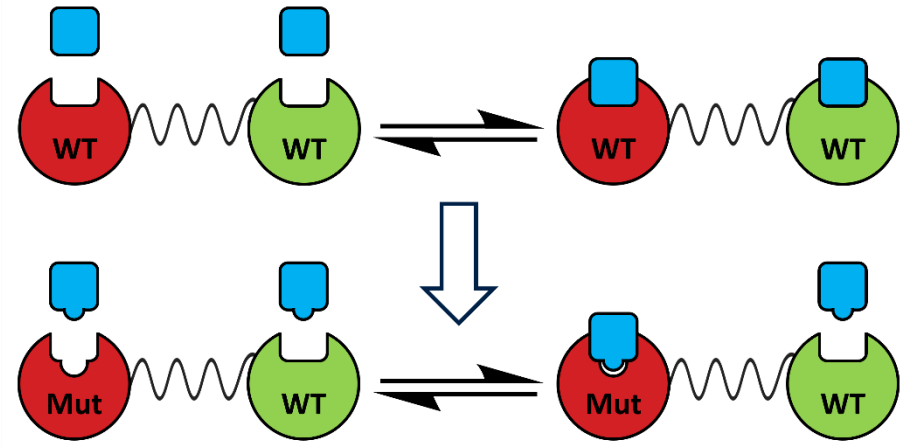


Figure 1.29. Schematic representation of how the 'bump & hole' approach works. Mutation of an amino acid residue for a smaller residue can form a 'hole' near the proteins binding pocket. A ligand is modified with a chemical 'bump', that can be accommodated by the 'hole'. The bulky 'bump' compound is unable to bind effectively to the WT and so displays selectivity for the mutant protein.

In 2018, Kharenko *et al*. designed and characterised covalent BRD4(1) ligands with the aim of BRD4(1) inhibition outlasting inhibition of other BET bromodomains in live cells (**Figure 1.30 A**).[135]

Crystallographic studies of the non-covalent DMI ligand **ZEN-2759** (**1.59**) bound to BRD4(1) revealed that the benzyl group was proximal to M149, which resides close to the binding pocket. By utilising a novel epoxide reactive group attached to the benzyl, Kharenko *et al*. were able to covalently label M149 with **ZEN-3129** (**1.60**) (**Figure 1.30 B**). Biophysical studies showed that the covalent analogues retained non-covalent binding to the BET bromodomains, however, only covalent binding to BRD4(1) was observed. By employing washout experiments with live cells, it was confirmed that covalent inhibition of BRD4(1) led to prolonged antiproliferative effects and decreased mRNA expression of BET-dependent genes.

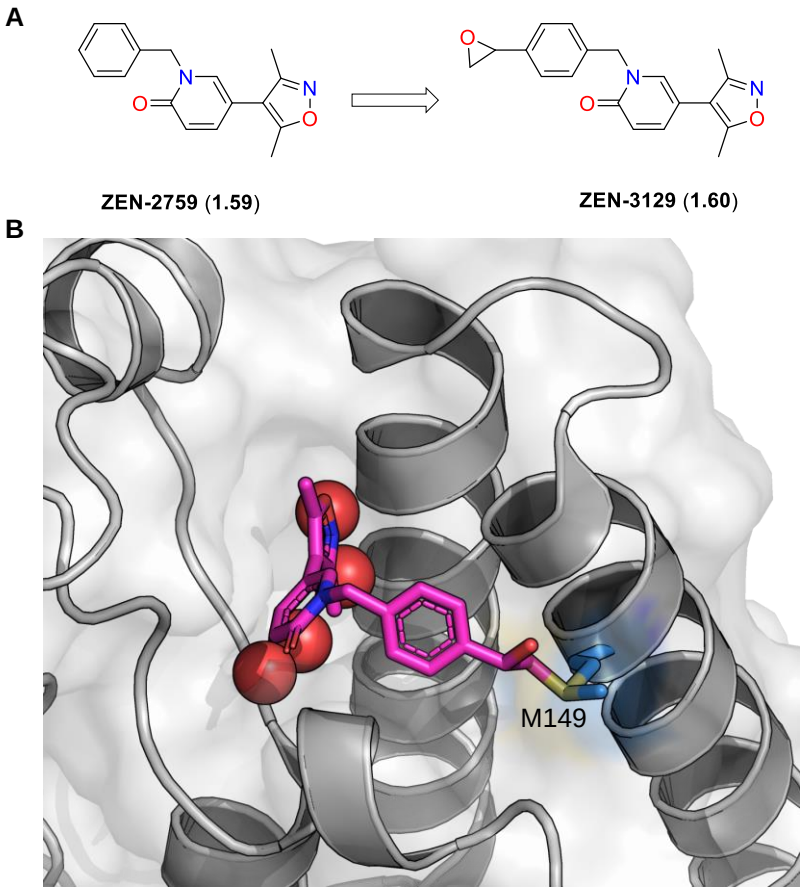


Figure 1.30. Development of covalent BRD4(1) ligands that selectively label M194. **A**) Kharenko *et al.* modified the non-covalent BRD4(1) ligand **1.59** by installing a reactive epoxide group **1.60**. **B**) X-Ray crystal structure of **1.60** (carbon = pink) covalently bound to BRD4(1) through the reaction with M149 (PDB ID: 6CZU).[135]

More recently GSK have published a number of papers surrounding the development of pan-BET first domain (pan-BD1) and pan-BET second domain probes (pan-BD2).[125,136–139] **iBET-BD1** (**1.61**), based on **I-BET151** (**1.62**),[140] uses a DMI KAc mimic and was developed as a low nanomolar pan-

BD1 inhibitor, with over 130-fold selectivity for BD1 (**Figure 1.31**). Similarly, **iBET-BD2** (**1.63**) utilises a simple acetamide KAc mimic and displays over 300-fold selectivity for BD2 with low nanomolar IC_{50} values (**Figure 1.31**).

Gilan *et al*. subsequently investigated the phenotypic effects of **1.61** and **1.63** on a range of human cancer cell lines.[125] It was found that treatment with **1.61** phenocopied the effects of treatment with pan-BET inhibitor **1.62** in cancer cells, leading to cell cycle arrest, inhibition of proliferation and apoptosis. This suggests that inhibition of the first bromodomains may be sufficient for anti-cancer therapies. For pan-BD2 inhibitor, **1.63**, the effects were less pronounced, further confirming that inhibition of the second BET bromodomains is insufficient for displacing BET bound chromatin. It was found, however, that the second BET bromodomains are vital for rapid gene induction after stimulus, raising the possibility that pan-BD2 inhibitors may be of use for the study or treatment of immunoinflammatory damage. Additionally, Gilan *et al*. found that all BET proteins, rather than just BRD4, were required for efficient gene induction.[125] Since this publication, GSK have released the structure of **GSK789** a pan-BD1 ligand with ~1500-fold selectivity for BD1,[136] and **GSK097** a pan-BD2 ligand with ~2000-fold selectivity for BD2.[138]

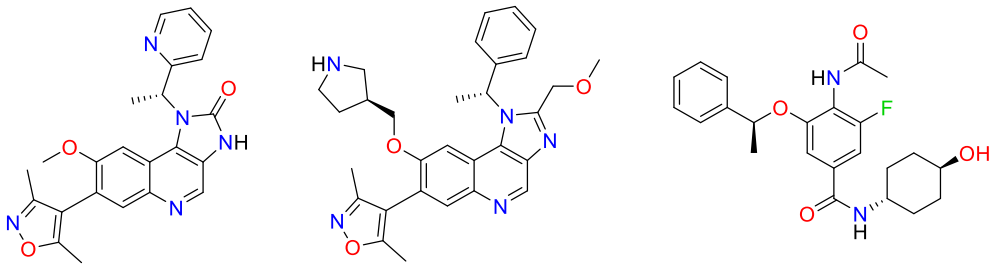


Figure 1.31. Chemical structures of **I-BET151**, **iBET-BD1**, and **iBET-BD2**.[140]

The chemical toolbox for probing the bromodomains has expanded greatly over the past decade, and the development of pan-BD1 and pan-BD2 ligands has helped further investigate the functions of bromodomains. However, the development of single domain-selective ligands remains the ultimate goal for probing the complex biology of the bromodomains.[133]

1.5 Project Aims

As outlined in **Chapter 1.1**, there is a clear need for robust target validation methods to enable faster drug discovery. The work in this thesis aimed to develop a method that allows for rapid target validation of epigenetic proteins. Termed 'mutate & conjugate', this method combines reactive covalent fragments with site-directed-mutagenesis of a target protein. The protein of interest is mutated to incorporate a carefully located nucleophilic cysteine residue close to the binding pocket, while also retaining wild type function and stability. Electrophilic reactive fragments that selectively and covalently bind to the mutant will therefore irreversibly inhibit the function of the protein, and provide information on the target proteins cellular function.

Using BRD4(1) as a proof-of-concept study, the work described here aimed to:

- Design, express, and validate a functional and stable BRD4(1) mutant capable of undergoing covalent labelling.
- Use both a structure guided, and a fragment screening approach to identify reactive fragments that are able to covalently inhibit the mutant, but not the wild type.
- Validate proteome-wide selectivity of the covalent fragments and confirm target engagement with the mutant BRD4(1).
- Explore the applicability of the 'mutate & conjugate' methodology to the development of selective protein degraders.

If this proof-of-concept study is found to work, it is envisioned that this method could be applicable to the rapid validation of other bromodomains, or poorly understood targets.

Chapter 2

Design, expression, and structural validation of a cysteine mutant bromodomain

2.1 Introduction and aims

The starting point for this project was to identify whether there were any reactive cysteine residues within BRD4(1)^{WT}. This was important to determine, as reactive native cysteines might make selective labelling of the mutant cysteine more challenging. Sequence analysis shows that there are two native cysteine residues, C125 and C136, within BRD4(1)^{WT} (**Figure 2.1 A**). C136 is highly conserved between all the BET bromodomains, while C125 is only conserved within the first BET bromodomains. To ascertain the locations of these cysteines, the X-ray crystal structure of BRD4(1)^{WT} (PDB ID: 6FSY) was examined.[132] Analysis shows that C125 is buried within the protein structure, however, C136 is located at the back of the KAc binding pocket and therefore might pose an issue for selectively labelling a mutant cysteine residue if the fragment interacts with the KAc binding pocket (**Figure 2.1 B** & **C**).

To ensure that C125 and C136 are not readily available to undergo covalent bond formation with a reactive fragment, Edward Coombs (University of Oxford) synthesised two small DMI-based reactive fragments, **2.1** and **2.2**, and incubated them separately with BRD4(1)^{WT} (**Figure 2.1 D**).[141] After 20 hours no covalent labelling of the protein was observed for either fragment by protein LC-MS. This indicated that neither of the native cysteines are hyper-reactive and that selective labelling of a mutant cysteine residue might be achievable. The work described in this chapter aimed to:

- Use computational methods to identify a residue suitable for mutation to a nucleophilic cysteine residue.
- Express and purify the mutant bromodomain.
- Validate the mutant as a structurally stable and functional bromodomain.

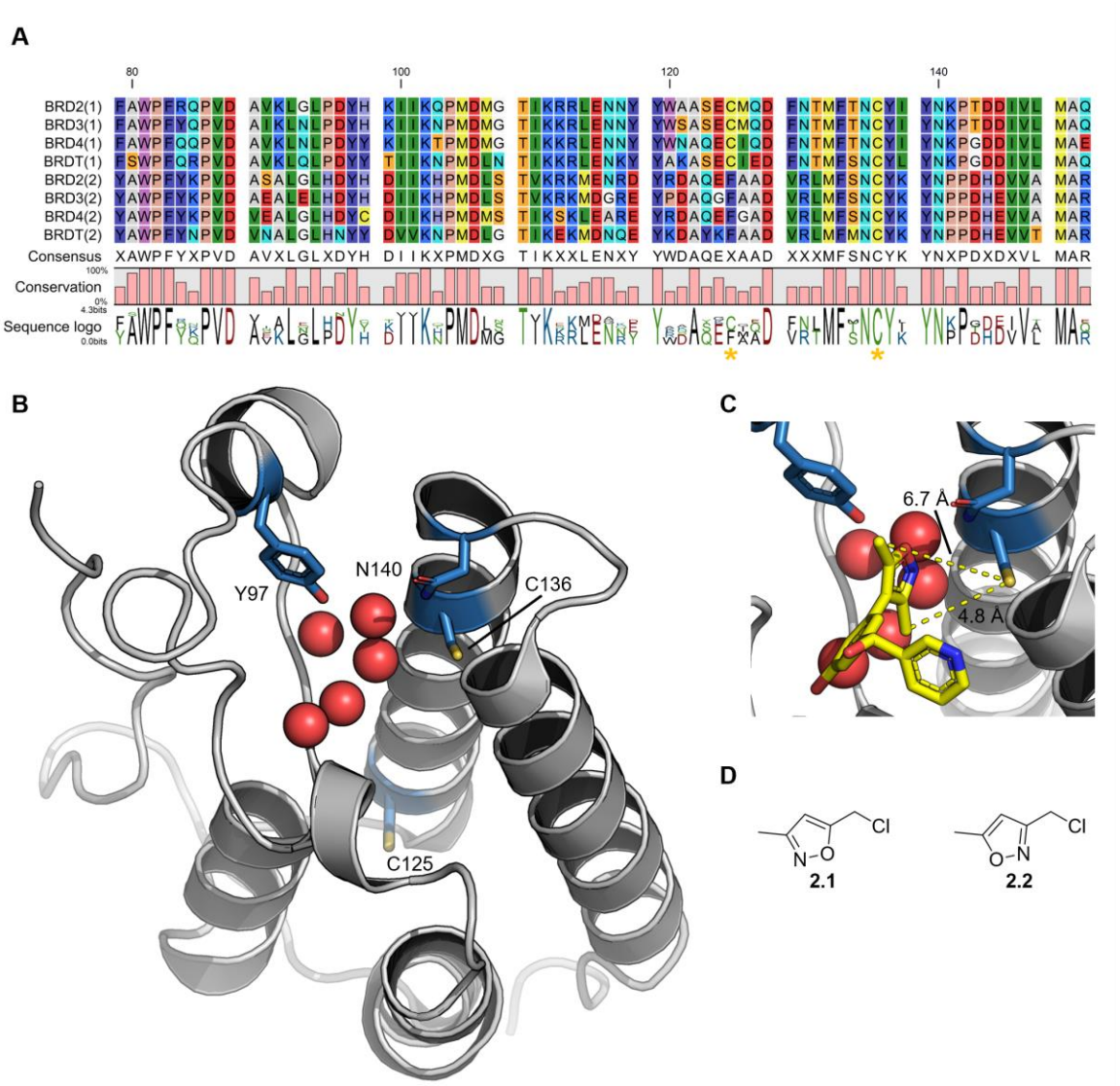


Figure 2.1. Initial work by Edward Coombs showed that BRD4(1)^{WT} did not possess any readily reactive cysteine residues. **A**) Protein sequence alignment of the BET bromodomains (BRD4(1) numbering). Sequences were aligned using CLC Genomics Workbench 22. The orange asterisks show C125 and C136. **B**) X-Ray crystal structure of BRD4(1)^{WT} with the key binding residues Y97 and N140 highlighted, as well as the two native cysteines, C125 and C136 (PDB ID: 6FSY).[132] **C**) The DMI-based ligand **1.58** bound to BRD4(1)^{WT} with the distances between the methyl groups and C136 highlighted. **D**) Small reactive isoxazole fragments **2.1** and **2.2** displayed no covalent labelling of BRD4(1)^{WT} in a protein LC-MS assay.

2.2 Mutant protein design

Whilst it was desirable for the mutant cysteine residue to be close to the KAc binding pocket, it was also considered how the mutation would alter the structural integrity of the protein, its KAc binding pocket, and ultimately its cellular function. Because of the roles BRD4 plays within a cell, it was vital

that the mutation did not significantly perturb chromatin binding to BRD4(1), as this would greatly impact overall BRD4 function. [125,128,142,143]

To determine possible sites of mutation, the X-ray crystal structure of **OXFBD04** (**1.58**) bound to BRD4(1)^{WT} (PDB ID: 6FSY) was examined (**Figure 2.2 A**).[132] Compound **1.58** is a BET BRD ligand, possessing a K_d for BRD4(1)^{WT} of 247 nM (determined by isothermal titration calorimetry (ITC)), and an IC_{50} value of 166 nM (determined by AlphaScreen™). The DMI acts as a KAc mimic and facilitates two key interactions within the pocket: a hydrogen bond with N140 and a water-mediated hydrogen bond with Y97 (**Figure 2.2 B**). While the key interactions occur in the binding pocket, the affinity of **1.58** is also driven by interactions with the WPF shelf and the ZA-channel waters. The WPF shelf is a lipophilic region which the pyridine ring of **1.58** occupies. SAR studies have shown that lipophilic interactions in this region boost ligand affinity. Hydrogen bonding between the phenol and the ZA-channel waters further increases the stability of the binding. To ensure that mutant BRD4(1) functionality would be a sufficient substitute for BRD4(1)^{WT} within a cellular environment, it was key that these essential KAc binding pocket interactions were retained.

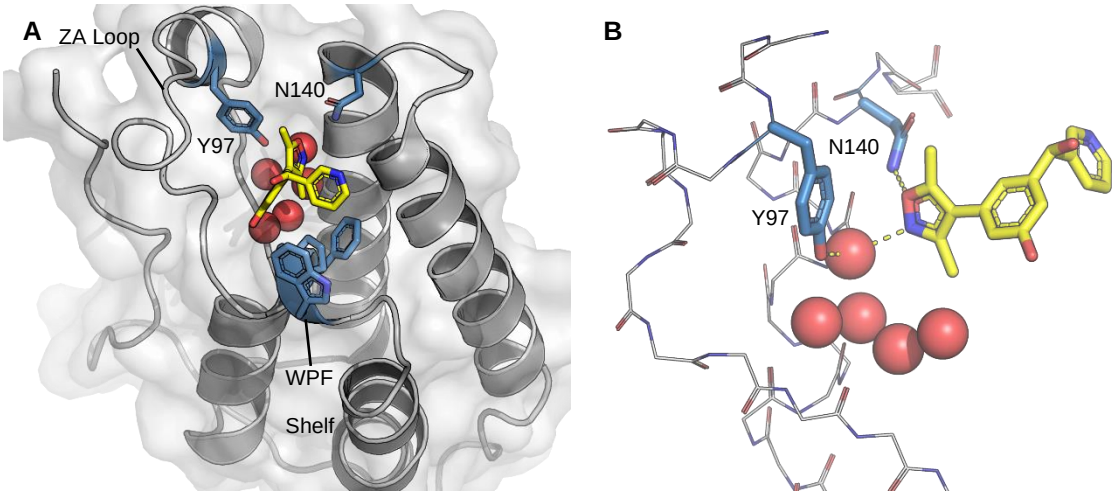


Figure 2.2. The DMI of **OXFBD04** (**1.58**) acts as a KAc mimic and binds to BRD4(1)^{WT}. **A**) **1.58** bound to BRD4(1)^{WT} (PDB ID: 6FSY).[132] **B**) Interactions within the KAc binding pocket. The isoxazole hydrogen bonds to N140 and makes a water mediated hydrogen bond to Y97.

2.2.1 PyMOL mutagenesis

Analysis of the X-ray structure (PDB ID: 6FSY) suggests that the ZA loop region, above the binding pocket, might be a suitable area for mutagenesis (**Figure 2.2 A**). There are two reasons for this: (i) a mutation in this region is less likely to perturb protein function compared to that of a mutation within the binding pocket, where alterations to the structure are more likely to impact KAc binding; and (ii) the region appears to align well with the trajectory of the isoxazoles 5-methyl group, suggesting that attachment of an electrophile at the 5-position might spatially position the cysteine and electrophile favourably for reaction. Using the X-ray crystal structure and the PyMOL mutagenesis wizard, mutation of the ZA loop residues were considered: L92, N93, L94, P95, and D96. Suitability of the mutation was based on the most likely rotamer, which was determined using the PyMOL relative strain scores, frequency in proteins, and visual inspection of steric clashes with neighbouring residues.

Due to proline residues' unique cyclic backbone, they are usually found in loops and β-turns because its structure dictates favourable ϕ and ψ angles. It was therefore hypothesised that mutation of P95 would have a detrimental impact on protein structure. D96 appeared to play a crucial role in local protein structure, indicated by side chain hydrogen bonding to the adjacent Y139, and formation of intramolecular hydrogen bonds to its amide backbone. Due to the perceived importance of the D96 side chain it was decided to not investigate it further. Whilst the solvent-exposed side chain of N93 appeared to not contribute to hydrogen bonding, mutagenesis in PyMOL anticipated that the most favourable cysteine rotamer would position the thiol pointing away from the KAc binding pocket, with a distance of 11.6 Å to the isoxazole. This did not seem like a feasible distance for formation of a covalent bond (**Figure 2.3 A** & **B**).

Because of leucine's innate hydrophobicity, these residues are usually buried within a protein's structure, where their isobutyl side chains can participate in hydrophobic interactions. Interestingly, both the side chains of the conserved L92 and L94 residues are solvent exposed,

suggesting that they may be less crucial to protein structure and stability, and thus making them attractive residues to investigate. When L92 was mutated in PyMOL a distance of 6.4 Å from the thiol sulfur to the 5-methyl carbon was predicted, however, the most likely rotamer was angled away from the KAc binding pocket, suggesting that it might be hard to target (**Figure 2.3 C** & **D**). L94 appeared to be positioned almost directly above the 5-methyl group of the isoxazole (**Figure 2.3 E** & **F**). Mutagenesis in PyMOL indicated that the most favourable rotamer would position the cysteine thiol 5 Å distance away from the 5-methyl of the isoxazole.

It is noteworthy that L94 is conserved across the BET proteins,[128] which may indicate a more prominent role in structure than initially expected. Interestingly, mutation of L94 to an alanine has been found to decrease tetra-acetylated H4 peptide binding affinity by up to 10-fold, but importantly does not abolish binding.[128,134,144] It is also important to note that these mutation studies have focused on an alanine substitution, which is smaller than the proposed cysteine, and so will have larger effects on ligand and peptide binding. For example, work by Runcie *et al*. found that a more conservative decrease in residue size, L94V mutation, resulted in only a two-fold decrease in peptide affinity.[134]

Overall, L94 offered the best opportunity for mutation; owing to its positioning near the KAc binding pocket, its alignment with the 5-methyl group of the DMI, and the fact that its mutation has been shown to be structurally tolerated. It was therefore decided to investigate BRD4(1)^{L94C} further.

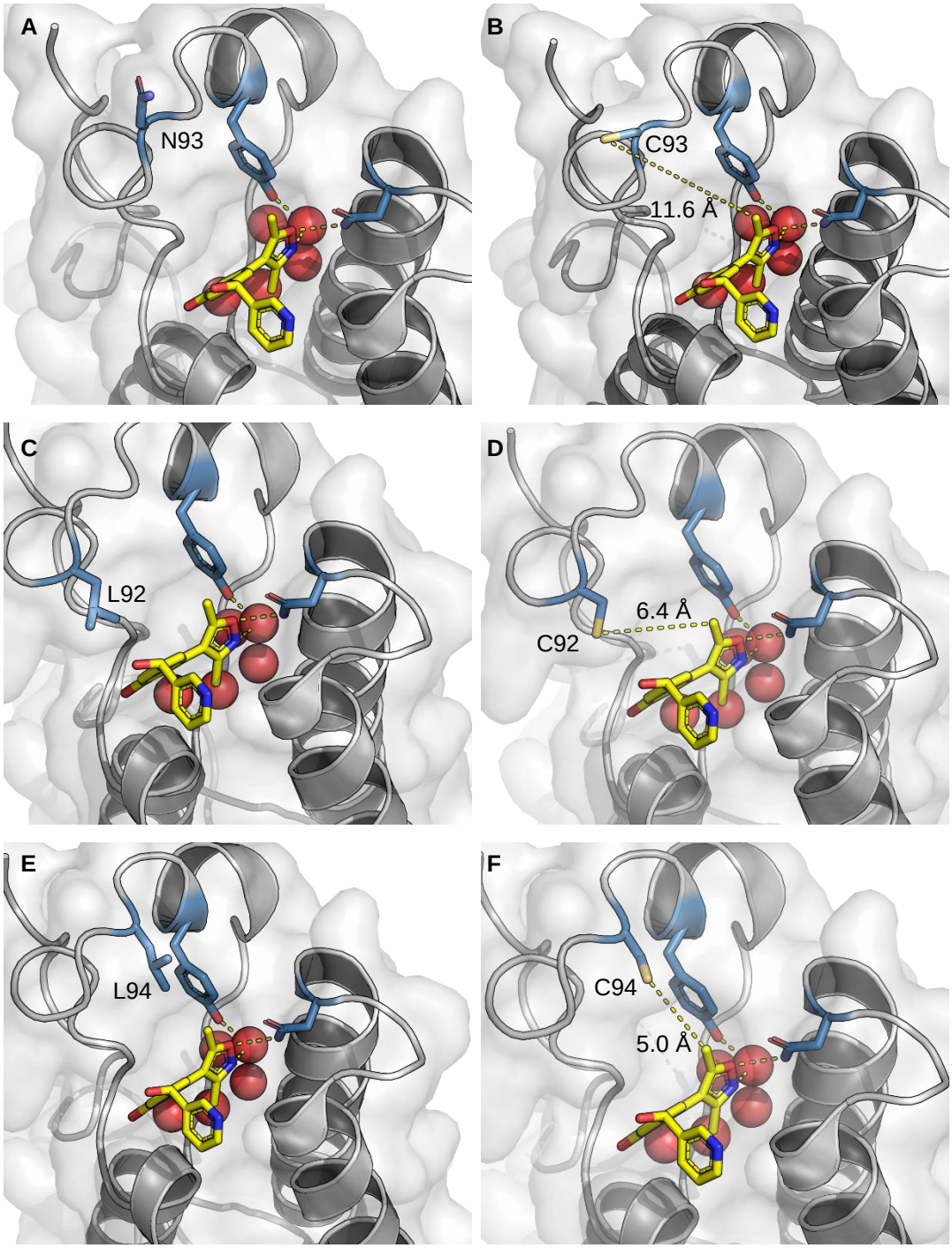


Figure 2.3. Investigation of the X-ray crystal structure of **(*R*)-OXFBD04** (**1.58**) (carbon = yellow) bound to BRD4(1)^{WT} (carbon = blue) (PDB ID:6FSY) suggested that the ZA loop region may be optimal for protein mutation. **A**) **1.58** bound to BRD4(1)^{WT} with N93 and the key binding residues highlighted. **B**) Mutagenesis in PyMOL predicted that the most likely cysteine rotamer would place the sulfur 11.6 Å away from the 5-methyl carbon. **C**) **1.58** bound to BRD4(1)^{WT} with L92 and the key binding residues highlighted. **D**) Mutagenesis in PyMOL predicted that the most likely cysteine rotamer would place the sulfur 6.4 Å away from the 5-methyl carbon. **E**) **1.58** bound to BRD4(1)^{WT} with L94 and the key binding residues highlighted. **F**) Mutagenesis in PyMOL predicted that the most likely cysteine rotamer would place the sulfur 5.0 Å away from the 5-methyl carbon.

2.2.2 Molecular dynamics simulations

To assess whether the proposed L94C mutation might result in an unstable protein, Dr Joseph Bluck (University of Oxford) carried out molecular dynamics (MD) simulations using GROMACS 2016.4. The BRD4(1)^{WT} (PDB ID: 3SVG) structure was mutated in PyMOL to give BRD4(1)^{L94C},[130] and simulations comparing the backbone root-mean-square deviation (RMSD) and backbone root-mean-square fluctuation (RMSF) of BRD4(1)^{WT} and BRD4(1)^{L94C} were performed over 50 ns. The RMSD value can be interpreted as the deviation of the protein's backbone compared to the original X-ray structure over time. From the simulation it can be seen that the calculated RMSD values for both BRD4(1)^{WT} and BRD4(1)^{L94C} were comparable, with deviations stable and remaining below 2.5 Å for BRD4(1)^{L94C} after equilibrium had been reached (**Figure 2.4**). The RMSF value allows for further investigation into the protein's structure and measures the fluctuation of each amino acid residue over the course of the simulation. Again, both proteins were comparable, with low levels of fluctuation overall (**Appendix A**). Taken together, the results from the MD simulations supported the hypothesis that the L94C mutation would not significantly disrupt the stability or structure of the protein.

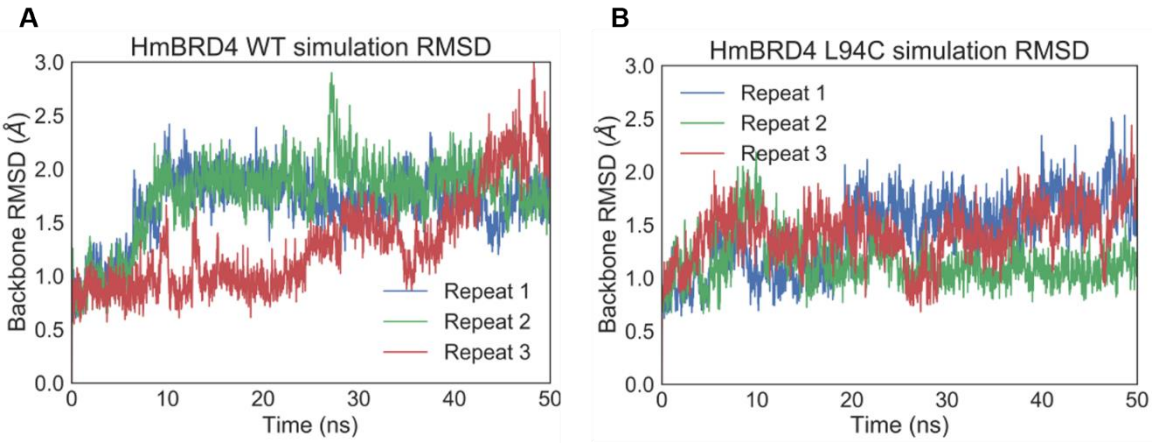


Figure 2.4. Molecular dynamics simulations were performed to compare the stability of BRD4(1)^{L94C} to BRD4(1)^{WT}. The experiments were carried out in triplicate, with each replica shown in either blue, green or red. **A**) RMSD simulation for BRD4(1)^{WT}. **B**) RMSD simulation for BRD4(1)^{L94C}.

2.3 Mutant protein expression and purification

Given the promising nature of the *in silico* data above, the expression and purification of BRD4(1)^{L94C} was consequently attempted. To aid with purification, a His_6-tag was fused to the N-terminal of the protein. Using the published BRD4(1)^{WT} construct (Addgene plasmid #38942) as a template, a DNA gene fragment coding for the amino acids between residues 44 and 168 with the desired L94C mutation was purchased (**Appendix A**). The mutation was induced by codon change from 'CTC' to 'TGT'. In collaboration with Dr Matthias Schiedel and Edward Coombs, the DNA was amplified using PCR, extracted, and then assembled into a pNIC28-Bsa4 vector using a Gibson assembly kit and transformed into *Escherichia coli* (*E.coli*) NEB 5-alpha competent cells. The plasmid was isolated from a single colony using a GeneJET Plasmid Minikit and submitted for sequencing, which confirmed that both the desired mutation and gene insertion into the plasmid had been achieved.[141] Using heat shock, competent *E. coli* BL21-Gold (DE3) cells were transformed with the isolated BRD4(1)^{L94C} plasmid. Colonies were grown overnight on agar plates treated with kanamycin, after which a single colony was picked and used to inoculate a 100 mL starter culture. The starter culture was incubated overnight before being used for large-scale protein expression. Once the large-scale cultures reached an optical density at 600 nm ($O.D_{600}$) of 0.8, isopropyl ß-D-1-thiogalactopyranoside (IPTG) was added to induce gene transcription and hence protein production of BRD4(1)^{L94C}.

The desired BRD4(1)^{L94C} protein was purified to homogeneity through sequential immobilised metal affinity chromatography (IMAC) and size exclusion chromatography (SEC) (**Figure 2.5 A** & **Appendix A**). Using protein mass spectrometry, the L94C mutation was confirmed by a major peak corresponding to the protein's expected mass of 17539 Da (**Figure 2.5 B**). A common peak adduct of +178 Da was observed, corresponding to the α-*N*-6-phosphogluconoylation of the His_6-tag.[145] This modification is important to note as it is within the mass range of the covalent fragments, and therefore should not be confused with covalent labelling of the protein by a fragment.

As a control, BRD4(1)^{WT} was expressed and purified by Dr Matthias Schiedel (University of Oxford) using the previously discussed method (**Figure 2.5 C**). Protein mass spectrometry confirmed the correct mass of BRD4(1)^{WT} as 17549 Da. This was a +10 Da difference compared to BRD4(1)^{L94C}, consistent with the L94C mutation (**Figure 2.5 D**). To determine whether a labelling reaction was occurring selectively at C94, or unselectively at a native residue, the reactive fragments were screened against both BRD4(1)^{WT} and BRD4(1)^{L94C}. Additionally, comparing the structural stability and functionality of BRD4(1)^{L94C} to BRD4(1)^{WT} would allow confirmation that the mutation had not been detrimental.

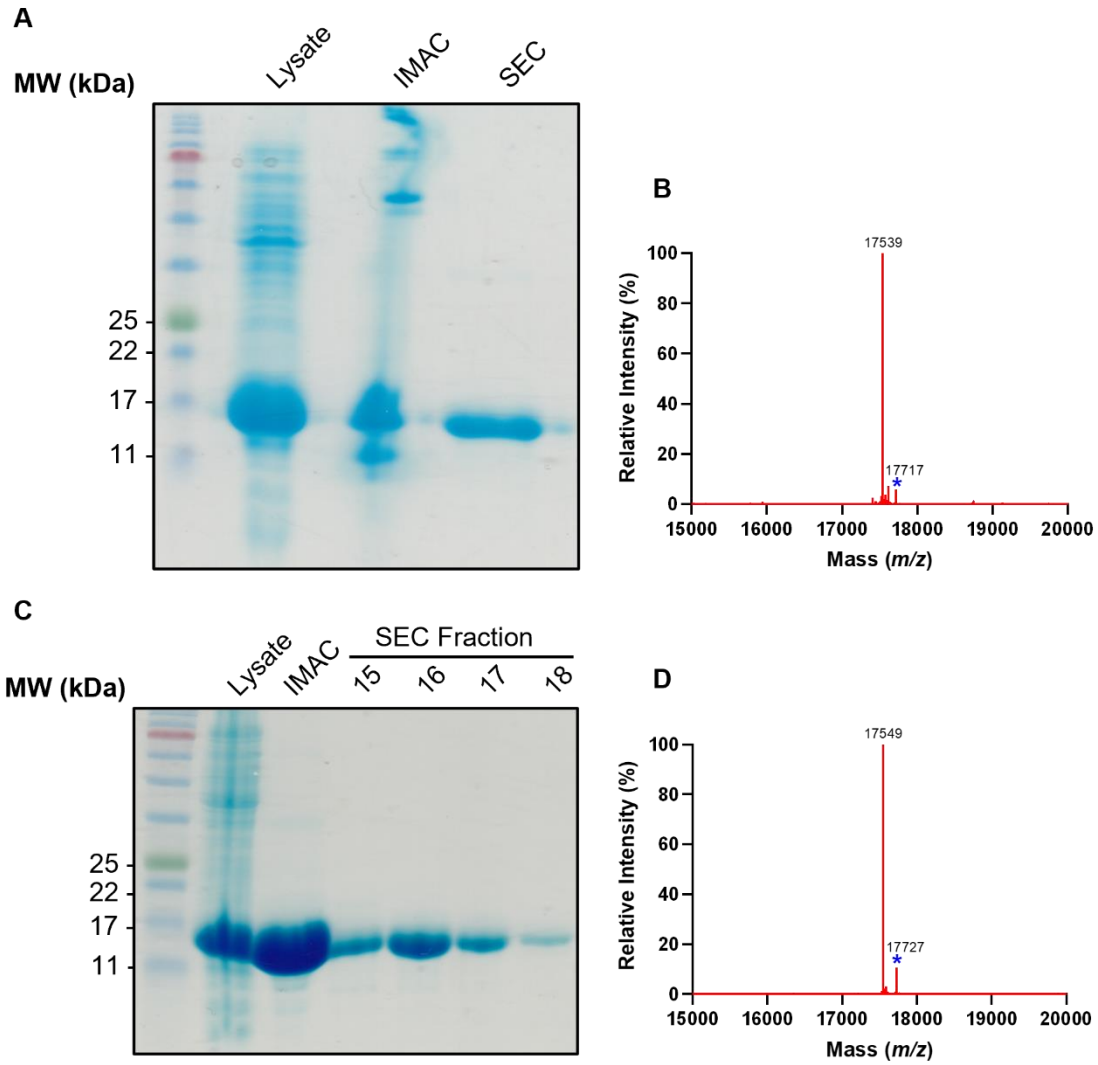


Figure 2.5. Purification of BRD4(1)^{L94C} and BRD4(1)^{WT}. Staining of gels with Coomassie was used to visualise protein. **A**) SDS-PAGE analysis of the cell lysate, IMAC fractions and SEC fractions for BRD4(1)^{L94C}. **B**) Protein mass spectrometry of BRD4(1)^{L94C} confirmed the expected mass. The blue asterisk highlights the +178 Da adduct for the His_6-tag modification. **C**) SDS-PAGE analysis of the cell lysate, IMAC fractions and SEC fractions for BRD4(1)^{WT}. **D**) Protein mass spectrometry of BRD4(1)^{WT} confirmed the expected mass. The blue asterisk highlights the +178 Da adduct for the His_6-tag modification.

2.4 Structural validation

2.4.1 Biophysical assays

When mutating proteins, it is important to validate that the protein is folded and functioning correctly. To evaluate the stability of BRD4(1)^{L94C} and compare it to that of BRD4(1)^{WT}, a differential scanning fluorimetry (DSF) experiment was carried out. DSF calculates the melting temperature (T_m) of proteins by measuring the change in fluorescence intensity of the dye SYPRO orange as the protein is heated. As the protein is heated and becomes denatured, it exposes hydrophobic regions typically buried within the protein. The SYPRO orange dye can then bind to these regions, leading to an increase in fluorescence emission.[146] From the DSF experiment, the mean calculated melting temperatures over six replicates for BRD4(1)^{WT} and BRD4(1)^{L94C} were 49.4 °C and 49.0 °C, respectively (**Figure 2.6**). The minimal discrepancy between the melting points of BRD4(1)^{WT} and BRD4(1)^{L94C} demonstrated that the L94C mutation had not led to any substantial perturbations in stability or protein folding.

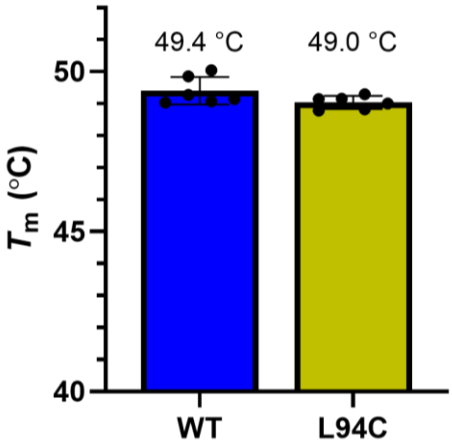


Figure 2.6. DSF experiments were performed to determine the melting temperature (T_m) and stability of BRD4(1)^{L94C} (yellow) in comparison to BRD4(1)^{WT} (blue) (n=6).

Once the stability of BRD4(1)^{L94C} had been verified by DSF, it was sought to investigate whether the protein structure, and specifically the structure of the binding pocket had been affected by the mutation. To investigate this, isothermal titration calorimetry (ITC) was employed, to measure the binding affinity (K_d) of the BRD4(1) ligand **OXFBD04** (**1.58**) to both BRD4(1)^{WT} and BRD4(1)^{L94C}. The

measured K_d value for BRD4(1)^{L94C} was 0.609 µM, and for BRD4(1)^{WT} was 0.194 µM, which was within error of the published value of 0.247 µM (**Figure 2.7**).[132] These data show that **1.58** still bound to BRD4(1)^{L94C} despite the mutation, indicating that the KAc binding pocket was not significantly altered by this change. Whilst there was a slight decrease in affinity, this was to be expected,[128,134,144] and is likely to be due to the loss of the hydrophobic interaction between the 5-methyl group of **1.58** and the isobutyl group of L94.

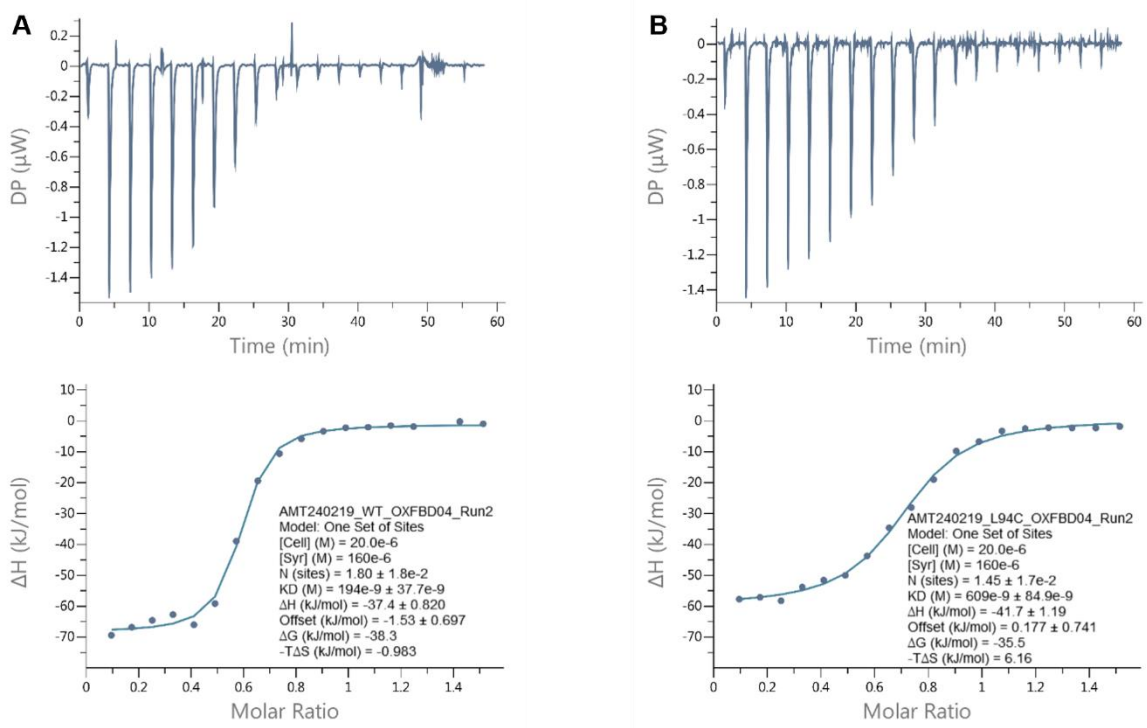


Figure 2.7. ITC experiments with **OXFBD04** (**1.58**) were performed to evaluate whether the mutation had compromised the vital KAc binding pocket. **A**) ITC trace for the binding of **1.58** to BRD4(1)^{WT}. **B**) ITC trace for the binding of **1.58** to BRD4(1)^{L94C}.

2.4.2 Crystallographic studies

To further confirm structural resemblance to BRD4(1)^{WT}, and to aid compound design, it was decided that, in collaboration with Dr Michael McDonough (University of Oxford), an X-ray structure of BRD4(1)^{L94C} should be obtained.

Despite the flexibility of the His_6-tag, it was decided that an attempt to crystalise BRD4(1)^{L94C} without cleavage of the His_6-tag should be carried out, as this had previously been achieved for BRD4(1)^{WT}.[132,147] Initial crystallisation trials aimed to crystalise the apo- structure of BRD4(1)^{L94C} by incubating the protein with a range of reagents and temperature conditions (**Table 2.1 Plate 1** & **2**). Unfortunately, after 2 months no crystals had formed. It was postulated that co-crystallisation with **OXFBD04** (**1.58**) might increase the chances of crystallisation through stabilisation of the protein. Therefore, BRD4(1)^{L94C} was incubated with a racemic mixture of **1.58** and applied the previous reagent and temperature conditions (**Table 2.1 Plate 3** & **4**).

Table 2.1. Conditions attempted for the crystallisation of BRD4(1)^{L94C}.

Plate	Crystal Screen	His_6-BRD4(1)^{L94C} (mg/mL)	OXFBD04 (mM)	Temp (°C)
1	JCSG-plus HT-96 (MD1-40)	8.56	0	4
2	JCSG-plus HT-96 (MD1-40)	8.56	0	20
3	JCSG-plus HT-96 (MD1-40)	8.56	1	4
4	JCSG-plus HT-96 (MD1-40)	8.56	1	20

After 6 weeks over 30 crystals had grown in plate 3, well G11.1 (2.0 M ammonium sulfate, 0.1 M Bis-Tris, pH 5.5, no precipitant), of which several crystals were collected and stored in liquid nitrogen (**Figure 2.8 A**). One crystal was mounted and diffracted in-house to confirm that the crystals were protein and not salts. After confirmation that the crystals were protein, diffraction data for one crystal was collected at Diamond Light Source to give a 1.88 Å resolution structure of **(*R*/*S*)-1.58** bound to BRD4(1)^{L94C} (**Figure 2.8 B** & **Appendix A**). This result was extremely pleasing and allowed visualisation of the spatial orientation of the cysteine residue and its thiol, which was measured to be 4.0–4.1 Å from the 5-methyl carbon. Previously, it has been shown that the both

the (*R*)- and (*S*)- enantiomer of **1.58** display similar IC_{50} values for BRD4(1)^{WT}.[132] Intriguingly, previous crystallisation of BRD4(1)^{WT} with a racemic mixture of **1.58** resulted in only the (*R*)-enantiomer being observed in the X-ray crystal structure (PDB ID: 6FSY),[132] however, in this X-ray crystal structure both the (*R*)- and (*S*)-enantiomer are observed in the electron density map. This phenomenon might be caused by a minor change in the ZA loop structure, owing to the slight difference in size of the cysteine residue compared to the leucine residue. Importantly, both enantiomers retain the hydrogen bonding interaction to N140 and the water-mediated hydrogen bond to Y97, suggesting minimal disruption to the KAc binding pocket.

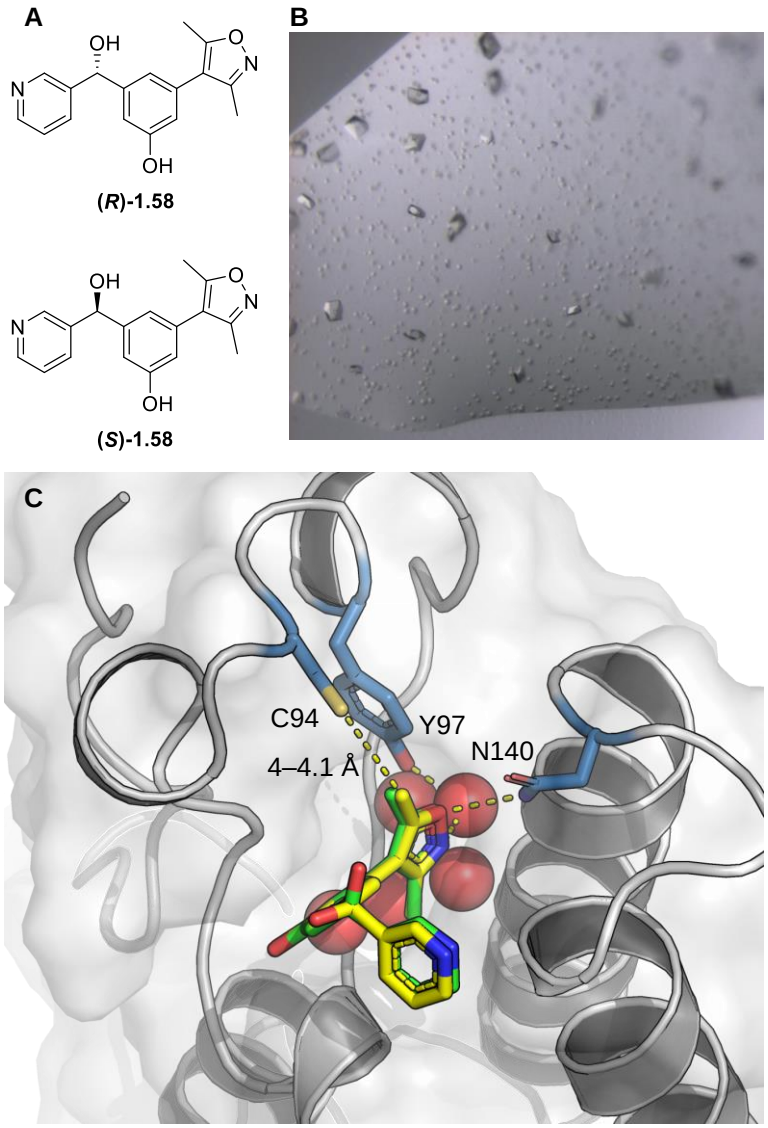


Figure 2.8. X-Ray crystallography was employed to further validate the structural integrity of BRD4(1)^{L94C}. **A**) Chemical structures of **(R/S)-1.58**. **B**) Microscopic images of the crystals after 6 weeks. **C**) X-ray crystal structure of **(*R*)-1.58** (carbon = yellow) or **(*S*)-1.58** (carbon = green) bound to BRD4(1)^{L94C} with C94, Y97 and N140 highlighted (carbon = blue).

Alignment and overlay of the structures of **(*R*)-1.58** bound to BRD4(1)^{WT} and BRD4(1)^{L94C} demonstrates the high structural resemblance between the two proteins (**Figure 2.9**). However, some variation in protein structure can be found in the ZA loop region where the mutation has occurred, as well as some movement of the binding pocket waters, with the largest shift being 0.7 Å (**Figure 2.9 B**). These observations may possibly contribute to the change in affinity for **1.58** by ITC, however at this resolution these discrepancies may not be significant, and rather just an artifact of the resolution.

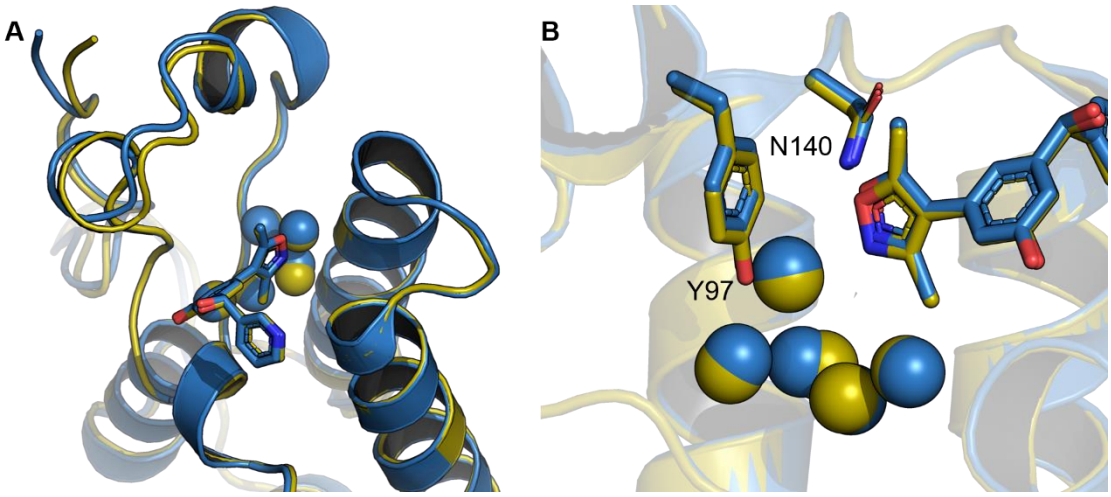


Figure 2.9. Aligned and overlaid structures of **(*R*)-1.58** bound to BRD4(1)^{WT} (blue) and **(*R*)-1.58** bound to BRD4(1)^{L94C} (yellow). **A**) Whilst the α-helices align perfectly there is some variation within the ZA loops. **B**) Inside the KAc binding pocket some differences between the water molecules can be observed.

2.5 Conclusions

In this chapter, the rational design of a stable and structurally comparable mutant BRD4(1) protein containing an accessible cysteine residue for covalent labelling has been discussed. X-Ray crystal structure analysis led to mutation of L94 in the ZA loop, owing to its proximity to the KAc binding pocket, and its likely minimal role in protein structure. Molecular dynamics simulations predicted that the mutant BRD4(1)^{L94C} would be stable and structurally similar to BRD4(1)^{WT}. Confident with these data, BRD4(1)^{L94C} was expressed in *E.coli*, and the protein was purified to homogeneity yielding 86 mg of BRD4(1)^{L94C} from 4 L of media. Protein mass spectrometry confirmed that the

mutation was present, with a −10 Da difference compared to the mass of BRD4(1)^{WT}, consistent with the mass difference between a leucine and cysteine residue. Using DSF and ITC, it was found that BRD4(1)^{L94C} was structurally equivalent to BRD4(1)^{WT}, and although there was some drop in the affinity of **1.58** for the mutant, it was hypothesised that BRD4(1)^{L94C} could function in the cellular environment.

An X-ray crystal structure of **(*R*/*S*)-1.58** bound to BRD4(1)^{L94C} was obtained with a 1.88 Å resolution, giving additional clarification that the mutation had not significantly altered the protein structure. With the aid of the X-ray crystal structure data, the design and synthesis of covalent fragments targeting C94 could then be the focus (**Chapter 3**).

Chapter 3

Identification of reactive fragments that covalently bind to BRD4(1)^{L94C}

3.1 Introduction and aims

Having validated BRD4(1)^{L94C} as a structurally stable mutant, as described in **Chapter 2**, the focus turned to the identification of covalent fragments targeting C94 of BRD4(1)^{L94C}. As it was envisioned that 'mutate & conjugate' would be best applicable to emerging targets, or poorly liganded proteins, it was decided that simple fragments that may be typical starting points for probe discovery should be employed. Using both a structure-guided approach and a diverse fragment-based approach, the work described in this chapter aimed to:

- Functionalise the DMI scaffold with a variety of electrophiles for optimal covalent labelling of C94.
- Discover novel covalent fragments targeting C94, using GSK's diverse reactive fragment screening platform.

3.2 Structure-guided covalent fragment design

To begin the structure-guided approach, a simple, known, KAc mimic fragment was employed. It was therefore decided to design a library of DMI-based covalent fragments. Analysis of the X-ray crystal structure of **OXFBD04** (**1.58**) bound to BRD4(1)^{L94C} obtained in **Chapter 2** shows that, as for **1.58** bound to BRD4(1)^{WT} (PDB ID: 6FSY),[132] the 3-methyl group is buried into the KAc binding pocket, facilitating hydrogen bonding interactions between the isoxazole and N140, as well as a water-mediated hydrogen bond to Y97 (**Figure 3.1**). With the 3-methyl group of the DMI pointing into the binding pocket, the 5-position is directed out of the binding pocket toward C94. Measurement using PyMOL estimated a distance of 4.0–4.1 Å between the 5-methyl carbon atom and the sulfur atom of the thiol (**Figure 3.1**). It was hypothesised that building the electrophile from the 5-position of the isoxazole would position the electrophile and C94 close, which was desirable for covalent bond formation. Therefore, it was decided to focus on the 3-methyl regioisomers by functionalising the

5-position of the DMI core with the electrophile, with the aim of retaining the key non-covalent interactions seen between **1.58** and BRD4(1)^{L94C}. Although it was believed that the 5-methyl regioisomers would be less effective at labelling C94, owing to the less favourable interactions with the KAc binding pocket, a small selection of the 5-methyl regioisomers were synthesised as control compounds.

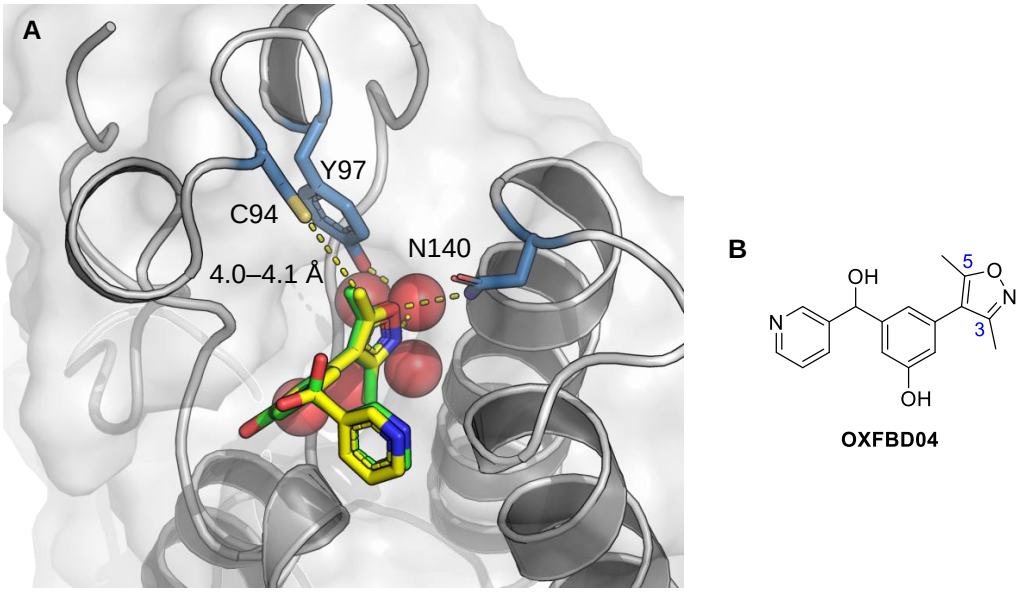
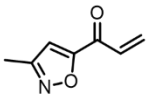
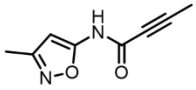
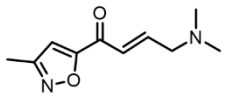
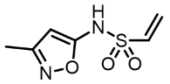
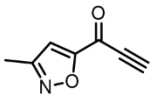
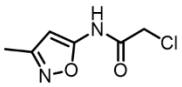
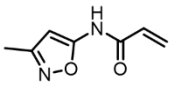
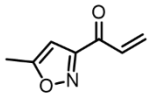
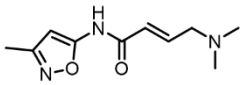
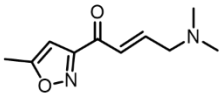
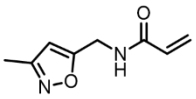
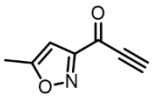
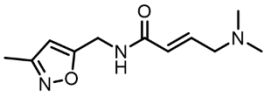
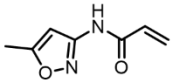
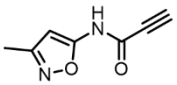
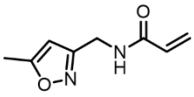


Figure 3.1. The X-ray crystal structure of ***(R)***- and ***(S)*-OXFBD04** (**1.58**) bound to BRD4(1)^{L94C}. **A**) Measurement using PyMOL estimates a 4.0–4.1 Å distance between the DMI 5-methyl carbon and C94 thiol sulfur. **B**) Chemical structure of **1.58**, with DMI numbering shown in blue.

With a diverse array of electrophilic groups available, fragments were designed with the emphasis on robust, short synthetic routes. It was postulated that the linker distance between the isoxazole core and the electrophile would be important for the reaction between C94 and the electrophile. If the linker were too long, then the covalent reaction might not occur owing to poor spatial orientation. If the linker were too short the fragment might not covalently bind, or the isoxazole might be pulled out of the binding pocket upon covalent bond formation, and therefore no longer inhibit the protein's function. Consequently, an initial library of isoxazole fragments, containing a variety of electrophilic groups and linkers were designed (**Table 3.1**).

Table 3.1. The initial target reactive fragments. A variety of electrophiles were placed at either the 3- or 5-position of the isoxazole KAc mimic, with the aim of understanding reactivity for BRD4(1)^{L94C}.

Compound	Structure	Compound	Structure
3.1		**3.9**	
3.2		**3.10**	
3.3		**3.11**	
3.4		**3.12**	
3.5		**3.13**	
3.6		**3.14**	
3.7		**3.15**	
3.8		**3.16**	

The initial fragment library was designed to incorporate electrophiles with a range of reactivities to enable probing of the optimum reactive fragment structure. Although likely to be very reactive, and possibly unselective due to their high electrophilicity, the α,β-unsaturated ketones, **3.1**–**3.3** and **3.12**–**3.14**, were designed to provide the shortest distance between the isoxazole and electrophile. Chloroacetamide **3.11** would also provide a similar distance and afforded covalent bond formation *via* an S_N2 reaction.

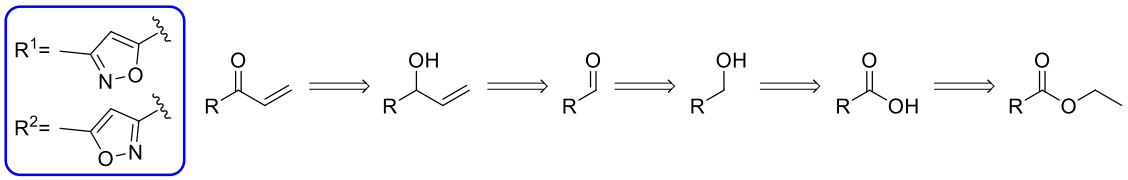
The amide-linked fragments, **3.4**, **3.5**, **3.8**, **3.9**, and **3.15**, increase the distance between the isoxazole and the reaction centre, while reducing the electrophilicity, compared to the α,β-unsaturated ketones. To provide a more flexible linker, an sp^3 hybridised carbon atom was added for compounds **3.6**, **3.7**, and **3.16**. It was also decided to explore how substitution of the

electrophile altered the reactivity and selectivity of the fragments, and so a selection of electrophiles bearing methyl, **3.9**, or DMAM substituents, **3.2**, **3.5**, **3.7**, and **3.13**, were designed. Vinyl sulfonamides have proven to be useful electrophiles for the covalent labelling of cysteine, and so fragment **3.10** functionalised with a vinyl sulfonamide was designed.

3.3 Structure-guided covalent fragment synthesis

3.3.1 Ketone-linked fragments

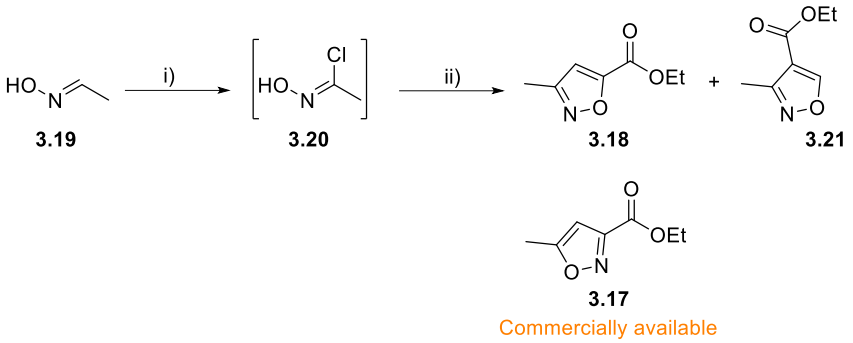
Retrosynthetic analysis shows that the formation of the α,β-unsaturated ketones could be obtained through a Grignard reaction, using vinyl- or ethynyl- magnesium bromide in the presence of the corresponding aldehyde (**Scheme 3.1**). The aldehyde could be obtained through sequential reduction and oxidation of the analogous esters.



Scheme 3.1. Retrosynthetic analysis to obtain the α,β-unsaturated ketone fragments.

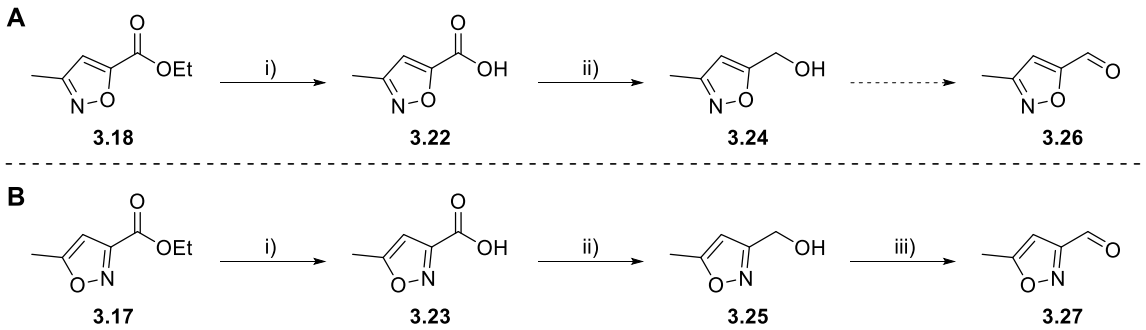
Whilst **3.17** was commercially available for a reasonable price, **3.18** was not, and so was synthesised *via* a two-step, 1,3-dipolar cycloaddition (**Scheme 3.2**).[148,149] The first step formed the chloroxime precursor to the nitrile oxide. This was achieved through an *N*-chlorosuccinimide (NCS) mediated chlorination of acetaldoxime **3.19**, which was used directly in next step due to the instability of this intermediate. The chloroxime **3.20** was added to a suspension of ethyl propiolate and potassium hydrogen carbonate in dichloromethane. The nitrile oxide was generated *in situ* through a base mediated γ-elimination of the chloride, followed by a 1,3-dipolar cycloaddition with ethyl propiolate to form **3.18** in a 43% yield. A common side product in this reaction is the formation of the regioisomer **3.21**, however, the desired regioisomer **3.18** could be separated from **3.21** using

silica gel column chromatography. This was confirmed by NMR analysis, where the 4-CH proton of **3.18** appears at 6.77 ppm, compared to 8.84 ppm for the 5-CH proton of **3.21**.[148]



Scheme 3.2. Synthesis of the ethyl ester intermediate **3.18**. *Reagents and conditions*: i) *N*-Chlorosuccinimide, DMF, 0 °C to rt, 3 h; ii) Ethyl propiolate, $KHCO_3$, CH_2Cl_2, H_2O, 16 h, 23–43%, n=2.

The carboxylic acids **3.22** and **3.23** were then obtained through the hydrolysis of **3.18** and **3.17**, using lithium hydroxide in a biphasic mixture of THF and water (**Scheme 3.3**).[150] The carboxylic acids were isolated in high yields by the removal of the THF and acidification to precipitate the carboxylic acid. Reduction of the carboxylic acids to the alcohols **3.24** and **3.25** was then achieved using borane dimethylsulfide complex, however, isolation of these compounds in high yields was hindered by the volatility of the alcohols (**Scheme 3.3**). To obtain **3.27**, oxidation of the alcohol was carried out with manganese dioxide. Although a very mild oxidising agent, manganese dioxide can selectively oxidise benzylic alcohols under mild conditions, making it a suitable choice for the oxidation of **3.25**.[151] Despite the reaction proceeding cleanly by TLC analysis, isolation of compound **3.27** was in poor yields due to its volatility and instability (**Scheme 3.3**).

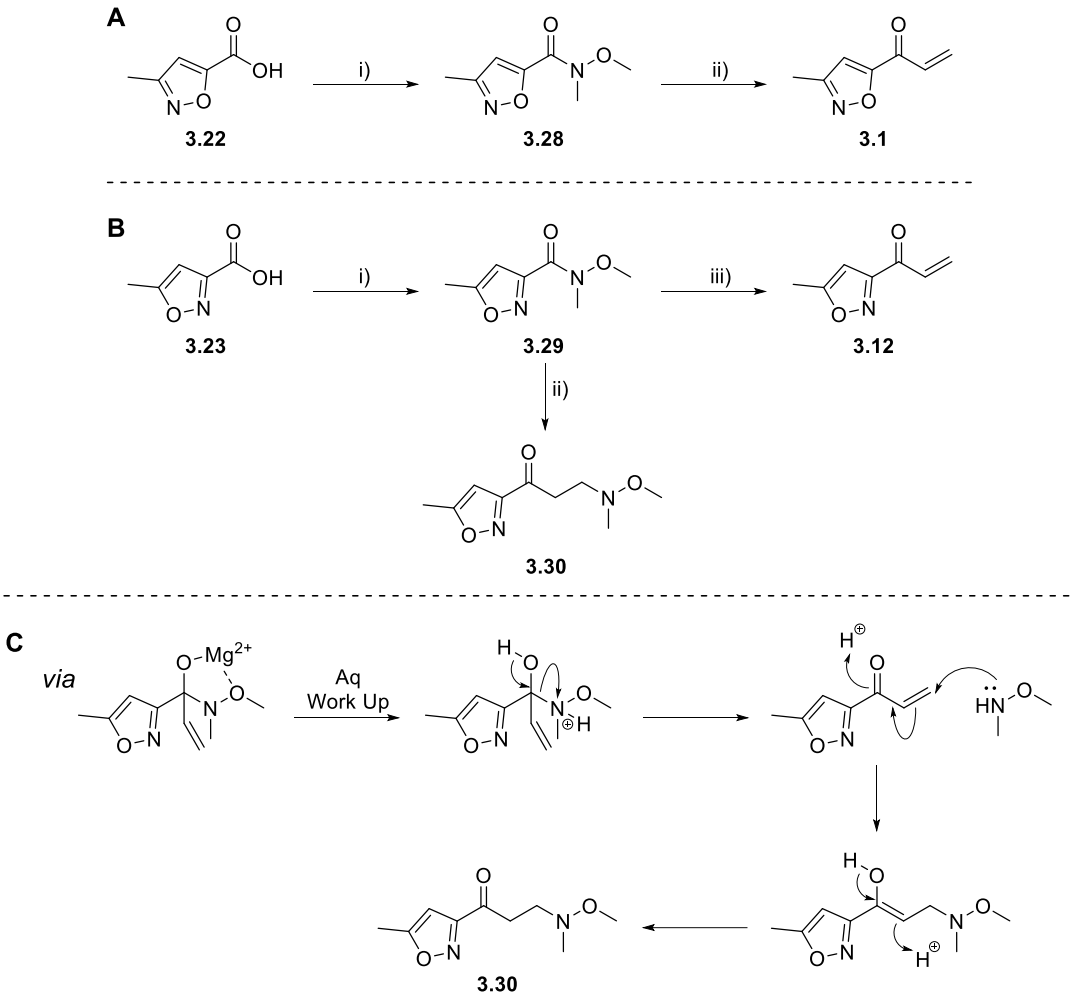


Scheme 3.3. Synthesis of the aldehyde intermediates **3.26** and **3.27**. *Reagents and conditions*: **A**) i) LiOH, THF, H_2O (2:1), rt, 16 h, 55–100%, n=3; ii) $BH_3{\cdot}SMe_2$ complex, THF, 0 °C to rt, 44 h, 18%, n=1. **B**) i) LiOH, THF, H_2O (2:1), rt, 16 h, 92–97%, n=3; ii) $BH_3{\cdot}SMe_2$ complex, THF, 0 °C to rt, 44 h, 67%, n=1; iii) MnO_2, CH_2Cl_2, rt, 60 h, 16%, n=1.

These observations led to the decision to use a Weinreb amide approach instead. This approach removes the need to re-oxidise the alcohols after the nucleophilic addition of the Grignard, and the products are less volatile. Using PyBOP as the coupling reagent, both **3.22** and **3.23** were coupled to *N,O*-dimethylhydroxylamine, in generally high yields, to form the required Weinreb amides **3.28** and **3.29** (**Scheme 3.4**).

During the initial attempts to synthesise **3.12,** the desired product was not formed, despite consumption of the starting material and appearance of a new spot by TLC. The new spot was isolated and characterised to determine where the reaction was failing. Analysis of NMR and MS data suggested that the undesired product was **3.30** (**Scheme 3.4**). This compound was likely formed during the aqueous work up, where the tetrahedral intermediate collapses to release the *N,O*-dimethylhydroxylamine leaving group and form the desired α,β-unsaturated ketone. However, if not all the amine is protonated and extracted into the aqueous phase, the *N,O*-dimethylhydroxylamine can act as a nucleophile and undergo a 1,4-conjugate addition, to yield **3.30**.

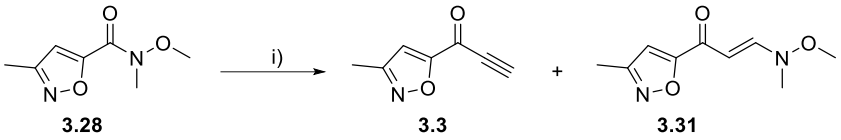
To reduce the formation of this side product, two modifications were made to the quenching procedure. Firstly, rather than quenching the reaction at room temperature, it was quenched and extracted at 0–5 °C using ice-cold aqueous 1 M HCl and EtOAc. It was hoped that lowering the temperature would slow the conjugate addition. Secondly, the reaction was quenched rapidly with excess ice-cold aqueous 1 M HCl to ensure the full protonation of the leaving group and removal of the nucleophile into the aqueous phase. These optimised conditions gave good results, with reactions proceeding cleanly and with no conjugate addition observed, and so were used for the synthesis of both regioisomers, **3.1** and **3.12** (**Scheme 3.4**).



Scheme 3.4. Synthesis of the **3.1** and **3.12**. *Reagents and conditions*: **A**) i) *N,O*-Dimethylhydroxylamine hydrochloride, PyBOP, Et_3N, CH_2Cl_2, 0 °C to rt, 16 h, 44–81%, n=4; ii) Vinylmagnesium bromide, THF, −35 °C to 0 °C, 2 h, 56%, n=1. **B**) i) *N,O*-Dimethylhydroxylamine hydrochloride, PyBOP, Et_3N, CH_2Cl_2, 0 °C to rt, 16 h, 62–82%, n=4; ii) Vinylmagnesium bromide, THF, 0 °C to rt, 4.5 h, 40%, n=1; iii) Vinylmagnesium bromide, THF, −35 °C to 0 °C, 2 h, 38–89%, n=2. **C**) The proposed mechanism for the formation of **3.30**. During the aqueous work up the tetrahedral intermediate collapses to release *N,O*-dimethylhydroxylamine which, if not protonated, can participate as a nucleophile in a 1,4-conjugate addition reaction with the enone.

With the Grignard conditions optimised, the same Weinreb amide approach was employed for the synthesis of the ynones **3.3** and **3.14**. Addition of ethynylmagnesium bromide to a solution of **3.28** in anhydrous THF at −35 °C formed the desired ynone **3.3**, with only minor amounts of the conjugate addition product, **3.31** (**Scheme 3.5**)**.** The conjugate addition product was isolated for testing due to its possible cysteine reactive properties. Upon purification of the ynone using silica gel column chromatography an inseparable impurity was formed, meaning that the biological testing standard of >95% purity was not met. Unfortunately, NMR and MS analysis did not allow for structure

determination of the impurity, but it was postulated that its formation was likely due to the reactivity of the ynone.

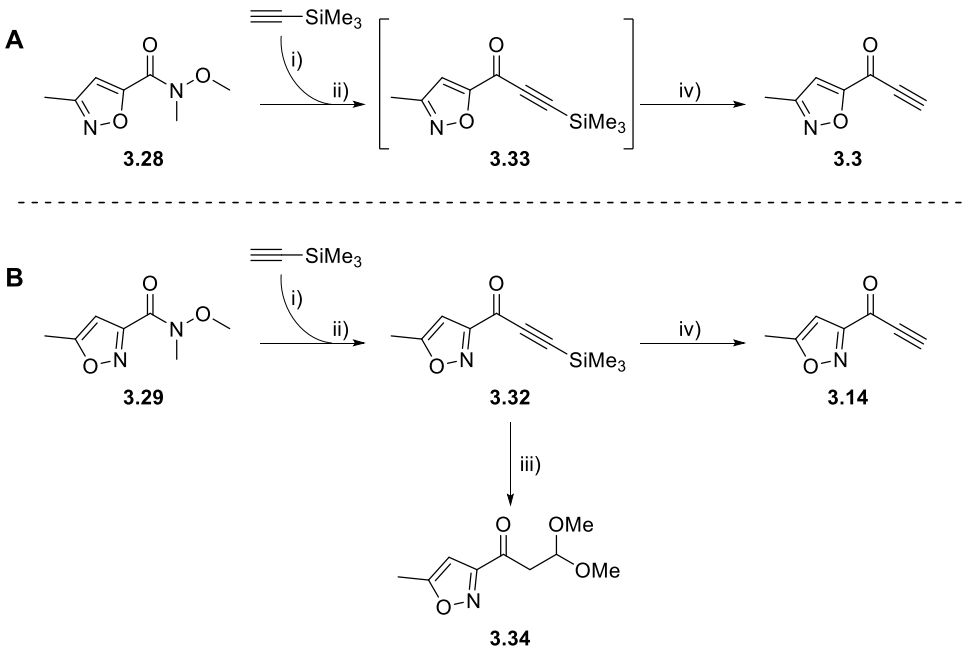


Scheme 3.5. Initial synthesis of **3.3**. *Reagents and conditions*: i) Ethynylmagnesium bromide, THF, −40 °C to 0 °C, 4 h, **3.3**: 34%, **3.31**: 5%, n=1.

To overcome the formation of the unknown impurity, the TMS protected alkyne, ethynyltrimethylsilane, was used, which could be deprotected as the final step. Isopropylmagnesium chloride was used to deprotonate the ethynyltrimethylsilane and form the Grignard reagent before dropwise addition to the Weinreb amide (**Scheme 3.6**). Similarly, to the enone synthesis, care was taken to rapidly quench the reaction at low temperatures to avoid any unwanted conjugate addition. For the reaction with **3.29** the product **3.32** was isolated in a 68% yield. Interestingly, upon purification of **3.33** using silica gel column chromatography, the crude silylated product decomposed to reveal the de-silylated ynone **3.3** which was isolated in a 27% yield (**Scheme 3.6**).

In an attempt to deprotect **3.32**, tetrabutylammonium fluoride (TBAF) was employed first . TLC analysis indicated that after 20 minutes of treatment with TBAF, at 0 °C, all starting material had been consumed, and a mixture of products had formed. Silica gel column chromatography was used to separate the mixture, however, ^{1}H NMR analysis revealed that none of the isolated compounds were the desired product. It was therefore decided to attempt potassium carbonate and methanol. The reaction was carried out at room temperature in methanol, and occurred rapidly, with complete consumption of the silylated starting material within minutes (**Scheme 3.6**). ^{1}H NMR and MS analysis confirmed that the product was not the desired ynone, but instead the product of the TMS deprotection followed by a double conjugate addition of the methoxy anions to yield **3.34**. It could be seen from the crude NMR that the singlet at 3.37 ppm integrating for six protons corresponded to the two methoxy groups. The addition of these two methoxy groups had saturated

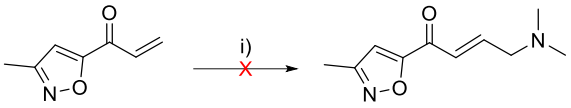
the α-position with two protons observed as a doublet at 3.35 ppm, which were coupling to the β-proton at 5.06 ppm with a *J* value of 5.8 Hz. MS analysis further verified this structure through an intense peak at 222 *m/z* corresponding to the mass of the sodium ion adduct. The deprotection conditions were modified for the second attempt. Instead of methanol, dichloromethane was used as the solvent, giving control over the amount of methoxy anions generated. The reaction was kept at 0 °C and small portions of methanol were added to the suspension, and the reaction monitored using TLC. Upon completion, the reaction was diluted and filtered to remove the potassium carbonate. A crude NMR confirmed that the afforded product was the de-silylated ynone **3.14**, and that minimal conjugate addition product had formed. After purification, **3.14** was isolated in a 46% yield (**Scheme 3.6**).



Scheme 3.6. Synthesis of **3.3** and **3.14**. *Reagents and conditions*: **A**) i) Isopropylmagnesium bromide, THF, −40 °C to 0 °C, 1.5 h; ii) THF, −40 °C to 0 °C, 1.5 h; iii) Silica gel, 27% over two steps, n=1. **B**) i) Isopropylmagnesium bromide, THF, −40 °C to 0 °C, 1.5 h; ii) THF, −40 °C to 0 °C, 1.5 h, 50–68%, n=2; iii) K_2CO_3, MeOH, rt, 5 min, 50%; iv) K_2CO_3, MeOH, CH_2Cl_2, 0 °C, 20 min, 46%, n=1.

With the unsubstituted α,β-unsaturated ketones in hand, the focus was turned to synthesising the substituted analogues, with the aim of modulating the reactivity of the fragments. To synthesise **3.2**, it was originally envisioned that utilising olefin metathesis to form the substituted enone might be viable (**Scheme 3.7**). Using Grubbs Second Generation catalyst,[152] **3.1** and *N,N*-

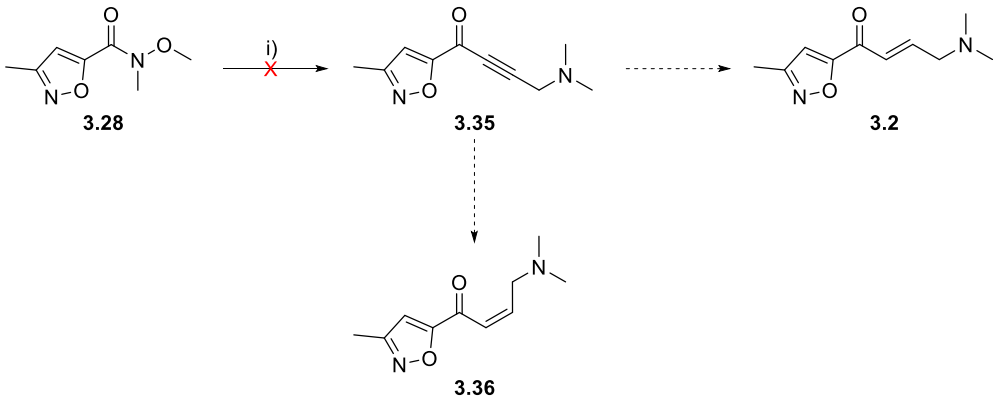
dimethylallylamine were reacted together. *p*-Toluenesulfonic acid was used to protonate the amine and prevent the amine lone pair from poisoning the catalyst. After 18 hours at room temperature, the reaction had formed a complex mixture of species, as judged by TLC analysis, and attempts to separate the components were unsuccessful.



3.1 **3.2**

Scheme 3.7. Attempted synthesis of substituted enone **3.2** using olefin metathesis. *Reagents and conditions*: i) *p*-TSOH·H_2O, *N*,*N*-dimethylallylamine, Grubbs 2nd generation catalyst, $CHCl_3$, rt, 18 h.

An alternative method to synthesise **3.2** using the previously synthesised Weinreb amide was investigated (**Scheme 3.8**). It was hypothesised that deprotonation of 3-dimethylamino-1-propyne with a strong base in the presence of the Weinreb amide would form the substituted ynone **3.35**. This could then be partially hydrogenated to selectively synthesise either the *E*- or *Z*- alkene. Attempts to synthesise ynone **3.35** were unsuccessful for both *n*-butyllithium as well as freshly prepared lithium diisopropylamide. In the case of *n*-butyllithium the major compound isolated was the butylated product.



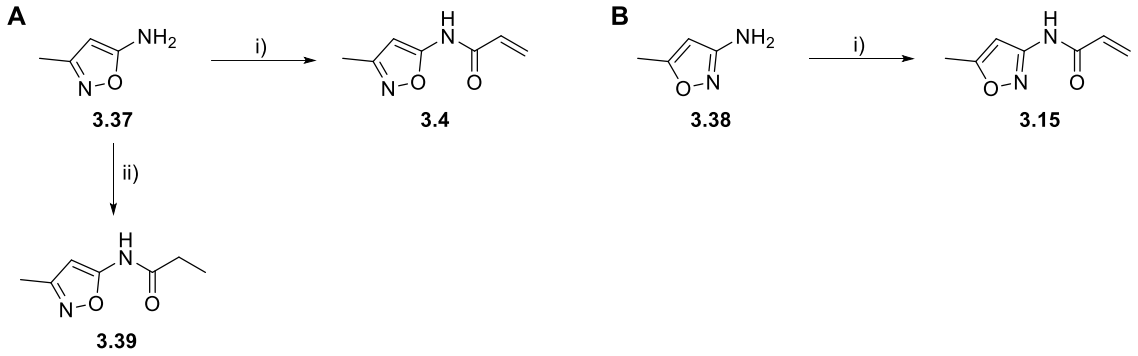
3.28 **3.35** **3.2** **3.36**

Scheme 3.8. Proposed route to access target substituted enones. *Reagents and conditions*: i) 3-Dimethylamino-1-propyne, *n*-BuLi or LDA, THF, −78 °C to 0 °C, 3 h.

Due to the challenges of synthesising the substituted α,β-unsaturated ketones **3.2** and **3.13**, and coupled with the likelihood of high reactivity and non-specific covalent binding, it was decided to focus on development of the amide fragments.

3.3.2 Amide-linked fragments

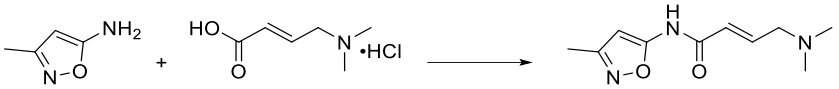
With both regioisomers of the amine, **3.37** and **3.38**, commercially available, functionalisation of the amine scaffolds to covalent fragments was achieved through coupling to the corresponding acid or acyl chloride. To synthesise the unsubstituted acrylamides **3.4** and **3.15**, acryloyl chloride was reacted with **3.37** and **3.38**, respectively. These coupling reactions typically proceeded well and in yields of over 70% (**Scheme 3.9**). To ensure that covalent labelling of the protein was through the reaction between the acrylamide and the thiol, a saturated control compound **3.39** was synthesised. This was obtained through the coupling of **3.37** and propionyl chloride, which proceeded in a 79% yield.



Scheme 3.9. Synthesis of amide fragments **3.4** and **3.15**, and negative control **3.39**. *Reagents and conditions*: **A**) i) Acryloyl chloride, Et_3N, CH_2Cl_2, 0 °C to rt, 16 h, 61%, n=1; ii) Propionyl chloride, Pyridine, CH_2Cl_2, 0 °C to rt, 2.5 h, 79%, n=1. **B**) i) Acryloyl chloride, Et_3N, CH_2Cl_2, 0 °C to rt, 16 h, 72%, n=1.

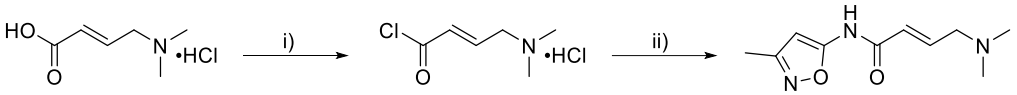
Synthesis of the DMAM substituted acrylamide **3.5** proved to be more challenging to synthesise than its unsubstituted counterpart. A variety of different amide coupling reagents and conditions were attempted to couple **3.37** and (*E*)-4-dimethylaminobut-2-enoic acid **3.40**, but with no success (**Table 3.2**). Instead, the acyl chloride was formed using oxalyl chloride, and coupled to **3.37** to synthesise **3.5**, in a 7% yield (**Scheme 3.10**).

Table 3.2. Amide coupling reagents and conditions investigated for the synthesis of **3.5**.



3.37 + **3.40** → **3.5**

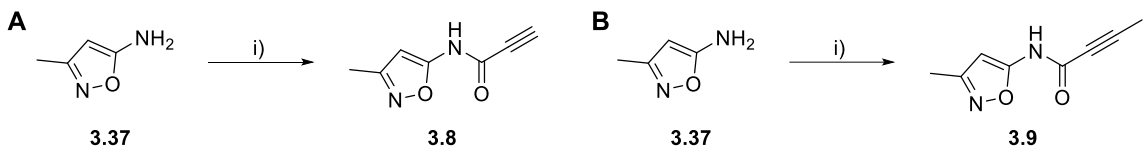
Coupling Reagent	Base	Solvent	Time (h)	3.5 detected (LC-MS)
HATU	DIPEA	DMF	16	✗
DCC	DMAP	CH_2Cl_2	16	✗
PyBOP	Et_3N	CH_2Cl_2	18	✗



3.40 → (i) → → (ii) → **3.5**

Scheme 3.10. Synthesis of **3.5.** *Reagents and conditions*: i) Oxalyl chloride, DMF, THF, 0 °C to rt, 2.5 h; ii) 3-Methylisoxazol-5-amine, Et_3N, THF, 0 °C to rt, 2 h, 7%, n=1.

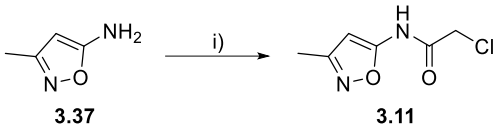
Similar to the ynone fragments, it was desirable to explore the spatial orientation of the electrophile so that optimal reactivity could be achieved. To achieve this, the propynamide **3.8** and 2-butynamide **3.9** derivatives, which would possess a linear vector, were synthesised (**Scheme 3.11**). PyBOP and DCC were investigated on a small scale for the coupling of propiolic acid to **3.37**. The reactions were carried out overnight, and after 16 hours TLC analysis indicated that a product had formed for the DCC coupling, but not for the PyBOP coupling. MS analysis indicated that the compound formed was the desired product, showing an intense peak at 149 *m/z*, corresponding to the mass of $[M-H]^-$. With suitable coupling conditions, both **3.8** and **3.9** were synthesised by coupling the corresponding acids to **3.37** using DCC and DMAP. These reactions were low yielding, however, adequate amounts of compound were obtained for biological testing (**Scheme 3.11**).



A: **3.37** → (i) → **3.8**; **B**: **3.37** → (i) → **3.9**

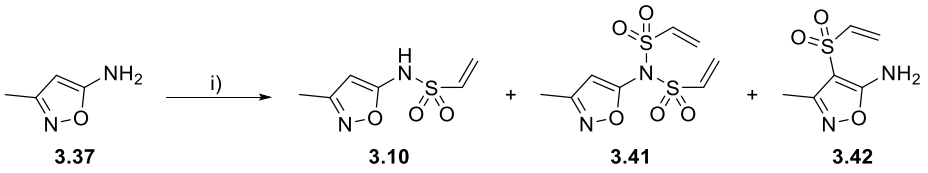
Scheme 3.11. Synthesis of fragments **3.8** and **3.9**. *Reagents and conditions*: **A**) i) Propiolic acid, DCC, DMAP, CH_2Cl_2, 0 °C to rt, 5 h, 31%, n=1. **B**) i) 2-Butynoic acid, DCC, DMAP, CH_2Cl_2, 0 °C to rt, 4 h, 16%, n=1.

Another frequently used electrophile for cysteine labelling is the chloroacetamide moiety, and so **3.11** was synthesised. Chloroacetyl chloride and **3.37** were coupled in a 46% yield to isolate **3.11** (**Scheme 3.12**). Despite the possibility of the amine attacking the α-position of the chloroacetyl chloride in an S_N2 fashion, no reaction following this synthetic pathway was observed.



Scheme 3.12. Synthesis of the chloroacetamide fragment **3.11**. *Reagents and conditions*: i) Chloroacetyl chloride, K_2CO_3, CH_2Cl_2, 0 °C to rt, 16 h, 46%, n=1.

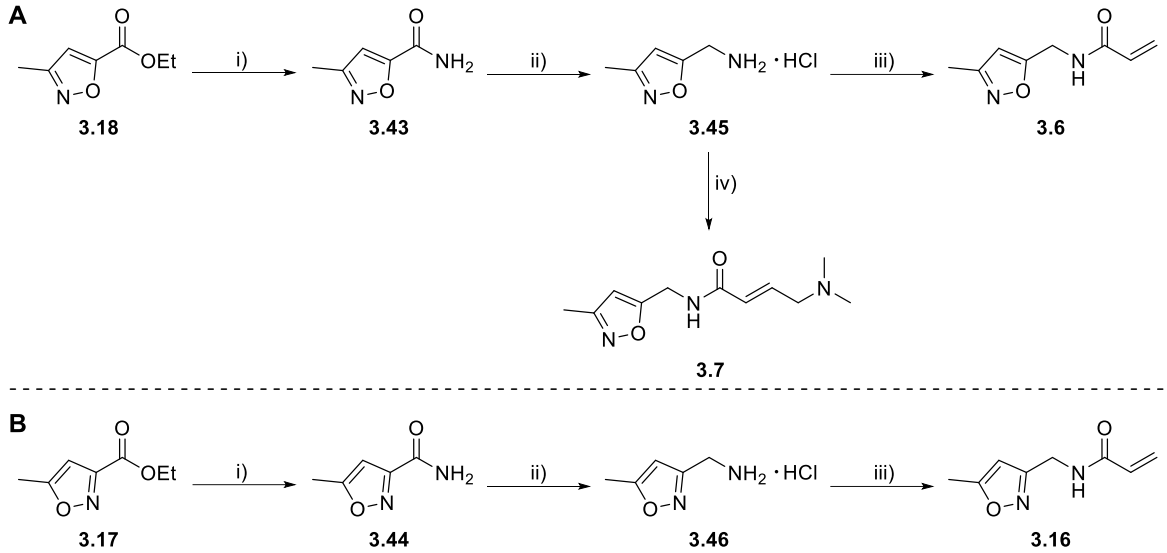
To synthesise the vinyl sulfonamide **3.10**, 2-chloroethanesulfonyl chloride was used in combination with triethylamine (**Scheme 3.13**). This reaction produced a mixture of compounds that were isolated and identified as **3.10**, **3.41**, and **3.42**. In an attempt to supress the formation of unwanted products, weaker, and bulkier bases such as pyridine and 2,6-lutidine were investigated. Unfortunately, it was found that the reaction only proceeded in the presence of triethylamine, albeit in a poor yield.



Scheme 3.13. Synthesis of the vinyl sulfonamide fragment **3.10**, and the structures of observed side products **3.41** and **3.42**. *Reagents and conditions*: i) 2-Chloroethanesulfonyl chloride, Et_3N, CH_2Cl_2, −15 °C to 0 °C, 3 h, **3.10**: 4%, **3.41**: 23%, **3.42**: 5%.

To probe the optimal positioning of the electrophile, it was decided that fragments with a longer linker should be investigated. It was therefore decided that a methylene linker would be inserted between the isoxazole KAc mimic and the amide. Retrosynthetic analysis suggested that synthesis of **3.6**, **3.7**, and **3.16** would be tractable through the reduction of a terminal amide at either the 3-, or 5-position of the isoxazole. To achieve this, the ethyl esters **3.18** and **3.17** were subjected to aminolysis with a methanolic solution of ammonia, which gave **3.43** and **3.44** (**Scheme 3.14**). The terminal amides, **3.43** and **3.44**, were then converted to the primary amines through reduction with borane, and were isolated as the hydrochloride salts **3.45** and **3.46**.[148] From these primary amine

intermediates **3.6**, **3.7** and **3.16** could be obtained using the methods described above (**Scheme 3.14**). Generally, these reactions were higher yielding, owing to the increased nucleophilicity of the amine.



Scheme 3.14. Synthesis of fragments **3.6**, **3.7** and **3.16**. *Reagents and conditions*: **A**) i) NH_3, MeOH, rt, 16 h, 38–89%, n=4; ii) a) $BH_3{\cdot}SMe_2$ complex, THF, reflux, 6 h, b) HCl aq, reflux, 1 h, 20–66%, n=2; iii) Acryloyl chloride, K_2CO_3, CH_2Cl_2, 0 °C to rt, 3h, 75%, n=1; iv) a) (*E*)-4-Dimethylaminobut-2-enoic acid, oxalyl chloride, DMF, THF, 0 °C to rt, 2.5 h; b) Et_3N, THF, 0 °C to rt, 16 h, 35%, n=1. **B**) i) NH_3, MeOH, rt, 16 h, 95%, n=1; ii) a) $BH_3{\cdot}SMe_2$ complex, THF, reflux, 6 h, b) HCl aq, reflux, 1 h, 33%, n=1; iii) Acryloyl chloride, Et_3N, CH_2Cl_2, 0 °C to rt, 3 h, 36%, n=1.

3.4 Evaluation of purified BRD4(1)^{L94C} labelling using protein mass spectrometry

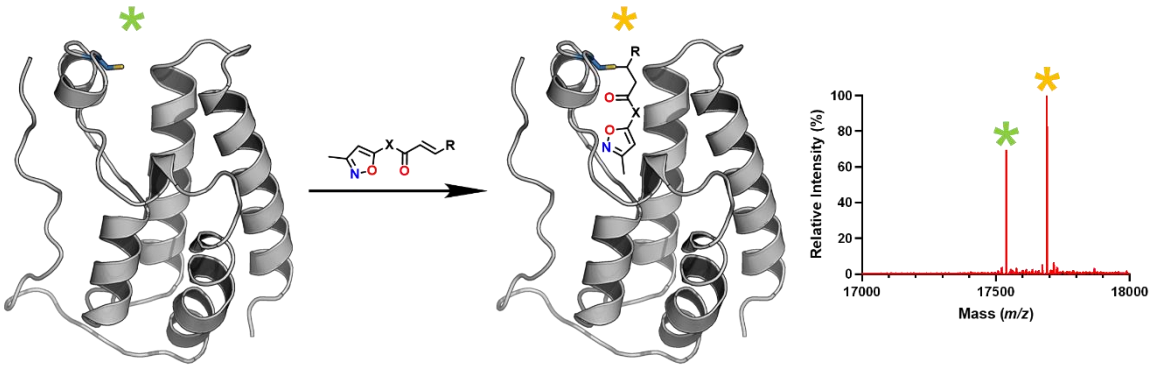


Figure 3.2. Graphical representation of how protein mass spectrometry can be used to identify compounds that covalently bind to proteins.

A protein LC-MS assay was developed to identify fragments able to covalently label BRD4(1)^{L94C}. If a fragment was to covalently bind to the protein, then a shift in the mass of the protein would be observed, corresponding to the mass of the fragment (**Figure 3.2**). It was postulated that if the mass of BRD4(1)^{WT}, possessing two native cysteines, remained unchanged, but a mass shift was observed for BRD4(1)^{L94C}, then the fragment was labelling the mutant C94 residue.

3.4.1 In-house screen

A 5:1 ratio of fragment (50 µM) to protein (10 µM) was determined to be suitable for the assays, as using 5 equivalents of fragment would allow evaluation of any non-specific covalent binding to native cysteines or other residues within the protein. As it was envisioned that these compounds would be taken into a cellular environment, it was decided that incubation at 37 °C would give the most relevant insight into their reactivity at a physiologically relevant temperature. Time points were taken at 0, 1 and 4 hours, diluted, flash-frozen, and stored at −80 °C until analysis. Samples were analysed using an Acquity UPLC system coupled to a Waters Xevo G2-S QTOF mass spectrometer (**Appendix B**). To compare how effectively fragments covalently bound to the protein,

the percentage of covalently labelled protein was calculated using **Equation 3.1**. For multiple binding events the intensity for all peaks were totalled.

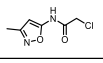
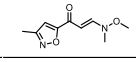
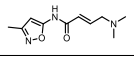
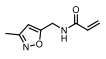
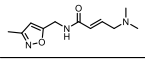
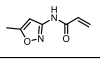
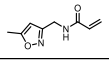
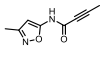
Equation 3.1.

$$Labelling\ (\%) = \left(\frac{Labelled\ Protein\ Intensity}{(Unlabelled\ Protein\ Intensity) + (Labelled\ Protein\ Intensity)}\right) \times 100$$

For a covalent labelling event to be considered a 'hit', a threshold of a 2% peak intensity, when compared to the major peak, was required. Fragments that were shown to covalently label BRD4(1)^{L94C} were then evaluated for covalent binding against BRD4(1)^{WT} to confirm labelling of the mutant C94.

Table 3.2. Results obtained from the protein LC-MS covalent binding assay. The table shows the percentage of labelled protein after incubation at 37 °C for given time points. Results are shown as the mean of three repeats for hits (n=3), or two repeats for non-hits (n=2), and the errors given as the standard error of mean (SEM). The colour scale is shown below, where low levels of protein labelling are shown in blue, and high levels shown in red. ND = Not detectable.

0% Labelling  **100% Labelling**

		Labelled Protein (%)					Labelled Protein (%)		
		L94C		WT			L94C		WT
Compound	Structure	1 h	4 h	4 h	Compound	Structure	1 h	4 h	4 h
3.1		100.0 ± 0.0	100.0 ± 0.0	82.8 ± 2.1	3.41		ND	ND	N/A
3.3		43.8 ± 0.5	42.3 ± 1.2	28.9 ± 0.8	3.11		10.2 ± 0.6	32.2 ± 0.7	ND
3.4		20.7 ± 1.5	55.5 ± 0.7	ND	3.31		ND	ND	N/A
3.5		15.8 ± 0.6	52.2 ± 0.8	ND	3.12		100.0 ± 0.0	100.0 ± 0.0	86.1 ± 7.4
3.6		ND	ND	N/A	3.14		59.6 ± 0.6	57.6 ± 1.5	22.2 ± 1.4
3.7		ND	ND	N/A	3.15		7.6 ± 0.3	9.9 ± 0.4	ND
3.8		3.4 ± 0.2	14.7 ± 1.1	ND	3.16		ND	ND	N/A
3.9		ND	7.8 ± 0.2	ND	3.39		ND	ND	N/A
3.10		ND	ND	N/A					

Analysis of these data revealed key reactivity patterns between the fragments (**Table 3.3**). Strikingly, as predicted, the 3-methyl regioisomer, **3.4**, was found to covalently bind to BRD4(1)^{L94C} to a greater extent than the corresponding 5-methyl regioisomer, **3.15**, where the percentage of

labelled BRD4(1)^{L94C} is 56% and 10%, respectively (**Figure 3.3 A & B)**. This is almost certainly due to the preferred orientation of the isoxazole in the binding pocket, where the 3-methyl group is buried within the binding pocket to facilitate the hydrogen bonding between the isoxazole and the protein, as observed in X-ray crystal structures.

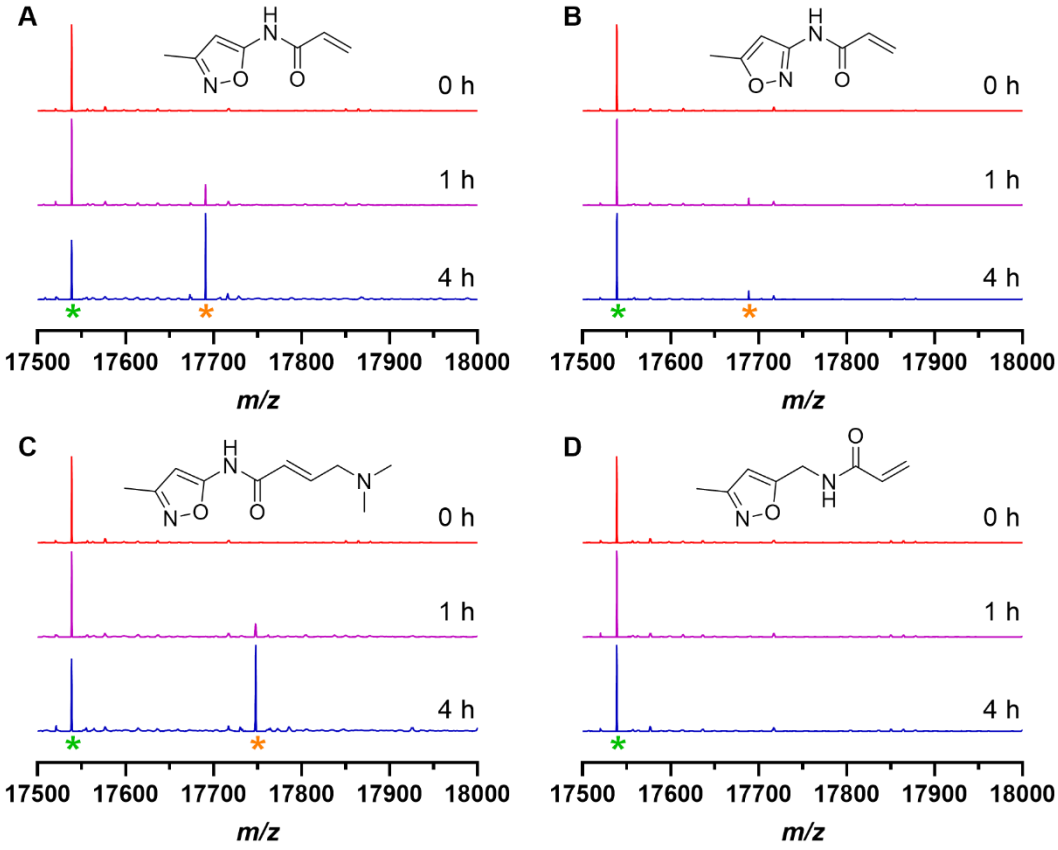


Figure 3.3. Time-dependent protein LC-MS spectra for fragments incubated with BRD4(1)^{L94C}. The green asterisk shows peaks corresponding to unlabelled protein, and the orange asterisks show peaks corresponding to labelled protein. The Y-axis shows relative intensity (%). **A**) Mass spectra for fragment **3.4**. **B**) Mass spectra for fragment **3.15**. **C**) Mass spectra for fragment **3.5**. **D**) Mass spectra for fragment **3.6**.

Compounds **3.4** and **3.5** were found to label over 50% of BRD4(1)^{L94C} (**Figure 3.3 A & C**). When incubated with BRD4(1)^{WT}, both **3.4** and **3.5** showed no covalent labelling, indicating selective labelling of C94. For fragments where the extra methylene group is incorporated, **3.6** and **3.7**, there was no observable labelling of BRD4(1)^{L94C} (**Figure 3.3 D**). These trends can be explained by the localisation of the nitrogen lone pair. For fragments **3.4** and **3.5** the lone pair is in conjugation with the amide and is also delocalised into the isoxazole π-system, leading to greater electrophilicity at the reacting acrylamide. For **3.6** and **3.7**, with the extra methylene linker, the nitrogen lone pair is

only in conjugation with the amide bond, making the acrylamide less electrophilic and therefore less reactive towards the cysteine. The introduction of the sp^3 centre, seen in **3.6** and **3.7**, should be noted. This extra methylene unit will have increased length and flexibility of the fragment, possibly causing the reactive group to be further away in space from the nucleophilic cysteine and such the reaction can not occur.

Although **3.8** selectively labelled BRD4(1)^{L94C}, it was far less reactive than **3.4**, with only 15% of BRD4(1)^{L94C} being labelled (**Figure 3.4 A)**. This is likely due to the C≡C bond being more electron-rich than the C=C bond and hence less reactive. Spatial orientation must also be considered as the C≡C bond orients on a different vector to the C=C bond, and so may align less favourably with the cysteine thiol.

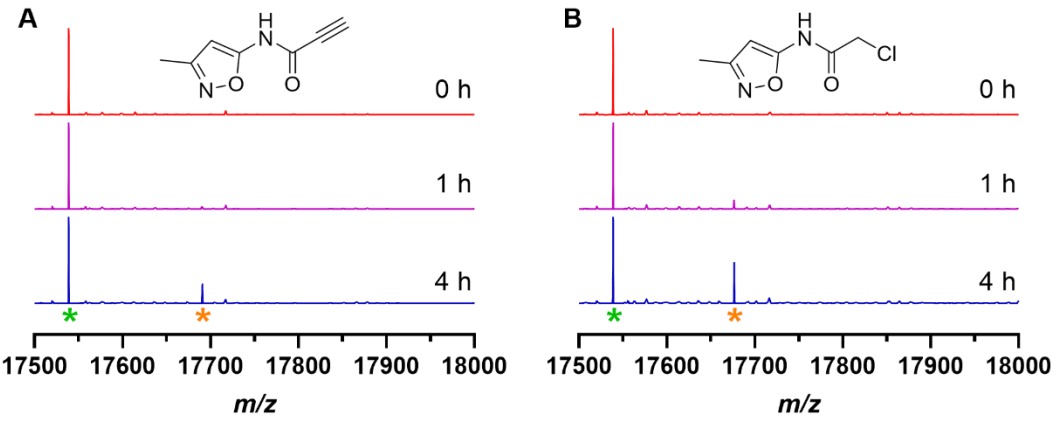


Figure 3.4. Time-dependent protein LC-MS spectra for fragments incubated with BRD4(1)^{L94C}. The green asterisk shows peaks corresponding to unlabelled protein, and the orange asterisks show peaks corresponding to labelled protein. The Y-axis shows relative intensity (%). **A**) Mass spectra for fragment **3.8**. **B**) Mass spectra for fragment **3.11**.

Substitution at the β-position of the electrophile was predicted to decrease the reactivity of the fragments. This was observed for **3.9** where the addition of a methyl group resulted in a 47% reduction in covalent binding when compared to **3.8**. However, for **3.5** such a decrease was not observed. For **3.5**, addition of the DMAM group led to only a 6% decrease in binding, compared to **3.4**. Despite **3.8** and **3.9** displaying selective labelling of C94, they did not produce high enough levels of labelling, and so were not considered fragments of interest.

The chloroacetamide fragment **3.11** was designed to assess the differences in reactivity between fragments binding *via* an S_N2 and a conjugate addition mechanism, while also offering a shorter linker length. Compound **3.11** was found to label 32% of BRD4(1)^{L94C} after 4 hours. Despite concerns about promiscuity, this compound displayed no covalent binding to BRD4(1)^{WT}, suggesting selective labelling of C94 (**Figure 3.4 B)**.

Even though the vinyl sulfonamide group has become a common reactive group in the field of covalent ligands of late, no covalent binding of either **3.10** or **3.41** to BRD4(1)^{L94C} was observed, possibly due to steric clashes caused by the larger sulfonamide group, compared to the amides.

The more electrophilic ketone-linked fragments, **3.1**, **3.3**, **3.12**, and **3.14** were more reactive and promiscuous than their amide counterparts. These compounds labelled both BRD4(1)^{L94C} and BRD4(1)^{WT} , with mass shifts corresponding to multiple binding events observed. For example, protein LC-MS data shows that **3.1** could alkylate BRD4(1)^{L94C} up to six times (**Figure 3.5**). These data reveal that these fragments are labelling more than just cysteine residues, as only three cysteine residues are present in BRD4(1)^{L94C}, indicating that reactions with other nucleophilic residues, such as lysine or serine,[60] must be occurring too. Similarly to the amides, fragments containing the C=C electrophile, **3.1** and **3.12**, were more reactive than the corresponding C≡C fragments, **3.3** and **3.14**.

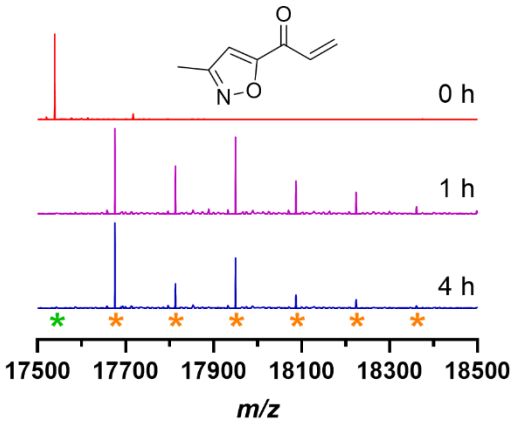


Figure 3.5. Time-dependent protein LC-MS spectra for **3.1** incubated with BRD4(1)^{L94C}. After 4 hours at least six labelling reactions can be observed in the spectra. The green asterisk shows peaks corresponding to unlabelled protein, and the orange asterisks show peaks corresponding to labelling events. The Y-axis shows relative intensity (%).

Compound **3.31** showed no covalent binding to BRD4(1)^{L94C}. This was almost certainly due to the nitrogen lone pair being in conjugation with the enone system, decreasing the electrophilicity of the enone.

The control compound **3.39** exhibited no covalent binding to BRD4(1)^{L94C}, while its unsaturated analogue **3.4** did. This result demonstrates that the covalent bond forming reaction between the fragments and the protein are occurring at the electrophilic reactive groups, and not elsewhere on the molecule.

From the data obtained through the protein LC-MS assay, it was decided that fragments **3.1**, **3.4**, **3.5**, and **3.11** would be progressed to the next stage of probe development with the aim of increasing selectivity and reactivity for the mutant cysteine. This was to be achieved by increasing the size and lipophilicity of the fragments to enhance non-covalent interactions with BRD4(1)^{L94C}. Whilst **3.4**, **3.5** and **3.11** showed a high level of selectivity for C94, **3.1** exhibited high reactivity to nucleophilic residues in both BRD4(1)^{L94C} and BRD4(1)^{WT}. However, it was decided to investigate whether the promiscuity of this hyper-reactive fragment could be reduced by increasing its steric bulk, and so it was also progressed for further development.

3.4.2 Diverse reactive fragment screening platform

To explore whether 'mutate & conjugate' could be used to rapidly discover probes for proteins with no known ligands, a collaboration with Dr Jacob Bush and Dr Emma Grant at GSK was established. BRD4(1)^{L94C} was screened against ~250 previously published reactive fragments, functionalised with either an acrylamide or chloroacetamide.[69]

Using GSK's cysteine reactive screening platform, BRD4(1)^{L94C} (1 μM) was incubated with either a chloroacetamide fragment (50 μM) or an acrylamide fragment (200 μM) at 4 °C for 24 hours. Samples were then analysed using protein LC-MS, where a positive hit was classified as a single

labelling event with a mass shift corresponding to the mass of the fragment (± 2 Da) and a labelled peak of >10% intensity. Because of the difference in screening conditions between GSK and Oxford, a selection of characterised isoxazole fragments (**3.4**, **3.5**, **3.11**, and **3.16**) were also screened to act as positive and negative controls (**Appendix B**).

Of the ~250 reactive fragments screened against BRD4(1)^{L94C}, only two chloroacetamide fragments and two acrylamide fragments were classified as hits (**Figure 3.6**). A further two acrylamide fragments showed significant labelling, however, due to mass discrepancies were omitted. All control compounds were screened and exhibited protein labelling trends that were expected. The low hit frequency also demonstrated that the location of cysteine mutation meant that sufficient affinity for the protein was required to covalently bind to C94. It also confirmed the in-house assay results, indicating that the isoxazole fragments had been designed well to allow for covalent binding to the cysteine mutation.

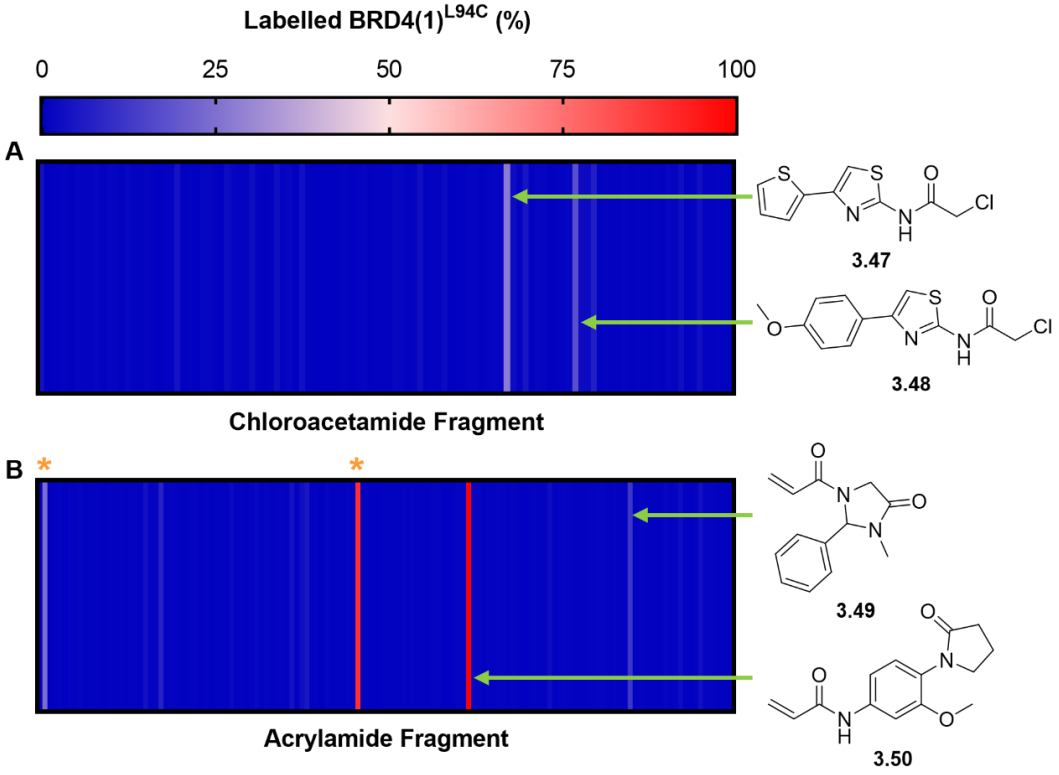


Figure 3.6. Heatmaps showing the results obtained from screening the GSK cysteine reactive fragment library against BRD4(1)^{L94C}, with 'hit' compound structures shown. The orange asterisks show fragments with significant labelling, but due to mass discrepancies did not fulfil the 'hit' criteria. **A**) Heatmap for the chloroacetamide library. **B**) Heatmap for the acrylamide library.

The two chloroacetamide hits **3.47** and **3.48** labelled 27% and 17% of BRD4(1)^{L94C}, respectively (**Figure 3.6 A**). Both **3.47** and **3.48** had been previously identified as promiscuous binders when assessed in a selection of human cell lines, and so appeared as frequent hits in screening assays.[69] Therefore, combined with their low to moderate levels of labelling, it was decided that these compounds would not progress further.

Of the binders identified from the acrylamide library, compound **3.50** was of note (**Figure 3.6 B**). When screened using GSK's reactive screening platform, **3.50** displayed a single covalent labelling event occurring at almost 90% of BRD4(1)^{L94C}. When followed up using the in-house conditions, **3.50** labelled 61% of BRD4(1)^{L94C}, and importantly exhibited no covalent binding to BRD4(1)^{WT}, after a 4 hour incubation at 37 °C (**Figure 3.7**). The lower percentage of protein labelling was to be expected due to the extended incubation period of 24 hours used in the GSK screen. The other fragment hit **3.49** labelled only 11% of BRD4(1)^{L94C}, and so was not progressed.

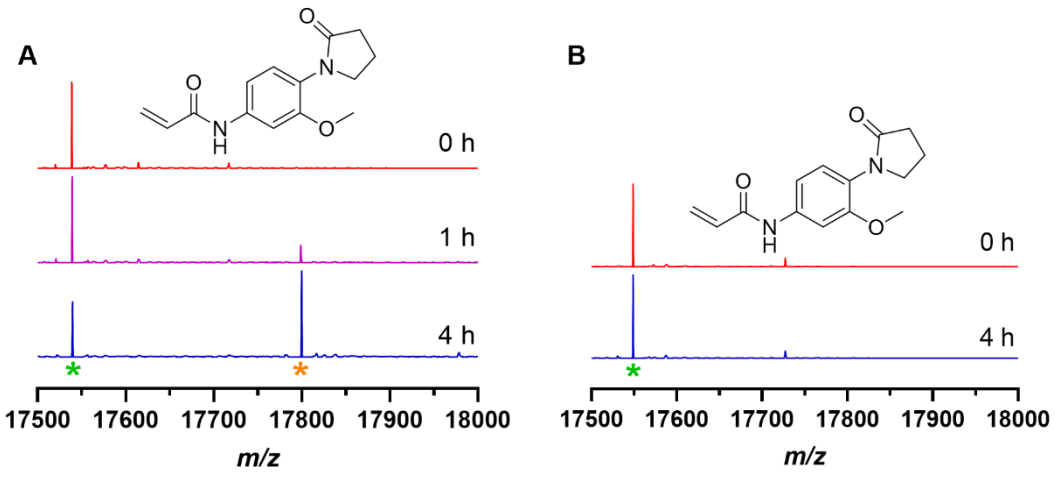


Figure 3.7. Time-dependent protein LC-MS spectra for **3.50** labelling of (**A**) BRD4(1)^{L94C} and (**B**) BRD4(1)^{WT}. The Y-axis shows relative intensity (%). The green asterisk shows peaks corresponding to unlabelled protein, and the orange asterisks show peaks corresponding to labelled protein.

Fragment **3.50** was an exciting find since it is structurally different to the structure-guided isoxazole fragments. Interestingly, both the binders identified from the acrylamide library possess a γ-lactam-type structure, a motif that is known as a KAc mimic, for example the commonly used solvent, *N*-methyl-2-pyrrolidone (NMP) is known to bind to BRD4 with an IC_{50} value of 6 mM.[120,130,153,154] Non-covalent docking studies using GOLD conducted by Bernadette Lee (University

of Oxford) shows three possible binding poses where the γ-lactam of **3.50** acts as a KAc mimic, forming a hydrogen bond to N140, and the phenyl group forming hydrophobic interactions with L92 (**Figure 3.8**). Two of the predicted poses (yellow and magenta) show the methoxy group pointing into a cleft in the binding pocket, while one pose (orange) shows the methoxy group pointing out of the pocket into a solvent exposed region. Interestingly, none of these poses position the acrylamide favourably for reaction with C94, however, rotation around the *N*-phenyl bond may lead to more favourable positioning.

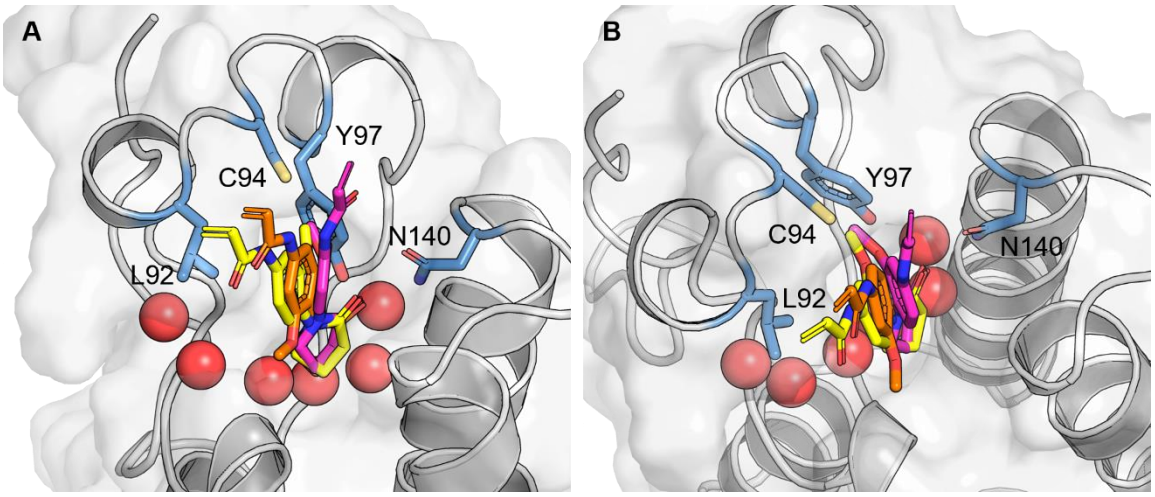


Figure 3.8. Non-covalent docking of **3.50** bound to BRD4(1)^{L94C}. **A**) Docking results predicted three binding poses where the γ-lactam mimics KAc. **B**) A top-down view shows two of the predicted binding poses (yellow and magenta) place the methoxy group pointing into a cleft, while one binding pose (orange) positions the methoxy group pointing out of the binding pocket and into a solvent exposed region.

Additionally, **3.50** does bear some resemblance to the BET bromodomain inhibitor **I-BET151** (**1.62**),[140] and if it is assumed that the γ-lactam of **3.50** is acting as the KAc mimic, then the phenyl ring and methoxy group mirror that of those in **1.62** (**Figure 3.9 A**). Overlay of the X-ray structure of **1.62** bound to BRD4(1) (PDB ID: 3ZYU) with the docking results indicates that, like for the methoxy group of **1.62**, **3.50** may preferentially bind with the methoxy group solvent exposed (**Figure 3.9 B**).[155]

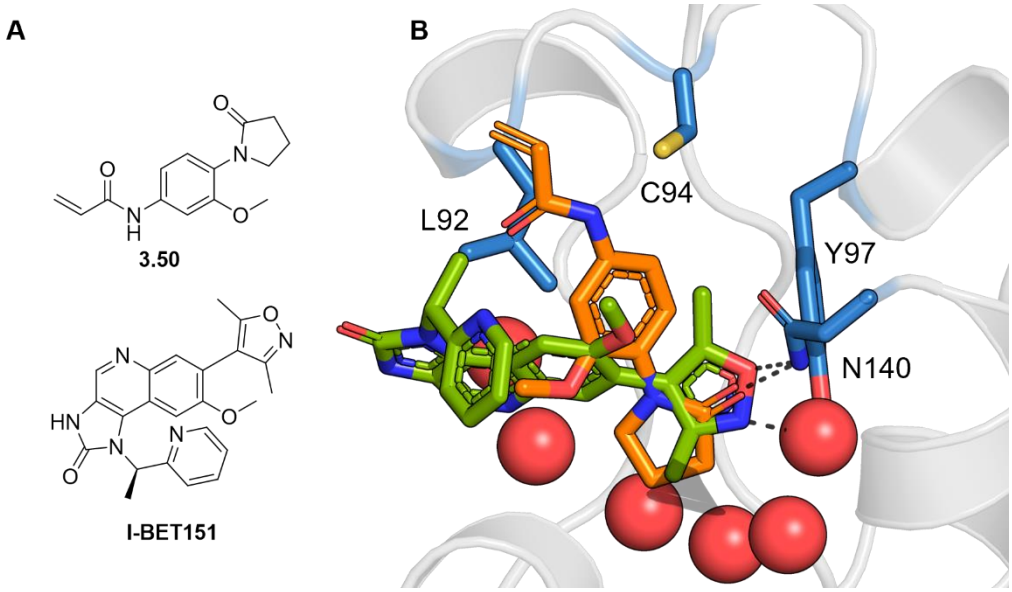


Figure 3.9. Comparing the binding poses of **3.50** (carbon = orange) and **I-BET151** (**1.62**) (carbon = green). **A**) Chemical structures of **3.50** and **1.62**. **B**) Docking results place the carbonyl oxygen and the isoxazole oxygen in the same region, allowing for hydrogen bonding with N140. The methoxy group of **1.62** is solvent exposed, suggesting a similar pose may be adopted for the methoxy group of **3.50**.

The results obtained suggested that **3.50** selectively labels high levels of C94, making it a fragment of interest. To investigate the cellular selectivity of this fragment, it was decided that a clickable analogue for in-cell selectivity studies would be synthesised and tested in human cell lines. Furthermore, the discovery of a selective hit was extremely encouraging, suggesting that 'mutate & conjugate' could be applied to emerging targets, or targets for which only weak binding motifs are known.

3.5 Conclusions

In conclusion, a structure-guided and a diverse fragment-based approach was employed to investigate ~270 reactive fragments for their ability to covalently label C94 of BRD4(1)^{L94C}, as a starting point for the development of a covalent probe. For the structure-guided approach, a library of 17 isoxazole-based fragments were designed and synthesised, comprising a range of electrophiles and linker lengths. In comparison, the diverse fragment-based approach screened ~250 fragments, functionalised with either a chloroacetamide or an acrylamide.

Using protein LC-MS assays, five fragments were identified to be of interest, owing to significant levels of BRD4(1)^{L94C} labelling. It was decided that the four isoxazole fragments, **3.1**, **3.4**, **3.5**, and **3.11**, should be developed further to increase their affinity and selectivity for the BRD4(1)^{L94C} binding pocket, with the hypothesis that this would facilitate greater covalent labelling and inhibition (**Chapter 4**).

Identification of **3.50** from the reactive fragment screening demonstrates that it is possible to discover covalent hits from fragment screening for a protein of interest harbouring a mutant cysteine residue. For fragment **3.50** it was decided that no compound optimisation would be carried out, and instead an 'clickable' analogue would be developed to evaluate in-cell selectivity (**Chapter 5**).

Chapter 4

Optimisation of reactive fragments targeting mutant BRD4(1)^{L94C}

4.1 Introduction and aims

When developing covalent ligands or probes, it is important to also consider the non-covalent interactions that occur between the protein of interest and the molecule. These non-covalent interactions, including hydrogen bonding and lipophilic interactions, are important during the initial stage of ligand binding, where the ligand must display sufficient affinity for the protein to allow both time and suitable positioning of the electrophile for the covalent bond to form (**Figure 4.1 A**). Building on the discovery of the DMI-based covalent fragments, as discussed in **Chapter 3**, the work described in this chapter aims to:

- Increase fragment affinity for BRD4(1)^{L94C} through enhanced non-covalent interactions with the protein.
- Improve the physicochemical properties of the fragments for in-cell assays.
- Optimise the installation of the DMAM modified acrylamide.

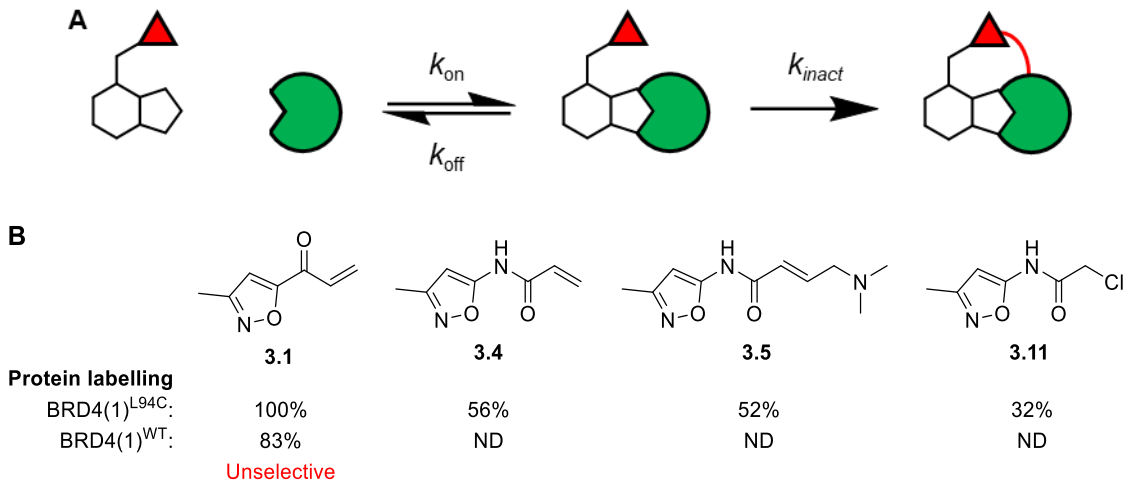


Figure 4.1. A) Mechanism of covalent inhibition. Binding to the target protein is initially driven by ligand affinity through non-covalent interactions. Non-covalent interactions are important and hold the ligand in position for the covalent bond formation to occur. **B**) The DMI-based covalent fragments of interest discovered in **Chapter 3**.

4.2 Design of optimised covalent fragments

To increase the affinity of the fragments for BRD4(1)^{L94C} the introduction of a simple modification to the isoxazole fragments was envisioned. With the essential methyl group occupying the 3-position of the isoxazole, and the electrophile at the 5-position, it was decided to functionalise the 4-position. To do this, a 4-phenyl modification was chosen— a group that has been shown to increase DMI affinity for BRD4(1) (**Figure 4.2 A**).[130–132] It was hypothesised that this lipophilic extension may enhance affinity through increased lipophilic interactions, while also providing extra steric bulk that may increase selectivity.

Additionally, with the aim of evaluating these optimised fragments in live cells, it was important to consider the physicochemical properties, such as lipophilicity.[23,24,156] The lipophilicity of a compound (often expressed as LogD) impacts its ability to permeate cell membranes and engage with intracellular targets. Compounds with LogD>5 are highly hydrophobic and typically exhibit poor aqueous solubility. Compounds that possess LogD<0 are very aqueous soluble, however are less cell permeant. Compounds with moderate LogD 0–3 generally show a desirable balance of solubility and cell permeability.[157] Calculated LogD at pH 7.4 ($cLogD_{pH7.4}$) (MarvinSketch 21.14), predicted that the 4-phenyl modification would generally lead to a ~1.5 increase in $cLogD_{pH7.4}$ compared to the parent fragments, potentially indicating a desirable increase in cell permeability (**Figure 4.2 A**).

Non-covalent docking studies using AutoDock4, in collaboration with Dr Joseph Bluck, evaluated the binding poses of **4.1**, **4.2**, and **4.3** to BRD4(1)^{L94C}. At the time of carrying out this work the X-ray crystal structure of BRD4(1)^{L94C} had not yet been obtained, and so BRD4(1)^{WT} (PDB ID: 6CZU) was mutated to BRD4(1)^{L94C} using the PyMOL mutagenesis wizard.[135] Results for the docking studies predicted that the addition of a phenyl group would retain the KAc mimicking interactions with N140 and Y97 while also allowing for the formation of lipophilic interactions between the phenyl group and the sidechain of L92, similar to those seen for other 4-phenyl DMI ligands (**Figure 4.2 B**,

C, **& D**).[130–132] Interestingly, **4.3** was also predicted to form a hydrogen bond to the backbone carbonyl of L92, through the protonated DMAM group, indicating that **4.3** might exhibit high affinity for BRD4(1)^{L94C} (**Figure 4.2 D**).

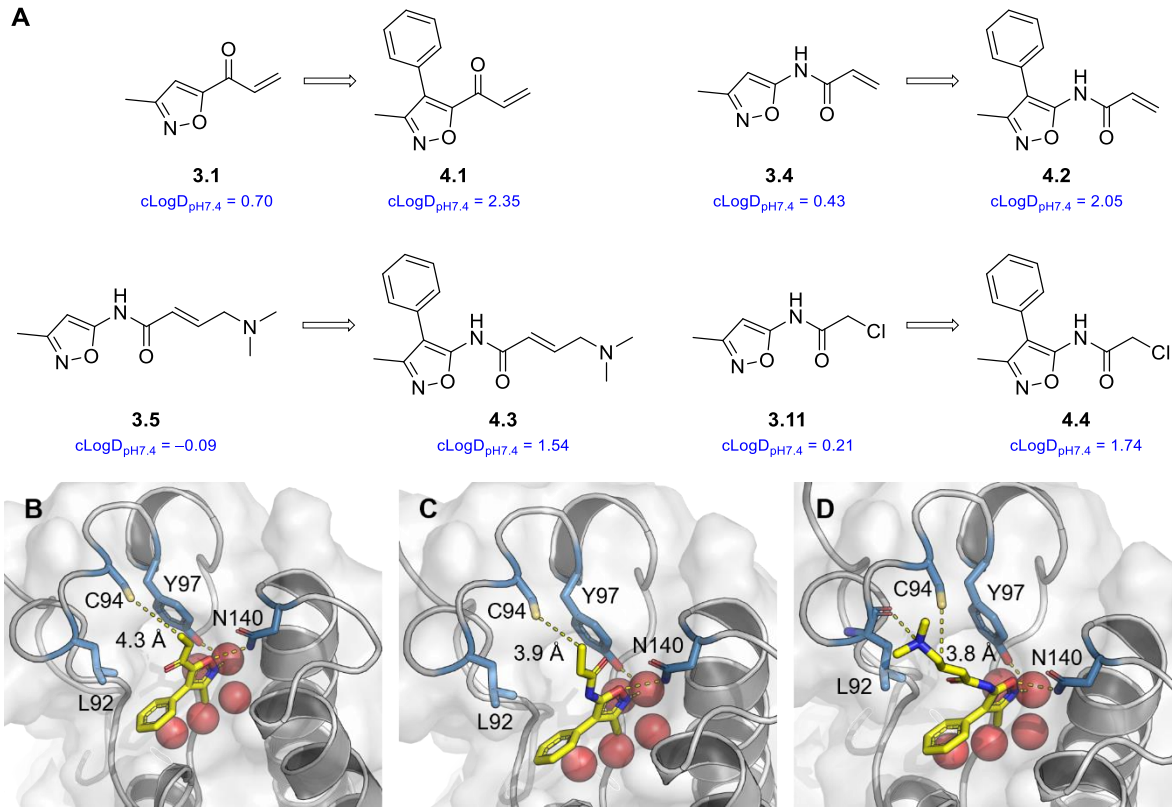


Figure 4.2. Optimisation of covalent fragments targeting C94 of BRD4(1)^{L94C}. **A**) Chemical structures of the 4-phenyl DMI covalent fragments with corresponding cLogD$_{pH7.4}$ values shown in blue. Docking studies show that the DMI core of **4.1** (**b**), **4.2** (**C**), and **4.3** (**D**) acts as a KAc mimic to BRD4(1)^{L94C}, and the 4-phenyl modification forms lipophilic interactions with L92. **4.3** (**D**) is predicted to hydrogen bond to the backbone carbonyl of L92.

Given the promising results from the initial acrylamide fragments **3.4** and **3.5**, it was desirable to gain a greater understanding of how substitution at the β-position of the acrylamide could influence reactivity and selectivity. Results from Lanning *et al.*,[64] (Described in **Chapter 1.2**) found that a β-methyl modification to the acrylamide of the Bruton's tyrosine kinase (BTK) inhibitor **Ibrutinib** (**1.19**) led to a significant increase in selectivity for BTK in proteome-wide selectivity studies (**Figure 4.3 A**). Inspired by these results a β-methyl substituted acrylamide analogue, **4.5**, was to be synthesised (**Figure 4.3 B**).

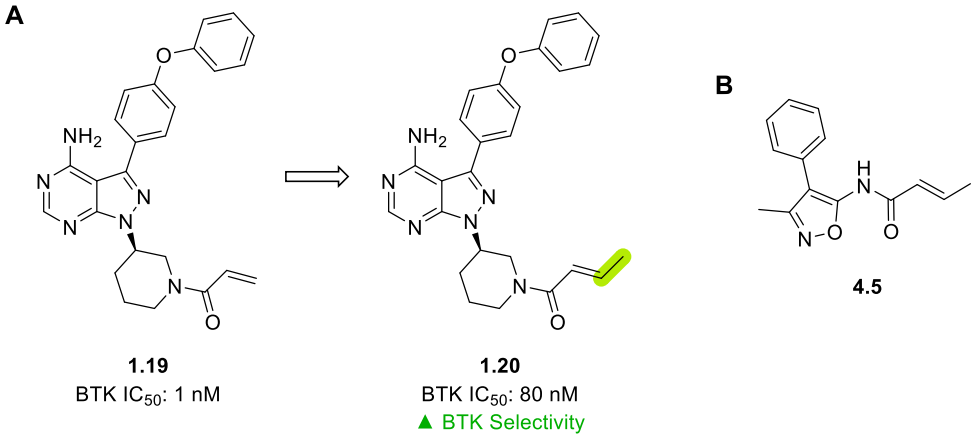


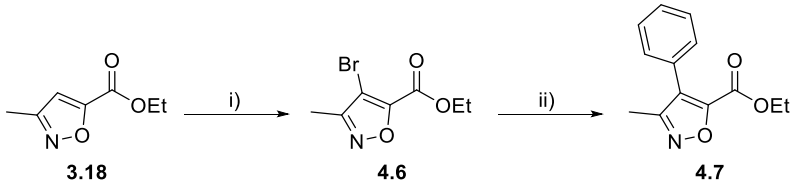
Figure 4.3. Substitution of acrylamides has shown to tune selectivity and reactivity of covalent inhibitors. **A**) Addition of a methyl group at the β-position of **Ibrutinib** (**1.19**) to give **1.20** greatly increased BTK selectivity.[64] **B**) The chemical structure of **4.5**.

4.3 Optimised covalent fragment synthesis

4.3.1 Ketone-linked fragment

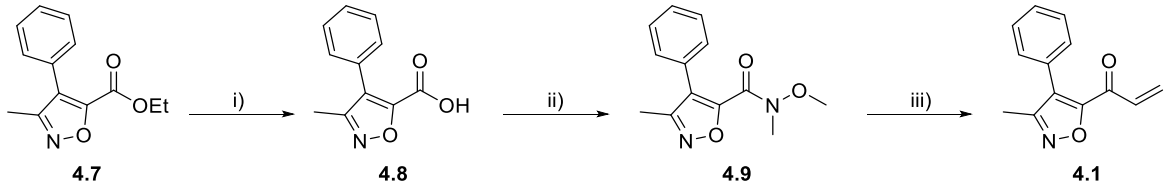
The first aim was to synthesise the ketone-linked fragment **4.1**. To install the phenyl group at the 4-position of the isoxazole a Suzuki-Miyaura coupling approach was employed (**Scheme 4.1**). This method is also compatible with simple changes to the aryl group, for example, coupling of an alkyne-bearing aryl boronic acid, to obtain a 'clickable' probe for evaluation of in-cell selectivity.

To synthesise the Suzuki-Miyaura coupling substrate **4.6**, the 4-position of **3.18** was brominated (**Scheme 4.1**). Due to the low reactivity of the 4-position to bromination, the reaction was carried out under forcing conditions using *N*-bromosuccinimide (NBS) in trifluoracetic acid (TFA),[148,158] and proceeded in a 65% yield. Using a palladium-catalysed cross coupling, phenylboronic acid was coupled to **4.6** to synthesise **4.7** in a 49% yield.



Scheme 4.1. Synthesis of intermediate **4.7**. *Reagents and conditions*: i) NBS, TFA, 150 °C, microwave, 50 min, 65%, n=1; ii) Phenylboronic acid, Pd(dppf)Cl_2·CH_2Cl_2, $NaHCO_3$, DME, H_2O (4:1), 90 °C, 2.5 h, 49%, n=1.

The ethyl ester **4.7** was hydrolysed to reveal the carboxylic acid **4.8**. This compound was coupled to *N*,*O*-dimethylhydroxylamine, using the previously described conditions, to obtain the Weinreb amide **4.9** in a 98% yield (**Scheme 4.2**). The final step of the synthesis used a Grignard reaction with vinylmagnesium bromide to form **4.1**. This proceeded cleanly in a 62% yield, with no impurities resulting from conjugate addition observed.



Scheme 4.2. Synthesis of compound **4.1**. *Reagents and conditions*: i) LiOH, THF, H_2O (2:1), rt, 2 h, 84%, n=1; ii) *N*,*O*-Dimethylhydroxylamine hydrochloride, PyBOP, Et_3N, CH_2Cl_2, 0 °C to rt, 16 h, 98%, n=1; iii) Vinylmagnesium bromide, THF, −35 °C to 0 °C, 1 h, 62%, n=1.

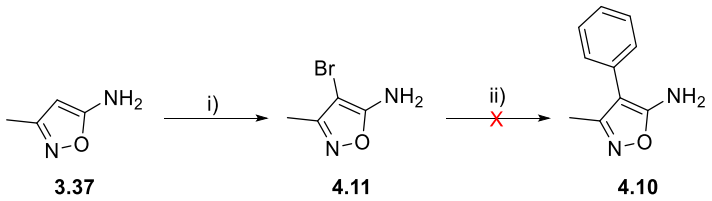
4.3.2 Amide-linked fragments

4.3.2.1 Synthesis of the amine-containing scaffold

Having employed the Suzuki-Miyaura coupling for the synthesis of **4.1**, this approach was attempted for the synthesis of the amine-containing scaffold **4.10** (**Scheme 4.3**). Using NBS, commercially available **3.37** was brominated to form **4.11** in a 67% yield following a known procedure (**Scheme 4.3**).[159] This bromination proceeded under milder conditions than those used to brominate the ethyl ester analogue, owing to the more electron rich C-4 position in **3.37**.

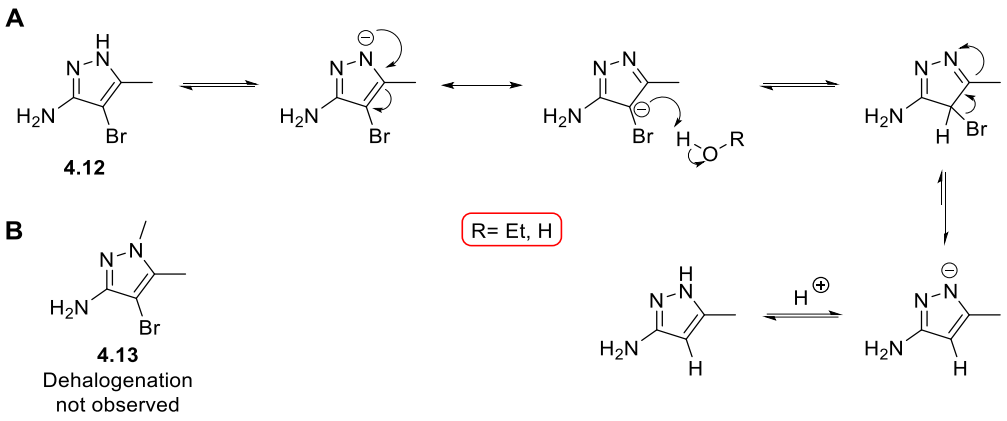
A literature search found that there were no published Suzuki-Miyaura couplings between an aryl boronic acid and a 3-methylisoxazol-5-amine halogenated at the 4-position. However, Jedinák *et al.* had previously reported optimisation of Suzuki-Miyaura coupling conditions for halogenated aminopyrazoles **4.12**, a problematic substrate due to the pyrazole NH and an unprotected amine.[160] This work overcame the problem of dehalogenation of the starting material. Following optimisation, they found that using an XPhos ligand, potassium carbonate as the base, degassed ethanol and water, and heating in a microwave reactor at 135 °C for 20 minutes, gave the greatest

yields. Inspired by these conditions, a small scale (50 mg, 0.284 mmol) test reaction using XPhos Pd G3 was undertaken (**Scheme 4.3**). It was found that for **4.11** there was no observable formation of **4.10**, however, significant dehalogenation of the isoxazole was seen. The temperature was lowered to 80 °C in the hope that this would decrease the rate of dehalogenation, but this was not the case.



Scheme 4.3. Attempted synthesis of **4.10**. *Reagents and conditions*: i) NBS, $CHCl_3$, 0 °C, 80 min, 67%, n=1; ii) Phenylboronic acid, XPhos Pd G3, XPhos, K_2CO_3, EtOH, H_2O (4:1), microwave, 135 °C, 15 min.

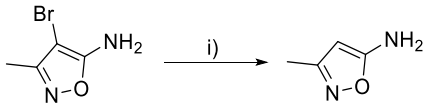
To investigate the mechanism of dehalogenation, Jedinák *et al.* monitored the stability of **4.12** in their proposed reaction conditions, but without catalyst and ligand. Whilst dehalogenation of the aminopyrazole substrate **4.12** was readily observed, this was completely suppressed for the corresponding *N*-methyl analogue **4.13**, leading them to propose a mechanism where dehalogenation was driven by deprotonation of the mildly acidic pyrazole NH (**Scheme 4.4**).



Scheme 4.4. A) The proposed mechanism for dehalogenation of aminopyrazole **4.12**.[160] **B**) Dehalogenation was not observed for the *N*-methyl analogue **4.13**.

This observation was intriguing, as substantial dehalogenation of the aminoisoxazole **4.11** substrate was also observed, despite the isoxazole lacking an acidic pyrazole NH. The stability of **4.11** in the reaction conditions was monitored by heating at 80 °C, in a 4:1 ethanol: water mix, and with

2 equivalents of potassium carbonate (**Scheme 4.5**). After 30 minutes degradation of **4.11** to **3.37** was observed by TLC, and after 24 hours complete degradation of **4.11** had occurred, with the majority of the degradation product corresponding to the dehalogenated amine **3.37**.

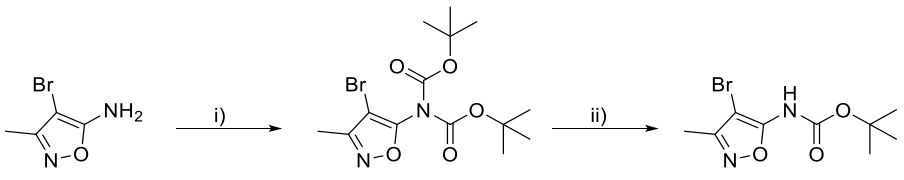


4.11 **3.37**

Scheme 4.5. TLC analysis revealed that dehalogenation of **4.11** occurred in the reaction conditions even without the presence of catalyst or ligand. After 24 hours **4.11** had completely degraded. *Reagents and conditions*: i) K_2CO_3, EtOH, H_2O (4:1), 80 °C, 24 h.

Next, it was investigated whether protecting the free amine would be beneficial for the cross coupling. It was hoped that protection of the free amine would alleviate coupling problems in two ways. Firstly, Boc protection of the amine should decrease the delocalisation of the amines lone pair into the isoxazole, and therefore decrease the amount of dehalogenation observed. Secondly, a known issue with palladium cross coupling reactions is that free amines can coordinate to the palladium and lower catalyst activity.

To synthesise the Boc protected amine **4.14**, compound **4.11** was heated under reflux in the presence of di-*tert*-butyl decarbonate (Boc_2O), and 4-dimethylaminopyridine (DMAP), in THF (**Scheme 4.6**).[148] After 24 hours the starting material had been consumed and a single product had formed (as judged by TLC). LC-MS analysis revealed a mass consistent with the di-Boc protected amine **4.15**. This was subjected to hydrolysis to form the mono-Boc protected amine **4.14** in a 93% yield, with no further purification required.

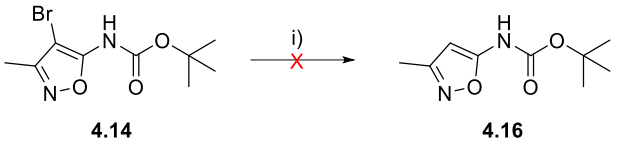


4.11 **4.15** **4.14**

Scheme 4.6. Synthesis of the Boc protected intermediate **4.14**. *Reagents and conditions*: i) Boc_2O, DMAP, THF, reflux, 24 h; ii) NaOH aq, MeOH, rt, 48 h, 93% over two steps, n=1.

To assess whether the protection of the amine had impacted the rate of dehalogenation, Boc protected **4.14** was subjected to the reaction conditions at 80 °C without the presence of catalyst

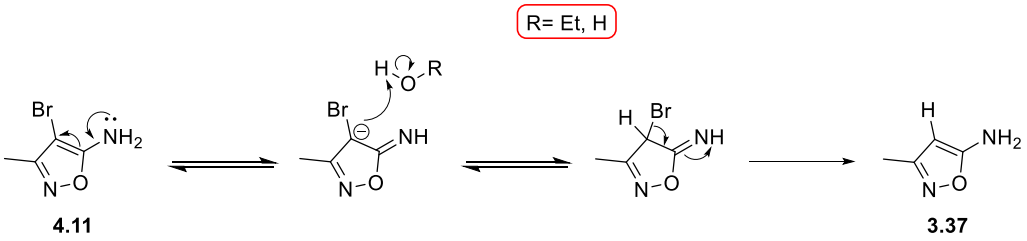
or ligand, as previously described. After 1 hour of heating no observable degradation had occurred (as judged by TLC), and after 24 hours only minor amounts of uncharacterised degradation products had formed (**Scheme 4.7**).



4.14 **4.16**

Scheme 4.7. TLC analysis was used to monitor dehalogenation of **4.14**. After 24 hours only minor amounts of degradation products had formed. *Reagents and conditions*: i) K_2CO_3, EtOH, H_2O (4:1), 80 °C, 24 h.

This result suggests the lone pair of the amine plays an important role in the dehalogenation of **4.11**. These data led to the proposed mechanism that helps explain the observed results (**Scheme 4.8**).

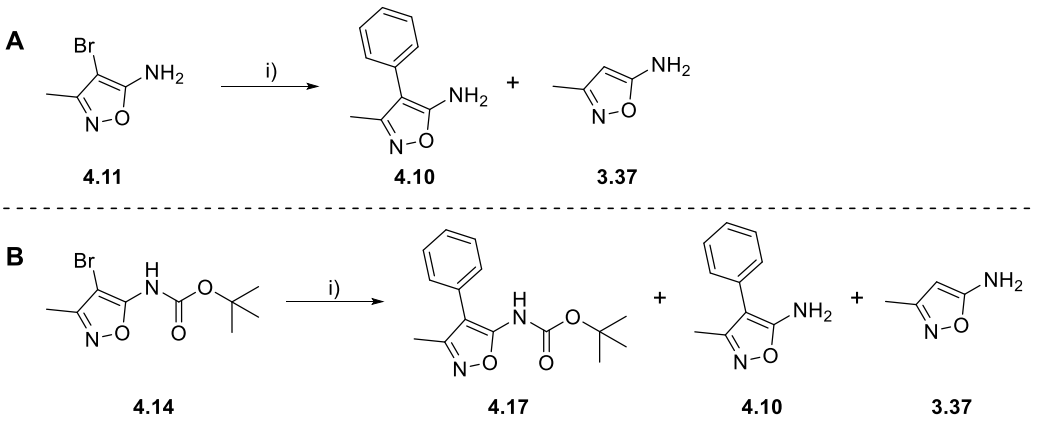


4.11 **3.37**

Scheme 4.8. The proposed mechanism for the dehalogenation of **4.11**. Significant dehalogenation was not observed for the Boc protected analogue **4.14** and therefore it was hypothesised that the amine lone pair is required for the formation of **3.37**.

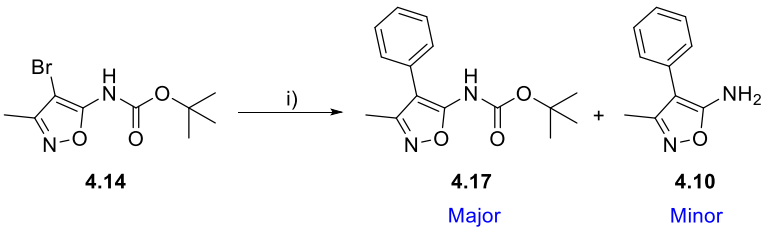
Encouraged by the increased stability of the Boc protected amine **4.14**, the Suzuki-Miyaura coupling was attempted, using the previously reported conditions with XPhos Pd G3.[160] After 20 minutes at 135 °C in a microwave reactor, a mixture of compounds was identified, however, only trace amounts of **4.17** had formed (**Table 4.1**). Interestingly **4.10** was isolated in a 10% yield. The presence of **4.10** suggests that the cross-coupling reaction had occurred, forming **4.17**, which presumably decomposed to form **4.10**, suggesting that the Boc group was not sufficiently stable at 135 °C. These data confirmed that the reaction did proceed, as long as the amine remained protected.

Table 4.1. Reaction conditions attempted for the Suzuki-Miyaura coupling of phenylboronic acid with **4.11** (**A**) and **4.14** (**B**). *Reagents and conditions*: i) Phenylboronic acid, XPhos Pd G3, XPhos, K_2CO_3, EtOH, H_2O (4:1).



		Reaction Products		
Substrate	**Temperature (°C)**	**4.10**	**3.37**	**4.17**
4.11 (A)	135	✗	✓	N/A
	80	✗	✓	N/A
4.14 (B)	135	✓	✓	✗
	80	✓	✓	✓

To decrease the rate of Boc deprotection, the temperature was lowered to 80 °C. After 26 hours a mixture of species had formed (**Table 4.1**). Although **4.17** was obtained in a 24% yield, the other species formed were consistent with those observed at 135 °C. Due to the low yield and complicated purification, other methods for synthesising scaffold **4.10** were explored.

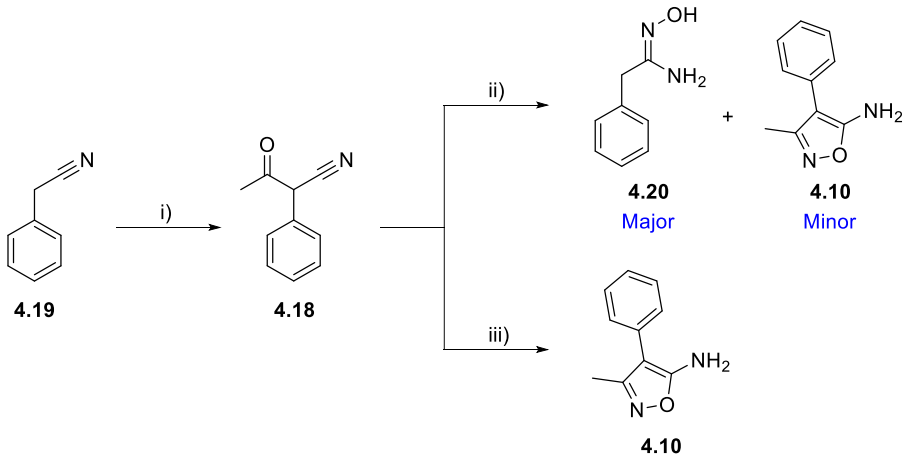


Scheme 4.9. Synthesis of **4.17** and **4.10** *via* a Suzuki-Miyaura cross coupling. *Reagents and conditions*: i) Phenylboronic acid, XPhos Pd G3, XPhos, K_2CO_3, EtOH, H_2O (4:1), 80 °C, 18 h, **4.17**: 24%, **4.10**: 10% .

A procedure from Krasavin *et al.* reported the reaction between **4.18** and hydroxylamine, to form the ketoxime *in situ*, which then undergoes an intramolecular cyclisation to form the desired product **4.10** (**Scheme 4.10**).[161] The publication reported yields over 90% with no purification of the product required.

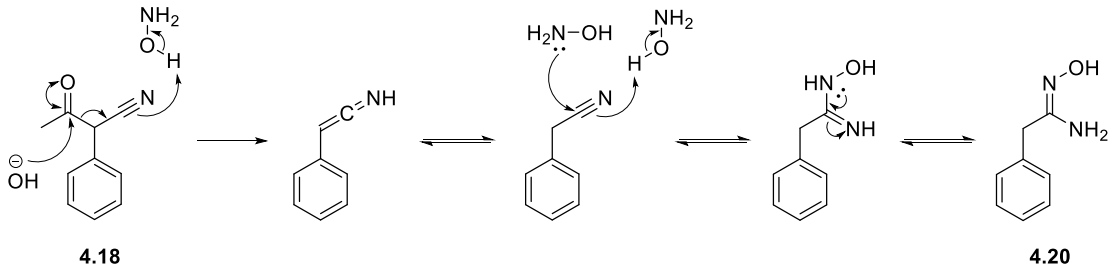
The ketonitrile starting material **4.18** was synthesised *via* a Claisen-type condensation reaction between phenylacetonitrile **4.19** and ethyl acetate, using sodium hydride, in THF (**Scheme 4.10)**.[162] It was found that using 10% DMF as a cosolvent increased the yield from 38% to 69%, likely due to increased solubility of the deprotonated phenylacetonitrile salt.

With the ketonitrile **4.18** in hand, and following the known procedure,[161] **4.18** was heated under reflux with 50% aqueous hydroxylamine solution, and 10% aqueous sodium carbonate, in ethanol (**Scheme 4.10 Conditions ii**). After 5 hours, TLC analysis indicated complete consumption of the starting material **4.18**, and the formation of two new products. Isolation of both compounds revealed a 7% yield of the desired isoxazole **4.10**, while an unknown side product was isolated in 73% yield. The side product was determined to be **4.20**. A key feature in determining the structure was a sharp peak in the ^{1}H NMR downfield at 8.89 ppm that was present in D_6-DMSO, corresponding to the exchangeable oxime proton.



Scheme 4.10. Synthesis of **4.10** *via* a condensation reaction with hydroxylamine. *Reagents and conditions*: i) EtOAc, NaH, THF, 10% *v/v* DMF, 0 °C to rt, 18 h, 60–69%, n=5; ii) Hydroxylamine (50% *w/v* in H_2O), 10% aq. Na_2CO_3, EtOH, reflux, 5 h, **4.20**: 73%, **4.10**: 7%, n=1; iii) Hydroxylamine hydrochloride, 10% aq. Na_2CO_3, EtOH, reflux, 1 h, 81–100%, n=6.

It was hypothesised that the ketonitrile undergoes a fragmentation reaction to re-form the nitrile through a retro-Claisen condensation,[163] which is accelerated by elevated temperatures and high pH. Nucleophilic attack by the hydroxylamine NH_2 on the nitrile then occurs to yield the observed amidoxime **4.20** (**Scheme 4.11**).



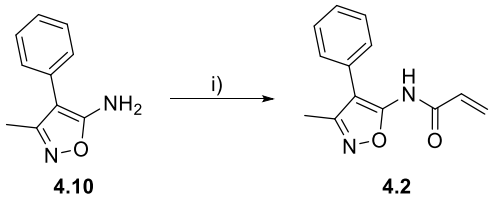
4.18 **4.20**

Scheme 4.11. Proposed mechanism for the formation of **4.20**.

The synthesis of **4.10** was attempted again, this time using hydroxylamine hydrochloride salt, rather than the 50% aqueous hydroxylamine solution. Using these conditions (**Scheme 4.10 Conditions iii**), yields of **4.10** ranged from 81–100%, and generally no purification was required. Scaffold **4.10** could now be synthesised on gram scales in high yields.

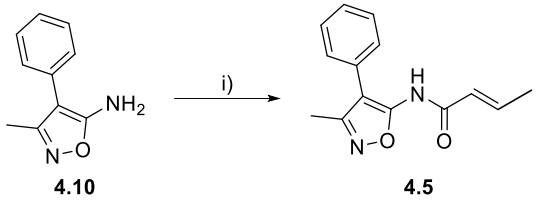
4.3.2.2 Installation of electrophilic groups

Compound **4.2** was synthesised from the appropriate acyl chloride which is commercially available. Using the conditions employed for parent fragment **3.4**, the synthesis of **4.2** was attempted. These conditions used triethylamine as the base. Despite dropwise addition of acryloyl chloride at 0 °C, diacylation of the amine was the major product, despite this phenomenon not being observed for **3.4**. It was suspected that the amide of **4.2** might be more acidic than **3.4**, and therefore may increase the chance of diacylation in the presence of triethylamine. The amide NH p*K*a values were calculated for both **4.2** and **3.4**, with predicted values of 7.77 and 7.85 (MarvinSketch 21.14), respectively. These data indicated that **4.2** was more acidic than **3.4**, and so pyridine was employed as a weaker base. Due to the poor nucleophilicity of **4.10** it was also proposed that the pyridine would act as a nucleophilic catalyst. Although changing the base decreased the amount of diacylation observed, **4.2** was only isolated in yields of 10–13% despite clean reaction profiles, as judged by TLC (**Scheme 4.12**). These low yields could be attributed to the compound's instability upon silica gel, however, this was not investigated further as adequate amounts of compound were obtained for biological testing.



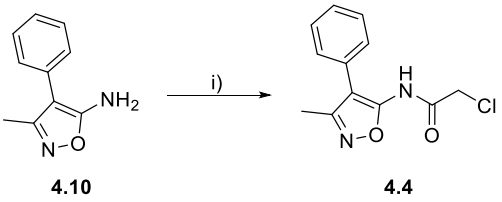
Scheme 4.12. Synthesis of compound **4.2**. *Reagents and conditions*: i) Acryloyl chloride, pyridine, CH_2Cl_2, 0 °C to rt, 20 h, 10–13%, n=3.

Similar to the synthesis of **4.2**, yields of 13% were obtained for the β-methyl analogue **4.5** using crotonoyl chloride. Again, enough compound meeting biological testing standards was obtained, and so no further optimisation was carried out.



Scheme 4.13. Synthesis of optimised covalent fragment **4.5**. *Reagents and conditions*: i) Crotonoyl chloride, pyridine, CH_2Cl_2, 0 °C to rt, 20 h, 13%, n=1.

Commercially available chloroacetyl chloride was coupled to **4.10** using the conditions employed for the synthesis of acrylamides **4.2** and **4.5**, to give **4.4** in a 24% yield (**Scheme 4.14**).



Scheme 4.14. Synthesis of optimised covalent fragment **4.4**. *Reagents and conditions*: i) Chloroacetyl chloride, pyridine, CH_2Cl_2, 0 °C to rt, 3 h, 24%, n=1.

Given the promising data obtained for fragment **3.5** it was decided to optimise the installation of the DMAM substituted acrylamide, as **3.5** was only obtained in a 7% yield and required multiple rounds of purification. To do this, three installation strategies were explored, to establish the most efficient route to **4.3** (**Figure 4.4**).

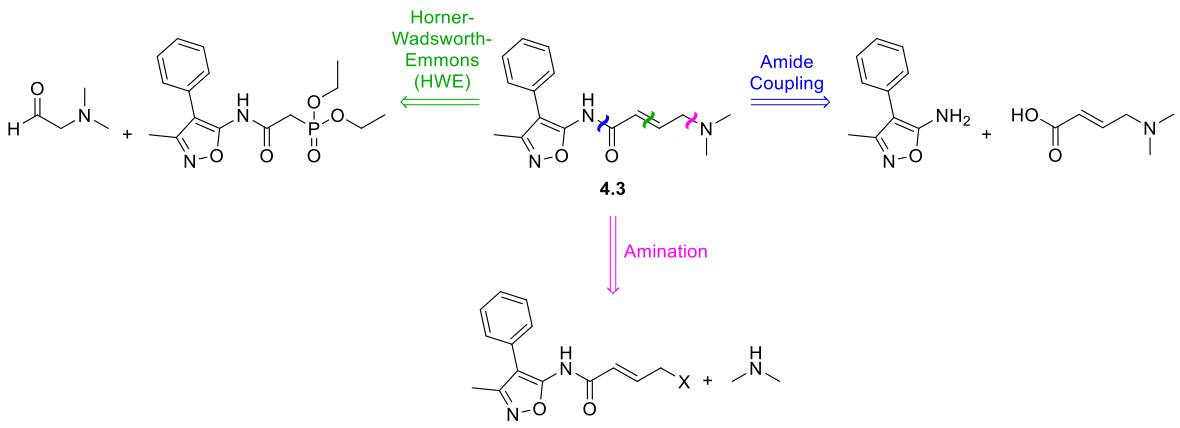
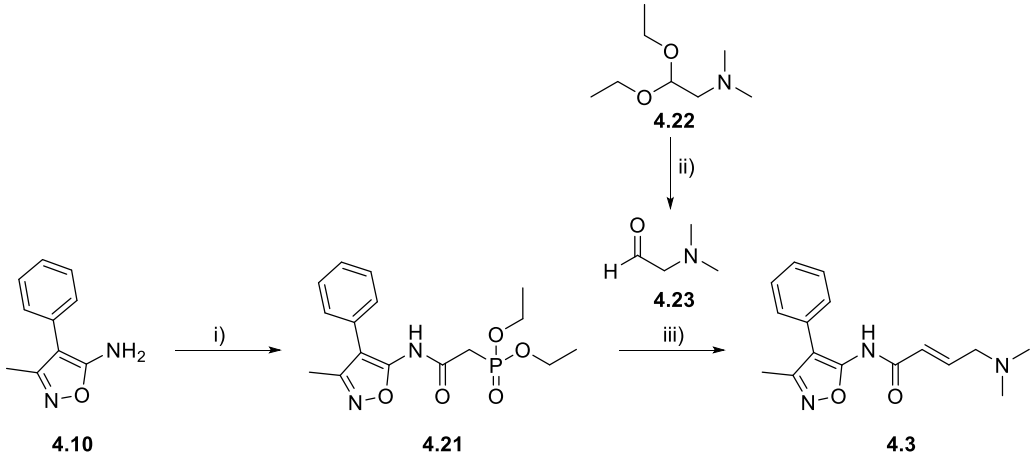


Figure 4.4. Retrosynthetic disconnections of **4.3**. To optimise the installation of the DMAM substituted acrylamide three synthetic routes were investigated, which were based on a direct amide coupling, an S_N2 amination, and a Horner-Wadsworth-Emmons reaction.

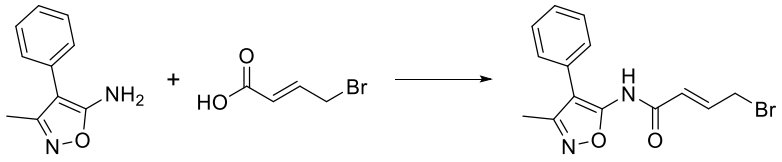
It was first decided to explore the synthesis of **4.3** through the formation of the alkene bond. There was literature precedence in process chemistry patents, for using the Horner–Wadsworth–Emmons (HWE) reaction to selectively install the *E*- substituted DMAM acrylamide.[164,165] To synthesise the require phosphonate **4.21**, amine **4.10** and diethylphosphonoacetic acid were coupled using carbonyldiimidazole (CDI) to form **4.21** in 85% yield (**Scheme 4.14**). To generate the aldehyde **4.23**, acetal **4.22** was hydrolysed under acidic conditions and added to the phosphonate under basic conditions. Despite the patents reporting high yields for the HWE step, the reaction proved challenging to push to completion, and although the reaction proceeded relatively cleanly (as judged by TLC), yields of only 5–17% were obtained for the *E*- substituted DMAM acrylamide **4.3** (**Scheme 4.15**).



Scheme 4.15. Synthesis of **4.3** *via* an HWE approach. *Reagents and conditions*: i) Diethylphosphonoacetic acid, CDI, THF, 40 °C, 38 h, 68–85%, n=2; ii) HCl aq (6.0 M), reflux, 4 h; iii) NaOH aq (50% *w*/*v*), LiCl, THF, 0 °C to rt, 26 h, 5–17%, n=2.

The attention was next turned to exploring a disconnection through the amine bond. Using the commercially available (*E*)-4-bromobut-2-enoic acid **4.24** a coupling reaction to **4.10** to form **4.25** was envisioned, which could then be subjected to an S_N2 amination with dimethylamine to install the amine (**Scheme 4.16**). It was first decided to optimise the coupling between (*E*)-4-bromobut-2-enoic acid **4.24** and **4.10**. To achieve this, seven different amide coupling reagents and conditions for the synthesis of **4.25** were explored (**Table 4.2**).

Table 4.2. Amide coupling reagents and conditions investigated for the synthesis of **4.25**.

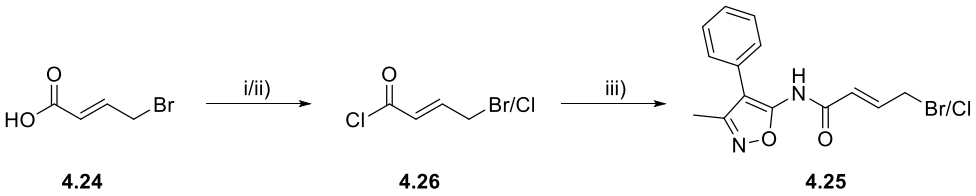


4.10 **4.24** **4.25**

Coupling Reagent	Base	Solvent	Time (h)	4.25 detected (LC-MS)
HATU	DIPEA	DMF	24	✗
CDI	DBU	THF	16	✗
DCC	DMAP	CH_2Cl_2	48	✗
EDC	DMAP	CH_2Cl_2	16	✗
CMPI	DIPEA	CH_2Cl_2	16	✗
DPPA	Et_3N	DMF	20	✗
BTFFH	DIPEA	CH_2Cl_2	19	✗

None of the conditions investigated gave the desired product, and so the acyl chloride of **4.24** was generated using thionyl chloride. This reaction of **4.26** with **4.10** produced **4.25** in a 39% yield, despite the diacylated product also being formed (**Scheme 4.16**). Interestingly, NMR and LC-MS analysis revealed that the isolated compound was a 1:1 mixture of the chloro- and bromo-substituted product. It was hypothesised that the formation of both the bromo- and chloro-substituted species might reduce the yield of **4.25**. To investigate this, the acyl bromide was generated using thionyl bromide, which would only produce the bromo-substituted product. This reaction proceeded poorly, likely due to the greater instability of the acyl bromide compared to the

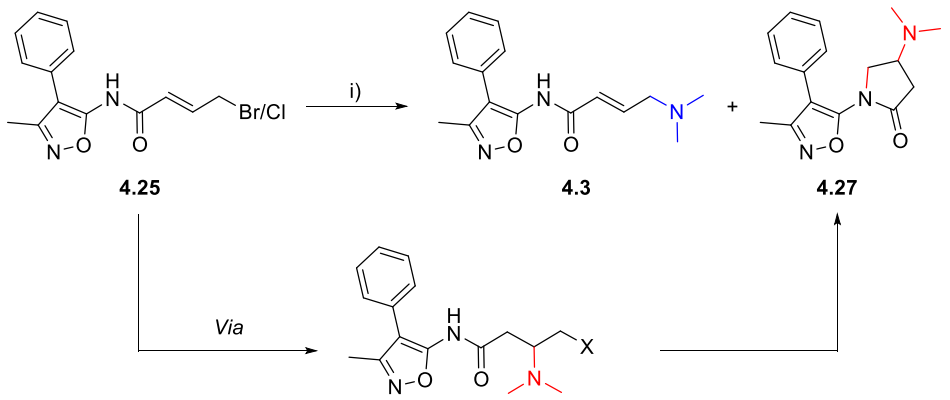
acyl chloride, and so it was decided that the acyl chloride would be used going forward. Later preparations, however, switched from using thionyl chloride to oxalyl chloride, as this gave a cleaner reaction and generated lower amounts of the chloro-derivative. Due to the inseparability of the bromo- and chloro-derivatives, all isolated mixtures were progressed to the amination step.



4.24 **4.26** **4.25**

Scheme 4.16. Synthesis of intermediate **4.25**. *Reagents and conditions*: i) Thionyl chloride, 0 °C to rt, 16 h; ii) Oxalyl chloride, DMF, THF, 0 °C to rt, 18 h; iii) **4.10**, pyridine, CH_2Cl_2, 0–5 °C , 4 h, 39%, n=1.

The final step in the synthesis of **4.3** was to install the substituted amine using dimethylamine. To suppress competing reactions, such as the 1,4-conjugate addition of the nucleophilic dimethylamine, sub-stoichiometric amounts of tetrabutylammonium iodide (TBAI) and sodium iodide were used to catalyse the substitution pathway (**Scheme 4.17**). Compound **4.3** was isolated in a 14% yield, and whilst attempts had been made to suppress competing reactions the product of a conjugate addition was also observed. Interestingly, upon the 1,4-conjugate addition of the dimethylamine nucleophile, the now saturated, and flexible, chain was able to undergo an intramolecular γ-lactamisation to form **4.27** (**Scheme 4.17**). This structure was confirmed by NMR and LC-MS analysis (**Appendix C**).

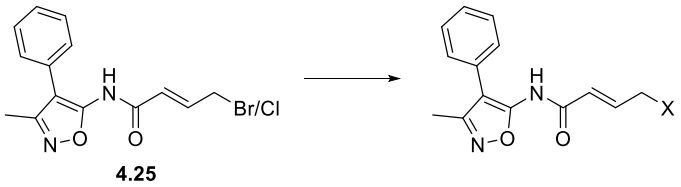


4.25 **4.3** **4.27**

Scheme 4.17. Initial synthesis of **4.3** *via* amination. *Reagents and conditions*: i) a) NaI, THF, rt, 30 min; b) Dimethylamine (2.0 M in THF), TBAI, K_2CO_3, 0 °C, 14 h, 14%, n=1. The product of a 1,4-conjugate addition followed by an intramolecular γ-lactamisation **4.27** was also isolated from the reaction.

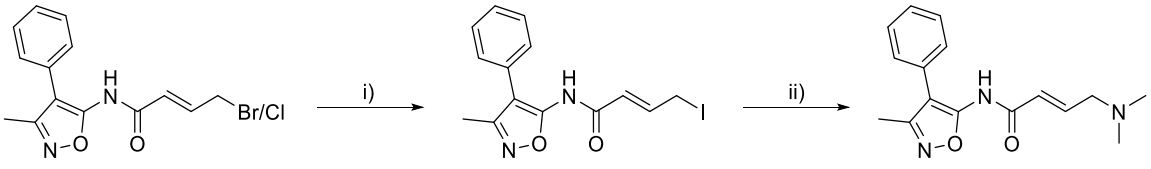
It was postulated that full conversion of **4.25** to the iodide would promote the amination and lead to a cleaner reaction. It was therefore desirable to understand whether using stoichiometric amounts of sodium iodide or TBAI in DMF would lead to full conversion to the iodide. Using LC-MS, the ratios of the halogenated products were determined, as these were separable by reverse phase column chromatography (**Table 4.3**). For both TBAI and sodium iodide in DMF the resulting species were a mixture of all three products. The classic Finkelstein reaction was next investigated as an alternative method, with acetone as the solvent. After 1 hour of heating at 50 °C full conversion to the iodide **4.28** was confirmed through LC-MS and NMR analysis. Isolation of the iodide **4.28** was achieved by filtration to remove the insoluble NaCl and NaBr salts. This intermediate was then taken directly onto the amination step with dimethylamine, which proceeded in an 31% yield over the two steps to give **4.3** (**Scheme 4.18**). The overall yield from scaffold **4.10** to **4.3** was 12%.

Table 4.3. Iodination conditions investigated for the formation of **4.28** as determined by LC-MS.



4.25

				Halogenated Products (% Conversion)		
Reagent	**Solvent**	**Temp (°C)**	**Time (h)**	**Cl**	**Br**	**I**
NaI	DMF	rt	1	28	29	43
TBAI	DMF	rt	1	25	37	38
NaI	Acetone	50	1	0	0	100



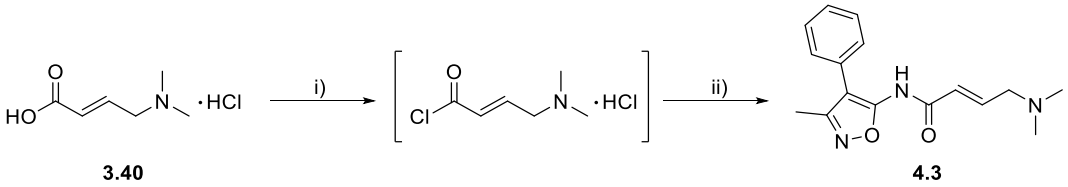
4.25 **4.28** **4.3**

Scheme 4.18. Synthesis of optimised covalent fragment **4.3** *via* amination with isolation of **4.28**. *Reagents and conditions*: i) NaI, acetone, 50 °C, 50 min, taken directly onto the next step; ii) Dimethylamine (2.0 M in THF), K_2CO_3, DMF, 0 °C, 1 h, 31%, n=1.

The final method for the synthesis of **4.3** was to explore the direct amide coupling of the electrophile to the isoxazole scaffold. Whilst this approach had previously been used for the low yielding synthesis of **3.5**, it was decided to optimise the conditions. The previous synthesis had

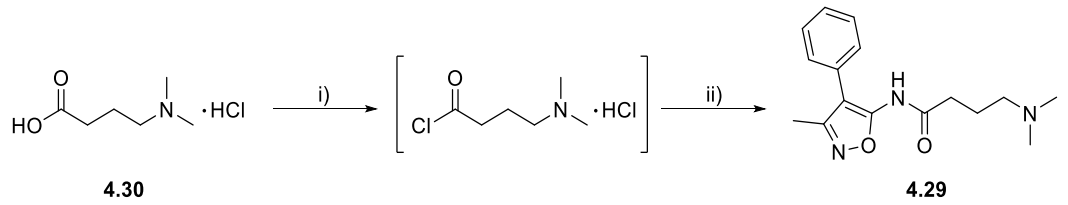
involved forming the acyl chloride and isolating it by removing the volatile components, before redissolving the acyl chloride and adding it to the amine solution. A detailed patent procedure from Chew *et al*. found higher coupling yields when the acyl chloride was not isolated,[166] and instead the amine solution added to the acyl chloride solution. They also noted that higher yields were observed when the amine was dissolved in NMP rather than THF or acetonitrile, they hypothesised that this was because of the increased solubility of the reactants in this solvent.

Encouraged by the patents results, the procedure for the synthesis of **4.3** was attempted. Whilst the coupling was slow and had to be kept at 5 °C overnight, a yield of 33% after purification was obtained (**Scheme 4.19**). This was a pleasing result and delivered the highest overall yield of **4.3** over the least number of steps.



Scheme 4.19. Synthesis of **4.3** *via* direct amide coupling. *Reagents and conditions*: i) Oxalyl chloride, DMF, THF, 0 °C to rt, 1.5 h; ii) **4.10**, NMP, 0–5 °C, 21 h, 33%, n=1.

Using the same procedure, the saturated control compound **4.29** was synthesised by coupling amine **4.10** to 4-(dimethylamino)butyric acid **4.30** in a 60% yield (**Scheme 4.20**).



Scheme 4.20. Synthesis of saturated control fragment **4.29**. *Reagents and conditions*: i) Oxalyl chloride, DMF, THF, 0 °C to rt, 2 h; ii) **410**, NMP, 0–5 °C, 20 h, 60%, n=1.

Overall, the synthesis of **4.10** and the installation of the DMAM substituted acrylamide were optimised to develop robust synthetic routes— important for future derivatisation.

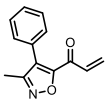
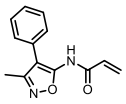
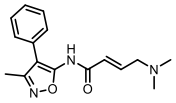
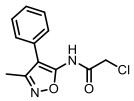
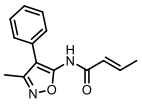
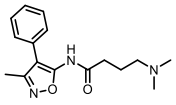
4.4 Biological evaluation

4.4.1 Evaluation of purified BRD4(1)^{L94C} labelling using protein mass spectrometry

The previously described protein mass spectrometry assay was used to evaluate the ability of the optimised fragments to covalently label C94 of BRD4(1)^{L94C} (**Table 4.4** & **Appendix C**).

Table 4.4. Results obtained from the protein LC-MS covalent binding assay. The table shows the percentage of labelled protein after incubation at 37 °C for given time points. Results are shown as the mean of three repeats for hits (n=3), or two repeats for non-hits (n=2), and the errors given as the standard error of mean (SEM). The colour scale is shown below, where low levels of protein labelling are shown in blue, and high levels shown in red. ND = Not detectable.



		Labelled Protein (%)			
		L94C		WT	
Compound	Structure	1 h	4 h	1 h	4 h
4.1		100.0 ± 0.0	100.0 ± 0.0	100.0 ± 0.0	N/A
4.2		69.8 ± 0.5	97.5 ± 0.1	ND	3.0 ± 0.2
4.3		26.9 ± 1.8	74.1 ± 2.1	ND	ND
4.4		5.9 ± 0.2	21.8 ± 0.3	ND	ND
4.5		ND	ND	N/A	N/A
4.29		ND	ND	N/A	N/A

The 4-phenyl modification of both **4.2** and **4.3** increased labelling compared to their parent fragments **3.4** and **3.5**, respectively. This increase in covalent labelling suggests that the additional phenyl group has increased the affinity of **4.2** and **4.3** for the KAc binding pocket, through lipophilic interactions with the protein, or has led to more favourable positioning for covalent bond

formation. Fragment **4.3** exhibited 74% labelling after 4 hours, an increase from 52% seen for **3.5**, whilst also maintaining selective labelling of C94. Fragment **4.2** labelled 97% of BRD4(1)^{L94C} over 4 hours, compared to 56% for **3.4**. However, low levels of non-specific labelling were also observed, and after 4 hours **4.2** had labelled 3% of BRD4(1)^{L94C} twice. This result was consistent with the 3% labelling observed for BRD4(1)^{WT}. While increasing the affinity for the KAc binding pocket, the increased lipophilicity of **4.2** may promote interactions and binding to another location on the proteins surface. It is hypothesised that this may be occurring at a site with a neighbouring nucleophilic residue that is capable of reacting with the acrylamide, and therefore leading to a second covalent binding event. Electronically, the phenyl modification could have also increased the electrophilicity of the acrylamide by withdrawing electrons from the acrylamide and hence slightly increasing its reactivity.

Interestingly, **4.5** displayed no covalent labelling of BRD4(1)^{L94C}. There are two plausible explanations for this: firstly, that the β-methyl substitution has significantly decreased the electrophilicity of the acrylamide through donation of the methyl group's electrons into the acrylamide. Or secondly, that the methyl group sterically blocks the reaction from occurring or the ligand from binding, however, this is unlikely, as substitution with the DMAM group does not abrogate protein labelling as seen for **4.3**. Therefore, it is most likely that the electronics argument is the favoured explanation.

In **Chapter 3** chloroacetamide fragment **3.11** was found to label 32% of BRD4(1)^{L94C}, however, fragment **4.4** displayed decreased labelling, with an average of 22% after 4 hours. This might be caused by the increased non-covalent interactions obstructing optimal positioning of the chloroacetamide for covalent bond formation.

Unsurprisingly, **4.1** displayed high reactivity towards both BRD4(1)^{WT} and BRD4(1)^{L94C}, resulting in multiple binding events observable by mass spectrometry. Intriguingly, after only 1 hour, BRD4(1)^{L94C} mass peaks corresponding to five covalent labelling events could be seen, whilst after

4 hours peaks of up to only three labelling events could be seen (**Figure 4.5**), causing a large decrease in total protein intensity, by LC-MS. These data suggest that covalent labelling of multiple sites by **4.1** causes protein instability, leading to these highly labelled proteins precipitating out of solution, and thus accounting for the decrease in the intensity of the mass peaks corresponding to three or more labelling events after 4 hours (**Figure 4.5**). Additionally, reliable quantification for the BRD4(1)^{WT} 4 hour time point could not be achieved due to distorted protein peaks and high levels of background noise, further indicating that multiple labelling by **4.1** causes protein instability.

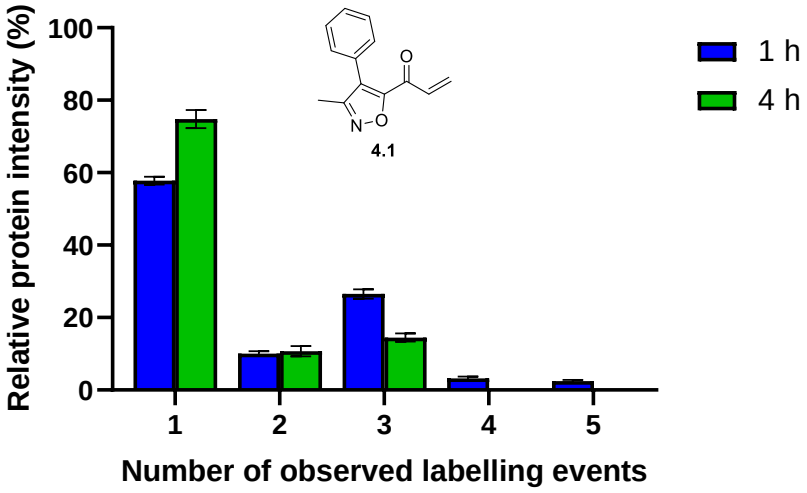


Figure 4.5. The number of labelling events observed by protein LC-MS and their relative respective intensities at 1 and 4 hours for **4.1** when incubated with BRD4(1)^{L94C}.

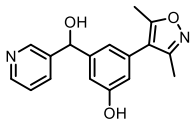
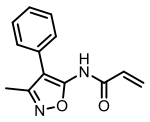
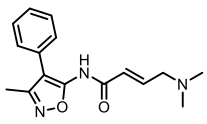
4.4.2 AlphaScreen™ assay

In collaboration with Dr Jessica Reynolds, compounds **4.2** and **4.3** were evaluated for inhibition of BRD4(1)^{L94C} and BRD4(1)^{WT} using a dose-response AlphaScreen™, starting at a compound concentration of 500 µM and serially diluting 1:2 to give 14 concentrations. The assay plate also included **OXFBD04** (**1.58**) as a positive control.[131,132] AlphaScreen™ is a displacement-based assay,[167,168] where ligand binding results in displacement of a biotinylated peptide from the protein giving a reduction in emission. In this case, the displacement of $H4_{1\text{-}20}(KAc)_4$ from BRD4(1)^{L94C} and BRD4(1)^{WT} by **4.2** and **4.3** was being evaluated.[167]

Compounds were incubated at room temperature for 3.5 hours in the presence of the biotinylated peptide $H4_{1-20}(KAc)_4$, and either BRD4(1)^{L94C} or BRD4(1)^{WT}. After this time the donor and acceptor beads were added, and the assay incubated for a further 30 minutes in the dark, before the assay was run. Compounds were run in triplicate, and the data analysed using GraphPad Prism 9. IC_{50} values were calculated for compounds which displayed sigmoidal curves, those that did not were deemed to have IC_{50} values of >100 µM (**Appendix C**).

Table 4.5. A table showing IC_{50} values for **OXFBD04** (**1.58**), **4.2**, and **4.3** against BRD4(1)^{WT} and BRD4(1)^{L94C}, determined using a dose-response AlphaScreen™. The colour scale is shown below, where low levels of protein inhibition are shown in blue, and high levels shown in red. The errors are shown as the standard error of the mean (SEM).

Weaker Binding — Stronger Binding

		AlphaScreen™ IC_{50} (µM)	
Compound	**Structure**	**L94C**	**WT**
OXFBD04		0.080 ± 0.009	0.183 ± 0.025
4.2		1.27 ± 0.243	>100
4.3		4.44 ± 0.785	>100

OXFBD04 (**1.58**) was found to inhibit both BRD4(1)^{L94C} and BRD4(1)^{WT}, as expected (**Table 4.5**). **1.58** displayed a greater inhibitory effect against BRD4(1)^{L94C} (IC_{50} = 80 nM) compared to BRD4(1)^{WT} (IC_{50} = 183 nM), which is likely due to a slight decrease in the affinity of the peptide for BRD4(1)^{L94C}.

Both **4.2** (IC_{50} = 1.27 µM) and **4.3** (IC_{50} = 4.44 µM) displayed inhibition of peptide binding to BRD4(1)^{L94C}, while also exhibiting IC_{50} values of >100 µM for BRD4(1)^{WT}. These data show that **4.2** is a more potent covalent inhibitor of BRD4(1)^{L94C} than **4.3**. These findings are in agreement with the mass spectrometry results, where **4.2** exhibits 97% protein labelling after 4 hours, compared to 74% for **4.3**.

4.5 Conclusions

The work in this chapter explored the development of the isoxazole fragments that bind to BRD4(1)^{L94C}. The focus was to introduce a simple phenyl modification at the 4-position of the isoxazole with the aim of increasing affinity for the protein through lipophilic interactions, while also improving cLogD$_{pH7.4}$ values. The designed compounds were synthesised, and significant optimisation and mechanistic investigation was carried out for the synthesis of the key scaffolds and electrophiles. Installation of the DMAM substituted acrylamide was optimised, increasing the yield from 7% to 33%. This optimisation will be useful for the synthesis of future **4.3** derivatives**,** for example, a clickable analogue for proteome selectivity studies.

The compounds were evaluated for their ability to covalently label C94 of BRD4(1)^{L94C} using the aforementioned protein mass spectrometry assay. Both **4.2** and **4.3** exhibited a marked increase in their ability to label C94 compared to their parent fragments. This was a strong indicator that ligand affinity for the KAc binding pocket had been increased. To confirm this, the ability of **4.2** and **4.3** to inhibit binding of the H4$_{1–20}$(KAc)$_4$ peptide to BRD4(1)^{WT} and BRD4(1)^{L94C} was assessed using an AlphaScreen™ assay. Both **4.2** and **4.3** exhibited low micromolar IC$_{50}$ values for BRD4(1)^{L94C}, 1.27 μM and 4.44 μM, respectively. No inhibition of BRD4(1)^{WT} was observed for either **4.2** or **4.3**.

These data are extremely promising and demonstrate how the combination of a carefully mapped cysteine mutation and a complementary covalent fragment could yield a compound that exhibits selective inhibition of the mutant protein over the wild type.

Chapter 5

Clickable probes to evaluate in-cell selectivity of reactive fragments targeting BRD4(1)^{L94C}

5.1 Introduction and aims

Work described in **Chapters 3** and **4** demonstrated that fragments **4.2**, **4.3**, and **3.50** exhibit high levels of BRD4(1)^{L94C} labelling, while displaying little or no labelling of BRD4(1)^{WT}. This indicates that these fragments might possess a degree of selectivity for BRD4(1)^{L94C} in a cellular environment, making them suitable for phenotypic studies (**Figure 5.1**). Building on these observations, the work described in this chapter aimed to:

- Assess the in-cell selectivity of fragments **4.2**, **4.3**, and **3.50** for BRD4(1)^{L94C} through the development of analogues possessing a clickable handle for fluorescent labelling studies.
- Confirm in-cell target engagement with BRD4(1)^{L94C} using competition assays and immunoprecipitation.

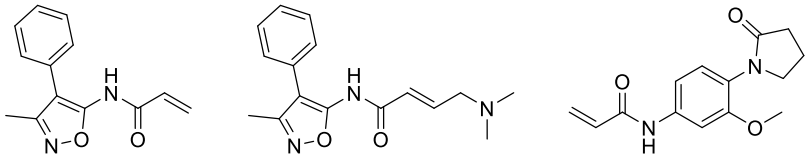


Figure 5.1. Chemical structures of covalent fragments targeting BRD4(1)^{L94C}.

5.2 Click probe design

Although there are numerous bioorthogonal click chemistry strategies, as described in **Chapter 1.2**, it was decided that a CuAAC approach would be employed, due to its robustness, versatility, and efficiency in aqueous solutions (**Figure 5.2**). Whilst the use of a copper catalyst is toxic to live cells,[93,94,169,170] it was envisioned that the click chemistry would be carried out in the cell lysate, and so copper toxicity was not an issue. Therefore, to develop the covalent fragments into clickable probes, **4.2**, **4.3**, and **3.50** would be functionalised with an alkyne moiety, suitable for CuAAC reaction with an azide-functionalised reporter molecule, such as a fluorescent dye or affinity tag.

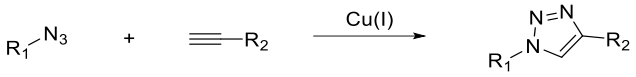


Figure 5.2. A copper-catalysed alkyne azide cycloaddition (CuAAC).[92]

5.2.1 Isoxazole scaffold

Initially, it was decided to build the alkyne handle directly attached to the 4-position of the phenyl ring. This could be achieved through a Sonogashira coupling between the corresponding aryl halide and TMS protected acetylene which could then be deprotected to reveal the alkyne handle (**Figure 5.3**).

Despite the simplicity of this approach, it was postulated that the alkynes rigidity and short distance from the main scaffold, might not be favourable for undergoing the click reaction, due to possible steric clashes with the protein. Therefore, a more flexible alkyne handle, located more remotely from the main ligand, was investigated. This analogue would be synthesised by an S_N2 reaction between propargyl bromide and a phenol bearing isoxazole (**Figure 5.3**).

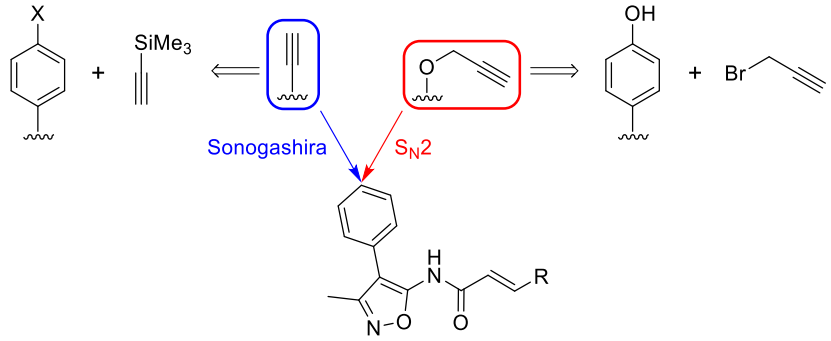


Figure 5.3. Design of clickable probes based on optimised fragments **4.3** and **4.2**. Two alkyne variations were explored, utilising either a Sonogashira coupling or an S_N2 approach in their synthesis.

From the data presented in **Chapter 4**, it appeared that the DMAM substituted acrylamide **4.3** showed greater selectivity for C94, compared to its unsubstituted counterpart **4.2**, likely due to decreased reactivity. These data, therefore, suggested that **4.3** might show greater in-cell selectivity, and so two clickable **4.3** analogues, **5.1** and **5.2** were the main focus (**Figure 5.4**). However, to compare in-cell selectivity, one clickable **4.2** analogue, **5.3**, was also designed, utilising

the propargyl handle. The corresponding saturated compounds **5.4**, **5.5**, and **5.6** would also be synthesised as negative controls to demonstrate the role of the Michael acceptor for target engagement (**Figure 5.4**).

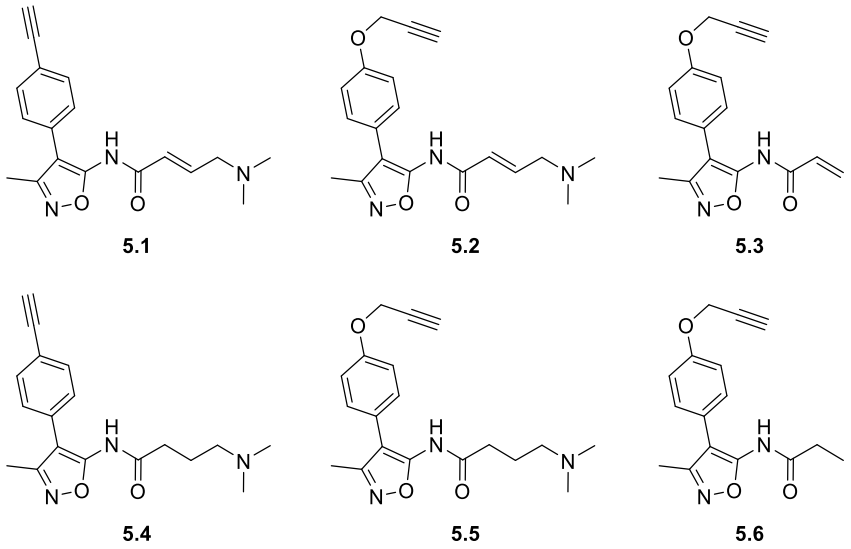


Figure 5.4. Target DMI-based clickable covalent probes and the corresponding saturated negative control compounds.

5.2.1 γ-Lactam scaffold

Assuming that the γ-lactam of **3.50** was acting as the KAc mimic, it seemed unwise to alter this structure as this might have significantly disrupted binding. Instead, the acrylamide and the methoxy group were identified as potentially modifiable areas that may not hinder binding and offered two distinct locations for determining the optimal alkyne placement (**Figure 5.5 A**).

The DMAM acrylamide substitution on the isoxazole fragments was well tolerated, with only a modest decrease in reactivity. Therefore, it was decided to incorporate an alkyne modified DMAM group on the acrylamide, **5.7** (**Figure 5.5 B**). The *N*-methylpropargylamine modification is well documented in the literature and could be installed as the final step through an S_N2 displacement of a halogen,[171–174] as described in **Chapter 4**. Modification of the methoxy moiety to an *O*-propargyl group was also an attractive strategy for alkyne placement. Results from non-covalent docking studies, conducted by Bernadette Lee, using GOLD, as discussed in **Chapter 3**, suggested that the

methoxy group of **3.50** might be solvent exposed, like for **I-BET151** (**1.62**) (**Figure 5.5 C** & **D**).[155] Unsure if modification of the methoxy group would alter binding to BRD4(1)^{L94C}, it was hypothesised that investigation into an *O*-propargyl analogue **5.8** would allow for a deeper understanding of how **3.50** engages with BRD4(1)^{L94C}.

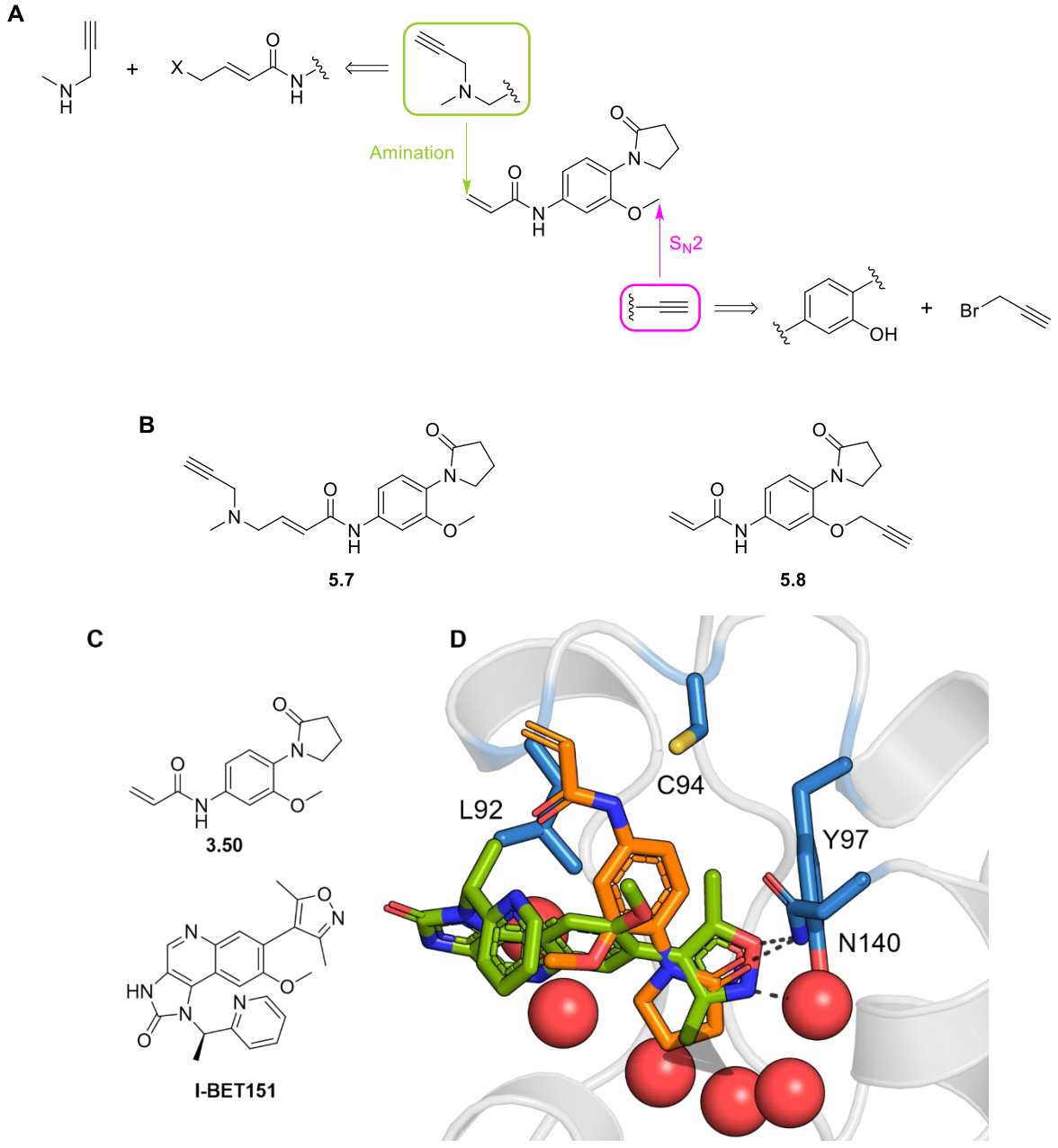


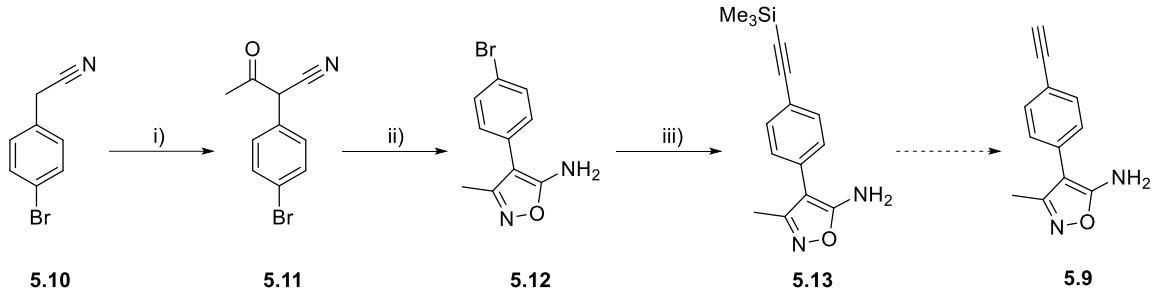
Figure 5.5. Design of clickable probes based on **3.50**, discovered from the GSK diverse fragment screen. **A**) Two alkyne variations were explored, utilising either an amination or an S_N2 approach in their synthesis. **B**) Target click probes. **C**) Chemical structures of **3.50** and **I-BET151** (**1.62**). **D**) Comparing the binding poses of **3.50** (carbon = orange) and **1.62** (carbon = green). Docking results place the carbonyl oxygen and the isoxazole oxygen in the same region, allowing for hydrogen bonding with N140. The methoxy group of **1.62** is solvent exposed, suggesting a similar pose may be adopted for the methoxy group of **3.50**.

5.3 Click probe synthesis

5.3.1 Isoxazole scaffold

*5.3.1.1 Synthesis of **5.1** and **5.4***

For the preparation of alkyne **5.9**, the 4-bromo intermediate **5.12** was initially chosen as the Sonogashira coupling partner. The synthesis started with a Claisen-type condensation between **5.10** and ethyl acetate to yield the ketonitrile **5.11** in a 79% yield.[162] The isoxazole ring was then formed using the condensation reaction with hydroxylamine, described in **Chapter 4**, which proceeded in an 83% yield to give **5.12** (**Scheme 5.1**).

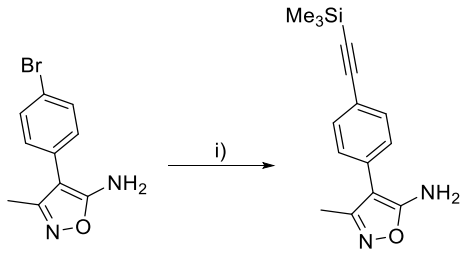


Scheme 5.1. Attempted synthesis of alkyne scaffold **5.9**. *Reagents and conditions*: i) EtOAc, NaH, THF, 10% *v/v* DMF, 0 °C to rt, 18 h, 79%, n=1; ii) Hydroxylamine hydrochloride, 10% aq. Na_2CO_3, EtOH, reflux, 1 h, 80–83%, n=3; iii) TMS acetylene, $Pd(PPh_3)_2Cl_2$, CuI, DMF, Et_3N, 80 °C, 2.5 h, 35% (conversion by LC-MS), n=1.

To install the alkyne, a Sonogashira coupling between **5.12** and TMS protected acetylene was attempted. Standard Sonogashira coupling conditions were employed, using $Pd(PPh_3)_2Cl_2$, copper iodide, and triethylamine as a base and solvent, with DMF acting as a cosolvent. Despite poor separation by TLC over a variety of solvent systems, it could be seen that product **5.13** had started to form after 1.5 hours at 80 °C. Due to poor separation by TLC, LC-MS analysis was employed for monitoring the reaction. After 1.5 hours at 80 °C LC-MS analysis indicated a 33% conversion to the protected alkyne **5.13**, however, by 2.5 hours only 35% conversion had occurred, indicative of a significant decrease in the reaction rate. The reaction was left for a further 18 hours in an attempt to push it to completion, as this would aid purification, owing to the poor separation between starting material **5.12** and product **5.13** by normal phase silica chromatography. After a total of

20 hours, LC-MS analysis indicated that a substantial amount of starting material remained, and that the desired product had degraded. None of the degradation products appeared to correspond to the deprotected alkyne **5.9**, implying that **5.13** was not sufficiently stable at 80 °C for an extended period of time. The reaction was attempted again, but this time the temperature was lowered to 60 °C to reduce degradation. Although **5.13** was found to be more stable at 60 °C, only low conversion to the product was observed after 20 hours, and so a more efficient method of synthesis was investigated.

Table 5.1. Reaction conditions attempted for the Sonogashira coupling between **5.12** and TMS acetylene. Conversion to the product was determined by LC-MS UV trace. *Reagents and conditions*: i) TMS acetylene, $Pd(PPh_3)_2Cl_2$, CuI, DMF, Et_3N.



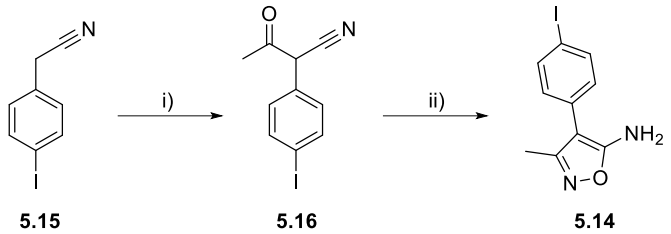
5.12 **5.13**

Temperature (°C)	Time (h)	Conversion to 5.13 (%)
80	1.5	33
	2.5	35
	20	0
60	3	27
	20	30

It was decided that the iodo-analogue **5.14** would be investigated as aryl iodides are better Sonogashira coupling partners than their corresponding bromides,[175] due to their weaker carbon-halogen bond, allowing for easier oxidative addition into the Pd^0 specie.

Ketonitrile **5.16** was synthesised from commercially available **5.15** using the previously described conditions. Again, the use of 10% *v/v* DMF as a cosolvent led to an increase in yield from 37% to 50% on a 2 g scale, and at a 4 g scale these conditions gave a yield of 86%. The isoxazole ring

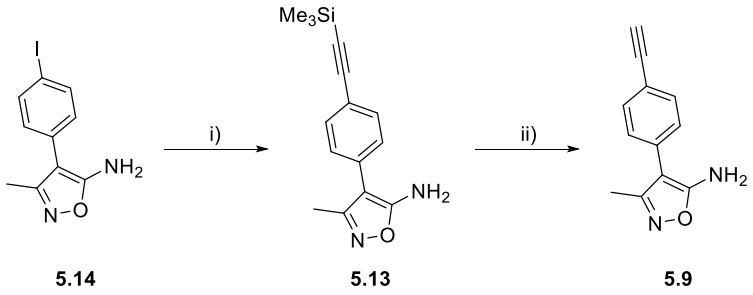
formation typically proceeded in high yields to give **5.14**, with a 2 g reaction giving a 99% yield (**Scheme 5.2**).



Scheme 5.2. Synthesis of the Sonogashira substrate **5.14**. *Reagents and conditions*: i) EtOAc, NaH, THF, 10% *v/v* DMF, 0 °C to rt, 18 h, 50–86%, n=2; ii) Hydroxylamine hydrochloride, 10% aq. Na_2CO_3, EtOH, reflux, 3 h, 51–99%, n=5.

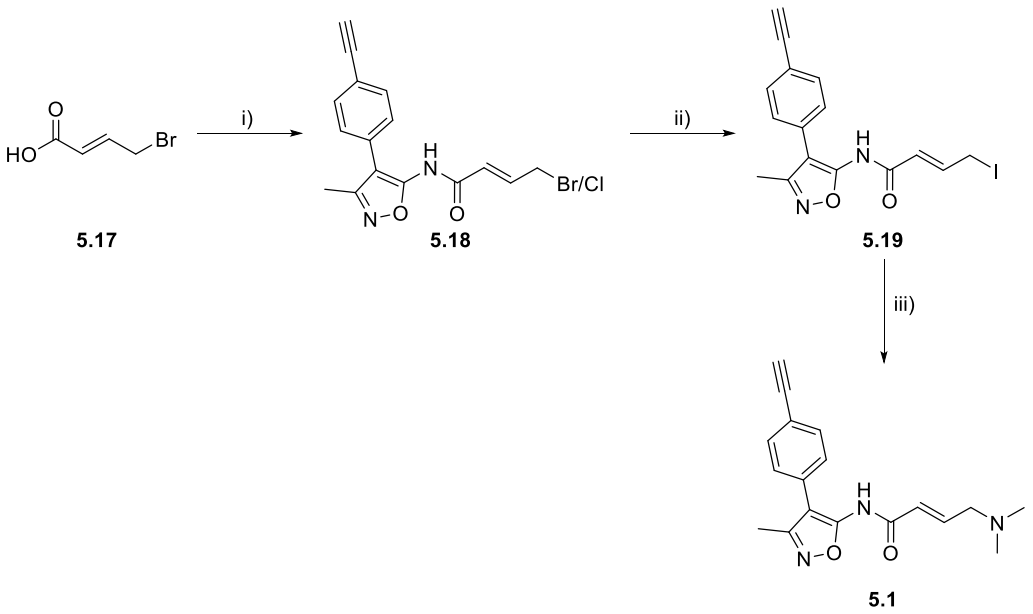
TMS acetylene was then coupled to **5.14** using Sonogashira conditions conducted at room temperature, affording **5.13** in yields of 67–74%. This reaction proceeded cleanly, with complete consumption of the starting material typically observed after 1.5 hours (**Scheme 5.3**).

With **5.13** in hand, it was decided that the deprotection of the TMS group would follow immediately, due to concerns that the deprotection conditions of potassium carbonate and methanol may react with the acrylamide if present, as seen for the ynone fragments in **Chapter 3**. Using potassium carbonate and methanol, the alkyne was revealed to give **5.9** in a 92% yield (**Scheme 5.3**).



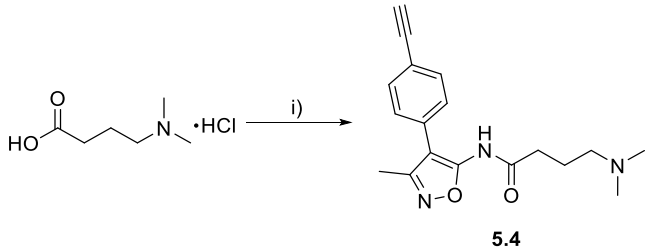
Scheme 5.3. Synthesis of alkyne scaffold **5.9**. *Reagents and conditions*: i) TMS acetylene, $Pd(PPh_3)_2Cl_2$, CuI, DMF, Et_3N, rt, 1.5 h, 67–74%, n=5; ii) K_2CO_3, MeOH, rt, 1 h, 87–92%, n=4.

Using the amination strategy described in Chapter 4, **5.9** was coupled to (*E*)-4-bromocrotonic acid **5.17** by generating the acyl chloride, to give **5.18** in a 36% yield as an ~8:1 mixture of the bromo- and chloro-substituted products. An S_N2 displacement gave the iodo-substituted **5.19**, which subsequently underwent amination with dimethylamine to synthesise **5.1** in a 28% yield over the two steps (**Scheme 5.4**).



Scheme 5.4. Synthesis of **5.1**. *Reagents and conditions*: i) a) Oxalyl chloride, DMF, CH_2Cl_2, 0 °C to rt, 15 h; b) **5.9**, pyridine, CH_2Cl_2, 0 °C, 5 h, 36%, n=1; ii) NaI, acetone, 50 °C, 1 h; iii) Dimethylamine (2.0 M in THF), K_2CO_3, DMF, 0 °C, 1.5 h, 28% over two steps, n=1.

The saturated negative control **5.4** was synthesised using the amide coupling method described in **Chapter 4**. The coupling between **5.9** and 4-(dimethylamino)butyric acid afforded **5.4** in an 81% yield (**Scheme 5.5**).



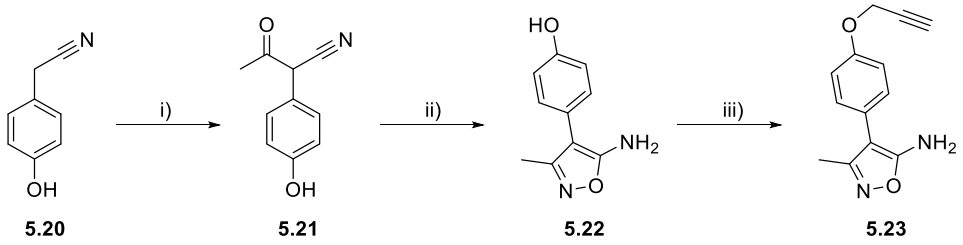
Scheme 5.5. Synthesis of the saturated negative control **5.4**. *Reagents and conditions*: i) a) Oxalyl chloride, DMF, THF, 0 °C to rt, 1 h, b) **5.9**, NMP, 0–5 °C, 16 h, 81%, n=1.

*5.3.1.2 Synthesis of **5.2**, **5.3**, **5.5**, and **5.6***

To synthesise **5.21**, commercially available **5.20** was acetylated using a modified version of the previously described conditions (**Scheme 5.6**). Here, due to the presence of the acidic phenol proton, twice the equivalents of sodium hydride were used to ensure full deprotonation of the α-carbon. It was predicted that the major product of the reaction would be the diacetylated compound, with the condensation reaction occurring at the phenolic oxygen too. TLC and LC-MS

analysis of the crude reaction mixture after an aqueous acidic work-up revealed that the major product was **5.21**, however, it is possible that the diacetylated ester could have hydrolysed under the work-up conditions. Repeated attempts to isolate **5.21** by silica gel chromatography were unsuccessful, with unknown impurities co-eluting with the desired product, and so the partially purified compound was progressed to the next step (**Scheme 5.6**).

The condensation reaction to form the isoxazole **5.22** generally proceeded in high yields, however, for some preparations, purification remained a challenge due to the similar polarities of the impurities. Installation of the alkyne handle was then conducted through an S_N2 reaction with propargyl bromide. To promote the propargylation at the phenol rather than the amine, **5.22** was pre-stirred with potassium carbonate to generate the phenolate. Addition of propargyl bromide then gave the desired **5.23** in a 92% yield (**Scheme 5.6**).

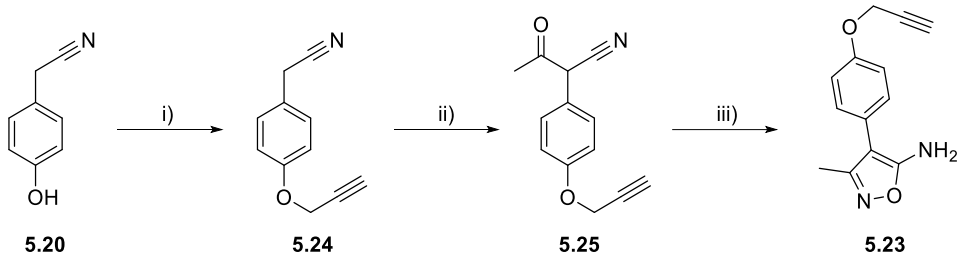


5.20 **5.21** **5.22** **5.23**

Scheme 5.6. Synthesis of alkyne scaffold **5.23**. *Reagents and conditions*: i) EtOAc, NaH, THF, 10% *v/v* DMF, 0 °C to rt, 16 h, 38%, n=3; ii) Hydroxylamine hydrochloride, 10% aq. Na_2CO_3, EtOH, reflux, 1 h, 63–99%, n=5; iii) Propargyl bromide, K_2CO_3, CH_2Cl_2, 0 °C to rt, 16 h, 85–92%, n=4.

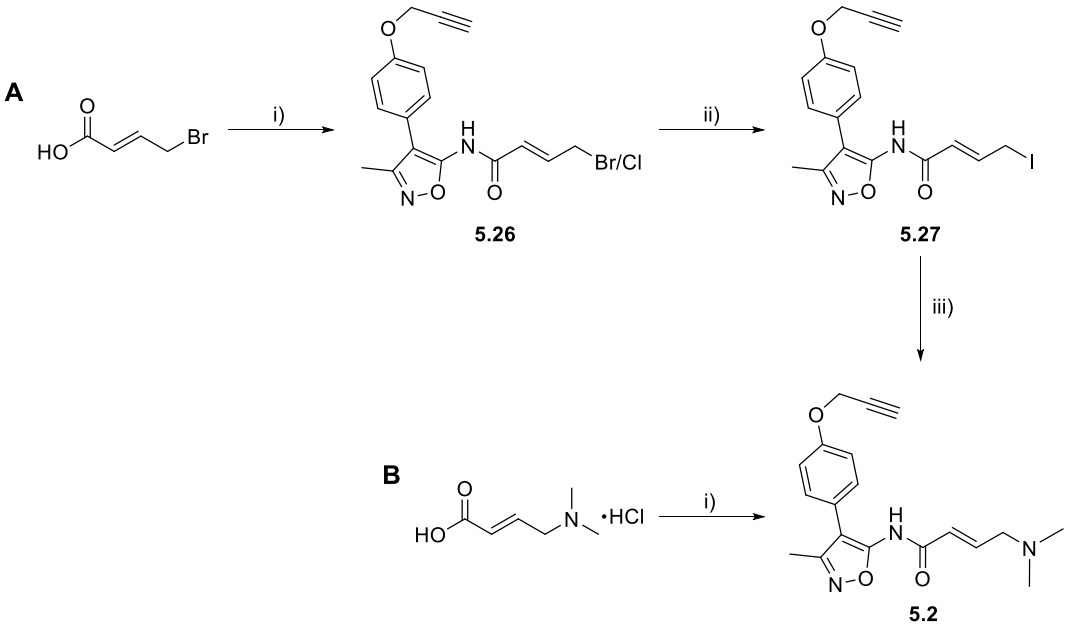
To assess whether a more efficient synthesis of **5.23** could be achieved, changing the order of the steps was explored. For this approach **5.20** was propargylated first, as it was postulated that issues arising with purification in the previous route were due to the acidic phenol. Propargylation of **5.20** with propargyl bromide afforded **5.24** in a 99% yield, with no silica gel chromatography required, this was a far more efficient first step compared to the previous route. Acetylation with ethyl acetate, this time using the standard 1.1 equivalents of sodium hydride, proceeded well to give **5.25** in a 91% yield. The ketonitrile was then treated with hydroxylamine to form the isoxazole **5.23** in an 83% yield. Overall, this route proved to be the most efficient way of synthesising **5.23** due to the

simpler purification required. This resulted in an overall yield of 75%, compared to approximately 35% for the initial route (**Scheme 5.7**).



Scheme 5.7. Alternative synthesis of alkyne scaffold **5.23**. *Reagents and conditions*: i) Propargyl bromide, K_2CO_3, CH_2Cl_2, 0 °C to rt, 3 h, 99%, n=1; ii) EtOAc, NaH, THF, 10% *v/v* DMF, 0 °C to rt, 4 h, 91%, n=1; iii) Hydroxylamine hydrochloride, 10% aq. Na_2CO_3, EtOH, reflux, 1 h, 83%, n=1.

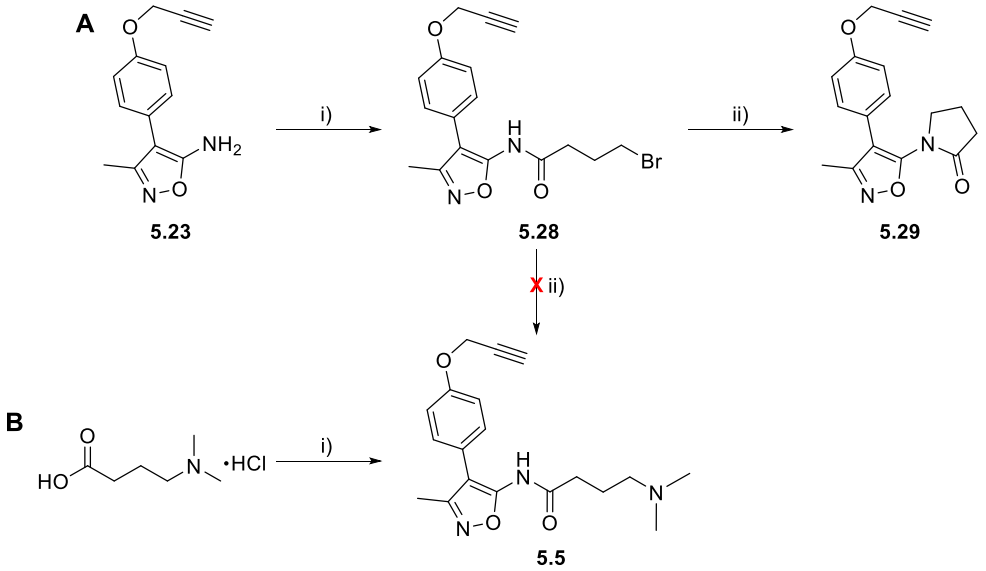
To synthesise **5.2,** two previously described routes were attempted, as discussed in **Chapter 4**: (i) amination to install the dimethyl amine group, and (ii) direct amide coupling. Whilst route **A** afforded **5.2** in an overall yield of 6% over three steps, route **B** was far more efficient and proceeded in a single step with a 37% yield (**Scheme 5.8**).



Scheme 5.8. The two different routes used to synthesise **5.2**. *Reagents and conditions*: **A**) i) a) Oxalyl chloride, DMF, CH_2Cl_2, 0 °C to rt, 3 h, b) **5.23**, pyridine, CH_2Cl_2, 0 °C, 2 h, 36%, n=1; ii) NaI, acetone, 50 °C, 1 h; iii) Dimethylamine (2.0 M in THF), K_2CO_3, DMF, 0 °C, 1 h, 18% over two steps, n=1; **B**) i) a) Oxalyl chloride, DMF, THF, 0 °C to rt, 1.5 h, b) **5.23**, NMP, 0–5 °C, 22 h, 37%, n=1.

The initial attempt to synthesise the saturated negative control compound **5.5** started with the coupling of **5.23** and 4-bromobutyryl chloride, to afford the desired amide **5.28** in an 80% yield (**Scheme 5.9 A**). During the attempted installation of the dimethylamine group, using TBAI as a

catalyst, no desired product was formed despite full consumption of the starting material **5.28**. Instead, NMR and LC-MS analysis revealed that the isolated compound was the product of an intramolecular γ-lactamisation to give **5.29** in an 80% yield (**Scheme 5.9 A**). NMR studies revealed the disappearance of the amide peak, and a shift in the triplet corresponding to the γ-CH_2 protons (3.42 ppm to 3.70 ppm) indicating the substitution of the bromide, however, the 6H singlet expected for the dimethylamine group was absent. HMBC analysis further confirmed the cyclic structure by revealing a strong correlation between the γ-CH_2 protons and the carbonyl carbon as well as a weak correlation to C-5 of the isoxazole (**Figure 5.6**). This major product was also observed when the reaction was run in the absence of potassium carbonate, suggesting that **5.29** was most likely formed upon addition of the dimethylamine which acted as a base rather than a nucleophile. This observation can be rationalised by the fact that the amide proton of **5.28** is predicted to have a p*K*a of 7.66 (MarvinSketch 21.14). To overcome this issue, the direct amide coupling approach was employed, and **5.5** was synthesised in a 33% yield (**Scheme 5.9 B**).



Scheme 5.9. The two different routes attempted for the synthesis of the saturated negative control compound **5.5**. *Reagents and conditions*: **A**) i) 4-Bromobutyryl chloride, pyridine, CH_2Cl_2, 0 °C, 1.5 h, 80%, n=1; ii) Dimethylamine (2.0 M in THF), TBAI, K_2CO_3, DMF, 0 °C to rt, 1 h, **5.5**: Not isolated, **5.29**: 80% n=1; **B**) i) a) Oxalyl chloride, DMF, THF, 0 °C to rt, 1 h, b) **5.23**, NMP, 0–5 °C, 16 h, 33%, n=1.

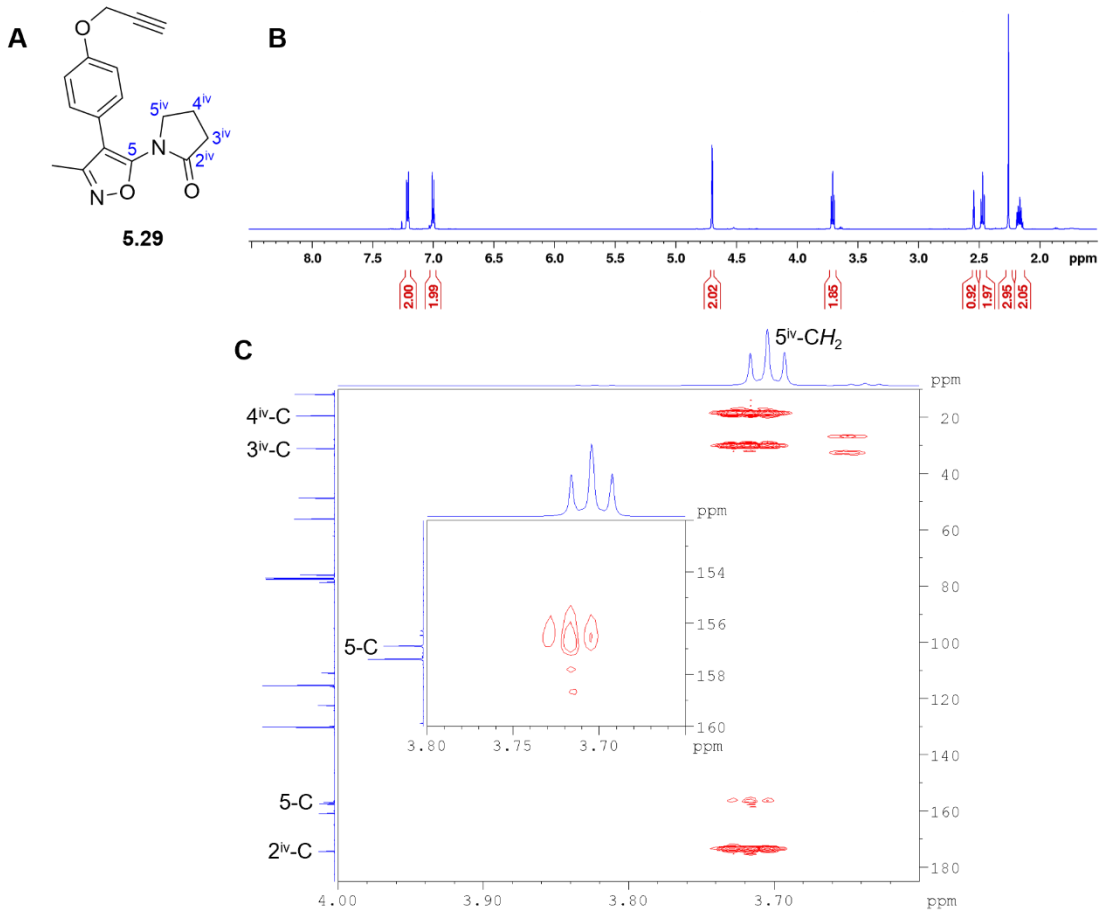
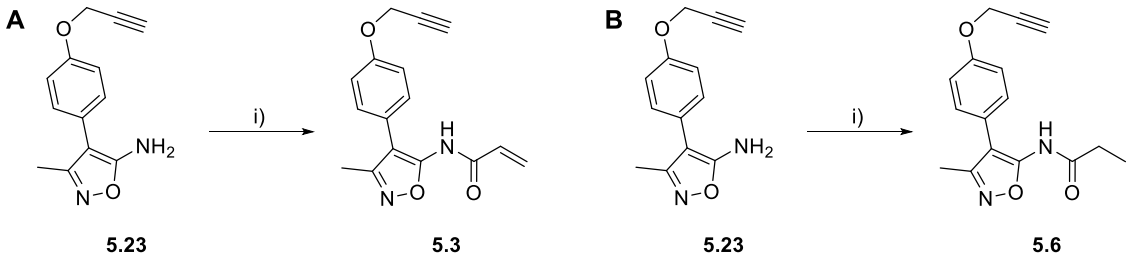


Figure 5.6. NMR analysis confirmed the formation of the γ-lactam **5.29**. **A**) Chemical structure of **5.29** with key atoms labelled. **B**) ^{1}H NMR spectrum of **5.29** revealed a large shift in the γ-CH_2 protons compared to **5.28,** indicative of a change in substituent. **C**) HMBC spectrum of **5.29** showing γ-CH_2 (5^{iv}) correlations to the carbonyl and the isoxazole 5-C and to the α- and β-carbons.

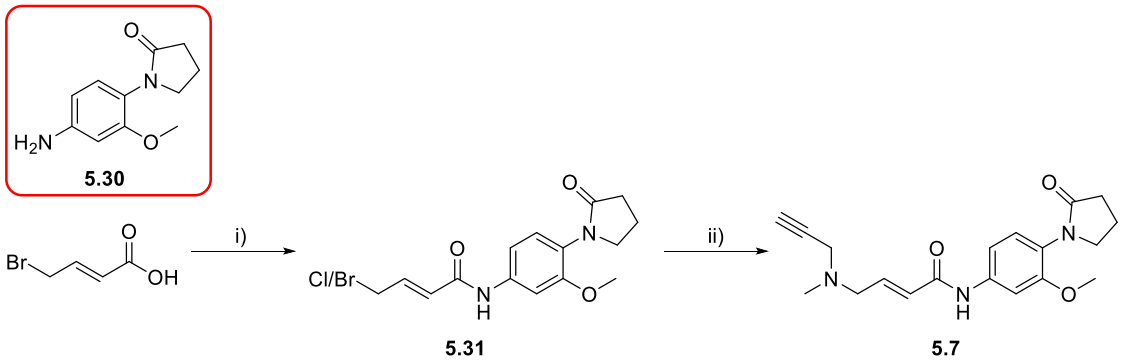
To synthesise the unsubstituted acrylamide **5.3**, acryloyl chloride was coupled to **5.23** to form **5.3** in an 18% yield (**Scheme 5.10 A**), using the previously described procedure in **Chapter 4**. Propionyl chloride was used for the synthesis of the saturated analogue **5.6**, which proceeded in a 56% yield (**Scheme 5.10 B**).



Scheme 5.10. Synthesis of **5.3** (**A**) and it's saturated negative control **5.6** (**B**). *Reagents and conditions*: **A**) i) Acryloyl chloride, pyridine, CH_2Cl_2, 0 °C to rt, 2.5 h, 18%, n=1; **B**) i) Propionyl chloride, pyridine, CH_2Cl_2, 0 °C, 1.5 h, 56%, n=1.

5.3.2 γ-Lactam scaffold

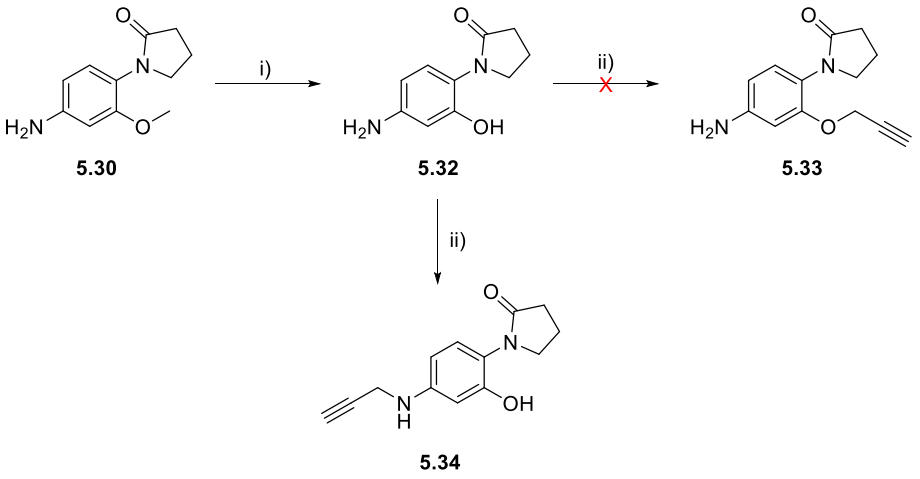
The basic γ-lactam scaffold was commercially available as the aniline **5.30** from Enamine. The first step of the substituted acrylamide analogue **5.7** synthesis was the coupling of **5.30** and (*E*)-4-bromocrotonic acid. This reaction proceeded cleanly and did not require silica gel chromatography, affording a 3:2 mixture of bromo- and chloro-substituted **5.31** in an 89% yield. Amination with *N*-methylpropargylamine afforded click probe **5.7**, in a 28% yield (**Scheme 5.11**).



Scheme 5.11. Synthesis of click probe **5.7**. *Reagents and conditions*: i) a) Oxalyl chloride, DMF, THF, 0 °C to rt, 45 min, b) **5.30**, NMP, 0 °C, 75 min, 89%, n=1; ii) *N*-Methylpropargylamine, KI, K_2CO_3, DMF, 0 °C to rt, 2 h, 28%, n=1.

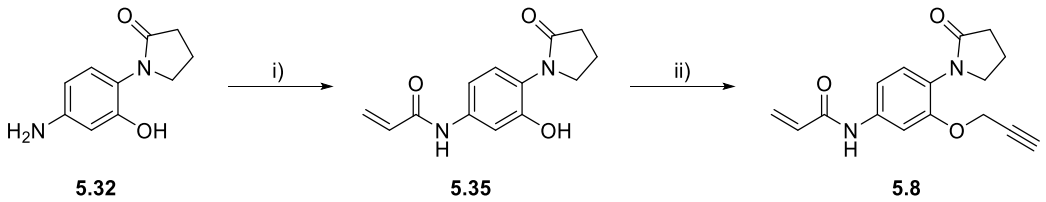
To synthesise the *O*-propargyl analogue **5.8**, boron tribromide was used for the removal of the methyl group from **5.30** to reveal the desired phenol **5.32** in an 11% yield (**Scheme 5.12**). This step was later optimised by Dr Marta Serafini (University of Oxford) to give yields of 78%.

Next, the installation of the propargyl group on the phenol of **5.32** was attempted (**Scheme 5.12**). Compound **5.32** was pre-stirred with 1.5 equivalents of potassium carbonate prior to the addition of propargyl bromide to first deprotonate the phenol. After 3.5 hours, consumption of the starting material **5.32** had occurred (by TLC) and the formation of a new, less polar, product was observed by TLC. However, NMR and LC-MS analysis of the isolated product suggested that propargylation had occurred on the amine forming **5.34** rather than **5.33**. A defining feature in the NMR was a broad singlet integrating for 1H, which was determined to be the NH of the secondary amine, as propargylation of the phenol would have resulted in the primary amine integrating for 2H (**Scheme 5.12**).



Scheme 5.12. Attempted synthesis of alkyne scaffold **5.33**. *Reagents and conditions*: i) BBr_3 (1.0 M in CH_2Cl_2), CH_2Cl_2, 0 °C to rt, 23 h, 11%, n=1; ii) Propargyl bromide, K_2CO_3, DMF, 0 °C to rt, 3.5 h, **5.33**: Not isolated. **5.34**: 75%, n=1.

It was hypothesised that rather than Boc protecting the amine to block propargylation at that site, installation of acrylamide first, and followed by propargylation of the phenol, might alleviate the need to protect and deprotect the amine. Due to time constraints, the remaining synthesis of **5.8** was conducted by Dr Marta Serafini. Using acryloyl chloride the acrylamide **5.35** was formed in a 40% yield. The phenol was then propargylated to afford the desired probe **5.8** in a 35% yield (**Scheme 5.13**).



Scheme 5.13. Synthesis of **5.8**, conducted by Dr Marta Serafini. *Reagents and conditions*: i) Acryloyl chloride, pyridine, CH_2Cl_2, 0 °C to rt, 7 h, 40%, n=1; ii) Propargyl bromide, K_2CO_3, DMF, rt, 16 h, 35%, n=1.

5.4 Click probe validation

Fluorescent tagging of the clickable probes after incubation with cells allows covalent interactions between the probe and proteins to be visualised by separation of cell lysate components on an SDS-PAGE gel, followed by irradiation at the desired excitation wavelength of the fluorophore. The rhodamine fluorophore, TAMRA, was chosen as the fluorophore due to its brightness and stability. A commercially available azide analogue with a PEG3 linker was used for the assays (**Figure 5.7**).

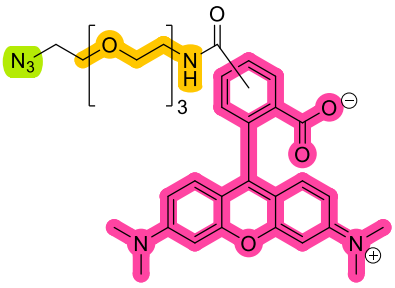


Figure 5.7. Chemical structure of the fluorescent dye TAMRA-PEG3-Azide, λ_{Abs}= 546 nm, λ_{Em}= 565 nm. The rhodamine dye scaffold is highlighted in pink, the PEG linker in orange, and the azide handle in green.

5.4.1 Click efficiency with TAMRA-PEG3-azide

To ensure that all the synthesised click probes could undergo a CuAAC reaction with TAMRA-PEG3-azide, an LC-MS assay was developed. The data collected would give an insight into the rate and efficiency of the click reaction, and would ultimately provide information on the optimum reaction time for fluorescent labelling in the cell lysates.

Click probes (50 µM) were diluted in LC-MS buffer (**Chapter 8**), followed by the addition of freshly prepared click mix (TAMRA-PEG3-azide [200 µM], TCEP·HCl [1 mM], $CuSO_4$ [1 mM], TBTA [100 µM]). Although the TAMRA-PEG3-azide is primarily the 5-isomer, low levels of the 6-isomer are also present, accounting for about 20%, and so for the LC-MS traces a small second peak can be observed, however, this did not interfere with the assay. The reaction was incubated in the dark at rt for 1 hour, after which a sample was submitted for LC-MS analysis.

All click probes underwent almost complete conversion to the triazole-TAMRA conjugate after 1 hour, resulting in the disappearance of the UV peak corresponding to the mass of the alkyne bearing click probe and the formation of a new UV peak corresponding to the mass of the click product. (**Figure 5.8**). Interestingly, both **5.1** and its control **5.4** did not go to completion during the hour. Whilst only low levels of the alkynes remained after 1 hour, it is noteworthy that both these compounds bear the alkyne directly on the phenyl ring, suggesting that this alkyne might be less accessible for CuAAC reaction than the corresponding *O*-propargyl alkynes, **5.2** and **5.5**. Full consumption of **5.3** can be observed after 1 hour, however the retention time for the triazole-linked product overlaps with that of the TAMRA-PEG3-azide. Confirmation of the triazole-linked product was determined by the desired mass being observed at the overlapping retention time. Additionally, it can be seen that the triazole product of the saturated analogue **5.6** also exhibits a retention time similar to that of TAMRA-PEG3-azide. Overall, these data were encouraging and suggested that 1 hour would be a reasonable duration for the click reaction to occur with good efficiency in the cell lysates.

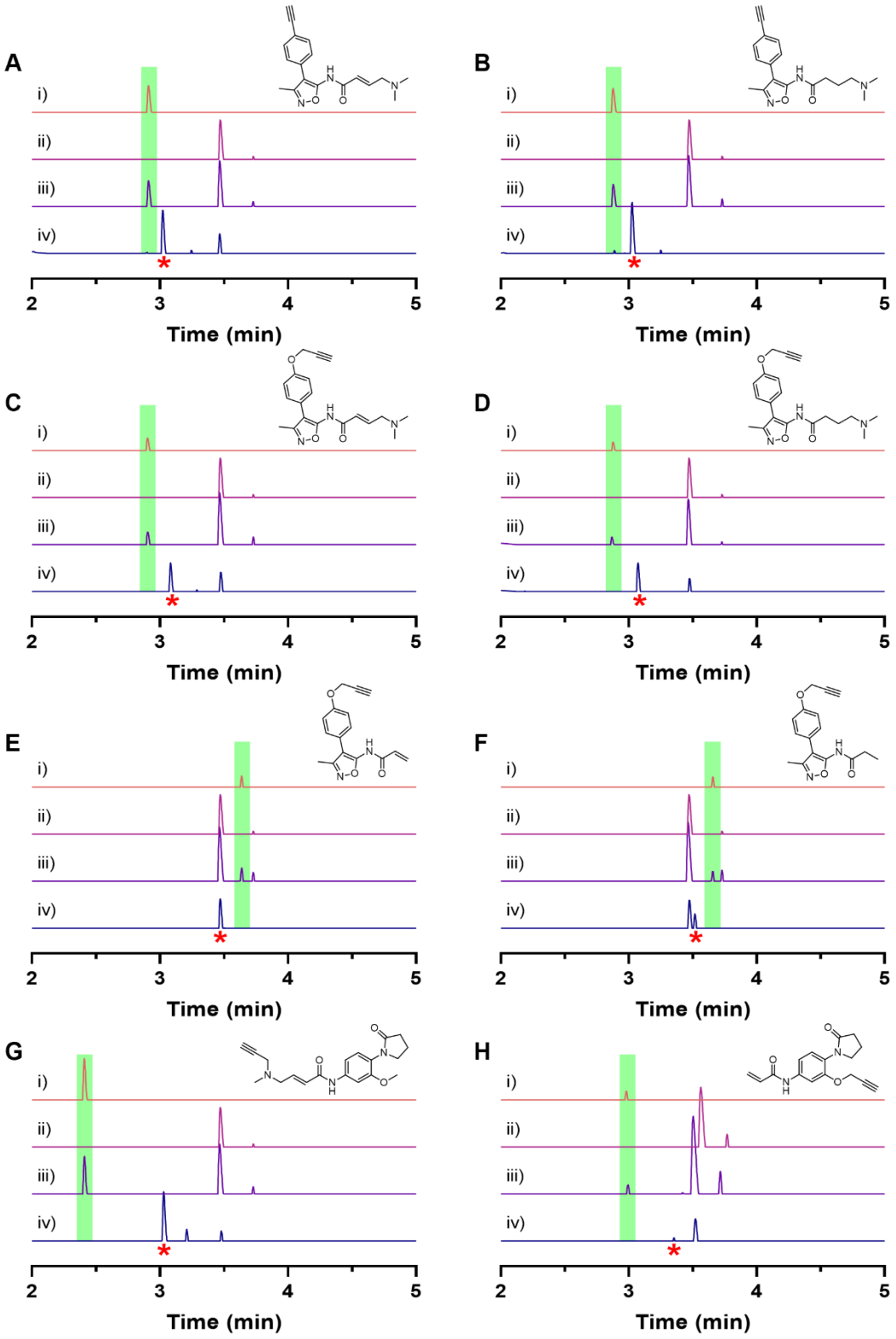


Figure 5.8. LC-MS UV Chromatograms for the CuAAC reaction between TAMRA-PEG3-azide and click probes **5.1** (**A**), **5.4** (**B**), **5.2** (**C**), **5.5** (**D**), **5.3** (**E**), **5.6** (**F**), **5.7** (**G**), and **5.8** (**H**). i) Click probe; ii) TAMRA-PEG3-azide; iii) Click probe and TAMRA-PEG3-azide, No catalyst; iv) CuAAC reaction, 1 h, rt. The green box highlights the UV peak corresponding to the click probe. The red asterisk shows the peak corresponding to the triazole-linked click probe-TAMRA-PEG3 conjugate.

5.4.2 Fluorescent labelling of purified BRD4(1)^{L94C}

With the knowledge that the triazole formation worked efficiently, the next step was to attempt the bioconjugation with purified BRD4(1)^{WT} and BRD4(1)^{L94C}. It was important that the probes were tested with purified protein before progression into cell work, as this would allow evaluation of whether the click reaction could still occur when the probe was covalently bound to the protein.

Click probes **5.1**, **5.2**, **5.3**, and **5.7** (50 µM) were individually incubated with either BRD4(1)^{WT} or BRD4(1)^{L94C} (10 µM) for 4 hours at 37 °C, before the addition of freshly prepared click mix (TAMRA-PEG3-azide (200 µM), TCEP·HCl [1 mM], $CuSO_4$ [1 mM], TBTA [100 µM]). The reaction mixture was incubated in the dark for a further 1 hour at rt, before being analysed using SDS-PAGE which allowed separation of the components.

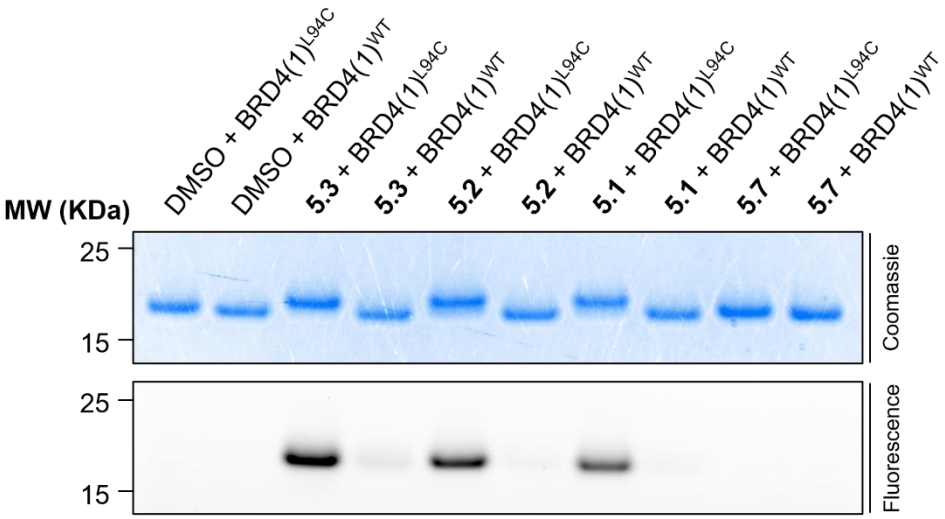


Figure 5.9. Fluorescence labelling of purified BRD4(1)^{L94C} and BRD4(1)^{WT} by **5.3**, **5.2**, **5.1**, and **5.7**.

Staining with Coomassie blue revealed a distinct shift in the protein band for BRD4(1)^{L94C} when incubated with **5.1**, **5.2**, and **5.3** (**Figure 5.9**). This observation is caused by the ~1 kDa increase in mass for the TAMRA labelled protein, when compared to the unlabelled protein, and was further confirmed upon irradiation, as only the heavier band fluoresced, while the lighter band did not. Compound **5.3** exhibited the highest level of BRD4(1)^{L94C} fluorescent labelling, however, some weak labelling of BRD4(1)^{WT} was observed, which is in agreement with the protein LC-MS data collected

for the parent compound **4.2**. Both **5.1** and **5.2** bound to, and fluorescently labelled, BRD4(1)^{L94C} at high levels. Some extremely weak fluorescence was observed for BRD4(1)^{WT} which is contrary to the results obtained for the parent compound **4.3** which exhibited no labelling of BRD4(1)^{WT}, by protein LC-MS. A possible explanation for this is that during the SDS-PAGE sample preparation, low levels of unspecific labelling had occurred. This is likely given the elevated temperatures used to denature the protein, and the excess concentration of unreacted click probe.

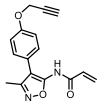
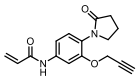
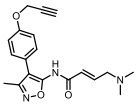
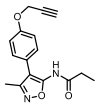
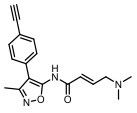
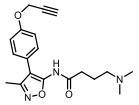
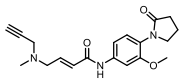
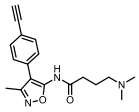
Compound **5.7** displayed no fluorescent labelling of BRD4(1)^{L94C}, suggesting that two issues may be at play. Firstly, it was possible that **5.7** was unable to covalently label BRD4(1)^{L94C}, and that substitution of the acrylamide had either decreased the reactivity of the electrophile or had sterically blocked the cysteine thiol and the acrylamide from reacting. Or secondly, the click reaction could not proceed when **5.7** was covalently bound to BRD4(1)^{L94C} which may have been caused by steric clashes. Unfortunately, at the time of this experiment the synthesis of **5.8** had not been completed and so was not tested in this experiment.

5.4.3 Evaluation of purified BRD4(1)^{L94C} labelling using protein mass spectrometry

To confirm and quantify the results of the fluorescent labelling of purified BRD4(1)^{L94C}, **5.1**, **5.2**, **5.3**, and **5.7** were evaluated for covalent labelling of BRD4(1)^{L94C} using the previously described protein LC-MS assay. Saturated control compounds **5.4**, **5.5**, and **5.6** were also assessed to ensure no covalent labelling could occur without the electrophilic acrylamide. In addition, with **5.8** now in hand its ability to covalently label BRD4(1)^{L94C} was evaluated, this was also tested against BRD4(1)^{WT} to confirm selective labelling of C94 (**Appendix D**).

Table 5.2. Results obtained from the protein LC-MS covalent binding assay for the click probes. The table shows the percentage of labelled protein after incubation at 37 °C for given time points. Results are shown as the mean of three repeats for hits (n=3), or two repeats for non-hits (n=2), and the errors given as the standard error of mean (SEM). The colour scale is shown below, where low levels of protein labelling are shown in blue, and high levels shown in red.



		Labelled Protein (%) L94C				Labelled Protein (%) L94C	
Compound	Structure	1 h	4 h	Compound	Structure	1 h	4 h
5.3		96.0 ± 0.2	100.0 ± 0.0	5.8		36.7 ± 1.2	82.4 ± 1.1
5.2		49.2 ± 2.0	91.8 ± 2.4	5.6		ND	ND
5.1		58.7 ± 1.9	96.3 ± 0.5	5.5		ND	ND
5.7		ND	2.0 ± 0.0	5.4		ND	ND

The mass spectrometry data verified the results from the fluorescent labelling gel, confirming that **5.1**, **5.2**, and **5.2** covalently label BRD4(1)^{L94C} (**Table 5.2**). **5.3** exhibited low levels of unspecific covalent labelling, in line with parent compound **4.2**, shown by a peak of 3% relative intensity corresponding to the mass of two labelling events. All three DMI-based covalent click probes displayed an increase in labelling compared to their respective parent compounds. It was hypothesised that this increase in labelling was driven by the additional modifications having

increased the affinity for BRD4(1)^{L94C} or increased the reactivity through more optimal positioning of the electrophile and C94, rather than a mass spectrometry artifact caused by the addition of the alkyne. The reasoning for this was that the alkyne modification was only a small change in mass and was not expected to greatly impact the ionisation of the protein, and also as the mass spectrometry method used was ESI+ it was deemed that the alkyne would be unlikely to significantly increase ionisation. This hypothesis was also in agreement with the fact that there was not an increase in the percentage of unspecific binding observed for **5.3**, which if the alkyne was increasing ionisation a proportional increase in this peak would have been expected. All three of the saturated negative controls, **5.4**, **5.5**, and **5.6** showed no labelling of BRD4(1)^{L94C}, confirming that they are suitable negative controls for the in-cell experiments.

The γ-lactam-based click probe, **5.7**, labelled only 2% of BRD4(1)^{L94C} after 4 hours, which was in agreement with the data obtained from the fluorescent labelling assay, confirming that the substitution had severely hindered the compound's ability to covalently label C94. For compound **5.8**, with the *O*-propargyl modification, the opposite was observed, and after 4 hours 82% of the protein was covalently labelled (**Figure 5.10 A**). The tolerance seen for the change from the methoxy group to the *O*-propargyl group suggests that this region of the molecule is solvent exposed— an important observation for future development of this compound. After 4 hours, no labelling of BRD4(1)^{WT} by **5.8** was observed, indicating selective labelling of C94 (**Figure 5.11 B**).

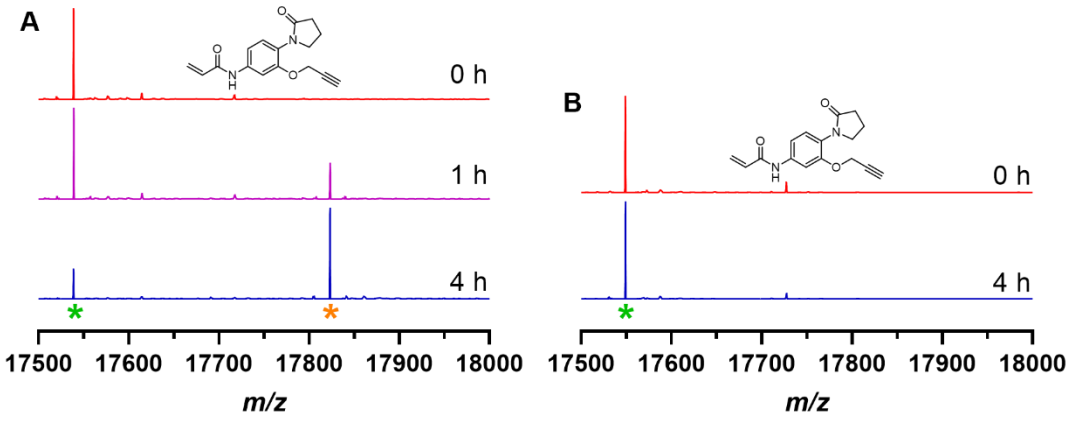


Figure 5.10. Time-dependent protein LC-MS spectra for **5.8** incubated with BRD4(1)^{L94C} (**A**), and BRD4(1)^{WT} (**B**). The green asterisk shows peaks corresponding to unlabelled protein , and the orange asterisks show peaks corresponding to labelled protein.

5.5 In-cell target engagement

Having verified that click probes **5.1**, **5.2**, **5.3**, and **5.8** were able to covalently label purified BRD4(1)^{L94C}, and were able to undergo a CuAAC reaction with TAMRA-PEG3-azide, the next step was to evaluate their proteome-wide selectivity in cells expressing BRD4(1)^{L94C}. For this proof-of-concept study it was decided that cells transiently expressing BRD4(1)^{L94C} would be used, due to the simplicity of validating engagement with the target. However, it is envisioned that this would be transferable to phenotypic studies using a CRISPR-Cas9 mutated cell line, expressing full length BRD4^{L94C}. Unfortunately, due to time constraints, the CRISPR-Cas9 work could not be carried out.

5.5.1 Plasmid design

In collaboration with Dr Philip Hublitz (MRC Weatherall Institute of Molecular Medicine, Genome Engineering Facility) and Prof. Tom Milne (MRC Weatherall Institute of Molecular Medicine, Molecular Haematology Unit), two constructs for transient expression in HEK293T cells were designed. It was decided that one construct would code for BRD4(1)^{L94C}, and the other code for BRD4(1)^{WT} as a control (**Appendix D**). Both constructs would have an N-terminal StrepII tag, which could be used for Western blotting and Immunoprecipitation (IP) allowing for protein detection (**Figure 5.11**). The C-terminus of the BRDs would be fused to eGFP by a flexible linker, followed by a self-cleaving T2A-mCherry to allow for protein degradation studies that will be discussed further in **Chapter 6**. The phosphoglycerate kinase (PGK) promoter was selected as a suitable promoter of the construct due to its weak to moderate expression levels in HEK293T cells. Using a promoter that was not too strong was vital to ensure that levels of the expressed BRDs did not reach cytotoxic levels. The plasmids were assembled and purified by Dr Philip Hublitz (MRC Weatherall Institute of Molecular Medicine, Genome Engineering Facility). Whilst HEK293T cells endogenously express full-length BRD4 (~152 kDa), this would not interfere with observing labelling of transiently expressed BRD4(1)^{WT} and BRD4(1)^{L94C} (~45.6 kDa) due to the large mass difference of the proteins.

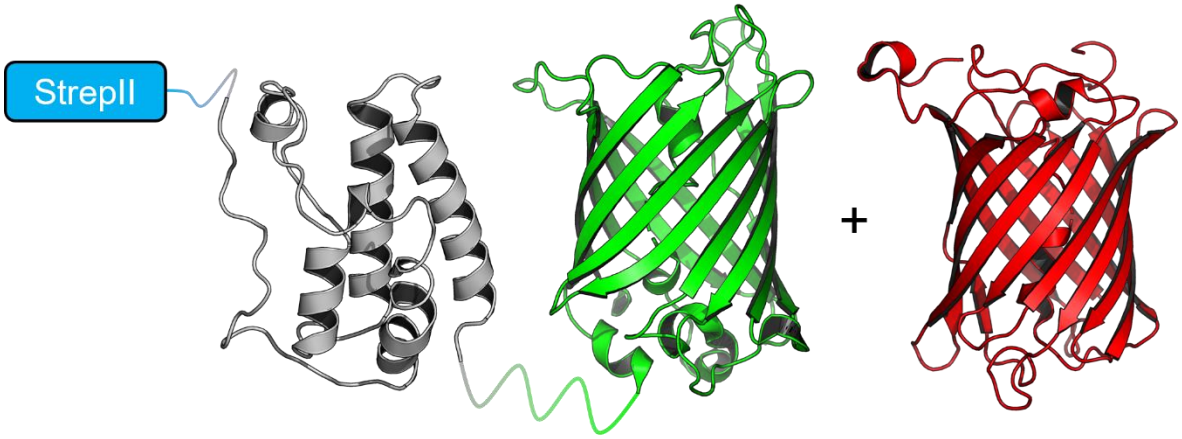


Figure 5.11. Graphical representation of the proteins coded for by the designed plasmids. The StrepII tag is shown in blue, BRD4(1)$^{WT/L94C}$ in grey, eGFP in green (PDB ID: 4KW4), and mCherry in red (PDB ID: 6YLM).[176]

5.5.2 Validation of protein expression

With the sequenced plasmids in hand, the next step was to determine transfection conditions and validate that the desired proteins were being expressed as expected. Using the transfection reagent Lipofectamine 2000, HEK293T cells were cultured and then transfected with either the plasmid coding for BRD4(1)^{L94C} or BRD4(1)^{WT}. The media was changed after 24 hours and the cells were allowed to grow for a further 24 hours before being harvested, lysed and Western blotted for StrepII-tag (**Figure 5.12 A**). Western blotting confirmed that the cell cultures transfected with the BRD4(1)^{L94C} or BRD4(1)^{WT} constructs were expressing the StrepII-tagged bromodomains (~45.6 kDa), while the untransfected control cultures displayed no expression of the StrepII-tagged bromodomains. These results were further verified by fluorescence microscopy, where transfected cultures were shown to express the bromodomain-eGFP fusion protein and mCherry (**Figure 5.12 B & C**).

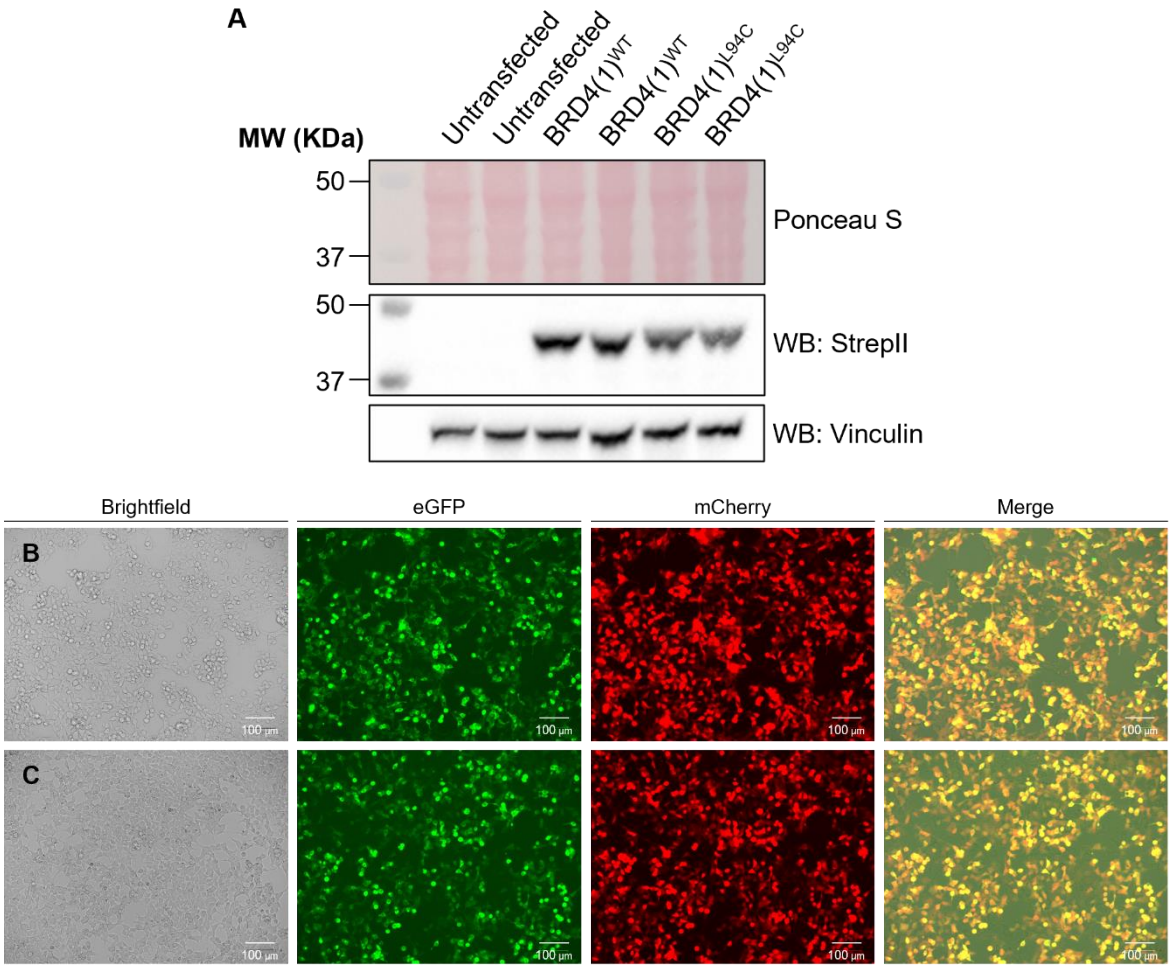


Figure 5.12. Validation of transfection and protein expression. **A**) Untransfected HEK293T cells (n=2), transfected HEK293T cells expressing BRD4(1)^{WT} (n=2), and transfected HEK293T cells expressing BRD4(1)^{L94C} (n=2) were analysed using Western blotting (WB) to evaluated protein expression of the StrepII-tagged fusion proteins. StrepII was blotted for and Vinculin was used as a loading control. **B**) HEK293T cells transiently expressing the BRD4(1)^{WT} construct. Microscope: Brightfield, eGFP (λ_{Abs}= 480 nm, λ_{Em}= 517 nm), mCherry (λ_{Abs}= 556 nm, λ_{Em}= 615 nm), Merge. **C**) HEK293T cells transiently expressing the BRD4(1)^{L94C} construct. Microscope: Brightfield, eGFP (λ_{Abs}= 480 nm, λ_{Em}= 517 nm), mCherry (λ_{Abs}= 556 nm, λ_{Em}= 615 nm), Merge.

5.5.3 In-cell protein labelling

The ultimate aim of this project was to investigate whether the combination of a reactive fragment and a specifically designed protein with a cysteine mutant would enable the rapid development of selective chemical probes suitable for use in phenotypic assays, to provide information on a protein's function. It is therefore vital to evaluate the selectivity of the compounds in live cells. Transfected HEK293T cells expressing BRD4(1)^{L94C} or BRD4(1)^{WT} were incubated with the click probes (10 µM), at a final DMSO concentration of 0.5%, for 1 hour, after which time cells were

harvested and lysed. The lysates were diluted to a total protein concentration of 2 mg/mL and then incubated with freshly prepared click mix (TAMRA-PEG3-azide [200 µM], TCEP·HCl [2 mM], $CuSO_4$ [2 mM], TBTA [200 µM]) for 1 hour. The components of the lysates were then separated using SDS-PAGE, and analysed by Coomassie blue staining, and in-gel fluorescent visualisation (**Figure 5.13**).

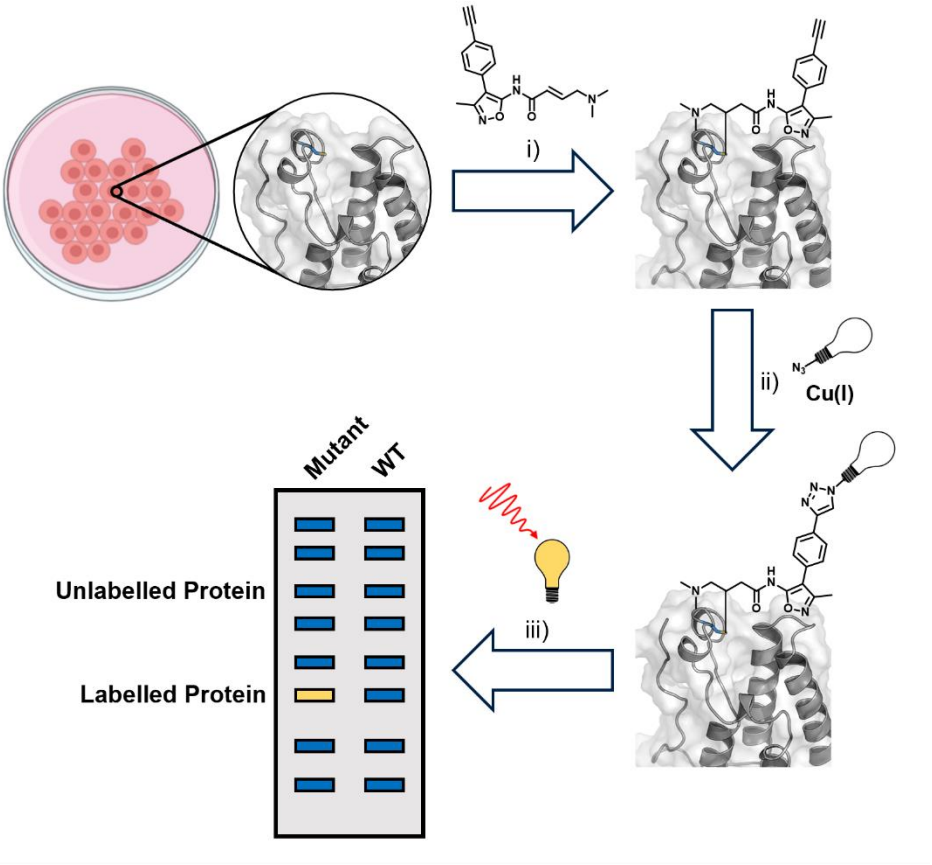


Figure 5.13. Graphical representation of the in-cell protein labelling assay; i) Transfected cells are incubated with the clickable probe. The probe covalently labels interacting proteins; ii) Cells are lysed, and a fluorescent dye functionalised with an azide undergoes a CuAAC reaction with the alkyne of the click probe— tethering the protein-probe conjugates to the fluorescent dye; iii) Fluorescently tagged proteins are then visualised using in-gel fluorescence imaging.

5.5.3.1 In-cell protein labelling by isoxazole-based click probes

All click probes were found to be cell permeable and suitable for in-cell protein labelling (**Figure 5.14**). Fluorescent labelling of proteins covalently bound to by **5.1**, **5.2**, and **5.3**, revealed high levels of off-target binding to endogenous proteins. Of the three probes, **5.3** appeared to exhibit the highest level of BRD4(1)^{L94C} labelling, demonstrated by a strong fluorescent band at the expected molecular weight, observed only in cells expressing BRD4(1)^{L94C}. Both DMAM substituted probes,

5.1 and **5.2**, exhibited lower levels of labelling for BRD4(1)^{L94C}, shown by weaker fluorescent bands corresponding to BRD4(1)^{L94C}. Probe **5.1** demonstrated the lowest level of BRD4(1)^{L94C} labelling. This was in line with previous results in purified protein and further suggests that the placement of the alkyne is not conducive to an optimal CuAAC reaction with TAMRA-PEG3-azide. It is interesting to note, that whilst results obtained in purified protein suggested that the DMAM substituted acrylamides, **5.1** and **5.2**, may exhibited greater proteome-wide selectivity compared to the unsubstituted analogue **5.3**, it was found that this was not the case in live cells. The three saturated control compounds **5.4**, **5.5**, and **5.6** exhibited no protein labelling in live cells, showing that covalent bond formation was vital (**Appendix D**).

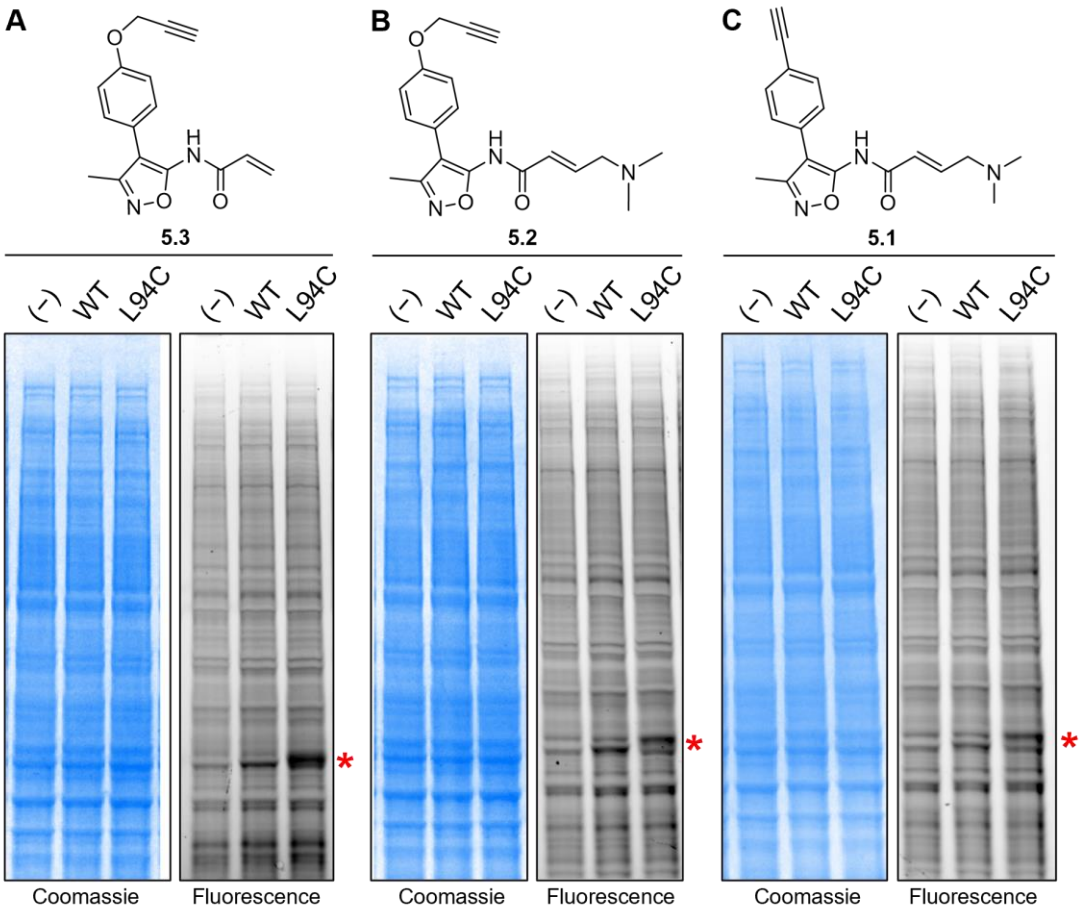


Figure 5.14. In-cell protein labelling of BRD4(1)^{L94C} in HEK293T cells by **5.3** (**A**), **5.2** (**B**), and **5.1** (**C**). The red asterisks indicate BRD4(1)^{L94C}. Coomassie staining was used to show the total protein. In-gel fluorescence imaging was used to identify proteins covalently interacting with the click probes.

To ensure that **5.3** was engaging with BRD4(1)^{L94C} in the expected manner, a competition assay using the known BRD4(1)^{L94C} ligand **OXFBD04** (**1.58**) was setup (**Figure 5.15**). Cells expressing BRD4(1)^{L94C} were preincubated with either **1.58** or DMSO for 30 minutes prior to the addition of **5.3** (10 µM) and incubated for a further 1 hour. Cells were harvested, lysed, and incubated with freshly prepared click mix, as described above. From the in-gel fluorescence it can be seen that preincubation with increasing concentrations of **1.58** decreases the levels of BRD4(1)^{L94C} labelling by **5.3**. This is consistent with **5.3** covalently engaging with C94 of BRD4(1)^{L94C}, driven by non-covalent interactions with the KAc binding pocket.

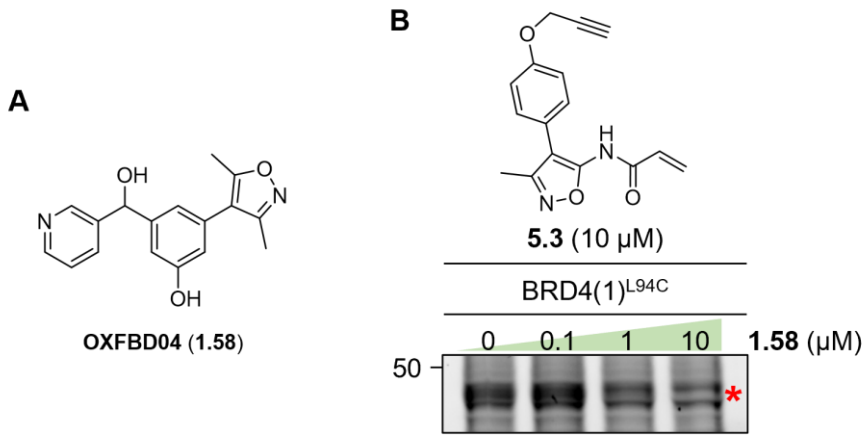


Figure 5.15. A competition assay with **OXFBD04** (**1.58**) confirmed in-cell protein labelling of BRD4(1)^{L94C}. **A**) Chemical structure of **1.58**. **B**) In-cell protein labelling of BRD4(1)^{L94C} by **5.3** with preincubation with increasing concentrations of **1.58**. The red asterisk indicates BRD4(1)^{L94C} which lies between two off-target labelled proteins. In-gel fluorescence imaging was used to identify proteins covalently interacting with **5.3**.

5.5.3.2 In-cell protein labelling by γ-lactam-based click probes

Both **5.7** and **5.8** were assessed for in-cell selectivity using the previously described in-cell protein labelling assay (**Figure 5.16**). Although **5.7** only displayed negligible labelling of purified BRD4(1)^{L94C} by protein mass spectrometry, it was decided that it would be assessed for any wider protein labelling. As suspected, **5.7** exhibited very low levels of labelling across all proteins within the cell, further confirming that the lack of labelling is due to a decrease in acrylamide reactivity when the substituent is installed. On the other hand, **5.8** displayed exquisite selectivity for BRD4(1)^{L94C}, with only very low levels of background fluorescence observed for the untransfected and BRD4(1)^{WT} transfected cells.

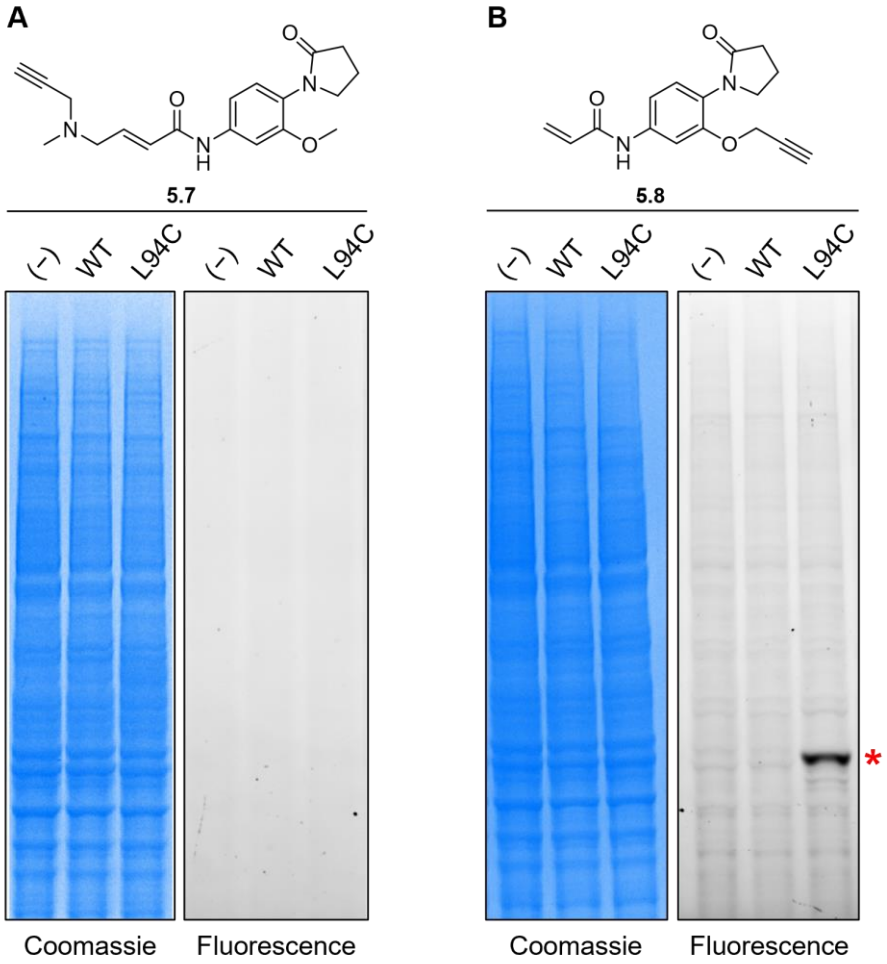


Figure 5.16. In-cell protein labelling of BRD4(1)^{L94C} in HEK293T cells by **5.7** (**A**), and **5.8** (**B**). The red arrow indicates BRD4(1)^{L94C}. Coomassie staining was used to show the total protein. In-gel fluorescence imaging was used to identify proteins covalently interacting with the click probes.

Using the previously described competition assay with **1.58**, fluorescent labelling of BRD4(1)^{L94C} by **5.8** was decreased when cells were preincubated with increasing concentrations of the competing ligand, suggesting that **5.8** binds to the same region of the protein as **1.58** (**Figure 5.17 A**).

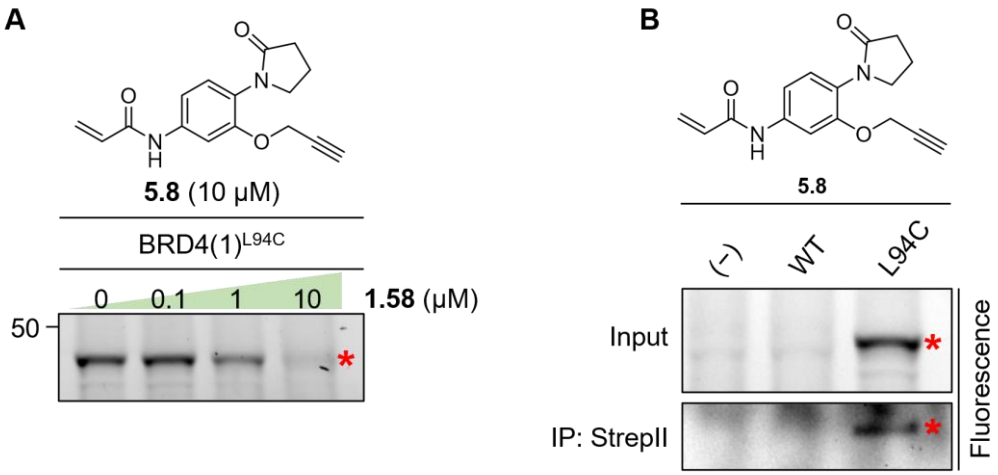


Figure 5.17. In-cell protein labelling of BRD4(1)^{L94C} by **5.8** was validated using a competition assay and IP. **A**) In-cell protein labelling of BRD4(1)^{L94C} by **5.8** with preincubation of increasing concentrations of **1.58**. The red asterisk indicates BRD4(1)^{L94C}. **B**) IP of StrepII-tagged proteins confirms labelling of BRD4(1)^{L94C}. In-gel fluorescence imaging was used to identify proteins covalently interacting with the **5.8**.

Finally, immunoprecipitation (IP) was employed to fully verify in-cell target engagement BRD4(1)^{L94C} (**Figure 5.17 B**). For the IP experiment, untransfected, BRD4(1)^{WT} transfected, and BRD4(1)^{L94C} transfected HEK293T cells were incubated with **5.8** (10 µM). A sample of each lysate was retained for the input, and the remainder of the lysate was incubated with anti-StrepII antibodies and protein A Dynabeads, which were used to pull out the StrepII tagged bromodomains. The beads were washed to elute unbound proteins, and then the click chemistry carried out in the presence of the beads, before the samples were prepared for SDS-PAGE separation and analysis. For the IP samples, a fluorescent band was only observed for cells expressing BRD4(1)^{L94C}, further verifying that the major protein engaged by **5.8** is unquestionably BRD4(1)^{L94C}.

5.6 Conclusions

Five covalent clickable analogues of fragments **4.2**, **4.3**, and **3.50** were designed and synthesised. In-cell selectivity experiments revealed that the isoxazole-based click probes, **5.1**, **5.2**, and **5.3**, exhibited high reactivity, resulting in non-specific labelling of endogenous proteins. Compound **5.8**, the clickable analogue of **3.50**, demonstrated remarkable selectivity for BRD4(1)^{L94C} with very low levels of background labelling observed. Despite probes **5.1**, **5.2**, **5.3**, and **5.8** displaying similar levels of BRD4(1)^{L94C} labelling by mass spectrometry, and being of similar size and molecular mass, there is a stark contrast between their observed in-cell selectivity. This is possibly caused by the nature of the acrylamide substituents, where the electron-withdrawing isoxazole likely increases the electrophilicity and reactivity of **5.1**, **5.2**, and **5.3**, while the electron-donating nature of the phenyl ring in **5.8** has the opposite effect.

Engagement with BRD4(1)^{L94C} by **5.8** was fully validated through competition assays using a known BRD4(1)^{L94C} ligand and protein pulldown by IP. This was a fascinating result considering that no optimisation of the initial hit fragment **3.50** had been carried out, apart from the installation of the alkyne for selectivity studies. Additionally, it was found that modification of the methoxy group was well tolerated, indicating that this region of the molecule is solvent exposed— an important finding for any further development of **3.50**. Whilst further inhibition data for **3.50** should be obtained, **3.50** appears to exhibit a high enough selectivity for BRD4(1)^{L94C} in cells for future phenotypic studies of BRD4(1) inhibition using a stable cell line expressing the full length mutant BRD4^{L94C}.

Furthermore, these data strongly suggest that site directed mutagenesis of a target protein to incorporate a cysteine residue could be combined with reactive fragments for the rapid development of selective, direct-to-biology, probes for the study of biological function.

Chapter 6

Investigating the applicability of ‘mutate & conjugate’ to the development of protein degraders

6.1 Introduction and aims

Most chemical probes exert their mechanism through the inhibition of the target, however, over the past decade there has been increasing research into chemical probes that induce degradation of the target protein.[177–180] As these chemical probes aim to remove the protein from the cell, they are complimentary to RNAi methods and offer potential for identifying off-target effects of RNAi.[26]

Whilst proteolysis-targeting chimeras (PROTACs) are usually the choice of degrader, their development typically require long synthetic routes and careful linker design to promote protein-protein interactions (PPI) between the E3 ligase and the protein of interest (POI).[181–185] Hydrophobic tagging of a protein for degradation, through the use of HyTags, offers an alternative strategy. HyTags consist of a ligand that binds the POI, joined to a hydrophobic moiety by a chemical linker.[186–189] Typically, the hydrophobic moiety is an adamantyl group. HyTags take advantage of the cellular machinery that patrols for misfolded proteins, such as chaperone proteins like 70 kDa heat shock proteins (HSP70). Proteins usually bury their hydrophobic residues within their structure and expose their hydrophilic residues on the surface. When a HyTag binds to a protein the hydrophobic moiety of the HyTag is exposed which chaperone proteins recognise as a misfolded protein and signal it for degradation (**Figure 7.1**). Another plausible mechanism for how HyTags exert their function is through destabilisation of the protein, whereby the hydrophobic group disrupts the local protein structure, leading to the recruitment of chaperones and then degradation.

HyTags potentially offer a faster method for the discovery of selective protein degraders. Firstly, this is because it is thought that optimal HyTag linker length is less vital for degradation compared to PROTACs, and secondly, the hydrophobic molecules that promote degradation are generally simple and require shorter synthetic routes than the ligands required for efficient E3 ligase engagement.

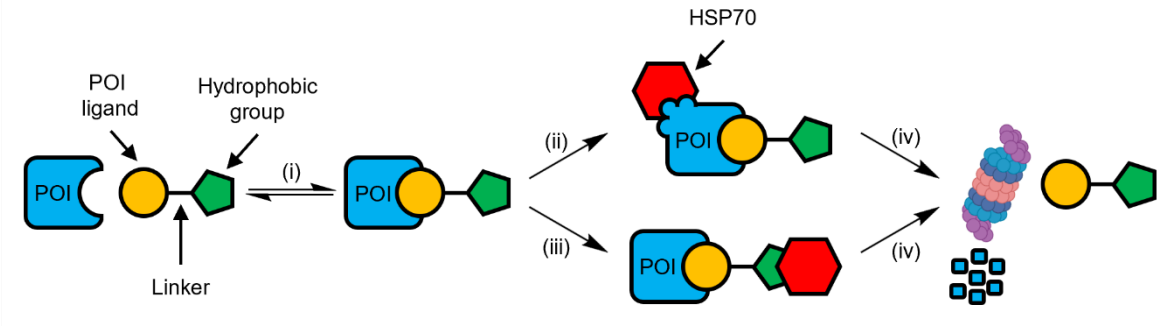


Figure 6.1. Schematic illustration of how HyTags function: (i) Binding of the HyTag to the POI; (ii) The hydrophobic group induces local disruption of protein structure, causing chaperone proteins to bind to the misfolded POI; (iii) Chaperone proteins such as HSP70 recognise the hydrophobic group as a misfolded protein; (iv) The POI is degraded by the proteasome and the HyTag is liberated.

It has therefore been investigated whether it is possible to promote selective protein degradation through the 'mutate and conjugate' methodology using a HyTag approach. The work described in this chapter aimed to:

- Develop covalent HyTag probes capable of selectively degrading BRD4(1)^{L94C} over BRD4(1)^{WT}.
- Validate covalent fragments as functionalisable moieties that are suitable for the rapid development of selective protein degraders.

Although this covalent method would not benefit from the catalytic nature exhibited by non-covalent protein degraders,[190] it was hypothesised that the covalent tethering of the HyTag to the desired mutant protein would ensure high levels of recognition by the chaperone protein, HSP70, leading to efficient protein degradation.

At the time of planning and carrying out this work, the in-cell selectivity work discussed in **Chapter 5** had not been completed, and so the majority of compounds in this chapter are based on the fragments **4.2** and **4.3**, described in **Chapter 4**.

6.2 Design of covalent HyTag probes

The work described in **Chapter 5** demonstrated that modification at the 4-position of the phenyl group of **4.2** and **4.3** is well tolerated and did not hinder covalent labelling of C94 of BRD4(1)^{L94C}. Therefore, it was decided to modify this position again as the site of attachment between the fragment and the linker (**Figure 6.2**).

A Sonogashira approach to installing the linker and hydrophobic tag was chosen owing to the robust synthesis developed for the alkyne-modification click probe **5.1** (**Figure 6.2**). Using the synthetic route from **Chapter 5**, the iodide substrate **5.14** could be reliably synthesised in high yields, and the installation of an alkyne could be achieved *via* the propargylation of a terminal alcohol on a PEG chain. An S_N2 approach using the phenol, as discussed in **Chapter 5**, was considered, however, it was postulated that issues may arise when using halogen substituted-PEG chains, such as base promoted elimination reactions forming the alkene.

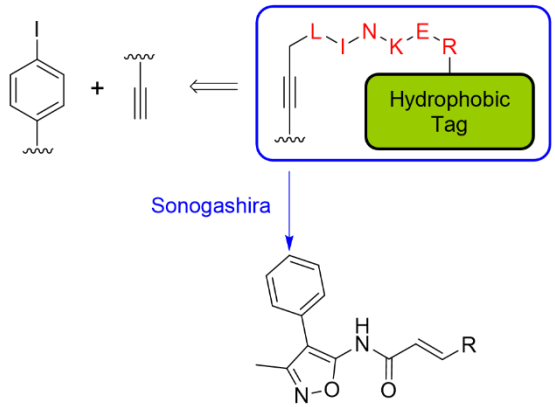


Figure 6.2. Design of protein degraders based on fragments **4.2** and **4.3** functionalised with a hydrophobic tag. A Sonogashira based approach was employed for linker installation due to the synthetic ease reported for the synthesis of **5.1**.

Due to the nature of the hydrophobic tags, it was likely that compound solubility could be an issue. To enhance the chances of retaining good aqueous solubility it was decided that a PEG type linker would be investigated. Whilst it is known that optimal linker length is important for protein degradation by PROTACs, it is far less understood how linker length impacts protein degradation by hydrophobic tagging. Encouraged by work from Tae *et al*. and Gustafson *et al*. it was decided that

a short 2-(2-aminoethoxy)-ethanol linker that was coupled *via* an amide bond to 1-adamantaneacetic acid would be employed (**Figure 6.3**). [187,188]

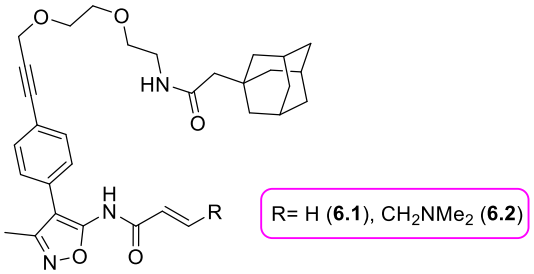
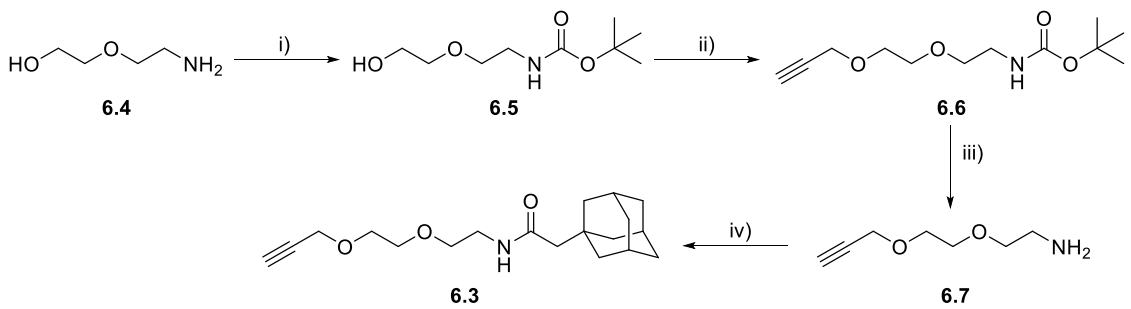


Figure 6.3. Structure of the proposed covalent HyTags, **6.1** and **6.2**, targeting BRD4(1)^{L94C}.

6.3 Synthesis of covalent HyTag probes

6.3.1 Synthesis of linker and hydrophobic moiety

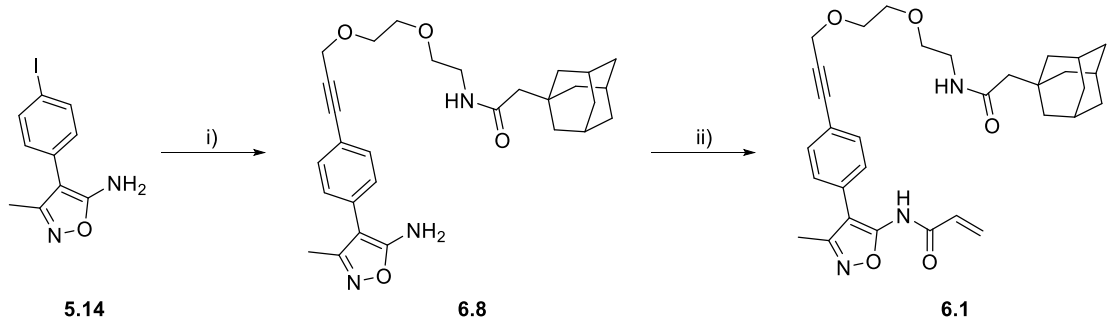
To synthesise **6.3**, an amine was required at one end of the linker for amide coupling to the hydrophobic 1-adamantaneacetic acid, while the other end required a terminal alkyne for the Sonogashira coupling. Following the procedure of Budhadev *et al*.,[191] the synthesis began with a Boc protection of commercially available 2-(2-aminoethoxy)-ethanol **6.4** which proceeded in a 99% yield to form **6.5** (**Scheme 6.1**). The alkyne was installed by alkylating the alcohol with propargyl bromide to give **6.6** in a 96% yield. To install the adamantane group, **6.6** was subjected to a TFA-mediated Boc deprotection to reveal the amine **6.7**, which was then coupled with 1-adamantaneacetic acid to synthesise **6.3** in a 70% yield over two steps.[186]



Scheme 6.1. Synthesis of **6.3**. *Reagents and conditions*: i) Boc_2O, $CHCl_3$, 0 °C to rt, 22 h, 99%, n=1; ii) Propargyl bromide, NaH, THF, 0 °C to rt, 15 h, 96%, n=1; iii) TFA, CH_2Cl_2, 0 °C to rt, 1.5 h; iv) 1-Adamantaneacetic acid, EDCI, HOBt, DIPEA, CH_2Cl_2, 0 °C to rt, 45 h, 70%, n=1.

6.3.2 Synthesis of unsubstituted acrylamide

The initial route to synthesise **6.1** employed the Sonogashira coupling, followed by the installation of the acrylamide, as it was thought that the reactive acrylamide might cause problems if installed before the final step (**Scheme 6.2**). Using the protocol described in **Chapter 5**, **5.14** was afforded in high yields. Amine **6.8** was synthesised *via* the palladium catalysed coupling of the iodide derivative **5.14** and linker **6.3**, in a 76% yield. Coupling of acryloyl chloride to **6.8** afforded **6.1**, as observed by crude NMR and LC-MS. Unfortunately, due to minimal differences in polarity between **6.8** and **6.1**, as shown by identical TLC R_f values and LC-MS retention times, **6.1** was not isolated in sufficient purity for biological testing.



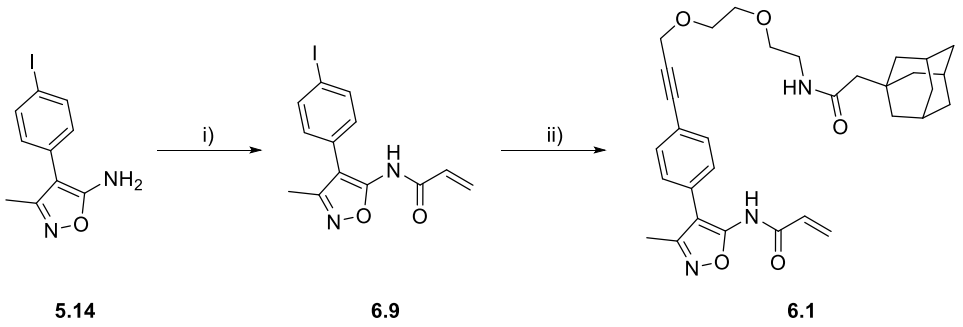
5.14 **6.8** **6.1**

Scheme 6.2. Initial synthesis of covalent HyTag **6.1**. *Reagents and conditions*: i) **6.3**, $Pd(PPh_3)_2Cl_2$, CuI, Et_3N, DMF, rt, 1 h, 76%, n=1; ii) Acryloyl chloride, pyridine, CH_2Cl_2, 24 h, not isolated, n=1.

It was therefore decided that an alternative approach would to be to install the acrylamide before carrying out the Sonogashira coupling as the final step (**Scheme 6.3**). This rearrangement of synthetic steps would mean there was a large difference in polarity between **6.9** and **6.1**, which would hopefully simplify purification.

Amine **5.14** was reacted with acryloyl chloride to afford acrylamide **6.9** in a 28% yield. As **6.9** possessed both an aryl iodide and an alkene it was hypothesised that formation of unwanted Heck reaction side products could occur. Whilst there was literature precedence for acrylamide moieties participating in Heck reactions with iodo-substituted aryls,[192,193] these were typically carried out at elevated temperatures and so it was decided that the Sonogashira would be attempted at rt to supress the Heck reaction. To improve purification, the isoxazole **6.9** was used in a 10% excess to

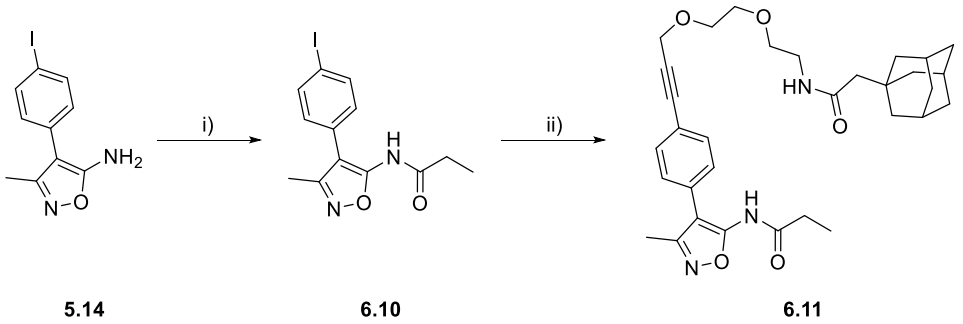
ensure complete consumption of the linker **6.3**, which possessed a similar R_f value to the product **6.1**. Using $Pd(PPh_3)_2Cl_2$ the reaction proceeded rapidly, with complete consumption of linker **6.3** after 40 minutes, affording **6.1** in a 67% yield. Unfortunately, due to the formation of the inseparable by-product, triphenylphospine oxide (TPPO), the purity obtained was only 94% which was not to the required standards for biological testing. To overcome this issue, the catalyst was changed to $Pd(dppf)Cl_2{\cdot}CH_2Cl_2$, which does not contain PPh_3, the precursor to TPPO. Switching the catalyst to $Pd(dppf)Cl_2{\cdot}CH_2Cl_2$ led to a slower reaction time of 20 hours, which required heating at 40 °C, as well as a decreased yield of 19%, however, the TPPO impurity was not formed and **6.1** was obtained in a >95% purity by HPLC for biological testing.



5.14 **6.9** **6.1**

Scheme 6.3. Synthesis of covalent HyTag **6.1**. *Reagents and conditions*: i) Acryloyl chloride, pyridine, CH_2Cl_2, 0 °C to rt, 6.5 h, 28%, n=1; ii) Attempt 1: **6.3**, $Pd(PPh_3)_2Cl_2$, CuI, Et_3N, DMF, rt, 1 h, 67%, n=1, Attempt 2: **6.3**, $Pd(dppf)Cl_2{\cdot}CH_2Cl_2$, CuI, Et_3N, DMF, rt to 40 °C, 20 h, 19%, n=1.

To synthesise the saturated negative control compound **6.11**, the same route for the optimised synthesis of **6.1** was employed (**Scheme 6.4**). The Sonogashira substrate **6.10** was obtained through the reaction of **5.14** and propionyl chloride in a 96% yield. The Sonogashira coupling was carried out between **6.10** and **6.3** to afford **6.11** in an 82% yield.

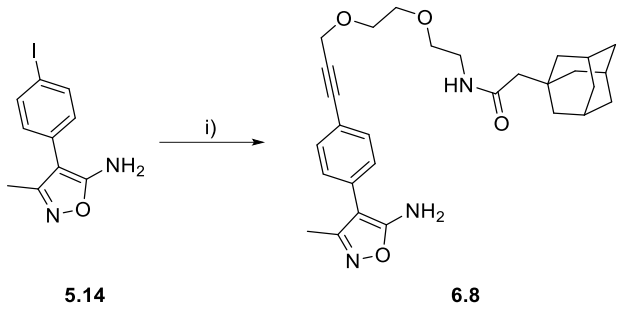


5.14 **6.10** **6.11**

Scheme 6.4. Synthesis of saturated negative control compound **6.11**. *Reagents and conditions*: i) Propionyl chloride, pyridine, CH_2Cl_2, 0 °C to rt, 24 h, 96%, n=1; ii) **6.3**, $Pd(dppf)Cl_2{\cdot}CH_2Cl_2$, CuI, Et_3N, DMF, 40 °C, 20 h, 82%, n=1.

6.3.3 Synthesis of substituted acrylamide

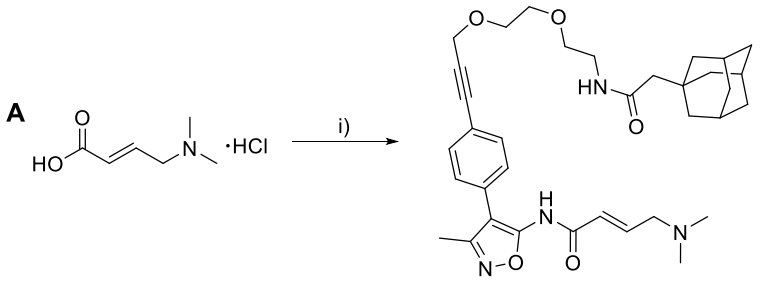
To synthesise **6.2**, coupling of the linker **6.3** to the scaffold **5.14** was performed prior to the installation of the electrophile. For the synthesis of **6.8** the catalyst was changed to $Pd(dppf)Cl_2 \cdot CH_2Cl_2$ from the previously used $Pd(PPh_3)Cl_2$, which increased the yield from 76% to 90% (**Scheme 6.5**).



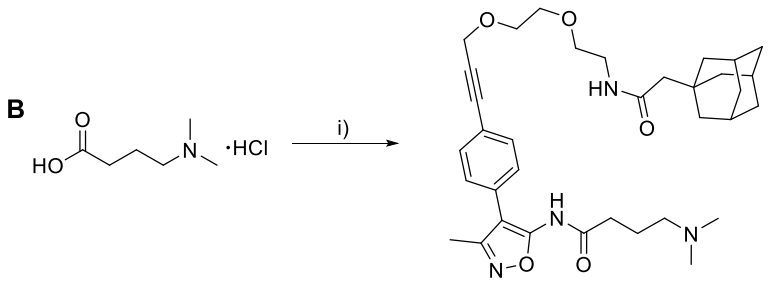
5.14 **6.8**

Scheme 6.5. Optimised synthesis of **6.8** to avoid TPPO contamination. *Reagents and conditions*: i) **6.3**, $Pd(dppf)Cl_2 \cdot CH_2Cl_2$, CuI, Et_3N, DMF, 40 °C, 3 h, 90%, n=1.

Using the previously described conditions, both the covalent HyTag **6.2** and negative control compound **6.12** were synthesised by coupling **6.8** and the corresponding carboxylic acids (**Scheme 6.6**).



6.2



6.12

Scheme 6.6. Synthesis of covalent HyTag **6.2** (**A**) and it's saturated negative control **6.12** (**B**). *Reagents and conditions*: **A**) i) a) Oxalyl chloride, DMF, THF, 0 °C to rt, 1.5 h, b) **6.8**, NMP, 0–5 °C, 20 h, 10%, n=1; **B**) i) a) Oxalyl chloride, DMF, THF, 0 °C to rt, 1.5 h, b) **6.8**, NMP, 0–5 °C, 21 h, 64%, n=1.

6.4 Evaluation of purified BRD4(1)^{L94C} labelling using protein mass spectrometry

To assess whether covalent HyTags **6.1** and **6.2** were able to label BRD4(1)^{L94C} the previously described protein LC-MS assay was employed. Covalent HyTags and negative control HyTags (50 µM) were incubated with BRD4(1)^{L94C} (10 µM) at 37 °C and time points taken at 0, 1, and 4 hours (**Appendix E**).

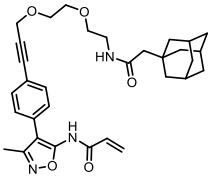
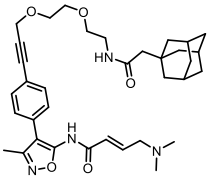
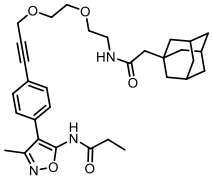
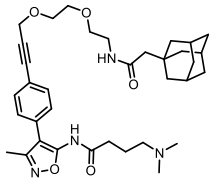
Both **6.1** and **6.2** demonstrated labelling of BRD4(1)^{L94C}, while the negative controls, as expected, did not (**Table 6.1**). Interestingly, **6.1** exhibited decreased levels of labelling compared to its parent fragment **4.2**, 32% compared to 98% respectively, and after 1 hour the rate of labelling had slowed substantially, shown by only a small increase in labelling between 1 and 4 hours (**Figure 6.4 A**). Presumably the low levels of labelling were due to poor solubility of **6.1** in aqueous buffer, meaning that the effective concentration of the covalent HyTag was far lower than expected, and that by 1 hour the majority of solubilised **6.1** had already labelled the protein. This hypothesis appears to be in agreement with the data obtained for **6.2**, which possesses a solubilising DMAM group, where **6.2** exhibited high levels of covalent labelling with a mean value of 97% after 4 hours (**Figure 6.4 B**). The calculated solubility forecast index (SFI) values (**Equation 6.1**), using cLogD values (Marvin Sketch 21.14), indicated that the aqueous solubility of **6.1** was lower than **6.2**, 5.92 compared to 4.39. According to Hill *et al*. compounds with an SFI value of <5 have a reasonable chance of possessing good solubility.[194] Taken together, these data indicate that the addition of the linker and hydrophobic tag to the fragment have decreased the aqueous solubility of the compounds, however, the DMAM group of **6.2** and **6.12** helps to counteract this increase in lipophilicity, and brings SFI values into line with compound that are likely to have good solubility.

Equation 6.1.

$$SFI = (cLogD_{pH7.4} + Number\ of\ aromatic\ rings)$$

Table 6.1. Results obtained from the protein LC-MS covalent binding assay for the HyTags. The table shows the percentage of labelled protein after incubation at 37 °C for given time points. Results are shown as the mean of three repeats for hits (n=3), or two repeats for non-hits (n=2), and the errors given as the standard error of mean (SEM). The colour scale is shown below, where low levels of protein labelling are shown in blue, and high levels shown in red. ND = Not detectable.

0% Labelling — 100% Labelling

		Labelled Protein (%)			
		L94C			
Compound	Structure	1 h	4 h	cLogD$_{pH7.4}$	SFI
6.1		26.5 ± 0.9	31.9 ± 0.4	3.92	5.92
6.2		67.7 ± 0.1	96.7 ± 0.6	2.39	4.39
6.11		ND	ND	3.83	5.83
6.12		ND	ND	1.77	3.77

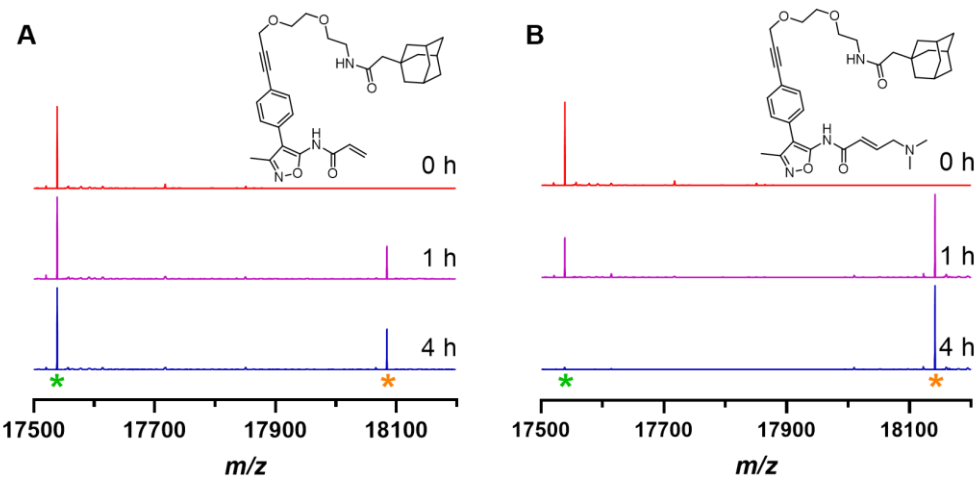


Figure 6.4. Time-dependent protein LC-MS spectra for (**A**) **6.1** and (**B**) **6.2** incubated with BRD4(1)^{L94C}. The Y-axis shows relative intensity (%). The green asterisk shows peaks corresponding to unlabelled protein, and the orange asterisks show peaks corresponding to labelled protein.

6.5 In-cell engagement and protein degradation by flow cytometry

As described in **Chapter 5**, two plasmids were designed for transient expression of BRD4(1)^{WT} and BRD4(1)^{L94C} in HEK293T cells. The expressed bromodomains would have an N-terminal StrepII tag, while the C-terminal would be fused to eGFP by a flexible linker followed by a 'self-cleaving' T2A-mCherry. The T2A peptide sequence 'EGRGSLLTCGDVEENPGP' induces ribosomal skipping preventing peptide bond formation between residues G and P,[195,196] and therefore results in the synthesis of two separate polypeptides— in this case the BRD4(1)-eGFP fusion protein and mCherry. Using this T2A technique, BRD4(1)-eGFP fusion proteins and the red fluorescent mCherry are expressed from the same strand of mRNA and hence theoretically yield with a 1:1 ratio of the BRD4(1)-eGFP fusion proteins and mCherry. This therefore allows the normalisation of eGFP: mCherry fluorescence,[197,198] where a decrease in eGFP fluorescence relative to mCherry fluorescence would indicate the degradation of the BRD4(1)-eGFP fusion proteins (**Figure 6.5**). Using this assay, therefore, enables a quantitative readout of BRD4(1)-eGFP degradation.

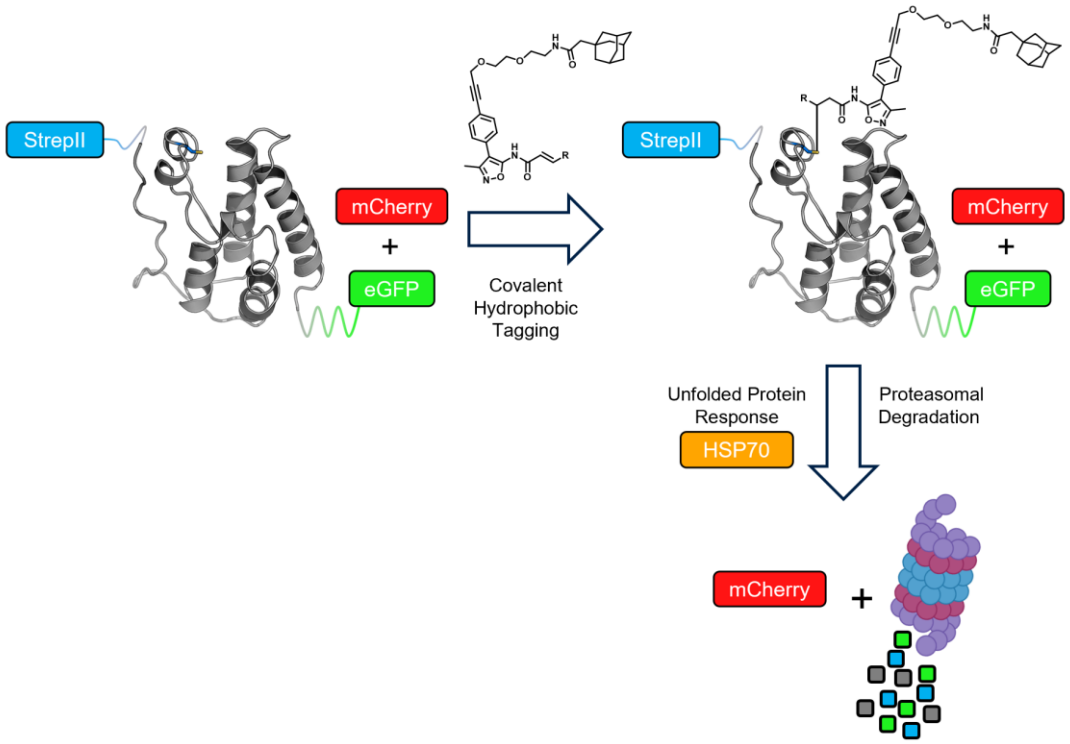


Figure 6.5. Schematic illustration of how protein degradation is monitored by the ratio of eGFP-bromodomain: mCherry fluorescence using flow cytometry.

HEK293T cells were transfected with either the plasmid coding for BRD4(1)^{WT} or BRD4(1)^{L94C}. The cells were then incubated with the HyTags (10 µM) for 16 hours before being harvested and analysed using flow cytometry. To measure the degradation of the bromodomain-eGFP fusion proteins, the ratio of the median eGFP signal to the median mCherry signal was calculated for successfully transfected cells and then normalised to the ratio calculated for the DMSO treated cells (**Figure 6.6**).

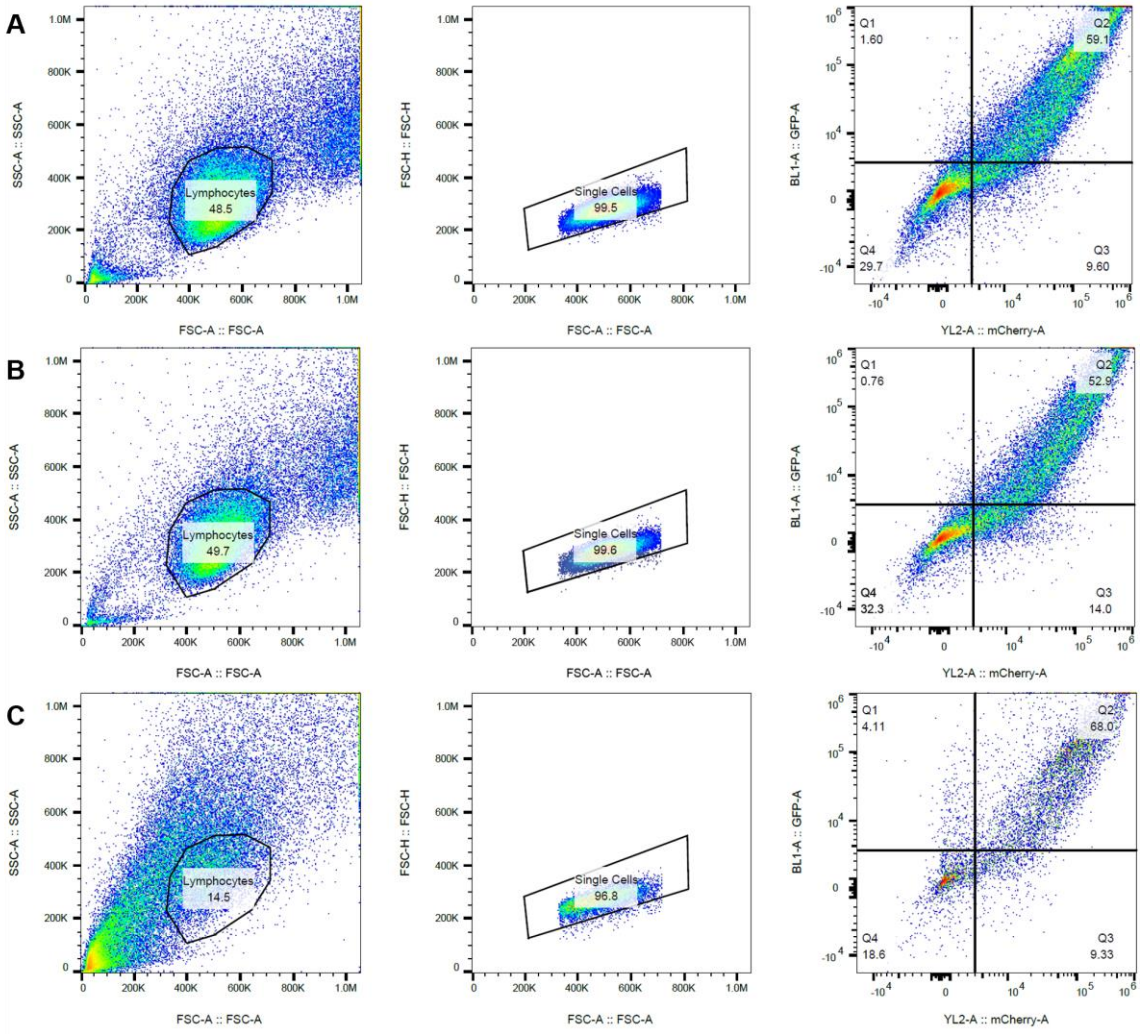


Figure 6.6. Flow cytometry data for the DMSO control (**A**) and HyTags **6.1** (**B**) and **6.2** (**C**) in cell lines expressing BRD4(1)^{L94C}-eGFP.

HyTag **6.2** displayed high levels of cytotoxicity in both cell lines, which was evident from the large proportion of dead cells observed under the microscope. This can also be seen in the flowcytometry data, shown by the densely populated bottom left corner of the plot and the low levels of forward scatter (**Figure 6.6 C & Appendix E**). For this reason, degradation results for BRD4(1)-eGFP degradation by **6.2** have not been included. This result was not unexpected given that the **6.2** clickable probe analogues, **5.1** and **5.2**, exhibited poor selectivity for BRD4(1)^{L94C} and proteome-wide reactivity in in-cell protein labelling experiments. It is thought that the reactive and promiscuous nature of this compound causes the high levels of observed cytotoxicity.

HyTag **6.1** did not display any obvious signs of high cytotoxicity by microscopy, and flow cytometry analysis revealed degradation of 21% of BRD4(1)^{L94C} and 27% of BRD4(1)^{WT} (**Figure 6.6 B, Figure 6.7 & Appendix E**). This was a surprising result considering that the click probe analogue, **5.3**, exhibits high levels of BRD4(1)^{L94C} labelling, while observed levels of non-specific BRD4(1)^{WT} labelling appeared far lower by in-cell protein labelling experiments. It is proposed that the difference in observed cytotoxicity for **6.2** and **6.1** might be due to the differing solubilities of the HyTags, where the lower solubility of **6.1** decreases the effective concentration compared to the more soluble **6.2** HyTag. It is presumed that the high promiscuity of **6.1** and **6.2**, demonstrated by their clickable derivatives, leads to low levels of BRD4(1)^{L94C} being labelled for degradation, and that the high levels of off-target labelling decrease the amount of the covalent HyTag 'available' for BRD4(1)^{L94C} labelling.

Intriguingly, the saturated negative controls **6.11** and **6.12** displayed degradation of both BRD4(1)^{L94C} and BRD4(1)^{WT}, suggesting that the non-covalent interactions between the isoxazole scaffold of the HyTags and the bromodomains may be enough to induce low levels of protein degradation by hydrophobic tagging (**Figure 6.7**). However, this experiment was only performed once, and should therefore be repeated, along with further experiments, to allow for significant interpretation.

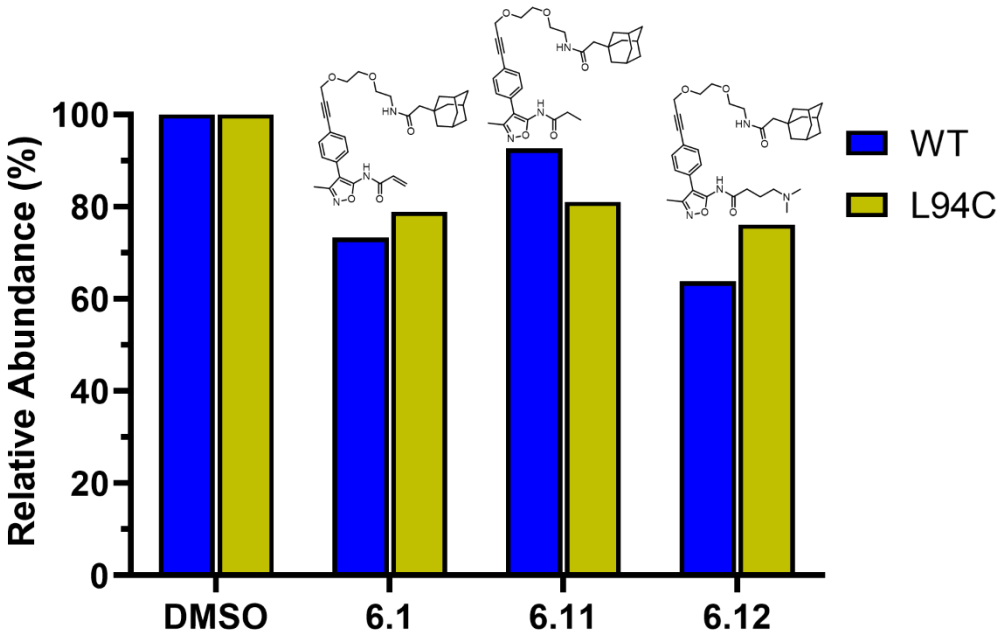


Figure 6.7. Flow cytometry data for the HyTags against BRD4(1)^{WT} and BRD4(1)^{L94C}. The ratio of median eGFP to mCherry signal for successfully transfected cells were calculated using FlowJo. Protein degradation was quantified relative to the ratio of median eGFP to mCherry signal for the DMSO controls (n=1).

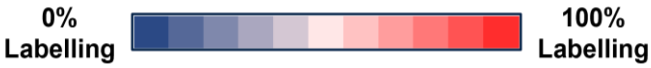
6.6 Current work

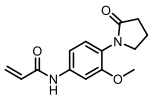
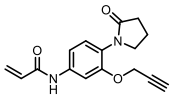
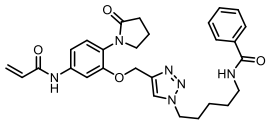
Given the promising selectivity data for **3.50** and clickable analogue **5.8**, discussed in **Chapter 5**, it was investigated whether a protein degrader based on the γ-lactam scaffold would promote selective degradation of BRD4(1)^{L94C}. It was envisioned that a triazole could connect the covalent fragment to the linker, which could be achieved by 'clicking' on the desired linker groups to **5.8**, using a CuAAC reaction. To determine whether linker installation would be tolerated at this position, a proof-of-concept compound, **6.13**, was designed (**Appendix E**). Inspired by the work of Min *et al*. a C-5 saturated linker was chosen.[199] For this proof-of-concept compound, a simple benzyl group attached to the linker *via* an amide bond was chosen, as this would allow for rapid synthesis and information on linker positioning. The synthesis of **6.13** was conducted by Dr. Marta Serafini.

Subsequent protein LC-MS assays revealed that **6.13** retained covalent labelling of BRD4(1)^{L94C}, confirming that this position was suitable for linker attachment (**Table 6.2 & Figure 6.8**). Interestingly, **6.13** exhibited higher levels of labelling than **3.50** and **5.8**, and after 1 hour of

incubation full labelling of BRD4(1)^{L94C} could be observed. This observation indicates a potential increase in affinity for BRD4(1)^{L94C}, possibly caused by interactions between the triazole and the WPF shelf, or increased reactivity through more optimal alignment of the acrylamide and the cysteine thiol.

Table 6.2. Results obtained from the protein LC-MS covalent binding assay for the γ-lactam-based compounds. The table shows the percentage of labelled protein after incubation at 37 °C for given time points. Results are shown as the mean of three repeats for hits (n=3), or two repeats for non-hits and **6.13** (n=2), and the errors given as the SEM. The colour scale is shown below, where low levels of protein labelling are shown in blue, and high levels shown in red.



		Labelled Protein (%)		Labelled Protein (%)	
		L94C		WT	
Compound	Structure	1 h	4 h	1 h	4 h
3.5		16.5 ± 0.8	60.9 ± 2.0	ND	ND
5.8		29.1 ± 1.2	75.5 ± 0.1	ND	ND
6.13		100.0 ± 0.0	100.0 ± 0.0	ND	ND

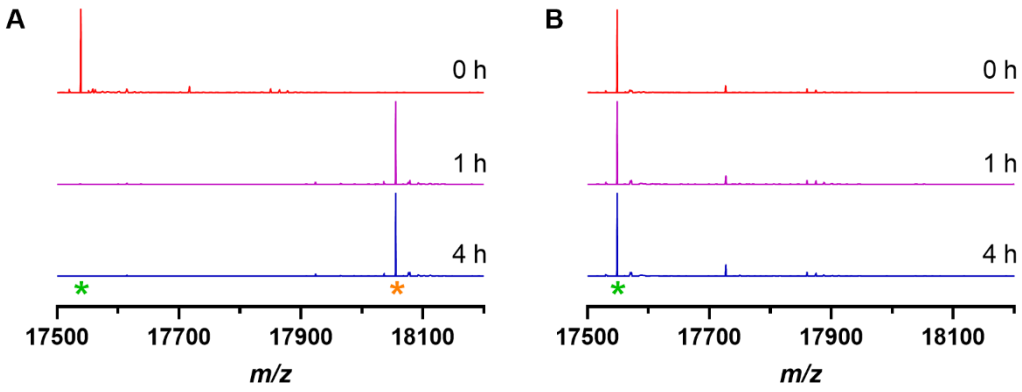


Figure 6.8. Time-dependent protein LC-MS spectra for **6.13** labelling of (**A**) BRD4(1)^{L94C} and (**B**) BRD4(1)^{WT}. The Y-axis shows relative intensity (%). The green asterisk shows peaks corresponding to unlabelled protein, and the orange asterisks show peaks corresponding to labelled protein.

6.7 Conclusions

In conclusion it was investigated whether a covalent fragment, targeting a specifically designed mutant protein, functionalised with a hydrophobic tag could lead to the rapid development of selective protein degraders. Two covalent HyTag analogues of fragments **4.2** and **4.3** were synthesised and tested, along with the two corresponding saturated controls.

Both covalent HyTags **6.1** and **6.2** were shown to covalently label BRD4(1)^{L94C}, however, **6.1** exhibited a decrease in labelling ability compared to its parent fragment **4.2**. Using flow cytometry, the HyTags were assessed for their ability to induce degradation of the BRD4(1)^{WT}-eGFP and BRD4(1)^{L94C}-eGFP fusion proteins. For the covalent HyTags **6.1** and **6.2**, high levels of off-target labelling, demonstrated in **Chapter 5** by the click probe analogues, meant that selective degradation of BRD4(1)^{L94C} was not achieved, and in the case of **6.2** led to cytotoxicity.

A proof-of-concept degrader **6.13**, based on fragment **3.50**, revealed that attachment of a triazole linker at the methoxy position was well tolerated, and resulted in an increase in BRD4(1)^{L94C} labelling, suggesting the modifications have increased affinity or reactivity for BRD4(1)^{L94C}. Therefore, future **3.50** protein degraders will be modified at this site for linker attachment.

Although this work targeted the transiently expressed BRD4(1)^{L94C} mutant, the ultimate aim for this work would be to use a genetically engineered cell-line endogenously expressing full length BRD4^{L94C}. Coupled with a selective covalent degrader, this system would allow for reliable phenotypic and target validation studies.

Chapter 7

Conclusions and future work

7.1 Conclusions

The primary aim of the work described in this thesis was to develop a 'mutate & conjugate' methodology for selective covalent modulation of epigenetic proteins. Using the BET bromodomain, BRD4(1), as a proof-of-concept target, this work combined reactive covalent fragments and site-directed-mutagenesis of BRD4(1) to develop covalent ligands that display selectivity for the mutant BRD4(1) over the wildtype and related bromodomains.

Analysis of X-ray crystal structures led to hypothesis that L94 of BRD4(1)^{WT} would be a suitable residue for the cysteine mutation. The mutant BRD4(1)^{L94C} was expressed, and was validated as a structurally stable, and functioning mutant, using biophysical assays and X-ray crystallography.

To identify reactive fragments capable of labelling C94 of BRD4(1)^{L94C} a structure guided, and a diverse fragment screening approach were employed. Of the ~270 reactive fragments screened only five fragments (**3.1**, **3.4**, **3.5**, **3.11**, and **3.50**) were progressed for further investigation due to their ability to effectively label C94. These data demonstrated that fragments required sufficient affinity for the protein in order to covalently label C94.

For fragments **3.4** and **3.5**, it was found that a modification of the 4-position with a phenyl group, to yield **4.2** and **4.3**, increased the fragment's affinity for the KAc binding pocket of BRD4(1)^{L94C}, shown by an increase in labelling of BRD4(1)^{L94C} (**Figure 7.1**). Furthermore, **4.2** and **4.3** displayed selective inhibition of BRD4(1)^{L94C} over BRD4(1)^{WT} with low micromolar IC_{50} values— indicating that these fragments could display high in-cell selectivity.

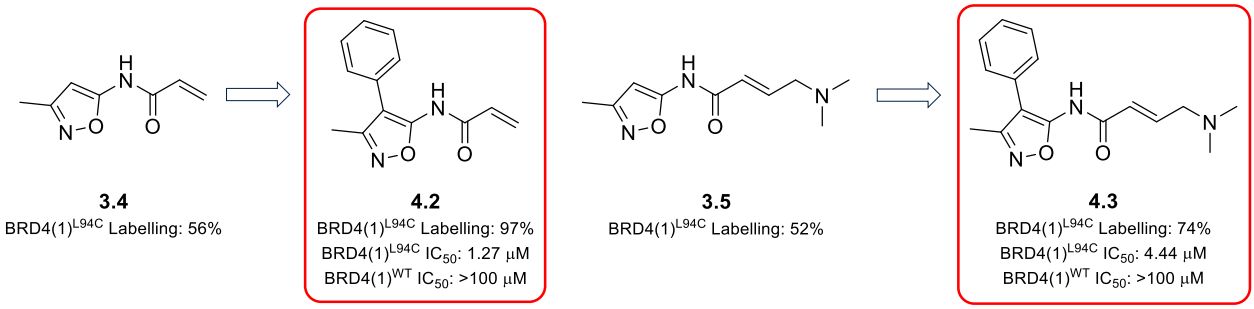


Figure 7.1. The development of covalent fragments, **3.4** and **3.5**, identified in **Chapter 3**.

To evaluate the in-cell selectivity of fragments **4.2**, **4.3**, and **3.50**, clickable probes were developed. In HEK293T cells transiently expressing BRD4(1)^{L94C} the isoxazole-based click probes (**5.1**, **5.2**, and **5.3**) displayed high levels of off-target labelling, however engagement with BRD4(1)^{L94C} by **5.3** was verified. On the other hand, **5.8** exhibited exquisite cellular engagement and selectivity for BRD4(1)^{L94C}, and without significant structural optimisation and SAR (**Figure 7.2**). The evidence suggests that **3.50**, and its clickable analogue **5.8**, display selectivity levels suitable for phenotypic studies of BRD4(1)^{L94C} inhibition in cells expressing endogenous BRD4^{L94C}. This result is substantial and indicates that the combination of site directed mutagenesis and covalent fragments could yield selective probes to interrogate complex biological systems.

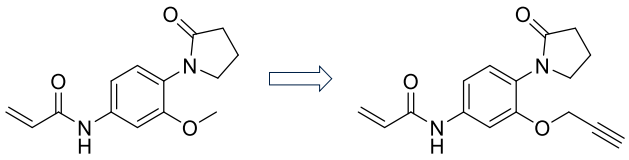


Figure 7.2. The development of **3.50** into clickable probe **5.8** revealed high in-cell selectivity for BRD4(1)^{L94C}.

It was then investigated whether 'mutate & conjugate' could be applied to the development of protein degraders, and hence induce selective degradation of mutant BRD4(1)^{L94C}. To accelerate the speed of degrader development, a HyTag approach was employed, with the fragments functionalised with a PEG-linker and an adamantane group. A flow cytometry assay was used to measure protein degradation. The covalent HyTags **6.1** and **6.2** did not induce selective degradation of BRD4(1)^{L94C}, and in the case of **6.2** high levels of cytotoxicity were observed. Following the promising selectivity data for **5.8**, a covalent proof-of-concept degrader, **6.13**, was designed. Protein LC-MS data showed that **6.13** displayed efficient labelling of C94 in just 1 hour, indicating that linker attachment by a triazole is well tolerated for future derivatisation.

7.2 Future work

The work described in this thesis led to the discovery of **3.50** as a selective covalent ligand for BRD4(1)^{L94C}, and therefore should be investigated to further validate the application of the 'mutate & conjugate' methodology.

Whilst **3.50,** and its clickable analogue **5.8,** exhibit no covalent labelling of BRD4(1)^{WT}, further studies could be conducted to ensure selective inhibition of BRD4(1)^{L94C} over BRD4(1)^{WT}. This selectivity data can be obtained by employing the well characterised AlphaScreen™ assay, previously described in **Chapter 4**. If **3.50** and **5.8** exhibit selective inhibition of BRD4(1)^{L94C} then a wider off-target screening method, such as BromoScan, could be employed.

Currently, docking studies have not allowed for full understanding of how **3.50** engages with BRD4(1)^{L94C}, and therefore an X-ray crystal structure of **3.50** covalently bound to BRD4(1)^{L94C} would be desirable. Structural data would help guide future functionalisation of **3.50**.

Having demonstrated that **6.13** is able to efficiently and covalently label BRD4(1)^{L94C}, future work could investigate synthesising a small library of PROTACs and HyTags to demonstrate the applicability of 'mutate and conjugate' to protein degradation.

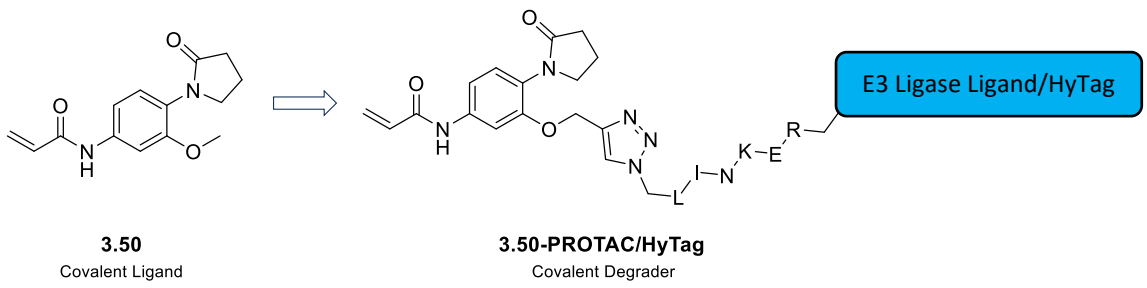


Figure 7.3. Proposed development of a **3.50** based protein degrader.

The ultimate aim of the 'mutate & conjugate' approach is to validate targets in the most physiologically relevant systems, as this will allow for the most reliable phenotypic data. Whilst the work described in this thesis demonstrates target engagement through the use of clickable probes and a transiently expressed protein system, meaningful phenotypic data cannot be guaranteed,

due to the overexpression of the target protein. Future work could aim to incorporate the mutation at the endogenous level, either through CRISPR-Cas9 mediated site directed mutagenesis or a knock-in/ knock-out system, as this would achieve a full length $BRD4^{L94C}$ mutant that is expressed and regulated as if it were $BRD4^{WT}$. Once this is achieved, BRD4(1) function can be fully interrogated with the use of the selective covalent probe, **3.50**.

Finally, the 'mutate & conjugate' methodology could be applied to newly emerging targets, or for targets which currently do not have selective probes. By applying 'mutate & conjugate' to these types of targets, it is envisaged that this will expand the ligandable proteome and drive the discovery and validation of targets through the rapid development of selective probes.

Chapter 8

Experimental

8.1 General chemical methods

Reagents and solvents were obtained from commercial sources (Sigma-Aldrich, Fluorochem and Alfa Aesar) and were used without further purification, unless otherwise stated. Anhydrous solvents were taken from MBRAUN Solvent Purification System 5, and stored over 3 Å molecular sieves, under an inert atmosphere of Ar. All reactions were carried out in a flame-dried flask, and under an inert atmosphere of Ar, unless otherwise stated. Concentration *in vacuo* refers to solvent removal under reduced pressure at 40 °C using a Buchi™ rotary evaporator, unless otherwise stated. Saturated aqueous solution of NaCl is referred to as brine. Petroleum ether refers to the fractions collected between 40–60 °C when distilled.

Analytical thin layer chromatography (TLC) was carried out on Merck silica gel 60 F_{254} aluminium-supported thin layer chromatography sheets. Visualisation was by absorption of UV light (λ_{max} 254 nm), or thermal development after staining in either an: aqueous basic solution of potassium permanganate, or ethanolic solution of ninhydrin.

Preparative thin layer chromatography (Prep TLC) was carried out on Macherey-Nagel precoated 60 silica TLC plates. Visualisation was by absorption of UV light (λ_{max} 254 nm).

Flash column chromatography was carried out using Merck silica gel (40–63 µm), eluting using solvents as supplied under a positive pressure of N_2 gas.

Semi-preparative high performance liquid chromatography (HPLC) was carried out an Agilent 1260 Infinity II® with an Agilent 5 Prep C18 column (5 µM, 21.2 × 50 mm); 95% H_2O/ 5% MeCN + 0.1% FA modifier (1 min), 95% H_2O/ 5% MeCN to 5% H_2O/ 95% MeCN + 0.1% FA modifier (10 min), hold (5 min); flow rate 20 mL/min.

Analytical high performance liquid chromatography (HPLC) was carried out on a PerkinElmer Flexar system with a Binary LC Pump and UV/Vis LC detector. For determination of compound purity, a Dionex Acclaim® 120 column (C18, 5 µm, 120 Å, 4.6 × 150 mm) was employed, with a

10 min gradient of 95% H_2O/ 5% MeCN + 0.1% TFA (solvent A) to 95% MeCN/ 5% H_2O + 0.1% TFA (solvent B), flow rate 1 mL/min. Samples injected were prepared in MeCN and filtered. All samples were run with 0.1% TFA added, unless otherwise stated.

^{1}H NMR spectra were recorded using a Bruker AVII HD 400 (400 MHz), Bruker AVIII HD 500 (500 MHz) or Bruker NEO 600 with broadband helium cryoprobe spectrometer using the stated solvents as a reference for internal deuterium lock. The chemical shift data for each signal are given as δ_H in units of parts per million (ppm). The spectra are calibrated using the solvent peak with the data provided by Fulmer *et al*.[200] The multiplicity of each signal is indicated by: s (singlet); br s (broad singlet); d (doublet); t (triplet); br t (broad triplet); q (quartet); quin (quintet) dd (doublet of doublets); dt (doublet of triplets); m (multiplet) or combinations thereof. The number of protons, n, for a given resonance signal is indicated by nH. Coupling constants (*J*) are expressed in Hz and are recorded to the nearest 0.1 Hz. Identical proton coupling constants (*J*) are averaged in each spectrum and reported to the nearest 0.1 Hz. MestReNova and TopSpin software were used for NMR analysis.

^{13}C NMR spectra were recorded using a Bruker AVIII HD 400 spectrometer (101 MHz), Bruker AVIII HD 500 (126 MHz) with broadband proton decoupling, or Bruker NEO 600 (151 MHz) with broadband helium cryoprobe spectrometer using the stated solvent as a reference for internal deuterium lock. The chemical shift data for each signal are quoted as δ_C in parts per million (ppm). The spectra are calibrated using the solvent peak with the data provided by Fulmer *et al*.[200]

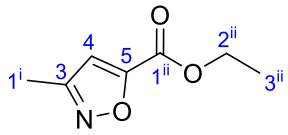
Mass spectra were recorded using either an Agilent 6120 (low resolution) or a Bruker microToF (high resolution) spectrometer using electrospray ionisation (ESI). Samples were submitted as solutions in either methanol or acetonitrile. *m/z* Values are given in Daltons (Da).

Melting points were obtained using a Leica Galen III hot stage microscope or a Griffin capillary tube melting apparatus and are uncorrected. The crystallisation solvent is given in parentheses.

Infrared (IR) spectra were obtained from neat samples. The spectra were recorded using a Bruker Tensor 27 spectrometer with a diamond ATR module. Absorption maxima are given in wavenumbers (cm^{-1}) and reported as s (strong), m (medium), or w (weak).

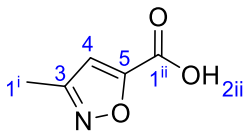
8.2 Synthetic procedures and compound data

Ethyl 3-methylisoxazole-5-carboxylate (3.18)[148,201]



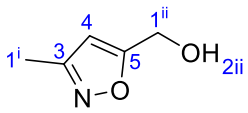
To a solution of acetaldoxime (**3.19**) (3.00 g, 50.8 mmol, 1.0 eq) in DMF (54 mL), was added *N*-chlorosuccinimide (8.82 g, 60.0 mmol, 1.3 eq) at 0 °C. The solution was slowly warmed to rt, stirred for 3 h, diluted with brine (200 mL) and extracted with Et_2O (5 × 150 mL). The organic components were combined, washed with brine (200 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo*. The residue was dissolved in CH_2Cl_2 (30 mL) and added to a suspension of ethyl propiolate (5.66 mL, 55.9 mmol, 1.1 eq), $KHCO_3$ (13.2 g), H_2O (0.5 mL), CH_2Cl_2 (160 mL) and stirred for 16 h. The reaction was quenched by the addition of H_2O (40 mL) to dissolve the $KHCO_3$. The phases were separated, organic components washed with H_2O (4 × 100 mL), brine (100 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo*. The crude oil was purified using silica gel column chromatography (0–20% Et_2O/petroleum ether), to isolate the title compound as a pale-yellow oil (3.36 g, 43%); R_f 0.26 (10% Et_2O/petroleum ether); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 6.77 (1H, s, 4-C*H*), 4.40 (2H, q, *J* 7.2, 2^{ii}-CH_2), 2.36 (3H, s, 1^{i}-CH_3), 1.39 (3H, t, *J* 7.2, 3^{ii}-CH_3); LRMS *m/z* (ESI$^+$) 156 ([M+H]$^+$, 100%), 178 ([M+Na]$^+$, 63%). These data are in good agreement with the literature.[148,201]

3-Methylisoxazole-5-carboxylic acid (3.22)[150]



LiOH (541 mg, 12.9 mmol, 2.0 eq) was added to a solution of ethyl 3-methylisoxazole-5-carboxylate (**3.18**) (1.00 g, 6.45 mmol, 1.0 eq) in THF/H_2O (2:1, 37.5 mL). The solution was stirred at rt for 16 h. The THF was removed *in vacuo*, and the remaining solution was acidified to pH 1 with aqueous HCl (6.0 M). The resulting solid was collected by filtration, washed with cold H_2O (35 mL), and dried *in vacuo*, to isolate the title compound as colourless solid (534 mg, 66%). R_f 0.13 (10% MeOH/EtOAc); m.p. 170–172 °C (from H_2O) [lit.[150] 170–171 °C]; ^{1}H NMR (400 MHz, DMSO-D_6) δ_H 7.03 (1H, s, 4-C*H*), 2.29 (3H, s, 1^i-CH_3); LRMS *m/z* (ESI$^-$) 126 ([M-H]$^-$, 100%). These data are in good agreement with the literature.[150]

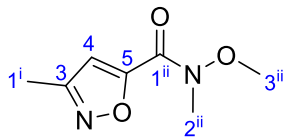
(3-Methylisoxazol-5-yl)methanol (3.24)[150,202]



$BH_3.SMe_2$ complex in THF (2.0 M, 1.58 mL, 3.15 mmol, 2.0 eq) was added dropwise to a solution of 3-methylisoxazole-5-carboxylic acid (**3.22**) (200 mg, 1.57 mmol, 1.0 eq) in anhydrous THF (3.0 mL) at 0 °C. The solution was slowly warmed to rt and stirred for 44 h, before being cooled to 0 °C and quenched by the addition of H_2O (2.0 mL), aqueous NaOH (3.8 M, 1.5 mL), and aqueous H_2O_2 (30%, 1.5 mL). The reaction mixture was extracted with 10% IPA in $CHCl_3$ (3 × 15 mL), dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude product was purified using silica gel column chromatography (50–70% EtOAc/petroleum ether), to isolate the title compound as a colourless oil (32 mg, 18%); R_f 0.17 (50% EtOAc/petroleum ether); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 6.06 (1H, s, 4-C*H*),

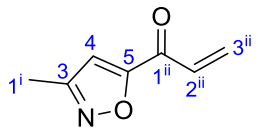
4.68 (2H, s, 1^{ii}-CH_2), 3.29 (1H, br s, 2^{ii}-O*H*), 2.26 (3H, s, 1^{i}-CH_3); LRMS *m/z* (ESI$^+$) 114 ([M+H]$^+$, 100%). These data are in good agreement with the literature. [150,202]

_N_-Methoxy-_N_-3-dimethylisoxazole-5-carboxamide (3.28)



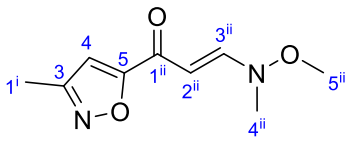
N,O-Dimethylhydroxylamine hydrochloride (299 mg, 3.07 mmol, 1.3 eq), and anhydrous triethylamine (981 µL, 7.08 mmol, 3.0 eq) were added to solution of 3-methylisoxazole-5-carboxylic acid (**3.22**) (300 mg, 2.36 mmol, 1.0 eq) in anhydrous CH_2Cl_2 (10 mL). In a separate flask, PyBOP (2.46 g, 4.72 mmol, 2.0 eq) was dissolved in anhydrous CH_2Cl_2 (10 mL), this solution was then added dropwise to the reaction mixture at 0 °C. The solution was slowly warmed to rt and stirred for 16 h. The reaction mixture was extracted with CH_2Cl_2 (20 mL) and washed with H_2O (2 × 30 mL), saturated aqueous $NaHCO_3$ solution (30 mL), and brine (30 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel column chromatography (40–60% EtOAc/petroleum ether), to isolate the title compound as a colourless solid (325 mg, 81%); R_f 0.35 (60% EtOAc/petroleum ether); m.p. 66–68 °C (from 40% EtOAc/petroleum ether); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 2982 (C-H, m), 1650 (C=O, s), 1578 (N-O, m); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 6.69 (1H, s, 4-C*H*), 3.79 (3H, s, 3^{ii}-CH_3), 3.35 (3H, s, 2^{ii}-CH_3), 2.35 (3H, s, 1^{i}-CH_3); ^{13}C NMR (101 MHz, $CDCl_3$) δ_C 161.7 (1^{ii}-C), 160.1 (3-C), 157.5 (5-C), 109.6 (4-C), 62.2 (3^{ii}-C), 33.2 (2^{ii}-C), 11.5 (1^{i}-C); HRMS *m/z* (ESI$^+$) [Found: 171.07643, $C_7H_{11}O_3N_2$ ([M+H]$^+$) requires 171.07642]; LRMS *m/z* (ESI+) 171 ([M+H]$^+$, 100%), 193 ([M+Na]$^+$, 21%); HPLC Retention time 5.7 min, >99.0%.

1^{ii}-(3-Methylisoxazol-5-yl)prop-2^{ii}-en-1^{ii}-one (3.1)



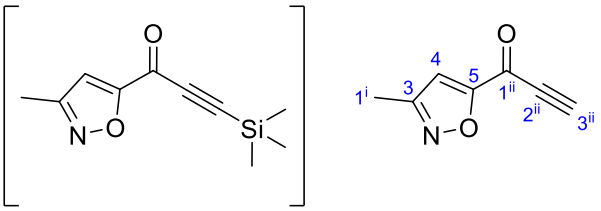
Vinyl magnesium bromide (0.7 M in THF, 1.90 mL, 1.32 mmol, 1.5 eq) was added dropwise over 5 min to a solution of *N*-methoxy-*N*-3-dimethylisoxazole-5-carboxamide (**3.28**) (150 mg, 0.881 mmol, 1.0 eq) in anhydrous THF (3 mL) at −35 °C. The solution was slowly warmed to 0 °C, and stirred for 2 h. The reaction was quenched by pouring over ice-cold aqueous HCl (2.0 M, 30 mL), and extracted with ice-cold EtOAc (3 × 30 mL). The extracts were washed with cold H_2O (50 mL), brine (50 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo*. The crude yellow solid was purified using silica gel column chromatography (20–40% EtOAc/petroleum ether), to isolate the title compound as a colourless solid (68 mg, 56%); R_f 0.62 (40% EtOAc/petroleum ether); m.p 148–150 °C (from CH_2Cl_2); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3117 (C-H, w), 2917 (C-H, w), 1674 (C=O, s); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 7.10 (1H, dd, *J* 17.2, 10.6, 2^{ii}-C*H*), 6.82 (1H, s, 4-C*H*), 6.65 (1H, dd, *J* 17.2, 1.6, 3^{ii}-C*H*$_x$H$_y$), 6.04 (1H, dd, *J* 10.6, 1.6, 3^{ii}-CH$_x$*H*$_y$), 2.39 (3H, s, 1^{i}-C*H*$_3$) ; ^{13}C NMR (101 MHz, $CDCl_3$) δ_C 178.6 (1^{ii}-C), 166.7 (5-C), 160.7 (3-C), 132.5 (3^{ii}-C), 131.5 (2^{ii}-C), 109.2 (4-C), 11.6 (1^{i}-C); HRMS *m/z* (ESI$^+$) [Found: 138.05508, $C_7H_8O_2N$ ([M+H]$^+$) requires 138.05496]; LRMS *m/z* (ESI$^+$) 138 ([M+H]$^+$, 100%); HPLC Retention time 7.0 min, >99.0%.

€-3^{ii}-[Methoxy(methyl)amino]-1^{ii}-(3-methylisoxazol-5-yl)prop-2^{ii}-en-1^{ii}-one (3.31)



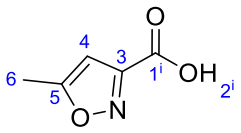
Ethynylmagnesium bromide (0.5 M in THF, 2.25 mL, 1.12 mmol, 2.0 eq) was added dropwise to a solution of *N*-methoxy-*N*-3-dimethylisoxazole-5-carboxamide (**3.28**) (95 mg, 0.558 mmol, 1.0 eq) in anhydrous THF (5 mL) at −40 °C. The solution was warmed to rt and stirred for 1 h, after which it was cooled to −40 °C, and ethynylmagnesium bromide (0.5 M in THF, 1.13 mL, 0.558 mmol, 1.0 eq) was added dropwise. The reaction was quenched with H_2O (20 mL) and extracted with EtOAc (3 × 20 mL). The organic components were combined, washed with brine (20 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel chromatography (50–100% EtOAc/petroleum ether) to afford the title compound as a yellow solid (60 mg, 55%); R_f 0.23 (EtOAc); m.p. 60–62 °C (from $CHCl_3$); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 2981 (C-H, m), 1646 (C=O, m), 1557 (N-O, s), 1167 (N-H, m); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 7.74 (1H, d, *J* 12.4, 3^{ii}-C*H*), 6.66 (1H, s, 4-C*H*), 6.04 (1H, d, *J* 12.4, 2^{ii}-C*H*), 3.75 (3H, s, 5^{ii}-C*H*$_3$), 3.29 (3H, s, 4^{ii}-C*H*$_3$), 2.35 (3H, s, 1^{i}-C*H*$_3$); ^{13}C NMR (101 MHz, $CDCl_3$) δ_C 176.5 (1^{ii}-C), 169.0 (5-C), 160.5 (3-C), 149.3 (3^{ii}-C), 106.6 (4-C), 92.6 (2^{ii}-C), 60.2 (5^{ii}-C), 39.8 (4^{ii}-C), 11.7 (1^{i}-C); HRMS *m/z* (ESI$^+$) [Found: 197.09203, $C_9H_{13}O_3N_2$ ([M+H]$^+$) requires 197.09207]; LRMS *m/z* (ESI$^+$) 219 ([M+Na]$^+$, 100%), 197 ([M+H]$^+$, 98%); HPLC Retention time 4.9 min, 98.8%.

1^{ii}-(3-Methylisoxazol-5-yl)prop-2^{ii}-yn-1^{ii}-one (3.3)



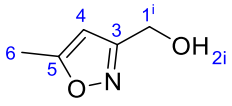
Isopropylmagnesium chloride (1.3 M, 1.22 mL, 1.59 mmol, 3.0 eq) was added dropwise over 5 min to a solution of ethynyltrimethylsilane (256 µL, 1.85 mmol, 3.5 eq) in anhydrous THF (5 mL) at −40 °C. The reaction was slowly warmed to 0 °C. After 60 min the solution was cooled to −40 °C and was added dropwise to a second flask containing *N*-methoxy-*N*-3-dimethylisoxazole-5-carboxamide (**3.28**) (90 mg, 0.529 mmol, 1.0 eq) in anhydrous THF (5.0 mL) at −40 °C. The solution was warmed to 0 °C and stirred for 1.5 h. The reaction was quenched slowly with aqueous NH_4Cl (15 mL), extracted with EtOAc (3 × 15 mL), washed with H_2O (20 mL), brine (20 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude yellow oil was purified by silica gel chromatography (100% toluene). Upon column chromatography the crude silylated product decomposed to reveal the de-silylated product as a pale-yellow solid (19 mg, 27%); R_f 0.16 (100% toluene); m.p 60-62 °C (from $CHCl_3$) ; $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 2982 (C-H, s), 1458 (C=O, w); ^{1}H NMR (500 MHz, $CDCl_3$) δ_H 6.91 (1H, s, 4-C*H*), 3.53 (1H, s, 3^{ii}-C*H*), 2.41 (3H, s, 1^{i}-C*H*$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ_C 165.5 (1^{ii}-C), 163.8 (5-C), 160.9 (3-C), 111.2 (4-C), 82.5 (3^{ii}-C), 79.5 (2^{ii}-C), 11.6 (1^{i}-C); HRMS *m/z* (ESI$^-$) [Found: 134.02429, $C_7H_4O_2N$ ([M-H]$^-$) requires 134.02365]; LRMS *m/z* (ESI$^-$) 134 ([M-H]$^-$, 100%); HPLC Retention time 6.5 min, >99.0%.

5-Methylisoxazole-3-carboxylic acid (3.23)[203]



LiOH (195 mg, 4.64 mmol, 1.2 eq) was added to a solution of ethyl 5-methylisoxazole-3-carboxylate (**3.17**) (600 mg, 3.87 mmol, 1.0 eq) in THF/H_2O (2:1, 45 mL). The solution was stirred for 16 h at rt. The THF was removed *in vacuo*, and the remaining solution was acidified to pH 1 with aqueous HCl (6.0 M). The resulting solid was collected by filtration, washed with cold H_2O (25 mL), and dried *in vacuo*, to isolate the title compound as colourless solid (451 mg, 92%); R_f 0.10 (10% MeOH/EtOAc); m.p. 168–170 °C (from H_2O) [lit.[203] 172-174 °C]; ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 6.47 (1H, s, 4-C*H*), 2.52 (3H, s, 6-C*H*$_3$); LRMS *m/z* (ESI$^+$) (128 [M+H]$^+$, 100%). These data are in good agreement with the literature.[203]

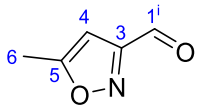
(5-Methylisoxazol-3-yl)methanol (3.25)[204]



BH_3.SMe_2 complex in THF (2.0 M, 1.58 mL, 3.15 mmol, 2.0 eq) was added dropwise to a solution of 5-methylisoxazole-3-carboxylic acid (**3.23**) (200 mg, 1.57 mmol, 1.0 eq) in anhydrous THF (3 mL) at 0 °C. The solution was slowly warmed to rt and stirred for 44 h, before being cooled to 0 °C and quenched by the addition of H_2O (2.0 mL), aqueous NaOH (3.8 M, 1.5 mL), and aqueous H_2O_2 (30%, 1.5 mL). The reaction mixture was extracted with 10% IPA in $CHCl_3$ (3 × 15 mL), washed with brine (30 mL), dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude product was purified using silica gel column chromatography (50–70% EtOAc/petroleum ether), to isolate the title compound as a pale-yellow oil (127 mg, 67%); R_f 0.19 (50% EtOAc/petroleum ether); ^{1}H NMR (400 MHz, $CDCl_3$)

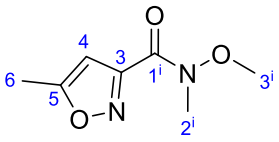
δ_H 6.02 (1H, d, *J* 0.9, 4-C*H*), 4.70 (2H, s, 1^i-CH_2), 2.98 (1H, s, 2^i-O*H*), 2.41 (3H, d, *J* 0.9, 6-CH_3); LRMS *m/z* (ESI^+) 114 ($[M+H]^+$, 100%). These data are in good agreement with the literature.[204]

5-Methylisoxazole-3-carbaldehyde (3.27)[205]



Manganese dioxide (385 mg, 4.42 mmol, 10 eq) was added to a solution of (5-methylisoxazol-3-yl)methanol (**3.25**) (50 mg, 0.442 mmol, 1.0 eq) in anhydrous CH_2Cl_2 (5 mL). The reaction was stirred at rt for 60 h, before being diluted with CH_2Cl_2 (15 mL), and filtered through a pad of Celite®, which was washed with CH_2Cl_2 (20 mL). The filtrate was concentrated *in vacuo* to isolate the title compound as a pale-brown oil (8 mg, 16%); R_f 0.60 (50% EtOAc/petroleum ether); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 9.93 (1H, s, 1^i-C*H*), 6.07 (1H, s, 4-C*H*), 2.40 (3H, s, 6-CH_3); LRMS *m/z* (ESI^+) 112 ($[M+H]^+$, 100%). These data are in good agreement with the literature.[205]

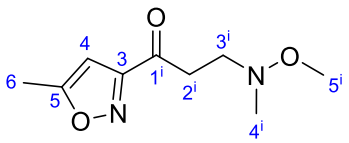
***N*-Methoxy-*N*-5-dimethylisoxazole-3-carboxamide (3.29)**



Anhydrous triethylamine (1.08 mL, 6.30 mmol, 4.0 eq) was added to a solution of 5-methylisoxazole-3-carboxylic acid (**3.23**) (200 mg, 1.57 mmol, 1.0 eq), and *N,O*-dimethylhydroxylamine hydrochloride (200 mg, 2.05 mmol, 1.3 eq), in anhydrous CH_2Cl_2 (10 mL). In a separate flask, PyBOP (2.13 g, 4.09 mmol, 2.6 eq) was dissolved in anhydrous CH_2Cl_2 (5 mL), this solution was then added dropwise to the other solution at 0 °C. The solution was warmed to rt and stirred for 16 h. The solution was extracted with CH_2Cl_2 (20 mL), washed with H_2O

(2 × 20 mL), saturated aqueous $NaHCO_3$ solution (20 mL), and brine (20 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude yellow oil was further purified using silica gel column chromatography (40–60% EtOAc/petroleum ether), to isolate the title compound as a colourless oil (218 mg, 82%); R_f 0.46 (60% EtOAc/petroleum ether); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 2938 (C-H, m), 1656 (C=O, s), 1599 (N-O, w); ^{1}H NMR (500 MHz, $CDCl_3$) δ_H 6.29 (1 H, s, 4-C*H*), 3.75 (3H, s, 3^i-CH_3), 3.35 (3H, s, 2^i-CH_3), 2.44 (3H, s, 6-CH_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ_C 169.8 (5-C), 161.3 (1^i-C), 157.9 (3-C), 102.5 (4-C), 62.4 (3^i-C), 32.9 (2^i-C), 12.2 (6-C); HRMS *m/z* (ESI$^+$) [Found: 171.07645, $C_7H_{11}O_3N_2$ ([M+H]$^+$) requires 171.07642]; LRMS *m/z* (ESI$^+$) 171 ([M+H]$^+$, 33%), 193 ([M+Na]$^+$, 100%); HPLC Retention time 6.0 min, >99.0%.

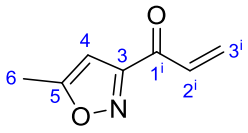
3^i-[Methoxy(methyl)amino]-1^i-(5-methylisoxazol-3-yl)propan-1^i-one (3.30)



Vinylmagnesium bromide (0.7 M in THF, 1.64 mL, 1.15 mmol, 1.3 eq) was added dropwise to a solution of *N*-methoxy-*N*-5-dimethylisoxazole-3-carboxamide (**3.29**) (150 mg, 0.881 mmol, 1.0 eq) in anhydrous THF (10 mL) at 0 °C. The solution was slowly warmed to rt and stirred for 4.5 h. The reaction was quenched with HCl (1.0 M) until pH 2 and extracted with Et_2O (3 × 15 mL), washed with brine (30 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel column chromatography (20–40% EtOAc/petroleum ether) to isolate the title compound as a yellow oil (69 mg, 40%). The product was immediately taken on to the next step; R_f 0.71 (20% EtOAc/petroleum ether); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 6.36 (1H, q, *J* 1.0, 4-C*H*), 3.43 (3H, s, 5^i-CH_3), 3.26 (2H, t, *J* 6.4, 2^i-CH_2), 3.08 (2H, t, *J* 6.4, 3^i-CH_2), 2.59 (3H, s, 4^i-CH_3), 2.48 (3H, d, *J* 1.0, 6-CH_3); ^{13}C NMR (101 MHz, $CDCl_3$) δ_C 193.8 (1^i-C), 171.1 (5-C), 162.1 (3-C), 100.3 (4-C), 59.9

(5^{i}-C), 55.4 (2^{i}-C), 45.2 (4^{i}-C), 38.1 (3^{i}-C), 12.41 (6-C); LRMS *m/z* (ESI^{+}) 199 ([M+H]$^{+}$, 21%), 221 ([M+Na]$^{+}$, 100%).

1^{i}-(5-Methylisoxazol-3-yl)prop-2^{i}-en-1^{i}-one (3.12)



Method 1:

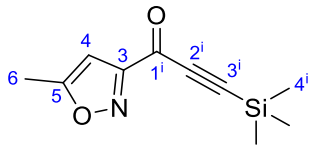
Using a procedure modified from Christine *et al.*,[206] methyl iodide (120 µL, 1.89 mmol, 15.0 eq) was added dropwise to a solution of 3^{i}-[methoxy(methyl)amino]-1^{i}-(5-methylisoxazol-3-yl)propan-1^{i}-one (**3.30**) (25 mg, 0.126 mmol, 1.0 eq) in anhydrous THF (2 mL) at rt, and the solution was stirred for 48 h. The solution was concentrated *in vacuo* to obtain the crude product which was purified using silica gel column chromatography (20–40% EtOAc/petroleum ether), to isolate the title compound as a colourless oil (3 mg, 17%), the data obtained are identical to those shown below.

Method 2:

Vinylmagnesium bromide (0.7 M in THF, 2.83 mL, 1.98 mmol, 1.7 eq) was added dropwise over 5 min to *N*-methoxy-*N*-5-dimethylisoxazole-3-carboxamide (**3.29**) (189 mg, 1.16 mmol, 1.0 eq) in anhydrous THF (4 mL), at −35 °C. The solution was slowly warmed to 0 °C, and stirred for 2 h. The reaction was quenched by pouring over ice-cold aqueous HCl (2.0 M, 30 mL), and extracted with ice-cold EtOAc (3 × 30 mL). The organic extracts were combined, washed with cold H_2O (50 mL), brine (50 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo*. The crude yellow solid was purified using silica gel column chromatography (20–40% EtOAc/petroleum ether), to isolate the title compound as a pale-yellow oil (103 mg, 65%); R_f 0.85 (40% EtOAc/petroleum ether); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 2981 (C-H, s), 2888 (C-H, m), 1658 (C=O, w), 955 (C=C, m); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 7.24 (1H, dd, *J* 17.3, 10.6, 2^{i}-C*H*), 6.65 (1H, dd, *J* 17.3, 1.5, 3^{i}-CH_xH_y), 6.43 (1H, q, *J* 1.0, 4-C*H*), 5.98

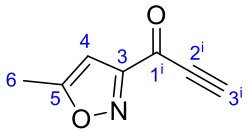
(1H, dd, *J* 10.6, 1.5, 3^{i}-CH$_x$*H*$_y$), 2.50 (3H, d, *J* 1.0, 6-C*H*$_3$); ^{13}C NMR (101 MHz, $CDCl_3$) δ_C 183.6 (1^{i}-C), 171.2 (5-C), 162.3 (3-C), 132.4 (2^{i}-C), 131.6 (3^{i}-C), 101.0 (4-C), 12.4 (6-C); HRMS *m/z* (ESI$^+$) [Found: 138.0549, $C_7H_8O_2N$ ([M+H]$^+$) requires 138.0550]; LRMS *m/z* (ESI$^+$) 138 [M+H]$^+$; HPLC Retention time 3.9 min, >99.0%.

1^{i}-(5-Methylisoxazol-3-yl)-3^{i}-(trimethylsilyl)prop-2^{i}-yn-1^{i}-one (3.32)



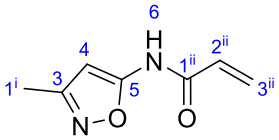
Isopropylmagnesium chloride (1.3 M, 3.38 mL, 4.41 mmol, 5.0 eq) was added dropwise over 5 min to a solution of ethynyltrimethylsilane (794 µL, 5.73 mmol, 6.5 eq) in anhydrous THF (10 mL) at −40 °C. The solution was slowly warmed to 0 °C. After 1 h the solution was cooled to −40 °C and was added dropwise to another flask containing *N*-methoxy-*N*-5-dimethylisoxazole-3-carboxamide (**3.29**) (150 mg, 0.881 mmol, 1.0 eq) dissolved in anhydrous THF (8 mL) at −40 °C. The solution was warmed to 0 °C and stirred for 1.5 h. The reaction was quenched slowly with aqueous NH_4Cl (20 mL), and extracted with EtOAc (3 × 25 mL). The combined organic components were washed with H_2O (20 mL), and brine (20 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel chromatography (0–10% Et_2O/petroleum ether) to isolate the title compound as a pale yellow solid (124 mg, 68%); R_f 0.80 (EtOAc); m.p. 42–44 °C (from $CHCl_3$); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 2981 (C-H, m), 2158 (C≡C, w), 1256 (Si-C, m); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 6.40 (1H, q, *J* 0.9, 4-C*H*), 2.49 (3H, d, *J* 0.9, 6-C*H*), 0.30 (9H, s, 4^{i}-C*H*$_3$); ^{13}C NMR (101 MHz, $CDCl_3$) δ_C 172.4 (1^{i}-C), 170.8 (3-C), 163.3 (5-C), 103.5 (2^{i}-C), 101.3 (4-C), 101.1 (3^{i}-C), 13.2 (6-C), 0.01 (4^{i}-C); HRMS *m/z* (ESI$^+$) [Found: 208.07895, $C_{10}H_{14}O_2N^{28}Si$ ([M+H]$^+$) requires 208.07883]; LRMS *m/z* (ESI$^+$) 230 ([M+Na]$^+$, 100%), 208 ([M+H]$^+$, 38%); HPLC Retention time 9.5 min, >99.0%.

1^{i}-(5-Methylisoxazol-3-yl)prop-2^{i}-yn-1^{i}-one (3.14)



MeOH (137 µL, 3.40 mmol, 35 eq) was added dropwise to a suspension of 1^{i}-(5-methylisoxazol-3-yl)-3^{i}-(trimethylsilyl)prop-2^{i}-yn-1^{i}-one (**3.32**) (20 mg, 0.097 mmol, 1.0 eq), and K_2CO_3 (13 mg, 0.097, 1.0 eq) in CH_2Cl_2 (2 mL) at 0 °C. The suspension was stirred for 20 min and diluted with CH_2Cl_2 (8 mL). The solution was filtered through a silica plug, eluted with EtOAc (10 mL) and concentrated *in vacuo*. The crude product was purified using silica gel chromatography (100% toluene) to isolate the title compound as a colourless solid (6 mg, 46%); R_f 0.24 (toluene); m.p. 51–53 °C (from $CHCl_3$); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 2982 (C-H, s), 1458 (C=O, w); ^{1}H NMR (500 MHz, $CDCl_3$) δ_H 6.44 (1H, q, *J* 0.9, 4-C*H*), 3.52 (1H, s, 3^{i}-C*H*), 2.51 (3H, d, *J* 0.9, 6-C*H*$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ_C 172.1 (1^{i}-C), 169.9 (5-C), 162.4 (3-C), 100.5 (4-C), 82.1 (3^{i}-C), 79.8 (2^{i}-C), 12.5 (6-C); HRMS *m/z* (ESI−) [Found: 134.02429, $C_7H_4O_2N$ ([M-H]$^{-}$) requires 134.02365]; LRMS *m/z* (ESI^{-}) 134 ([M-H]$^{-}$, 100%); HPLC Retention time 6.4 min, 95.4%.

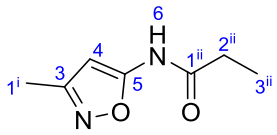
N-(3-Methylisoxazol-5-yl)prop-2^{ii}-enamide (3.4)



Acryloyl chloride (83 µL, 1.02 mmol, 1.0 eq) was added dropwise to a solution of 3-methylisoxazol-5-amine (**3.37**) (100 mg, 1.02 mmol, 1.0 eq), and anhydrous triethylamine (284 µL, 2.04 mmol, 2.0 eq) in anhydrous CH_2Cl_2 (10 mL) at 0 °C. The solution was warmed to rt and stirred for 16 h. The reaction was quenched upon the addition of H_2O (10 mL), and then extracted with CH_2Cl_2 (3 × 15 mL). The organic components were combined and washed with brine (20 mL),

dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel chromatography (0–50% EtOAc/petroleum ether) to isolate the title compound as a colourless solid (94 mg, 61%); R_f 0.44 (50% EtOAc/petroleum ether); m.p. 125–126 °C (from 50% EtOAc/petroleum ether); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3213 (N-H, w), 3048 (C-H, w), 1680 (C=O, m), 1615 (C=C, m); ^{1}H NMR (500 MHz, $CDCl_3$) δ_H 8.16 (1H, s, 6-N*H*), 6.51 (1H, dd, *J* 16.9, 0.9, 3^{ii}-CH_xH$_y$), 6.30 (1H, s, 4-C*H*), 6.25 (1H, dd, *J* 16.9, 10.4, 2^{ii}-C*H*), 5.91 (1H, dd, *J* 10.4, 0.9, 3^{ii}-CH$_x$*H*$_y$), 2.29 (3H, s, 1^{i}-C*H*$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ_C 161.8 (1^{ii}-C), 161.0 (3-C), 159.7 (5-C), 130.5 (3^{ii}-C), 129.2 (2^{ii}-C), 90.2 (4-C), 12.0 (1^{i}-C); HRMS *m/z* (ESI$^+$) [Found: 153.06580, $C_7H_9O_2N_2$ ([M+H]$^+$) requires 153.06585]; LRMS *m/z* (ESI$^+$) 153 ([M+H]$^+$, 100%); HPLC Retention time 4.9 min, >99.0%.

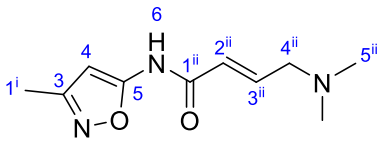
_N_-(3-Methyl-isoxazol-5-yl)propanamide (3.39)



Propionyl chloride (196 µL, 2.24 mmol, 1.1 eq) was added dropwise to a solution of 3-methyl-isoxazol-5-amine (**3.37**) (200 mg, 2.04 mmol, 1.0 eq) and pyridine (182 µL, 2.24 mmol, 1.1 eq) in anhydrous CH_2Cl_2 (20 mL) at 0 °C. The reaction was stirred at rt for 2.5 h, after which another portion of propionyl chloride (53 µL, 0.612 mmol, 0.3 eq) and pyridine (33 µL, 0.408 mmol, 0.2 eq) were added dropwise. The reaction was stirred for a further 1 h, before being quenched by the addition of aqueous HCl (1.0 M, 20 mL), and extracted with CH_2Cl_2 (3 × 30 mL). The organic extracts were combined, washed with H_2O (30 mL) and brine (30 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel column chromatography (10–30% EtOAc/ petroleum ether) to isolate the title compound as a colourless solid (247 mg, 79%); R_f 0.58 (100% EtOAc); m.p 90–92 °C (from $CHCl_3$); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3203 (N-H, m), 3025 (C-H, m), 1690 (C=O, s); ^{1}H NMR (500 MHz, $CDCl_3$) δ_H 8.25 (1H, br s, 6-N*H*), 6.21 (1H, s, 4-C*H*), 2.46 (2H, q,

J 7.5, 2^{ii}-C*H*$_2$), 2.26 (3H, s, 1^{i}-C*H*$_3$), 1.25 (3H, t, *J* 7.5, 3^{ii}-C*H*$_2$); ^{13}C NMR (126 MHz, $CDCl_3$) δ_C 169.9 (1^{ii}-C), 161.8 (3-C), 160.0 (5-C), 89.6 (4-C), 30.0 (2^{ii}-C), 12.0 (1^{i}-C), 9.2 (3^{ii}-C); HRMS *m/z* (ESI$^+$) [Found: 177.06355, $C_7H_{10}O_2N_2{}^{23}Na$ ([M+Na]$^+$) requires 177.06345]; LRMS *m/z* (ESI$^+$) 155 ([M+H]$^+$, 56%), 177 ([M+Na]$^+$, 100%); HPLC Retention time 6.0 min, >99.0%.

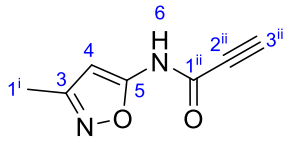
(*E*)-4^{ii}-(Dimethylamino)-*N*-(3-methylisoxazol-5-yl)but-2^{ii}-enamide (3.5)



Oxalyl chloride (54 µL, 0.636 mmol, 1.05 eq) was added dropwise to a suspension of (*E*)-4-dimethylaminobut-2-enoic acid hydrochloride (100 mg, 0.606 mmol, 1.0 eq) in anhydrous THF (3 mL), and anhydrous DMF (1 drop) at 0 °C. The suspension was stirred at 0 °C for a further 30 min and then warmed to rt and stirred for 2 h. The solution was then cooled to 0 °C, and 3-methylisoxazol-5-amine (**3.37**) (65 mg, 0.666 mmol, 1.1 eq) dissolved in anhydrous THF (3 mL) and anhydrous triethylamine (100 µL) was added dropwise. The solution was stirred for 2 h at rt. The reaction was quenched with saturated aqueous $NaHCO_3$ solution until pH 10, and extracted with EtOAc (3 × 30 mL). The organic components were combined and washed with brine (50 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo*. The crude product was purified by preparative TLC (92% CH_2Cl_2/7% EtOH /1% aqueous NH_4OH), to isolate the title compound as a colourless solid (9 mg, 7%); R_f 0.10 (92% CH_2Cl_2/7% EtOH /1% aqueous NH_4OH); m.p 96–98 °C (from $CHCl_3$); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 2980 (C-H, m), 1686 (C=O, m), 1613 (C=C, m), 1267 (C-N, w); ^{1}H NMR (500 MHz, $CDCl_3$) δ_H 9.13 (1H, br s, 6-N*H*), 7.05 (1H, dt, *J* 15.4, 6.0, 3^{ii}-C*H*), 6.28 (1H, s, 4-C*H*), 6.15 (1H, dt, *J* 15.4, 1.6, 2^{ii}-C*H*), 3.14 (2H, dd, *J* 6.0, 1.6, 4^{ii}-C*H*$_2$), 2.29 (6H, s, 5^{ii}-C*H*$_3$), 2.27 (3H, s, 1^{i}-C*H*$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ_C 161.8 (1^{ii}-C), 161.3 (3-C), 160.2 (5-C), 145.0 (3^{ii}-C), 123.9 (2^{ii}-C), 90.0

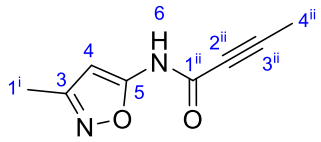
(4-C), 60.3 (4^{ii}-C), 45.5 (5^{i}-C), 11.9 (1^{i}-C); HRMS *m/z* (ESI$^+$) [Found: 210.12378, $C_{10}H_{16}O_2N_3$ ([M+H]$^+$) requires 210.12370]; LRMS *m/z* (ESI$^+$) 210 ([M+H]$^+$, 100%); HPLC Retention time 7.8 min, 96.0%.

N-(3-Methylisoxazol-5-yl)prop-2^{ii}-ynamide (3.8)



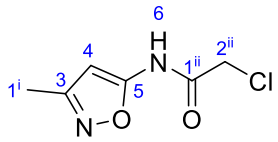
3-Methylisoxazol-5-amine (**3.37**) (100 mg, 1.02 mmol, 1.0 eq), DMAP (1.30 mg, 0.01 mmol, 0.01 eq), and *N,N'*-dicyclohexylcarbodiimide (273 mg, 1.33 mmol, 1.3 eq) were dissolved in anhydrous CH_2Cl_2 (5 mL). The solution was cooled to 0 °C. A solution of propiolic acid (82 µL, 1.33 mmol, 1.3 eq) in anhydrous CH_2Cl_2 (5 mL) was added dropwise to the previous solution. Upon addition the reaction turned to a bright yellow suspension. The reaction was slowly warmed to rt, stirred for 5 h, diluted with CH_2Cl_2 (20 mL) and filtered. The resulting filtrate was concentrated *in vacuo*, and purified using silica gel column chromatography (10–30% EtOAc/petroleum ether) to isolate the title compound as a colourless solid (47 mg, 31%); R_f 0.53 (50% EtOAc/petroleum ether); m.p 125 °C dec (from $CHCl_3$); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3219 (C-H, m), 2109 (C≡C, m), 1684 (C=O, s); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 8.71 (1H, br s, 6-N*H*), 6.26 (1H, s, 4-C*H*), 3.06 (1H, s, 3^{ii}-C*H*), 2.28 (3H, s, 1^{i}-CH_3); ^{13}C NMR (101 MHz, $CDCl_3$) δ_C 161.8 (3-C), 158.7 (5-C), 147.1 (1^{ii}-C), 91.4 (4-C), 76.7 (3^{ii}-C), 76.2 (2^{ii}-C), 12.0 (1^{i}-C); HRMS *m/z* (ESI$^+$) [Found 151.05035, $C_7H_7O_2N_2$ ([M+H]$^+$) requires 151.05020]; LRMS *m/z* (ESI$^+$) 151 ([M+H]$^+$, 100%); HPLC Retention time 6.1 min, 95.8%.

_N_-(3-Methylisoxazol-5-yl)but-2^{ii}-ynamide (3.9)



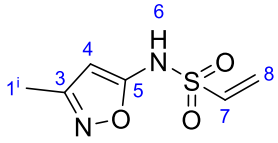
3-Methylisoxazol-5-amine (**3.37**) (150 mg, 1.53 mmol, 1.0 eq), 2-butynoic acid (167 mg, 1.99 mmol, 1.3 eq), and DMAP (18 mg, 0.153 mmol, 0.1 eq), were dissolved in CH_2Cl_2 (13 mL). In a separate flask *N,N'*-dicyclohexylcarbodiimide (410 mg, 1.99 mmol, 1.3 eq) was dissolved in CH_2Cl_2 (2 mL). The *N,N'*-dicyclohexylcarbodiimide solution was then added dropwise to the first solution at 0 °C. The solution was stirred at 0 °C for 15 min, and then warmed to rt to stir for 4 h. The reaction mixture was filtered through Celite® which was washed with EtOAc (20 mL). The filtrate was washed with H_2O (20 mL), and extracted with EtOAc (3 × 30 mL). The combined organic components were washed with brine (30 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel chromatography (10–40% EtOAc/petroleum ether) to isolate the title compound as a pale-yellow solid (40 mg, 16%); R_f 0.32 (40% EtOAc/petroleum ether); m.p 123–125 °C; $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3185 (C-H, w), 2925 (C-H, m), 2248 (C≡C, m), 1647 (C=O, m), 1292 (C-N, s); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 8.66 (1H, s, 6-N*H*), 6.22 (1H, s, 4-C*H*), 2.27 (3H, s, 1^i-C*H*$_3$), 2.02 (3H, s, 4^{ii}-C*H*$_3$); ^{13}C NMR (101 MHz, $CDCl_3$) δ_C 161.6 (3-C), 159.4 (5-C), 148.5 (1^{ii}-C), 90.8 (4-C), 87.6 (2^{ii}-C), 74.2 (3^{ii}-C), 12.0 (1^i-C), 4.0 (4^{ii}-C); HRMS *m/z* (ESI$^+$) [Found: 165.0660, $C_8H_9O_2N_2$ ([M+H]$^+$) requires 165.0659]; LRMS *m/z* (ESI$^+$) 165 ([M+H]$^+$, 100%); HPLC Retention time 5.5 min, >99.0%.

2^{ii}-Chloro-*N*-(3-methylisoxazol-5-yl)acetamide (3.11)



Chloroacetyl chloride (81 µL, 1.02 mmol, 1.0 eq) was added dropwise to a suspension of 3-methylisoxazol-5-amine (**3.37**) (100 mg, 1.02 mmol, 1.0 eq), K_2CO_3 (563 mg, 4.08 mmol, 4.0 eq), in anhydrous CH_2Cl_2 (10 mL) at 0 °C, and the mixture stirred for 16 h at rt. The reaction was quenched by the addition of H_2O (20 mL), extracted with CH_2Cl_2 (3 × 20 mL), and concentrated *in vacuo*. The crude product was purified using silica gel chromatography (20–50% EtOAc/petroleum ether) to isolate the title compound as a colourless solid (81 mg, 46%); R_f 0.58 (50% EtOAc/petroleum ether); m.p 110–111 °C (from $CHCl_3$); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3296 (N-H, m), 2991 (C-H, w), 1702 (C=O, s), 1624 (N-H, m), 1404 (C-H, m), 790 (C-Cl, m); ^{1}H NMR (500 MHz, $CDCl_3$) δ_H 9.25 (1H, br s, 6-N*H*), 6.27 (1H, s, 4-C*H*), 4.22 (2H, s, 2^{ii}-CH_2), 2.28 (3H, s, 1^i-CH_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ_C 162.5 (1^{ii}-C), 161.8 (3-C), 159.1 (5-C), 90.5 (4-C), 42.4 (2^i-C), 12.0 (1^i-C); HRMS *m/z* (ESI$^-$) [Found: 173.01201, $C_6H_6O_2N_2{}^{35}Cl^-$ ([M-H]$^-$) requires 173.01233, Found: 175.00903 $C_6H_6O_2N_2{}^{37}Cl$ ([M-H]$^-$) requires 175.00938]; LRMS *m/z* (ESI$^-$) 173 ([^{35}M-H]$^-$, 100%), 175 ([^{37}M-H]$^-$, 33%); HPLC Retention time 5.5 min, >99.0%.

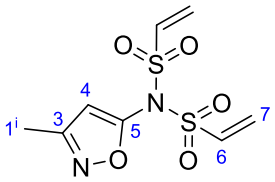
***N*-(3-Methylisoxazol-5-yl)ethenesulfonamide (3.10)**



2-Chloroethanesulfonyl chloride (176 µL, 1.68 mmol, 1.1 eq) and anhydrous triethylamine (466 µL, 3.36 mmol, 2.2 eq) were added dropwise to a solution of 3-methylisoxazol-5-amine (**3.37**) (150 mg, 1.53 mmol, 1.0 eq) in anhydrous CH_2Cl_2 (15 mL), at −15 °C. The solution was slowly warmed to 0 °C,

and stirred for 3 h. The reaction was quenched with aqueous HCl (1.0 M, 3.0 mL), extracted with CH_2Cl_2 (3 × 20 mL), washed with H_2O (30 mL), and brine (30 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel column chromatography (0–10% EtOAc/CH_2Cl_2 and 0.1% AcOH) to isolate the title compound as a yellow oil (10 mg, 4%); R_f 0.11 (10% EtOAc/CH_2Cl_2 + 0.1% AcOH); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3057 (C-H, m), 1615 (N-H, m), 1352 (S=O, m); ^{1}H NMR (500 MHz, $CDCl_3$) δ_H 9.10 (1H, s, 6-N*H*), 6.60 (1H, dd, *J* 16.5, 9.9, 7-C*H*), 6.38 (1H, dd, *J* 16.5, 0.6, 8-CH_xH$_y$), 6.07 (1H, dd, *J* 9.9, 0.6, CH$_x$*H*$_y$), 5.77 (1H, s, 4-C*H*), 2.28 (3H, s, 1^i-C*H*$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ_C 162.0 (3-C), 160.1 (5-C), 134.6 (7-C), 129.5 (8-C), 89.6 (4-C), 11.9 (1^i-C); LRMS *m/z* (ESI$^+$) 189 ([M+H]$^+$, 100%); HRMS *m/z* (ESI$^+$) [Found 189.03285, $C_6H_9O_3N_2{}^{32}S$ ([M+H]$^+$) requires 189.03284]; HPLC Retention time 5.5 min, 96.0%.

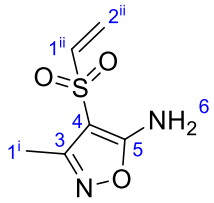
_N_-(Ethenesulfonyl)-_N_-(3-methylisoxazol-5-yl)ethenesulfonamide (3.41)



2-Chloroethanesulfonyl chloride (117 µL, 1.12 mmol, 1.1 eq), and anhydrous triethylamine (155 µL, 1.12 mmol, 1.1 eq) were added to a solution of 3-methylisoxazol-5-amine (**3.37**) (100 mg, 1.02 mmol, 1.0 eq) in anhydrous CH_2Cl_2 (10 mL) at 0 °C. The solution was stirred at 0 °C for a further 1 h, after which another portion of anhydrous triethylamine (155 µL, 1.12 mmol, 1.1 eq) was added. The solution was warmed to rt and left to stir for 1 h. The reaction was quenched with H_2O (20 mL), and extracted with CH_2Cl_2 (3 × 15 mL). The organic components were combined, washed with aqueous HCl (1.0 M 10 mL), brine (50 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel column chromatography (0–30% Et_2O/toluene) to isolate the title compound as a colourless solid (65 mg, 23%); R_f 0.63 (30% Et_2O/toluene); m.p 100–102 °C (from $CHCl_3$); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3088 (C-H, w), 1614 (C=C, m), 1384 (S=O, s), 1365 (S=O,

s); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 7.00 (2H, dd, *J* 16.5, 9.9, 6-C*H*), 6.46 (2H, dd, *J* 16.5, 1.2, 7-CH_xH_y), 6.27 (2H, dd, *J* 9.9, 1.2, 7-CH_xH_y), 6.20 (1H, s, 4-C*H*), 2.32 (3H, s, 1^{i}-CH_3); ^{13}C NMR (101 MHz, $CDCl_3$) δ_C 162.0 (3-C), 155.5 (5-C), 135.5 (6-C), 131.54 (7-C), 105.3 (4-C), 12.3 (1^{i}-C); HRMS *m/z* (ESI$^-$) [Found: 276.99587, $C_8H_9O_5N_2{}^{32}S_2$ ([M-H]$^-$) requires 276.99584]; LRMS *m/z* (ESI$^+$) 279 ([M+H]$^+$, 100%), 189 ([M-$C_2H_3O_2{}^{32}$S+2H]$^+$, 86%), (ESI$^-$) 277 ([M-H]$^-$, 100%); HPLC Retention time 7.2 min, 92.2%.

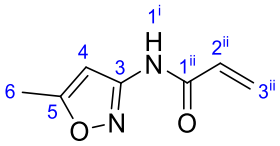
3-Methyl-4-(vinylsulfonyl)isoxazol-5-amine (3.42)



2-Chloroethanesulfonyl chloride (117 µL, 1.12 mmol, 1.1 eq), and anhydrous triethylamine (155 µL, 1.12 mmol, 1.1 eq) were added to a solution of 3-methylisoxazol-5-amine (**3.37**) (100 mg, 1.02 mmol, 1.0 eq) in anhydrous CH_2Cl_2 (10 mL) at 0 °C. The solution was stirred at 0 °C for a further 1 h, after which another portion of anhydrous triethylamine (155 µL, 1.12 mmol, 1.1 eq) was added. The solution was warmed to rt and left to stir for 1 h. The reaction was quenched with H_2O (20 mL), and extracted with CH_2Cl_2 (3 × 15 mL). The organic components were combined, washed with aqueous HCl (1.0M 10 mL), brine (50 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel column chromatography (0–30% Et_2O/ toluene) to isolate the title compound as a colourless oil (10 mg, 5%); R_f 0.32 (50% EtOAc/petroleum ether); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3334 (C-H, m), 1643 (C=C, s), 1572 (N-H, s), 1305 (S=O, m); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 6.64 (1H, dd, *J* 16.5, 9.8, 1^{ii}-C*H*), 6.35 (1H, d, *J* 16.5, 2^{ii}-CH_xH_y), 6.00 (1H, d, *J* 9.8, 2^{ii}-CH_xH_y), 5.84 (2H, br s, 6-NH_2), 2.23 (3H, s, 1^{i}-CH_3); ^{13}C NMR (101 MHz, $CDCl_3$) δ_C 169.3 (5-C), 158.4 (3-C), 138.9 (1^{ii}-C), 126.4 (2^{ii}-C), 91.1 (4-C), 10.8 (1^{i}-C); HRMS *m/z* (ESI$^+$) [Found 189.03284, $C_6H_9O_3N_2{}^{32}$S

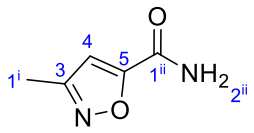
([M+H]$^+$) requires 189.03284]; LRMS *m/z* (ESI$^+$) 189 ([M+H]$^+$, 100%); HPLC Retention time 4.8 min, 98.6%.

_N_-(5-Methylisoxazol-3-yl)prop-2^{ii}-enamide (3.15)



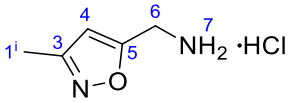
Acryloyl chloride (83 µL, 1.02 mmol, 1.0 eq) was added dropwise to a solution of 5-methylisoxazol-3-amine (**3.37**) (100 mg, 1.02 mmol, 1.0 eq), and anhydrous triethylamine (284 µL, 2.04 mmol, 2.0 eq) in anhydrous CH_2Cl_2 (10 mL) at 0 °C. The solution was warmed to rt and stirred for 16 h. The reaction was quenched upon the addition of H_2O (10 mL), and extracted with CH_2Cl_2 (3 × 15 mL). The organic components were combined, washed with brine (20 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude material was purified using silica gel chromatography (0–50% EtOAc/petroleum ether) to isolate the title compound as a colourless solid (112 mg, 72%); R_f 0.52 (50% EtOAc/petroleum ether); m.p. 146–148 °C (from EtOAc); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 2924 (C-H, m), 1697 (C=O, s), 1622 (C=C, s); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 10.04 (1H, s, 1^i-N*H*), 6.83 (1H, s, 4-C*H*), 6.51 (1H, dd, *J* 17.0, 1.4, 3^{ii}-CH_xH$_y$), 6.39 (1H, dd, *J* 17.0, 10.1, 2^{ii}-C*H*), 5.87 (1H, dd, *J* 10.1, 1.4, 3^{ii}-CH$_x$*H$_y$*), 2.43 (3H, s, 6-C*H*); ^{13}C NMR (101 MHz, $CDCl_3$) δ_C 170.2 (5-C), 163.6 (1^{ii}-C), 158.6 (3-C), 130.4 (2^{ii}-C), 129.4 (3^{ii}-C), 97.0 (4-C), 12.8 (6-C); HRMS *m/z* (ESI$^+$) [Found: 153.06583, $C_7H_9O_2N_2$ ([M+H]$^+$) requires 153.06585]; LRMS *m/z* (ESI$^+$) 175 ([M+Na]$^+$, 100%); HPLC Retention time 4.5 min, >99.0%.

3-Methylisoxazole-5-carboxamide (3.43)[148]



Ethyl 3-methylisoxazole-5-carboxylate (**3.18**) (800 mg, 5.16 mmol, 1.0 eq) was dissolved in a methanolic solution of NH_3 (3.0 M, 20 mL) and stirred for 16 h at rt. The solution was then concentrated *in vacuo* to obtain the crude product, which was purified using silica gel column chromatography (20–50% EtOAc/petroleum ether) to afford the title compound as a colourless solid (575 mg, 89%); R_f 0.28 (50% EtOAc/petroleum ether); m.p. 142–145 °C dec. (from MeOH); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 6.78 (1H, s, 4-C*H*), 6.45 (1H, br s, 2^{ii}-NH_AH_B), 5.83 (1H, br s, 2^{ii}-NH$_A$$H_B$), 2.37 (3H, s, 1^{i}-C$H_3$); LRMS *m/z* (ESI$^+$) 127 ([M+H]$^+$, 100%). These data are in good agreement with the literature.[148]

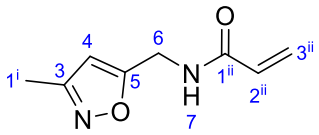
(3-Methylisoxazol-5-yl)methanamine hydrochloride (3.45)[148]



Following a modified procedure from Dannhardt *et al.*,[207] 3-methylisoxazole-5-carboxamide (**3.43**) (350 mg, 2.78 mmol, 1.0 eq) was added to a microwave vial, dissolved in anhydrous THF (28 mL), and the vial sealed. $BH_3.SMe_2$ complex in THF (2.0 M, 6.95 mL, 13.9 mmol, 5.0 eq) was added dropwise at 0 °C. The solution was heated under reflux for 6 h. The solution was then cooled to 0 °C, and MeOH (16.5 mL) was added dropwise with vigorous stirring. The solution was further stirred for 16 h at rt, and then concentrated *in vacuo*. Aqueous HCl (1.0 M, 16 mL) was added to the resultant oil and the solution heated under reflux for 1 h. The solution was cooled to 0 °C, basified to pH 10 with NaOH (1.0 M), extracted with CH_2Cl_2 (3 × 30 mL), washed with brine (30 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The resulting yellow oil was dissolved in Et_2O

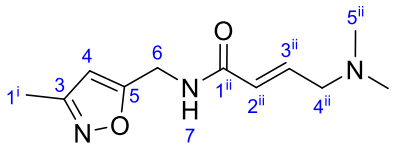
(5 mL) and HCl (2.0 M in Et_2O, 3 mL) was added. The resulting solid was collected by filtration and dried *in vacuo*, to isolate the title compound as colourless solid (270 mg, 66%). R_f 0.06 (10% MeOH/EtOAc); m.p. 172–174 °C dec. (from Et_2O) [lit.[148] 175 °C] ^{1}H NMR (400 MHz, DMSO-D_6) δ_H 8.70 (3H, br s, 7-NH_3), 6.50 (1H, s, 4-C*H*), 4.21 (2H, s, 6-C*H*), 2.25 (3H, s, 1^i-C*H*); LRMS *m/z* (ESI+) 113 ([M+H]$^+$, 100%). These data are in good agreement with the literature.[148]

_N_-[(3-Methylisoxazol-5-yl)methyl]prop-2^{ii}-enamide (3.6)



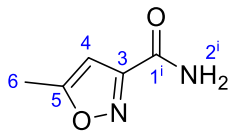
K_2CO_3 (92 mg, 0.664 mmol, 1.2 eq) was added to a solution of (3-methylisoxazol-5-yl)methanamine hydrochloride (**3.45**) (90 mg, 0.608 mmol, 1.1 eq) in anhydrous CH_2Cl_2 (6 mL). The solution was cooled to 0 °C and acryloyl chloride (56 µL, 0.553 mmol, 1.0 eq) was added dropwise. The solution was slowly warmed to rt, and stirred for 3 h. The reaction was diluted with CH_2Cl_2 (10 mL) and washed with H_2O (10 mL). The organic component was extracted with CH_2Cl_2 (3 × 10 mL). The organic components were combined, washed with brine (20 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo.* The resultant yellow oil was purified using silica gel column chromatography (0–50% EtOAc/petroleum ether) to isolate the title compound as a colourless solid (76 mg, 75%); R_f 0.19 (50% EtOAc/ petroleum ether); m.p 55–57 °C (from $CHCl_3$) ; $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3277 (N-H, s br), 2980 (C-H, m), 1662 (C=O, s), 1608 (C=C, s); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 6.43 (1H, s, 7-N*H*), 6.32 (1H, dd, *J* 17.0, 1.4, 3^{ii}-CH_xH$_y$), 6.13 (1H, dd, *J* 17.0, 10.3, 2^{ii}-C*H*), 6.04 (1H, s, 4-C*H*), 5.69 (1H, dd, *J* 10.3, 1.4, 3^{ii}-CH$_x$*H$_y$*), 4.57 (2H, d, *J* 6.0, 6-C*H*), 2.25 (3H, s, 1^i-C*H*); ^{13}C NMR (101 MHz, $CDCl_3$) δ_C 168.4 (1^{ii}-C), 165.6 (5-C), 160.2 (C-3), 130.1 (2^{ii}-C), 127.7 (3^{ii}-C), 103.3 (4-C), 35.3 (6-C), 11.5 (1^i-C); HRMS *m/z* (ESI$^+$) [Found: 167.08151, $C_8H_{11}O_2N_2$ ([M+H]$^+$) requires 167.08150]; LRMS *m/z* (ESI$^+$) 167 ([M+H]$^+$, 100%), 189 ([M+Na]$^+$, 29%); HPLC Retention time 4.2 min, 98.6%.

(*E*)-4^{ii}-(Dimethylamino)-*N*-(3-methylisoxazol-5-yl)but-2^{ii}-enamide (3.7)



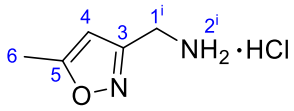
Oxalyl chloride (74 µL, 0.861 mmol, 1.6 eq) was added dropwise to a suspension of (*E*)-4-dimethylaminobut-2-enoic acid hydrochloride (124 mg, 0.754 mmol, 1.4 eq) in anhydrous THF (5 mL) at 0 °C. The suspension was stirred at 0 °C for a further 30 min and then warmed to rt and stirred for 2 h. The solution was concentrated *in* vacuo at 0 °C to remove the volatile components. The resulting brown oil was dissolved in anhydrous THF (2 mL) and cooled to 0 °C. A solution of (3-methylisoxazol-5-yl)methanamine hydrochloride (**3.45**) (80 mg, 0.538 mmol, 1.0 eq) dissolved in anhydrous THF (3 mL), and anhydrous triethylamine (186 µL, 1.35 mmol, 2.5 eq) was added dropwise to the previous solution. The solution was stirred for 16 h at rt. The reaction was quenched with saturated aqueous $NaHCO_3$ solution, and extracted with EtOAc (3 × 30 mL). The organic components were combined and washed with brine (30 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel column chromatography (0–100% 50:8:1 CH_2Cl_2: EtOH: aqueous NH_4OH/ CH_2Cl_2) to isolate the title compound as a colourless oil (40 mg, 35%); R_f 0.37 (100:8:1 CH_2Cl_2: EtOH: aqueous NH_4OH); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3274 (N-H, s br), 2974 (C-H, m), 1675 (C=O, s), 1641 (C=C, s), 1609 (C=C, m); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 6.84 (1H, dt, *J* 15.4, 6.0, 3^{ii}-C*H*), 6.56 (1H, br t, *J* 6.0, 7-N*H*), 6.01 (1H, s, 4-C*H*), 5.98 (1H, dt, *J* 15.4, 1.6, 2^{ii}-C*H*), 4.60 (2H, d, *J* 6.0, 6-CH_2), 3.08 (2H, dd, *J* 6.0, 1.6, 4^{ii}-CH_2), 2.29 (3H, s, 1^{i}-CH_3), 2.27 (6H, s, 5^{ii}-CH_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ_C 168.5 (1^{ii}-C), 165.5 (5-C), 160.2 (3-C), 142.7 (3^{ii}-C), 124.5 (2^{ii}-C), 103.2 (4-C), 60.5 (4^{ii}-C), 45.7 (5^{ii}-C), 35.3 (6-C), 11.5 (1^{i}-C); HRMS *m/z* (ESI$^+$) [Found: 224.13930, $C_{11}H_{18}O_2N_3$ ([M+H]$^+$) requires 224.13935]; LRMS *m/z* (ESI$^+$) 224 ([M+H]$^+$, 100%), 246 ([M+Na]$^+$, 20%); HPLC Retention time 7.4 min, 96.8%.

5-Methylisoxazole-3-carboxamide (3.44)[208]



Ethyl 5-methylisoxazole-3-carboxylate (**3.17**) (500 mg, 3.22 mmol, 1.0 eq) was dissolved in a methanolic solution of NH_3 (3.0 M, 12 mL) and stirred for 16 h at rt. The solution was then concentrated *in vacuo* to afford the crude product, which was purified using silica gel column chromatography (50% EtOAc/petroleum ether) to isolate the title compound as colourless solid (386 mg, 95%); R_f 0.42 (50% EtOAc/petroleum ether); m.p. 164–166 °C (from CH_2Cl_2) [lit.[208] 166 °C]; ^{1}H NMR (400 MHz, DMSO-D_6) δ_H 8.03 (1H, br s, 2^i-NH_AH$_B$), 7.75 (1H, br s, 2^i-NH$_A$$H_B$), 6.50 (1H, s, 4-C*H*), 2.45 (3H, s, 6-CH_3); LRMS *m/z* (ESI$^+$) 127 ([M+H]$^+$, 92%), 149 ([M+Na]$^+$, 100%). These data are in good agreement with the literature.[208]

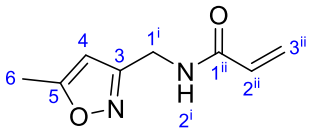
1^i-(5-Methylisoxazol-3-yl methanamine) hydrochloride (3.46)[209]



Following a procedure modified from Dannhardt *et al.*,[207] 5-methylisoxazole-3-carboxamide (**3.44**) (100 mg, 0.793 mmol, 1.0 eq) was added to a microwave vial, dissolved in anhydrous THF (5 mL), and the vial sealed. BH_3.SMe_2 complex in THF (2.0 M, 1.70 mL, 3.17 mmol, 4.0 eq) was then added at 0 °C. The solution was heated under reflux for 6 h. The reaction was cooled to 0 °C, and MeOH (2.5 mL) was added dropwise with vigorous stirring. The vial was sealed and stirred for 16 h at rt. The solution was concentrated *in vacuo*. Aqueous HCl (1.0 M, 2.5 mL) was added to the resultant oil, and the solution was heated under reflux for 1 h. The solution was cooled to 0 °C, basified to pH 10 with aqueous NaOH (1.0 M) and extracted with CH_2Cl_2 (3 × 15 mL). The organic components were combined, washed with brine (20 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*.

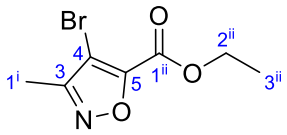
The resulting yellow oil was dissolved in Et_2O (2 mL) and HCl (2.0 M in Et_2O, 0.70 mL) was added. The colourless solid was collected by filtration and dried *in vacuo*, to isolate the title compound as colourless solid (39 mg, 33%); R_f 0.08 (10% MeOH/EtOAc); m.p. 198–201 °C dec. (from Et_2O) [lit.[209] 202–203 °C (dec.)]; ^{1}H NMR (400 MHz, methanol-D_4) δ_H 6.28 (1 H, s, 4-C*H*), 4.23 (2H, s, 1^i-C*H*$_2$), 2.47 (3H, s, 6-C*H*$_3$); LRMS *m/z* (ESI$^+$) 113 ([M+H]$^+$, 100%).

_N_-[(5-Methylisoxazol-3-yl)methyl]prop-2^{ii}-enamide (3.16)



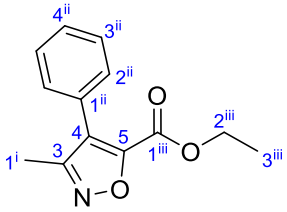
Triethylamine (50 µL, 0.504 mmol, 3.0 eq), and acryloyl chloride (27 µL, 0.336 mmol, 2.0 eq) were added to a solution of 1^i-(5-methylisoxazol-3-yl methanamine) hydrochloride (**3.46**) (25 mg, 0.168 mmol, 1.0 eq) in anhydrous CH_2Cl_2 (2 mL) at 0 °C. The solution was warmed to rt, stirred for 18 h, and then concentrated *in vacuo*, to obtain the crude product which was purified using silica gel column chromatography (0–50% EtOAc/petroleum ether) to isolate the title compound as a colourless solid (10 mg, 36%); R_f 0.21 (50% EtOAc/petroleum ether); m.p. 158–160 °C (from CH_2Cl_2); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3281 (N-H, m), 2927 (C-H, w), 1655 (C=O, s) 1610 (C=C, s); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 6.32 (1H, dd, *J* 17.0, 1.4, 3^{ii}-C*H*$_y$H$_x$), 6.20 (1H, br s, 2^i-N*H*), 6.14 (1H, dd, *J* 17.0, 10.3, 2^{ii}-C*H*), 5.99 (1H, q, *J* 0.8, 4-C*H*), 5.69 (1H, dd, *J* 10.3, 1.4, 3^{ii}-CH$_x$*H*$_y$), 4.54 (2H, d, *J* 5.7, 1^i-C*H*$_2$), 2.29 (3H, d, *J* 0.8, 6-C*H*$_3$); ^{13}C NMR (101 MHz, $CDCl_3$) δ_C 170.3 (5-C), 165.7 (1^{ii}-C), 161.1 (3-C), 130.4 (3^{ii}-C), 127.4 (2^{ii}-C), 101.3 (4-C), 35.7 (1^i-C), 12.4 (6-C); LRMS *m/z* (ESI$^+$) 167 ([M+H]$^+$, 7%), 189 ([M+Na]$^+$, 100%); HRMS *m/z* (ESI$^+$) [Found: 167.08150, $C_8H_{11}O_2N_2$ ([M+H]$^+$) requires 167.08150]; HPLC Retention time 5.1 min, 97.6%.

Ethyl 4-bromo-3-methylisoxazole-5-carboxylate (4.6)[141,148]



Following the procedure of Li *et al.,*[210] ethyl 3-methylisoxazole-5-carboxylate (**3.18**) (300 mg, 1.93 mmol, 1.0 eq), and *N*-bromosuccinimide (412 mg, 2.32 mmol, 1.2 eq) were dissolved in TFA (8 mL) in a microwave vial. The vial was sealed, and the solution was heated in a microwave reactor at 150 °C for 50 min. The solution was cooled, and TFA was removed by azeotropic distillation with toluene. The crude product was purified using silica gel chromatography (0–10% Et_2O/petroleum ether) to isolate the title compound as a colourless oil (292 mg, 65%); R_f 0.30 (10% Et_2O/petroleum ether); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 4.44 (2H, q, *J* 7.2, 2ii-C*H*$_2$), 2.35 (3H, s, 1^i-C*H*$_3$), 1.42 (3H, t, *J* 7.2, 3^{ii}-C*H*$_3$); LRMS *m/z* (ESI$^+$) 234 ([^{79}M+H]$^+$, 100%), 236 ([^{81}M+H]$^+$, 95%). These data are in good agreement with the literature. [141,148]

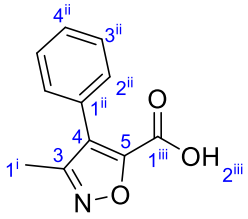
Ethyl 3-methyl-4-phenylisoxazole-5-carboxylate (4.7)[141]



Using a modified procedure from Coombs,[141] ethyl 4-bromo-3-methylisoxazole-5-carboxylate (**4.6**) (350 mg, 1.50 mmol, 1.0 eq), phenylboronic acid (219 mg, 1.79 mmol, 1.2 eq), $NaHCO_3$ (378 mg, 4.50 mmol, 3.0 eq), and Pd(dppf)Cl_2·CH_2Cl_2 (37 mg, 0.045 mmol, 0.03 eq) were added to a microwave vial. The vial was sealed and purged with Ar. A degassed solution of in DME (4 mL), and H_2O (1 mL) was then added to the purged vial. The reaction was heated at 90 °C for 2.5 h. The reaction mixture was then cooled and filtered through a pad of Celite® and silica, washed with Et_2O (100 mL), and the filtrate concentrated *in vacuo*. The crude product was purified using silica gel chromatography (2–10% Et_2O/petroleum ether) to isolate the title compound as a colourless solid

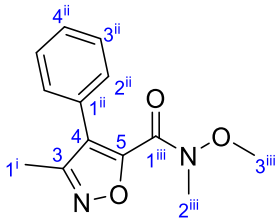
(170 mg, 49%); R_f 0.18 (15% Et_2O/petroleum ether); m.p. 56–58 °C (from $CHCl_3$) [lit.[141] 53 °C]; ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 7.50–7.39 (3H, m, 3^{ii}, 4^{ii}-C*H*), 7.38–7.30 (2H, m, 2^{ii}-C*H*), 4.32 (2H, q, *J* 7.1, 2^{iii}-CH_2), 2.29 (3H, s, 1^{i}-CH_3), 1.27 (3H, d, *J* 7.1, 3^{iii}-CH_3); LRMS *m/z* (ESI$^+$) 232 ([M+H]$^+$, 100%). These data are in good agreement with the literature.[141]

3-Methyl-4-phenylisoxazole-5-carboxylic acid (4.8)



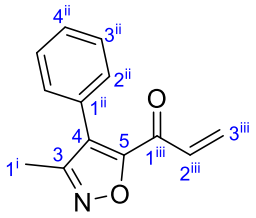
To a solution of ethyl 3-methyl-4-phenylisoxazole-5-carboxylate (**4.7**) (170 mg, 0.735 mmol, 1.0 eq) in THF (4 mL), and H_2O (2 mL), was added LiOH (70 mg, 2.94 mmol, 4.0 eq) at 0 °C. The reaction was stirred and allowed to warm to rt. After 2 h the THF was removed *in vacuo*, and the remaining solution was acidified to pH 2 with aqueous HCl (1.0 M). The suspension was cooled to 0 °C, filtered, and the solid washed with cold H_2O (10 mL). The solid was eluted off the filter with EtOAc (30 mL). The filtrate was concentrated *in vacuo* to isolate the title compound as a colourless solid (125 mg, 84%); R_f 0.63 (100% EtOH); m.p 145–147 °C (from $CHCl_3$); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 1704 (C=O, m), 773 (C-H, s); ^{1}H NMR (500 MHz, $CDCl_3$) δ_H 7.46–7.41 (3H, m, 3^{ii}, 4^{ii}-C*H*), 7.34–7.32 (2H, m, 2^{ii}-C*H*), 2.29 (3H, s, 1^{i}-CH_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ_C 160.9 (1^{ii}-C), 160.8 (3-C), 153.8 (5-C), 129.8 (2^{ii}-C), 129.1 (4^{ii}-C), 128.6 (3^{ii}-C), 127.9 (1^{ii}-C), 126.7 (4-C), 10.8 (1^{i}-C); HRMS *m/z* (ESI$^+$) [Found: 204.06566, $C_{11}H_{10}O_3N$ ([M+H]$^+$) requires 204.06552]; LRMS *m/z* (ESI$^+$) 204 ([M+H]$^+$, 100%); HPLC Retention time 8.0 min, >99.0%.

_N_-Methoxy-_N_,3-dimethyl-4-phenylisoxazole-5-carboxamide (4.9)



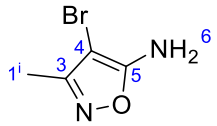
N,O-Dimethylhydroxylamine hydrochloride (68 mg, 0.691 mmol, 1.3 eq), and anhydrous triethylamine (221 µL, 1.60 mmol, 3.0 eq) were added to solution of 3-methyl-4-phenylisoxazole-5-carboxylic acid (**4.8**) (108 mg, 0.532 mmol, 1.0 eq) in anhydrous CH_2Cl_2 (3 mL). In a separate flask, PyBOP (554 mg, 1.06 mmol, 2.0 eq) was dissolved in anhydrous CH_2Cl_2 (2 mL), this solution was then added dropwise to the reaction mixture at 0 °C. The solution was slowly warmed to rt and stirred for 16 h. The reaction mixture was quenched with H_2O (20 mL), and extracted with CH_2Cl_2 (3 × 20 mL). The organic extracts were combined, washed with H_2O (20 mL), and brine (20 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel chromatography (10% EtOAc/ petroleum ether) to isolate the title compound as a colourless oil (128 mg, 98%); R_f 0.63 (50% EtOAc/ petroleum ether); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 2937 (C-H, m), 1662 (C=O); ^{1}H NMR (500 MHz, $CDCl_3$) δ_H 7.44–7.41 (2H, m, 3^{ii}-C*H*), 7.39–7.36 (3H, m, 2^{ii}, 4^{ii}-C*H*), 3.71 (3H, s, 3^{iii}-C*H*$_3$), 3.26 (3H, s, 2^{iii}-C*H*$_3$), 2.34 (3H, s, 1^{i}-C*H*$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ_C 160.0 (1^{iii}-C), 159.1 (3-C), 158.1 (5-C), 129.3 (2^{ii}-C), 128.8 (3^{ii}-C), 128.7 (1^{ii}-C), 128.5 (4^{ii}-C), 121.5 (4-C), 62.4 (3^{iii}-C), 32.8 (2^{iii}-C), 10.8 (1^{i}-C); HRMS *m/z* (ESI$^+$) [Found: 247.10777, $C_{13}H_{15}O_3N_2$ ([M+H]$^+$) requires 247.10772]; LRMS *m/z* (ESI$^+$) 247 ([M+H]$^+$, 91%), 269 ([M+Na]$^+$, 100%); HPLC Retention time 9.2 min, >99.0%.

1^{iii}-(3-Methyl-4-phenylisoxazol-5-yl)prop-2^{iii}-en-1^{iii}-one (4.1)



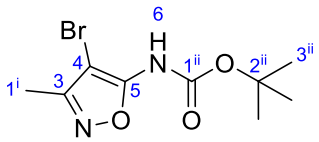
Vinyl magnesium bromide (1.0 M in THF, 446 µL, 0.447 mmol, 2.0 eq) was added dropwise over 5 min to a solution of *N*-methoxy-*N*,3-dimethyl-4-phenylisoxazole-5-carboxamide (**4.9**) (55 mg, 0.223 mmol, 1.0 eq) in anhydrous THF (3 mL) at −35 °C. The solution was slowly warmed to 0 °C, and stirred for 1 h. The reaction was quenched by pouring over ice-cold aqueous HCl (1.0 M, 5 mL), and extracted with ice-cold EtOAc (2 × 20 mL). The organic extracts were washed with brine (10 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel column chromatography (12–100% EtOAc/petroleum ether) to isolate the title compound as a colourless solid (24 mg, 62%); R_f 0.75 (50% EtOAc/ petroleum ether); m.p 37–39 °C (from $CHCl_3$); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3062 (C-H, m), 1703 (C=O, s), 1607 (C=C, s); ^{1}H NMR (500 MHz, $CDCl_3$) δ_H 7.48–7.42 (3H, m, 3^{ii}, 4^{ii}-C*H*), 7.38–7.35 (2H, m, 2^{ii}-C*H*), 7.05 (1H, dd, *J* 17.2, 10.5, 2^{iii}-C*H*), 6.54 (1H, dd, *J* 17.2, 1.5, 3^{iii}-CH_xH$_y$), 5.92 (1H, dd, *J* 10.5, 1.5, 3^{iii}-CH$_x$$H_y$), 2.32 (3H, s, 1^{i}-C$H_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ_C 179.3 (1^{iii}-C), 160.7 (5-C), 160.6 (3-C), 132.1 (2^{iii}-C), 131.8 (3^{iii}-C), 129.9 (2^{ii}-C), 129.1 (4^{ii}-C), 128.9 (3^{ii}-C), 128.4 (1^{ii}-C), 124.7 (4-C), 10.8 (1^{i}-C); HRMS *m/z* (ESI$^+$) [Found: 236.0681, $C_{13}H_{11}O_2NNa$ ([M+Na]$^+$ requires 236.0682]; LRMS *m/z* (ESI$^+$) 214 ([M+H]$^+$, 100%); HPLC Retention time 11.0 min, 96.4%.

4-Bromo-3-methylisoxazol-5-amine (4.11)[148,211,212]



Following the procedure from Chan *et al.*,[211] 5-amino-3-methylisoxazole (**3.37**) (1.00 g, 10.2 mmol, 1.0 eq), was dissolved in anhydrous $CHCl_3$ (15 mL). The solution was cooled to 0 °C, and *N*-bromosuccinimide (1.82 g, 10.2 mmol, 1.0 eq) was added portion wise over 20 min. The solution was stirred at 0 °C for 1 h, diluted with $CHCl_3$ (50 mL), and washed with H_2O (2 × 50 mL). The crude product was purified using silica gel chromatography (10–20% EtOAc/petroleum ether), to isolate the title compound as a yellow solid (1.2 g, 67%); R_f 0.53 (50% EtOAc/petroleum ether); m.p 58–59 °C (from $CHCl_3$) [lit.[148] 58–60 °C]; ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 4.54 (2H, br s, 6-N*H*$_2$), 2.17 (3H, s, 1^i-C*H*$_3$); LRMS *m/z* (ESI$^+$) 177 ([^{79}M+H]$^+$, 100%), 179 ([^{81}M+H]$^+$, 98%). These data are in good agreement with the literature.[148,211,212]

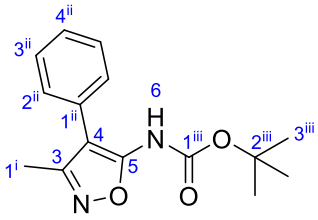
***tert*-Butyl (4-bromo-3-methylisoxazol-5-yl)carbamate (4.14)**[148]



Following the procedure from Hewings,[148] 4-bromo-3-methylisoxazol-5-amine (**4.11**) (250 mg, 1.41 mmol, 1.0 eq), was dissolved in anhydrous THF (7 mL). In a separate flask di-*tert*-butyl dicarbonate (617 mg, 2.83 mmol, 2.0 eq), and DMAP (9 mg, 0.071 mmol, 0.05 eq) were dissolved in anhydrous THF (1.0 mL). This solution was added dropwise to the isoxazole solution. The combined solution was heated under reflux for 22 h, after which additional di-*tert*-butyl dicarbonate (462 mg, 2.12 mmol, 1.5 eq) was added. After a further 2 h under reflux the solution was cooled and concentrated *in vacuo*. The residue was dissolved in MeOH (10 mL) and aqueous NaOH (2.0 M, 1 mL) was added. This solution was stirred at rt for 48 h before the addition of H_2O (5 mL), and extraction with EtOAc (3 × 20 mL). The organic components were combined, washed

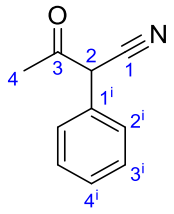
with brine (20 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. No further purification was required, and the title compound was isolated as a colourless solid (361 mg, 93%); R_f 0.32 (Et_2O/petroleum ether); m.p. 72–74 °C (from $CHCl_3$) [lit.[148] 74–75 °C]; ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 6.70 (1H, s, 6-N*H*), 2.25 (1H, s, 1^i-C*H*$_3$), 1.52 (9H, s, 3^{ii}-C*H*$_3$); LRMS *m/z* (ESI$^+$) 299 ([^{79}M+Na]$^+$, 75%), 301 ([^{81}M+Na]$^+$, 100%). These data are in good agreement with the literature.[148]

tert-Butyl (3-methyl-4-phenylisoxazol-5-yl)carbamate (4.17)



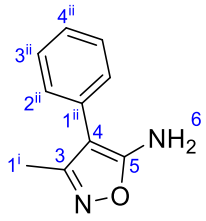
Using a procedure modified from Jedinák *et al.*,[160] *tert*-butyl (4-bromo-3-methylisoxazol-5-yl)carbamate (**4.14**) (50 mg, 0.180 mmol, 1.0 eq), phenylboronic acid (48 mg, 0.397 mmol, 2.2 eq), K_2CO_3 (50 mg, 0.360 mmol, 2.0 eq), XPhos Pd G3 (5 mg, 0.005 mmol, 0.03 eq), and XPhos (3 mg, 0.005 mmol, 0.03 eq) were added to a microwave vial and the vial sealed and purged with Ar. Degassed EtOH/H_2O (4:1, 1.8 mL) was added, and the solution heated at 80 °C for 18 h. The solution was then cooled, and filtered through a pad of Celite® which was washed with Et_2O (20 mL). The filtrate was concentrated *in vacuo*, and the crude product purified using silica gel chromatography (10–30% EtOAc/petroleum ether) to isolate the title compound as a colourless oil (12 mg, 24%); R_f 0.59 (30% EtOAc/petroleum ether); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 7.73–7.14 (5H, m, 2^{ii},3^{ii},4^{ii}-C*H*), 6.98 (1H, s, 6-N*H*), 2.23 (3H, s, 1^i-C*H*), 1.52 (9H, s, 3^{iii}-C*H*); HRMS *m/z* (ESI$^+$) [Found: 275.13913, $C_{15}H_{18}O_3N_2$ ([M+H]$^+$) requires 275.13902]; LRMS *m/z* (ESI$^+$) 275 ([M+H]$^+$, 100%).

3-Oxo-2-phenylbutanenitrile (4.18) [162,213–215]



Using a modified procedure from Smith *et al.,*[162] sodium hydride (60% dispersion in mineral oil, 1.04 g, 26.0 mmol, 1.5 eq) was added portion-wise to a solution of benzyl cyanide (**4.19**) (2.00 mL, 17.3 mmol, 1.0 eq), and anhydrous DMF (3.5 mL) in anhydrous THF (35 mL) at 0 °C. The yellow suspension was stirred for 10 min at 0 °C before the dropwise addition of anhydrous EtOAc (6.77 mL, 69.2 mmol, 4.0 eq) at 0 °C. The suspension was stirred for 18 h at rt. The reaction was quenched at 0 °C by the addition of aqueous HCl (1.0 M, 15 mL), and extracted with EtOAc (4 × 50 mL). The organic components were combined, washed with brine (50 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo* to isolate the title compound as a pale-yellow solid. No further purification was required (1.87 g, 68%); R_f 0.50 (50% EtOAc/petroleum ether); m.p. 73–75 °C (from EtOAc) [lit.[215] 75–77 °C, lit.[214] 89–90 °C, lit.[213] 97 °C]; ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 7.45–7.39 (5H, m, 2^i, 3^i, 4^i-C*H*), 4.67 (1H, s, 2-C*H*), 2.27 (3H, s, 4-CH_3); LRMS *m/z* (ESI$^+$) 160 ([M+H]$^+$, 100%). These data are in good agreement with the literature.[162,213–215]

3-Methyl-4-phenylisoxazol-5-amine (4.10) [161,216,217]



Method 1:

Following the procedure of Krasavin *et al.*,[161] aqueous Na_2CO_3 (10% *w/v*, 1.38 mL), and hydroxylamine solution (50% *w/v* in H_2O, 1.26 mL, 19.1 mmol, 15.1 eq) were added to a solution of 3-oxo-2-phenylbutanenitrile (**4.18**) (200 mg, 1.26 mmol, 1.0 eq) dissolved in EtOH (3 mL). The solution was heated under reflux for 5 h, then cooled and concentrated *in vacuo*. The crude product was purified using silica gel chromatography (10–50% EtOAc/petroleum ether) to afford the title compound as an orange solid (16 mg, 7%), the data obtained are identical to those shown below.

Method 2:

Using a modified procedure from Krasavin *et al.*,[161] aqueous Na_2CO_3 (10% *w/v*, 28 mL), and hydroxylamine hydrochloride (6.18 g, 89.0 mmol, 2.0 eq) were added to a solution of 3-oxo-2-phenylbutanenitrile (**4.18**) (7.08 g, 44.5 mmol, 1.0 eq) dissolved in EtOH (108 mL). The solution was heated under reflux for 40 min, then cooled and concentrated *in vacuo*. The crude material was dissolved in H_2O (20 mL) and extracted with EtOAc (3 × 100 mL). The organic components were combined, washed with brine (50 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo* to isolate the title compound as an orange solid. No further purification was required (6.81 g, 88%), the data obtained are identical to those shown below.

Method 3:

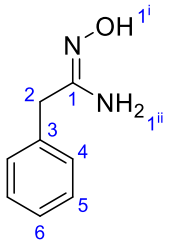
Using a procedure modified from Jedinák *et al.*,[160] *tert*-butyl(4-bromo-3-methylisoxazol-5-yl)carbamate (**4.17**) (50 mg, 0.180 mmol, 1.0 eq), phenylboronic acid (48 mg, 0.397 mmol, 2.2 eq),

K_2CO_3 (50 mg, 0.360 mmol, 2.0 eq), XPhos Pd G3 (5 mg, 0.005 mmol, 0.03 eq), and XPhos (3 mg, 0.005 mmol, 0.03 eq) were added to a microwave vial and the vial sealed and purged with Ar. Degassed EtOH/H_2O (4:1, 1.8 mL) was added to, and the solution heated at 80 °C for 18 h. The solution was then cooled, and filtered through a pad of Celite® which was washed with Et_2O (20 mL). The filtrate was concentrated *in vacuo*, and the crude product purified using silica gel chromatography (10–30% EtOAc/petroleum ether) to isolate the title compound as an orange solid (4 mg, 13%), the data obtained are identical to those shown below.

Method 4:

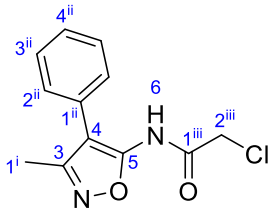
3-Oxo-2-phenylbutanenitrile (**4.18**) (100 mg, 0.628 mmol, 1.0 eq), hydroxylamine hydrochloride (87 mg, 1.26 mmol, 2.0 eq), pyridine (127 µL, 1.57 mmol, 2.5 eq) were dissolved in EtOH (6 mL). The reaction was heated under reflux for 3 h. The solution was cooled to room temperature and concentrated *in vacuo*. The crude product was dissolved in CH_2Cl_2 (3 mL), cooled to 0 °C, and filtered to remove the precipitated *trans*-oxime. The filtrate was concentrated *in vacuo*, and purified using silica gel chromatography (10–40% EtOAc/petroleum ether) to isolate the title compound as an orange solid (42 mg, 38%); R_f 0.53 (50% EtOAc/petroleum ether); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 7.47–7.38 (2H, m, 2^{ii}-C*H*), 7.34–7.26 (3H, m, 3^{ii}, 4^{ii}-C*H*), 4.51 (2H, s, 6-NH_2), 2.24 (3H, s, 1^{i}-CH_3); LRMS *m/z* (ESI$^+$) 175 ([M+H]$^+$, 100%). These data are in good agreement with the literature.[161,216,217]

N^{i}-Hydroxy(phenyl)ethanimidamide (4.20)[218]



Following the procedure of Krasavin *et al.*,[161] aqueous Na_2CO_3 (10% *w/v*, 1.38 mL), and hydroxylamine solution (50% *w/v* in H_2O, 1.26 mL, 19.1 mmol, 15.1 eq) were added to a solution of 3-oxo-2-phenylbutanenitrile (**4.18**) (200 mg, 1.26 mmol, 1.0 eq) dissolved in EtOH (3 mL). The solution was heated under reflux for 5 h, then cooled and concentrated *in vacuo*. The crude product was purified using silica gel chromatography (10–50% EtOAc/petroleum ether) to afford the title compound as a colourless solid (139 mg, 73%); R_f 0.12 (50% EtOAc/petroleum ether); m.p. 60–62 °C (from DMSO) [lit.[218] 64–66 °C]; ^{1}H NMR (400 MHz, DMSO-D_6) δ_H 8.89 (1H, s, 1^{i}-NO*H*), 7.29–7.18 (5H, m, 4,5,6-C*H*), 5.37 (2H, s, 1^{ii}-NH_2), 3.26 (2H, s, 2-CH_2); LRMS *m/z* (ESI$^+$) 151 ([M+H]$^+$, 100%). These data are in good agreement with the literature.[218]

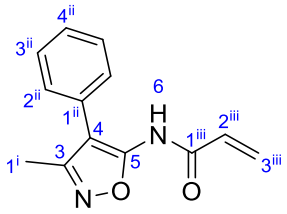
2^{iii}-Chloro-*N*-(3-methyl-4-phenylisoxazol-5-yl)acetamide (4.4)



Chloroacetyl chloride (50 µL, 0.631 mmol, 1.1 eq) was added dropwise to a solution of 3-methyl-4-phenylisoxazol-5-amine (**4.10**) (100 mg, 0.574 mmol, 1.0 eq), and pyridine (51 µL, 0.631 mmol, 1.1 eq) in anhydrous CH_2Cl_2 (5 mL) at 0 °C. The reaction was stirred at 0 °C for 1 h, and a further 2 h at rt. The reaction was quenched with the addition of H_2O (20 mL), and extracted with CH_2Cl_2 (3 × 20 mL). The organic extracts were washed with brine (20 mL), dried over Na_2SO_4, filtered,

and concentrated *in vacuo.* The crude orange solid was purified using silica gel chromatography (20–50% Et_2O/petroleum ether) to isolate the title compound as a colourless solid (34 mg, 24%); R_f 0.21 (50% Et_2O/petroleum ether); m.p 77–78 °C (from $CHCl_3$); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3214 (N-H, w), 1701 (C=O, m); ^1H NMR (500 MHz, $CDCl_3$) δ_H 8.48 (1H, br s, 6-N*H*), 7.46–7.43 (2H, m, 3^{ii}-C*H*), 7.40–7.36 (1H, m, 4^{ii}-C*H*), 7.29–7.27 (2H, m, 2^{ii}-C*H*), 4.15 (2H, s, 2^{iii}-CH_2), 2.31 (3H, s, 1^{i}-CH_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ_C 163.9 (1^{iii}-C), 160.4 (3-C), 154.9 (5-C), 129.2 (3^{ii}-C), 128.7 (1^{ii}-C), 128.7 (2^{ii}-C), 128.4 (4^{ii}-C), 110.0 (4-C), 42.4 (2^{iii}-C), 11.5 (1^{i}-C); HRMS *m/z* (ESI^+) [Found: 251.0583, $C_{12}H_{12}O_2N_2Cl$ ($[M+H]^+$) requires 251.0582]; LRMS *m/z* (ESI^+) 251 ($[^{35}M+H]^+$, 100%), 253 ($[^{37}M+H]^+$, 33%); HPLC Retention time 7.5 min, >99.0%.

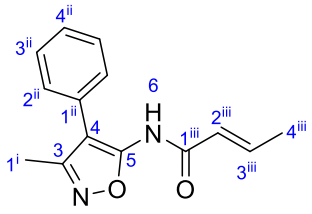
_N_-(3-methyl-4-phenylisoxazol-5-yl)prop-2^{iii}-enamide (4.2)



Acryloyl chloride (77 µL, 0.947 mmol, 1.1 eq) was added dropwise to a solution of 3-methyl-4-phenylisoxazol-5-amine (**4.10**) (150 mg, 0.861 mmol, 1.0 eq), and pyridine (77 µL, 0.947 mmol, 1.1 eq) in anhydrous CH_2Cl_2 (4.0 mL) at 0 °C. The solution was stirred at 0 °C for 2 h, and was then warmed to rt. After 8 h the reaction was cooled back down to 0 °C, and a further portion of acryloyl chloride (35 µL, 0.431 mmol, 0.5 eq) and pyridine (35 µL, 0.431 mmol, 0.5 eq) were added. The reaction was stirred at 0 °C for a further 10 h, after which was quenched with the addition of aqueous HCl (1.0 M, 10 mL), and extracted with CH_2Cl_2 (2 × 20 mL). The organic extracts were washed with H_2O (20 mL), and brine (20 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel chromatography (25–50% Et_2O/petroleum ether) to isolate the title compound as a colourless solid (25 mg, 13%); R_f 0.11 (50% Et_2O/petroleum ether); m.p 85–87 °C (from $CHCl_3$); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3228 (N-H, m), 3035 (C-H, w), 1685 (C=O,

s), 1642 (C=C, s); ^{1}H NMR (500 MHz, DMSO-D_6) δ_H 10.74 (1H, s, 6-N*H*), 7.46–7.40 (2H, m, 3^{ii}-C*H*), 7.40–7.34 (3H m, 2^{ii}, 4^{ii}-C*H*), 6.37 (1H, dd, *J* 17.0, 10.1, 2^{iii}-C*H*), 6.27 (1H, dd, *J* 17.0, 1.8, 3^{iii}-CH_xH_y), 5.83 (1H, dd, *J* 10.1, 1.8, 3^{iii}-CH_xH_y), 2.28 (3H, s, 1^{i}-CH_3); ^{13}C NMR (126 MHz, DMSO-D_6) δ_C 163.8 (1^{iii}-C), 159.8 (3-C), 156.9 (5-C), 129.7 (2^{iii}-C), 129.2 (3^{iii}-C), 129.0 (1^{ii}-C), 128.7 (3^{ii}-C), 128.4 (2^{ii}-C), 127.7 (4^{ii}-C), 110.0 (4-C), 11.3 (1^{i}-C); HRMS *m/z* (ESI$^+$) [Found: 229.0973, $C_{13}H_{13}O_2N_2$ ([M+H]$^+$) requires 229.0972]; LRMS *m/z* (ESI$^+$) 229 ([M+H]$^+$, 100%); HPLC Retention time 8.3 min, >99.0%.

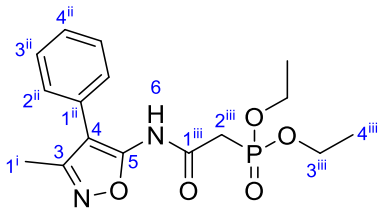
(*E*)-*N*-(3-methyl-4-phenylisoxazol-5-yl)but-2^{iii}-enamide (4.5)



Crotonoyl chloride (319 µL, 3.00 mmol, 1.5 eq) was added dropwise to a solution of 3-methyl-4-phenylisoxazol-5-amine (**4.10**) (348 mg, 2.00 mmol, 1.0 eq), and pyridine (324 µL, 4.00 mmol, 2.0 eq) in anhydrous CH_2Cl_2 (10 mL) at 0 °C. The solution was stirred at 0 °C for 2 h, before being quenched by the addition of aqueous HCl (1.0 M, 20 mL), and extracted with CH_2Cl_2 (3 × 30 mL). The organic extracts were combined, washed with aqueous NaOH (1.0 M, 20 mL) and brine (30 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel column chromatography (20–70% Et_2O/petroleum ether) to isolate the title compound as a colourless oil (25 mg, 13%); R_f 0.24 (60% Et_2O/petroleum ether); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3220 (N-H, m br), 2967 (C-H, m), 1690 (C=O, m), 1643 (C=C, s); ^{1}H NMR (500 MHz, $CDCl_3$) δ_H 7.41–7.38 (2H, m, 3^{ii}-C*H*), 7.34–7.31 (2H, m, 6-N*H*, 4^{ii}-C*H*), 7.26–7.23 (2H, m, 2^{ii}-C*H*), 6.98 (1H, dq, *J* 15.1, 6.9, 3^{iii}-C*H*), 5.98 (1H, d, *J* 15.1, 2^{ii}-C*H*), 2.27 (3H, s, 1^{i}-CH_3), 1.85 (3H, dd, *J* 6.9, 1.6, 4^{iii}-CH_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ_C 163.7 (1^{iii}-C), 160.3 (3-C), 156.4 (5-C), 145.3 (3^{iii}-C), 129.2 (3^{ii}-C), 129.1 (1^{ii}-C), 128.9 (2^{ii}-C), 128.2 (4^{ii}-C), 122.8 (2^{ii}-C), 108.9 (4-C), 18.2 (4^{iii}-C), 11.6 (1^{i}-C); HRMS *m/z* (ESI$^+$)

[Found: 243.1130, $C_{14}H_{15}O_2N_2$ ([M+H]$^+$) requires 243.1128]; LRMS *m/z* (ESI$^+$) 243 ([M+H]$^+$, 100%); HPLC Retention time 8.7 min, 97.5%.

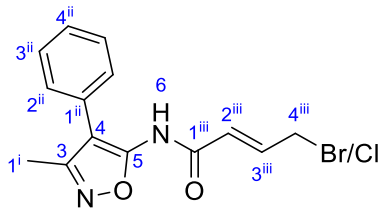
Diethyl {2^{iii}-[(3-methyl-4-phenylisoxazol-5-yl)amino]-2^{iii}-oxoethyl}phosphonate (4.21)



To a solution CDI (782 mg, 4.82 mmol, 3.0 eq) in anhydrous THF (10 mL), was added a solution of diethylphosphonoacetic acid (775 µL, 4.82 mmol, 3.0 eq) in anhydrous THF (5.0 mL). The solution was heated at 40 °C for 30 min, after which a solution of 3-methyl-4-phenylisoxazol-5-amine (**4.10**) (280 mg, 1.61 mmol, 1.0 eq) in anhydrous THF (3 mL) was added dropwise. The solution was heated at 40 °C for 18 h, after which a further solution of CDI (465 mg, 2.87 mmol, 1.8 eq), and diethylphosphonoacetic acid (461 µL, 2.87 mmol, 1.8 eq) in anhydrous THF (2 mL) was added. After 38 h at 40 °C the THF was removed *in vacuo*. The residue was suspended in aqueous NaOH (1.0 M, 20 mL), and extracted with EtOAc (3 × 20 mL). The organic extracts were washed with brine (15 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel chromatography (20–100% EtOAc/petroleum ether) to isolate the title compound as a colourless solid (480 mg, 85%); R_f 0.35 (100% EtOAc); m.p 68–70 °C (from EtOAc); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3161 (N-H, m), 2988 (C-H, m), 1713 (C=O, s), 1237 (P=O, s), 1025 (P-O, s); ^{1}H NMR (500 MHz, $CDCl_3$) δ_H 9.10 (1H, br s, 6-N*H*), 7.44–7.40 (2H, m, 3^{ii}-C*H*), 7.37–7.33 (1H, m, 4^{ii}-C*H*), 7.32–7.29 (2H, m, 2^{ii}-C*H*), 4.10 (4H, dq, *J* 8.2, 7.1 3^{iii}-C*H*$_2$), 2.97 (2H, d, *J* 20.5, 2^{iii}-C*H*$_2$), 2.30 (3H, s, 1^{i}-C*H*$_3$), 1.30 (6H, t, *J* 7.1, 4^{iii}-C*H*$_3$); ^{13}C NMR (126 MHz, $CDCl_3$) δ_C 162.3 (1^{iii}-C), 160.3 (3-C), 155.7 (5-C), 129.0 (3^{ii}-C), 128.9 (2^{ii}-C), 128.1 (4^{ii}-C), 110.0 (4-C), 63.4 (d, *J* 6.6, 3^{iii}-C), 35.4 (d, *J* 131.0, 2^{iii}-C), 16.4 (d, *J* 6.1, 4^{iii}-C), 11.5 (1^{i}-C); ^{31}P NMR (202 MHz; CDCl3) δ_P 21.6; HRMS *m/z* (ESI$^+$) [Found: 353.1259,

$C_{16}H_{22}O_5N_2P$ ([M+H]$^+$) requires 353.1261]; LRMS *m/z* (ESI$^+$) 353 ([M+H]$^+$, 100%); HPLC Retention time 7.8 min, >99.0%.

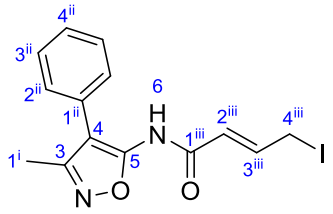
(*E*)-4^{iii}-Bromo/Chloro-*N*-(3-methyl-4-phenylisoxazol-5-yl)but-2^{iii}-enamide (4.25)



Oxalyl chloride (1.07 mL, 12.6 mmol, 2.2 eq) was added dropwise to a solution of (*E*)-4-bromocrotonic acid (1.42 g, 8.61 mmol, 1.5 eq) in anhydrous CH_2Cl_2 (10 mL), and anhydrous DMF (2 drops) at 0 °C. The solution was stirred at 0 °C for a further 30 min and then warmed to rt and stirred for 20 h. The solution was then concentrated *in vacuo*, and the pale-brown oil redissolved in anhydrous CH_2Cl_2 (8.6 mL).

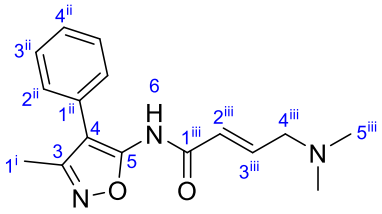
The acyl chloride solution was then added dropwise to a solution of 3-methyl-4-phenylisoxazol-5-amine (**4.10**) (1.00 g, 5.74 mmol, 1.0 eq) and pyridine (557 µL, 6.89 mmol, 1.3 eq) in anhydrous CH_2Cl_2 (25 mL), at −15 °C. The reaction was stirred at −15 °C for 40 min before being quenched with the addition of aqueous HCl (1.0 M, 30 mL), and extracted with CH_2Cl_2 (3 × 100 mL). The organic extracts were washed with H_2O (50 mL), and brine (50 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude brown oil was purified using silica gel chromatography (40–70% Et_2O/petroleum ether) to isolate a mixture of the bromo- or chloro-substituted title compound as a pale-brown solid (450 mg, 24%); R_f 0.29 (60% Et_2O/petroleum ether); ^{1}H NMR (500 MHz, $CDCl_3$) δ_H 7.76 (1H, s, 6-N*H*), 7.44–7.41 (2H, m, 3^{ii}-C*H*), 7.37–7.34 (1H, m, 4^{ii}-C*H*), 7.28–7.26 (2H, m, 2^{ii}-C*H*), 7.06–6.99 (2H, m, 3^{iii}-C*H*), 6.28 (0.7H, d, *J* 15.0, 2^{iii}-C$H_{(Cl)}$), 6.22 (0.3H, d, *J* 15.0, 2^{iii}-C$H_{(Br)}$), 4.17 (1.5H, dd, *J* 5.6, 1.3, 4^{iii}-C$H_{2(Cl)}$), 4.00 (0.5H, dd, *J* 7.0, 1.3, 4^{iii}-C$H_{2(Br)}$), 2.29 (3H, s, 1^{i}-CH_3); LRMS *m/z* (ESI$^+$) 277, 279 ([M+H]$^+$; ^{35}Cl, 100%; ^{37}Cl, 30%), 321, 323 ([M+H]$^+$; ^{79}Br, 100%; ^{81}Br, 95%).

(*E*)-4^{iii}-Iodo-*N*-(3-methyl-4-phenylisoxazol-5-yl)but-2^{iii}-enamide (4.28)



Sodium iodide (105 mg, 0.701 mmol, 1.5 eq) was added to a solution of (*E*)-4^{iii}-bromo/chloro-*N*-(3-methyl-4-phenylisoxazol-5-yl)but-2^{iii}-enamide (**4.25**) (150 mg, 0.467 mmol, 1.0 eq) in acetone (4.0 mL). The reaction was heated at 50 °C for 40 min, after which it was diluted with acetone (10 mL), filtered and concentrated *in vacuo* to isolate the title compound as a brown oil. This intermediate was taken directly on to the next step without further purification; R_f 0.29 (60% Et_2O/petroleum ether); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 7.57 (1H, br s, 6-N*H*), 7.46–7.41 (2H, m, 3^{ii}-C*H*), 7.39–7.35 (1H, m, 4^{iii}-C*H*), 7.30–7.27 (2H, m, 2^{ii}-C*H*), 7.09 (1H, dt, *J* 14.9, 8.3, 3^{iii}-C*H*), 6.13 (1H, d, *J* 14.9, 2^{iii}-C*H*), 3.92 (2H, dd, *J* 8.3, 1.0, 4^{iii}-CH_2), 2.30 (3H, s, 1^{i}-CH_3); LRMS *m/z* (ESI$^+$) 369 ([M+H]$^+$, 100%).

(*E*)-4^{iii}-(Dimethylamino)-*N*-(3-methyl-4-phenylisoxazol-5-yl)but-2^{iii}-enamide (4.3)



Method 1:

The crude (*E*)-4^{iii}-iodo-*N*-(3-methyl-4-phenylisoxazol-5-yl)but-2^{iii}-enamide (**4.28**) was dissolved in anhydrous DMF (3.0 mL), and K_2CO_3 (86 mg, 0.622 mmol, 2.0 eq) was added. The suspension was cooled to 0 °C and dimethylamine solution (2.0 M in THF, 156 µL, 0.311 mmol, 1.0 eq) was added dropwise. After 1 h the reaction was quenched by the addition of saturated aqueous $NaHCO_3$ solution (10 mL), extracted with EtOAc (3 × 30 mL), washed with brine (30 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel column

chromatography (0–50% 12:2:1 EtOAc: EtOH: aqueous NH_4OH/EtOAc) to isolate the title compound as a colourless oil (27 mg, 31% over two steps), the data obtained are identical to those shown below.

Method 2:

(Dimethylamino)acetaldehyde diethyl acetal (1.00 mL, 5.47 mmol, 9.6 eq) was added dropwise to aqueous HCl (6.0 M, 5 mL) at 0 °C. The solution was stirred at 0 °C for 15 min before being heated under reflux for 4 h, after which it was cooled and concentrated *in vacuo*. The residue was then resuspended in THF (2 mL).

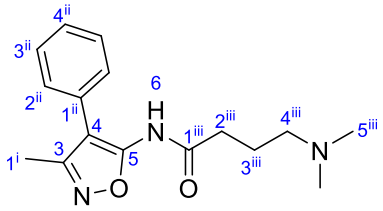
In a separate flask, aqueous NaOH (50% *w/v*, 1.1 mL) was added to a solution of diethyl {2iii-[(3-methyl-4-phenylisoxazol-5-yl)amino]-2iii-oxoethyl}phosphonate (**4.21**) (200 mg, 0.568 mmol, 1.0 eq), and LiCl (24 mg, 0.568 mmol, 1.0 eq) in THF (5 mL) at 0 °C. The first suspension was then added dropwise at 0 °C, slowly warmed to rt, and stirred for 26 h. The reaction was then concentrated *in vacuo*, and extracted with EtOAc (3 × 30 mL). The organic components were combined, washed with brine (30 mL), dried over Na_2SO4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel chromatography (2–10% MeOH/CH_2Cl_2) to isolate the title compound as a colourless oil (27 mg, 17%), the data obtained are identical to those shown below.

Method 3:

Using a modified procedure from Chew *et al*.,[166] oxalyl chloride (284 µL, 3.35 mmol, 1.95 eq) was added dropwise to a suspension of (*E*)-4-(dimethylamino)but-2-enoic acid hydrochloride (570 mg, 3.44 mmol, 2.0 eq) in anhydrous THF (4.0 mL), and anhydrous DMF (2 drops) at 0 °C. The suspension was stirred at 0 °C for a further 20 min and then warmed to rt and stirred for 60 min. The solution was then cooled back down to 0 °C. 3-Methyl-4-phenylisoxazol-5-amine (**4.10**) (300 mg, 1.72 mmol, 1.0 eq) in anhydrous *N*-methyl-2-pyrrolidone (2 mL) was then added dropwise at 0 °C, and the

reaction stirred for 21 h at 5 °C. The reaction was quenched by the addition of aqueous HCl (1.0 M, 20 mL), and washed with EtOAc (30 mL). The acidic aqueous extract was basified with aqueous NaOH (2.0 M) until pH 11, extracted with EtOAc (3 × 30 mL), washed with brine (30 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel column chromatography (0–50% 12:2:1 EtOAc: EtOH: aqueous NH_4OH/EtOAc) followed by semi-preparative HPLC (see general chemical methods) to isolate the title compound as the formic acid salt, which was then dissolved in aqueous NaOH (1.0 M), extracted with EtOAc (3 × 20 mL), and concentrated in vacuo to isolate the free base as a colourless oil (164 mg, 33%); R_f 0.21 (50% 12:2:1 EtOAc: EtOH: aqueous NH_4OH/EtOAc); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3211 (N-H, w br), 3035 (C-H, w), 1675 (C=O); ^{1}H NMR (500 MHz, $CDCl_3$) δ_H 7.44–7.40 (2H, m, 3^{ii}-C*H*), 7.37–7.33 (1H, m, 4^{ii}-C*H*), 7.29–7.27 (2H, m, 2^{ii}-C*H*), 6.97 (1H, dt, *J* 15.5, 5.9, 3^{iii}-CH_2), 6.14 (1H, d, *J* 15.5, 2^{iii}-C*H*), 3.06 (2H, dd, *J* 5.9, 1.6, 4^{iii}-CH_2), 2.30 (3H, s, 1^{i}-CH_3), 2.22 (6H, s, 5^{iii}-CH_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ_C 163.5 (1^{iii}-C), 160.3 (3-C), 156.3 (5-C), 146.0 (3^{iii}-C), 129.2 (3^{ii}-C), 129.1 (1^{ii}-C), 128.8 (2^{ii}-C) 128.2 (4^{ii}-C), 122.9 (2^{ii}-C), 108.9 (4-C), 60.4 (4^{ii}-C), 45.7 (5^{ii}-C), 11.6 (1^{i}-C); HRMS *m/z* (ESI$^+$) [Found: 286.1548, $C_{16}H_{20}O_2N_3$ ([M+H]$^+$) requires 285.1561]; LRMS *m/z* (ESI$^+$) 286 ([M+H]$^+$, 100%); HPLC Retention time 6.2 min, >99.0%.

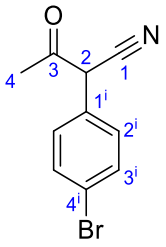
4^{iii}-(Dimethylamino)-*N*-(3-methyl-4-phenylisoxazol-5-yl)butanamide (4.29)



Using a modified procedure from Chew *et al*.,[166] oxalyl chloride (284 µL, 3.36 mmol, 1.95 eq) was added dropwise to a suspension of 4-(dimethylamino)butyric acid hydrochloride (577 mg, 3.44 mmol, 2.0 eq) in anhydrous THF (3.5 mL), and anhydrous DMF (2 drops) at 0 °C. The suspension was stirred at 0 °C for a further 30 min and then warmed to rt and stirred for 30 min. The solution

was then cooled back down to 0 °C. 3-Methyl-4-phenylisoxazol-5-amine (**4.10**) (300 mg, 1.72 mmol, 1.0 eq) in anhydrous *N*-methyl-2-pyrrolidone (2 mL) was then added dropwise at 0 °C, and the reaction stirred for 20 h at 5 °C. The reaction was quenched by the addition of aqueous HCl (1.0 M, 30 mL), and washed with EtOAc (30 mL). The acidic aqueous extract was basified with aqueous NaOH (2.0 M) until pH 11, extracted with EtOAc (3 × 30 mL), washed with brine (20 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel column chromatography (0–60% 12:2:1 EtOAc: EtOH: aqueous NH_4OH/EtOAc) to isolate the title compound as a pale-yellow oil (295 mg, 60%); R_f 0.20 (50% 12:2:1 EtOAc: EtOH: aqueous NH_4OH/EtOAc); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3197 (N-H, w), 3049 (C-H, m), 1661 (C=O, s); ^{1}H NMR (600 MHz, $CDCl_3$) δ_H 7.41–7.38 (2H, m, 3^{ii}-C*H*), 7.33–7.30 (1H, m, 4^{ii}-C*H*), 7.29–7.27 (2H, m, 2^{ii}-C*H*), 2.52–2.48 (2H, m, 2^{iii}-C*H*$_2$), 2.40 (2H, t, *J* 5.7, 4^{iii}-C*H*$_2$), 2.25 (3H, s, 1^{i}-C*H*$_3$), 2.04 (6H, s, 5^{iii}-C*H*$_3$), 1.80–1.76 (2H, m, 3^{ii}-C*H*$_2$); ^{13}C NMR (151 MHz, $CDCl_3$) δ_C 171.7 (1^{iii}-C), 160.1 (3-C), 157.8 (5-C), 130.0 (1^{ii}-C), 129.0 (2^{ii}-C), 128.8 (3^{ii}-C), 127.7 (4^{ii}-C), 108.0 (4-C), 59.6 (4^{iii}-C), 44.5 (5^{iii}-C), 37.2 (2^{iii}-C), 22.0 (3^{iii}-C), 11.5 (1^{i}-C); HRMS *m/z* (ESI$^+$) [Found: 288.1707, $C_{16}H_{22}O_2N_3$ ([M+H]$^+$) requires 288.1707]; LRMS *m/z* (ESI$^+$) 288 ([M+H]$^+$, 100%); HPLC Retention time 6.1 min, >99.0%.

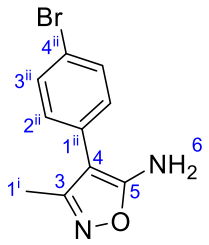
2-(4^{i}-Bromophenyl)-3-oxobutanenitrile (5.11)



Using a modified procedure from Smith *et al.,*[162] sodium hydride (60% dispersion in mineral oil, 816 mg, 20.4 mmol, 2.0 eq) was added portion-wise to a solution of 4-bromophenylacetonitrile (**5.10**) (2.00 g, 10.2 mmol, 1.0 eq), and anhydrous DMF (2 mL) in anhydrous THF (20 mL) at 0 °C. The yellow suspension was stirred for 10 min at 0 °C before the dropwise addition of anhydrous EtOAc (4.00 mL, 40.8 mmol, 4.0 eq) at 0 °C. The suspension was stirred for 18 h at rt. The reaction was

quenched at 0 °C by the addition of aqueous HCl (1.0 M, 15 mL), and extracted with EtOAc (4 × 50 mL). The organic components were combined, washed with aqueous LiCl solution (0.5 M, 50 mL), and brine (50 mL), dried over Na_2SO4, filtered, and concentrated *in vacuo*. The crude brown oil was purified using silica gel chromatography (10–100% EtOAc/petroleum ether) to isolate the title compound as a pale-yellow solid (1.74 g, 72%); R_f 0.31 (50% EtOAc/petroleum ether); m.p 105–107 °C (from EtOAc); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 2218 (C≡N, m), 1639 (C=O, m); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 7.58–7.54 (2H, m, 3^i-C*H*), 7.27–7.24 (2H, m, 2^i-C*H*), 4.62 (1H, s, 2-C*H*), 2.27 (3H, s, 4-C*H*$_3$); ^{13}C NMR (101 MHz, $CDCl_3$) δ_C 195.9 (3-C), 133.0 (3^i-C), 129.7 (2^i-C), 128.7 (1^i-C), 123.9 (4^i-C), 115.9 (1-C), 50.9 (2-C), 27.2 (4-C); HRMS *m/z* (ESI$^-$) [Found: 235.9712, $C_{10}H_7ON^{79}Br$ ([^{79}M-H]$^-$) requires 235.9716, Found: 237.9691, $C_{10}H_7ON^{81}Br$ ([^{81}M-H]$^-$) requires 237.9696]; LRMS *m/z* (ESI$^+$) 238 ([^{79}M+H]$^+$, 100%); HPLC Retention time 9.8 min, >99.0%.

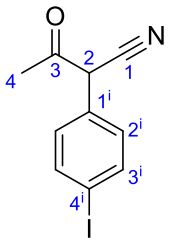
4-(4^{ii}-Bromophenyl)-3-methylisooxazol-5-amine (5.12)



Using a modified procedure from Krasavin *et al.,*[161] aqueous Na_2CO_3 (10% *w/v*, 1 mL), and hydroxylamine hydrochloride (147 mg, 2.11 mmol, 2.0 eq) were added to a solution of 2-(4'-bromophenyl)-3-oxobutanenitrile (**5.11**) (250 mg, 1.06 mmol, 1.0 eq) dissolved in EtOH (4 mL). The solution was heated under reflux for 1 h, then cooled and concentrated *in vacuo*. The crude material was dissolved in H_2O (10 mL) and extracted with CH_2Cl_2 (3 × 20 mL). The organic components were combined, washed with brine (20 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo.* The crude product was purified using silica gel chromatography (5–40% EtOAc/petroleum ether) to isolate the title compound as a pale-yellow solid (220 mg, 83%); R_f 0.61 (50% EtOAc/petroleum ether); m.p 94–96 °C (from $CHCl_3$); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3313 (N-H, m), 3173

(C-H, m), 1638 (C-N, s), 1508 (C-C, m); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 7.56–7.52 (2H, m, 3^{ii}-C*H*), 7.18–7.14 (2H, m, 2^{ii}-C*H*), 4.56 (1H, br s, 6-NH_2), 2.20 (3H, s, 1^{i}-CH_3); ^{13}C NMR (101 MHz, $CDCl_3$) δ_C 164.8 (5-C), 159.6 (3-C), 132.4 (3^{ii}-C), 130.0 (1^{ii}-C), 129.9 (2^{ii}-C), 120.7 (4^{ii}-C), 94.4 (4-C), 11.3 (1^{i}-C); HRMS *m/z* (ESI^{+}) [Found: 252.9972, $C_{10}H_{10}ON_2{}^{79}Br$ ([^{79}M+H]$^{+}$) requires 252.9971, Found: 254.9952, $C_{10}H_{10}ON_2{}^{81}Br$ ([^{81}M+H]$^{+}$) requires 254.9951]; LRMS *m/z* (ESI^{+}) 253 ([^{79}M+H]$^{+}$, 100%), 255 ([^{81}M+H]$^{+}$, 92%); HPLC Retention time 9.5 min, >99.0%.

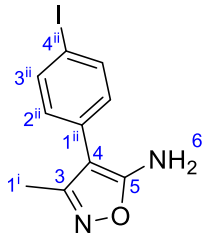
2-(4^{i}-Iodophenyl)-3-oxobutanenitrile (5.16)



Using a modified procedure from Smith *et al.,*[162] sodium hydride (60% dispersion in mineral oil, 757 mg, 18.9 mmol, 1.15 eq) was added portion-wise to a solution of 4-iodophenylacetonitrile (**4.15**) (4.00 g, 16.5 mmol, 1.0 eq), and anhydrous DMF (5.5 mL) in anhydrous THF (49.5 mL) at 0 °C. The dark purple suspension was stirred for 30 min at 0 °C before the dropwise addition of anhydrous EtOAc (6.44 mL, 65.8 mmol, 4.0 eq) at 0 °C. The suspension was stirred for 4 h at rt, after which an extra portion of anhydrous EtOAc (6.44 mL, 65.8 mmol, 4.0 eq) was added dropwise. The reaction was stirred for a further 50 min and was then quenched at 0 °C by the addition of aqueous HCl (1.0 M, 30 mL), and extracted with EtOAc (3 × 100 mL). The organic components were combined, washed with aqueous LiCl solution (0.5 M, 3 × 30 mL), and brine (50 mL), dried over Na_2SO4, filtered, and concentrated *in vacuo*. The crude orange solid was purified using silica gel chromatography (10–50% EtOAc/petroleum ether) to isolate the title compound as a yellow solid (4.06 g, 86%); R_f 0.35 (40% EtOAc/petroleum ether); m.p 113–115 °C (from $CHCl_3$); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 2215 (C≡N, s), 1678 (C=O, m); ^{1}H NMR (400 MHz, methanol-D_4) δ_H 7.68–7.64 (2H, m, 3^{i}-C*H*), 7.48–7.44 (2H, m, 2^{ii}-C*H*), 4.85 (1H, s, 2-C*H*), 2.35 (3H, s, 4-CH_3); HRMS *m/z* (ESI^{-}) [Found:

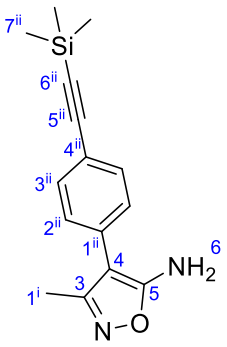
283.9575, $C_{10}H_7ONI$ ([M−H]$^-$) requires 283.9578]; LRMS *m/z* (ESI$^-$) 284 ([M−H]$^-$, 100%); HPLC Retention time 10.1 min, >99.0%.

4-(4^{ii}-Iodophenyl)-3-methylisoxazol-5-amine (5.14)



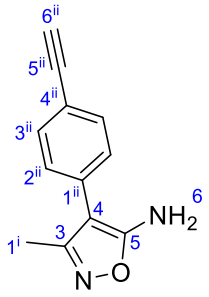
Using a modified procedure from Krasavin *et al.,*[161] aqueous Na_2CO_3 (10% *w/v*, 5 mL), and hydroxylamine hydrochloride (975 mg, 14.0 mmol, 2.0 eq) were added to a solution of 2-(4'-iodophenyl)-3-oxobutanenitrile (**5.16**) (2.00 g, 7.02 mmol, 1.0 eq) dissolved in EtOH (20 mL). The solution was heated at 60 °C for 75 min, then cooled and concentrated *in vacuo*. The crude material was dissolved in H_2O (30 mL) and extracted with EtOAc (3 × 50 mL). The organic components were combined, washed with brine (30 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo* to isolate the title compound as a pale-yellow solid. No further purification was required (2.10 g, 99%); R_f 0.55 (50% EtOAc/petroleum ether); m.p 106–108 °C (from $CHCl_3$); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3310 (N-H, m), 3174 (C-H, m), 1637 (C-N, s), 1506 (C-C, m); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 7.76–7.72 (2H, m, 3^{ii}-C*H*), 7.05–7.01 (2H, m, 2^{ii}-C*H*), 4.57 (1H, br s, 6-N*H*$_2$), 2.20 (3H, s, 1^{i}-C*H*$_3$); ^{13}C NMR (101 MHz, $CDCl_3$) δ_C 164.7 (5-C), 159.6 (3-C), 138.4 (3^{ii}-C), 130.6 (1^{ii}-C), 130.1 (2^{ii}-C), 94.4 (4-C), 91.9 (4^{ii}-C), 11.3 (1^{i}-C); HRMS *m/z* (ESI$^+$) [Found: 300.9833, $C_{10}H_{10}ON_2I$ ([M+H]$^+$) requires 300.9832]; LRMS *m/z* (ESI$^+$) 301 ([M+H]$^+$, 100%); HPLC Retention time 9.8 min, >99.0%

3-Methyl-4-{4^{ii}-[(trimethylsilyl)ethynyl]phenyl}isoxazol-5-amine (5.13)



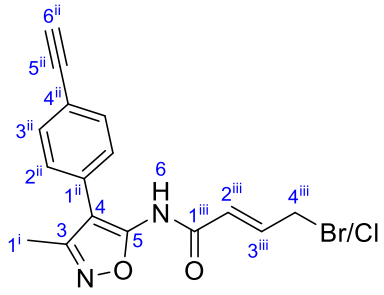
4-(4^{ii}-Iodophenyl)-3-methylisoxazol-5-amine (**5.14**) (1.00 g, 3.33 mmol, 1.0 eq), $Pd(PPh_3)_2Cl_2$ (140 mg, 0.20 mmol, 0.06 mmol), and CuI (76 mg, 0.40 mmol, 0.12 eq) were added to a microwave vial, and the vial sealed and purged with Ar. Degassed and anhydrous triethylamine (4.0 mL) and DMF (1.0 mL) were added, followed by ethynyltrimethylsilane (1.39 mL, 10.0 mmol, 3.0 eq). The reaction was stirred for 1 h at rt, after which a constant stream of N_2 was applied to remove the triethylamine, and the residue filtered through a pad of Celite® which was washed with Et_2O (50 mL). The filtrate was concentrated *in vacuo*, and the crude brown oil purified using silica gel chromatography (6–35% EtOAc/petroleum ether) to isolate the title compound as a pale-orange solid (666 mg, 74%); R_f 0.63 (50% Et_2O/toluene); m.p 134–135 °C (from EtOAc); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3299 (N-H, w), 2155 (C≡C), 1641 (C-N, s), 1520 (C-C, m); ^{1}H NMR (600 MHz, $CDCl_3$) δ_H 7.51–7.49 (2H, m, 3^{ii}-C*H*), 7.23–7.21 (2H, m, 2^{ii}-C*H*), 4.59 (1H, br s, 6-N*H*$_2$), 2.22 (3H, s, 1^{i}-C*H*$_3$), 0.26 (9H, s, 7^{ii}-C*H*$_3$); ^{13}C NMR (151 MHz, $CDCl_3$) δ_C 164.8 (5-C), 159.7 (3-C), 132.8 (3^{ii}-C), 131.4 (1^{ii}-C), 127.8 (2^{ii}-C), 121.5 (4^{ii}-C), 104.8 (5^{ii}-C), 95.0 (6^{ii}-C), 94.9 (4-C), 11.4 (1^{i}-C), 0.1 (7^{ii}-C); HRMS *m/z* (ESI$^+$) [Found: 271.1262, $C_{15}H_{19}ON_2Si$ ([M+H]$^+$) requires 271.1261]; LRMS *m/z* (ESI$^+$) 271 ([M+H]$^+$, 100%); HPLC Retention time 11.9, >99.0%.

4-(4^{ii}-Ethynylphenyl)-3-methylisoxazol-5-amine (5.9)



K_2CO_3 (309 mg, 2.24 mmol, 1.1 eq) was added portion-wise to a solution of 3-ethyl-4-{4^{ii} [(trimethylsilyl)ethynyl]phenyl}isoxazol-5-amine (**5.13**) (550 mg, 2.03 mmol, 1.0 eq) in MeOH (20 mL), stirred at rt for 50 min, and then concentrated *in vacuo*. The crude brown solid was purified using silica gel chromatography (5–30% EtOAc/petroleum ether) to isolate the title compound as a colourless solid (372 mg, 92%); R_f 0.24 (25% EtOAc/petroleum ether); m.p 137–138 °C (from EtOAc); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3307 (N-H, m), 3255 (C-H, s), 1638 (C-N, s), 1511 (C-C, m); ^{1}H NMR (600 MHz, $CDCl_3$) δ_H 7.54–7.52 (2H, m, 3^{ii}-C*H*), 7.26–7.24 (2H, m, 2^{ii}-C*H*), 4.61 (1H, br s, 6-N*H*$_2$), 3.11 (1H, s, 6^{ii}-C*H*), 2.22 (3H, s, 1^{i}-C*H*$_3$); ^{13}C NMR (151 MHz, $CDCl_3$) δ_C 164.9 (5-C), 159.7 (3-C), 133.0 (3^{ii}-C), 131.8 (1^{ii}-C), 127.9 (2^{ii}-C), 120.4 (4^{ii}-C), 94.8 (4-C), 83.4 (5^{ii}-C), 77.8 (6^{ii}-C), 11.4 (1^{i}-C); HRMS *m/z* (ESI$^+$) [Found: 199.0866, $C_{12}H_{11}ON_2$ ([M+H]$^+$) requires 199.0866]; LRMS *m/z* (ESI$^+$) 199 ([M+H]$^+$, 100%); HPLC Retention time 8.8 min, >99.0%.

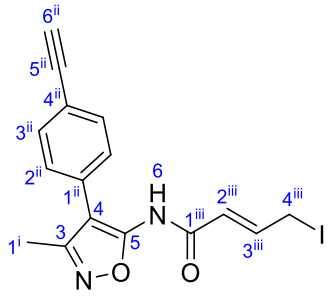
(*E*)-4^{iii}-Bromo/Chloro-*N*-[4-(4^{ii}-ethynylphenyl)-3-methylisoxazol-5-yl]but-2^{iii}-enamide (5.18)



Oxalyl chloride (238 µL, 2.82 mmol, 1.5 eq) was added dropwise to a solution of (*E*)-4-bromocrotonic acid (403 mg, 2.44 mmol, 1.3 eq) in anhydrous CH_2Cl_2 (5 mL), and anhydrous DMF (1 drop) at 0 °C. The solution was stirred at 0 °C for a further 30 min and then warmed to rt and stirred for 15 h. The solution was then concentrated *in vacuo*, and the pale-brown oil redissolved in anhydrous CH_2Cl_2 (5.6 mL).

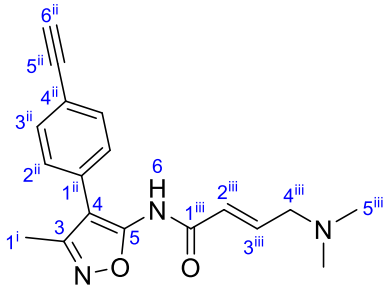
The acyl chloride solution was then added dropwise to a solution of 4-(4^{ii}-ethynylphenyl)-3-methylisoxazol-5-amine (**5.9**) (372 mg, 1.88 mmol, 1.0 eq) and pyridine (228 µL, 2.85 mmol, 1.3 eq) in anhydrous CH_2Cl_2 (9 mL), at 0 °C. After 5 h the reaction was quenched with HCl (1.0 M, 20 mL), and extracted with CH_2Cl_2 (3 × 50 mL). The organic extracts were combined, washed with saturated aqueous $NaHCO_3$ solution (20 mL), and brine (20 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel chromatography (25–60% Et_2O/petroleum ether) to isolate a mixture of the bromo- or chloro-substituted title compound as a pale-yellow solid (235 mg, 36%); R_f 0.25 (50% Et_2O/petroleum ether); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 7.61 (1H, s, 6-N*H*), 7.56–7.53 (2H, m, 3^{ii}-C*H*), 7.26–7.23 (2H, m, 2^{ii}-C*H*), 7.08–7.00 (2H, m, 3^{iii}-C*H*), 6.27 (0.2H, d, *J* 15.0, 2^{iii}-C*H*$_{(Cl)}$), 6.20 (0.8H, d, *J* 14.8, 2^{iii}-C*H*$_{(Br)}$), 4.18 (0.4H, dd, *J* 5.5, 1.7, 4^{iii}-C*H*$_{2(Cl)}$), 4.00 (1.6H, dd, *J* 7.1, 1.3, 4^{iii}-C*H*$_{2(Br)}$), 3.13 (1H, s, 6^{ii}-C*H*), 2.30 (3H, s, 1^{i}-C*H*$_3$); LRMS *m/z* (ESI$^+$) 301, 303 ([M+H]$^+$; ^{35}Cl, 100%; ^{37}Cl, 30%), 345, 347 ([M+H]$^+$; ^{79}Br, 100%; ^{81}Br, 95%).

(*E*)-*N*-[4-(4^{ii}-Ethynylphenyl)-3-methylisoxazol-5-yl]-4^{iii}-iodobut-2^{iii}-enamide (5.19)



Sodium iodide (154 mg, 1.03 mmol, 1.5 eq) was added to a solution of (*E*)-4^{iii}-bromo/chloro-*N*-[4-(4^{ii}-ethynylphenyl)-3-methylisoxazol-5-yl]but-2^{iii}-enamide (**5.18**) (235 mg, 0.683 mmol, 1.0 eq) in acetone (4 mL). The reaction was heated at 50 °C for 1 h, after which it was diluted with acetone (10 mL), filtered and concentrated *in vacuo* to isolate the title compound as a yellow solid. This intermediate was taken directly on to the next step without further purification. R_f 0.25 (50% Et_2O/petroleum ether); LRMS *m/z* (ESI$^+$) 393 ([M+H]$^+$, 100%).

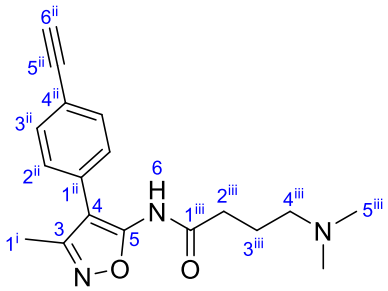
(*E*)-4^{iii}-(Dimethylamino)-*N*-[4-(4^{ii}-ethynylphenyl)-3-methylisoxazol-5-yl]but-2^{iii}-enamide (5.1)



The crude (*E*)-*N*-[4-(4^{ii}-ethynylphenyl)-3-methylisoxazol-5-yl]-4^{iii}-iodobut-2^{iii}-enamide (**5.19**) was dissolved in anhydrous DMF (4 mL), and K_2CO_3 (189 mg, 1.37 mmol, 2.0 eq) was added. The suspension was cooled to 0 °C and dimethylamine solution (2.0 M in THF, 376 µL, 0.751 mmol, 1.1 eq) was added dropwise. After 1 h at 0 °C a further portion of dimethylamine solution (2.0 M in THF, 68 µL, 0.137 mmol, 0.2 eq) was added dropwise. The reaction was stirred for a further 30 min before it was quenched by the addition of aqueous HCl (1.0 M, 30 mL), and washed with EtOAc (30 mL). The acidic aqueous extract was basified with aqueous NaOH (2.0 M) until pH 11, extracted

with EtOAc (3 × 30 mL), washed with brine (30 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel column chromatography (0–50% 12:2:1 EtOAc: EtOH: aqueous NH_4OH/EtOAc) to isolate the title compound as a pale-yellow solid (60 mg, 28% over two steps); R_f 0.33 (10% MeOH/CH_2Cl_2); m.p 132–134 °C (from $CHCl_3$); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3276 (N-H, w), 3184 (C-H, w), 1670 (C=O, m), 1646 (C=C, m); ^{1}H NMR (500 MHz, $CDCl_3$) δ_H 7.54–7.52 (2H, m, 3^{ii}-C*H*), 7.26–7.24 (2H, m, 2^{ii}-C*H*), 6.96 (1H, dt, *J* 15.3, 5.8, 3^{iii}-C*H*), 6.12 (1H, dt, *J* 15.3, 1.6, 2^{iii}-C*H*), 3.12 (1H, s, 6^{ii}-C*H*), 3.06 (2H, dd, *J* 5.8, 1.6, 4^{iii}-CH_2), 2.29 (3H, s, 1^{i}-C*H*), 2.22 (6H, s, 5^{iii}-CH_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ_C 163.3 (1^{iii}-C), 160.2 (3-C), 156.5 (5-C), 146.2 (3^{iii}-C), 132.8 (3^{ii}-C), 129.8 (1^{ii}-C), 128.7 (2^{ii}-C), 122.9 (2^{iii}-C), 121.9 (4^{iii}-C), 108.5 (4-C), 83.2 (5^{ii}-C), 78.4 (6^{ii}-C), 60.4 (4^{iii}-C), 45.7 (5^{iii}-C), 11.6 (1^{i}-C); HRMS *m/z* (ESI^+) [Found: 310.1550, $C_{18}H_{20}O_2N_3$ ($[M+H]^+$) requires 310.1550]; LRMS *m/z* (ESI^+) 310 ($[M+H]^+$, 100%); HPLC Retention time 6.7 min, >99.0%.

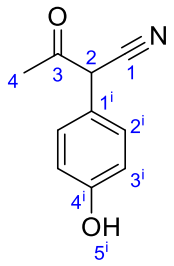
4^{iii}-(Dimethylamino)-*N*-[4-(4^{ii}-ethynylphenyl)-3-methylisoxazol-5-yl]butanamide (5.4)



Using a modified procedure from Chew *et al.*,[166] oxalyl chloride (72 µL, 0.854 mmol, 1.95 eq) was added dropwise to a suspension of 4-(dimethylamino)butyric acid hydrochloride (147 mg, 0.876 mmol, 2.0 eq) in anhydrous THF (2 mL), and anhydrous DMF (1 drop) at 0 °C. The suspension was stirred at 0 °C for 15 min and then warmed to rt and stirred for a further 40 min. The solution was then cooled back down to 0 °C. 3-Methyl-4-{4^{ii}-[(prop-2^{iii}-yn-1^{iii}-yl)oxy]phenyl}-isoxazol-5-amine (**5.9**) (100 mg, 0.438 mmol, 1.0 eq) in anhydrous *N*-methyl-2-pyrrolidone (2 mL) was then added dropwise at 0 °C, and the reaction stirred for 16 h at 5 °C. The reaction was quenched by the addition of aqueous HCl (1.0 M, 30 mL), and washed with EtOAc (2 × 20 mL). The acidic aqueous

extract was basified with aqueous NaOH (2.0 M) until pH 11, extracted with EtOAc (3 × 30 mL), washed with aqueous LiCl solution (0.5 M, 2 × 20 mL), and brine (20 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel column chromatography (0–100% 12:2:1 EtOAc: EtOH: aqueous NH_4OH/EtOAc) to isolate the title compound as an off-white solid (127 mg, 81%); R_f 0.29 (70% 12:2:1 EtOAc: EtOH: aqueous NH_4OH/EtOAc); m.p 114–116 °C (from $CHCl_3$); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3254 (N-H, w), 3037 (C-H, w), 1672 (C=O, s); ^{1}H NMR (600 MHz, $CDCl_3$) δ_H 7.52–7.50 (2H, m, 3^{ii}-C*H*), 7.25–7.23 (2H, m, 2^{ii}-C*H*), 3.10 (1H, s, 6^{ii}-C*H*), 2.50–2.46 (2H, m, 2^{iii}-CH_2), 2.42 (2H, t, *J* 5.7, 4^{iii}-CH_2), 2.29 (3H, s, 1^{i}-C*H*), 2.22 (6H, s, 5^{iii}-CH_3), 1.80–1.75 (2H, m, 3^{iii}-CH_2); ^{13}C NMR (151 MHz, $CDCl_3$) δ_C 171.5 (1^{iii}-C), 160.0 (3-C), 158.1 (5-C), 132.5 (3^{ii}-C), 130.9 (1^{ii}-C), 128.8 (2^{ii}-C), 121.3 (4^{ii}-C), 107.3 (4-C), 83.4 (5^{ii}-C), 78.0 (6^{ii}-C), 59.7 (4^{iii}-C), 44.6 (5^{iii}-C), 37.4 (2^{iii}-C), 22.0 (3^{iii}-C), 11.5 (1^{i}-C); HRMS *m/z* (ESI^+) [Found: 312.1706, $C_{18}H_{22}O_2N_3$ ($[M+H]^+$) requires 310.1707]; LRMS *m/z* (ESI^+) 312 ($[M+H]^+$, 100%); HPLC Retention time 6.7 min, >99.0%.

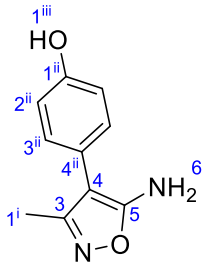
2-(4^{i}-Hydroxyphenyl)-3-oxobutanenitrile (5.21)



Using a modified procedure from Smith *et al.*,[162] sodium hydride (60% dispersion in mineral oil, 1.32 g, 33.0 mmol, 2.2 eq) was added portion-wise to a solution of 4-hydroxyphenylacetonitrile (**5.20**) (2.00 g, 15.0 mmol, 1.0 eq), and anhydrous DMF (4 mL) in anhydrous THF (36 mL) at 0 °C. The suspension was stirred for 10 min at 0 °C before the dropwise addition of anhydrous EtOAc (5.88 mL, 60.1 mmol, 4.0 eq) at 0 °C. The suspension was stirred for 16 h at rt before being quenched at 0 °C by the addition of aqueous HCl (1.0 M, 30 mL), and the THF removed *in vacuo*. The solution was extracted with EtOAc (3 × 50 mL) and the organic components combined, washed

with aqueous LiCl solution (0.5 M, 30 mL), and brine (50 mL), dried over Na_2SO4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel chromatography (20–45% EtOAc/petroleum ether) to isolate the title compound as a pale-yellow solid (2.10 g); R_f 0.56 (100% EtOAc); m.p 74–76 °C (from EtOAc); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3132 (O-H, s br), 2226 (C≡N, m), 1639 (C=O, m); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 7.28–7.24 (2H, m, 2^i-C*H*), 6.91–6.87 (2H, m, 3^i-C*H*), 5.27 (1H, br s, 5^i-O*H*), 4.60 (1H, s, 2-C*H*), 2.25 (3H, s, 4-C*H*$_3$); ^{13}C NMR (101 MHz, $CDCl_3$) δ_C 196.7 (3-C), 156.7 (4^i-C), 129.6 (2^i-C), 121.8 (1^i-C), 116.8 (3^i-C), 116.5 (1-C), 50.9 (2-C), 26.9 (4-C); HRMS *m/z* (ESI$^-$) [Found: 174.0552, $C_{10}H_8O_2N$ ([M−H]$^-$) requires 174.0550]; LRMS *m/z* (ESI$^-$) 174 ([M−H]$^-$, 100%); HPLC Retention time 6.6 min, >99.0%.

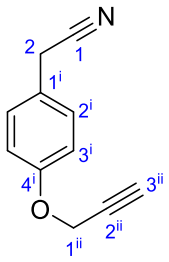
4^{ii}-(5-Amino-3-methylisoxazol-4-yl)phenol (5.22)



Using a modified procedure from Krasavin *et al.*,[161] aqueous Na_2CO_3 (10% *w/v*, 3.0 mL), and hydroxylamine hydrochloride (794 mg, 11.4 mmol, 2.0 eq) were added to a solution of 2-(4^i-hydroxyphenyl)-3-oxobutanenitrile (**5.21**) (1.00 g, 5.71 mmol, 1.0 eq) dissolved in EtOH (14 mL). The solution was heated at 60 °C for 1.5 h, then cooled and concentrated *in vacuo*. The crude material was dissolved in H_2O (30 mL) and extracted with EtOAc (3 × 50 mL). The organic components were combined, washed with brine (30 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo* to isolate the title compound as a pale-yellow solid. No further purification was required (1.08 g, 99%); R_f 0.46 (100% EtOAc); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3331 (O-H, m br), 3186 (N-H, w), 2920 (C-H, s); ^{1}H NMR (400 MHz, DMSO-D_6) δ_H 9.33 (1H, s, 1^{iii}-O*H*), 7.12–7.08 (2H, m, 3^{ii}-C*H*), 6.81–6.77 (2H, m, 2^{ii}-C*H*), 6.43 (2H, br s, 6-N*H*$_2$), 2.06 (3H, s, 1^i-C*H*$_3$); ^{13}C NMR (101 MHz, DMSO-D_6)

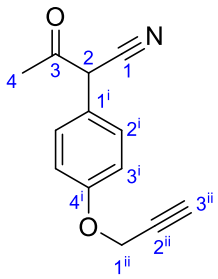
δ_C 166.8 (1^{ii}-C), 158.6 (3-C), 155.5 (5-C), 129.3 (3^{ii}-C), 121.6 (4^{ii}-C), 115.5 (2^{ii}-C), 91.9 (4-C), 11.0 (1^{i}-C); HRMS *m/z* (ESI^+) [Found: 191.0817, $C_{10}H_{11}O_2N_2$ ($[M+H]^+$) requires 191.0815]; LRMS *m/z* (ESI^+) 191 ($[M+H]^+$, 100%); HPLC Retention time 6.1 min, 98.9%.

2-(4^{i}-Prop-2^{ii}-ynoxyphenyl)acetonitrile (5.24)



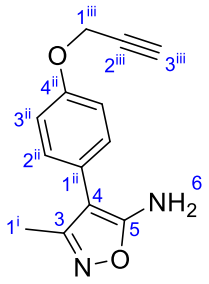
K_2CO_3 (10.4 g, 75.1 mmol, 4.0 eq) was added to a solution of 4-hydroxyphenylacetonitrile (**5.20**) (2.5 g, 18.8 mmol, 1.0 eq) in anhydrous DMF (50 mL). The suspension was stirred at rt for 30 min before being cooled to 0 °C. Propargyl bromide solution (80 wt. % in toluene, 3.14 mL, 28.2 mmol, 1.5 eq) was then added dropwise to the suspension at 0 °C, and stirred for 15 min before being raised to rt. After 2.5 h the reaction was concentrated *in vacuo*, resuspended in aqueous NaOH (1.0M, 50 mL), and extracted with EtOAc (3 × 75 mL). The organic extracts were combined, washed with aqueous LiCl solution (0.5 M, 3 × 50 mL), and brine (50 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo* to isolate the title compound as a yellow solid. No further purification was required (3.21 g, 99%); R_f 0.62 (40% EtOAc/petroleum ether); m.p 47–48 °C (from $CHCl_3$); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3289 (C-H, m), 2251 (C≡N, w), 2121 (C≡C, w); ^{1}H NMR (600 MHz, $CDCl_3$) δ_H 7.17–7.13 (2H, m, 2^{i}-C*H*), 6.89–6.86 (2H, m, 3^{i}-C*H*), 4.59 (2H, d, *J* 2.4, 1^{ii}-CH_2), 3.58 (2H, s, 2-CH_2), 2.43 (1H, t, *J* 2.4, 3^{ii}-C*H*); ^{13}C NMR (151 MHz, $CDCl_3$) δ_C 157.3 (4^{i}-C), 129.2 (2^{i}-C), 122.9 (1^{i}-C), 118.2 (1-C), 115.6 (3^{i}-C), 78.3 (2^{ii}-C), 75.9 (3^{ii}-C), 55.9 (1^{ii}-C), 22.9 (2-C); LRMS *m/z* (ESI^-) 170 ($[M-H]^-$, 100%); HPLC Retention time 9.0 min, >99.0%.

3-Oxo-2-(4^{i}-prop-2^{ii}-ynoxyphenyl)butanenitrile (5.25)



Using a modified procedure from Smith *et al.,*[162] sodium hydride (60% dispersion in mineral oil, 257 mg, 6.43 mmol, 1.1 eq) was added portion-wise to a solution of 2-(4^{i}-prop-2^{ii}-ynoxyphenyl)acetonitrile (**5.24**) (1.00 g, 5.84 mmol, 1.0 eq), and anhydrous DMF (1.6 mL) in anhydrous THF (14.4 mL) at 0 °C. The suspension was stirred for 30 min at 0 °C before the dropwise addition of anhydrous EtOAc (4.57 mL, 46.7 mmol, 9.0 eq) at 0 °C. The suspension was stirred for 3.5 h at rt. The reaction was quenched at 0 °C by the addition of aqueous HCl (1.0 M, 20 mL), and the THF removed *in vacuo*. The solution was extracted with EtOAc (3 × 30 mL) and the organic components combined, washed with H_2O (50 mL), and brine (50 mL), dried over Na_2SO4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel chromatography (30–80% Et_2O/petroleum ether) to isolate the title compound as a pale-yellow solid (1.14 g, 91%); R_f 0.31 (40% EtOAc/petroleum ether); m.p 54–56 °C (from $CHCl_3$); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3289 (C-H, m), 2208 (C≡N, w), 2123 (C≡C, w), 1728 (C=O, m); ^{1}H NMR (600 MHz, $CDCl_3$) δ_H 7.32–7.29 (2H, m, 2^{i}-C*H*), 7.03–7.00 (2H, m, 3^{i}-C*H*), 4.70 (2H, d, *J* 2.4, 1^{ii}-C*H*$_2$), 4.64 (1H, br s, 2-C*H*), 2.54 (1H, t, *J* 2.4, 3^{ii}-C*H*), 2.23 (3H, s, 4-C*H*$_3$); ^{13}C NMR (151 MHz, $CDCl_3$) δ_C 196.8 (3-C), 158.3 (4^{i}-C), 129.4 (2^{i}-C), 122.6 (1^{i}-C), 116.4 (1-C), 116.1 (3^{i}-C), 78.1 (2^{ii}-C), 76.2 (3^{ii}-C), 56.0 (1^{ii}-C), 50.7 (2-C), 27.0 (4-C); HRMS *m/z* (ESI^{+}) [Found: 212.0710, $C_{13}H_{10}O_2N$ ([M+H]$^{+}$) requires 212.0706]; LRMS *m/z* (ESI^{+}) 212 ([M+H]$^{+}$, 100%); HPLC Retention time 8.9 min, >99.0%.

3-Methyl-4-(4^{ii}-prop-2^{iii}-ynoxyphenyl)-isoxazol-5-amine (5.23)



Method 1:

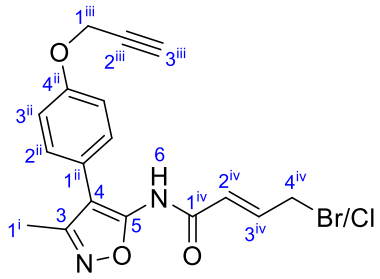
K_2CO_3 (1.60 g, 11.6 mmol, 2.0 eq) was added to a solution of 4^{ii}-(5-amino-3-methylisoxazol-4-yl)phenol (**5.22**) (1.10 g, 5.79 mmol, 1.0 eq) in anhydrous DMF (15 mL). The bright red suspension was stirred at rt for 10 min before being cooled to 0 °C. Propargyl bromide solution (80 wt. % in toluene, 967 µL, 8.68 mmol, 1.5 eq) was then added dropwise to the suspension at 0 °C, and stirred for 10 min before being raised slowly to rt. The reaction was stirred for 50 h, after which it was quenched by the addition of aqueous LiCl solution (0.5 M, 50 mL), and extracted with EtOAc (3 × 50 mL). The organic extracts were combined, washed with aqueous LiCl solution (0.5 M, 3 × 30 mL), brine (30 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel chromatography (10–50% EtOAc/petroleum ether) to isolate the title compound as a yellow solid (1.15 g, 89%).

Method 2:

Using a modified procedure from Krasavin *et al.,*[161] aqueous Na_2CO_3 (10% *w/v*, 3.0 mL), and hydroxylamine hydrochloride (525 mg, 7.55 mmol, 2.0 eq) were added to a solution of 3-oxo-2-(4^{i}-prop-2^{ii}-ynoxyphenyl)butanenitrile (**5.24**) (805 mg, 3.78 mmol, 1.0 eq) dissolved in EtOH (12 mL). The solution was heated at 80 °C for 1 h, then cooled and concentrated *in vacuo*. The crude material was dissolved in H_2O (30 mL) and extracted with EtOAc (3 × 30 mL). The organic components were combined, washed with brine (30 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The

crude product was purified using silica gel chromatography (0–5% EtOAc/CH_2Cl_2) to isolate the title compound as a yellow solid (711 mg, 83%); R_f 0.51 (50% EtOAc/petroleum ether); m.p 57–59 °C (from EtOAc); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3359 (N-H, w), 2121 (C≡C, w), 1642 (C-N, s), 1516 (C-C, m); ^{1}H NMR (500 MHz, $CDCl_3$) δ_H 7.24–7.21 (2H, m, 2^{ii}-C*H*), 7.06–7.03 (2H, m, 3^{ii}-C*H*), 4.72 (2H, d, *J* 2.4, 1^{iii}-CH_2), 4.44 (2H, br s, 6-NH_2), 2.54 (1H, t, *J* 2.4, 3^{iii}-C*H*), 2.20 (3H, s, 1^{i}-CH_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ_C 164.6 (5-C), 160.0 (3-C), 156.6 (4^{ii}-C), 129.7 (2^{ii}-C), 124.1 (1^{ii}-C), 115.8 (3^{ii}-C), 95.0 (4-C), 78.6 (2^{iii}-C), 75.9 (3^{iii}-C), 56.0 (1^{iii}-C), 11.3 (1^{i}-C); HRMS *m/z* (ESI^+) [Found: 229.0973 $C_{13}H_{13}O_2N_2$ ($[M+H]^+$) requires 229.0983]; LRMS *m/z* (ESI^+) 229 ($[M+H]^+$, 100%); HPLC Retention time 8.5 min, >99.0%.

(*E*)-4^{iv}-Bromo/Chloro-*N*-[3-methyl-4-(4^{ii}-prop-2^{iii}-ynoxyphenyl)-isoxazol-5-yl]but-2^{iv}-enamide (5.26)

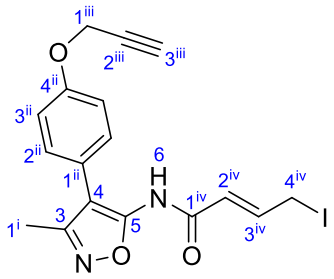


Oxalyl chloride (668 µL, 7.89 mmol, 3.0 eq) was added dropwise to a solution of (*E*)-4-bromocrotonic acid (868 mg, 5.26 mmol, 1.6 eq) in anhydrous CH_2Cl_2 (5 mL), and anhydrous DMF (2 drops) at 0 °C. The solution was stirred at 0 °C for a further 15 min and then warmed to rt and stirred for 3 h. The solution was then concentrated *in vacuo*, and the pale-brown oil redissolved in anhydrous CH_2Cl_2 (2 mL).

The acyl chloride solution (1.1 mL) was then added dropwise to a solution of 3-methyl-4-(4^{ii}-prop-2^{iii}-ynoxyphenyl)-isoxazol-5-amine (**5.23**) (600 mg, 2.63 mmol, 1.0 eq) and pyridine (319 µL, 3.95 mmol, 1.5 eq) in anhydrous CH_2Cl_2 (13 mL), at 0 °C. After 2 h the reaction was quenched with

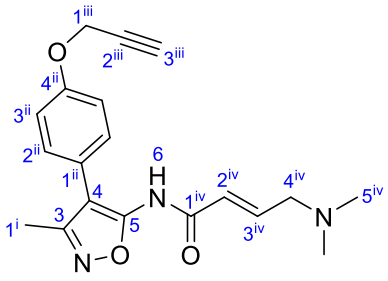
aqueous HCl (1.0M, 20 mL), and extracted with CH_2Cl_2 (3 × 50 mL). The organic extracts were combined, washed with saturated aqueous $NaHCO_3$ solution (20 mL), and brine (20 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel chromatography (40–80% Et_2O/petroleum ether) to isolate a mixture of bromo- or chloro-substituted title compound as a pale-brown oil (350 mg, 36%); R_f 0.19 (60% Et_2O/petroleum ether); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 8.52 (1H, br s, 6-N*H*), 7.20–7.16 (2H, m, 2^{ii}-C*H*), 7.00–6.90 (3H, m, 3^{ii}-C*H*, 3^{iv}-C*H*), 6.26 (0.3H, dt, *J* 15.0, 1.6, 2^{iv}-C$H_{(Cl)}$), 6.20 (0.7H, dt, *J* 15.0, 1.3, 2^{iv}-C$H_{(Br)}$), 4.68 (2H, d, *J* 2.3, 1^{iii}-CH_2), 4.15 (0.6H, dd, *J* 5.7, 1.6, 4^{iv}-C$H_{2(Cl)}$), 3.99 (1.4H, dd, *J* 7.2, 1.3, 4^{iv}-C$H_{2(Br)}$), 2.54 (1H, t, *J* 2.3, 3^{iii}-C*H*), 2.23 (3H, s, 1^{i}-CH_3); LRMS *m/z* (ESI^+) 331, 333 ($[M+H]^+$; ^{35}Cl, 100%; ^{37}Cl, 30%), 375, 377 ($[M+H]^+$; ^{79}Br, 100%; ^{81}Br, 97%).

(*E*)-4^{iv}-Iodo-*N*-[3-methyl-4-(4^{ii}-prop-2^{iii}-ynoxyphenyl)-isoxazol-5-yl]but-2^{iv}-enamide (5.27)



Sodium iodide (120 mg, 0.800 mmol, 1.5 eq) was added to a solution of (*E*)-4^{iv}-bromo/chloro-*N*-[3-methyl-4-(4^{ii}-prop-2^{iii}-ynoxyphenyl)-isoxazol-5-yl]but-2^{iv}-enamide (**5.26**) (200 mg, 0.533 mmol, 1.0 eq) in acetone (5 mL). The reaction was heated at 50 °C for 1 h, after which it was diluted with acetone (10 mL), filtered and concentrated *in vacuo* to isolate the title compound as a yellow oil. This intermediate was taken directly on to the next step without further purification; R_f 0.19 (60% Et_2O/petroleum ether); LRMS *m/z* (ESI^+) 423 ($[M+H]^+$, 100%).

(*E*)-4^{iv}-(Dimethylamino)-*N*-[3-methyl-4-(4^{ii}-prop-2^{iii}-ynoxyphenyl)-isoxazol-5-yl]but-2^{iv}-enamide (5.2)



Method 1:

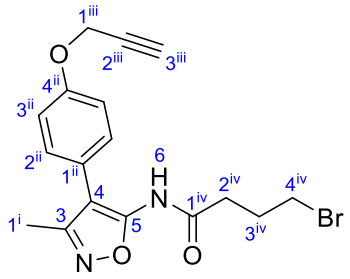
The crude (*E*)-4^{iv}-iodo-*N*-[3-methyl-4-(4^{ii}-prop-2^{iii}-ynoxyphenyl)-isoxazol-5-yl]but-2^{iv}-enamide (**5.27**) was dissolved in anhydrous DMF (5 mL), and K_2CO_3 (147 mg, 1.07 mmol, 2.0 eq) was added. The suspension was cooled to 0 °C and dimethylamine solution (2.0 M in THF, 293 µL, 0.586 mmol, 1.1 eq) was added dropwise. After 1 h the reaction was quenched by the addition of aqueous NaOH (1.0 M, 30 mL), and extracted with EtOAc (3 × 30 mL). The organic extracts were combined, washed with brine (2 × 30 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel column chromatography (1 – 12% EtOH/CH_2Cl_2) followed by another round of silica gel column chromatography (0–50% 12:2:1 EtOAc: EtOH: aqueous NH_4OH/EtOAc) to isolate the title compound as a pale-yellow oil (30 mg, 18% over two steps).

Method 2:

Using a modified procedure from Chew *et al.*,[166] oxalyl chloride (98 µL, 1.16 mmol, 1.95 eq) was added dropwise to a suspension of (*E*)-4-(dimethylamino)but-2-enoic acid hydrochloride (198 mg, 1.19 mmol, 2.0 eq) in anhydrous THF (3 mL), and anhydrous DMF (1 drop) at 0 °C. The suspension was stirred at 0 °C for a further 15 min and then warmed to rt and stirred for 80 min. The solution was then cooled back down to 0 °C. 3-Methyl-4-(4^{ii}-prop-2^{iii}-ynoxyphenyl)-isoxazol-5-amine (**5.23**) (136 mg, 0.596 mmol, 1.0 eq) in anhydrous *N*-methyl-2-pyrrolidone (1.2 mL) was then added dropwise at 0 °C, and the reaction stirred for 22 h at 5 °C. The reaction was quenched by the

addition of aqueous HCl (1.0 M, 30 mL), and washed with EtOAc (30 mL). The acidic aqueous extract was basified with aqueous NaOH (2.0 M) until pH 11, extracted with EtOAc (3 × 30 mL), washed with brine (20 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel column chromatography (0–50% 12:2:1 EtOAc: EtOH: aqueous NH_4OH/EtOAc) followed by semi-preparative HPLC (see general chemical methods) to isolate the title compound as the formic acid salt, which was then dissolved in aqueous NaOH (1.0 M), extracted with EtOAc (3 × 20 mL), and concentrated in vacuo to isolate the free base as a pale-yellow oil (75 mg, 37%); R_f 0.31 (50% 12:2:1 EtOAc: EtOH: aqueous NH_4OH/EtOAc); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3263 (N-H, w), 3036 (C-H, w), 2122 (C≡C, w), 1673 (C=O, s); ^{1}H NMR (600 MHz, $CDCl_3$) δ_H 7.23–7.21 (2H, m, 2^{ii}-C*H*), 7.04–7.01 (2H, m, 3^{ii}-C*H*), 6.97 (1H, dt, *J* 15.4, 5.9, 3^{iv}-C*H*), 6.14 (1H, d, *J* 15.4, 2^{iv}-C*H*), 4.71 (2H, d, *J* 2.4, 1^{iii}-CH_2), 3.06 (2H, dd, *J* 5.9, 1.5, 4^{iv}-CH_2), 2.54 (1H, t, *J* 2.4, 3^{iii}-C*H*), 2.28 (3H, s, 1^{i}-CH_3), 2.22 (6H, s, 5^{iv}-CH_3); ^{13}C NMR (151 MHz, $CDCl_3$) δ_C 163.6 (1^{iv}-C), 160.4 (3-C), 157.5 (4^{ii}-C), 156.1 (5-C), 145.9 (3^{iv}-C), 130.1 (2^{ii}-C), 123.0 (2^{iv}-C), 122.1 (1^{ii}-C), 115.6 (3^{ii}-C), 108.7 (4-C), 78.4 (2^{iii}-C), 76.0 (3^{iii}-C), 60.4 (4^{iv}-C), 56.0 (1^{iii}-C), 45.6 (5^{iv}-C), 11.6 (1^{i}-C); HRMS *m/z* (ESI$^+$) [Found: 340.1656, $C_{19}H_{22}O_3N_3$ ([M+H]$^+$) requires 340.1656]; LRMS *m/z* (ESI$^+$) 340 ([M+H]$^+$, 100%); HPLC Retention time 6.6 min, 96.9%.

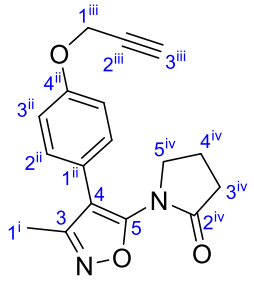
4^{iv}-Bromo-*N*-[3-methyl-4-(4^{ii}-prop-2^{iii}-ynoxyphenyl)-isoxazol-5-yl]butanamide (5.28)



4-Bromobutyryl chloride (61 µL, 0.329 mmol, 1.5 eq) was added dropwise to a solution of 3-methyl-4-(4^{ii}-prop-2^{iii}-ynoxyphenyl)-isoxazol-5-amine (**5.23**) (50 mg, 0.219 mmol, 1.0 eq) and pyridine (20 µL, 0.241 mmol, 1.1 eq) in anhydrous CH_2Cl_2 (2 mL) at 0 °C. The reaction was stirred at 0 °C for

90 min before being quenched by the addition of aqueous HCl (1.0 M, 10 mL), and extracted with EtOAc (3 × 15 mL). The organic extracts were combined, washed with brine (20 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel column chromatography (50–70% Et_2O/petroleum ether) to isolate the title compound as a yellow solid (66 mg, 80%); R_f 0.43 (80% Et_2O/petroleum ether); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 7.76 (1H, br s, 6-N*H*), 7.23–7.20 (2H, m, 2^{ii}-C*H*), 7.05–7.02 (2H, m, 3^{ii}-C*H*), 4.72 (2H, d, *J* 2.4, 1^{iii}-CH_2), 3.42 (2H, t, *J* 6.5, 4^{iv}-CH_2), 2.58–2.52 (3H, m, 3^{iii}-C*H*, 2^{iv}-CH_2), 2.27 (3H, s, 1^{i}-CH_3), 2.16 (2H, quint, *J* 6.5, 3^{iv}-CH_2); LRMS *m/z* (ESI$^+$) 377 ([^{79}M+H]$^+$, 100%), 379 ([^{81}M+H]$^+$, 95%).

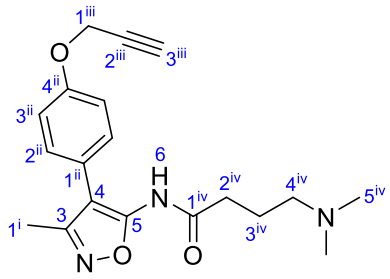
1^{iv}-[3-Methyl-4-(4-prop-2^{iii}-ynoxyphenyl)-isoxazol-5-yl]pyrrolidin-2^{iv}-one (5.29)



TBAI (9 mg, 0.024 mmol, 0.2 eq) was added to a solution of 4^{iv}-bromo-*N*-[3-methyl-4-(4^{ii}-prop-2^{iii}-ynoxyphenyl)-isoxazol-5-yl]butanamide (**5.28**) (45 mg, 0.120 mmol, 1.0 eq) in anhydrous DMF (1.2 mL) at rt. The reaction was stirred at rt for 15 min, after which it was cooled to 0 °C. Dimethylamine solution (2.0 M in THF, 66 µL, 0.132 mmol, 1.1 eq) was added then added dropwise at 0 °C, followed by K_2CO_3 (33 mg, 0.240 mmol, 2.0 eq), and the suspension stirred at 0 °C for 5 min before slowly warming to rt. After 1 h at rt, the reaction was diluted with H_2O (10 mL), and extracted with EtOAC (3 × 20 mL). The organic extracts were combined, washed with brine (2 × 20 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel column chromatography (10–50% EtOAc/petroleum ether) to isolate the title compound as a colourless oil (33 mg, 81%); R_f 0.22 (80% Et_2O/petroleum ether); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3286 (N-H,

m), 2122 (C≡C, w), 1673 (C=O, s); ^{1}H NMR (600 MHz, $CDCl_3$) δ_H 7.23–7.20 (2H, m, 2^{ii}-C*H*), 7.02–6.98 (2H, m, 3^{ii}-C*H*), 4.70 (2H, d, *J* 2.4, 1^{iii}-CH_2), 3.70 (2H, t, *J* 7.2, 5^{iv}-CH_2), 2.54 (1H, t, *J* 2.4, 3^{iii}-C*H*), 2.47 (2H, t, *J* 7.9, 3^{iv}-CH_2), 2.25 (3H, s, 1^{i}-CH_3), 2.18 (2H, tt, *J* 7.2, 7.9, 4^{iv}-CH_2); ^{13}C NMR (151 MHz, $CDCl_3$) δ_C 174.2 (2^{iv}-C), 160.8 (3-C), 157.4 (4^{ii}-C), 156.9 (5-C), 130.1 (2^{ii}-C), 122.3 (1^{ii}-C), 115.2 (3^{ii}-C), 110.8 (4-C), 78.5 (2^{iii}-C), 75.9 (3^{iii}-C), 55.9 (1^{iii}-C), 48.5 (5^{iv}-C), 30.8 (3^{iv}-C), 19.2 (4^{iv}-C), 11.5 (1^{i}-C); HRMS *m/z* (ESI$^+$) [Found: 297.1235, $C_{17}H_{17}O_3N_2$ ([M+H]$^+$) requires 297.1234]; LRMS *m/z* (ESI$^+$) 297 ([M+H]$^+$, 100%); HPLC Retention time 9.0 min, 84.5%.

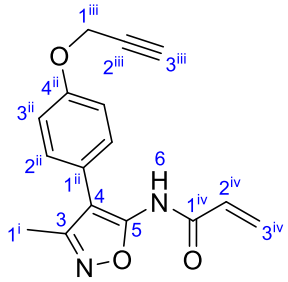
4^{iv}-(Dimethylamino)-*N*-[3-methyl-4-(4^{ii}-prop-2^{iii}-ynoxyphenyl)-isoxazol-5-yl]butanamide (5.5)



Using a modified procedure from Chew *et al.*,[166] oxalyl chloride (72 µL, 0.854 mmol, 1.95 eq) was added dropwise to a suspension of 4-(dimethylamino)butyric acid hydrochloride (147 mg, 0.876 mmol, 2.0 eq) in anhydrous THF (2.0 mL), and anhydrous DMF (1 drop) at 0 °C. The suspension was stirred at 0 °C for 15 min and then warmed to rt and stirred for a further 40 min. The solution was then cooled back down to 0 °C. 3-Methyl-4-(4^{ii}-prop-2^{iii}-ynoxyphenyl)-isoxazol-5-amine (**5.23**) (100 mg, 0.438 mmol, 1.0 eq) in anhydrous *N*-methyl-2-pyrrolidone (2.0 mL) was then added dropwise at 0 °C, and the reaction stirred for 16 h at 5 °C. The reaction was quenched by the addition of aqueous HCl (1.0 M, 30 mL), and washed with EtOAc (2 × 20 mL). The acidic aqueous extract was basified with aqueous NaOH (2.0 M) until pH 11, extracted with EtOAc (3 × 30 mL), washed with aqueous LiCl solution (0.5 M, 2 × 20 mL), and brine (20 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel column chromatography

(0–100% 12:2:1 EtOAc: EtOH: aqueous NH_4OH/EtOAc) followed by semi-preparative HPLC (see general chemical methods) to isolate the title compound as the formic acid salt, which was then dissolved in aqueous NaOH (1.0 M), extracted with EtOAc (3 × 20 mL), and concentrated in vacuo to isolate the free base as a pale-yellow oil (50 mg, 33%); R_f 0.20 (50% 12:2:1 EtOAc: EtOH: aqueous NH_4OH/EtOAc); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3251 (N-H, m), 2118 (C≡C, w), 1675 (C=O, s); ^{1}H NMR (600 MHz, $CDCl_3$) δ_H 7.25–7.22 (2H, m, 2^{ii}-C*H*), 7.03–7.00 (2H, m, 3^{ii}-C*H*), 4.71 (2H, d, *J* 2.4, 1^{iii}-CH_2), 2.55–2.51 (3H, m, 3^{iii}-C*H*, 2^{iv}-CH_2), 2.47 (2H, t, *J* 5.9, 4^{iv}-CH_2), 2.24 (3H, s, 1^{i}-CH_3), 2.13 (6H, br s, 5^{iv}-CH_3), 1.85–1.80 (2H, m, 3^{iv}-CH_2); ^{13}C NMR (151 MHz, $CDCl_3$) δ_C 171.7 (1^{iv}-C), 160.3 (3-C), 157.5 (5-C), 157.2 (4^{ii}-C), 130.3 (2^{ii}-C), 123.1 (1^{ii}-C), 115.3 (3^{ii}-C), 108.0 (4-C), 78.5 (2^{iii}-C), 75.8 (3^{iii}-C), 59.4 (4^{iv}-C), 56.0 (1^{ii}-C), 44.5 (5^{iv}-C), 36.9 (2^{iv}-C), 22.0 (3^{iv}-C), 11.5 (1^{i}-C); HRMS *m/z* (ESI^+) [Found: 342.1814, $C_{19}H_{24}O_3N_3$ ($[M+H]^+$) requires 342.1812]; LRMS *m/z* (ESI^+) 342 ($[M+H]^+$, 100%); HPLC Retention time 6.5 min, >99.0%.

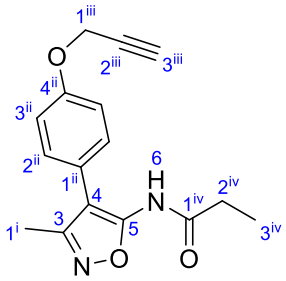
_N_-[3-Methyl-4-(4^{ii}-prop-2^{iii}-ynoxyphenyl)-isoxazol-5-yl]prop-2^{iv}-enamide (5.3)



Acryloyl chloride (108 µL, 1.34 mmol, 1.3 eq) was added dropwise to a solution of 3-methyl-4-(4^{ii}-prop-2^{iii}-ynoxyphenyl)-isoxazol-5-amine (**5.23**) (235 mg, 1.03 mmol, 1.0 eq) and pyridine (125 µL, 1.55 mmol, 1.5 eq) in anhydrous CH_2Cl_2 (5 mL) at 0 °C. The reaction was stirred at 0 °C for 90 min, after which a further portion of acryloyl chloride (108 µL, 1.34 mmol, 1.3 eq) was added dropwise at 0 °C. The reaction was stirred for another 1 h before being quenched by the addition of aqueous HCl (1.0 M, 20 mL), and extracted with CH_2Cl_2 (3 × 30 mL). The organic extracts were combined, washed with H_2O (30 mL) and brine (30 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*.

The crude product was purified using silica gel column chromatography (45–70% Et_2O/petroleum ether) to isolate the title compound as a yellow oil (52 mg, 18%); R_f 0.16 (60% Et_2O/petroleum ether); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3284 (N-H, m), 3030 (C-H, w), 2118 (C≡C, w), 1683 (C=O, m), 1646 (C=C, m); ^{1}H NMR (600 MHz, $CDCl_3$) δ_H 8.21 (1H, br s, 6-N*H*), 7.21–7.17 (2H, m, 2^{ii}-C*H*), 7.01–6.97 (2H, m, 3^{ii}-C*H*), 6.37 (1H, dd, *J* 16.9, 1.0, 3^{iv}-CH_xH_y), 6.26 (1H, dd, *J* 16.9, 10.4, 2^{iv}-C*H*), 5.77 (1H, dd, *J* 10.4, 1.0, 3^{iv}-CH_xH_y), 4.68 (2H, d, *J* 2.3, 1^{iii}-CH_3), 2.53 (1H, t, *J* 2.3, 3^{iii}-C*H*), 2.24 (3H, s, 1^{i}-CH_3); ^{13}C NMR (151 MHz, $CDCl_3$) δ_C 163.8 (1^{iv}-C), 160.4 (3-C), 157.4 (4^{ii}-C), 155.9 (5-C), 130.4 (3^{iv}-C), 130.1 (2^{ii}-C), 128.8 (2^{iv}-C), 122.0 (1^{ii}-C), 115.4 (3^{ii}-C), 109.2 (4-C), 78.4 (2^{iii}-C), 75.9 (3^{iii}-C), 55.9 (1^{iii}-C), 11.5 (1^{i}-C); HRMS *m/z* (ESI^+) [Found: 283.1078, $C_{16}H_{15}O_3N_2$ ($[M+H]^+$) requires 283.1077]; LRMS *m/z* (ESI^+) 283 ($[M+H]^+$, 100%); HPLC Retention time 8.7 min, >99.0%.

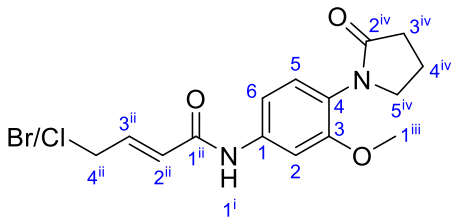
_N_-[3-Methyl-4-(4^{ii}-prop-2^{iii}-ynoxyphenyl)-isoxazol-5-yl]propenamide (5.6)



Propionyl chloride (86 µL, 0.986 mmol, 1.5 eq) was added dropwise to a solution of 3-methyl-4-(4^{ii}-prop-2^{iii}-ynoxyphenyl)-isoxazol-5-amine (**5.23**) (150 mg, 0.657 mmol, 1.0 eq) and pyridine (80 µL, 0.986 mmol, 1.5 eq) in anhydrous CH_2Cl_2 (6 mL) at 0 °C. The reaction was stirred at 0 °C for 90 min before being quenched by the addition of aqueous HCl (1.0 M, 20 mL), and extracted with CH_2Cl_2 (3 × 30 mL). The organic extracts were combined, washed with H_2O (30 mL) and brine (30 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel column chromatography (60–90% Et_2O/petroleum ether) to isolate the title compound as a colourless oil (105 mg, 56%); R_f 0.27 (80% Et_2O/petroleum ether); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3284 (N-H, m), 3028 (C-H, w), 2121 (C≡C, w), 1688 (C=O, m); ^{1}H NMR (600 MHz, $CDCl_3$) δ_H 7.63 (1H, br s, 6-N*H*),

7.22–7.18 (2H, m, 2^{ii}-C*H*), 7.04–7.00 (2H, m, 3^{ii}-C*H*), 4.71 (2H, d, *J* 2.3, 1^{iii}-CH_3), 2.54 (1H, t, *J* 2.3, 3^{iii}-C*H*), 2.38 (2H, q, *J* 7.4, 2^{iv}-C*H*), 2.26 (3H, s, 1^{i}-CH_3), 1.13 (3H, t, *J* 7.4, 3^{iv}-CH_3); ^{13}C NMR (151 MHz, $CDCl_3$) δ_C 172.8 (1^{iv}-C), 160.4 (3-C), 157.4 (4^{ii}-C), 156.2 (5-C), 130.0 (2^{ii}-C), 122.0 (1^{ii}-C), 115.6 (3^{ii}-C), 108.7 (4-C), 78.4 (2^{iii}-C), 76.0 (3^{iii}-C), 56.0 (1^{iii}-C), 29.1 (2^{iv}-C), 11.5 (1^{i}-C), 9.1 (3^{iv}-C); HRMS *m/z* (ESI^+) [Found: 285.1230, $C_{16}H_{17}O_3N_2$ ([M+H]$^+$) requires 285.1234]; LRMS *m/z* (ESI^+) 285 ([M+H]$^+$, 100%); HPLC Retention time 8.7 min, >99.0%.

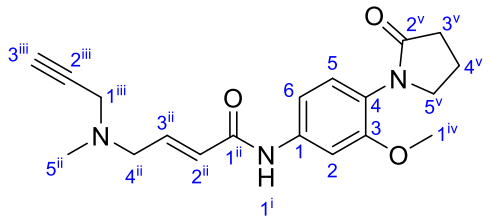
(*E*)-4^{ii}-Bromo/Chloro-*N*-(3-methoxy-4-(2^{iv}-oxopyrrolidin-1^{iv}-yl)phenyl)but-2^{ii}-enamide (5.31)



Using a modified procedure from Chew *et al.*,[166] oxalyl chloride (40 µL, 0.472 mmol, 1.95 eq) was added dropwise to a solution of (*E*)-4-bromocrotonic acid (80 mg, 0.484 mmol, 2.0 eq) in anhydrous THF (1.5 mL), and anhydrous DMF (1 drop) at 0 °C. The reaction was stirred at 0 °C for 15 min and then warmed to rt and stirred for a further 30 min. The solution was then cooled back down to 0 °C. 1-(4-Amino-2-methoxyphenyl)pyrrolidin-2^{i}-one (**5.30**) (50 mg, 0.242 mmol, 1.0 eq) in anhydrous *N*-methyl-2-pyrrolidone (1 mL) was then added dropwise at 0 °C, and the reaction stirred for 75 min at 0 °C. The reaction was quenched by the addition of H_2O (20 mL), and extracted with EtOAc (3 × 30 mL). The organic extracts were combined, washed with aqueous NaOH (1.0 M, 20 mL), aqueous LiCl solution (0.5 M, 2 × 30 mL), and brine (30 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo* to isolate the title compound as a pale purple solid. No further purification was required (76 mg, 89%); R_f 0.38 (10% MeOH/CH_2Cl_2); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 9.02 (1H, br s, 1^{i}-N*H*), 7.40 (1H, d, *J* 2.1, 2-C*H*), 7.03–6.92 (1H, m, 3^{ii}-C*H*), 6.88 (1H, dd, *J* 8.4, 3.4, 5-C*H*), 6.59 (1H, td, *J* 8.4, 2.2, 6-C*H*), 6.26 (0.4H, dt, *J* 15.0, 1.6, 2^{ii}-C$H_{(Cl)}$), 6.19 (0.6H, dt, *J* 15.0, 1.1, 2^{ii}-C$H_{(Br)}$), 4.22

(0.8H, dd, *J* 5.9, 1.5, 4^{ii}-C*H*$_{2(Cl)}$), 4.07 (1.2H, dd, *J* 7.3, 1.1, 4^{ii}-C*H*$_{2(Br)}$), 3.73–3.67 (5H, m, 1^{iii}-C*H*$_3$, 5^{iv}-C*H*$_2$), 2.62 (2H, t, *J* 8.1, 3^{iv}-C*H*$_2$), 2.26–2.17 (2H, m, 4^{iv}-C*H*$_2$); LRMS *m/z* (ESI$^+$) 309, 311 ([M+H]$^+$; ^{35}Cl, 100%; ^{37}Cl, 33%), 353, 355 ([M+H]$^+$; ^{79}Br, 100%; ^{81}Br, 97%).

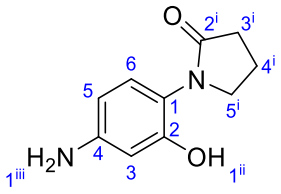
(*E*)-*N*-[3-Methoxy-4-(2^{v}-oxopyrrolidin-1^{v}-yl)phenyl]-4^{ii}-[methyl(prop-2^{iii}-ynyl)amino]but-2^{ii}-enamide (5.7)



Potassium Iodide (71 mg, 0.426 mmol, 2.0 eq) was added to a solution of (*E*)-4^{ii}-bromo/chloro-*N*-(3-methoxy-4-(2^{iv}-oxopyrrolidin-1^{iv}-yl)phenyl)but-2^{ii}-enamide (**5.31**) (75 mg, 0.213 mmol, 1.0 eq) in anhydrous DMF (2 mL) at rt. The reaction was stirred at rt for 10 min, after which it was cooled to 0 °C. *N*-Methylpropargylamine (36 µL, 0.426 mmol, 2.0 eq) was added then added dropwise at 0 °C, followed by K_2CO_3 (59 mg, 0.426 mmol, 2.0 eq), and the suspension stirred at 0 °C for 5 min before slowing warming to rt. The reaction was stirred at rt for 2 h before being quenched by the addition of aqueous HCl (1.0 M, 20 mL), and washed with EtOAc (20 mL). The acidic aqueous extract was basified with aqueous NaOH (2.0 M) until pH 11, extracted with EtOAc (3 × 20 mL), washed with aqueous LiCl solution (0.5 M, 20 mL), and brine (20 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using semi-preparative HPLC (see general chemical methods) to isolate the formic acid salt, which was then dissolved in aqueous NaOH (1.0 M), extracted with EtOAc (3 × 20 mL), and concentrated *in vacuo* to isolate the free base as a colourless solid (20 mg, 28%); R_f 0.30 (10% MeOH/CH_2Cl_2); m.p 130–132 °C (from $CHCl_3$); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3279 (N-H, m), 2944 (C-H, w), 1667 (C=O, s), 1609 (C=C, s); ^{1}H NMR (600 MHz, $CDCl_3$) δ_H 8.77 (1H, br s, 1^{i}-N*H*), 7.45 (1H, d, *J* 1.9, 2-C*H*), 6.91 (1H, d, *J* 8.4, 5-C*H*), 6.88 (1H, dt, *J* 15.3, 6.3, 3^{ii}-C*H*), 6.64 (1H, dd, *J* 8.2, 1.9, 6-C*H*), 6.16 (1H, dt, *J* 15.3, 1.5, 2^{ii}-C*H*), 3.71–3.67 (5H, m, 1^{iv}-C*H*$_3$,

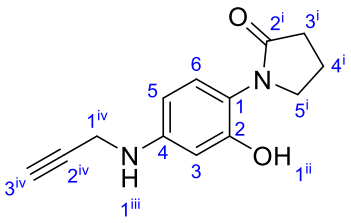
5^{v}-CH_2), 3.39 (2H, d, *J* 2.2, 1^{iii}-CH_2), 3.25 (2H, dd, *J* 6.3, 1.5, 4^{ii}-C*H*), 2.59 (2H, t, *J* 8.1, 3^{v}-CH_2), 2.37 (3H, s, 5^{ii}-CH_3), 2.27 (1H, t, *J* 2.2, 3^{ii}-C*H*), 2.22–2.16 (2H, m, 4^{v}-CH_2); ^{13}C NMR (151 MHz, $CDCl_3$) δ_C 176.3 (2^{v}-C), 163.8 (1^{ii}-C), 154.9 (3-C), 140.7 (3^{ii}-C), 139.5 (1-C), 128.3 (5-C), 127.1 (2^{ii}-C), 122.0 (4-C), 112.3 (6-C), 104.6 (2-C), 78.3 (2^{iii}-C), 73.8 (3^{iii}-C), 56.7 (4^{ii}-C), 55.6 (1^{iv}-C), 50.5 (5^{v}-C), 45.5 (1^{iii}-C), 41.9 (5^{ii}-C), 31.5 (3^{v}-C), 18.9 (4^{v}-C); HRMS *m/z* (ESI^+) [Found: 342.1814, $C_{19}H_{24}O_3N_3$ ($[M+H]^+$) requires 342.1812]; LRMS *m/z* (ESI^+) 342 ($[M+H]^+$, 100%); HPLC Retention time 5.6 min, >99.0%.

1-(4-Amino-2-hydroxyphenyl)pyrrolidin-2^{i}-one (5.32)



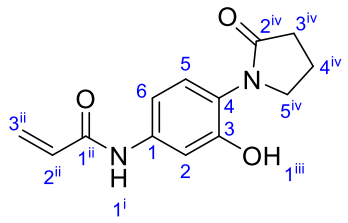
Boron tribromide (1.0 M in CH_2Cl_2, 728 µL, 0.728 mmol, 1.5 eq) was added dropwise to a solution of 1-(4-amino-2-methoxyphenyl)pyrrolidin-2^{i}-one (**5.30**) (100 mg, 0.485 mmol, 1.0 eq) in CH_2Cl_2 (4 mL) at 0 °C. The reaction was slowly warmed to rt and stirred for 2.5 h before being cooled back to 0 °C and a further portion of boron tribromide (1.0 M in CH_2Cl_2, 728 µL, 0.728 mmol, 1.5 eq) added dropwise. The reaction was then stirred at rt for a further 20 h before being quenched by the addition of saturated aqueous $NaHCO_3$ solution (30 mL) and extracted with EtOAc (3 × 30 mL). The organic extracts were combined, washed with brine (30 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel column chromatography (0–5% EtOH/EtOAc) to isolate the title compound as a pale pink oil (10 mg, 11%); R_f 0.45 (10% MeOH/CH_2Cl_2); ^{1}H NMR (400 MHz, methanol-D_4) δ_H 6.84 (1H, d, *J* 8.4, 6-C*H*), 6.28 (1H, d, *J* 2.4, 3-C*H*), 6.24 (1H, dd, *J* 8.4, 2.4, 5-C*H*), 3.74 (2H, t, *J* 7.2, 5^{i}-CH_2), 2.52 (2H, t, *J* 8.0, 3^{i}-CH_2), 2.22–2.13 (2H, m, 4^{i}-CH_2); LRMS *m/z* (ESI^+) 193 ($[M+H]^+$, 100%).

1-[2-Hydroxy-4-(prop-2^{iv}-ynylamino)phenyl]pyrrolidin-2^{i}-one (5.34)



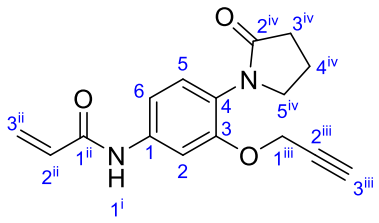
K_2CO_3 (11 mg, 0.078 mmol, 1.5 eq) was added to a solution of 1-(4-amino-2-hydroxyphenyl)pyrrolidin-2^{i}-one (**5.32**) (10 mg, 0.052 mmol, 1.0 eq) in anhydrous DMF (2 mL). The suspension was stirred at rt for 20 min before being cooled to 0 °C. Propargyl bromide solution (80 wt. % in toluene, 9 µL, 0.062 mmol, 1.2 eq) was then added dropwise to the suspension at 0 °C, and stirred for 15 min before being raised slowly to rt. After 3.5 h the reaction was cooled to 0 °C and a further portion of propargyl bromide solution (80 wt. % in toluene, 18 µL, 0.124 mmol, 2.4 eq) was added. The reaction was stirred at rt for a further 24 h, after which it was quenched by the addition of saturated aqueous $NaHCO_3$ solution (10 mL), and extracted with EtOAc (3 × 15 mL). The organic extracts were combined, washed with H_2O (20 mL), and brine (20 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel chromatography (30–100% EtOAc/petroleum ether) to isolate the title compound as a pale yellow oil (9 mg, 75%); R_f 0.37 (100% EtOAc); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 8.71(1H, br s, 1^{iii}-N*H*), 6.91 (1H, d, *J* 8.6, 6-C*H*), 6.43 (1H, d, *J* 2.7, 3-C*H*), 6.31 (1H, dd, *J* 8.6, 2.7, 5-C*H*), 3.93–3.88 (4H, m, 5^{i}-CH_2, 1^{iv}-CH_2), 2.66 (2H, t, *J* 8.0, 3^{i}-CH_2), 2.30–2.21 (3H, m, 4^{i}-CH_2, 3^{iv}-C*H*); LRMS *m/z* (ESI$^+$) 231 ([M+H]$^+$, 100%).

_N_-[3-Hydroxy-4-(2-oxopyrrolidin-1-yl)phenyl]prop-2-enamide (5.35)



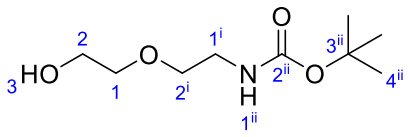
Acryloyl chloride (9.9 µL, 0.12 mmol, 1.1 eq) was added dropwise to a solution of 1-(4-amino-2-hydroxyphenyl)pyrrolidin-2-one (**5.32**) (21.3 mg, 0.11 mmol, 1.0 eq) and pyridine (9.8 µL, 0.12 mmol, 1.1 eq) in dry CH_2Cl_2 (0.6 mL) at 0 °C. The reaction was slowly warmed to rt and stirred for 2 h before being cooled back to 0 °C and a further portion of pyridine (9.8 µL, 0.12 mmol, 1.1 eq) and acryloyl chloride (9.93 µL, 0.12 mmol, 1.1 eq) were added dropwise. The reaction was then stirred at rt for a further 2 h before being quenched by the addition of saturated aqueous $NaHCO_3$ solution (5 mL) and extracted with CH_2Cl_2 (2 × 5 mL). The combined organic phases were then washed with aqueous HCl (1.0 M, 2 × 5 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo* to isolate the title compound as a yellow solid (12 mg, 43%); R_f 0.26 (EtOAc); m.p. 162–164 °C (from CH_2Cl2); $\tilde{\upsilon}_{max}$ (neat)/cm^{-1} 3109 (O-H, w), 2920 (C-H, w), 2850 (C-H, w), 1660 (C=O, s), 1609 (C=O, s), 1524 (C=C, s), 1421 (s), 1246 (N-O, w), 975 (C=C-H, w), 855 (w); ^{1}H NMR (600 MHz, DMSO-D_6) δ_H 10.07 (1H, s, 1^i-N*H*), 9.61 (1H, s, 1^{iii}-O*H*), 7.42 (1H, d, *J* 2.1, 2-C*H*), 7.06–7.01 (2H, m, 5-C*H*, 6-C*H*), 6.44–6.40 (1H, m, 2^{ii}-C*H*), 6.24 (1H, dd, *J* 16.9, 1.9, 3^{ii}-CH_xH$_y$), 5.74 (1H, dd, *J* 10.0, 1.9, 3^{ii}-CH$_x$$H_y$), 3.66 (2H, t, *J* 6.9, 5^{iv}-CH_2), 2.37 (2H, t, *J* 8.1, 3^{iv}-CH_2), 2.07 (2H, quint, *J* 7.7, 4^{iv}-CH_2); ^{13}C (NMR 151 MHz; DMSO-D_6) δ_C 174.2 (2^{iv}-C), 163.0 (1^{ii}-C), 152.5 (3-C), 138.6 (1-C), 131.9 (2^{ii}-C), 127.8 (5-C), 126.8 (3^{ii}-C), 121.7 (4-C), 110.3 (6-C), 107.7 (2-C), 49.3 (5^{iv}-C), 30.7 (3^{iv}-C), 18.3 (4^{iv}-C); HRMS *m/z* (ESI^+) [Found: 247.1078, $C_{13}H_{15}N_2O_3$ requires $[M+H]^+$ 247.1077]; LRMS *m/z* (ESI^+) 247 ($[M+H]^+$, 93%), 269 ($[M+Na]^+$, 100%); HPLC Retention time 5.5 min, 83.5%.

_N_-[4-(2-oxopyrrolidin-1-yl)-3-prop-2-ynoxyphenyl]prop-2-enamide (5.8)



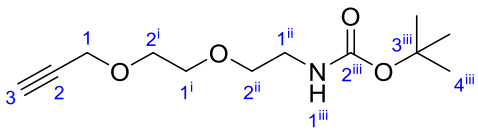
To a solution of *N*-[3-hydroxy-4-(2-oxopyrrolidin-1-yl)phenyl]prop-2-enamide (**5.35**) (25 mg, 0.10 mmol, 1.0 eq) in dry DMF (2 mL), K_2CO_3 (21 mg, 0.15 mmol, 1.5 eq) was added, followed by the addition of propargyl bromide (80% v/v in toluene, 11.5 µL, 0.12 mmol, 1.2 eq) after 10 mins. The mixture was stirred at rt for 20 h, then diluted with water (5 mL) and extracted with EtOAc (3 × 5 mL). The organic phases were combined, dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified by silica gel column chromatography (0–10% EtOAc/petroleum ether) to isolate the title compound as a yellow solid (10 mg, 35%); R_f 0.51 (10% MeOH/EtOAc); m.p. 110–112 °C (from EtOAc); $\tilde{\upsilon}_{max}$ (neat)/cm^{-1} 3291 (N-H, w), 2918 (C-H, w), 2850 (C-H, w), 2114 (C≡C, w), 1673 (C=O, s), 1609 (C=O, s), 1546 (C=C, s), 1516 (w), 1416 (s), 1269 (N-O, w), 1036 (C=C-H, w), 851 (w), 637 (C≡C-H, w); ^{1}H NMR (600 MHz, $CDCl_3$) δ_H 8.93 (1H, s, 1^i-N*H*), 7.50 (1H, s, 2-C*H*), 6.93 (1H, d, *J* 8.2, 5-C*H*), 6.74 (1H, d, *J* 8.2, 6-C*H*), 6.39 (1H, d, *J* 16.3, 3^{ii}-C*H*$_x$H$_y$), 6.30–6.25 (1H, m, 2^{ii}-C*H*), 5.69 (1H, d, *J* 9.8, 3^{ii}-CH$_x$*H*$_y$), 4.56 (2H, d, *J* 2.4, 1^{iii}-C*H*$_2$), 3.73 (2H, t, *J* 7.0, 5^{iv}-C*H*$_2$), 2.61 (2H, t, *J* 8.0, 3^{iv}-C*H*$_2$), 2.50 (1H, s, 3^{iii}-C*H*), 2.21 (2H, quint, *J* 7.8, 4^{iv}-C*H*$_2$); ^{13}C NMR (151 MHz; $CDCl_3$) δ_C 176.5 (2^{iv}-C), 164.0 (1^{ii}-C), 152.8 (3-C), 139.4 (1-C), 131.5 (2^{ii}-C), 128.5 (5-C), 127.4 (3^{ii}-C), 122.5 (4-C), 113.5 (6-C), 105.9 (2-C), 78.1 (2^{iii}-C), 75.9 (3^{iii}-C), 56.3 (1^{iii}-C), 50.8 (5^{iv}-C), 31.5 (3^{iv}-C), 19.0 (4^{iv}-C); HRMS *m/z* (ESI$^+$) [Found: 285.1231, $C_{16}H_{17}N_2O_3$ requires $[M+H]^+$ 285.1234]; LRMS *m/z* (ESI$^+$) 285 ($[M+H]^+$, 100%), 307 ($[M+Na]^+$, 71%); HPLC Retention time 6.5 min, 96.0%.

tert-Butyl [2^{i}-(2-hydroxyethoxy)ethyl]carbamate (6.5)[191]



Following the procedure from Budhadev *et al*.,[191] di-*tert*-butyl decarbonate (6.22 g, 28.3 mmol, 1.0 eq) in $CHCl_3$ (10 mL) was added portion-wise to a solution of 2-(2-aminoethoxy)-ethanol (2.86 mL, 28.53 mmol, 1.0 eq) in $CHCl_3$ (28 mL) at 0 °C. The reaction was slowly warmed to rt, and stirred for 22 h, after which it was diluted with H_2O (30 mL) and extracted with $CHCl_3$ (2 × 50 mL). The organic components were combined, washed with H_2O (50 mL), and brine (50 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo* to isolate the title compound as a colourless oil. No further purification was required (5.83 g, 99%); R_f 0.23 (75% EtOAc/petroleum ether); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 5.20 (1H, br s, 1^{ii}-N*H*), 3.71–3.66 (2H, m, 2-C*H*$_2$), 3.54–3.48 (4H, m, 1-C*H*$_2$, 2^{i}-C*H*$_2$), 3.28 (2H, dt, *J* 5.2, 5.1, 1^{i}-C*H*$_2$), 2.96 (1H, t, *J* 5.7, 3-O*H*), 1.40 (9H, s, 4^{ii}-C*H*$_3$); LRMS *m/z* (ESI$^+$) 228 ([M+Na]$^+$, 100%). These data are in good agreement with the literature.[191]

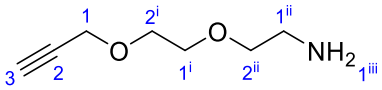
tert-Butyl (2^{ii}-{2^{i}-[(prop-2-yn-1-yl)oxy]ethoxy}ethyl)carbamate (6.6)[191]



Following the procedure of Budhadev *et al*.,[191] sodium hydride (60% dispersion in mineral oil, 643 mg, 16.1 mmol, 1.1 eq) was added portion-wise to a solution of *tert*-butyl [2^{i}-(2-hydroxyethoxy)ethyl]carbamate (**6.5**) (3.00 g, 14.6 mmol, 1.0 eq) in anhydrous THF (30 mL) at 0 °C. The suspension was stirred for 10 min at 0 °C before the dropwise addition of propargyl bromide solution (80 wt. % in toluene, 1.79 mL, 16.1 mmol, 1.1 eq) at 0 °C. The reaction was slowly warmed to rt, and stirred for 15 h, after which it was quenched by the addition of aqueous HCl (1.0 M, 30 mL), and the THF removed *in vacuo*. The solution was extracted with EtOAc (3 × 50 mL), and the

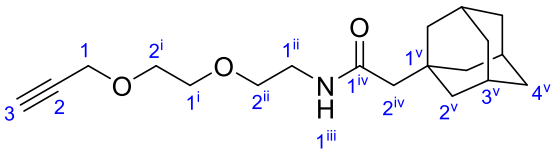
organic components combined, washed with H_2O (50 mL), brine (50 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo* to isolate the title compound as an orange oil. No further purification was required (3.41 g, 96%); R_f 0.70 (75% EtOAc/petroleum ether); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 4.98 (1H, br s, 1^{iii}-N*H*), 4.17 (2H, d, *J* 2.4, 1-CH_2), 3.67–3.63 (2H, m, 2^{i}-CH_2), 3.62–3.59 (2H, m, 1^{i}-CH_2), 3.51 (2H, t, *J* 5.2, 2^{ii}-CH_2), 3.28 (2H, dt, *J* 5.2, 5.2, 1^{ii}-CH_2), 2.42 (1H, t, *J* 2.4, 3-C*H*), 1.41 (9H, s, 4^{iii}-CH_3); LRMS *m/z* (ESI^+) 244 ($[M+H]^+$, 100%). These data are in good agreement with the literature.[191]

2^{ii}-{2^{i}-[(Prop-2-yn-1-yl)oxy]ethoxy}ethan-1^{ii}-amine (6.7)



Following the procedure of Neklesa *et al*.,[186] trifluoroacetic acid (4.0 mL) was added portion-wise to a solution of *tert*-butyl *tert*-butyl (2^{ii}-{2^{i}-[(prop-2-yn-1-yl)oxy]ethoxy}ethyl)carbamate (**6.6**) (1.00 g, 4.11 mmol, 1.0 eq) in CH_2Cl_2 (20 mL) at 0 °C. The reaction was slowly warmed to rt, and stirred for 1.5 h, after which it was concentrated *in vacuo*, and excess TFA removed by azeotropic distillation with toluene, to isolate the title compound as an orange oil. This intermediate was taken directly on to the next step without further purification; R_f 0.05 (75% EtOAc/petroleum ether).

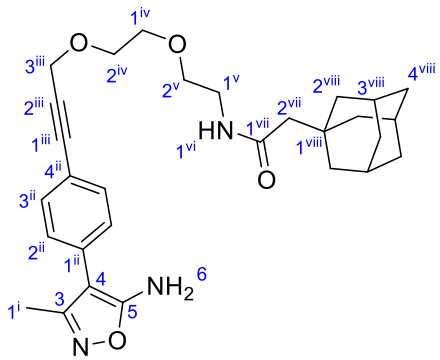
2^{iv}-(Adamantan-1^{v}-yl)-*N*-(2^{ii}-{2^{i}-[(prop-2-yn-1-yl)oxy]ethoxy}ethyl)acetamide (6.3)



Following the procedure of Neklesa *et al*.,[186] EDCI (946 mg, 4.93 mmol, 1.2 eq) was added to a suspension of 2^{ii}-{2^{i}-[(prop-2-yn-1-yl)oxy]ethoxy}ethan-1^{ii}-amine (**6.7**) (589 mg, 4.11 mmol, 1.0 eq), 1-adamantaneacetic acid (878 mg, 4.52 mmol, 1.1 eq), hydroxybenzotriazole (666 mg, 4.93 mmol, 1.2 eq), and DIPEA (2.86 mL, 16.4 mmol, 4.0 eq) in CH_2Cl_2 (40 mL) at 0 °C. The reaction was slowly warmed to rt, and stirred for 45 h, after which it was diluted with saturated aqueous $NaHCO_3$

solution (30 mL), and extracted with CH_2Cl_2 (3 × 50 mL). The organic extracts were combined, washed with H_2O (30 mL), and brine (30 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel column chromatography (50–90% EtOAc/petroleum ether) to isolate the title compound as a pale-yellow oil (920 mg, 70% over two steps); R_f 0.17 (50% EtOAc/petroleum ether); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3307 (C-H, m br), 2902 (C-H, s), 2114 (C≡C, w), 1645 (C=O, s), 1101 (C-O, s); ^{1}H NMR (400 MHz, $CDCl_3$) δ_H 5.84 (1H, br s, 1^{iii}-N*H*), 4.21 (2H, d, *J* 2.4, 1-C*H*$_2$), 3.71–3.68 (2H, m, 2^{i}-C*H*$_2$), 3.66–3.63 (2H, m, 1^{i}-C*H*$_2$), 3.56 (2H, t, *J* 5.2, 2^{ii}-C*H*$_2$), 3.45 (2H, dt, *J* 5.2, 5.2, 1^{ii}-C*H*$_2$), 2.44 (1H, t, *J* 2.4, 3-C*H*), 1.99–1.95 (3H, m, 3^{v}-C*H*), 1.93 (2H, s, 2^{iv}-C*H*$_2$), 1.73–1.63 (6H, m, 4^{v}-C*H*$_2$), 1.62 (6H, d, *J* 2.3, 2^{v}-C*H*$_2$); ^{13}C NMR (151 MHz, $CDCl_3$) δ_C 171.2 (1^{iv}-C), 79.6 (2-C), 74.0 (3-C), 70.2 (1^{i}-C), 70.2 (2^{ii}-C), 69.2 (2^{i}-C), 58.6 (1-C), 51.9 (2^{iv}-C), 42.8 (2^{v}-C), 39.1 (1^{ii}-C), 36.9 (4^{v}-C), 32.9 (1^{v}-C), 28.8 (3^{v}-C); HRMS *m/z* (ESI$^+$) [Found: 320.2220, $C_{19}H_{30}O_3N$ ([M+H]$^+$ requires 320.2220]; LRMS *m/z* (ESI$^+$) 320 ([M+H]$^+$, 100%).

2^{vii}-(Adamantan-1^{viii}-yl)-*N*-{2^{v}-[2^{iv}-({3^{iii}-[4^{ii}-(5-amino-3-methylisoxazol-4-yl)phenyl]prop-2^{iii}-yn-1^{iii}-yl}oxy)ethoxy]ethyl}acetamide (6.8)



Method 1:

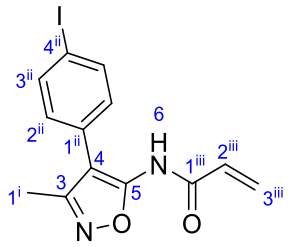
4-(4^{ii}-Iodophenyl)-3-methylisoxazol-5-amine (**5.14**) (100 mg, 0.333 mmol, 1.0 eq), 2^{iv}-(adamantan-1^{v}-yl)-*N*-(2^{ii}-{2-[(prop-2^{i}-yn-1^{i}-yl)oxy]ethoxy}ethyl)acetamide (**6.3**) (213 mg, 0.666 mmol, 2.0 eq), $Pd(PPh_3)_2Cl_2$ (14 mg, 0.020 mmol, 0.06 eq), and CuI (7.60 mg, 0.040 mmol, 0.12 eq) were added to a microwave vial, and the vial sealed and purged with Ar. Degassed triethylamine (1.0 mL) and DMF

(0.25 mL) were added, and the reaction stirred at rt for 1 h, after which a constant stream of N_2 was applied to remove the triethylamine. The crude product was purified using silica gel chromatography (30–100% EtOAc/petroleum ether) to isolate the title compound as a dark yellow oil (124 mg, 76%)

Method 2:

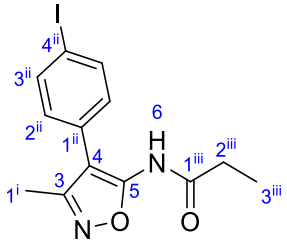
4-(4^{ii}-Iodophenyl)-3-methylisoxazol-5-amine (**5.14**) (300 mg, 1.00 mmol, 1.1 eq), 2^{iv}-(adamantan-1^{v}-yl)-*N*-(2^{ii}-{2-[(prop-2^{i}-yn-1^{i}-yl)oxy]ethoxy}ethyl)acetamide (**6.3**) (290 mg, 0.909 mmol, 1.0 eq), Pd(dppf)Cl_2.CH_2Cl_2 (45 mg, 0.055 mmol, 0.06 eq), and CuI (21 mg, 0.109 mmol, 0.12 eq) were added to a microwave vial, and the vial sealed and purged with Ar. Degassed triethylamine (4.0 mL) and DMF (1.0 mL) were added, and the reaction stirred at 40 °C for 3 h, after which it was diluted with aqueous HCl (1.0 M, 20 mL) and extracted with EtOAc (3 × 30 mL). The organic extracts were combined, washed with aqueous LiCl solution (0.5 M, 20 mL), and brine (20 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel chromatography (40–100% EtOAc/petroleum ether) to isolate the title compound as a pale-brown oil (400 mg, 90%); R_f 0.37 (EtOAc); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3304 (N-H, w br), 2845 (C-H, w), 1640 (C=O, m); ^{1}H NMR (600 MHz, $CDCl_3$) δ_H 7.45–7.41 (2H, m, 3^{ii}-C*H*), 7.23–7.19 (2H, m, 2^{ii}-C*H*), 5.98 (1H, t, *J* 5.3, 1^{vi}-N*H*), 4.88 (2H, s, 6-NH_2), 4.40 (2H, s, 3^{iii}-CH_2), 3.75–3.71 (2H, m, 2^{iv}-CH_2), 3.67–3.64 (2H, m, 1^{iv}-CH_2), 3.55 (2H, t, *J* 5.4, 2^{v}-CH_2), 3.42 (2H, dt, *J* 5.4, 5.3, 1^{v}-CH_2), 2.19 (3H, s, 1^{i}-CH_3), 1.94–1.89 (3H, m, 3^{viii}-C*H*), 1.87 (2H, s, 2^{vii}-CH_2), 1.67–1.57 (6H, m, 4^{viii}-CH_2), 1.56 (6H, d, *J* 2.3, 2^{viii}-CH_2); ^{13}C NMR (151 MHz, $CDCl_3$) δ_C 171.2 (1^{vii}-C), 165.1 (5-C), 159.5 (3-C), 132.5 (3^{ii}-C), 131.6 (1^{ii}-C), 127.8 (2^{ii}-C), 120.5 (4^{ii}-C), 94.3 (4-C), 86.2 (1^{iii}-C), 85.4 (2^{iii}-C), 70.1 (1^{iv}-C), 70.1 (2^{v}-C), 69.2 (2^{iv}-C), 59.3 (3^{iii}-C), 51.7 (2^{vii}-C), 42.6 (2^{viii}-C), 39.1 (1^{v}-C), 36.8 (4^{viii}-C), 32.7 (1^{viii}-C), 28.7 (3^{viii}-C), 11.4 (1^{i}-C); HRMS *m/z* (ESI$^+$) [Found: 492.2856, $C_{29}H_{38}O_4N_3$ ([M+H]$^+$) requires 492.2857]; LRMS *m/z* (ESI$^+$) 492 ([M+H]$^+$, 100%); HPLC Retention time 10.6 min, >99.0%.

_N_-[4-(4^{ii}-Iodophenyl)-3-methylisoxazol-5-yl]prop-2^{iii}-enamide (6.9)



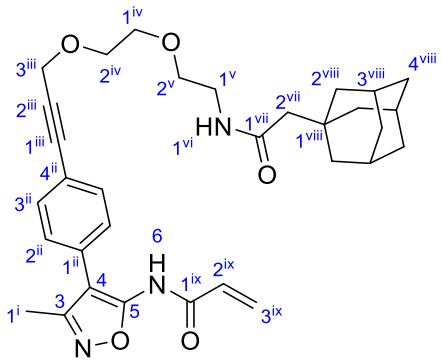
Acryloyl chloride (269 µL, 3.33 mmol, 2.0 eq) was added dropwise to a solution of 4-(4^{ii}-iodophenyl)-3-methylisoxazol-5-amine (**5.14**) (500 mg, 1.67 mmol, 1.0 eq), and pyridine (269 µL, 3.33 mmol, 2.0 eq) in CH_2Cl_2 (10 mL) at 0 °C. The reaction was stirred at 0 °C for 1.5 h, after which a further portion of acryloyl chloride (202 µL, 2.51 mmol, 1.5 eq) was added dropwise. The reaction was stirred for a further 3.5 h at 0 °C before being raise to rt for another 1.5 h. The reaction was quenched by the addition of aqueous HCl (1.0 M, 20 mL), and extracted with CH_2Cl_2 (3 × 30 mL). The organic extracts were combined, washed with H_2O (30 mL) and brine (30 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified by using silica gel column chromatography (30–80% Et_2O/petroleum ether) to isolate the title compound as a colourless solid (165 mg, 28%); R_f 0.24 (60% Et_2O/petroleum ether); m.p 115–117 °C (from $CHCl_3$); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3228 (N-H, m br), 3004 (C-H, m), 1682 (C=O, s), 1640 (C=C, s); ^{1}H NMR (600 MHz, $CDCl_3$) δ_H 7.83 (1H, br s, 6-N*H*), 7.75–7.72 (2H, m, 3^{ii}-C*H*), 7.03–7.00 (2H, m, 2^{ii}-C*H*), 6.41 (1H, dd, *J* 17.0, 0.9, 3^{iii}-C*H*$_x$H$_y$), 6.25 (1H, dd, *J* 17.0, 10.4, 2^{iii}-C*H*), 5.83 (1H, dd, *J* 10.4, 0.9, 3^{iii}-CH$_x$*H*$_y$), 2.27 (3H, s, 1^{i}-C*H*$_3$); ^{13}C NMR (151 MHz, $CDCl_3$) δ_C 163.2 (1^{iii}-C), 160.3 (3-C), 156.0 (5-C), 138.2 (3^{ii}-C), 130.9 (3^{iii}-C), 130.6 (2^{ii}-C), 128.8 (1^{ii}-C), 128.6 (2^{iii}-C), 108.6 (4-C), 94.0 (4^{ii}-C), 11.5 (1^{i}-C); HRMS *m/z* (ESI$^+$) [Found: 354.9938, $C_{13}H_{12}O_2N_2I$ ([M+H]$^+$) requires 354.9949]; LRMS *m/z* (ESI$^+$) 355 ([M+H]$^+$, 100%); HPLC Retention time 9.8 min, 98.5%.

_N_-[4-(4^{ii}-Iodophenyl)-3-methylisoxazol-5-yl]propionamide (6.10)



Propionyl chloride (66 µL, 0.750 mmol, 1.5 eq) was added dropwise to a solution of 4-(4^{ii}-iodophenyl)-3-methylisoxazol-5-amine (**5.14**) (150 mg, 0.500 mmol, 1.0 eq) and pyridine (61 µL, 0.750 mmol, 1.5 eq) in anhydrous CH_2Cl_2 (5 mL) at 0 °C. The reaction was warmed to rt and stirred for 2 h before an additional portion of propionyl chloride (88 µL, 1.00 mmol, 2.0 eq) and pyridine (20 µL, 0.250 mmol, 0.5 eq) were added at 0 °C. The reaction was stirred at rt for further 22 h before being quenched by the addition of aqueous HCl (1.0 M, 15 mL), and extracted with EtOAc (3 × 30 mL). The organic extracts were combined, washed with H_2O (30 mL) and brine (30 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel column chromatography (30–85% Et_2O/petroleum ether) to isolate the title compound as a colourless oil (170 mg, 96%); R_f 0.42 (80% Et_2O/petroleum ether); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3241 (N-H, m br), 2940 (C-H, m), 1690 (C=O, s); ^{1}H NMR (600 MHz, $CDCl_3$) δ_H 7.77–7.73 (2H, m, 3^{ii}-C*H*), 7.50 (1H, br s, 6-N*H*), 7.03–6.99 (2H, m, 2^{ii}-C*H*), 2.39 (2H, q, *J* 7.5, 2^{iii}-C*H*$_2$), 2.26 (3H, s, 1^{i}-C*H*$_3$), 1.15 (3H, t, *J* 7.5, 3^{iii}-C*H*$_3$); ^{13}C NMR (151 MHz, $CDCl_3$) δ_C 172.23 (1^{iii}-C), 160.2 (3-C), 156.3 (5-C), 138.3 (3^{ii}-C), 130.5 (2^{ii}-C), 128.8 (1^{ii}-C), 108.3 (4-C), 93.9 (4^{ii}-C), 29.3 (2^{iii}-C), 11.5 (1^{i}-C), 9.1 (3^{iii}-C); HRMS *m/z* (ESI$^+$) [Found: 357.0094, $C_{13}H_{14}O_2N_2I$ ([M+H]$^+$) requires 357.0094]; LRMS *m/z* (ESI$^+$) 357 ([M+H]$^+$, 100%); HPLC Retention time 9.9 min, >99.0%.

_N_-[4-(4^{ii}-{3^{iii}-[2^{v}-(2^{iv}-{2^{vii}[Adamantan-1^{viii}-yl]acetamido}ethoxy)ethoxy]prop-1^{iii}-yn-1^{iii}-yl}phenyl)-3-methylisoxazol-5-yl]prop-2^{ix}-enamide (6.1)



Method 1:

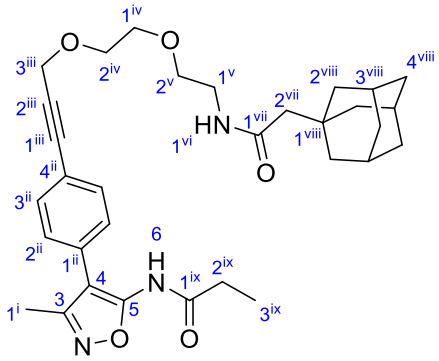
_N_-[4-(4^{ii}-Iodophenyl)-3-methylisoxazol-5-yl]prop-2^{iii}-enamide (**6.9**) (34 mg, 0.096 mmol, 1.1 eq), 2^{iv}-(adamantan-1^{v}-yl)-_N_-(2^{ii}-{2-[(prop-2^{i}-yn-1^{i}-yl)oxy]ethoxy}ethyl)acetamide (**6.3**) (28 mg, 0.087 mmol, 1.0 eq), $Pd(PPh_3)_2Cl_2$ (4 mg, 0.005 mmol, 0.06 eq), and CuI (2 mg, 0.010 mmol, 0.12 eq) were added to a microwave vial, and the vial sealed and purged with Ar. Degassed triethylamine (0.5 mL) and DMF (0.1 mL) were added, and the reaction stirred at rt for 1 h, after which it was diluted with aqueous HCl (1.0 M, 15 mL) and extracted with EtOAc (3 × 20 mL). The organic extracts were combined, washed with aqueous LiCl solution (0.5 M, 2 × 20 mL), and brine (20 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel chromatography (50–100% EtOAc/petroleum ether) to isolate the title compound as a colourless oil (32 mg, 67%).

Method 2:

_N_-[4-(4^{ii}-Iodophenyl)-3-methylisoxazol-5-yl]prop-2^{iii}-enamide (**6.9**) (75 mg, 0.212 mmol, 1.05 eq), 2^{iv} (adamantan-1^{v}-yl)-_N_-(2^{ii}-{2-[(prop-2^{i}-yn-1^{i}-yl)oxy]ethoxy}ethyl)acetamide (**6.3**) (64 mg, 0.202 mmol, 1.0 eq), $Pd(dppf)Cl_2.CH_2Cl_2$ (10 mg, 0.012 mmol, 0.06 eq), and CuI (5 mg, 0.024 mmol, 0.12 eq) were added to a microwave vial, and the vial sealed and purged with Ar. Degassed triethylamine (1 mL) and DMF (0.25 mL) were added, and the reaction stirred at rt for 3 h, after

which it was heated to 40 °C and stirred for a further 17 h. A constant stream of N_2 was applied to remove the triethylamine, and the remaining residue dissolved in H_2O (20 mL), and extracted with EtOAc (3 × 30 mL). The organic extracts were combined, washed with aqueous LiCl solution (0.5 M, 2 × 30 mL), and brine (30 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel chromatography (50–100% EtOAc/petroleum ether) to isolate the title compound as a colourless oil (21 mg, 19%); R_f 0.39 (EtOAc); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3270 (N-H, m br), 3016 (C-H, w), 2903 (C-H, s), 1692 (C=O, m), 1640 (C=C, s), 1097 (C-O, s); ^{1}H NMR (600 MHz, $CDCl_3$) δ_H 8.68 (1H, br s, 6-N*H*), 7.46–7.42 (2H, m, 3^{ii}-C*H*), 7.25–7.21 (2H, m, 2^{ii}-C*H*), 6.40 (1H, dd, *J* 17.0, 1.2, 3^{ix}-C*H*$_x$H$_y$), 6.32 (1H, dd, *J* 17.0, 10.1, 2^{ix}-C*H*), 5.93 (1H, t, *J* 5.2, 1^{vi}-N*H*), 5.78 (1H, dd, *J* 10.0, 1.2, 3^{ix}-CH$_x$*H*$_y$), 4.42 (2H, s, 3^{ii}-C*H*$_2$), 3.77–3.73 (2H, m, 2^{iv}-C*H*$_2$), 3.69–3.66 (2H, m, 1^{iv}-C*H*$_2$), 3.56 (2H, t, *J* 5.3, 2^{v}-C*H*$_2$), 3.42 (2H, dt, *J* 5.3, 5.2, 1^{v}-C*H*$_2$), 2.28 (3H, s, 1^{i}-C*H*$_3$), 1.95–1.91 (3H, m, 3^{viii}-C*H*), 1.87 (2H, s, 2^{vii}-C*H*$_2$), 1.69–1.58 (6H, m, 4^{viii}-C*H*$_2$), 1.57 (6H, d, *J* 2.3, 2^{viii}-C*H*$_2$); ^{13}C NMR (151 MHz, $CDCl_3$) δ_C 171.4 (1^{vii}-C), 163.4 (1^{ix}-C), 160.1 (3-C), 156.6 (5-C), 132.3 (3^{ii}-C), 130.3 (3^{ix}-C), 129.7 (1^{ii}-C), 129.0 (2^{ix}-C), 128.7 (2^{ii}-C), 122.1 (4^{ii}-C), 108.7 (4-C), 86.2 (1^{iii}-C), 86.0 (2^{iii}-C), 70.2 (1^{iv}-C), 70.0 (2^{v}-C), 69.2 (2^{iv}-C), 59.3 (3^{iii}-C), 51.9 (2^{vii}-C), 42.7 (2^{viii}-C), 39.2 (1^{v}-C), 36.9 (4^{viii}-C), 32.8 (1^{viii}-C), 28.8 (3^{viii}-C), 11.6 (1^{i}-C); HRMS *m/z* (ESI$^+$) [Found: 546.2963, $C_{32}H_{40}O_5N_3$ ([M+H]$^+$) requires 546.2962]; LRMS *m/z* (ESI$^+$) 546 ([M+H]$^+$, 100%); HPLC Retention time 10.6 min, >99.0%.

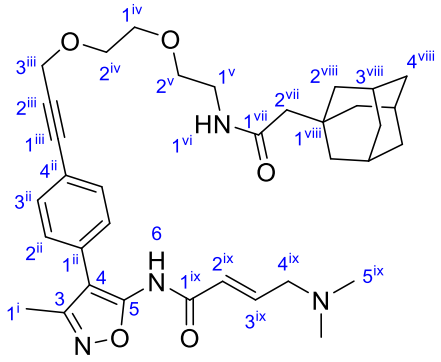
N-[4-(4^{ii}-{3^{iii}-[2^{v}-(2^{iv}-{2^{vii}[Adamantan-1^{viii}-yl]acetamido}ethoxy)ethoxy]prop-1^{iii}-yn-1^{iii}-yl}phenyl)-3-methylisoxazol-5-yl]propionamide (6.11)



N-[4-(4^{ii}-Iodophenyl)-3-methylisoxazol-5-yl]propionamide (**6.10**) (154 mg, 0.432 mmol, 1.05 eq), 2^{iv}-(adamantan-1^{v}-yl)-N-(2^{ii}-{2-[(prop-2^{i}-yn-1^{i}-yl)oxy]ethoxy}ethyl)acetamide (**6.3**) (132 mg, 0.412 mmol, 1.0 eq), Pd(dppf)Cl_2 (18 mg, 0.025 mmol, 0.06 eq), and CuI (9 mg, 0.049 mmol, 0.12 eq) were added to a microwave vial, and the vial sealed and purged with Ar. Degassed triethylamine (2 mL) and DMF (0.5 mL) were added, and the reaction stirred at 40 °C for 20 h. A constant stream of N_2 was applied to remove the triethylamine, and the remaining residue was dissolved in EtOAc, and filtered through a pad of Celite® and silica which was washed with EtOAc. The filtrate was washed with aqueous LiCl solution (0.5 M, 2 × 20 mL), and brine (30 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel chromatography (25–100% EtOAc/petroleum ether) to isolate the title compound as a yellow oil (185 mg, 82%); R_f 0.39 (100% EtOAc); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3333 (N-H, m br), 2904 (C-H, s), 1698 (C=O, m), 1100 (C-O, s); ^{1}H NMR (600 MHz, $CDCl_3$) δ_H 8.46 (1H, br s, 6-N*H*), 7.47–7.43 (2H, m, 2H, m, 3^{ii}-C*H*), 7.24–7.20 (2H, m, 2H, m, 2^{ii}-C*H*), 5.94 (1H, t, *J* 5.3, 1^{vi}-N*H*), 4.41 (2H, s, 3^{ii}-CH_2), 3.76–3.73 (2H, m, 2^{iv}-CH_2), 3.68–3.65 (2H, m, 1^{iv}-CH_2), 3.55 (2H, t, *J* 5.3, 2^{v}-CH_2), 3.42 (2H, dt, *J* 5.3, 5.2, 1^{v}-CH_2), 2.37 (2H, q, *J* , 7.5, 2^{ix}-CH_2), 2.26 (3H, s, 1^{i}-CH_3), 1.95–1.91 (3H, m, 3^{viii}-C*H*), 1.88 (2H, s, 2^{vii}-CH_2), 1.69–1.58 (6H, m, 4^{viii}-CH_2), 1.57 (6H, d, *J* 2.3, 2^{viii}-CH_2), 1.12 (3H, t, *J* 7.5, 3^{ix}-CH_3); ^{13}C NMR (151 MHz, $CDCl_3$) δ_C 172.5 (1^{ix}-C), 171.3 (1^{vii}-C), 160.0 (3-C), 156.8 (5-C), 132.3 (3^{ii}-C), 129.7 (1^{ii}-C), 128.6 (2^{ii}-C), 122.1 (4^{ii}-C), 108.6 (4-C), 86.1 (1^{iii}-C), 86.0 (2^{iii}-C), 70.2 (1^{iv}-C), 70.1 (2^{v}-C), 69.2 (2^{iv}-C), 59.3 (3^{iii}-C), 51.8 (2^{vii}-C),

42.7 (2^{viii}-C), 39.2 (1^{v}-C), 36.8 (4^{viii}-C), 32.8 (1^{viii}-C), 29.2 (2^{ix}-C), 28.7 (3^{viii}-C), 11.6 (1^{i}-C), 9.2 (3^{ix}-C); HRMS *m/z* (ESI^+) [Found: 548.3114, $C_{32}H_{42}O_5N_3$ ($[M+H]^+$) requires 548.3119]; LRMS *m/z* (ESI^+) 548 ($[M+H]^+$, 100%); HPLC Retention time 10.8 min, 99.0%.

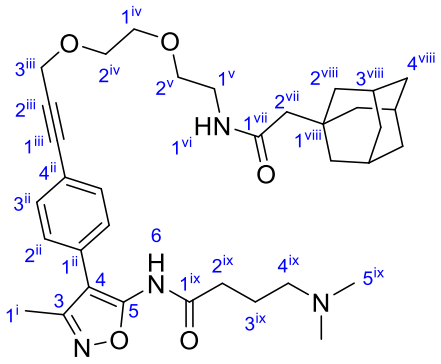
(*E*)-*N*-[4-(4^{ii}-{3^{iii}-[2^{v}-(2^{iv}-{2^{vii}[Adamantan-1^{viii}-yl]acetamido}ethoxy)ethoxy]prop-1^{iii}-yn-1^{iii}-yl}phenyl)-3-methylisoxazol-5-yl]-4^{ix}-(dimethylamino)but-2^{ix}-enamide (6.2)



Using a modified procedure from Chew *et al*.,[166] oxalyl chloride (42 µL, 0.495 mmol, 1.95 eq) was added dropwise to a suspension of (*E*)-4-(dimethylamino)but-2-enoic acid hydrochloride (84 mg, 0.509 mmol, 2.0 eq) in anhydrous THF (2 mL), and anhydrous DMF (1 drop) at 0 °C. The suspension was stirred at 0 °C for 20 min and then warmed to rt and stirred for a further 50 min. The solution was then cooled back down to 0 °C. 2^{vii}-(Adamantan-1^{viii}-yl)-*N*-{2^{v}-[2^{iv}-({3^{iii}-[4^{ii}-(5-amino-3-methylisoxazol-4-yl)phenyl]prop-2^{iii}-yn-1^{iii}-yl}oxy)ethoxy]ethyl}acetamide (**6.8**) (125 mg, 0.254 mmol, 1.0 eq) in anhydrous *N*-methyl-2-pyrrolidone (1 mL) was then added dropwise at 0 °C, and the reaction stirred for 20 h at 5 °C. The reaction was quenched by the addition of aqueous HCl (1.0 M, 30 mL), and washed with EtOAc (30 mL). The acidic aqueous extract was basified with aqueous NaOH (2.0 M) until pH 11, extracted with EtOAc (3 × 30 mL), washed with aqueous LiCl solution (0.5 M, 2 × 30 mL), and brine (30 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel column chromatography (0–100% 12:2:1 EtOAc: EtOH: aqueous NH_4OH/EtOAc) followed by semi-preparative HPLC (see general chemical methods) to isolate the title compound as the formic acid salt, which was then dissolved in aqueous

NaOH (1.0 M), extracted with EtOAc (3 × 20 mL), and concentrated in vacuo to isolate the free base as a colourless oil (15 mg, 10%); R_f 0.16 (50% 12:2:1 EtOAc: EtOH: aqueous NH_4OH/EtOAc); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3307 (N-H, m br), 2905 (C-H, s), 1676 (C=O, s), 1642 (C=C, s), 1096 (C-O, s); ^{1}H NMR (600 MHz, $CDCl_3$) δ_H 8.08 (1H, br s, 6-N*H*), 7.48–7.45 (2H, m, 3^{ii}-C*H*), 7.25–7.22 (2H, m, 2^{ii}-C*H*), 6.97 (1H, dt, *J* 15.5, 5.8, 3^{ix}-C*H*), 6.17 (1H, d, *J* 15.5, 2^{ix}-C*H*), 5.87 (1H, t, *J* 5.5, 1^{vi}-N*H*), 4.43 (2H, s, 3^{iii}-CH_2), 3.77–3.74 (2H, m, 2^{iv}-CH_2), 3.69–3.66 (2H, m, 1^{iv}-CH_2), 3.57 (2H, t, *J* 5.1, 2^{v}-CH_2), 3.43 (2H, dt, *J* 5.3, 5.1, 1^{v}-CH_2), 3.06 (2H, d, *J* 5.8, 4^{ix}-CH_2), 2.29 (3H, s, 1^{i}-CH_3), 2.23 (6H, s, 5^{ix}-CH_3), 1.96–1.92 (3H, m, 3^{viii}-C*H*), 1.89 (2H, s, 2^{vii}-CH_2), 1.70–1.59 (6H, m, 4^{viii}-CH_2), 1.59 (6H, d, *J* 2.3, 2^{viii}-CH_2); ^{13}C NMR (151 MHz, $CDCl_3$) δ_C 171.2 (1^{vii}-C), 163.3 (1^{ix}-C), 160.1 (3-C), 156.7 (5-C), 145.9 (3^{ix}-C), 132.4 (3^{ii}-C), 129.7 (1^{ii}-C), 128.7 (2^{ii}-C), 123.2 (2^{ix}-C), 122.3 (4^{ii}-C), 108.3 (4-C), 86.2 (1^{iii}-C), 86.1 (2^{iii}-C), 70.3 (1^{iv}-C), 70.1 (2^{v}-C), 69.3 (2^{iv}-C), 60.4 (4^{ix}-C), 59.4 (3^{iii}-C), 51.9 (2^{vii}-C), 45.6 (5^{ix}-C), 42.8 (2^{viii}-C), 39.2 (1^{v}-C), 36.9 (4^{viii}-C), 32.9 (1^{viii}-C), 28.8 (3^{viii}-C), 11.6 (1^{i}-C); HRMS *m/z* (ESI^+) [Found: 603.3536, $C_{35}H_{47}O_5N_4$ ($[M+H]^+$) requires 603.3541]; LRMS *m/z* (ESI^+) 603 ($[M+H]^+$, 100%); HPLC Retention time 8.7 min, >99.0%.

N-[4-(4^{ii}-{3^{iii}-[2^{v}-(2^{iv}-{2^{vii}[Adamantan-1^{viii}-yl]acetamido}ethoxy)ethoxy]prop-1^{iii}-yn-1^{iii}-yl}phenyl)-3-methylisoxazol-5-yl]-4^{ix}-(dimethylamino)butanamide (6.12)



Using a modified procedure from Chew *et al.*,[166] oxalyl chloride (57 µL, 0.675 mmol, 1.95 eq) was added dropwise to a suspension of 4-(dimethylamino)butyric acid hydrochloride (116 mg, 0.692 mmol, 2.0 eq) in anhydrous THF (2 mL), and anhydrous DMF (1 drop) at 0 °C. The suspension

was stirred at 0 °C for 15 min and then warmed to rt and stirred for a further 1 h. The solution was then cooled back down to 0 °C. 2^{vii}-(Adamantan-1^{viii}-yl)-*N*-{2^{v}-[2^{iv}-({3^{iii}-[4^{ii}-(5-amino-3-methylisoxazol-4-yl)phenyl]prop-2^{iii}-yn-1^{iii}-yl}oxy)ethoxy]ethyl}acetamide (**6.8**) (170 mg, 0.346 mmol, 1.0 eq) in anhydrous *N*-methyl-2-pyrrolidone (3 mL) was then added dropwise at 0 °C, and the reaction stirred for 21 h at 5 °C. The reaction was quenched by the addition of aqueous HCl (1.0 M, 30 mL), and washed with EtOAc (2 × 30 mL). The acidic aqueous extract was basified with aqueous NaOH (2.0 M) until pH 11, extracted with EtOAc (3 × 30 mL), washed with aqueous LiCl solution (0.5 M, 2 × 20 mL), and brine (20 mL), dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The crude product was purified using silica gel column chromatography (0–100% 12:2:1 EtOAc: EtOH: aqueous NH_4OH/EtOAc) to isolate the title compound as a yellow oil (134 mg, 64%); R_f 0.10 (50% 12:2:1 EtOAc: EtOH: aqueous NH_4OH/EtOAc); $\tilde{\nu}_{max}$ (thin film)/cm^{-1} 3344 (N-H, m br), 2905 (C-H, s), 1675 (C=O, s), 1096 (C-O, s); ^{1}H NMR (600 MHz, $CDCl_3$) δ_H 7.47–7.44 (2H, m, 3^{ii}-C*H*), 7.25–7.22 (2H, m, 2^{ii} C*H*), 5.88 (1H, t, *J* 5.2, 1^{vi}-N*H*), 4.43 (2H, s, 3^{iii}-CH_2), 3.76–3.73 (2H, m, 2^{iv}-CH_2), 3.69–3.66 (2H, m, 1^{iv}-CH_2), 3.57 (2H, t, *J* 5.2, 2^{v}-CH_2), 3.43 (2H, dt, *J* 5.2, 5.1, 1^{v}-CH_2), 2.50–2.48 (2H, m, 2^{ix}-C*H*), 2.43 (2H, t, *J* 5.5, 4^{ix}-CH_2), 2.25 (3H, s, 1^{i}-CH_3), 2.12 (6H, s, 5^{ix}-CH_3), 1.96–1.93 (3H, m, 3^{viii}-C*H*), 1.92 (2H, s, 2^{vii}-CH_2), 1.81–1.76 (2H, m, 3^{ix}-C*H*), 1.70–1.60 (6H, m, 4^{viii}-CH_2), 1.60 (6H, d, *J* 2.4, 2^{viii}-CH_2); ^{13}C NMR (151 MHz, $CDCl_3$) δ_C 171.5 (1^{ix}-C), 171.1 (1^{vii}-C), 160.0 (3-C), 158.1 (5-C), 132.1 (3^{ii}-C), 130.7 (1^{ii}-C), 128.8 (2^{ii}-C), 121.7 (4^{ii}-C), 107.3 (4-C), 86.3 (1^{iii}-C), 85.7 (2^{iii}-C), 70.3 (1^{iv}-C), 70.2 (2^{v}-C), 69.2 (2^{iv}-C), 59.7 (4^{ix}-C), 59.3 (3^{iii}-C), 51.9 (2^{vii}-C), 44.6 (5^{ix}-C), 42.7 (2^{viii}-C), 39.1 (1^{v}-C), 37.5 (2^{ix}-C), 36.9 (4^{viii}-C), 32.8 (1^{viii}-C), 28.8 (3^{viii}-C), 22.0 (3^{ix}-C), 11.6 (1^{i}-C); HRMS *m/z* (ESI$^+$) [Found: 605.3698, $C_{35}H_{49}O_5N_4$ ([M+H]$^+$) requires 605.3697]; LRMS *m/z* (ESI$^+$) 605 ([M+H]$^+$, 100%); HPLC Retention time 8.6 min, >99.0%.

8.3 General biology methods

8.3.1 Chemicals and cell culture reagents

Materials were obtained from the following suppliers:

Table 8.1. General reagents.

Material	Supplier
Acetic Acid	Sigma Aldrich
Acetonitrile	Sigma Aldrich
Agarose	Sigma Aldrich
Bacto agar	Fisher Scientific
β-Mercaptoethanol	Fisher Scientific
Bovine serum albumin (BSA)	Sigma Aldrich
Bromophenol blue	Fisher Scientific
Calcium chloride	Sigma Aldrich
Copper Sulfate ($CuSO_4$)	Sigma Aldrich
DMSO	Sigma Aldrich
DMEM (Dulbecco's modified eagle medium)	Gibco
Ethanol	Sigma Aldrich
Ethylenediaminetetraacetic acid (EDTA)	Sigma Aldrich
Foetal bovine serum (FBS)	Gibco
GlutaMAX™	Gibco
Glycerol	Sigma Aldrich
Glycine	Sigma Aldrich
Imidazole	Sigma Aldrich
InstantBlue® coomassie protein stain	Abcam
Lipofectamine 2000	Invitrogen
Methanol	Sigma Aldrich
MOPS SDS running buffer	Invitrogen
Nickel sulfate hexahydrate ($NiSO_4 \cdot 6H_2O$)	Sigma Aldrich
Opti-MEM™ Reduced Serum Medium	Gibco
Phosphate-buffered saline (PBS)	Gibco
Ponceau S	Sigma Aldrich
SOC media	New England Biolabs
Sodium chloride (NaCl)	Sigma Aldrich

Material	Supplier
Sodium dodecyl sulphate (SDS)	Fisher Scientific
SYBR Safe™	Invitrogen
TAMRA-PEG3-Azide	Sigma Aldrich
Tris(benzyltriazolylmethyl)amine (TBTA)	Sigma Aldrich
Tris(2-carboxyethyl)phosphine hydrochloride (TCEP)	Sigma Aldrich
Tris(hydroxymethyl)aminomethane hydrochloride) (Tris-HCl)	Thermo Fisher Scientific
Triton X-100	Thermo Fisher Scientific
Trypsin 0.5% EDTA	Gibco
Tryptone	Sigma Aldrich
Tween 20	Fisher Scientific
Urea	Fisher Scientific
Yeast Extract	Sigma Aldrich
Zinc	Sigma Aldrich

8.3.2 Bacterial cell culture

Standard sterile techniques were followed thoroughly. Media and equipment were sterilised by autoclaving at 121 °C for 20 min. Solutions of IPTG, antibiotics and other labile compounds were sterilised by filtration through 0.22 µm filters (Millex®GP, Merck Millipore).

Kanamycin was prepared as a stock solution in water of 50 mg/mL. Final concentrations of antibiotics, unless otherwise stated, were; Kanamycin 50 µg/mL.

Incubation of bacteria was performed in a New Brunswick G-25 or Innova-44 Shaker Incubator.

Optical density at 600 nm (OD_{600}) was measured with a Novaspec III visible spectrophotometer.

Escherichia coli (*E. coli*) strains of the following genotypes were used:

- **BL21 (DE3) Gold**: F^- ompT hsdSB (rB^-mB^-) gal dcm (DE3). DE3 denotes a chromosomal copy of the T7 RNA polymerase gene. – For protein expression
- **NEB 5-alpha** (New England Biolabs Inc.); genotype: *fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17* – For plasmid amplification

Cells were grown in 2×TY broth, according to the experimental requirements.

8.3.3 Mammalian cell culture

HEK293T cells were cultured in DMEM supplemented with 10% FBS and GlutaMAX in a standard humidified incubator for mammalian tissue culture maintained at 37 °C, 21% O_2 and 5% CO_2.

Cells were passaged 2–3 per week. In each case, the medium was removed, and the cells washed with rt PBS. Cells were detached by incubating with 37 °C Trypsin-EDTA solution for 2–4 min. The trypsin was quenched by the addition of fresh 37 °C DMEM supplemented with 10% FBS and GlutaMAX. Cells were counted using a Neubauer haemocytometer chamber.

8.4 Protein expression and purification

8.4.1 Cloning

8.4.1.1 DNA amplification - PCR

Amplification was carried out in a thermal cycler (Labnet Multigene II Personal Thermal Cycler). An initial denaturation step at 98 °C for 3 min was followed by 30 cycles of denaturing at 98 °C for 10 sec, annealing at 63 °C for 30 sec, and extension at 72 °C for 30 sec was carried out. A final extension of 2 min at 72 °C was followed by the resulting samples being cooled to 4 °C.

Table 8.2. Components for DNA amplification.

Component	Volume per 50 µL Reaction (µL)	Final Concentration
5× Q5 Reaction Buffer	10	1×
10 mM dNTPs	1	200 µM
10 µM Forward Primer	2.5	0.5 µM
10 µM Reverse Primer	2.5	0.5 µM
Template DNA	2	< 1,000 ng
Q5 High-Fidelity DNA Polymerase	0.5	0.02 U/µL
Nuclease-Free Water	31.5	n.a.

8.4.1.2 DNA purification - Agarose gel electrophoresis

DNA samples were analysed and purified by 2% agarose gel electrophoresis prepared with SYBR® safe DNA gel stain (Invitrogen). DNA samples were loaded after mixing with 6 × loading buffer and run in a TAE buffer (**Table 8.3**). The gels were run at 100 V until the desired fragments had separated. Gels were visualised using a UV transilluminator (Gel logic 200). A GeneRuler DNA Ladder, 1 kb (Thermo Scientific) was used molecular weight reference.

Digested DNA was purified using a GeneJET Gel Extraction Kit (Thermo Scientific) according to manufacturer's instructions. Plasmids were purified using a GeneJET Plasmid Miniprep Kit (Thermo Scientific) according to manufacturer's instructions. DNA concentration was estimated by measuring absorbance at 260 nm using a Nanodrop® ND-1000 spectrophotometer (Nanodrop® Technologies Inc.).

Table 8.3. Buffer compositions for DNA purification.

Buffer	**Reagent**	**Amount required**
50× TAE Running Buffer (per 1 L)	Tris·HCl	242.0 g
	Acetic acid (glacial)	57.0 mL
	EDTA (0.5 M)	100 mL
6× DNA Loading Buffer (per 10 mL)	Milli-Q H_2O	6.7 mL
	Bromophenol blue	25.0 mg
	Glycerol	3.30 mL
	Xylene cyanol FF	25.0 mg
DNA electrophoresis gel (per gel)	Agarose	1.0 g
	TAE buffer (1×)	50.0 mL
	SYBR® Safe DNA gel stain	5.0 µL

8.4.1.3 Gibson Assembly

Cloning was performed using a one-step isothermal in vitro recombination reaction consisting of T5 exonuclease, Phusion DNA polymerase and Taq DNA ligase. Linear vectors (~0.03 pmol) and gene inserts (~0.1 pmol) with complementary regions were incubated at 50 °C for 15-60 min in a 1:3 ratio

with Gibson Assembly master mix (NEB): 20 µL reaction volume, 10 µL 2 × GA master mix. 2 µL of the reaction mixture was used directly to transform NEB 5-alpha (#C2987) competent cells.

8.4.1.4 DNA Sequencing

Plasmid sequencing was performed by Source BioScience Sequencing facility in the Department of Biochemistry (University of Oxford). T7 promoter and terminator sequence primers were used, on a 3730xl DNA Analyser-Titania.

8.4.2 Competent cells

Competent cells were thawed slowly on ice for 20 min prior to use. 1–5 µL containing 50–100 ng of plasmid DNA was added to 20 µL of the competent cells. The tubes were placed on ice for 30 min before incubated at 42 °C for 30 sec. The tubes were returned to ice for 5 min. Gently 0.5 mL of SOC outgrowth medium (New England Biolabs Inc.) was added to each tube. The tubes were incubated at 37 °C for 1 h prior to 100-400 µL of the transformation mixture being spread over agar plates containing kanamycin (50 µg/mL) and incubated overnight, under aseptic conditions.

8.4.3 Fermentation conditions for human bromodomain containing proteins

For the expression and purification of the human bromodomain containing proteins the published procedure from Filippakopoulos *et al.* was followed.[147] Single colonies of transformed *E.Coli* BL21 (DE3) were transferred, under aseptic conditions, to 100 mL of 2× TY medium containing kanamycin (50 µg/mL) (**Table 8.4**). Cultures were incubated over night at 37 °C and 180 rpm. From the overnight night preculture 10 mL was transferred to 1 L of 2× TY medium containing kanamycin (50 µg/mL), and incubated at 37 °C and 180 rpm until an $O.D_{600}$ of 0.8 was reached. $O.D_{600}$ readings were taken in 1.6 mL cuvettes with 2× dilution in 2× TY medium against a reference of 2× TY medium, using a Novaspec® II spectrophotometer measuring at 600 nm. The cultures were then

cooled to 16 °C, 60 rpm, and IPTG (final concentration of 0.1 mM) was added. The cultures were then further incubated overnight at 16 °C and 180 rpm. Cells were pelleted by centrifugation (11,325 ×g, 4 °C, 10 min), and the pellets collected, combined, and stored at −80 °C.

Table 8.4. Media composition for bacterial cell culture.

Media	Reagent	Amount per litre (g)
2× TY	Tryptone	16.0
	Yeast Extract	10.0
	NaCl	5.0

8.4.4 Purification conditions for human bromodomain containing proteins

Cell pellets were re-suspended in extraction buffer (100 mL per 20 g cell mass) with 250 Units (U) of Benzonase at 4 °C (**Table 8.5**). The cells were lysed by sonication (60 × 5 sec burst, 10 sec pause), followed by the addition of polyethyleneimine (PEI, 0.15% v/v) to precipitate DNA. The lysate was incubated over ice for 15 min. The lysate was centrifuged (25,000 ×g, 4 °C, 15 min) to pellet cell debris, the resulting supernatant was decanted, filtered through a 0.45 µm filter and applied to a purification column. The protein was initially purified at 4 °C by immobilised metal affinity chromatography (IMAC) with a HisTrap™ (5 mL) column (GE Healthcare) rinsed with 10 CV of Milli-Q water, then charged with 4 CV of 100 mM $NiSO_4$, and subsequently equilibrated with 10 CV of binding buffer (supplemented with 10 mM β-ME). The column was then loaded with the lysate at a rate of 1 mL/min. The column was washed with a step gradient of 10 CV 9% elution buffer (45 mM imidazole), and 10 CV 12% elution buffer (60 mM imidazole). The polyhistidine-tagged protein of interest was batch eluted with 4 CV of 60% elution buffer (300 mM imidazole) followed by 2 CV of 100% elution buffer (500 mM imidazole). Eluted protein was collected by fractionation in 2 mL volumes. Fractions containing the highest levels of pure protein, as determined by UV trace and SDS-PAGE gels, were collected, combined, and concentrated using 5k molecular weight cut-off concentrators (GE Healthcare).

Further purification of the desired protein was achieved using size exclusion chromatography (SEC). Concentrated IMAC fractions (2 mL) were loaded onto a 120 mL Superdex 75 size exclusion column (Amersham) equilibrated with 150 mL of gel filtration buffer, via a 2 mL injection loop. The protein was eluted with a further 150 mL of gel filtration buffer, and collected by fractionation in 2 mL volumes. Fractions containing protein were analysed by SDS-PAGE gel, and the fractions containing the highest levels of pure protein were collected, combined, and concentrated using 5k molecular weight cut-off concentrators (GE Healthcare). Protein concentration was determined using a Nanodrop® ND-1000 spectrophotometer (Nanodrop® Technologies Inc.) measuring absorbance at 280 nm. Protein molecular weight and molar extinction coefficient were calculated using ProtParam on the ExPASy Bioinformatics Resource Portal. Proteins were stored at −80 °C.

Table 8.5. Buffer compositions for protein purification.

Buffer[α]	**Reagent**	**Concentration**	**Amount per 1 L**
Extraction Buffer (pH = 7.6)	HEPES	50 mM	11.9 g
	NaCl	500 mM	29.2 g
	Glycerol	10% (v/v)	100 mL
	Sigma*FAST*[β]	1 tab/ 100 mL	10 tabs
Binding Buffer (pH = 7.6)	HEPES	50 mM	11.9 g
	NaCl	500 mM	29.2 g
	Glycerol	10% (v/v)	100 mL
Elution Buffer (pH = 7.6)	HEPES	50 mM	11.9 g
	NaCl	500 mM	29.2 g
	Glycerol	10% (v/v)	100 mL
	Imidazole	500 mM	34.0 g
Re-Charge	$NiSO_4 \cdot 6H_2O$	100 mM	26.3 g
Strip Buffer (pH = 7.6)	HEPES	50 mM	11.9 g
	NaCl	500 mM	29.2 g
	EDTA	50 mM	18.6 g
Gel Filtration Buffer	HEPES	50 mM	11.9 g
	NaCl	500 mM	29.2 g

[α] All buffers were filtered through a 0.22 µm filter and degassed

[β] Sigma*FAST* Protease Inhibitor Cocktail Tablet, EDTA Free

8.4.5 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

Protein purity was analysed by SDS-PAGE gel electrophoresis. Gels were prepared using 70 mm × 100 mm glass plates, with 0.75 mm spacers. TEMED and freshly prepared ammonium persulfate were added just prior to pouring the gels (**Table 8.6**). The resolving gel was cast with the addition of a separate layer of isopropyl alcohol (to ensure a level surface). Once the resolving layer had polymerized, isopropyl alcohol was removed, and the stacking gel was cast. For protein denaturation samples were prepared by mixing with sample loading buffer (4×) and incubation at 100 °C for 10 min (**Table 8.7**). A colour protein standard broad range (NEB) molecular weight ladder was used a reference. Gels were run on a mini-PROTEAN Tetra Electrophoresis System (Bio-Rad) at a constant potential of 100 V. Following electrophoresis, gels were stained for 30 min with InstantBlue™ Coomasie® stain (Abcam) and afterwards destained for 5 h.

Table 8.6. Gel compositions for a single gel.

Reagent	Resolving Gel (16%)	Stacking Gel (4%)
Tris-HCl (1.5 M, pH 8.8)	1.125 mL	-
Tris-HCl (0.5 M, pH 6.8)	-	0.65 mL
Milli-Q Water	1.05 mL	1.45 mL
30% (w/v) Acrylamide	2.56 mL	0.35 mL
SDS (10% w/v)	50 µL	25.0 µL
10 % APS	50 µL	25.0 µL
TEMED	4 µL	2.5 µL

Table 8.7. Buffer compositions for SDS-PAGE.

Buffer	Reagent	Composition	Amount per 1 L
SDS-PAGE Running Buffer	Tris-HCl	-	30.0 g
	Glycine	-	144 g
	SDS	-	10.0 g
SDS-PAGE Sample Loading Buffer (pH 6.8)	Tris-HCl	0.5 M	10.0 mL
	Bromophenol Blue	0.2% w/v	0.20 g
	SDS	10% w/v	20.0 mL
	Glycerol	-	12.0 mL
	B-mercaptoethanol	-	5.0 mL
SDS-PAGE Stain	Coomassie® Brilliant Blue	0.25%	2.5 mL
	Acetic Acid	10%	100 mL
	Methanol	30%	300 mL
SDS-PAGE Destain	Acetic Acid	10%	100 mL
	Methanol	40%	400 mL

8.5 Mammalian protein expression

8.5.1 DNA transfection

HEK293T cells at ~40% confluency were transfected using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. The media was changed after 24 h, and the cells left to grow for a further 24 h before being used at a confluency of ~90%. For transfection with either BRD4(1)^{WT} or BRD4(1)^{L94C} these ratios were used:

Table 8.8. Quantities of Lipofectamine 2000 and DNA used for transfection of HEK293T cells.

Mixture	Component	10 cm Dish	12 Well Plate
Diluted Lipofectamine	Opti-MEM	500 µL	72 µL
	Lipofectamine 2000	45 µL	3 µL
Diluted DNA	Opti-MEM	500 µL	74 µL
	DNA (~1 µg/mL)	15 µL	1 µL

8.5.2 Cell harvesting and lysis

To harvest adherent HEK293T cells, the medium was removed, and the cells washed with rt PBS. Cells were detached by incubating with 37 °C Trypsin-EDTA solution for 2–4 min. The trypsin was quenched by the addition of fresh 37 °C DMEM supplemented with 10% FBS and GlutaMAX. The detached cells were then pelleted (500 ×g, rt, 5 min) and washed with PBS, before being pelleted again (500 ×g, rt, 5 min).

Cell pellets were lysed by resuspending the pellet in lysis buffer (1% Triton X-100, 0.1% SDS, and 1:200 Protease inhibitor cocktail) and incubating at 4 °C for 20–30 min. The lysate was centrifuged (13,000 ×g, 4 °C, 10 min) to pellet cell debris and the resulting supernatant was transferred to fresh Eppendorf tubes.

8.5.3 Determination of total protein concentration

Total protein concentration of the lysate was determined by a BCA protein assay (Thermo Scientific) and the absorbance at 562 nm measured and plotted using a NanoDrop One. Cell lysates were diluted to 2 mg/mL with lysis buffer, unless otherwise stated.

8.5.4 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

Cell lysate proteins were separated by SDS-PAGE gel electrophoresis. 4× Loading buffer (4× LDS sample buffer and 2% β-mercaptoethanol) was added to the diluted protein samples and then heated at 95 °C for 5 min. Protein samples (15–20 µL) were loaded into NuPAGE 4-12% Bis-Tris Gels and electrophoresed for 70–80 min at 180V in MOPS buffer. The Precision Plus Protein Kaleidoscope Prestained Protein Standards (Bio-rad) (5 µL) was used as a molecular weight reference.

8.5.5 Western blotting

Proteins in the electrophoresed gels were transferred onto a PVDF membrane for 1 h at 100 V, at 4 °C. The primary antibodies were diluted (**Table 8.9**) in 5% milk in TBS-Tween (TBS-T, 0.1% Tween). The membrane was incubated with the primary antibodies for 16 h at 4 °C. The membrane was washed with TBS-T (3 × 5 min), and then treated with the appropriate secondary antibody diluted (**Table 8.10**) in 5% milk in TBS-Tween (TBS-T, 0.1% Tween), and then incubated for 1 h at rt. The membrane was washed with TBS-T (3 × 5 min). The membrane was then covered with ECL Prime Western Blotting Detection Reagent for 1−3 min before being imaged on a Bio-rad molecular imager. Bio-rad Image Lab 6.1 was used to process the data.

Table 8.9. Primary antibodies used for Western blotting.

Name	Species	Dilution	Supplier
Strep-tag	Mouse	1:1000	Qiagen/34850
Vinculin	Rabbit	1:10,000	Abcam/129002

Table 8.10. Secondary antibodies used for Western blotting.

Name	Dilution	Supplier
Anti-Mouse IgG Peroxidase	1:15,000	Sigma/A4416
Anti-Rabbit IgG Peroxidase	1:15,000	Sigma/A6667

8.6 Assay procedures

8.6.1 Differential scanning fluorimetry (DSF)

The differential scanning fluorimetry or thermal stability assays were carried out following the protocol described by Niesen *et al.*[146] and data was analysed as described previously.[146] Thermal melting experiments were carried out using the Mx3005p real-time PCR machine (Agilent) and employing a protein concentration of 2 µM for BRD4(1)^{WT} and BRD4(1)^{L94C}. Buffer conditions were 50 mM HEPES buffer, pH 7.5, 500 mM NaCl, 0.5% DMSO and a 1:1000 dilution of SYPROR Orange (Invitrogen, CA) for BRD4(1)^{WT} and BRD4(1)^{L94C}. The 96-well PCR plates were sealed and centrifuged

for 2 min at 1000 rpm at 25 °C before measurement. Temperature was raised with a step of 1 °C per 30 s from 25 °C to 85 °C, and fluorescence readings were taken at each interval. Excitation and emission filters for the SYPROR Orange dye were set to 465 and 590 nm, respectively.

8.6.2 Isothermal titration calorimetry (ITC)

All calorimetric experiments were performed on a MicroCal PEAQ-ITC Automated (Malvern) and analysed with the MicroCal PEAQ-ITC Analysis software (Malvern 1.1.0.1262) using a single binding site model. The first data point was excluded from the analysis. The proteins were dialysed at 4 °C overnight in a Slide-A-Lyzer MINI Dialysis Device (2000 MWCO; Thermo Scientific Life Technologies) into 50 mM HEPES, 150 mM NaCl containing 0.2% DMSO; pH 7.4. Proteins were centrifuged to remove aggregates (3 min, 3000 rpm, 4 °C). Protein concentrations were determined by measuring the absorbance at 280 nm using a NanoDrop Lite spectrophotometer (NanodropR Technologies Inc.). Small molecule ligands were dissolved as DMSO stock solution and diluted to the required concentration using dialysis buffer. The cell was stirred at 750 rpm, reference power set to 5 µcal/s and temperature held at 25 °C. After an initial delay of 60 s, 19 x 2 µL injections (first injection 0.4 µL) were performed with a spacing of 180 s. Small molecule solutions in the calorimetric cell (420 µL, 20 µM) were titrated with the protein solutions in the syringe (150 µL, 160 µM).

8.6.3 Protein LC-MS assays

8.6.3.1 In-house

Assays were performed in 1.5 mL Eppendorf tubes at 37 °C and gently shaken, unless otherwise stated. Experiments were performed in a protein LC-MS assay buffer, with 10 µM protein, and 50 µM of the desired compound with a final DMSO concentration of 0.5% (**Table 8.11**). Time points were taken at 0, 1 and 4 h, unless otherwise stated, where a 20 µL sample of the reaction mix was

taken, diluted with 100 µL Milli-Q water, and flash frozen. Samples were thawed immediately prior to LC-MS analysis on a Waters Xevo G2-S system (Waters) fitted with a ProSwift™ RP-2H PSDVB Reversed Phase HPLC Column (Thermofisher). Data was analysed and processed using MaxEnt. Covalent labelling was calculated using the intensities of the unlabelled and labelled protein (**Equation 8.1**). Figures were made using GraphPad Prism 7.0.

Table 8.11. Buffer composition for protein LC-MS.

Buffer	Reagent	Composition	Amount per 1 L
Protein LC-MS Assay Buffer (pH 7.4)	HEPES	25 mM	6.0 g
	NaCl	100 mM	5.8 g

Equation 8.1.

$$Protein\ Labelling = \frac{Intensity\ of\ Labelled\ Peak}{(Intensity\ of\ Unlabelled\ Peak + \ Intensity\ of\ Labelled\ Peak)} \times 100$$

8.6.3.2 GSK

Chloroacetamide (50 µM) or acrylamide (200 µM) fragments were incubated with BRD4(1)^{L94C} (1 µM) in buffer (25 mM HEPES, 50 mM NaCl, pH 7.5) at 4 °C for 24 h. Intact protein masses were recorded by LC-MS using an Agilent G6224 time-of-flight (ToF) Accurate Mass Series mass spectrometer, interfaced with an Agilent 1200 series liquid chromatography and sample handling system. The protein sample was injected using an Agilent 1200 series AutoSampler (Model No. G1367B) with a 10 µL injection volume and maintained at a temperature of 10 °C. Chromatography was carried out on an Agilent Bio-HPLC PLRP-S (1000Å, 5 µm × 50 mm × 1.0 mm, PL1312-1502) reverse phase HPLC column at 70 °C. Using an Agilent 1200 series binary pump system (Model No. G1312B). Data acquisition was carried out in 2 GHz Extended Dynamic range mode. Spectra were processed using Mass Hunter Qualitative Analysis™ B06.00 (Agilent) software with the Maximum Entropy method employed. The total ion chromatograms (TIC) were extracted (region containing protein) and the summed scans were deconvoluted (using a maximum entropy algorithm).[219]

8.6.4 Dose response AlphaScreen™

The peptides (**Table 8.12**) were supplied at >95% purity by either Cambridge Biochemicals or GenScript and were stored as either 1 mM or 100 µM stocks in Milli-Q water at –80 °C.

Table 8.12. Amino acid sequences of peptides used in AlphaScreen™.

Peptide	Amino Acid Sequence
$H4KAc_4$	H_2N-YSGRGK(Ac)GGK(Ac)GLGK(Ac)GGAK(Ac)RHRK-$CONH_2$

The assay buffer (**Table 8.13**, pH 7.6) was filtered through a 0.22 µm filter before use. The buffer is stored at 4 °C as a 5× solution without bovine serum albumin (BSA), which was added on the day of use. The compounds, proteins, and peptides were dispensed into a ProxiPlate-384 Plus (Perkin Elmer) using a Thermo-Fisher electronic multi-channel pipette. For the incubation steps, the plates were sealed, shaken for 10 seconds at 600 rpm on a plate oscillator, and incubated at room temperature in darkness for 4 h. Ni^{2+} chelate acceptor and streptavidin donor beads were prepared as a mixture in a 1:300 dilution (0.007 mg/mL final assay concentration, added to the plates and incubated for a further 1h at room temperature. Plates were read using a SynergyTM 2 MultiMode Microplate Reader using the inbuilt AlphaScreen™ 384 ProxiPlate function: excitation 680 nm, 0.18 sec; emission 570 nm, 0.37 sec.

Table 8.13. Composition of assay buffer used in AlphaScreen™ experiments, adjusted to pH 7.6 using an aqueous NaOH solution.

Reagent	Composition	Mass (mg) for 50 mL
HEPES	25 mM	297.9
NaCl	100 mM	292.2
CHAPS	0.05% (w/v)	25
BSA	0.1% (w/v)	50

50 mM DMSO stock solutions of the compounds were serially diluted, and dispensed into wells (5 µL). The protein-peptide mixes (7 µL) for each combination were prepared following the specified final assay concentration (**Table 8.14**), and dispensed into wells. Assay beads (8 µL) were added to

attain a final assay volume of 20 µL. Dose response curves were obtained for each compound in triplicate, with serial 1:2 dilutions. Data was processed by fitting a four-parameter equation (**Equation 8.2**) to calculate IC_{50} values using Prism software.

Table 8.14. Final assay concentrations of each peptide-protein combination.

Protein	Peptide	Protein FAC (nM)	Peptide FAC (nM)
BRD4(1)^{WT}	H4KAc$_4$	10	4
BRD4(1)^{L94C}	H4KAc$_4$	10	4

Equation 8.2.

$$\text{Response} = \text{Bottom} + \frac{\text{Top} - \text{Bottom}}{1 + 10^{((\log IC50 - X) \times \text{HillSlope})}}$$

X = Log of concentration

8.6.5 Click probe efficiency

Assays were performed in 1.5 mL Eppendorf tubes at room temperature, in the dark, and gently shaken, unless otherwise stated. The desired click probe (50 µM) was incubated with TAMRA-PEG3-Azide (200 µM), TCEP·HCl (1 mM), TBTA (100 µM), and $CuSO_4$ (1 mM) in protein LC-MS buffer. Time points were taken at 0 and 1 h, unless otherwise stated, where a 25 µL sample of the reaction mix was taken, diluted with 25 µL Milli-Q water, and flash frozen. Samples were thawed immediately prior to LC-MS analysis on a Waters LCT Premier bench-top orthogonal acceleration time-of-flight LC-MS system. Data was analysed and processed using MassLynx. Figures were made using GraphPad Prism 7.0.

8.6.6 Fluorescent labelling of purified BRD4(1)^{L94C}

Assays were performed in 1.5 mL Eppendorf tubes at 37 °C, and gently shaken, unless otherwise stated. The desired click probe (50 µM) was incubated with protein (10 µM) in protein LC-MS buffer for 4 h, after which TAMRA-PEG3-Azide (200 µM), TCEP·HCl (1 mM), TBTA (100 µM), and $CuSO_4$ (1 mM) were added and incubated for a further 1 h at room temperature in the dark. A 25 µL sample of the reaction mix was taken and 8 µL of Lamelli buffer added, and the samples denatured by heating at 95 °C for 10 min. Samples were separated by gel electrophoresis on NuPAGE™ 4–12%, Bis-Tris precast polyacrylamide gels (Thermofisher). Following electrophoresis, gels were stained for 15 min with InstantBlue™ Coomasie® stain (Abcam), destained, and imaged using Bio-rad molecular imager. Bio-rad Image Lab 6.1 was used to process the data.

8.6.7 In-cell target engagement

8.6.7.1 In-cell protein labelling

HEK293T cells at ~90% confluency were incubated for 1 h with a DMSO control (0.5% DMSO final concentration) or a click probe (10 µM, 0.5% DMSO final concentration) diluted in Opti-MEM media.

Cells were then harvested and lysed using the previously described protocol in **section 8.5.2**. The resulting cell lysates were diluted to 2 mg/mL using the previously described protocol in **section 8.5.3**.

Cell lysates (50 µL) were then incubated with TAMRA-PEG3-Azide (200 µM), TCEP·HCl (2 mM), TBTA (200 µM), and $CuSO_4$ (2 mM), for 1 h at rt in the dark. Samples were then prepared and separated by electrophoresis using the previously described protocol in **section 8.5.4**.

Gels were imaged for fluorescence using Bio-rad molecular imager and then stained for 15 min with InstantBlue™ Coomasie® stain (Abcam), destained, and imaged using Bio-rad molecular imager.

Typically, the very bottom of the gel was removed prior to imaging due to residual unreacted TAMRA-PEG3-Azide, which would cause oversaturation and poor image quality.

8.6.7.2 OXFBD04 Competition assay

HEK293T cells at ~90% confluency were incubated for 30 min with a DMSO control or varying concentrations of OXFBD04 diluted in Opti-MEM media. Click probes (10 µM) were then added to the media (0.5% DMSO concentration) and incubated for a further 1 h. Cells were then harvested, processed, and analysed using the previously described protocol in **section 8.6.7.1.**

8.6.7.3 Immunoprecipitation

HEK293T cells at ~90% confluency were incubated for 1 h with a DMSO control (0.5% DMSO final concentration) or a click probe (10 µM, 0.5% DMSO final concentration) diluted in Opti-MEM media. Cells were then harvested and lysed using the previously described protocol in **section 8.5.2**. Samples of the resulting cell lysates were diluted to make a 30 µL input sample at 2 mg/mL, using the previously described protocol in **section 8.5.3**. The remainder of the cell lysates were diluted to 500 µL using 4 °C IP buffer (0.1% SDS, 1% Triton X-100, 10 mM Tris·HCl pH 8, 200 mM NaCl, 2 mM EDTA, 1× Protease inhibitor cocktail) and 5 µL of Strep-tag antibody (Qiagen/34850) added to each sample and incubated at 4 °C for 16 h. After this time, 20 µL of Protein A Dynabeads (Invitrogen) were added and incubated at 4 °C for a further 5 h. The beads were then washed with RIPA buffer (3 × 1 mL), and wash buffer (10 mM Tris·HCl pH 8, 50 mM NaCl) (1 mL), before being pelleted (1000 ×g, 2 min). The beads were resuspended in lysis buffer (15 µL) and incubated with TAMRA-PEG3-Azide (200 µM), TCEP·HCl (2 mM), TBTA (200 µM), and $CuSO_4$ (2 mM) for 1 h at rt in the dark.

4× Loading buffer (4× LDS sample buffer and 2% β-mercaptoethanol) was added to the IP samples and then heated at 95 °C for 5 min. The beads were pelleted and the IP supernatant transferred to

a fresh Eppendorf tube. 4× Loading buffer (4× LDS sample buffer and 2% β-mercaptoethanol) was added to the input samples and then heated at 95 °C for 5 min. Samples were then separated by electrophoresis using the previously described protocol in **section 8.5.4**. The gels were then imaged and stained using the previously described protocol in **section 8.6.7.1.**

8.6.8 Protein degradation

HEK293T cells expressing either BRD4(1)^{WT} or BRD4(1)^{L94C} at ~90% confluency were incubated for 16 h with a DMSO control (0.5% DMSO final concentration) or a HyTag probe (10 µM, 0.5% DMSO final concentration) diluted in DMEM supplemented with 10% FBS and GlutaMAX. Cells were washed with rt PBS and were detached by incubating with 37 °C Trypsin-EDTA solution for 2–4 min. The trypsin was quenched by the addition of fresh 37 °C DMEM supplemented with 10% FBS and GlutaMAX. The cells were then submitted to flow cytometry to analyse protein degradation. Cells were analysed using an Attune NxT Flow Cytometer (Thermo Fisher). Cells were identified and gated using Forward Scatter (FSC) and Side Scatter (SCC) and further gated to select singlets using a FSC area v FSC height. Singlets were then assessed for expression of mCherry using a 561nm laser and associated emitted light was collected using a 620/15nm bandpass filter. eGFP expression was measured using a 488nm laser and emitted light was collected through a 530/30nm band pass filter. Data was analysed using FlowJo software.

8.7 Protein X-ray crystallography

To obtain an apo- or co-crystal structure, the JCSG-plus MD1-37 (Molecular Dynamics) was screened to find appropriate conditions for crystallization, incubating at room temperature or 4 °C. For co-crystal structures purified aliquots of BRD4(1)^{L94C} (9.0 mg/mL) were spiked with a DMSO stock of OXFBD04 (50 mM) to a final concentration of 1 mM. A Phoenix crystallography system by

Art Robbins Instruments was used to dispense the protein samples as well as the additives in 96-well IntelliPlates. For apo BRD4(1)^{L94C} no crystals were observed at either room temperature or 4 °C after 2 months. For the co-crystal structure, many small crystals were observed in well G11.1 after 6 weeks in the presence of 0.15 M Ammonium sulfate, 0.1 M Bis-Tris, pH 5.5, no precipitant. To the wells containing crystals to be fished, one drop of a solution containing 75% (*v/v*) of the reservoir solution and 25% (*v/v*) of glycerol were added. Crystals were mounted with a nylon loop (loop size was chosen according to crystal size) and snap frozen with liquid nitrogen. Crystals were shot at the Diamond Light Source X-ray synchrotron by Dr Michael McDonough and diffracted to a resolution of 1.9 Å to generate a co-crystal structure of OXFBD04 bound to BRD4(1)^{L94C}. The resulting X-ray crystal structure was solved *via* molecular replacement and multiple rounds of refinement using Coot/WinCoot CCP4 and Phenix (version 1.17) software, by Dr Michael McDonough.

8.8 Computational Studies

8.8.1 Molecular Dynamics (MD) Simulations

Performed by Dr Joseph Bluck, the protein co-ordinates were taken from the crystal structure BRD4(1) (PDB ID: 3SVG). The protein was parameterized using the AMBER99SB-IDLN forcefield.[220] The L94C mutant version of BRD4(1) was generated using the mutation wizard in PyMOL (Version 1.8). All crystallographic water molecules were retained and used the TIP3P (TIP 3-Point) water model.[221] The system was solvated within a dodecahedral box, with a minimum boundary distance of 1.2 nm. Water molecules were substituted with a sodium ion to neutralize the net charge and to maintain an overall salt concentration of 150 nM sodium chloride.

The systems were subject to energy minimization using the steepest decent algorithm, with a maximum force cut off of 100 kJ mol^{-1} nm^{-1}. They then continued equilibration under the NVT and

NPT ensembles for a total of 200 ps. Simulations were then carried out for 50 ns using GROMACS 2016.4, in triplicate.[222]

8.8.2 Ligand Docking Studies

Docking was performed on a holo structure of BRD4(1) (PDB ID: 6CZV). All organic molecules and crystallographic waters outside of the binding site were subsequently removed before the protein structure was checked using the WHAT-IF web-interface.[223] The protein was protonated at pH 7.4 and the water network optimised using the MOE software package (v. 2018.2). Finally the structure was converted to the required PDBQT file type using AutoDockTools within the MGLTools package (v. 1.5.6). Ligand models were protonated and given an initial geometry using cxcalc from ChemAxon (https://www.chemaxon.com) (v. 16.7.11.0), before conversion to the PDBQT file type using the prepare_ligand4.py script available within the MGLTools package.[224,225] Molecular docking was performed using AutoDock4,[225] with 2,500,000 energy evaluations and 100 poses generated. Docked ligands were clustered using a 2 Å cut-off and underwent manual inspection.

References

1 S. M. Paul, D. S. Mytelka, C. T. Dunwiddie, C. C. Persinger, B. H. Munos, S. R. Lindborg and A. L. Schacht, *Nature Reviews Drug Discovery 2010 9:3*, 2010, **9**, 203–214.

2 M. E. Bunnage, *Nature Chemical Biology 2011 7:6*, 2011, **7**, 335–339.

3 J. W. Scannell, A. Blanckley, H. Boldon and B. Warrington, *Nature Reviews Drug Discovery 2012 11:3*, 2012, **11**, 191–200.

4 M. Ringel, P. Tollman, G. Hersch and U. Schulze, *Nature Reviews Drug Discovery 2013 12:12*, 2013, **12**, 901–902.

5 E. May, S. Shah, W. Miranda, M. Joao Cruz and K. Taylor, *Nurturing growth Measuring the return from pharmaceutical innovation 2021 Contents*, 2022.

6 A. D. Hingorani, V. Kuan, C. Finan, F. A. Kruger, A. Gaulton, S. Chopade, R. Sofat, R. J. MacAllister, J. P. Overington, H. Hemingway, S. Denaxas, D. Prieto and J. P. Casas, *Scientific Reports 2019 9:1*, 2019, **9**, 1–25.

7 M. Hay, D. W. Thomas, J. L. Craighead, C. Economides and J. Rosenthal, *Nature Biotechnology 2014 32:1*, 2014, **32**, 40–51.

8 H. Dowden and J. Munro, *Nat Rev Drug Discov*, 2019, **18**, 495–496.

9 R. M. Plenge, E. M. Scolnick and D. Altshuler, *Nature Reviews Drug Discovery 2013 12:8*, 2013, **12**, 581–594.

10 R. K. Harrison, *Nature Reviews Drug Discovery*, 2016, **15**, 817–818.

11 P. Morgan, P. H. van der Graaf, J. Arrowsmith, D. E. Feltner, K. S. Drummond, C. D. Wegner and S. D. A. Street, *Drug Discovery Today*, 2012, **17**, 419–424.

12 R. F. Sweis, *ACS Medicinal Chemistry Letters*, 2015, **6**, 618–621.

13 C. H. Emmerich, L. M. Gamboa, M. C. J. Hofmann, M. Bonin-Andresen, O. Arbach, P. Schendel, B. Gerlach, K. Hempel, A. Bespalov, U. Dirnagl and M. J. Parnham, *Nature Reviews Drug Discovery 2020 20:1*, 2020, **20**, 64–81.

14 D. Cook, D. Brown, R. Alexander, R. March, P. Morgan, G. Satterthwaite and M. N. Pangalos, *Nature Reviews Drug Discovery 2014 13:6*, 2014, **13**, 419–431.

15 P. Morgan, D. G. Brown, S. Lennard, M. J. Anderton, J. C. Barrett, U. Eriksson, M. Fidock, B. Hamrén, A. Johnson, R. E. March, J. Matcham, J. Mettetal, D. J. Nicholls, S. Platz, S. Rees, M. A. Snowden and M. N. Pangalos, *Nature Reviews Drug Discovery 2018 17:3*, 2018, **17**, 167–181.

16 J. D. Moore, *Drug Discovery Today*, 2015, **20**, 450–457.

17 W. A. Weiss, S. S. Taylor and K. M. Shokat, *Nature Chemical Biology 2007 3:12*, 2007, **3**, 739–744.

18 *Created with BioRender.com*.

19 D. R. Spring, *Chemical Society Reviews*, 2005, **34**, 472–482.

20 C. H. Arrowsmith, J. E. Audia, C. Austin, J. Baell, J. Bennett, J. Blagg, C. Bountra, P. E. Brennan, P. J. Brown, M. E. Bunnage, C. Buser-Doepner, R. M. Campbell, A. J. Carter, P. Cohen, R. A. Copeland, B. Cravatt, J. L. Dahlin, D. Dhanak, A. M. Edwards, S. v. Frye, N. Gray, C. E. Grimshaw, D. Hepworth, T. Howe, K. V. M. Huber, J. Jin, S. Knapp, J. D. Kotz, R. G. Kruger, D. Lowe, M. M. Mader, B. Marsden, A. Mueller-Fahrnow, S. Müller, R. C. O'Hagan, J. P. Overington, D. R. Owen, S. H. Rosenberg, B. Roth, R. Ross, M. Schapira, S. L. Schreiber, B. Shoichet, M. Sundström, G. Superti-Furga, J. Taunton, L. Toledo-Sherman, C. Walpole, M. A. Walters, T. M. Willson, P. Workman, R. N. Young and W. J. Zuercher, *Nature Chemical Biology 2015 11:8*, 2015, **11**, 536–541.

21 T. C. Gust, L. Neubrandt, C. Merz, K. Asadullah, U. Zügel and A. von Bonin, *Cell Communication and Signaling*, 2008, **6**, 3.

22 A. L. Jackson, S. R. Bartz, J. Schelter, S. v. Kobayashi, J. Burchard, M. Mao, B. Li, G. Cavet and P. S. Linsley, *Nature Biotechnology 2003 21:6*, 2003, **21**, 635–637.

23 M. E. Bunnage, E. L. P. Chekler and L. H. Jones, *Nature Chemical Biology*, 2013, **9**, 195–199.

24 J. Blagg and P. Workman, *Current Opinion in Pharmacology*, 2014, **17**, 87–100.

25 S. v. Frye, *Nature Chemical Biology 2010 6:3*, 2010, **6**, 159–161.

26 A. C. Runcie, K. H. Chan, M. Zengerle and A. Ciulli, *Current Opinion in Chemical Biology*, 2016, **33**, 186–194.

27 H. James Benns, J. Wincott, E. W. Tate and M. A. Child, *Current Opinion in Chemical Biology*, 2021, **60**, 20–29.

28 W. P. Heal, T. H. T. Dang and E. W. Tate, *Chemical Society Reviews*, 2010, **40**, 246–257.

29 B. Ghosh and L. H. Jones, *Medchemcomm*, 2014, **5**, 247–254.

30 A. Abdeldayem, Y. S. Raouf, S. N. Constantinescu, R. Moriggl and P. T. Gunning, *Chemical Society Reviews*, 2020, **49**, 2617–2687.

31 J. Singh, R. C. Petter, T. A. Baillie and A. Whitty, *Nature Reviews Drug Discovery*, 2011, **10**, 307–317.

32 T. D. Warner and J. A. Mitchell, *Proc Natl Acad Sci U S A*, 2002, **99**, 13371–13373.

33 D. J. Waxman and J. L. Strominger, *https://doi.org/10.1146/annurev.bi.52.070183.004141*, 2003, **Vol. 52**, 825–869.

34 T. Zhang, J. M. Hatcher, M. Teng, N. S. Gray and M. Kostic, *Cell Chemical Biology*, 2019, **26**, 1486–1500.

35 R. Lonsdale and R. A. Ward, *Chemical Society Reviews*, 2018, **47**, 3816–3830.

36 F. Sutanto, M. Konstantinidou and A. Dömling, *RSC Medicinal Chemistry*, 2020, **11**, 876–884.

37 R. Lagoutte, R. Patouret and N. Winssinger, *Current Opinion in Chemical Biology*, 2017, **39**, 54–63.

38 E. Resnick, A. Bradley, J. Gan, A. Douangamath, T. Krojer, R. Sethi, P. P. Geurink, A. Aimon, G. Amitai, D. Bellini, J. Bennett, M. Fairhead, O. Fedorov, R. Gabizon, J. Gan, J. Guo, A.

Plotnikov, N. Reznik, G. F. Ruda, L. Díaz-Sáez, V. M. Straub, T. Szommer, S. Velupillai, D. Zaidman, Y. Zhang, A. R. Coker, C. G. Dowson, H. M. Barr, C. Wang, K. V. M. Huber, P. E. Brennan, H. Ovaa, F. von Delft and N. London, *J Am Chem Soc*, 2019, **141**, 8951–8968.

39 W. Lu, M. Kostic, T. Zhang, J. Che, M. P. Patricelli, L. H. Jones, E. T. Chouchani and N. S. Gray, *RSC Chemical Biology*, 2021, **2**, 354–367.

40 S. G. Kathman, Z. Xu and A. v. Statsyuk, *Journal of Medicinal Chemistry*, 2014, **57**, 4969–4974.

41 H. Johansson, Y. C. Isabella Tsai, K. Fantom, C. W. Chung, S. Kümper, L. Martino, D. A. Thomas, H. C. Eberl, M. Muelbaier, D. House and K. Rittinger, *J Am Chem Soc*, , DOI:10.1021/jacs.8b13193.

42 T. Barf and A. Kaptein, *Journal of Medicinal Chemistry*, 2012, **55**, 6243–6262.

43 F. M. Ferguson and N. S. Gray, *Nature Reviews Drug Discovery*, 2018, **17**, 353–376.

44 J. M. Ostrem, U. Peters, M. L. Sos, J. A. Wells and K. M. Shokat, *Nature 2013 503:7477*, 2013, **503**, 548–551.

45 P. A. Jackson, J. C. Widen, D. A. Harki and K. M. Brummond, *Journal of Medicinal Chemistry*, 2017, **60**, 839–885.

46 M. Gehringer and S. A. Laufer, *Journal of Medicinal Chemistry*, 2019, **62**, 5673–5724.

47 G. Roos, N. Foloppe and J. Messens, *Antioxid Redox Signal*, 2013, **18**, 94–127.

48 E. Awoonor-Williams and C. N. Rowley, *Journal of Chemical Theory and Computation*, 2016, **12**, 4662–4673.

49 K. Brocklehurst, M. Resmini and C. M. Topham, *Methods*, 2001, **24**, 153–167.

50 A. Miseta and P. Csutora, *Molecular Biology and Evolution*, 2000, **17**, 1232–1239.

51 J. van der Reest, S. Lilla, L. Zheng, S. Zanivan and E. Gottlieb, *Nature Communications 2018 9:1*, 2018, **9**, 1–16.

52 D. A. Erlanson, A. C. Braisted, D. R. Raphael, M. Randal, R. M. Stroud, E. M. Gordon and J. A. Wells, *Proc Natl Acad Sci U S A*, 2000, **97**, 9367–9372.

53 S. G. Kathman and A. v. Statsyuk, *Medchemcomm*, 2016, **7**, 576–585.

54 J. D. Sadowsky, M. A. Burlingame, D. W. Wolan, C. L. McClendon, M. P. Jacobson and J. A. Wells, *Proc Natl Acad Sci U S A*, 2011, **108**, 6056–6061.

55 J. M. Lodge, T. Justin Rettenmaier, J. A. Wells, W. C. Pomerantz and A. K. Mapp, *Medchemcomm*, 2014, **5**, 370–375.

56 A. C. Braisted, J. D. Oslob, W. L. Delano, J. Hyde, R. S. McDowell, N. Waal, C. Yu, M. R. Arkin and B. C. Raimundo, *J Am Chem Soc*, 2003, **125**, 3714–3715.

57 R. Lonsdale and R. A. Ward, *Chemical Society Reviews*, 2018, **47**, 3816–3830.

58 A. Birkholz, D. J. Kopecky, L. P. Volak, M. D. Bartberger, Y. Chen, C. M. Tegley, T. Arvedson, J. D. McCarter, C. Fotsch and V. J. Cee, *Journal of Medicinal Chemistry*, 2020, **63**, 11602–11614.

59 H. R. Tsou, N. Mamuya, B. D. Johnson, M. F. Reich, B. C. Gruber, F. Ye, R. Nilakantan, R. Shen, C. Discafani, R. DeBlanc, R. Davis, F. E. Koehn, L. M. Greenberger, Y. F. Wang and A. Wissner, *Journal of Medicinal Chemistry*, 2001, **44**, 2719–2734.

60 J. S. Martin, C. J. MacKenzie, D. Fletcher and I. H. Gilbert, *Bioorganic & Medicinal Chemistry*, 2019, **27**, 2066–2074.

61 V. J. Cee, L. P. Volak, Y. Chen, M. D. Bartberger, C. Tegley, T. Arvedson, J. McCarter, A. S. Tasker and C. Fotsch, *Journal of Medicinal Chemistry*, 2015, **58**, 9171–9178.

62 R. Lonsdale, J. Burgess, N. Colclough, N. L. Davies, E. M. Lenz, A. L. Orton and R. A. Ward, *Journal of Chemical Information and Modeling*, 2017, **57**, 3124–3137.

63 R. H. Nonoo, A. Armstrong and D. J. Mann, *ChemMedChem*, 2012, **7**, 2082–2086.

64 B. R. Lanning, L. R. Whitby, M. M. Dix, J. Douhan, A. M. Gilbert, E. C. Hett, T. O. Johnson, C. Joslyn, J. C. Kath, S. Niessen, L. R. Roberts, M. E. Schnute, C. Wang, J. J. Hulce, B. Wei, L. O. Whiteley, M. M. Hayward and B. F. Cravatt, *Nature Chemical Biology 2014 10:9*, 2014, **10**, 760–767.

65 J. Clayden, N. Greeves and S. Warren, *Organic Chemistry*, 2nd edn., 2012.

66 E. Weerapana, C. Wang, G. M. Simon, F. Richter, S. Khare, M. B. D. Dillon, D. A. Bachovchin, K. Mowen, D. Baker and B. F. Cravatt, *Nature*, 2010, **468**, 790–795.

67 C. Wang, E. Weerapana, M. M. Blewett and B. F. Cravatt, *Nat Methods*, 2014, **11**, 79.

68 K. M. Backus, B. E. Correia, K. M. Lum, S. Forli, B. D. Horning, G. E. González-Páez, S. Chatterjee, B. R. Lanning, J. R. Teijaro, A. J. Olson, D. W. Wolan and B. F. Cravatt, *Nature 2016 534:7608*, 2016, **534**, 570–574.

69 M. Kuljanin, D. C. Mitchell, D. K. Schweppe, A. S. Gikandi, D. P. Nusinow, N. J. Bulloch, E. v. Vinogradova, D. L. Wilson, E. T. Kool, J. D. Mancias, B. F. Cravatt and S. P. Gygi, *Nature Biotechnology 2021 39:5*, 2021, **39**, 630–641.

70 S. G. Kathman, I. Span, A. T. Smith, Z. Xu, J. Zhan, A. C. Rosenzweig and A. v. Statsyuk, *J Am Chem Soc*, 2015, **137**, 12442–12445.

71 B. W. Zaro, L. R. Whitby, K. M. Lum and B. F. Cravatt, *J Am Chem Soc*, 2016, **138**, 15841–15844.

72 U. P. Dahal, A. M. Gilbert, R. S. Obach, M. E. Flanagan, J. M. Chen, C. Garcia-Irizarry, J. T. Starr, B. Schuff, D. P. Uccello and J. A. Young, *Medchemcomm*, 2016, **7**, 864–872.

73 S. M. Hacker, K. M. Backus, M. R. Lazear, S. Forli, B. E. Correia and B. F. Cravatt, *Nature Chemistry 2017 9:12*, 2017, **9**, 1181–1190.

74 M. E. Abbasov, M. E. Kavanagh, T. A. Ichu, M. R. Lazear, Y. Tao, V. M. Crowley, C. W. am Ende, S. M. Hacker, J. Ho, M. M. Dix, R. Suciu, M. M. Hayward, L. L. Kiessling and B. F. Cravatt, *Nature Chemistry 2021 13:11*, 2021, **13**, 1081–1092.

75 A. Narayanan and L. H. Jones, *Chemical Science*, 2015, **6**, 2650–2659.

76 K. Gilbert, A. Vuorinen, A. Aatkar, P. Pogany, J. Pettinger, E. Grant, J. Kirkpatrick, K. Rittinger, D. House, G. Burley and J. Bush, , DOI:10.26434/CHEMRXIV-2022-J8B8D-V2.

77 J. Pettinger, M. Carter, K. Jones and M. D. Cheeseman, *Journal of Medicinal Chemistry*, 2019, **62**, 11383–11398.

78 N. N. Gushwa, S. Kang, J. Chen and J. Taunton, *J Am Chem Soc*, 2012, **134**, 20214–20217.

79 G. H. Iyer, S. Garrod, V. L. Woods and S. S. Taylor, *Journal of Molecular Biology*, 2005, **351**, 1110–1122.

80 S. E. Dalton, L. Dittus, D. A. Thomas, M. A. Convery, J. Nunes, J. T. Bush, J. P. Evans, T. Werner, M. Bantscheff, J. A. Murphy and S. Campos, *J Am Chem Soc*, 2018, **140**, 932–939.

81 J. C. L. Fournier, J. P. Evans, F. Zappacosta, D. A. Thomas, V. K. Patel, G. v. White, S. Campos and N. C. O. Tomkinson, *ACS Chemical Biology*, 2021, **16**, 1644–1653.

82 Y. Qian, M. Schürmann, P. Janning, C. Hedberg and H. Waldmann, *Angewandte Chemie International Edition*, 2016, **55**, 7766–7771.

83 P. Martín-Gago, E. K. Fansa, M. Winzker, S. Murarka, P. Janning, C. Schultz-Fademrecht, M. Baumann, A. Wittinghofer and H. Waldmann, *Cell Chemical Biology*, 2017, **24**, 589-597.e5.

84 K. Chen and L. Kurgan, *PLOS ONE*, 2009, **4**, e4473.

85 V. H. B. Serrão and J. F. Scortecci, *Frontiers in Molecular Biosciences*, 2020, **7**, 2.

86 J. K. Eaton, L. Furst, R. A. Ruberto, D. Moosmayer, A. Hilpmann, M. J. Ryan, K. Zimmermann, L. L. Cai, M. Niehues, V. Badock, A. Kramm, S. Chen, R. C. Hillig, P. A. Clemons, S. Gradl, C. Montagnon, K. E. Lazarski, S. Christian, B. Bajrami, R. Neuhaus, A. L. Eheim, V. S. Viswanathan and S. L. Schreiber, *Nature Chemical Biology 2020 16:5*, 2020, **16**, 497–506.

87 M. Weïwer, J. A. Bittker, T. A. Lewis, K. Shimada, W. S. Yang, L. MacPherson, S. Dandapani, M. Palmer, B. R. Stockwell, S. L. Schreiber and B. Munoz, *Bioorganic & Medicinal Chemistry Letters*, 2012, **22**, 1822–1826.

88 H. Fang, B. Peng, S. Y. Ong, Q. Wu, L. Li and S. Q. Yao, *Chemical Science*, 2021, **12**, 8288–8310.

89 H. Kolb, M. Finn and B. Sharpless, *Angew Chem Int Ed Engl*, 2001, **40**, 2004–2021.

90 C. J. Pickens, S. N. Johnson, M. M. Pressnall, M. A. Leon and C. J. Berkland, *Bioconjugate Chemistry*, 2018, **29**, 686–701.

91 N. Z. Fantoni, A. H. El-Sagheer and T. Brown, *Chemical Reviews*, 2021, **121**, 7122–7154.

92 V. V. Rostovtsev, L. G. Green, V. V. Fokin and K. B. Sharpless, *Angew Chem Int Ed Engl.*, 2002, **41**, 2596–2599.

93 E. M. Sletten, C. R. Bertozzi, C. R. Bertozzi and E. M. Sletten, *Angewandte Chemie International Edition*, 2009, **48**, 6974–6998.

94 N. J. Agard, J. A. Prescher and C. R. Bertozzi, *J Am Chem Soc*, 2004, **126**, 15046–15047.

95 J. M. Baskin, J. A. Prescher, S. T. Laughlin, N. J. Agard, P. v. Chang, I. A. Miller, A. Lo, J. A. Codelli and C. R. Bertozzi, *Proc Natl Acad Sci U S A*, 2007, **104**, 16793–16797.

96 A. J. Sterling, R. C. Smith, E. A. Anderson and F. Duarte, , DOI:10.26434/CHEMRXIV-2021-N0XM9-V2.

97 F. Thalhammer, U. Wallfahrer and J. Sauer, *Tetrahedron Letters*, 1990, **31**, 6851–6854.

98 M. L. Blackman, M. Royzen and J. M. Fox, *J Am Chem Soc*, 2008, **130**, 13518–13519.

99 R. Selvaraj and J. M. Fox, *Current Opinion in Chemical Biology*, 2013, **17**, 753–760.

100 B. L. Oliveira, Z. Guo and G. J. L. Bernardes, *Chemical Society Reviews*, 2017, **46**, 4895–4950.

101 J. Yang, J. Šečkute, C. M. Cole and N. K. Devaraj, *Angewandte Chemie International Edition*, 2012, **51**, 7476–7479.

102 J. M. J. M. Ravasco, C. M. Monteiro and A. F. Trindade, *Organic Chemistry Frontiers*, 2017, **4**, 1167–1198.

103 N. R. Burton, P. Kim and K. M. Backus, *Organic & Biomolecular Chemistry*, 2021, **19**, 7792–7809.

104 E. Smith and I. Collins, *Future Medicinal Chemistry*, 2015, **7**, 159–183.

105 M. H. Wright and S. A. Sieber, *Natural Product Reports*, 2016, **33**, 681–708.

106 A. Dugan, C. Y. Majmudar, R. Pricer, S. Niessen, J. K. Lancia, H. Y. H. Fung, B. F. Cravatt and A. K. Mapp, *J Am Chem Soc*, 2016, **138**, 12629–12635.

107 M. G. Lougee, V. V. Pagar, H. J. Kim, S. X. Pancoe, R. H. Mach, B. A. Garcia and E. J. Petersson, *bioRxiv*, 2021, 2021.12.23.473999.

108 J. Zhu, J. Mo, H. zhi Lin, Y. Chen and H. peng Sun, *Bioorganic & Medicinal Chemistry*, 2018, **26**, 3065–3075.

109 R. T. Howard, P. Hemsley, P. Petteruti, C. N. Saunders, J. A. Molina Bermejo, J. S. Scott, J. W. Johannes and E. W. Tate, *ACS Chemical Biology*, 2020, **15**, 325–333.

110 J. W. Johannes, L. Almeida, K. Daly, A. D. Ferguson, S. E. Grosskurth, H. Guan, T. Howard, S. Ioannidis, S. Kazmirski, M. L. Lamb, N. A. Larsen, P. D. Lyne, K. Mikule, C. Ogoe, B. Peng, P. Petteruti, J. A. Read, N. Su, M. Sylvester, S. Throner, W. Wang, X. Wang, J. Wu, Q. Ye, Y. Yu, X. Zheng and D. A. Scott, *Bioorganic & Medicinal Chemistry Letters*, 2015, **25**, 5743–5747.

111 Z. Wang, S. E. Grosskurth, T. Cheung, P. Petteruti, J. Zhang, X. Wang, W. Wang, F. Gharahdaghi, J. Wu, N. Su, R. T. Howard, M. Mayo, D. Widzowski, D. A. Scott, J. W. Johannes, M. L. Lamb, D. Lawson, J. R. Dry, P. D. Lyne, E. W. Tate, M. Zinda, K. Mikule, S. E. Fawell, C. Reimer and H. Chen, *Cancer Research*, 2018, **78**, 6691–6702.

112 J. A. Ward, A. Pinto-Fernandez, L. Cornelissen, S. Bonham, L. Díaz-Sáez, O. Riant, K. V. M. Huber, B. M. Kessler, O. Feron and E. W. Tate, *Journal of Medicinal Chemistry*, 2020, **63**, 3756–3762.

113 S. L. Berger, T. Kouzarides, R. Shiekhattar and A. Shilatifard, *Genes & Development*, 2009, **23**, 781–783.

114 C. H. Arrowsmith, C. Bountra, P. v. Fish, K. Lee and M. Schapira, *Nature Reviews Drug Discovery 2012 11:5*, 2012, **11**, 384–400.

115 S. J. Conway, *ACS Medicinal Chemistry Letters*, 2012, **3**, 691–694.

116 T. Jenuwein and C. D. Allis, *Science (1979)*, 2001, **293**, 1074–1080.

117 A. J. Bannister and T. Kouzarides, *Cell Research 2011 21:3*, 2011, **21**, 381–395.

118 L. E. Jennings, A. R. Measures, B. G. Wilson and S. J. Conway, *Future Medicinal Chemistry*, 2014, **6**, 179–204.

119 S. R. Haynes, C. Dollard, F. Winston, S. Beck, J. Trowsdale and I. B. Dawid, *Nucleic Acids Res*, 1992, **20**, 2603.

120 P. Filippakopoulos, S. Picaud, M. Mangos, T. Keates, J. P. Lambert, D. Barsyte-Lovejoy, I. Felletar, R. Volkmer, S. Müller, T. Pawson, A. C. Gingras, C. H. Arrowsmith and S. Knapp, *Cell*, 2012, **149**, 214–231.

121 D. S. Hewings, T. P. C. Rooney, L. E. Jennings, D. A. Hay, C. J. Schofield, P. E. Brennan, S. Knapp and S. J. Conway, *Journal of Medicinal Chemistry*, 2012, **55**, 9393–9413.

122 W. Zhang, C. Prakash, C. Sum, Y. Gong, Y. Li, J. J. T. Kwok, N. Thiessen, S. Pettersson, S. J. M. Jones, S. Knapp, H. Yang and K. C. Chin, *Journal of Biological Chemistry*, 2012, **287**, 43137–43155.

123 J. Shi and C. R. Vakoc, *Molecular Cell*, 2014, **54**, 728–736.

124 A. Tsherniak, F. Vazquez, P. G. Montgomery, B. A. Weir, G. Kryukov, G. S. Cowley, S. Gill, W. F. Harrington, S. Pantel, J. M. Krill-Burger, R. M. Meyers, L. Ali, A. Goodale, Y. Lee, G. Jiang, J. Hsiao, W. F. J. Gerath, S. Howell, E. Merkel, M. Ghandi, L. A. Garraway, D. E. Root, T. R. Golub, J. S. Boehm and W. C. Hahn, *Cell*, 2017, **170**, 564-576.e16.

125 O. Gilan, I. Rioja, K. Knezevic, M. J. Bell, M. M. Yeung, N. R. Harker, E. Y. N. Lam, C. Chung, P. Bamborough, M. Petretich, M. Urh, S. J. Atkinson, A. K. Bassil, E. J. Roberts, D. Vassiliadis, M. L. Burr, A. G. S. Preston, C. Wellaway, T. Werner, J. R. Gray, A. M. Michon, T. Gobbetti, V. Kumar, P. E. Soden, A. Haynes, J. Vappiani, D. F. Tough, S. Taylor, S. J. Dawson, M. Bantscheff, M. Lindon, G. Drewes, E. H. Demont, D. L. Daniels, P. Grandi, R. K. Prinjha and M. A. Dawson, *Science (1979)*, 2020, **368**, 387–394.

126 P. Filippakopoulos, J. Qi, S. Picaud, Y. Shen, W. B. Smith, O. Fedorov, E. M. Morse, T. Keates, T. T. Hickman, I. Felletar, M. Philpott, S. Munro, M. R. McKeown, Y. Wang, A. L. Christie, N. West, M. J. Cameron, B. Schwartz, T. D. Heightman, N. la Thangue, C. A. French, O. Wiest, A. L. Kung, S. Knapp and J. E. Bradner, *Nature 2010 468:7327*, 2010, **468**, 1067–1073.

127 E. Nicodeme, K. L. Jeffrey, U. Schaefer, S. Beinke, S. Dewell, C. W. Chung, R. Chandwani, I. Marazzi, P. Wilson, H. Coste, J. White, J. Kirilovsky, C. M. Rice, J. M. Lora, R. K. Prinjha, K. Lee and A. Tarakhovsky, *Nature 2010 468:7327*, 2010, **468**, 1119–1123.

128 M. Baud, E. Lin-Shiao, T. Cadote, C. Tallant, A. Pschibul, K.-H. Chan, M. Zengerle, J. Garcia, T. Kwan, F. Ferguson and A. Ciulli, *Science (1979)*, 2014, **346**, 0–4.

129 Z. Li, D. Wang, L. Li, S. Pan, Z. Na, C. Y. J. Tan and S. Q. Yao, *J Am Chem Soc*, 2014, **136**, 9990–9998.

130 D. S. Hewings, M. Wang, M. Philpott, O. Fedorov, S. Uttarkar, P. Filippakopoulos, S. Picaud, C. Vuppusetty, B. Marsden, S. Knapp, S. J. Conway and T. D. Heightman, *Journal of Medicinal Chemistry*, 2011, **54**, 6761–6770.

131 D. S. Hewings, O. Fedorov, P. Filippakopoulos, S. Martin, S. Picaud, A. Tumber, C. Wells, M. M. Olcina, K. Freeman, A. Gill, A. J. Ritchie, D. W. Sheppard, A. J. Russell, E. M. Hammond, S. Knapp, P. E. Brennan and S. J. Conway, *Journal of Medicinal Chemistry*, 2013, **56**, 3217–3227.

132 L. E. Jennings, M. Schiedel, D. S. Hewings, S. Picaud, C. M. C. Laurin, P. A. Bruno, J. P. Bluck, A. R. Scorah, L. See, J. K. Reynolds, M. Moroglu, I. N. Mistry, A. Hicks, P. Guzanov, J. Clayton, C. N. G. Evans, G. Stazi, P. C. Biggin, A. K. Mapp, E. M. Hammond, P. G. Humphreys, P. Filippakopoulos and S. J. Conway, *Bioorganic and Medicinal Chemistry*, 2018, **26**, 2937–2957.

133 J. Chen, P. Tang, Y. Wang, J. Wang, C. Yang, Y. Li, G. Yang, F. Wu, J. Zhang and L. Ouyang, *Journal of Medicinal Chemistry*, 2022, acs.jmedchem.1c01835.

134 A. C. Runcie, M. Zengerle, K. H. Chan, A. Testa, L. van Beurden, M. G. J. Baud, O. Epemolu, L. C. J. Ellis, K. D. Read, V. Coulthard, A. Brien and A. Ciulli, *Chemical Science*, 2018, **9**, 2452–2468.

135 O. A. Kharenko, R. G. Patel, S. D. Brown, C. Calosing, A. White, D. Lakshminarasimhan, R. K. Suto, B. C. Duffy, D. B. Kitchen, K. G. McLure, H. C. Hansen, E. H. van der Horst and P. R. Young, *Journal of Medicinal Chemistry*, 2018, **61**, 8202–8211.

136 R. J. Watson, P. Bamborough, H. Barnett, C. W. Chung, R. Davis, L. Gordon, P. Grandi, M. Petretich, A. Phillipou, R. K. Prinjha, I. Rioja, P. Soden, T. Werner and E. H. Demont, *Journal of Medicinal Chemistry*, 2020, **63**, 9045–9069.

137 J. T. Seal, S. J. Atkinson, H. Aylott, P. Bamborough, C. W. Chung, R. C. B. Copley, L. Gordon, P. Grandi, J. R. J. Gray, L. A. Harrison, T. G. Hayhow, M. Lindon, C. Messenger, A. M. Michon, D. Mitchell, A. Preston, R. K. Prinjha, I. Rioja, S. Taylor, I. D. Wall, R. J. Watson, J. M. Woolven and E. H. Demont, *Journal of Medicinal Chemistry*, 2020, **63**, 9093–9126.

138 L. A. Harrison, S. J. Atkinson, A. Bassil, C. W. Chung, P. Grandi, J. R. J. Gray, E. Levernier, A. Lewis, D. Lugo, C. Messenger, A. M. Michon, D. J. Mitchell, A. Preston, R. K. Prinjha, I. Rioja, J. T. Seal, S. Taylor, I. D. Wall, R. J. Watson, J. M. Woolven and E. H. Demont, *Journal of Medicinal Chemistry*, 2021, **64**, 10742–10771.

139 C. R. Wellaway, P. Bamborough, S. G. Bernard, C. W. Chung, P. D. Craggs, L. Cutler, E. H. Demont, J. P. Evans, L. Gordon, B. Karamshi, A. J. Lewis, M. J. Lindon, D. J. Mitchell, I. Rioja, P. E. Soden, S. Taylor, R. J. Watson, R. Willis, J. M. Woolven, B. S. Wyspiańska, W. J. Kerr and R. K. Prinjha, *Journal of Medicinal Chemistry*, 2020, **63**, 9020–9044.

140 J. Seal, Y. Lamotte, F. Donche, A. Bouillot, O. Mirguet, F. Gellibert, E. Nicodeme, G. Krysa, J. Kirilovsky, S. Beinke, S. McCleary, I. Rioja, P. Bamborough, C. W. Chung, L. Gordon, T. Lewis, A. L. Walker, L. Cutler, D. Lugo, D. M. Wilson, J. Witherington, K. Lee and R. K. Prinjha, *Bioorganic & Medicinal Chemistry Letters*, 2012, **22**, 2968–2972.

141 E. Coombs, 2018.

142 O. Gilan, I. Rioja, K. Knezevic, M. J. Bell, M. M. Yeung, N. R. Harker, E. Y. N. Lam, C. Chung, P. Bamborough, M. Petretich, M. Urh, S. J. Atkinson, A. K. Bassil, E. J. Roberts, D. Vassiliadis, M. L. Burr, A. G. S. Preston, C. Wellaway, T. Werner, J. R. Gray, A. M. Michon, T. Gobbetti, V. Kumar, P. E. Soden, A. Haynes, J. Vappiani, D. F. Tough, S. Taylor, S. J. Dawson, M. Bantscheff, M. Lindon, G. Drewes, E. H. Demont, D. L. Daniels, P. Grandi, R. K. Prinjha and M. A. Dawson, *Science (1979)*, 2020, **368**, 387–394.

143 T. C. R. Miller, B. Simon, V. Rybin, H. Grötsch, S. Curtet, S. Khochbin, T. Carlomagno and C. W. Müller, *Nature Communications 2016 7:1*, 2016, **7**, 1–13.

144 M. Jung, M. Philpott, S. Müller, J. Schulze, V. Badock, U. Eberspächer, D. Moosmayer, B. Bader, N. Schmees, A. Fernández-Montalván and B. Haendler, *The Journal of Biological Chemistry* , 2014, **289**, 9304–9319.

145 K. F. Geoghegan, H. B. F. Dixon, P. J. Rosner, L. R. Hoth, A. J. Lanzetti, K. A. Borzilleri, E. S. Marr, L. H. Pezzullo, L. B. Martin, P. K. Lemotte, A. S. McColl, A. v. Kamath and J. G. Stroh, *Analytical Biochemistry*, 1999, **267**, 169–184.

146 F. H. Niesen, H. Berglund and M. Vedadi, *Nature Protocols*, 2007, **2**, 2212–2221.

147 P. Filippakopoulos, J. Qi, S. Picaud, Y. Shen, W. B. Smith, O. Fedorov, E. M. Morse, T. Keates, T. T. Hickman, I. Felletar, M. Philpott, S. Munro, M. R. McKeown, Y. Wang, A. L. Christie, N. West, M. J. Cameron, B. Schwartz, T. D. Heightman, N. la Thangue, C. A. French, O. Wiest, A. L. Kung, S. Knapp and J. E. Bradner, *Nature*, 2010, **468**, 1067–1073.

148 D. S. Hewings, University of Oxford, 2014.

149 D. Bell, E. A. Crowe, N. J. Dixon, G. R. Geen, I. S. Mann and M. R. Shipton, *Tetrahedron*, 1994, **50**, 6643–6652.

150 A. R. Sekirnik (née Measures), D. S. Hewings, N. H. Theodoulou, L. Jursins, K. R. Lewendon, L. E. Jennings, T. P. C. Rooney, T. D. Heightman and S. J. Conway, *Angewandte Chemie - International Edition*, 2016, **55**, 8353–8357.

151 T. Meng, D. Zhang, Z. Xie, T. Yu, S. Wu, L. Wyder, U. Regenass, K. Hilpert, M. Huang, M. Geng and J. Shen, *Journal of Medicinal Chemistry*, 2014, **57**, 9832–9843.

152 M. Scholl, S. Ding, C. W. Lee and R. H. Grubbs, *Organic Letters*, 1999, **1**, 953–956.

153 M. Philpott, J. Yang, T. Tumber, O. Fedorov, S. Uttarkar, P. Filippakopoulos, S. Picaud, T. Keates, I. Felletar, A. Ciulli, S. Knapp and T. D. Heightman, *Molecular BioSystems*, 2011, **7**, 2899–2908.

154 J. P. Hilton-Proctor, O. Ilyichova, Z. Zheng, I. G. Jennings, R. W. Johnstone, J. Shortt, S. J. Mountford, M. J. Scanlon and P. E. Thompson, *European Journal of Medicinal Chemistry*, 2020, **191**, 112120.

155 M. A. Dawson, R. K. Prinjha, A. Dittmann, G. Giotopoulos, M. Bantscheff, W. I. Chan, S. C. Robson, C. W. Chung, C. Hopf, M. M. Savitski, C. Huthmacher, E. Gudgin, D. Lugo, S. Beinke, T. D. Chapman, E. J. Roberts, P. E. Soden, K. R. Auger, O. Mirguet, K. Doehner, R. Delwel, A. K. Burnett, P. Jeffrey, G. Drewes, K. Lee, B. J. P. Huntly and T. Kouzarides, *Nature 2011 478:7370*, 2011, **478**, 529–533.

156 K. D. Freeman-Cook, R. L. Hoffman and T. W. Johnson, *Future Medicinal Chemistry*, 2013, **5**, 113–115.

157 L. Di and E. H. Kerns, *Current Opinion in Chemical Biology*, 2003, **7**, 402–408.

158 G. Li, R. Kakarla and S. W. Gerritz, *Tetrahedron Letters*, 2007, **48**, 4595–4599.

159 US5571821 (A), 1996.

160 L. Jedinák, R. Zátopková, H. Zemánková, A. Šustková and P. Cankař, *Journal of Organic Chemistry*, 2017, **82**, 157–169.

161 M. Krasavin, M. Korsakov, Z. Zvonaryova, E. Semyonychev, T. Tuccinardi, S. Kalinin, M. Tanç and C. T. Supuran, *Bioorganic and Medicinal Chemistry*, 2017, **25**, 1914–1925.

162 WO2019032720 (A1), 2019.

163 J. Velcicky, A. Soicke, R. Steiner and H. G. Schmalz, *J Am Chem Soc*, 2011, **133**, 6948–6951.

164 WO2013052157 (A1), 2013.

165 WO2010104406 (A1), 2010.

166 WO2006127207 (A1), 2006.

167 M. Philpott, J. Yang, T. Tumber, O. Fedorov, S. Uttarkar, P. Filippakopoulos, S. Picaud, T. Keates, I. Felletar, A. Ciulli, S. Knapp and T. D. Heightman, *Molecular BioSystems*, 2011, **7**, 2899–2908.

168 M. Schiedel, M. Moroglu, D. M. H. Ascough, A. E. R. Chamberlain, J. J. A. G. Kamps, A. R. Sekirnik and S. J. Conway, *Angewandte Chemie International Edition*, 2019, **58**, 17930–17952.

169 Kenry and B. Liu, *Trends in Chemistry*, 2019, **1**, 763–778.

170 N. Z. Fantoni, A. H. El-Sagheer and T. Brown, *Chemical Reviews*, 2021, **121**, 7122–7154.

171 C. Guo, Y. Chang, X. Wang, C. Zhang, P. Hao, K. Ding and Z. Li, *Chemical Communications*, 2019, **55**, 834–837.

172 L. Qian, S. Pan, J. S. Lee, J. Ge, L. Li and S. Q. Yao, *Chemical Communications*, 2019, **55**, 1092–1095.

173 R. Rezende Miranda, Y. Fu, X. Chen, J. Perino, P. Cao, J. Carpten, Y. Chen and C. Zhang, *Journal of Medicinal Chemistry*, 2020, **63**, 11484–11497.

174 H. Deng, Q. Lei, W. Shang, Y. Li, L. Bi, N. Yang, Z. Yu and W. Li, *European Journal of Medicinal Chemistry*, 2022, **230**, 114100.

175 R. Chinchilla and C. Nájera, *Chemical Reviews*, 2007, **107**, 874–922.

176 J. Myšková, O. Rybakova, J. Brynda, P. Khoroshyy, A. Bondar and J. Lazar, *Proc Natl Acad Sci U S A*, 2020, **117**, 32395–32401.

177 M. Békés, D. R. Langley and C. M. Crews, *Nature Reviews Drug Discovery 2022 21:3*, 2022, **21**, 181–200.

178 A. C. Lai and C. M. Crews, *Nature Reviews Drug Discovery 2016 16:2*, 2016, **16**, 101–114.

179 K. Raina and C. M. Crews, *Current Opinion in Chemical Biology*, 2017, **39**, 46–53.

180 T. K. Neklesa, J. D. Winkler and C. M. Crews, *Pharmacology & Therapeutics*, 2017, **174**, 138–144.

181 K. Cyrus, M. Wehenkel, E. Y. Choi, H. J. Han, H. Lee, H. Swanson and K. B. Kim, *Molecular BioSystems*, 2011, **7**, 359–364.

182 M. S. Gadd, A. Testa, X. Lucas, K. H. Chan, W. Chen, D. J. Lamont, M. Zengerle and A. Ciulli, *Nature Chemical Biology 2017 13:5*, 2017, **13**, 514–521.

183 B. E. Smith, S. L. Wang, S. Jaime-Figueroa, A. Harbin, J. Wang, B. D. Hamman and C. M. Crews, *Nature Communications 2019 10:1*, 2019, **10**, 1–13.

184 M. A. Troy Bemis, J. J. la Clair and M. D. Burkart, , DOI:10.1021/acs.jmedchem.1c00482.

185 P. Pfaff, K. T. G. Samarasinghe, C. M. Crews and E. M. Carreira, *ACS Central Science*, 2019, **5**, 1682–1690.

186 T. K. Neklesa, H. S. Tae, A. R. Schneekloth, M. J. Stulberg, T. W. Corson, T. B. Sundberg, K. Raina, S. A. Holley and C. M. Crews, *Nature Chemical Biology 2011 7:8*, 2011, **7**, 538–543.

187 H. S. Tae, T. B. Sundberg, T. K. Neklesa, D. J. Noblin, J. L. Gustafson, A. G. Roth, K. Raina and C. M. Crews, *ChemBioChem*, 2012, **13**, 538–541.

188 J. L. Gustafson, T. K. Neklesa, C. S. Cox, A. G. Roth, D. L. Buckley, H. S. Tae, T. B. Sundberg, D. B. Stagg, J. Hines, D. P. McDonnell, J. D. Norris and C. M. Crews, *Angewandte Chemie*, 2015, **127**, 9795–9798.

189 T. K. Neklesa and C. M. Crews, *Nature 2012 487:7407*, 2012, **487**, 308–309.

190 W. H. Guo, X. Qi, X. Yu, Y. Liu, C. I. Chung, F. Bai, X. Lin, D. Lu, L. Wang, J. Chen, L. H. Su, K. J. Nomie, F. Li, M. C. Wang, X. Shu, J. N. Onuchic, J. A. Woyach, M. L. Wang and J. Wang, *Nature Communications 2020 11:1*, 2020, **11**, 1–16.

191 D. Budhadev, E. Poole, I. Nehlmeier, Y. Liu, J. Hooper, E. Kalverda, U. S. Akshath, N. Hondow, W. B. Turnbull, S. Pöhlmann, Y. Guo and D. Zhou, *J Am Chem Soc*, 2020, **142**, 18022–18034.

192 X. Wang, H. Sun, J. Liu, W. Zhong, M. Zhang, H. Zhou, D. Dai and X. Lu, *Organic Letters*, 2019, **21**, 719–723.

193 B. Mondal, P. Ghosh, M. Kundu, T. K. Das and S. Das, *Organic & Biomolecular Chemistry*, 2021, **19**, 360–364.

194 A. P. Hill and R. J. Young, *Drug Discovery Today*, 2010, **15**, 648–655.

195 M. L. L. Donnelly, L. E. Hughes, G. Luke, H. Mendoza, E. ten Dam, D. Gani and M. D. Ryan, *Journal of General Virology*, 2001, **82**, 1027–1041.

196 Z. Liu, O. Chen, J. B. J. Wall, M. Zheng, Y. Zhou, L. Wang, H. Ruth Vaseghi, L. Qian and J. Liu, *Scientific Reports 2017 7:1*, 2017, **7**, 1–9.

197 R. P. Nowak, S. L. Deangelo, D. Buckley, Z. He, K. A. Donovan, J. An, N. Safaee, M. P. Jedrychowski, C. M. Ponthier, M. Ishoey, T. Zhang, J. D. Mancias, N. S. Gray, J. E. Bradner and E. S. Fischer, *Nature Chemical Biology 2018 14:7*, 2018, **14**, 706–714.

198 R. P. Nowak, Y. Xiong, N. Kirmani, J. Kalabathula, K. A. Donovan, N. A. Eleuteri, J. C. Yuan and E. S. Fischer, *Journal of Medicinal Chemistry*, 2021, **64**, 11637–11650.

199 J. Min, A. Mayasundari, F. Keramatnia, B. Jonchere, S. W. Yang, J. Jarusiewicz, M. Actis, S. Das, B. Young, J. Slavish, L. Yang, Y. Li, X. Fu, S. H. Garrett, M. K. Yun, Z. Li, S. Nithianantham, S. Chai, T. Chen, A. Shelat, R. E. Lee, G. Nishiguchi, S. W. White, M. F. Roussel, P. R. Potts, M. Fischer and Z. Rankovic, *Angewandte Chemie International Edition*, 2021, **60**, 26663–26670.

200 G. R. Fulmer, A. J. M. Miller, N. H. Sherden, H. E. Gottlieb, A. Nudelman, B. M. Stoltz, J. E. Bercaw and K. I. Goldberg, *Organometallics*, 2010, **29**, 2176–2179.

201 G. Giacomelli, L. de Luca and A. Porcheddu, *Tetrahedron*, 2003, **59**, 5437–5440.

202 O. Zhurakovskyi, Y. E. Türkmen, L. E. Löffler, V. A. Moorthie, C. C. Chen, M. A. Shaw, M. R. Crimmin, M. Ferrara, M. Ahmad, M. Ostovar, J. v. Matlock and V. K. Aggarwal, *Angewandte Chemie International Edition*, 2018, **57**, 1346–1350.

203 P. L. Jackson, C. D. Hanson, A. K. Farrell, R. J. Butcher, J. P. Stables, N. D. Eddington and K. R. Scott, *European Journal of Medicinal Chemistry*, 2012, **51**, 42–51.

204 G. Shakya, H. Rivera, D. J. Lee, M. J. Jaremko, J. J. la Clair, D. T. Fox, R. W. Haushalter, A. J. Schaub, J. Bruegger, J. F. Barajas, A. R. White, P. Kaur, E. R. Gwozdziowski, F. Wong, S. C. Tsai and M. D. Burkart, *J Am Chem Soc*, 2014, **136**, 16792–16799.

205 I. K. Khanna, Y. Yu, R. M. Huff, R. M. Weier, X. Xu, F. J. Koszyk, P. W. Collins, J. N. Cogburn, P. C. Isakson, C. M. Koboldt, J. L. Masferrer, W. E. Perkins, K. Seibert, A. W. Veenhuizen, J. Yuan, D. C. Yang and Y. Y. Zhang, *Journal of Medicinal Chemistry*, 2000, **43**, 3168–3185.

206 C. Christine, K. Ikhiri, A. Ahond, A. al Mourabit, C. Poupat and P. Potier, *Tetrahedron*, 2000, **56**, 1837–1850.

207 G. Dannhardt, W. Kiefer, G. Lambrecht, S. Laufer, E. Mutschler, J. Schweiger and H. G. Striegep, *European Journal of Medicinal Chemistry*, 1995, **30**, 839–850.

208 R. H. Good, G. Jones and J. R. Phipps, *J. Chem. Soc., Perkin Trans. 1*, 1972, 2441–2445.

209 A. Quilico and L. Panizzi, *Gazzetta Chimica Italiana*, 1938, **68**, 625–635.

210 G. Li, R. Kakarla and S. W. Gerritz, *Tetrahedron Letters*, 2007, **48**, 4595–4599.

211 US5571821, 1996.

212 C. Gibson, K. Schnatbaum, J. R. Pfeifer, E. Locardi, M. Paschke, U. Reimer, U. Richter, D. Scharn, A. Faussner and T. Tradler, *Journal of Medicinal Chemistry*, 2009, **52**, 4370–4379.

213 T. Itoh, T. Fukuda and T. Fujisawa, *Bull. Chem. Soc. Jpn.*, 1989, 62, 3851–3855.

214 C. K. Chan and M. Y. Chang, *Tetrahedron*, 2017, **73**, 5207–5213.

215 B. R. Kim, H. G. Lee, S. B. Kang, K. J. Jung, G. H. Sung, J. J. Kim, S. G. Lee and Y. J. Yoon, *Tetrahedron*, 2013, **69**, 10331–10336.

216 J. Lasri, S. Mukhopadhyay and M. A. J. Charmier, *Journal of Heterocyclic Chemistry*, 2008, **45**, 1385–1389.

217 Z. Pei, R. Mendonca, L. Gazzard, R. Pastor, L. Goon, A. Gustafson, E. Vanderporten, G. Hatzivassiliou, K. Dement, R. Cass, P. W. Yuen, Y. Zhang, G. Wu, X. Lin, Y. Liu and B. D. Sellers, *ACS Medicinal Chemistry Letters*, 2018, **9**, 417–421.

218 C. C. Lin, T. H. Hsieh, P. Y. Liao, Z. Y. Liao, C. W. Chang, Y. C. Shih, W. H. Yeh and T. C. Chien, *Organic Letters*, 2014, **16**, 892–895.

219 E. K. Grant, D. J. Fallon, M. M. Hann, K. G. M. Fantom, C. Quinn, F. Zappacosta, R. S. Annan, C. wa Chung, P. Bamborough, D. P. Dixon, P. Stacey, D. House, V. K. Patel, N. C. O. Tomkinson and J. T. Bush, *Angewandte Chemie International Edition*, 2020, **59**, 21096–21105.

220 K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror and D. E. Shaw, *Proteins: Structure, Function and Bioinformatics*, 2010, **78**, 1950–1958.

221 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *Journal of Chemical Physics*, 1983, **79**, 926– 935.

222 M. J. Abraham, T. Murtola, R. Schulz, S. Páall, J. C. Smith, B. Hess and E. Lindah, *SoftwareX*, 2015, **1–2**, 19–25.

223 G. Vriend., R. Krause, M. L. Hekkelman and J. E. Nielsen, WHAT-IF web interface, https://swift.cmbi.umcn.nl/servers/html/index.html.

224 M. F. Sanner, *J. Mol. Graphics Mod.*, 1999, **17**, 57–61.

225 G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson, *J. Computational Chemistry*, 2009, **16**, 2785–2791.

Appendices

Appendix A: Chapter 2

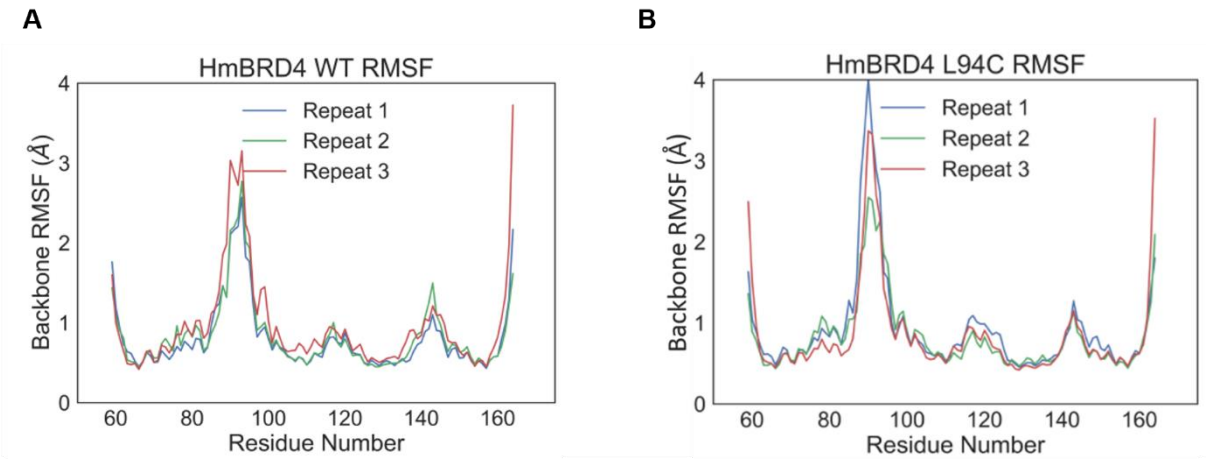


Figure A.1. A) RMSF simulation for BRD4(1)^{WT}. **B**) RMSF simulation for BRD4(1)^{L94C}.

Table A.1. Gene inserts were generated by PCR from parental DNA templates using the following primers:

Primer	Direction	Sequence (5' to 3')	Tm (°C)
GA_f	Forward	AGAACCTGTACTTCCAATCC	61
GA_r	Reverse	CGGAGCTCGAATTCG	60

Table A.2. Characterisation of His_6-BRD4(1)^{WT}.

Protein:	**His_6-BRD4(1)^{WT}**
Plasmid vector:	pNIC28-Bsa4 (Kanamycin)
DNA Sequence:	ATG AAC CCC CCG CCC CCA GAG ACC TCC AAC CCT AAC AAG CCC AAG AGG CAG ACC AAC CAA CTG CAA TAC CTG CTC AGA GTG GTG CTC AAG ACA CTA TGG AAA CAC CAG TTT GCA TGG CCT TTC CAG CAG CCT GTG GAT GCC GTC AAG CTG AAC CTC CCT GAT TAC TAT AAG ATC ATT AAA ACG CCT ATG GAT ATG GGA ACA ATA AAG AAG CGC TTG GAA AAC AAC TAT TAC TGG AAT GCT CAG GAA TGT ATC CAG GAC TTC AAC ACT ATG TTT ACA AAT TGT TAC ATC TAC AAC AAG CCT GGA GAT GAC ATA GTC TTA ATG GCA GAA GCT CTG GAA AAG CTC TTC TTG CAA AAA ATA AAT GAG CTA CCC ACA GAA GAA
Protein Sequence:	MNPPPPETSN PNKPKRQTNQ LQYLLRVVLK TLWKHQFAWP FQQPVDAVKL NLPDYYKIIK TPMDMGTIKK RLENNYYWNA QECIQDFNTM FTNCYIYNKP GDDIVLMAEA LEKLFLQKIN ELPTEE
Molecular Weight:	17549 Da
Extinction coefficient:	28545 M^{-1} cm^{-1}

Table A.3. Characterisation of His_6-BRD4(1)^{L94C}. The bold red letters show the mutation.

Protein:	**His_6-BRD4(1)^{L94C}**
Plasmid vector:	pNIC28-Bsa4 (Kanamycin)
DNA Sequence:	ATG AAC CCC CCG CCC CCA GAG ACC TCC AAC CCT AAC AAG CCC AAG AGG CAG ACC AAC CAA CTG CAA TAC CTG CTC AGA GTG GTG CTC AAG ACA CTA TGG AAA CAC CAG TTT GCA TGG CCT TTC CAG CAG CCT GTG GAT GCC GTC AAG CTG AAC **TGT** CCT GAT TAC TAT AAG ATC ATT AAA ACG CCT ATG GAT ATG GGA ACA ATA AAG AAG CGC TTG GAA AAC AAC TAT TAC TGG AAT GCT CAG GAA TGT ATC CAG GAC TTC AAC ACT ATG TTT ACA AAT TGT TAC ATC TAC AAC AAG CCT GGA GAT GAC ATA GTC TTA ATG GCA GAA GCT CTG GAA AAG CTC TTC TTG CAA AAA ATA AAT GAG CTA CCC ACA GAA GAA
Protein Sequence:	MNPPPPETSN PNKPKRQTNQ LQYLLRVVLK TLWKHQFAWP FQQPVDAVKL N**C**PDYYKIIK TPMDMGTIKK RLENNYYWNA QECIQDFNTM FTNCYIYNKP GDDIVLMAEA LEKLFLQKIN ELPTEE
Molecular Weight:	17539 Da
Extinction coefficient:	28545 M^{-1} cm^{-1}

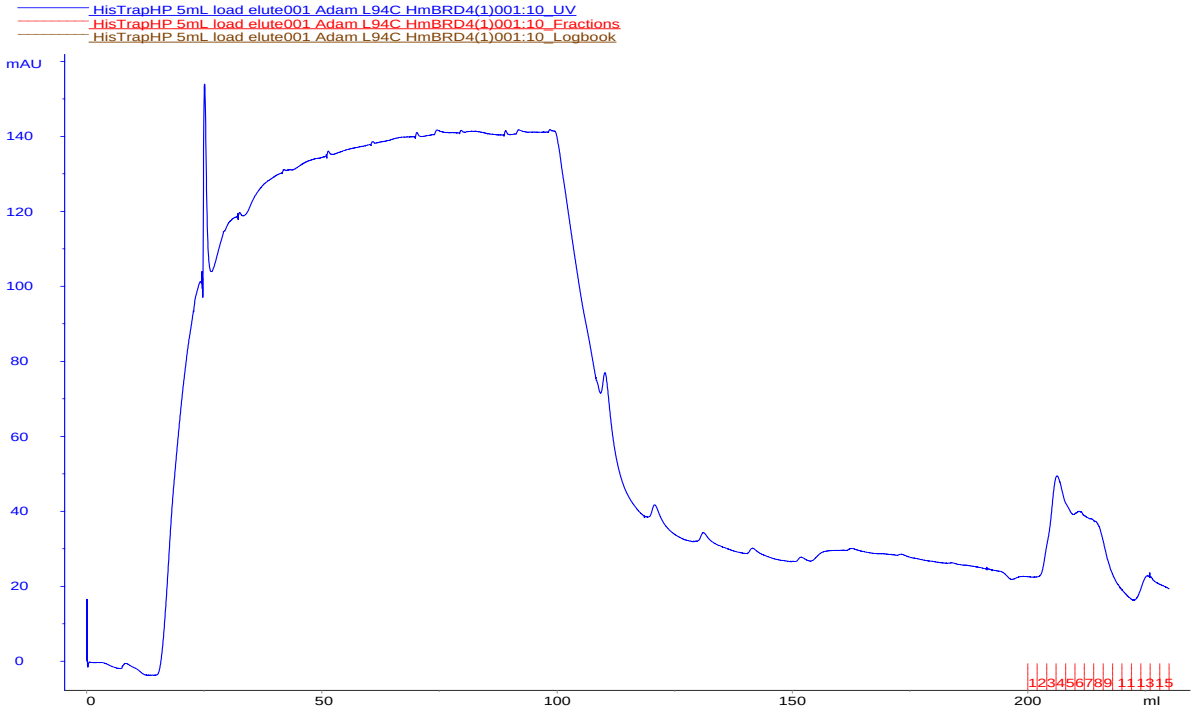


Figure A.2. IMAC trace of BRD4(1)^{L94C} purification. Fractions 2–5 combined.

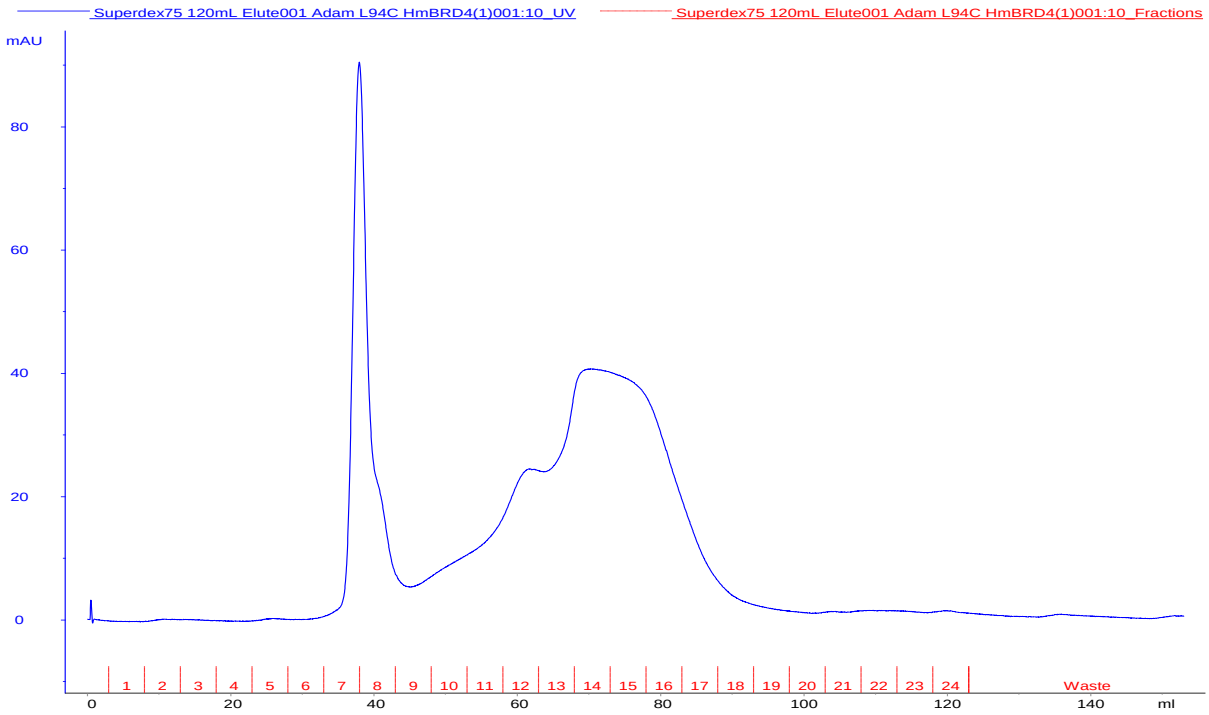


Figure A.3. SEC trace of BRD4(1)^{L94C} purification. Fractions 13–16 combined.

Preliminary PDB report



Preliminary Full wwPDB X-ray Structure Validation Report (i)

Dec 17, 2021 – 09:44 AM EST

This is a Preliminary Full wwPDB X-ray Structure Validation Report.

This report is produced by the standalone wwPDB validation server.
The structure in question has not been deposited to the wwPDB.
This report should not be submitted to journals.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

PRELIMINARY VALIDATION REPORT

The following versions of software and data (see references (i)) were used in the production of this report:

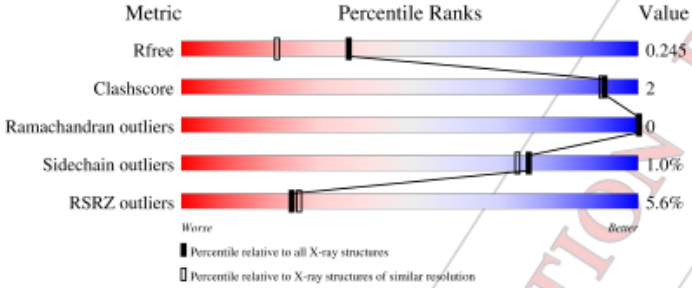
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.24
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.24

1 Overall quality at a glance

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.88 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	9470 (1.90-1.86)
Clashscore	141614	10282 (1.90-1.86)
Ramachandran outliers	138981	10152 (1.90-1.86)
Sidechain outliers	138945	10152 (1.90-1.86)
RSRZ outliers	127900	9303 (1.90-1.86)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	124	6% 94% 6%

PRELIMINARY VALIDATION REPORT

2 Entry composition

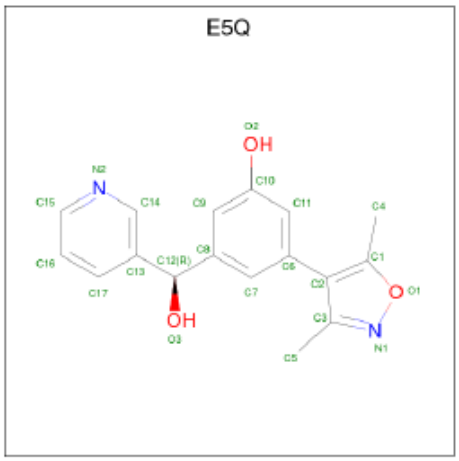
There are 4 unique types of molecules in this entry. The entry contains 2122 atoms, of which 997 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	124	1967	654	965	159	181	8	0	4	0

- Molecule 2 is 3-(3,5-dimethyl-1,2-oxazol-4-yl)-5-[({R})-oxidanyl(pyridin-3-yl)methyl]phenol (three-letter code: E5Q) (formula: $C_{17}H_{16}N_2O_3$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	H	N	O		
2	A	1	38	17	16	2	3	0	1

- Molecule 3 is a ligand with the chemical component id E5R but its atom names do not match the existing wwPDB Chemical Component Dictionary definition for E5R. Consequently no firm identification of ligand chemistry can be made. Once the structure is annotated then an identification and diagram will be given here.

PRELIMINARY VALIDATION REPORT

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	H	N	O		
3	B	1	38	17	16	2	3	0	1

- Molecule 4 is water.

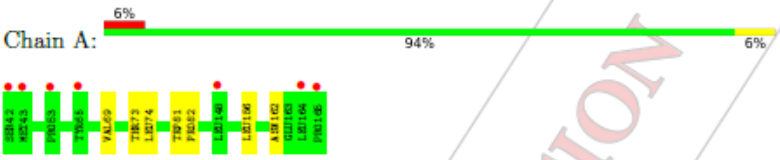
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	S	79	79	79	0	0

PRELIMINARY VALIDATION REPORT

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1:



PRELIMINARY VALIDATION REPORT

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	33.54Å 46.89Å 78.58Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.29 – 1.88 39.29 – 1.88	Depositor EDS
% Data completeness (in resolution range)	99.5 (39.29-1.88) 99.7 (39.29-1.88)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.18 (at 1.88Å)	Xtriage
Refinement program	phenix.refine 1.19.1_4122, PHENIX 1.19.1_4122	Depositor
R, R_{free}	0.206 , 0.245 0.205 , 0.245	Depositor DCC
R_{free} test set	510 reflections (4.83%)	wwPDB-VP
Wilson B-factor (Å^2)	24.7	Xtriage
Anisotropy	1.375	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.42 , 47.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.52$, $< L^2 > = 0.36$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	2122	wwPDB-VP
Average B, all atoms (Å^2)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.67% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

PRELIMINARY VALIDATION REPORT

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: E5Q, E5R

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| > 5$	RMSZ	$\#\|Z\| > 5$
1	A	0.47	0/1045	0.54	0/1430

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1002	965	945	3	0
2	A	22	16	0	0	0
3	B	22	16	0	0	0
4	S	79	0	0	0	2
All	All	1125	997	945	3	2

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 2.

All (3) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

PRELIMINARY VALIDATION REPORT

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:69:VAL:O	1:A:73[A]:THR:HG23	2.03	0.59
1:A:81:TRP:CG	1:A:82:PRO:HD3	2.49	0.48
1:A:74:LEU:HD23	1:A:156:LEU:HD23	2.01	0.42

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:S:13:HOH:O	4:S:29:HOH:O[3_655]	1.92	0.28
4:S:40:HOH:O	4:S:56:HOH:O[4_545]	2.01	0.19

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	126/124 (102%)	125 (99%)	1 (1%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	107/117 (92%)	106 (99%)	1 (1%)	78	76

All (1) residues with a non-rotameric sidechain are listed below:

PRELIMINARY VALIDATION REPORT

Mol	Chain	Res	Type
1	A	162	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 2 ligands modelled in this entry, 1 could not be matched to an existing wwPDB Chemical Component Dictionary definition at this stage - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	E5Q	A	203[A]	-	20,24,24	0.80	0	24,34,34	0.71	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	E5Q	A	203[A]	-	-	0/12/12/12	0/3/3/3

PRELIMINARY VALIDATION REPORT

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

PRELIMINARY VALIDATION REPORT

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2		OWAB(Å^2)	Q<0.9
1	A	124/124 (100%)	0.34	7 (5%)	24 26	27, 36, 49, 67	0

All (7) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	42	SER	8.4
1	A	164	LEU	3.1
1	A	65	TYR	2.9
1	A	53	PRO	2.3
1	A	148	LEU	2.1
1	A	165	PRO	2.0
1	A	43	MET	2.0

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	E5R	B	204[B]	22/?	0.86	0.19	42,42,50,50	38

Continued on next page...

PRELIMINARY VALIDATION REPORT

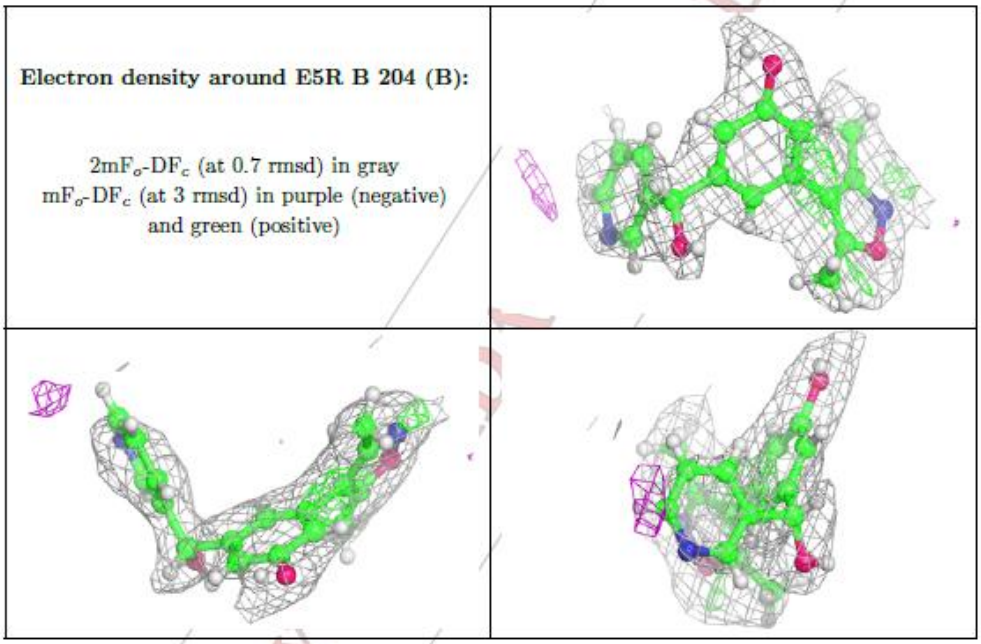
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	E5Q	A	203[A]	22/?	0.87	0.18	34,39,48,53	38

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around E5R B 204 (B):

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



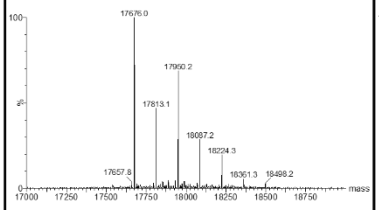
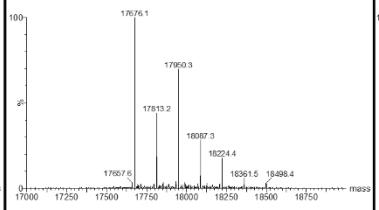
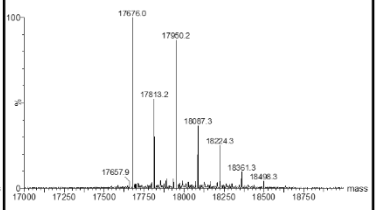
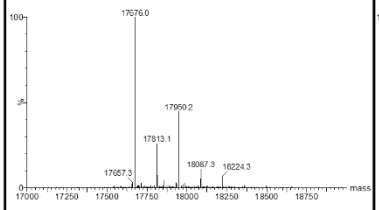
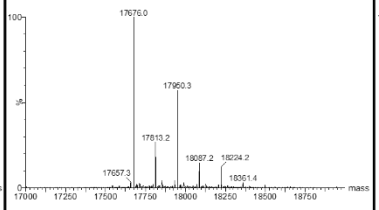
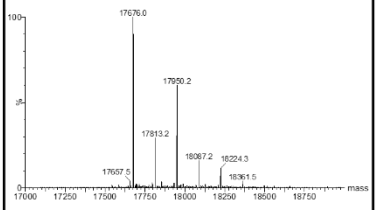
6.5 Other polymers (i)

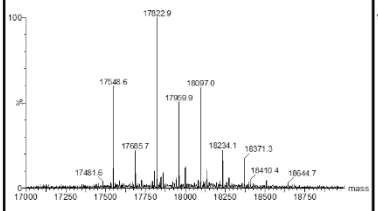
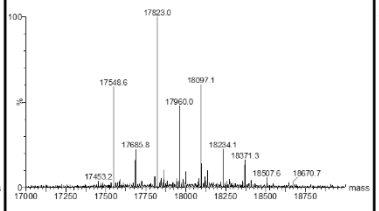
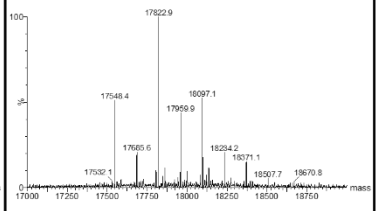
There are no such residues in this entry.

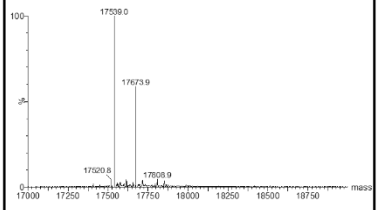
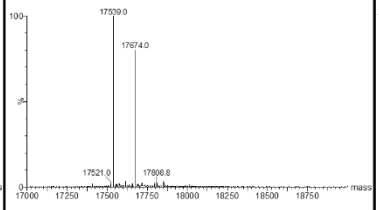
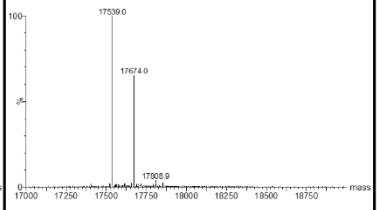
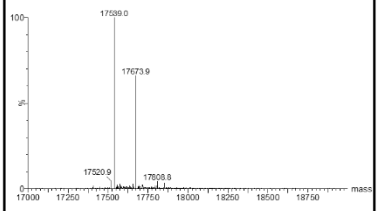
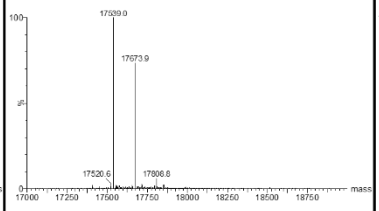
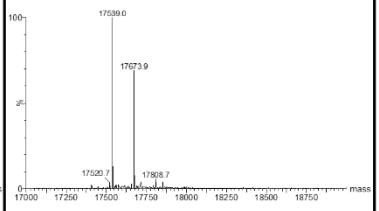
PRELIMINARY REPORT

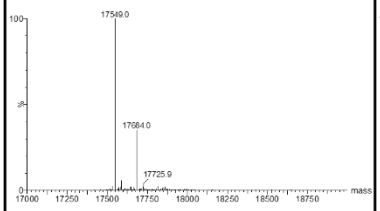
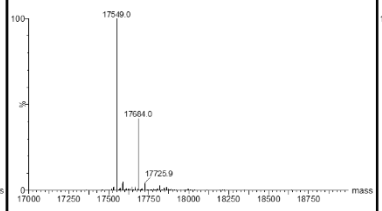
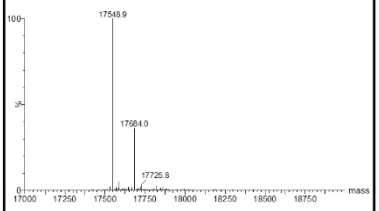
Appendix B: Chapter 3

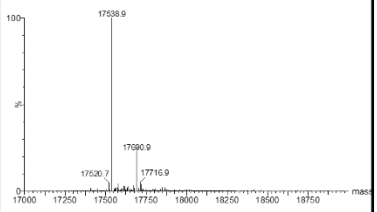
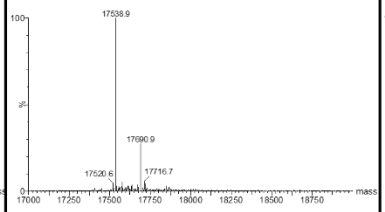
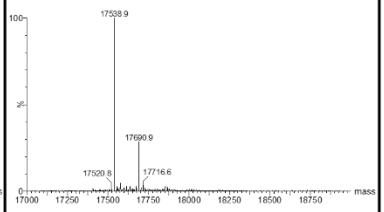
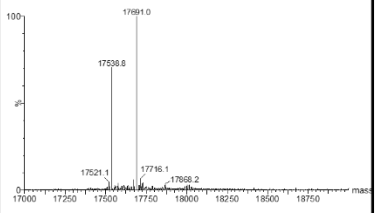
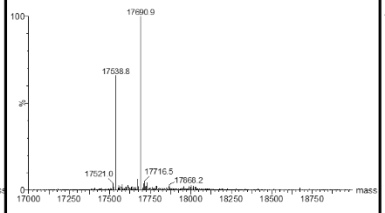
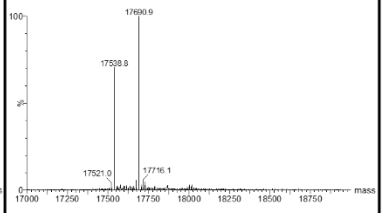
Deconvoluted mass spectra for in-house protein LC-MS assay of DMI-based fragments

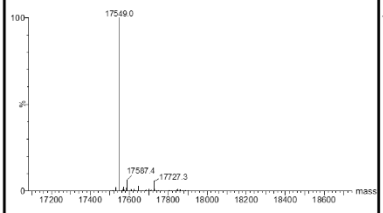
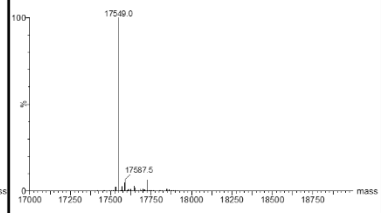
	3.1		
	BRD4(1)^{L94C}		
1 h			
4 h			

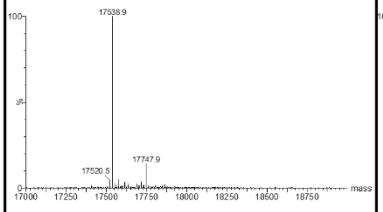
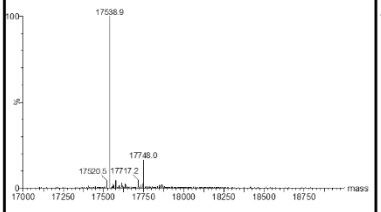
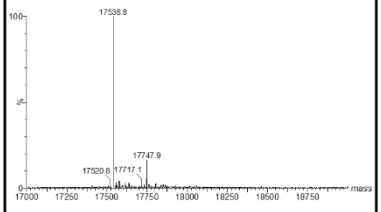
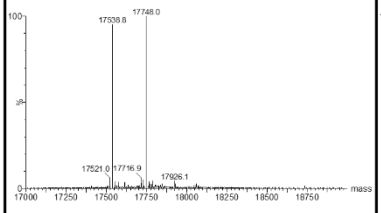
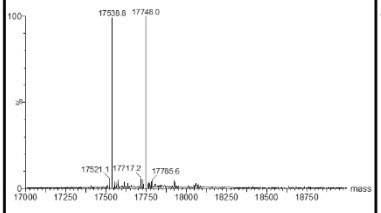
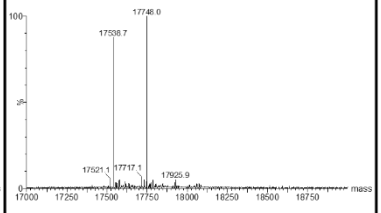
	3.1		
	BRD4(1)^{WT}		
4 h			

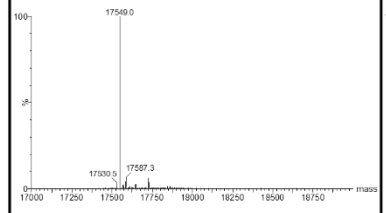
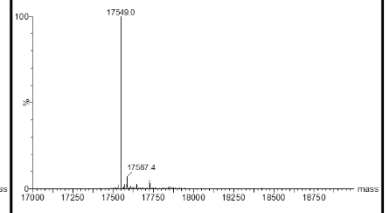
	3.3		
	BRD4(1)^{L94C}		
1 h			
4 h			

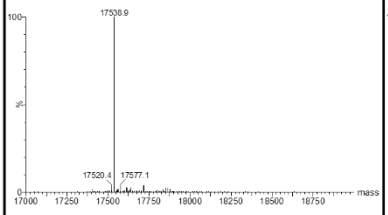
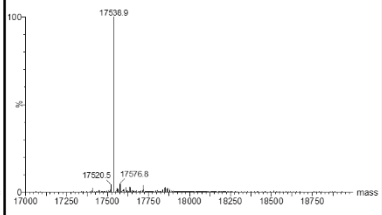
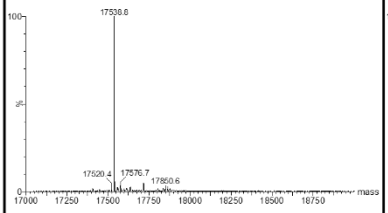
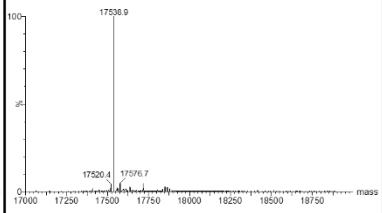
	3.3		
	BRD4(1)^{WT}		
4 h			

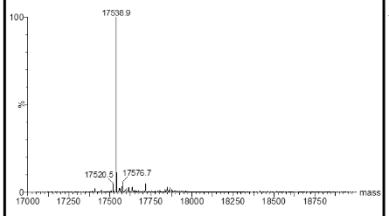
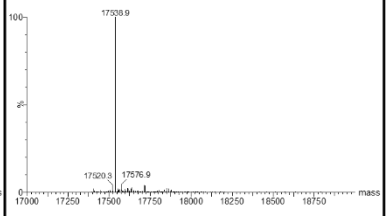
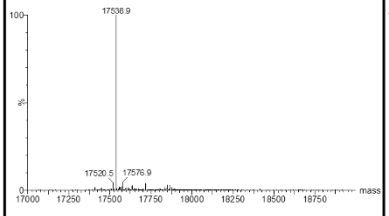
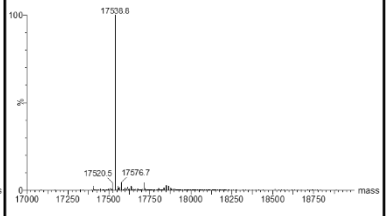
	3.4		
	BRD4(1)^{L94C}		
1 h			
4 h			

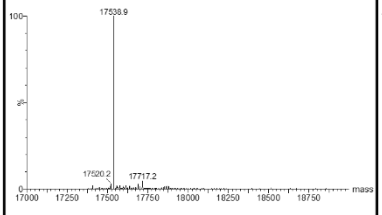
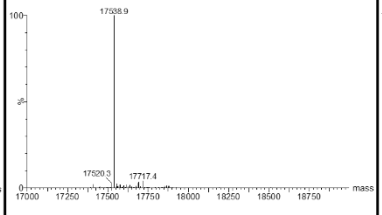
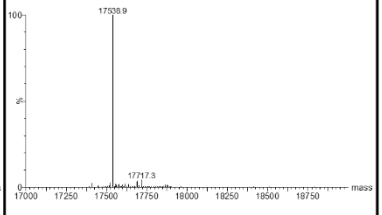
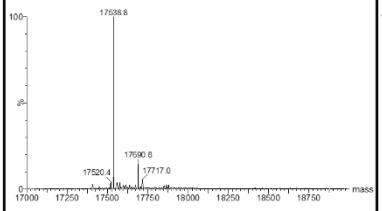
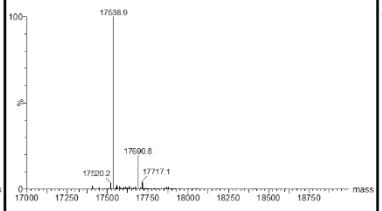
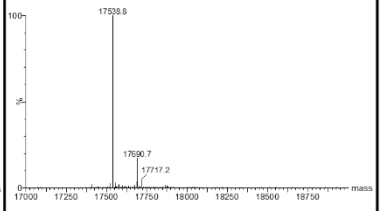
	3.4	
	BRD4(1)^{WT}	
4 h		

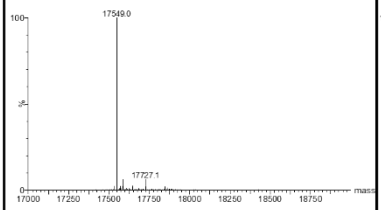
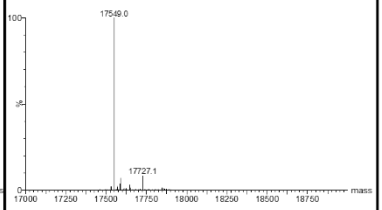
	3.5		
	BRD4(1)^{L94C}		
1 h			
4 h			

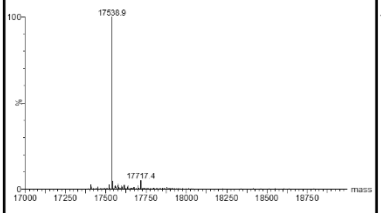
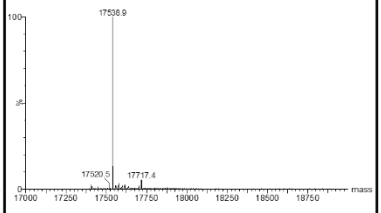
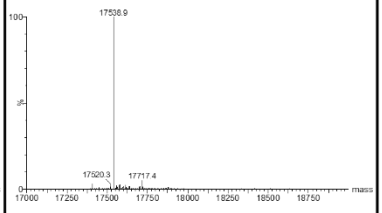
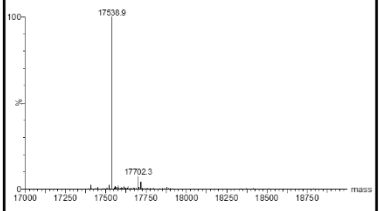
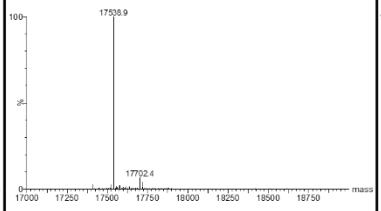
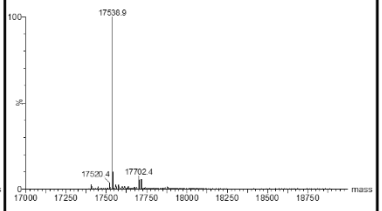
	3.5	
	BRD4(1)^{WT}	
4 h		

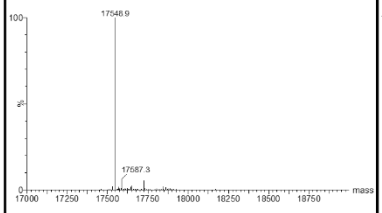
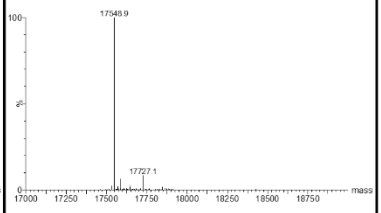
	3.6	
	BRD4(1)^{L94C}	
1 h		
4 h		

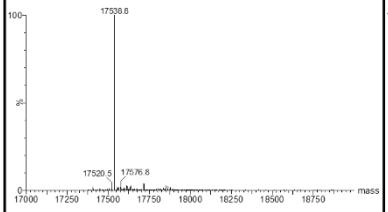
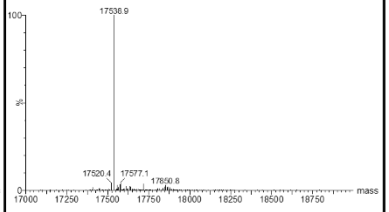
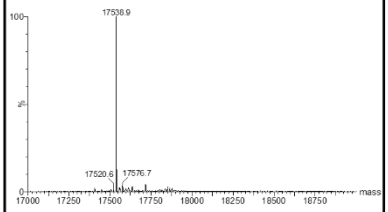
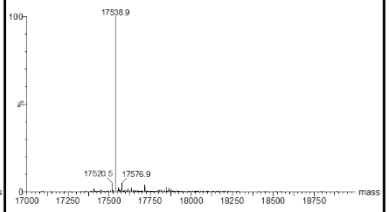
	3.7	
	BRD4(1)^{L94C}	
1 h		
4 h		

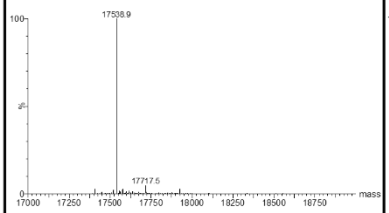
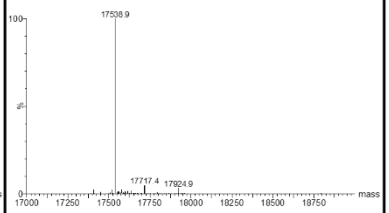
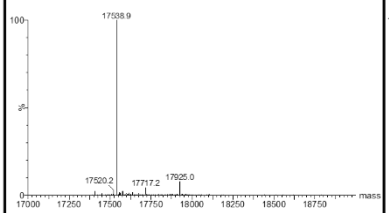
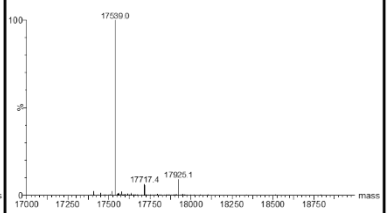
	3.8		
	BRD4(1)^{L94C}		
1 h			
4 h			

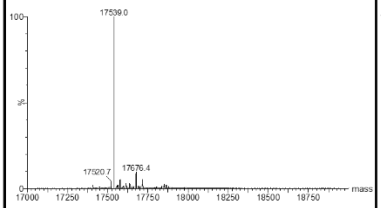
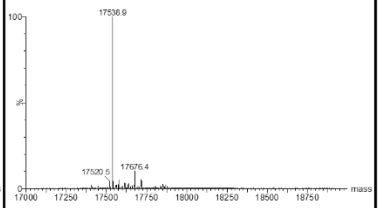
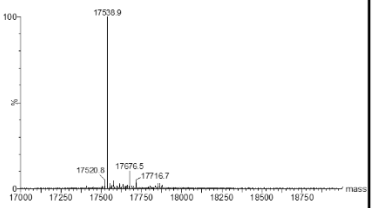
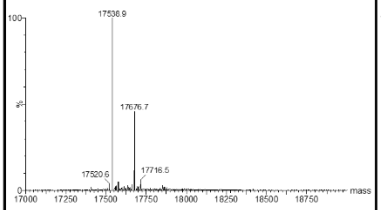
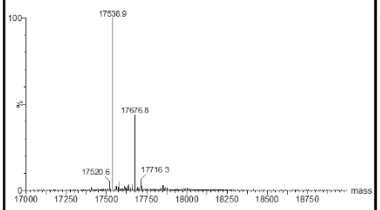
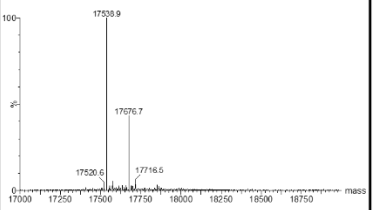
	3.8	
	BRD4(1)^{WT}	
4 h		

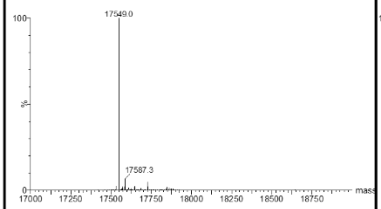
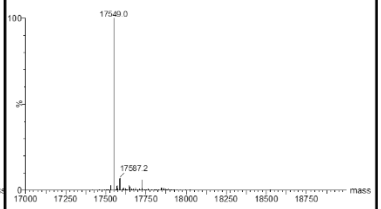
	3.9		
	BRD4(1)^{L94C}		
1 h			
4 h			

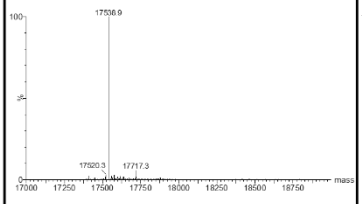
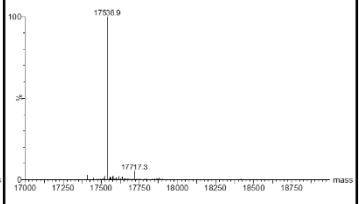
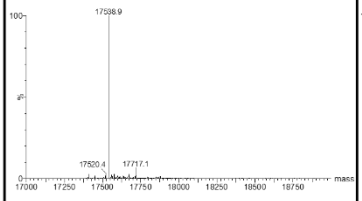
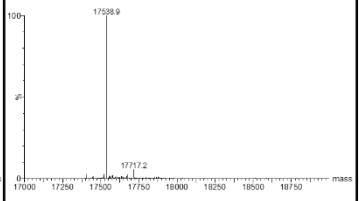
	3.9	
	BRD4(1)^{WT}	
4 h		

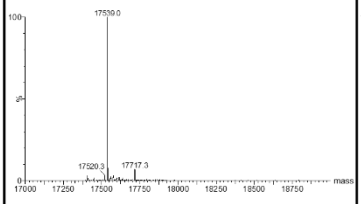
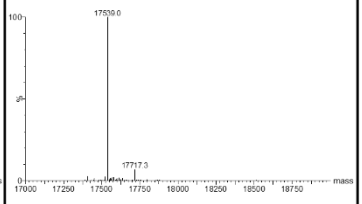
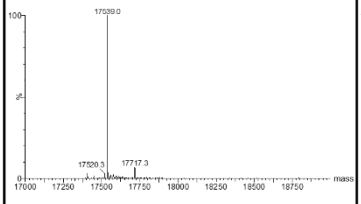
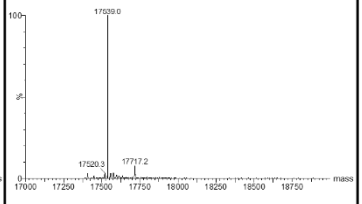
	3.10	
	BRD4(1)^{L94C}	
1 h		
4 h		

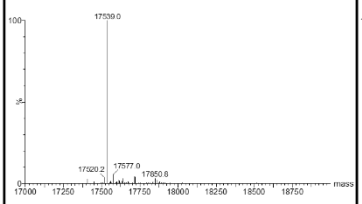
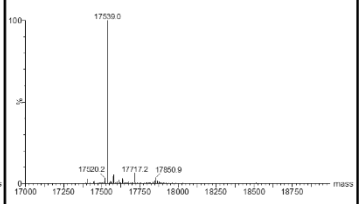
	3.41	
	BRD4(1)^{L94C}	
1 h		
4 h		

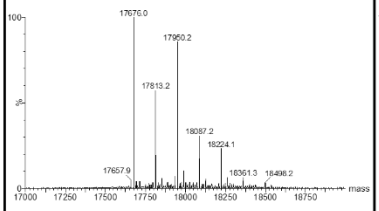
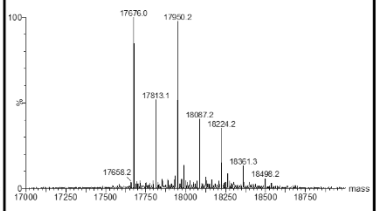
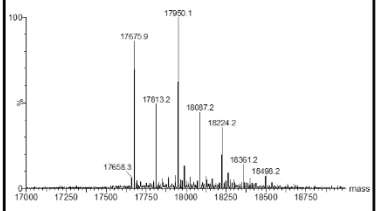
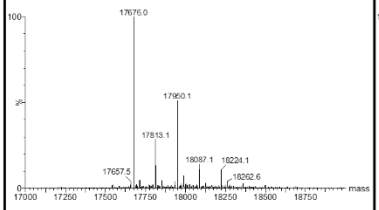
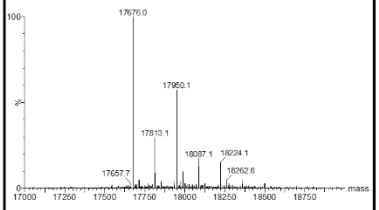
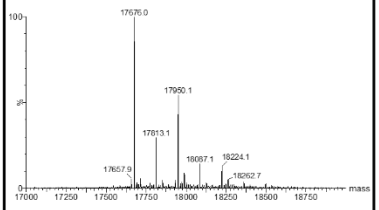
	3.11		
	BRD4(1)^{L94C}		
1 h			
4 h			

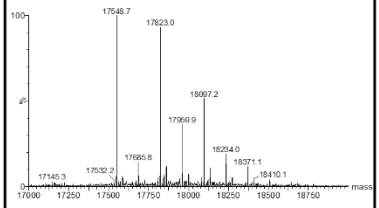
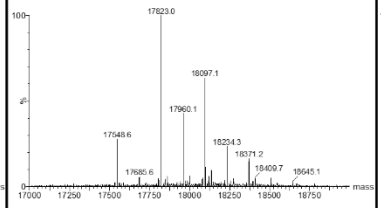
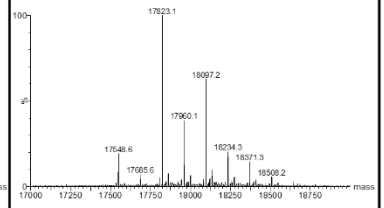
	3.11	
	BRD4(1)^{WT}	
4 h		

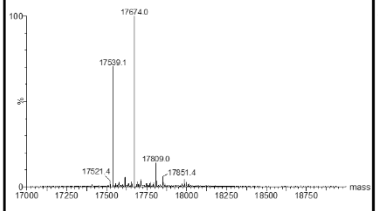
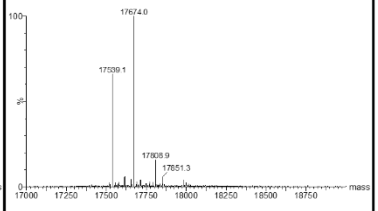
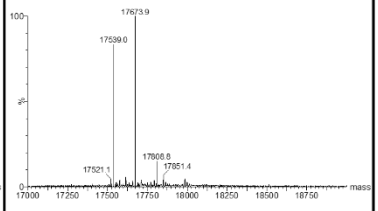
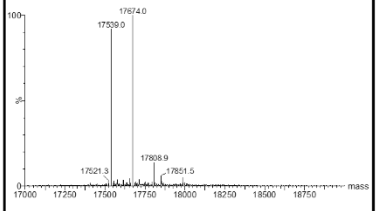
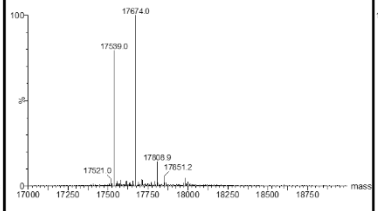
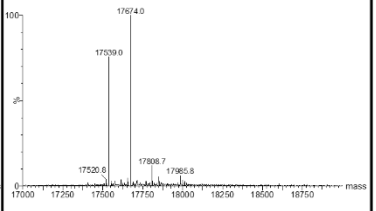
	3.31	
	BRD4(1)^{L94C}	
1 h		
4 h		

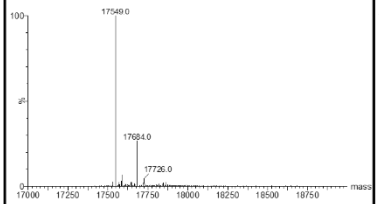
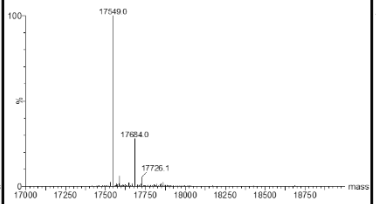
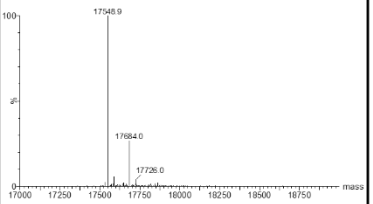
	3.16	
	BRD4(1)^{L94C}	
1 h		
4 h		

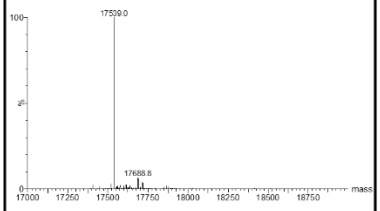
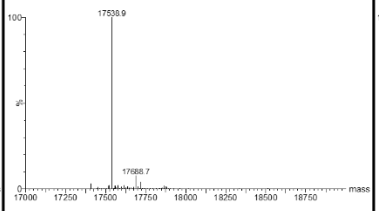
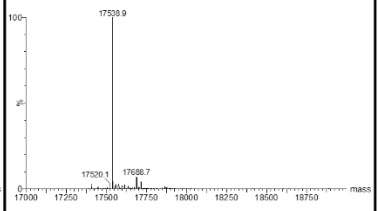
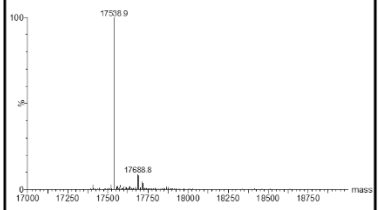
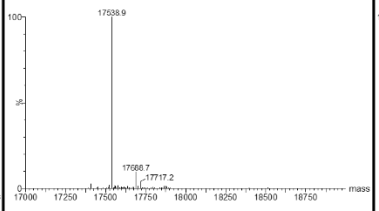
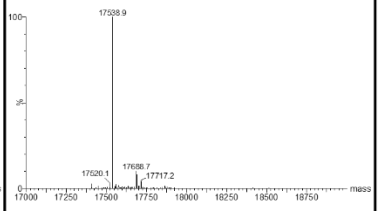
	3.39	
	BRD4(1)^{L94C}	
4 h		

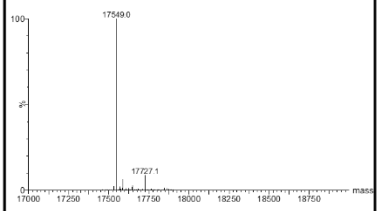
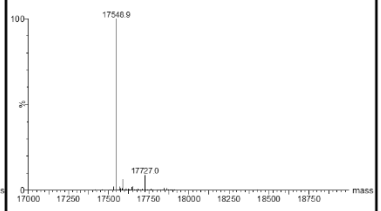
	3.12		
	BRD4(1)^{L94C}		
1 h			
4 h			

	3.12		
	BRD4(1)^{WT}		
4 h			

	3.14		
	BRD4(1)^{L94C}		
1 h			
4 h			

	3.14		
	BRD4(1)^{WT}		
4 h			

	3.15		
	BRD4(1)^{L94C}		
1 h			
4 h			

	3.15	
	BRD4(1)^{WT}	
4 h		

Deconvoluted mass spectra for GSK diverse fragment screen

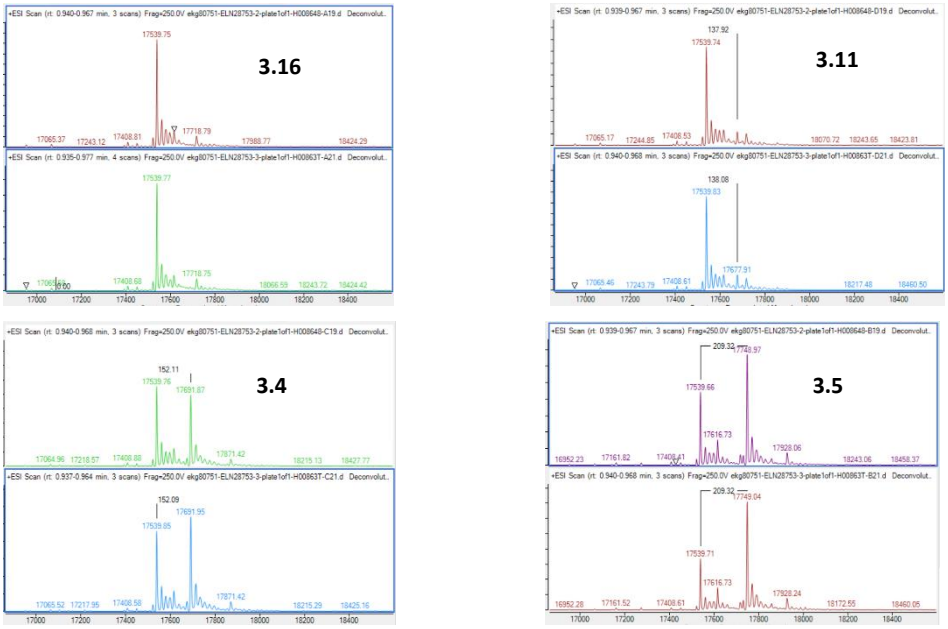


Figure B.1. Isoxazole fragments **3.16**, **3.11**, **3.4**, and **3.5** were used as positive and negative controls for the GSK screening.

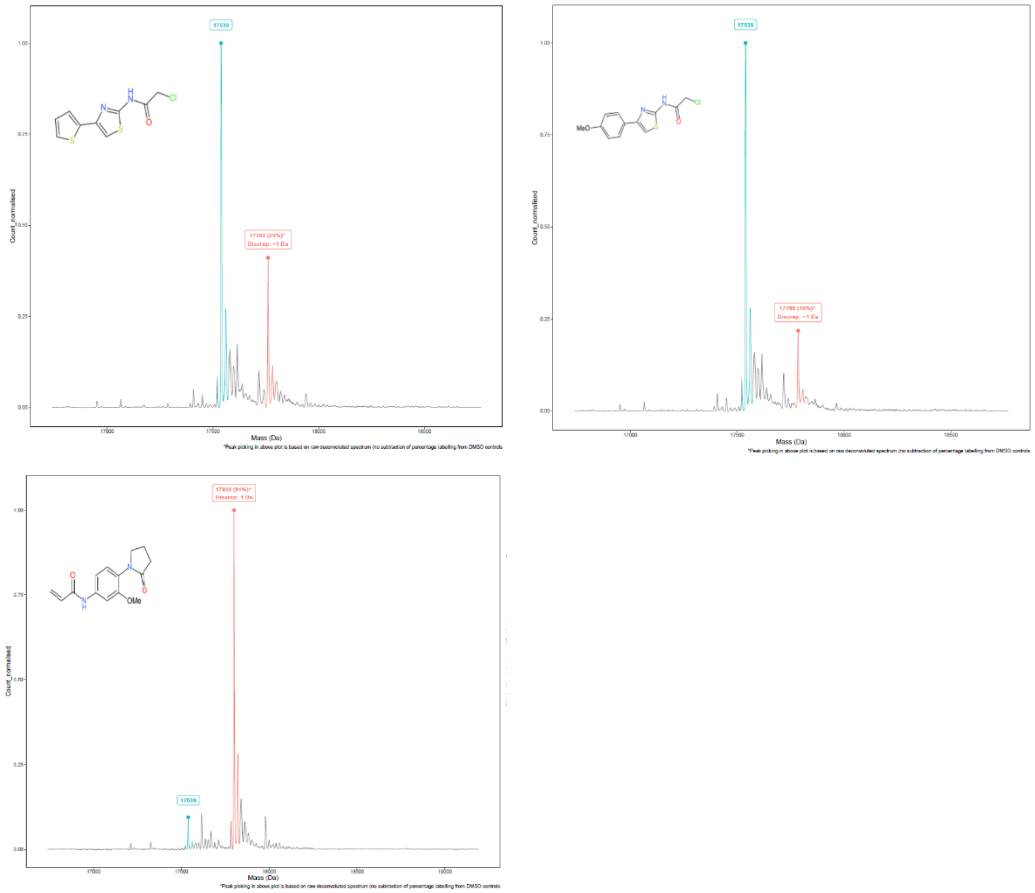


Figure B.2. Protein LC-MS spectra for compounds **3.47**, **3.48** and **3.50**, using the GSK screening conditions.

Appendix C: Chapter 4

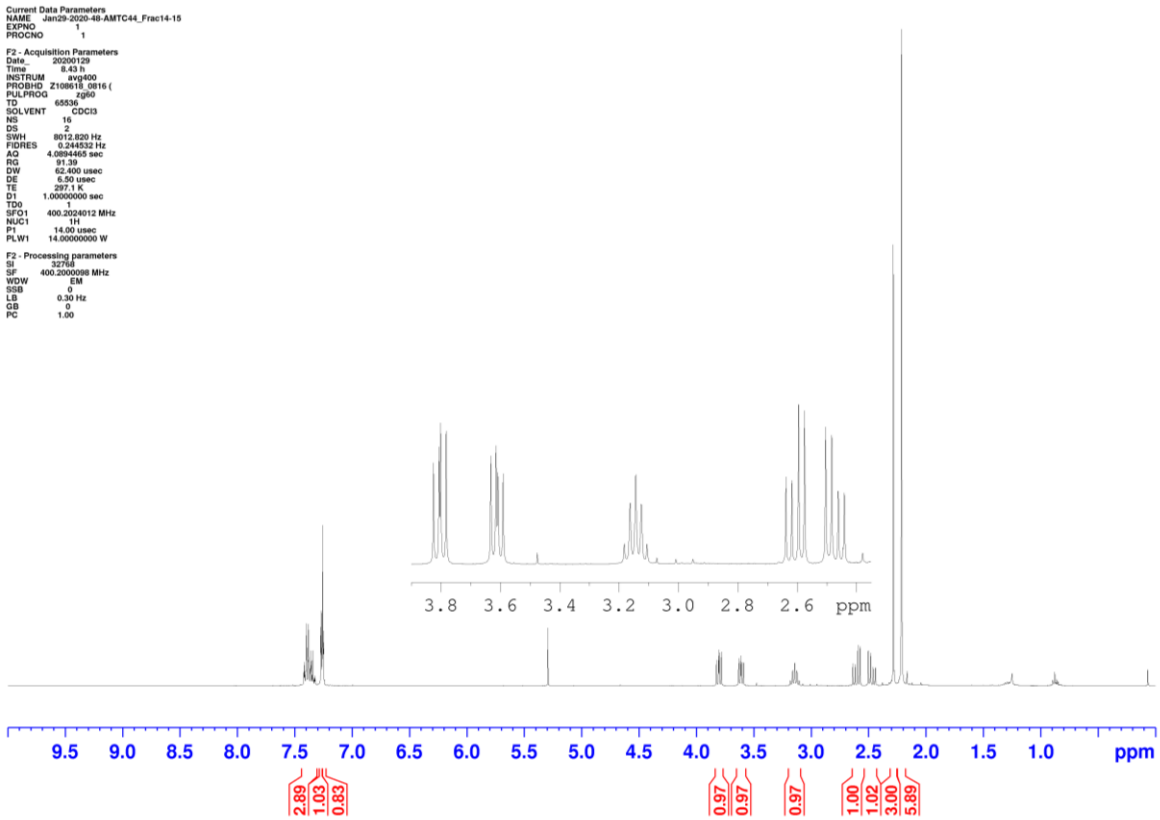


Figure C.1. ^{1}H NMR of **4.27**.

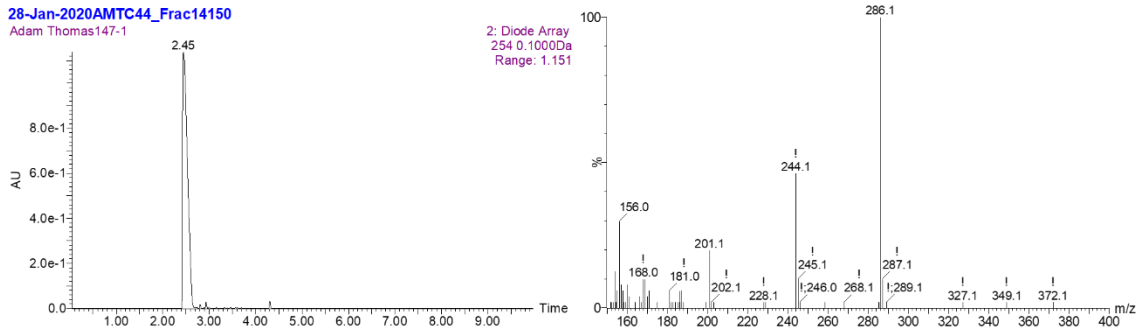
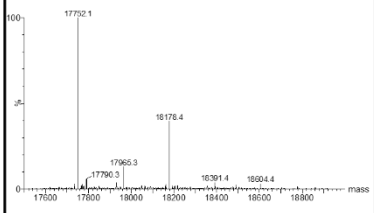
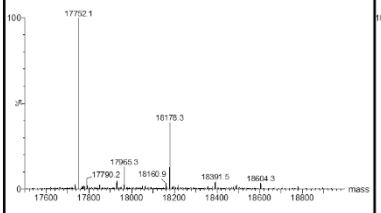
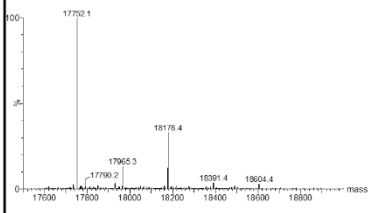
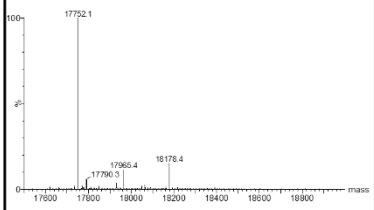
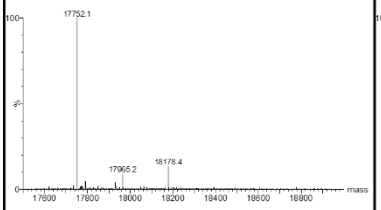
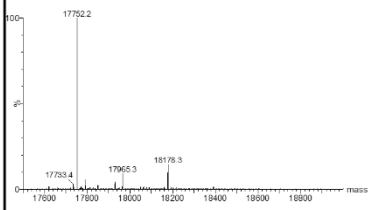
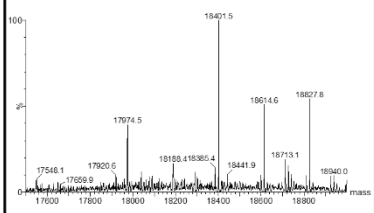
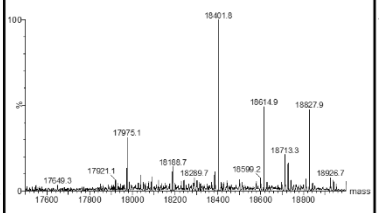
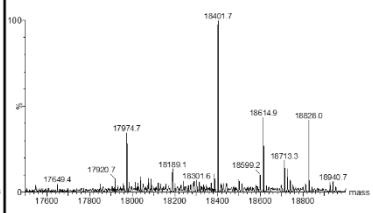
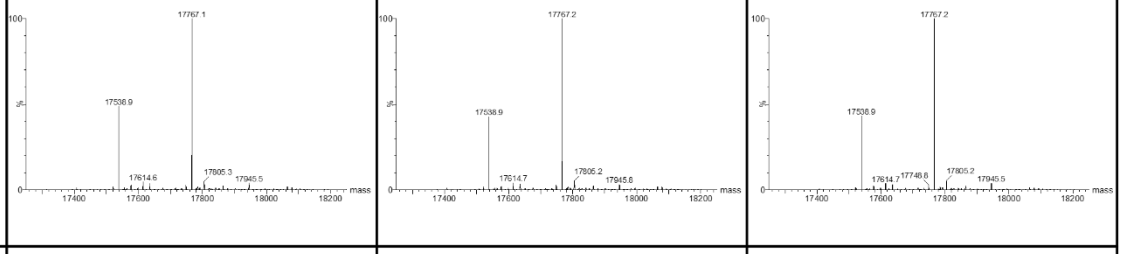
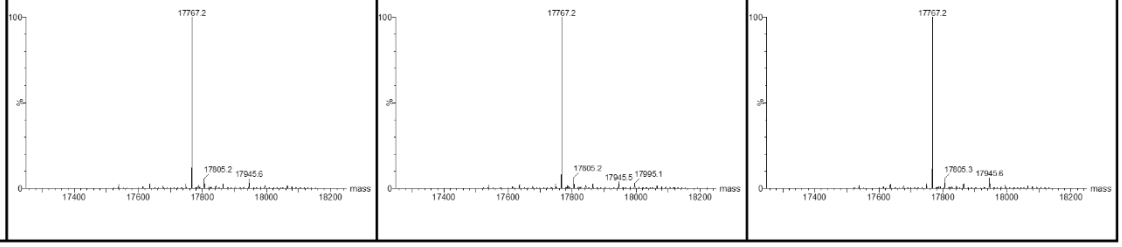


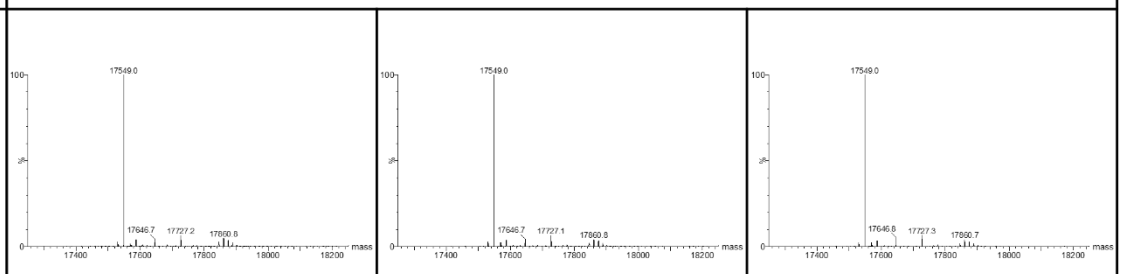
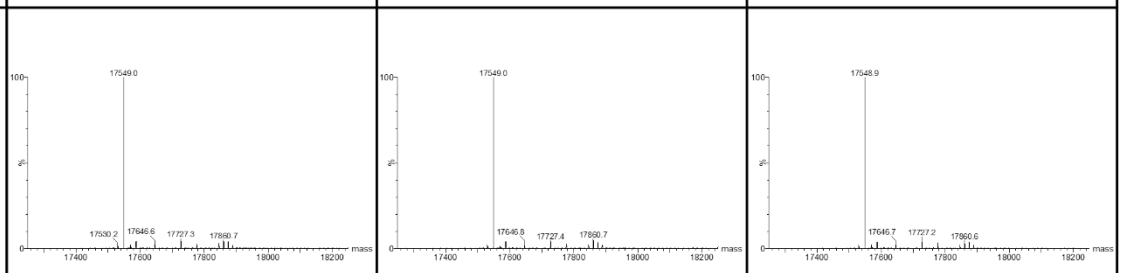
Figure C.2. LC-MS data for **4.27**.

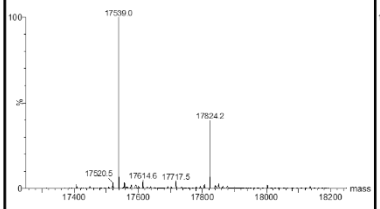
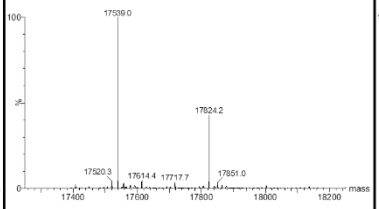
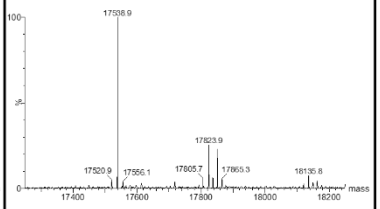
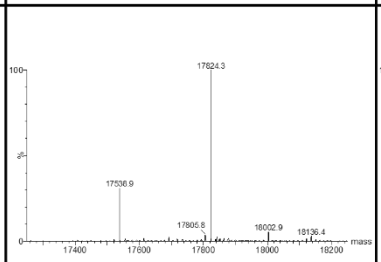
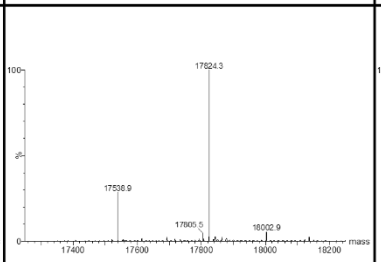
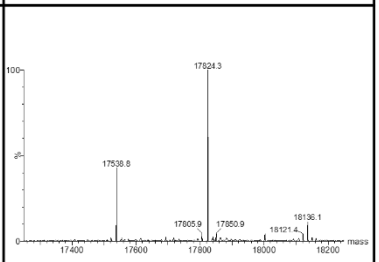
Deconvoluted mass spectra for the optimised DMI fragments

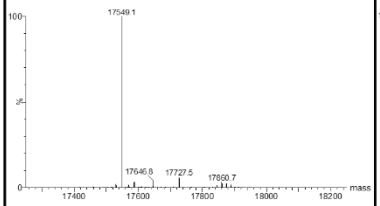
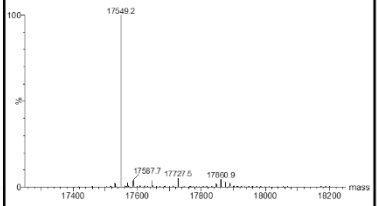
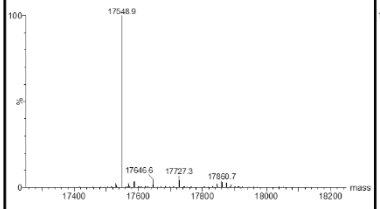
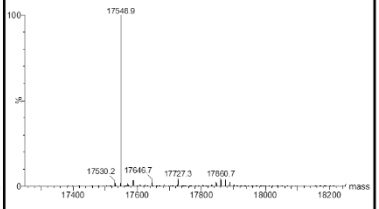
	4.1		
	BRD4(1)^{L94C}		
1 h			
4 h			

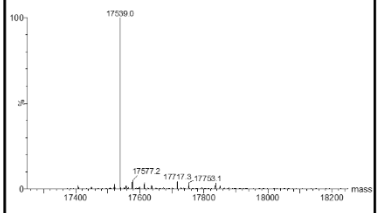
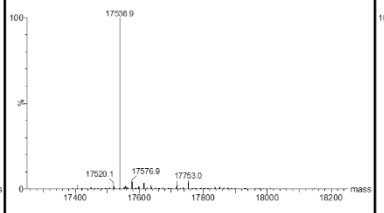
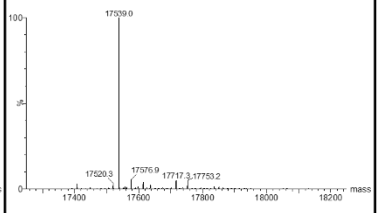
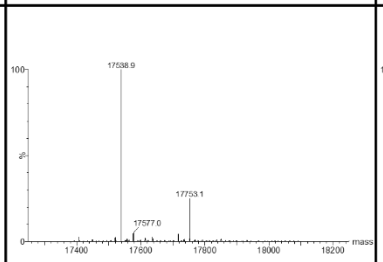
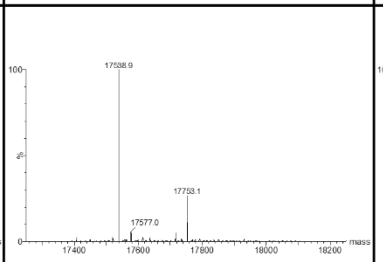
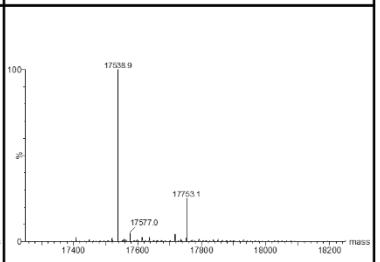
	4.1		
	BRD4(1)^{WT}		
1 h			

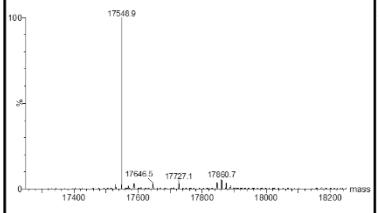
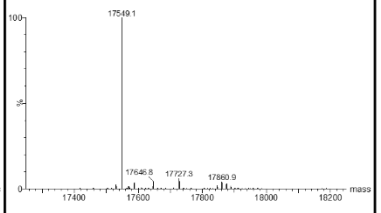
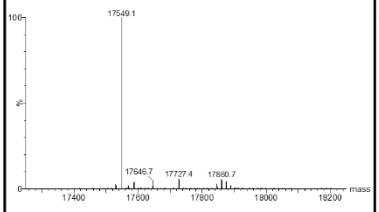
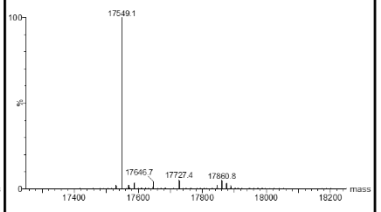
	4.2
	BRD4(1)^{L94C}
1 h	
4 h	

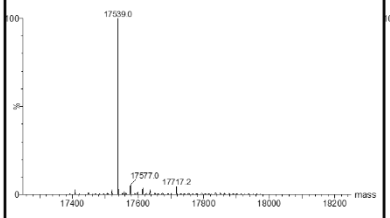
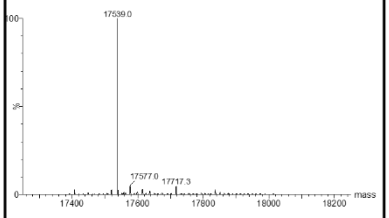
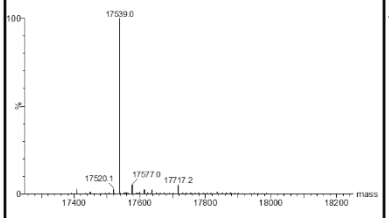
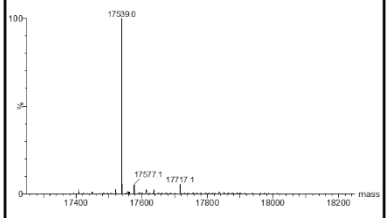
	4.2
	BRD4(1)^{WT}
1 h	
4 h	

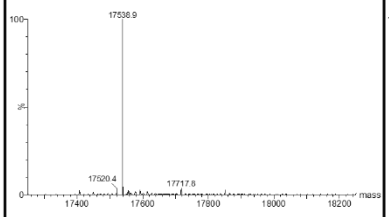
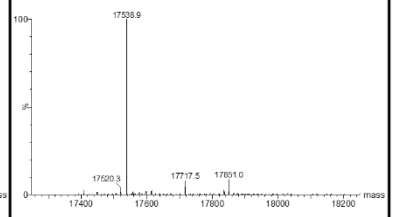
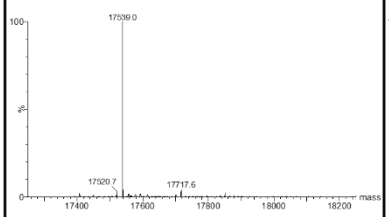
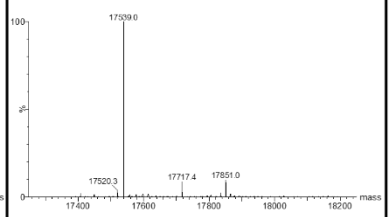
	4.3		
	BRD4(1)^{L94C}		
1 h			
4 h			

	4.3	
	BRD4(1)^{L94C}	
1 h		
4 h		

	4.4		
	BRD4(1)^{L94C}		
1 h			
4 h			

	4.4	
	BRD4(1)^{WT}	
1 h		
4 h		

	4.5	
	BRD4(1)^{L94C}	
1 h		
4 h		

	4.29	
	BRD4(1)^{L94C}	
1 h		
4 h		

AlphaScreen™ data

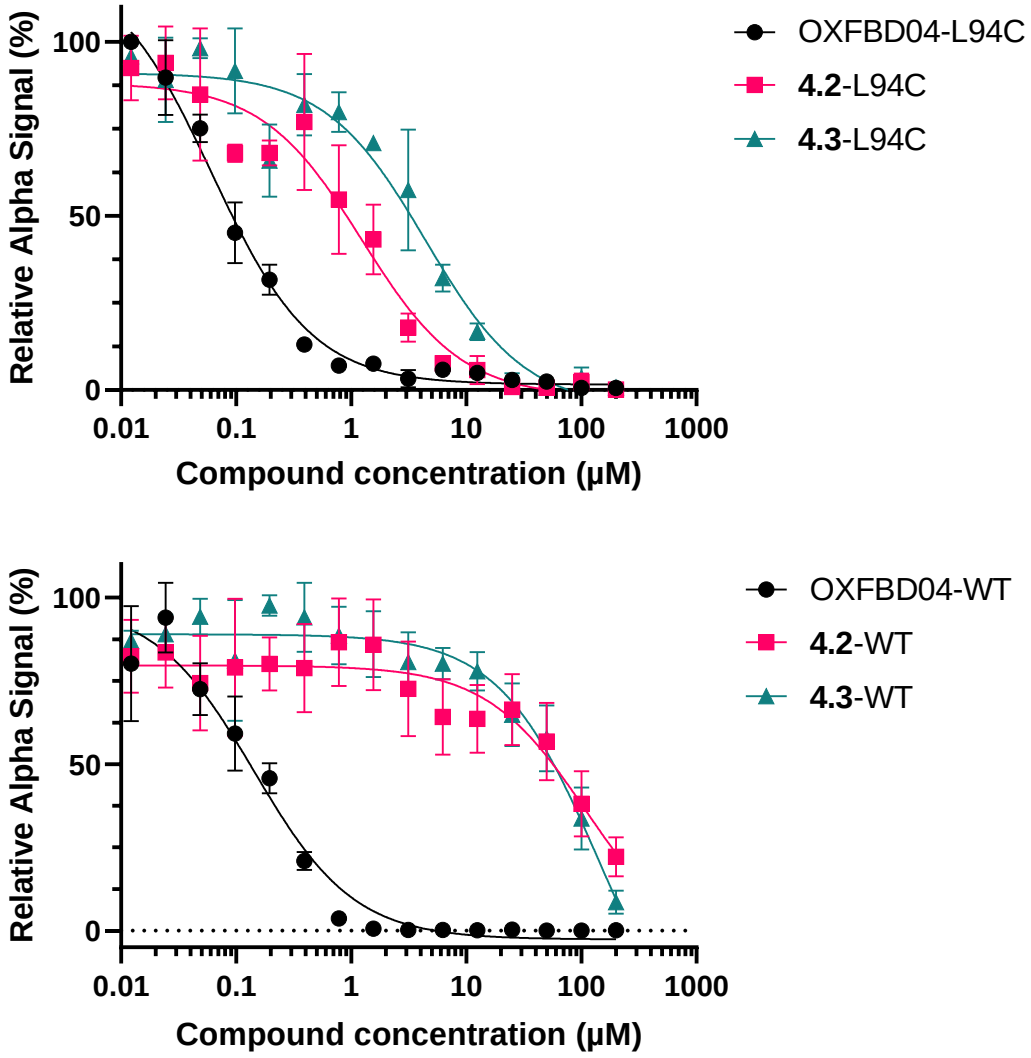
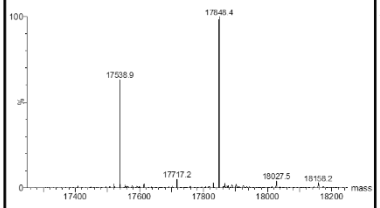
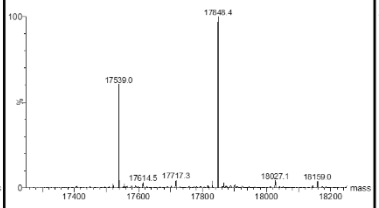
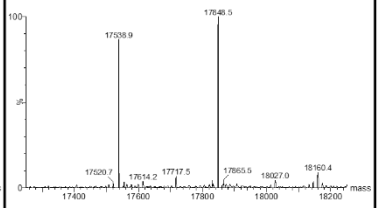
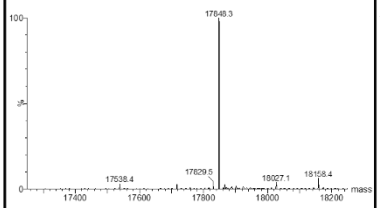
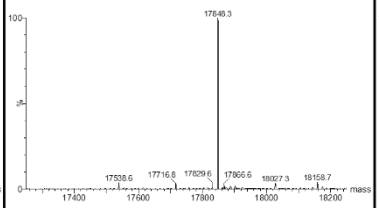
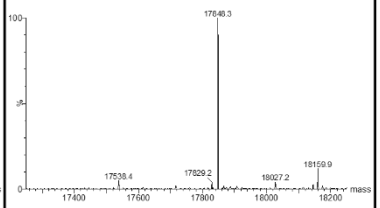
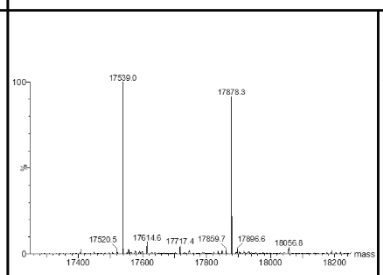
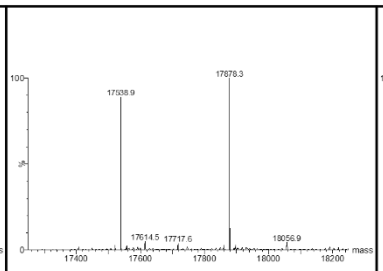
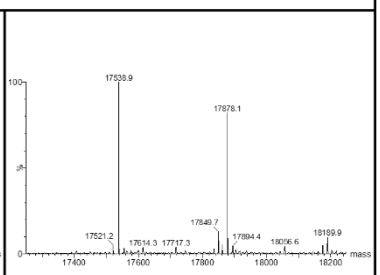
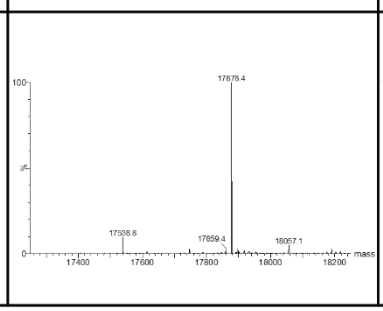
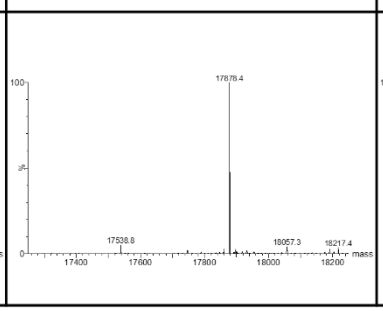
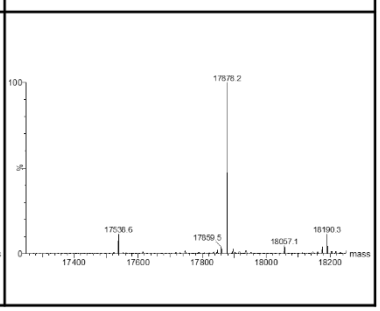


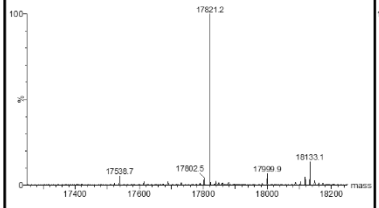
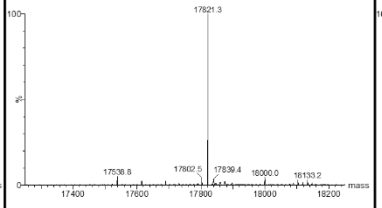
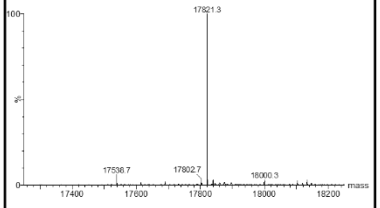
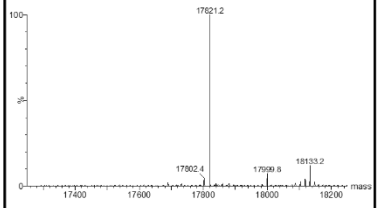
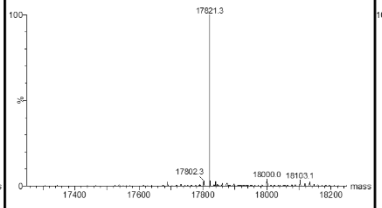
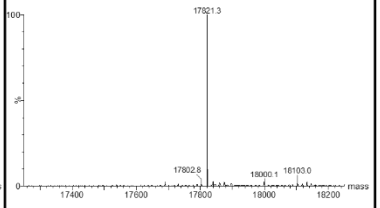
Figure C.3. IC_{50} curves for compounds against $BRD4(1)^{L94C}$ and $BRD4(1)^{WT}$.

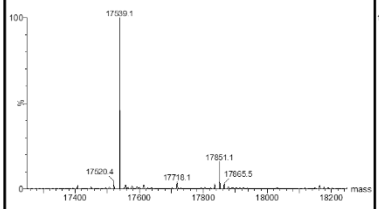
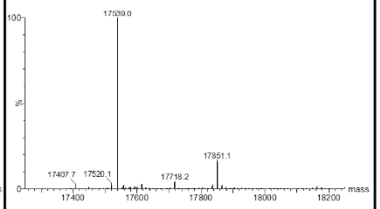
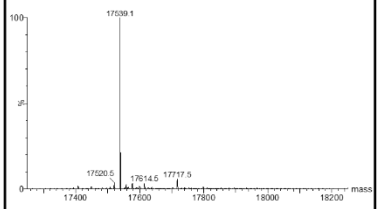
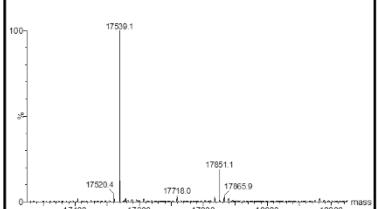
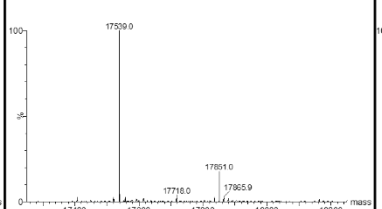
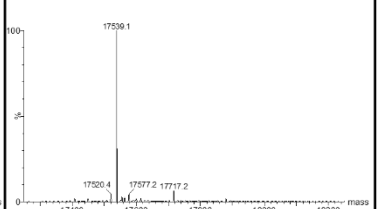
Appendix D: Chapter 5

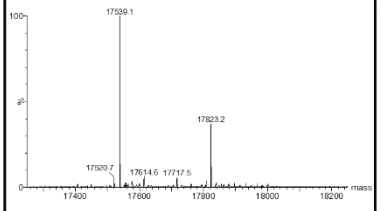
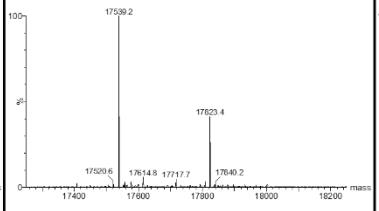
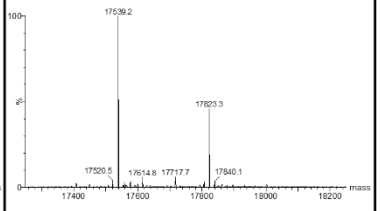
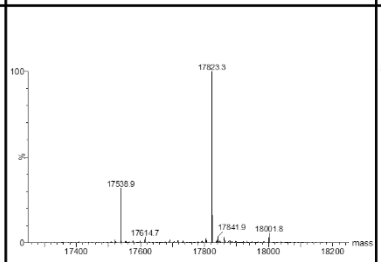
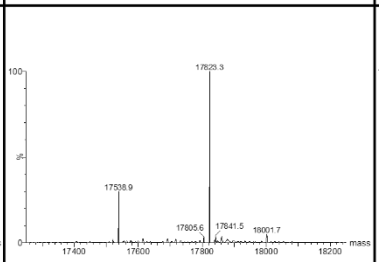
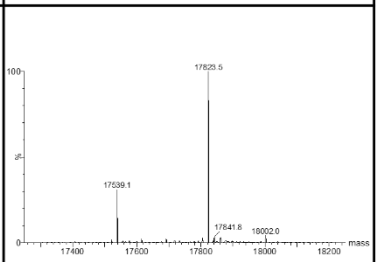
Deconvoluted mass spectra for the click probes

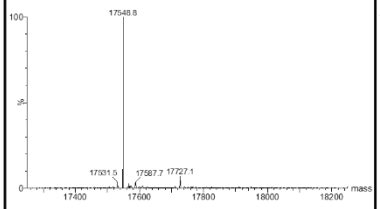
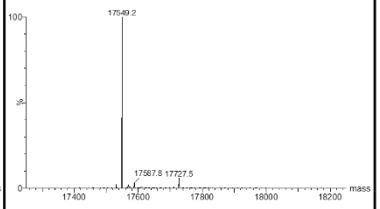
	5.1		
	BRD4(1)^{L94C}		
1 h			
4 h			

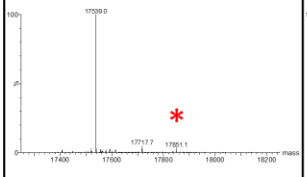
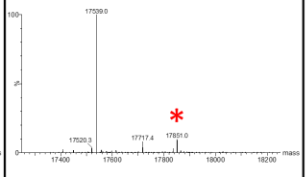
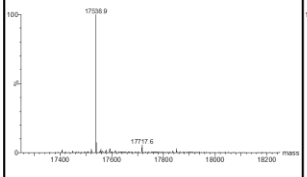
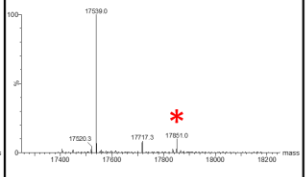
	5.2		
	BRD4(1)^{L94C}		
1 h			
4 h			

	5.3		
	BRD4(1)^{L94C}		
1 h			
4 h			

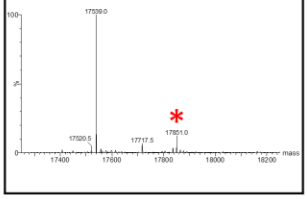
	5.7		
	BRD4(1)^{L94C}		
1 h			
4 h			

	5.8		
	BRD4(1)^{L94C}		
1 h			
4 h			

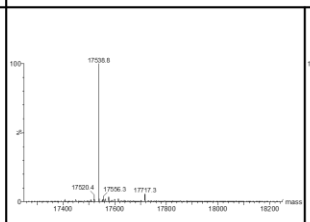
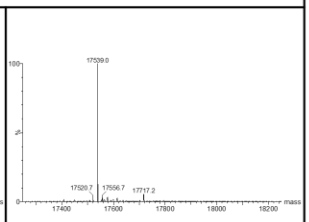
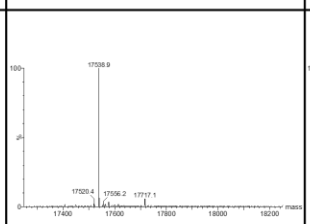
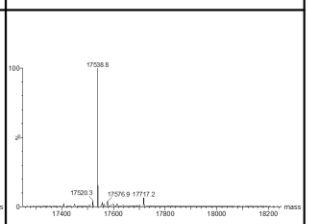
	5.8	
	BRD4(1)^{WT}	
4 h		

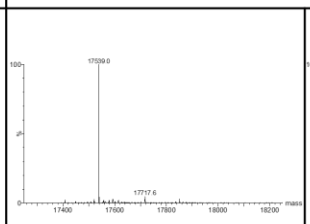
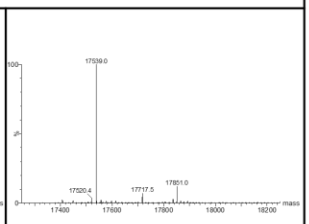
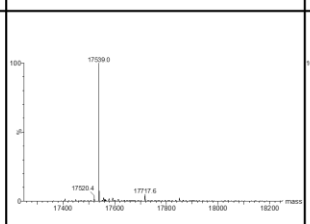
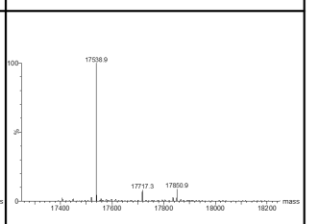
	5.4	
	BRD4(1)^{L94C}	
1 h		
4 h		

BRD4(1)^{L94C} 0 h Control



* Unknown +312 Da adduct.
Not related to **5.4**.
Seen in protein only sample.

	5.5	
	BRD4(1)^{L94C}	
1 h		
4 h		

	5.6	
	BRD4(1)^{L94C}	
1 h		
4 h		

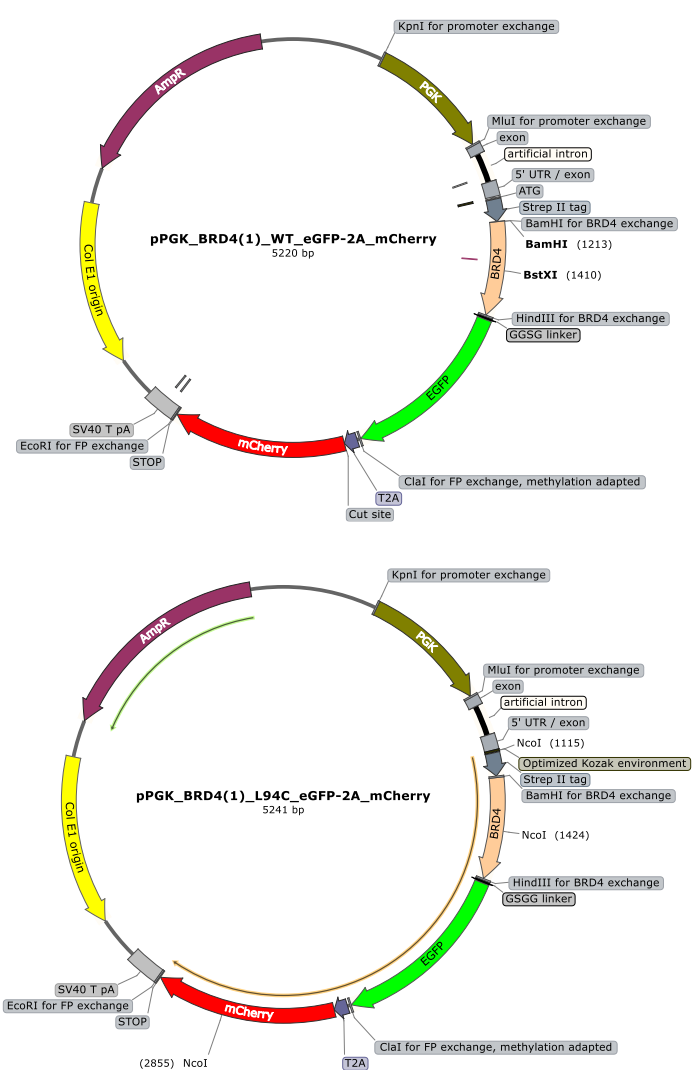


Figure D.1. Plasmid maps for the constructs used to transiently express BRD4(1) in HEK293T cells.

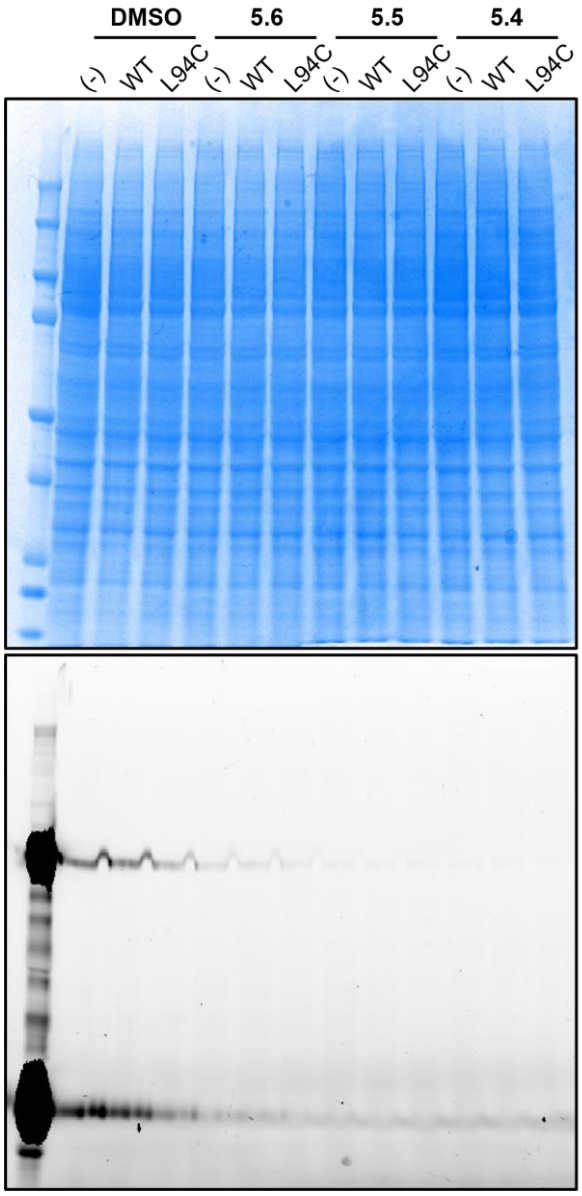
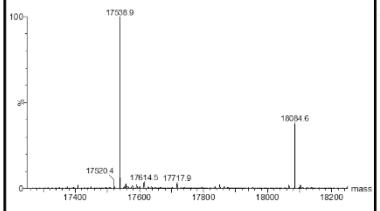
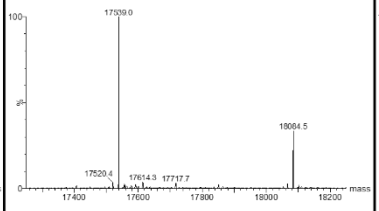
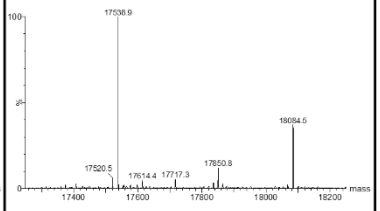
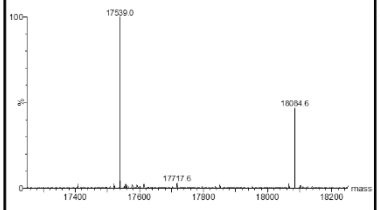
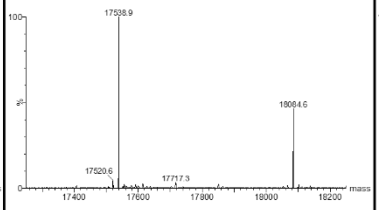
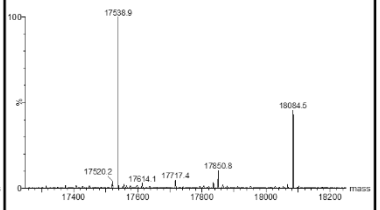
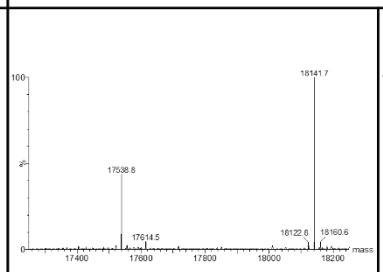
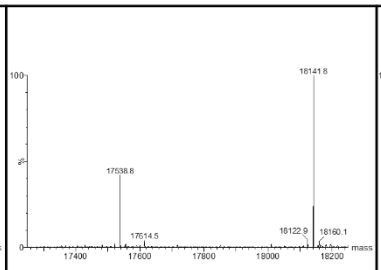
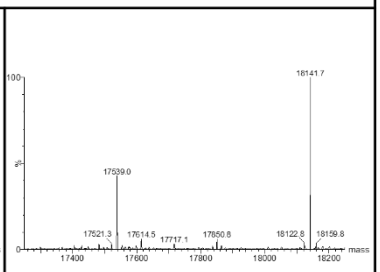
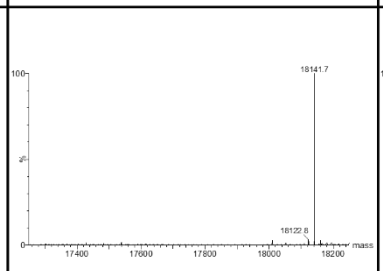
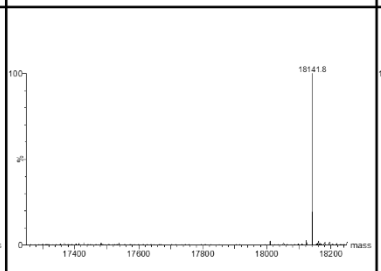
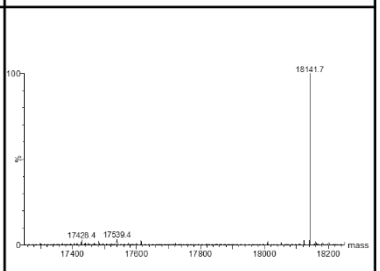


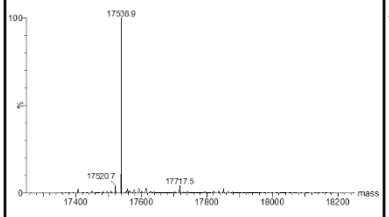
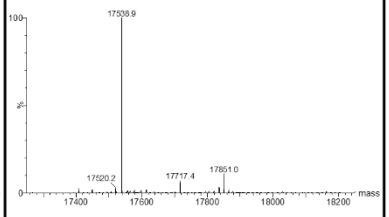
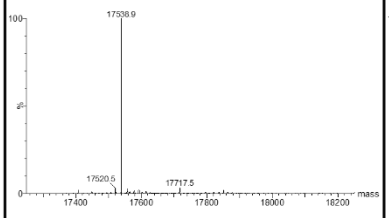
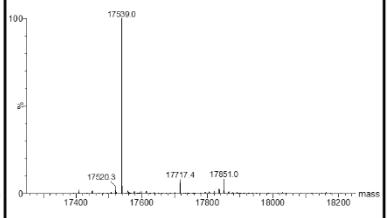
Figure D.2. In-cell protein labelling by saturated negative control compounds **5.4**, **5.5**, and **5.6**.

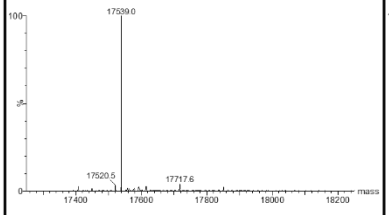
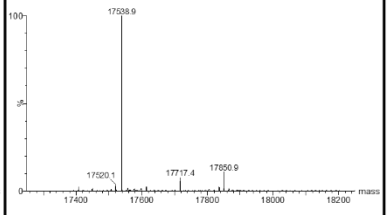
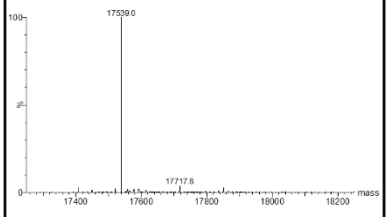
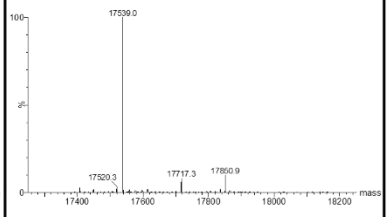
Appendix E: Chapter 6

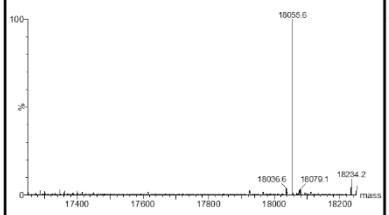
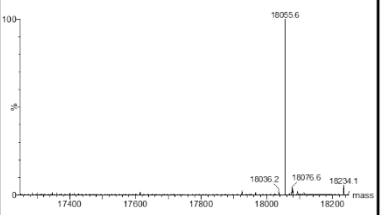
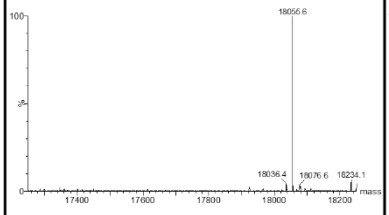
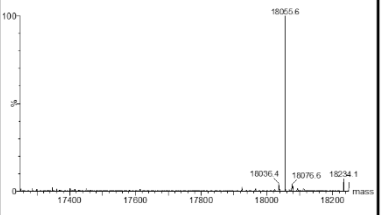
Deconvoluted mass spectra for the HyTags

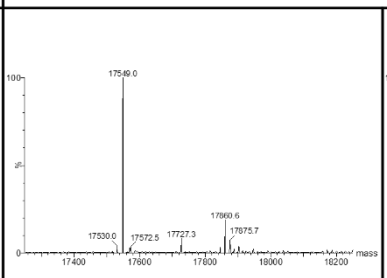
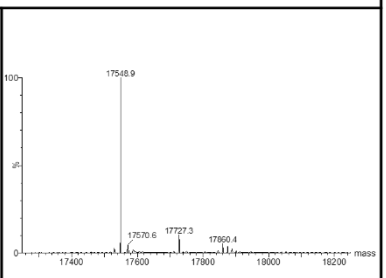
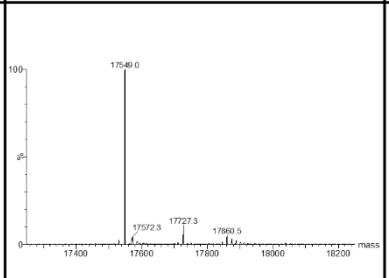
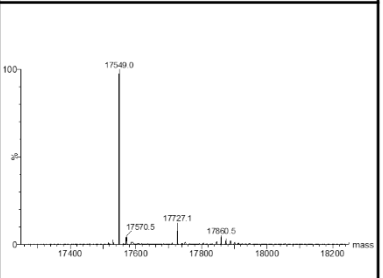
	6.1		
	BRD4(1)^{L94C}		
1 h			
4 h			

	6.2		
	BRD4(1)^{L94C}		
1 h			
4 h			

	6.11	
	BRD4(1)^{L94C}	
1 h		
4 h		

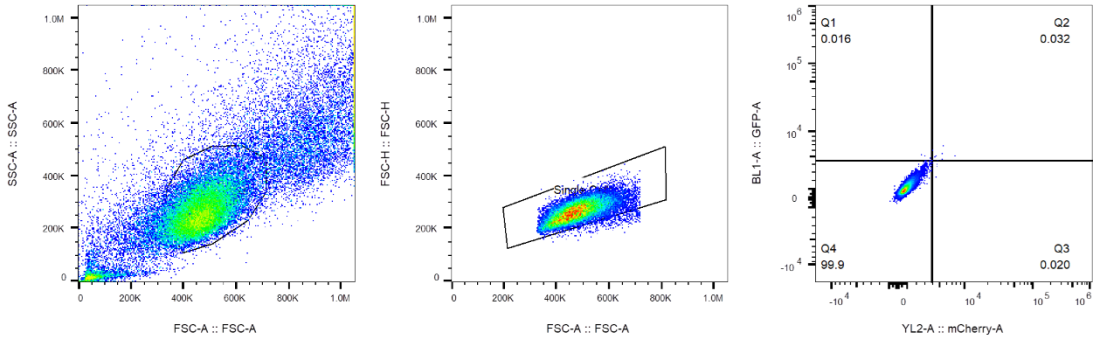
	6.12	
	BRD4(1)^{L94C}	
1 h		
4 h		

	6.13	
	BRD4(1)^{L94C}	
1 h		
4 h		

	6.13	
	BRD4(1)^{WT}	
1 h		
4 h		

Flow cytometry data for the HyTags

DMSO (-)

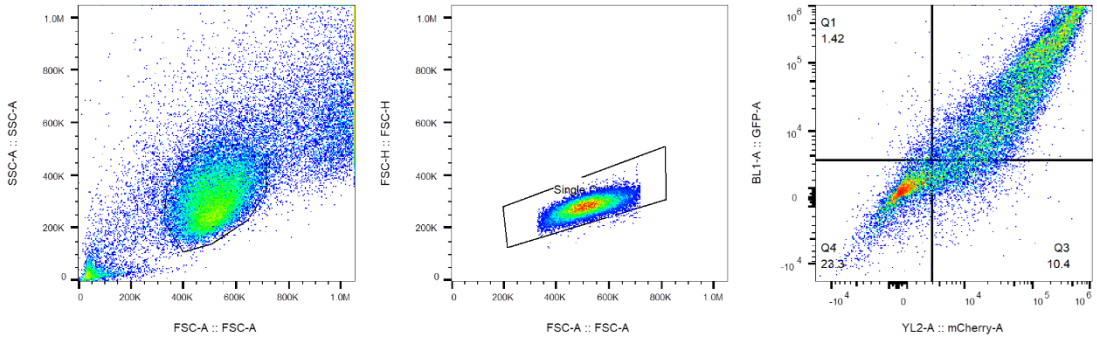


DMSO-.fcs
Ungated
55085

DMSO-.fcs
Lymphocytes
25263

DMSO-.fcs
Single Cells
24761

DMSO WT

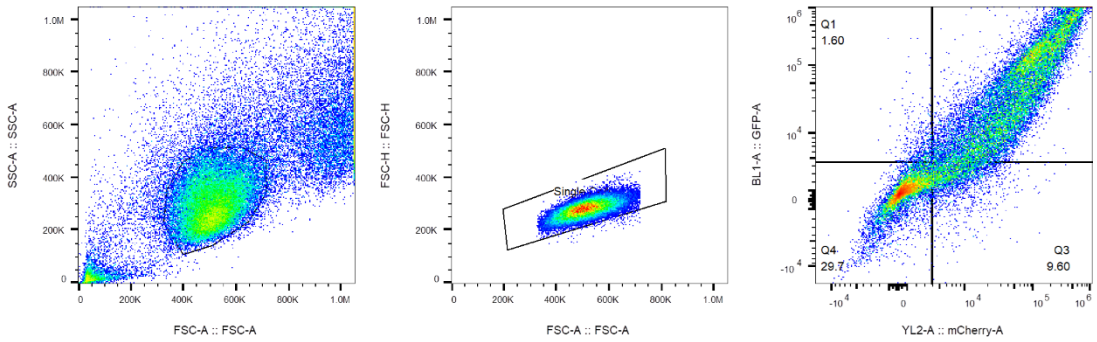


DMSO WT-.fcs
Ungated
53793

DMSO WT-.fcs
Lymphocytes
24524

DMSO WT-.fcs
Single Cells
24424

DMSO L94C

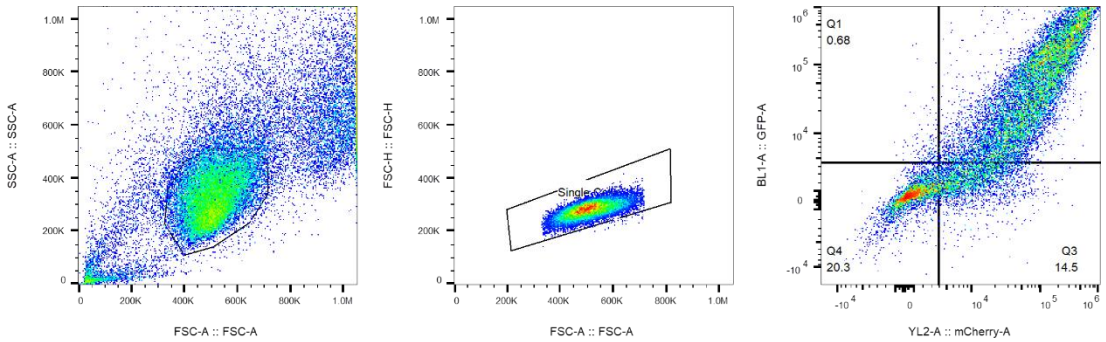


DMSO LC-.fcs
Ungated
70819

DMSO LC-.fcs
Lymphocytes
34345

DMSO LC-.fcs
Single Cells
34190

6.1 WT

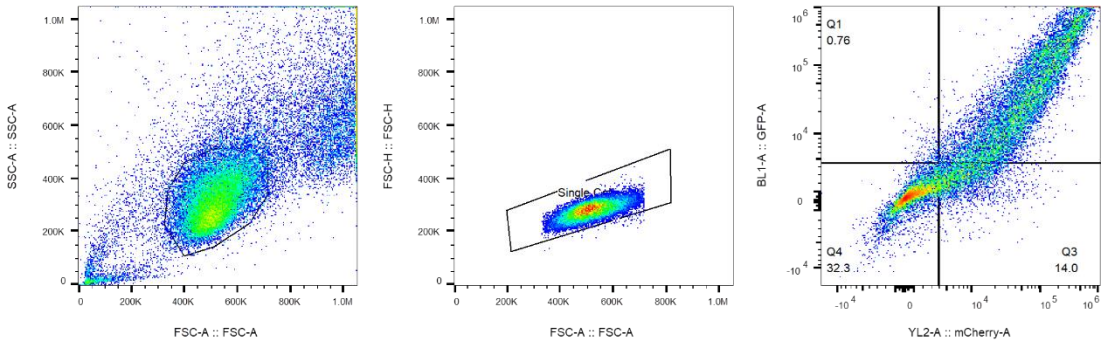


E69 WT-.fcs
Ungated
50000

E69 WT-.fcs
Lymphocytes
22317

E69 WT-.fcs
Single Cells
22250

6.1 L94C

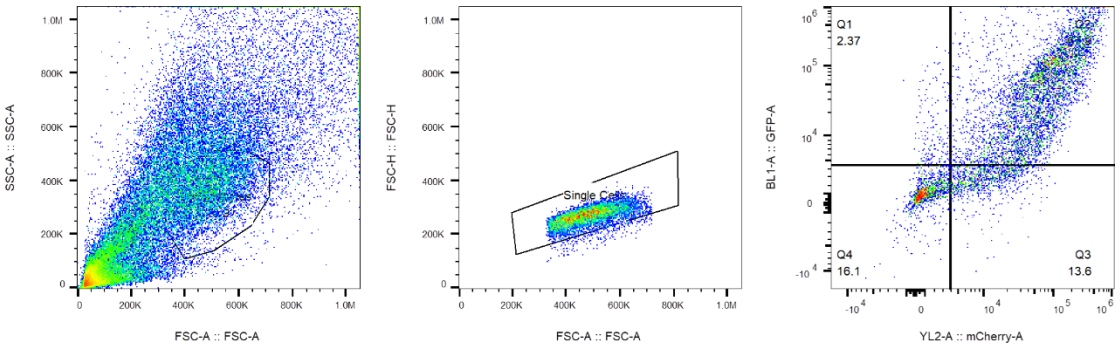


E69 LC-.fcs
Ungated
50000

E69 LC-.fcs
Lymphocytes
24843

E69 LC-.fcs
Single Cells
24746

6.2 WT

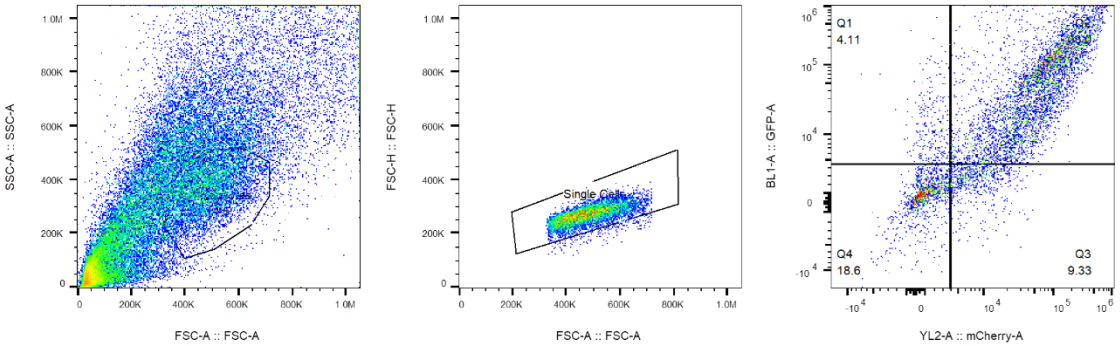


E86 WT-.fcs
Ungated
50000

E86 WT-.fcs
Lymphocytes
7962

E86 WT-.fcs
Single Cells
7632

6.2 L94C

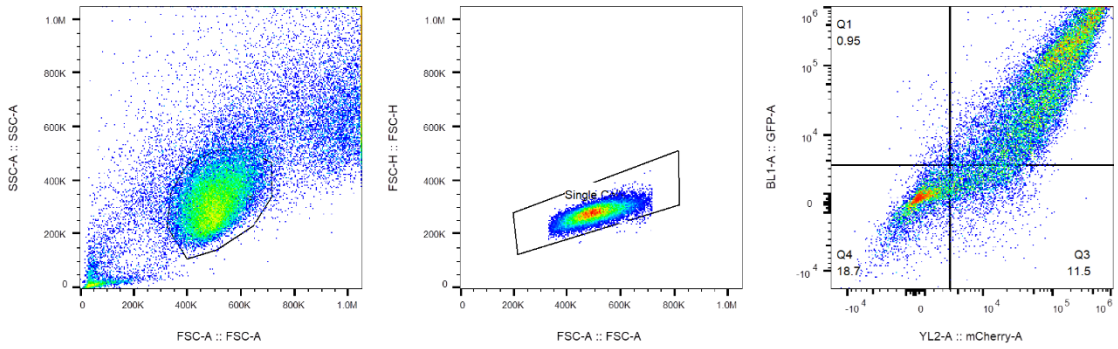


E86 LC-.fcs
Ungated
50000

E86 LC-.fcs
Lymphocytes
7265

E86 LC-.fcs
Single Cells
7030

6.11 WT

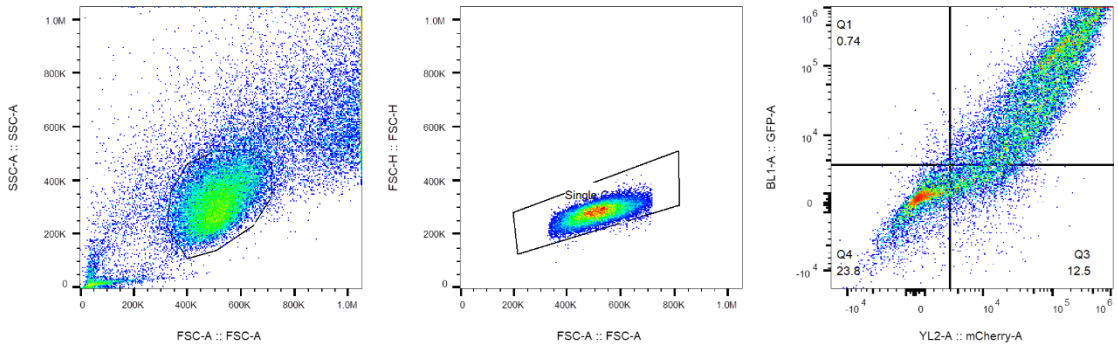


E79 WT-.fcs
Ungated
50000

E79 WT-.fcs
Lymphocytes
25214

E79 WT-.fcs
Single Cells
25161

6.11 L94C

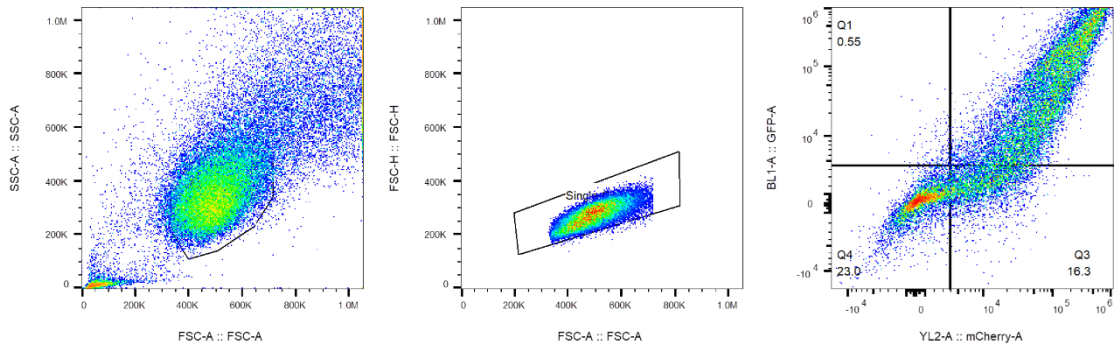


E79 LC-.fcs
Ungated
50000

E79 LC-.fcs
Lymphocytes
21665

E79 LC-.fcs
Single Cells
21564

6.12 WT

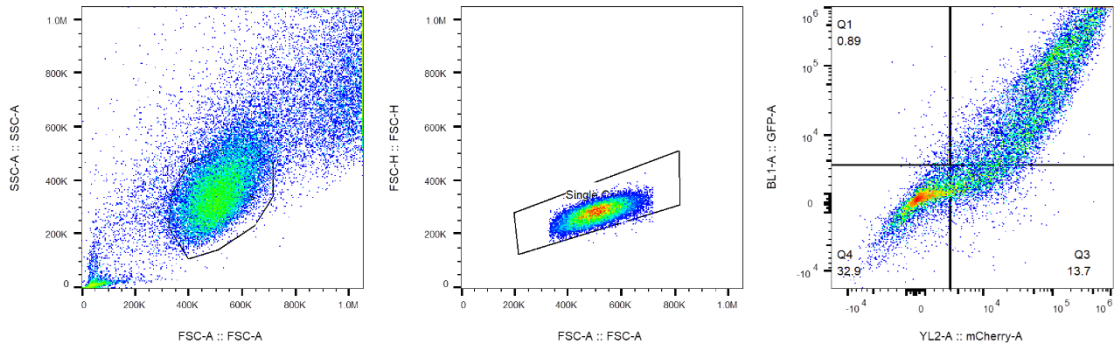


E81 WT.fcs
Ungated
50000

E81 WT.fcs
Lymphocytes
25505

E81 WT.fcs
Single Cells
25280

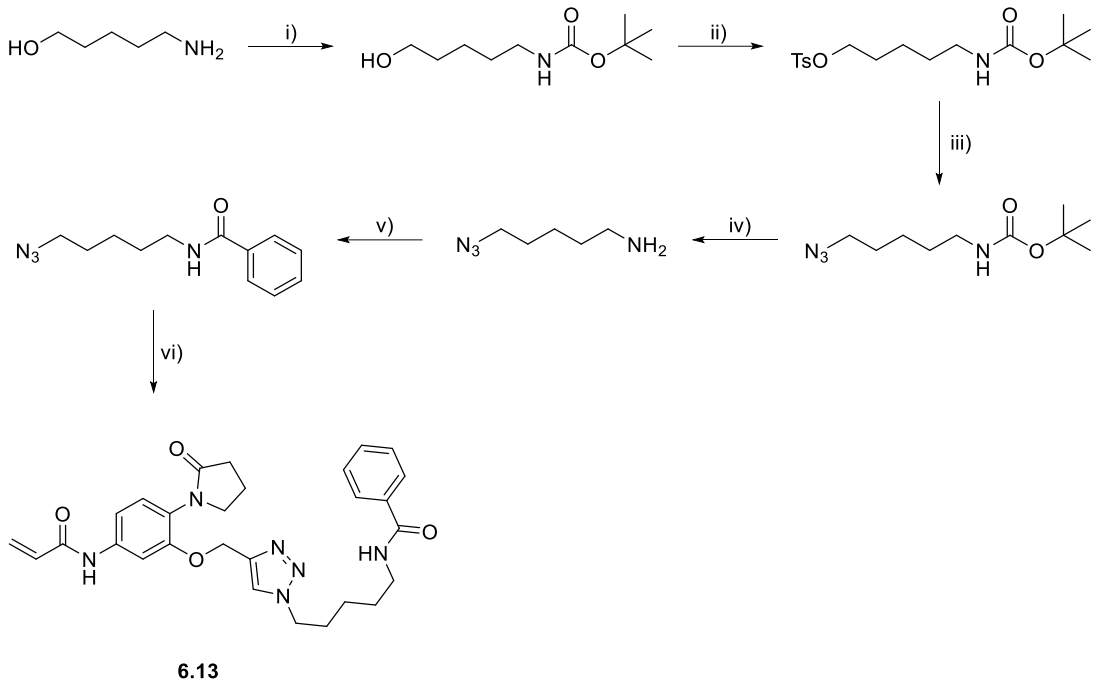
6.12 L94C



E81 LC-.fcs
Ungated
50000

E81 LC-.fcs
Lymphocytes
18481

E81 LC-.fcs
Single Cells
18406



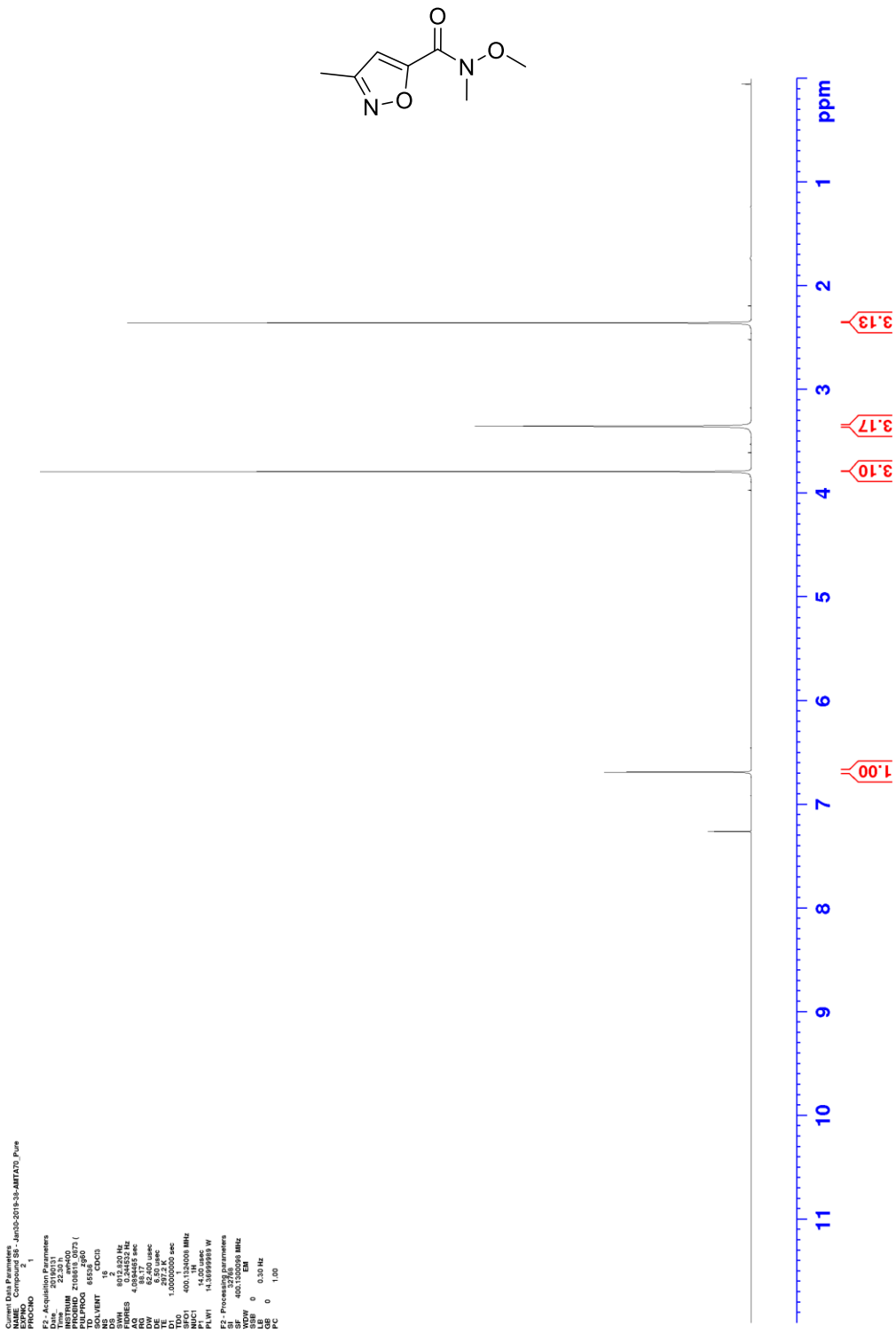
6.13

Scheme E.1. Synthesis of **E63** conducted by Dr Marta Serafini. *Reagents and conditions*: i) Boc_2O, CH_2Cl_2, rt, 16 h, 99%; ii) TsCl, Et_3N, CH_2Cl_2, rt, 16 h, 77%; iii) NaN_3, EtOH, 70 °C, 7 h, 87%; iv) HCl (4.0 M in dioxane), CH_2Cl_2, rt, 1 h; v) Benzoic acid, HATU, DIPEA, DMF, 16 h, 50%; vi) **3.50**, NaAsc, $CuSO_4 \cdot 5H_2O$, H_2O, *t*BuOH, rt, 16 h, 99%.

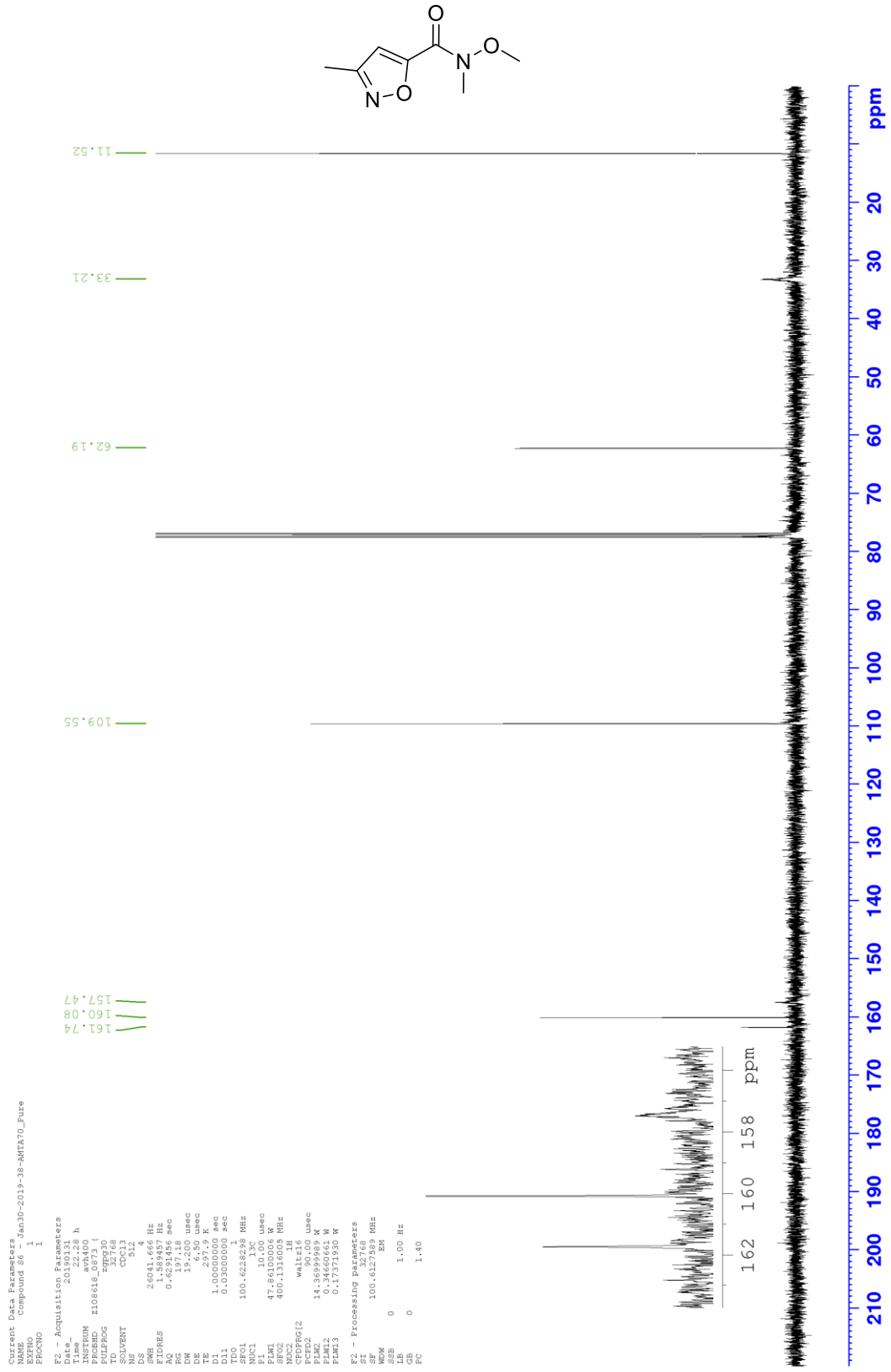
Appendix F: NMR spectra of novel compounds

Chapter 3 spectra

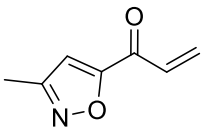
***N*-Methoxy-*N*-3-dimethylisoxazole-5-carboxamide**

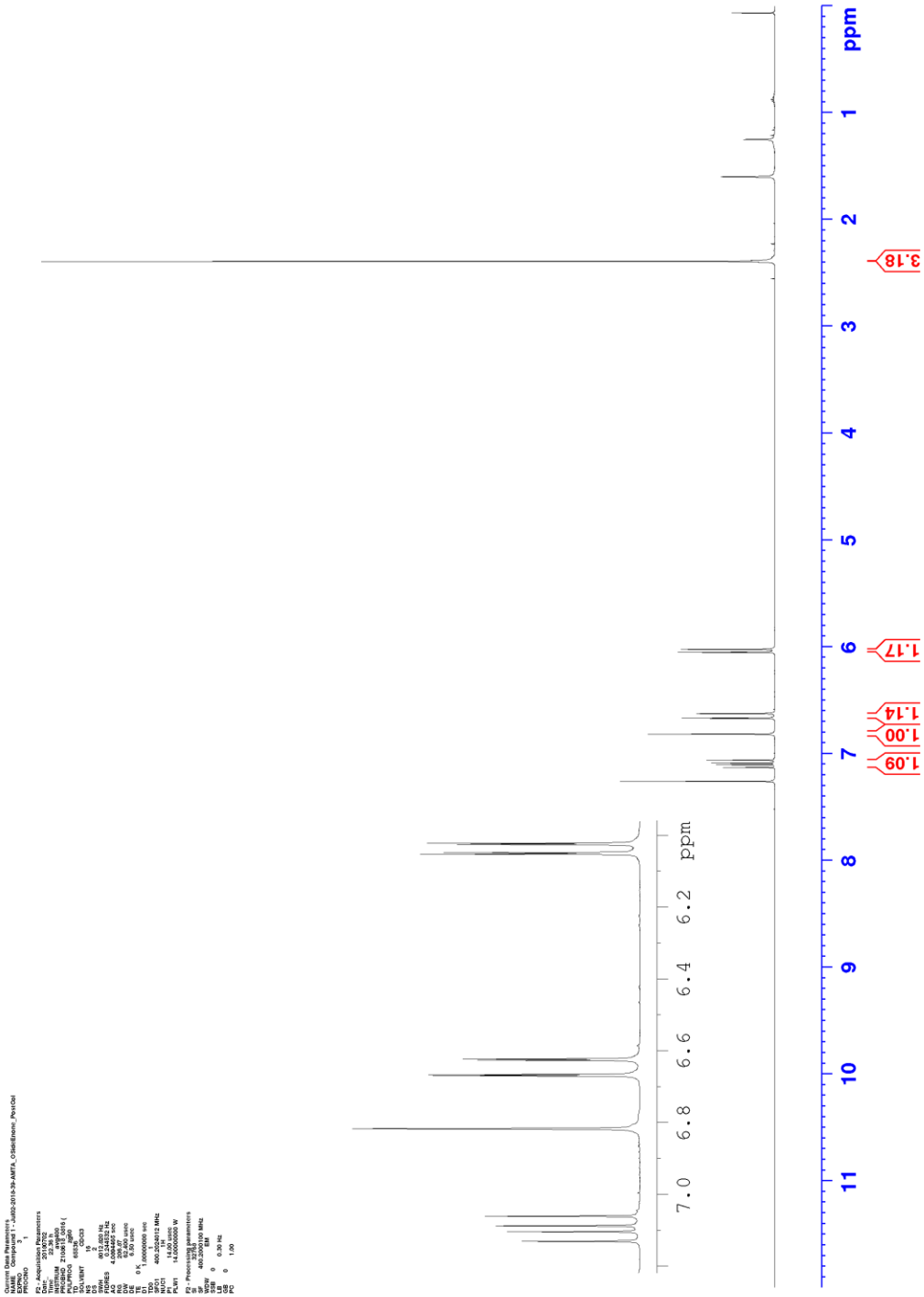


_N_-Methoxy-_N_-3-dimethylisoxazole-5-carboxamide

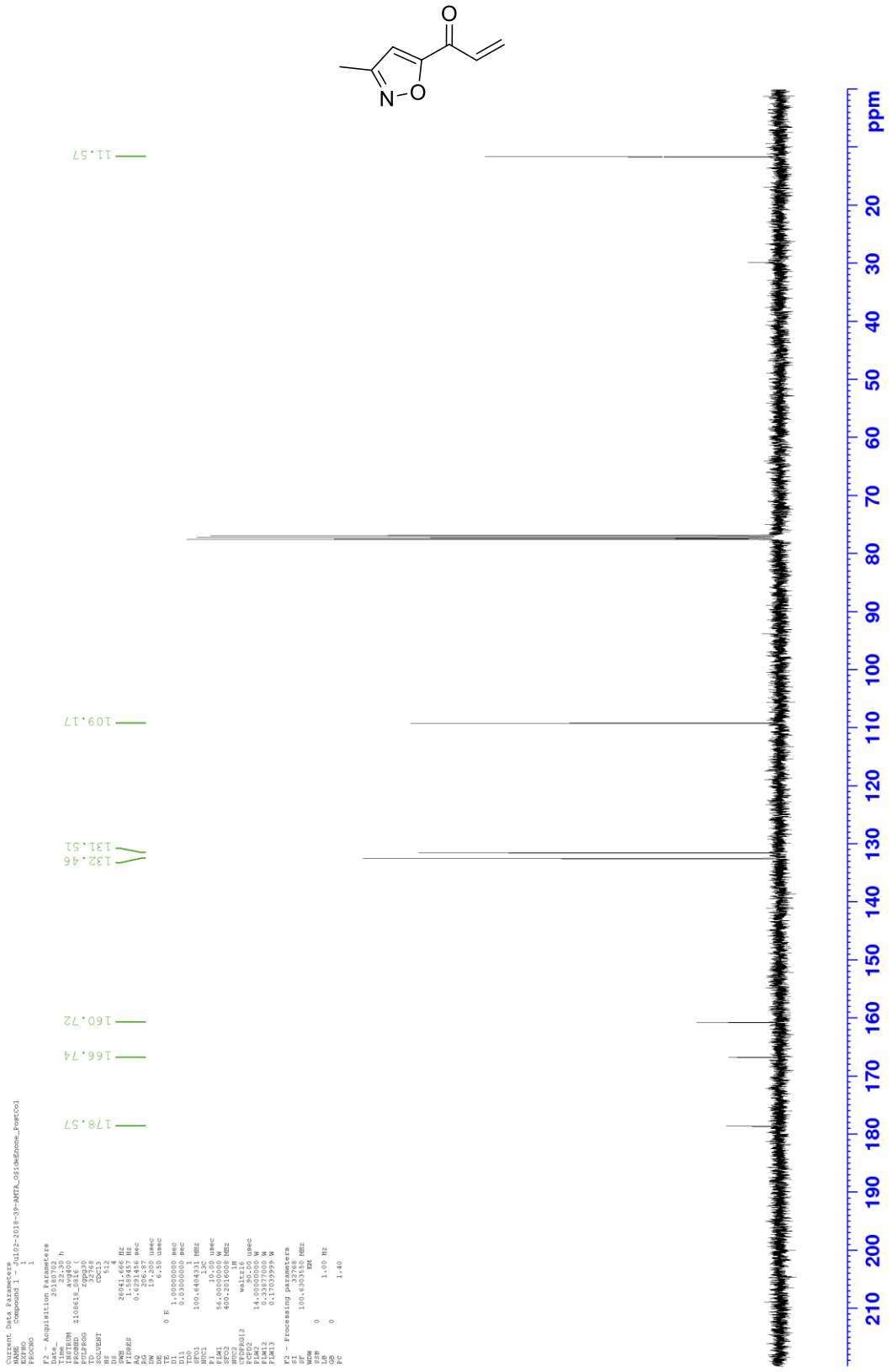


1^{ii}-(3-Methylisoxazol-5-yl)prop-2^{ii}-en-1^{ii}-one

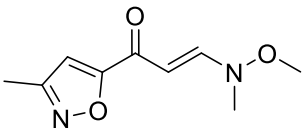


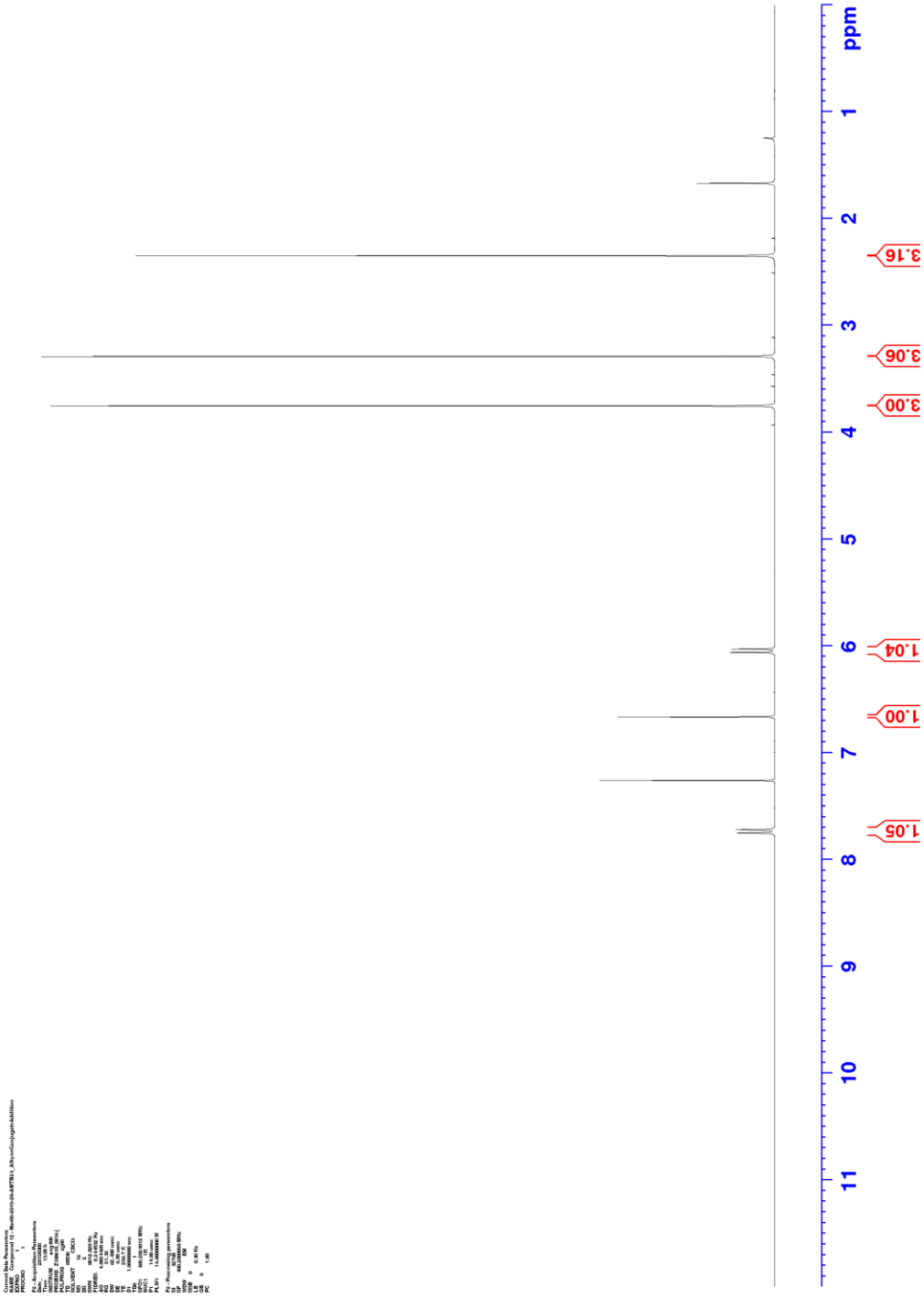


1^{ii}-(3-Methylisoxazol-5-yl)prop-2^{ii}-en-1^{ii}-one

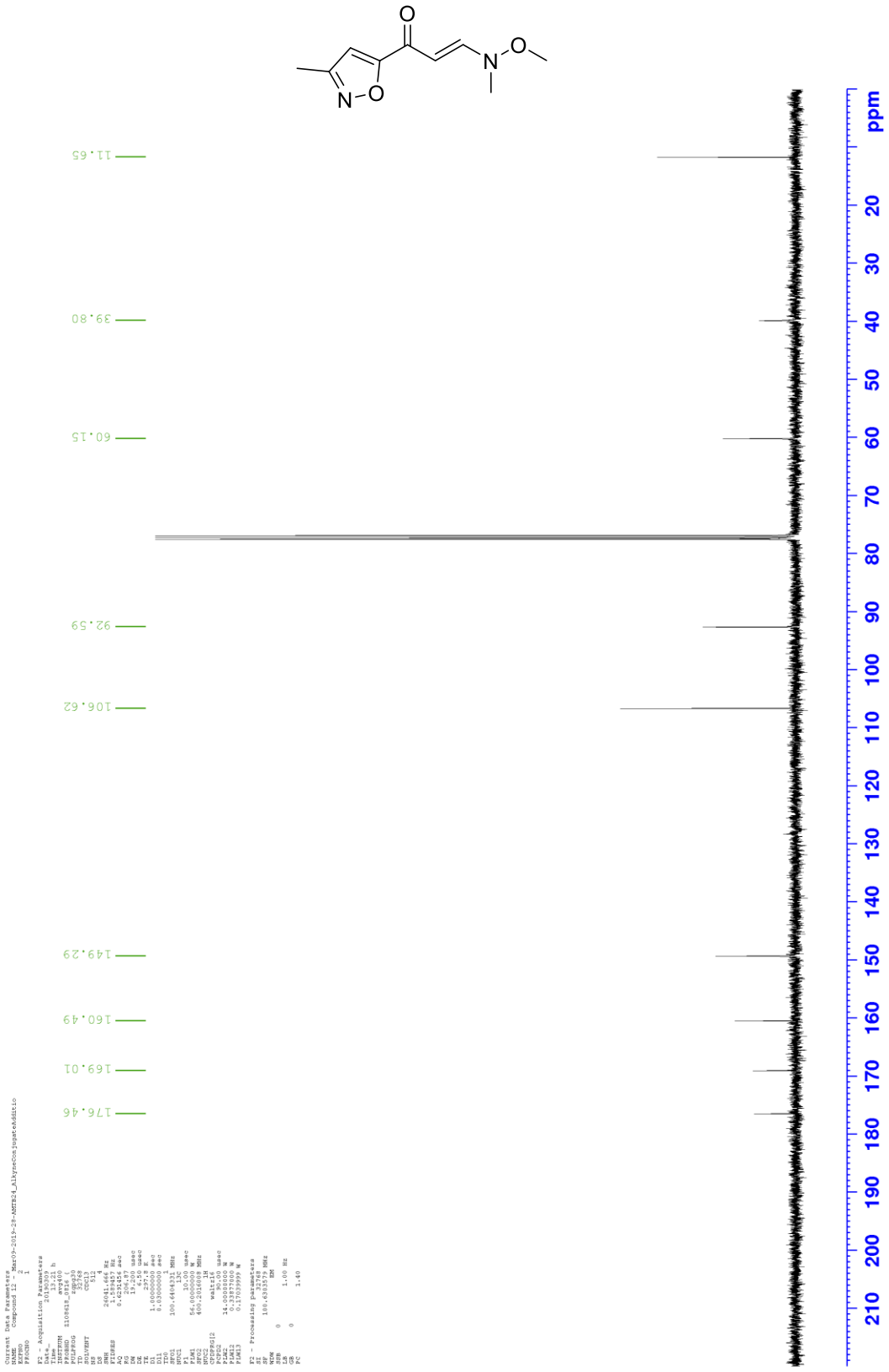


(*E*)-3[ii]-[Methoxy(methyl)amino]-1[ii]-(3-methylisoxazol-5-yl)prop-2[ii]-en-1[ii]-one

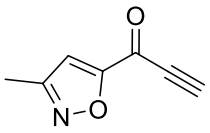


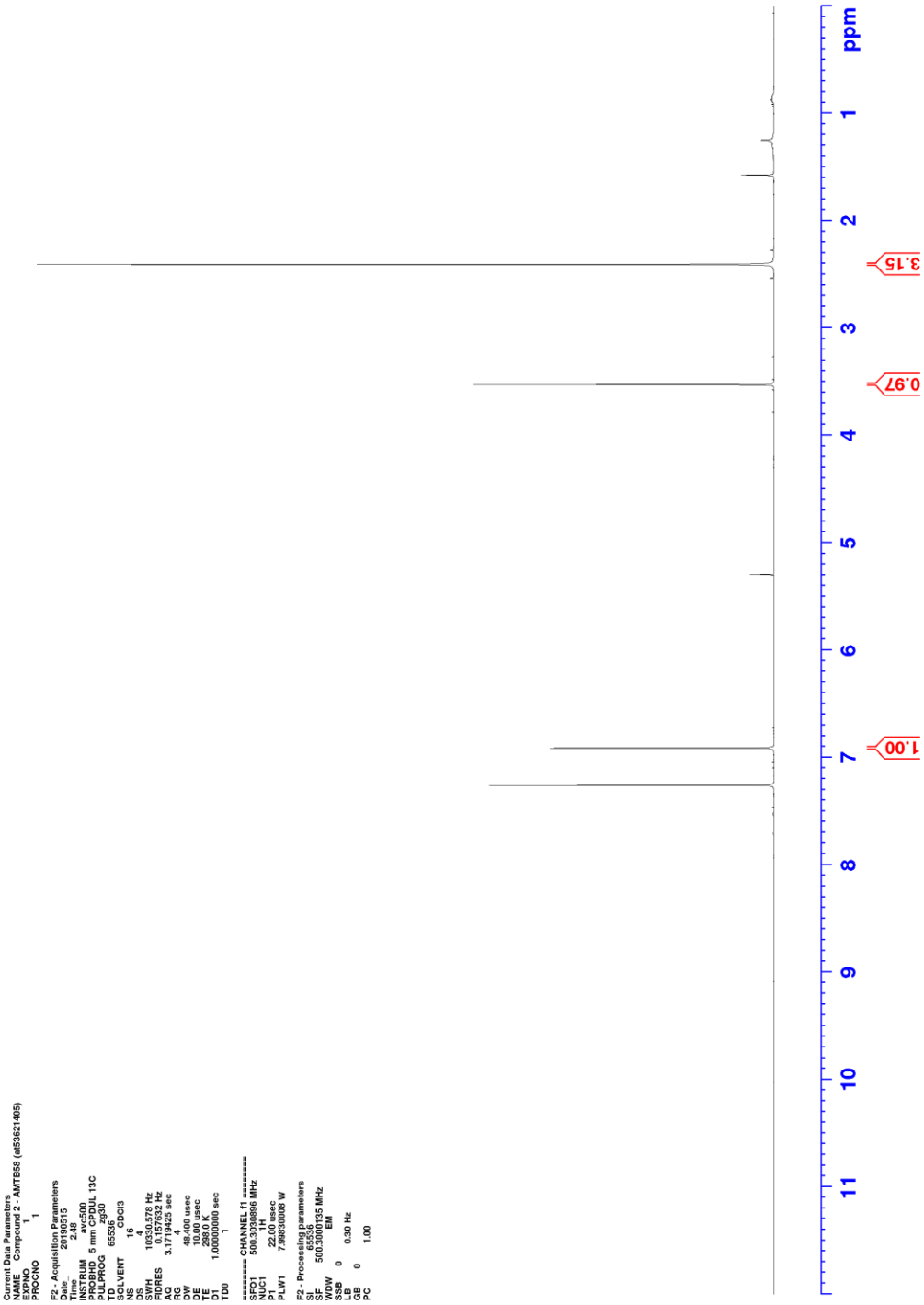


(*E*)-3^{ii}-[Methoxy(methyl)amino]-1^{ii}-(3-methylisoxazol-5-yl)prop-2^{ii}-en-1^{ii}-one



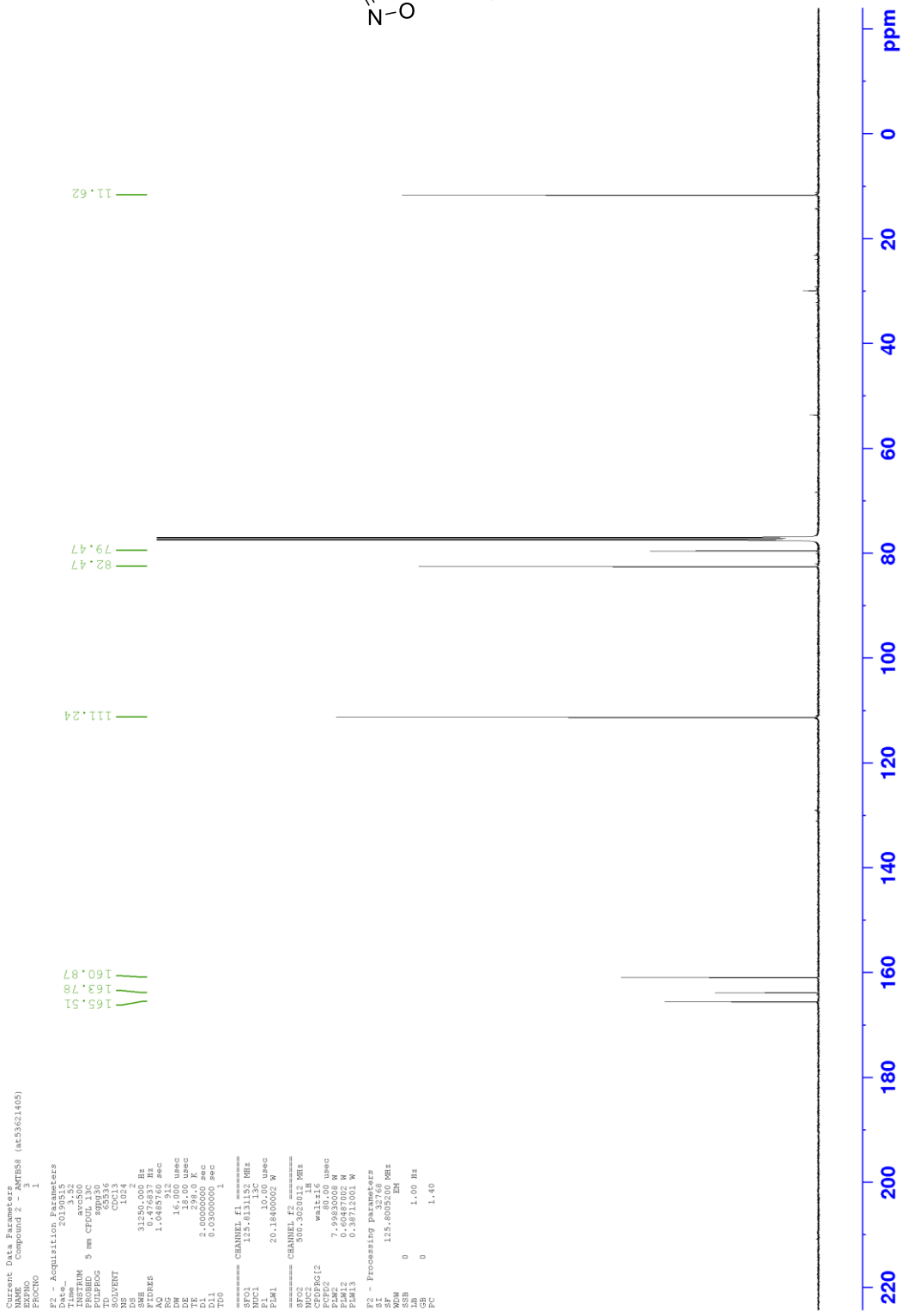
1^{ii}-(3-Methylisoxazol-5-yl)prop-2^{ii}-yn-1^{ii}-one



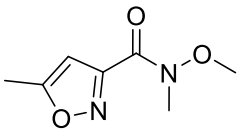


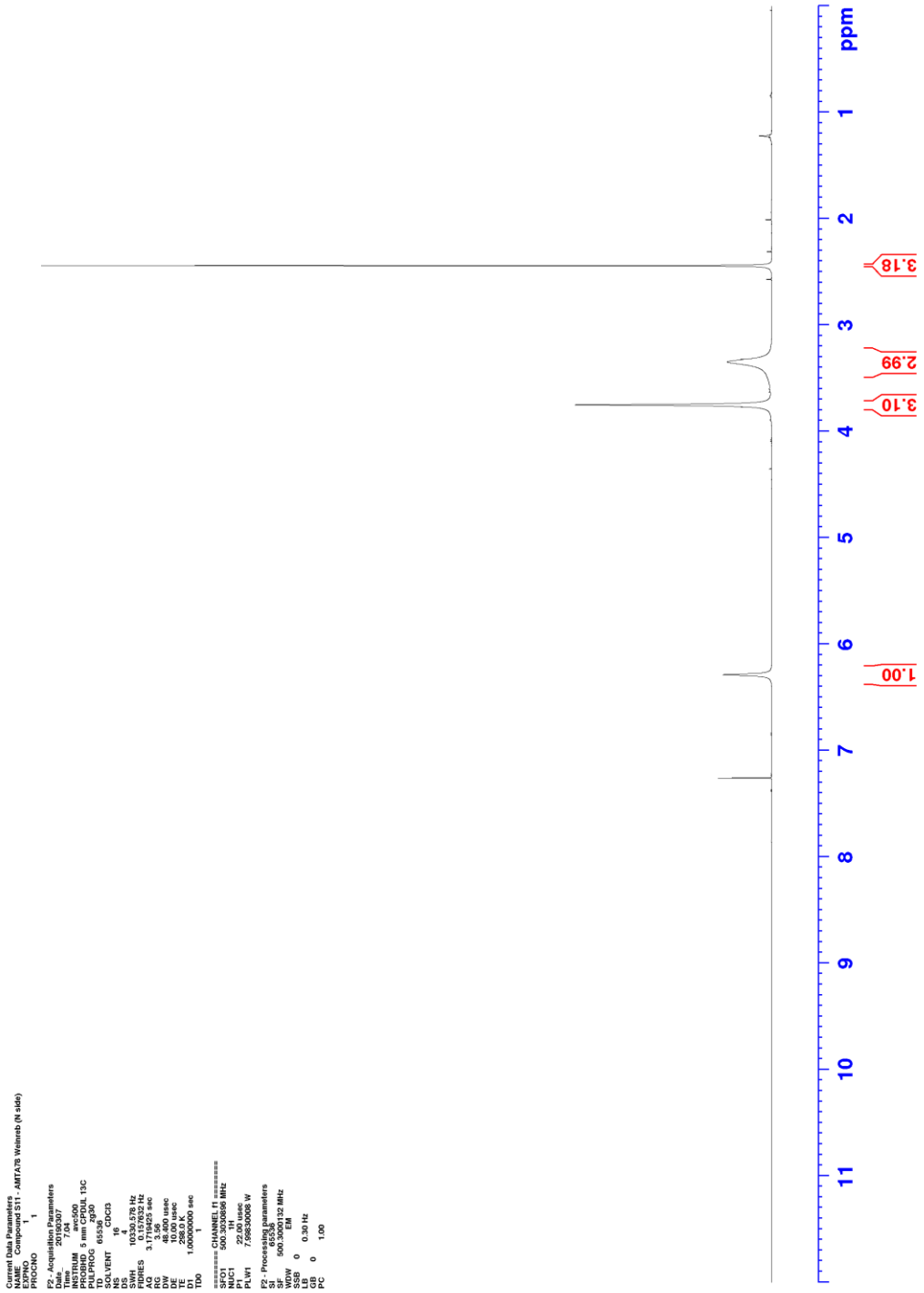
1ii-(3-Methylisoxazol-5-yl)prop-2ii-yn-1ii-one



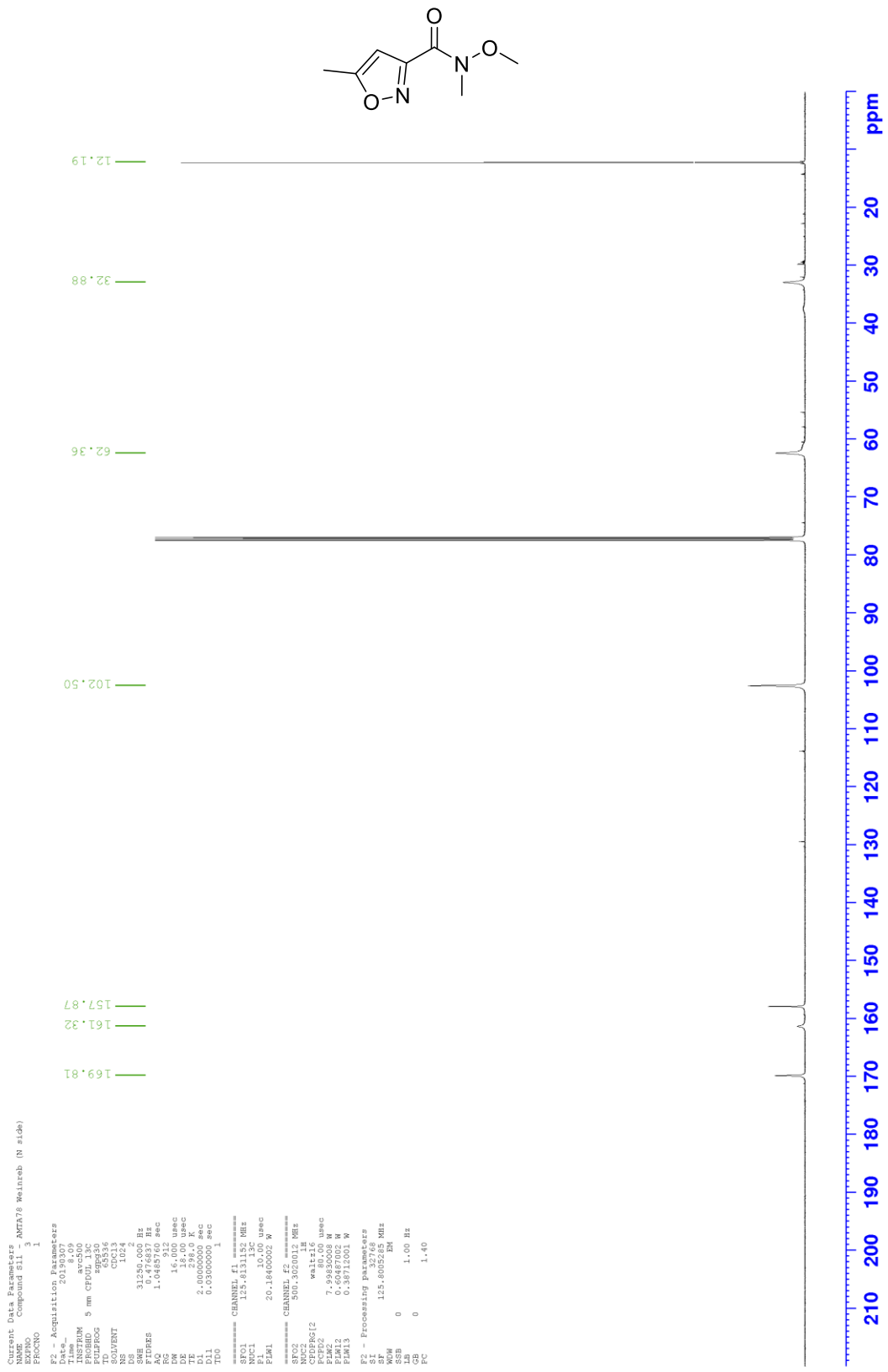


N-Methoxy-*N*-5-dimethylisoxazole-3-carboxamide

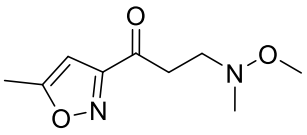




_N_-Methoxy-_N_-5-dimethylisoxazole-3-carboxamide



3^{i}-[Methoxy(methyl)amino]-1^{i}-(5-methylisoxazol-3-yl)propan-1^{i}-one

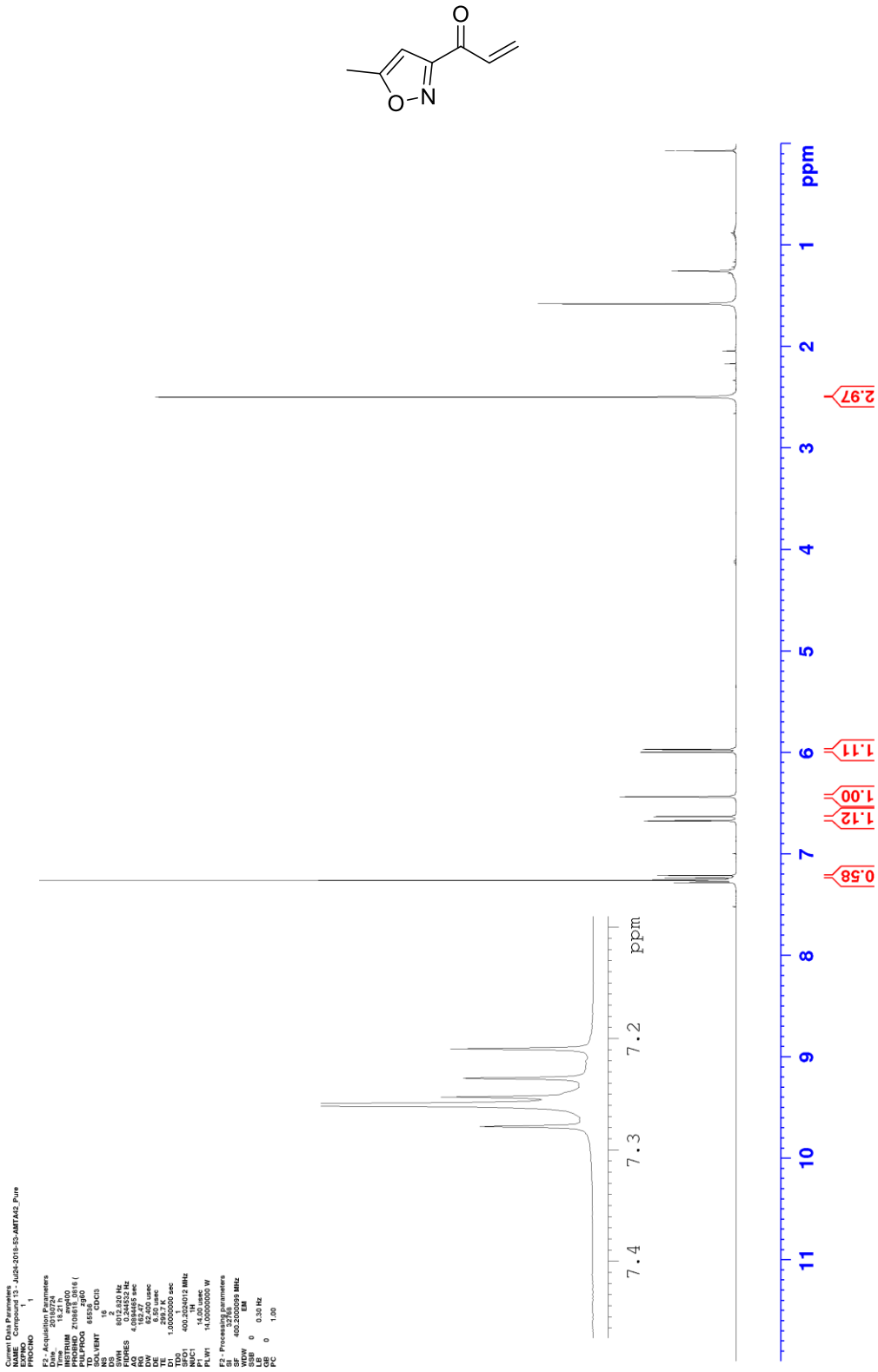




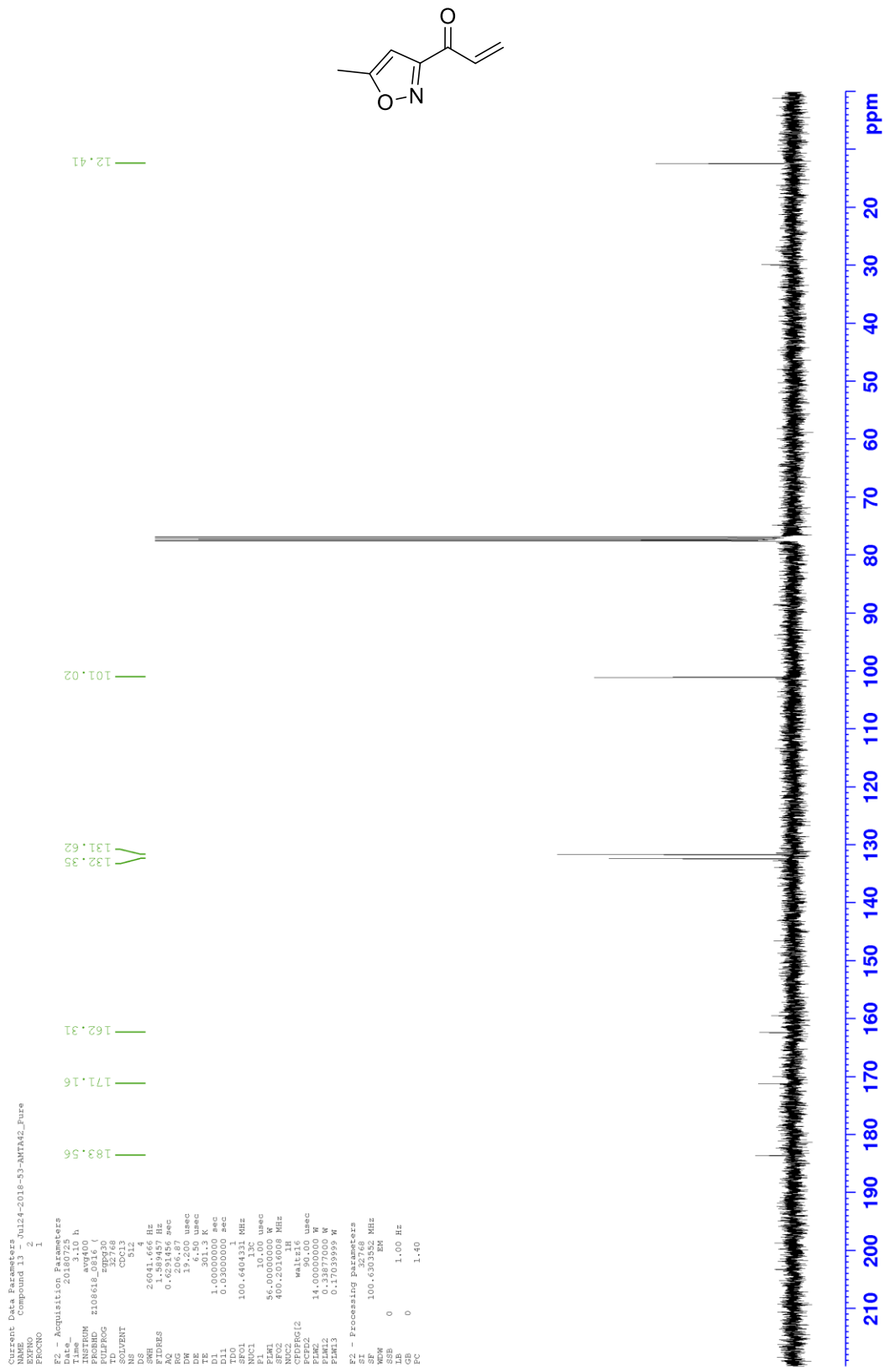
3[i]-[Methoxy(methyl)amino]-1[i]-(5-methylisoxazol-3-yl)propan-1[i]-one



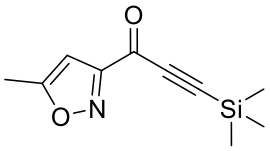
1^{i}-(5-Methylisoxazol-3-yl)prop-2^{i}-en-1^{i}-one



1^{i}-(5-Methylisoxazol-3-yl)prop-2^{i}-en-1^{i}-one

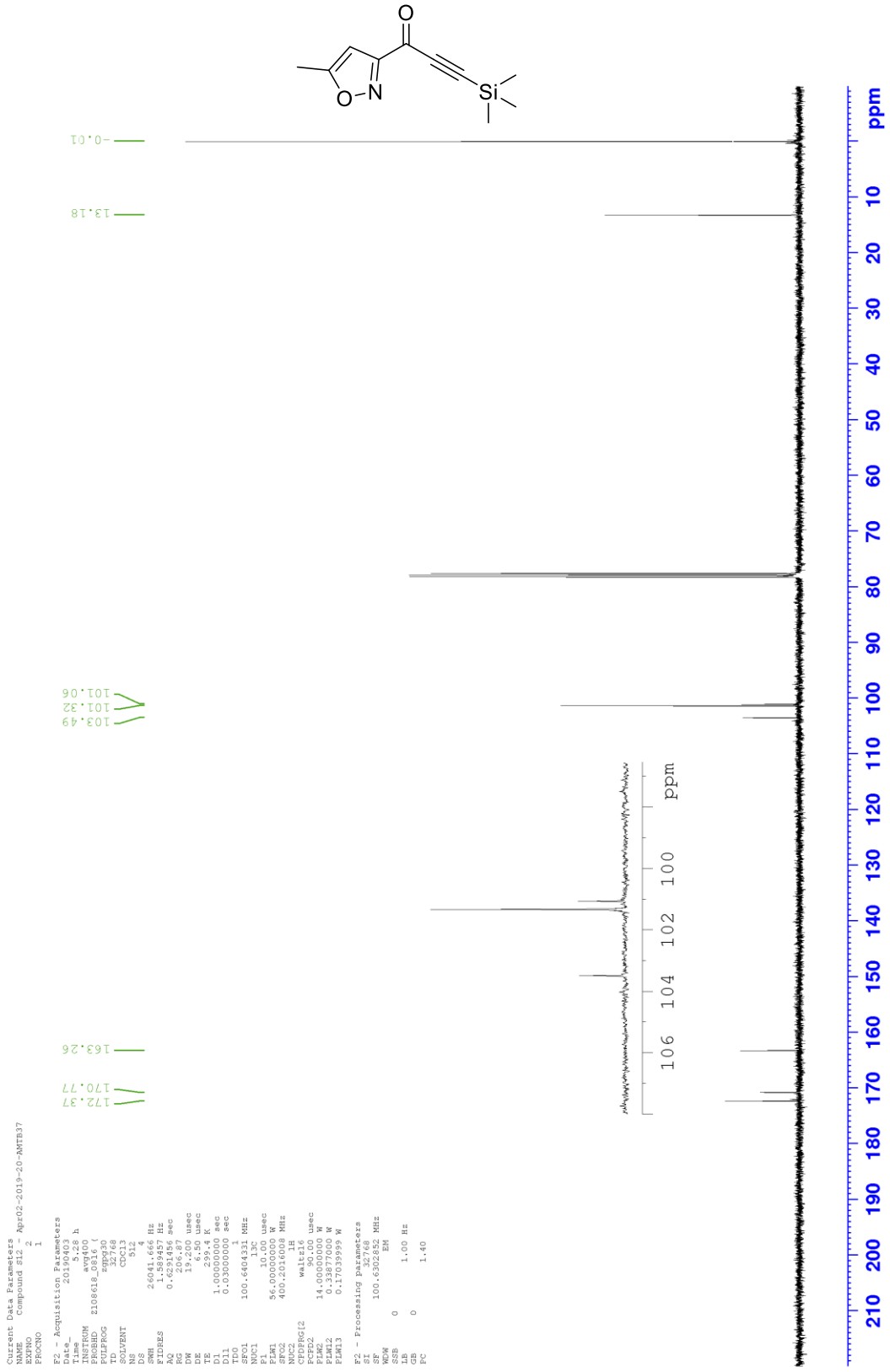


1^{i}-(5-Methylisoxazol-3-yl)-3^{i}-(trimethylsilyl)prop-2^{i}-yn-1^{i}-one

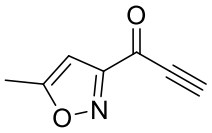


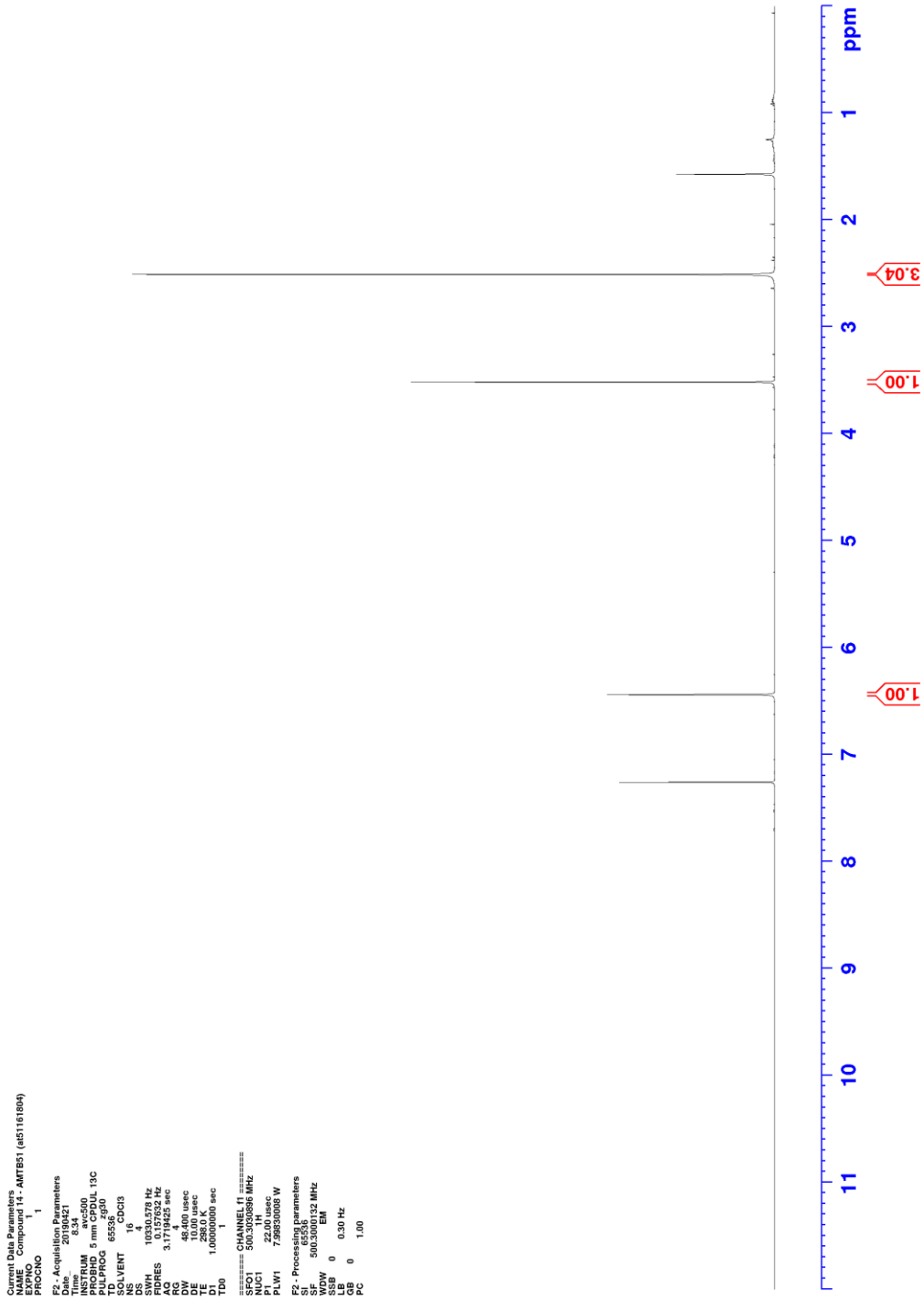


1^{i}-(5-Methylisoxazol-3-yl)-3^{i}-(trimethylsilyl)prop-2^{i}-yn-1^{i}-one

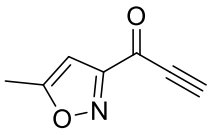


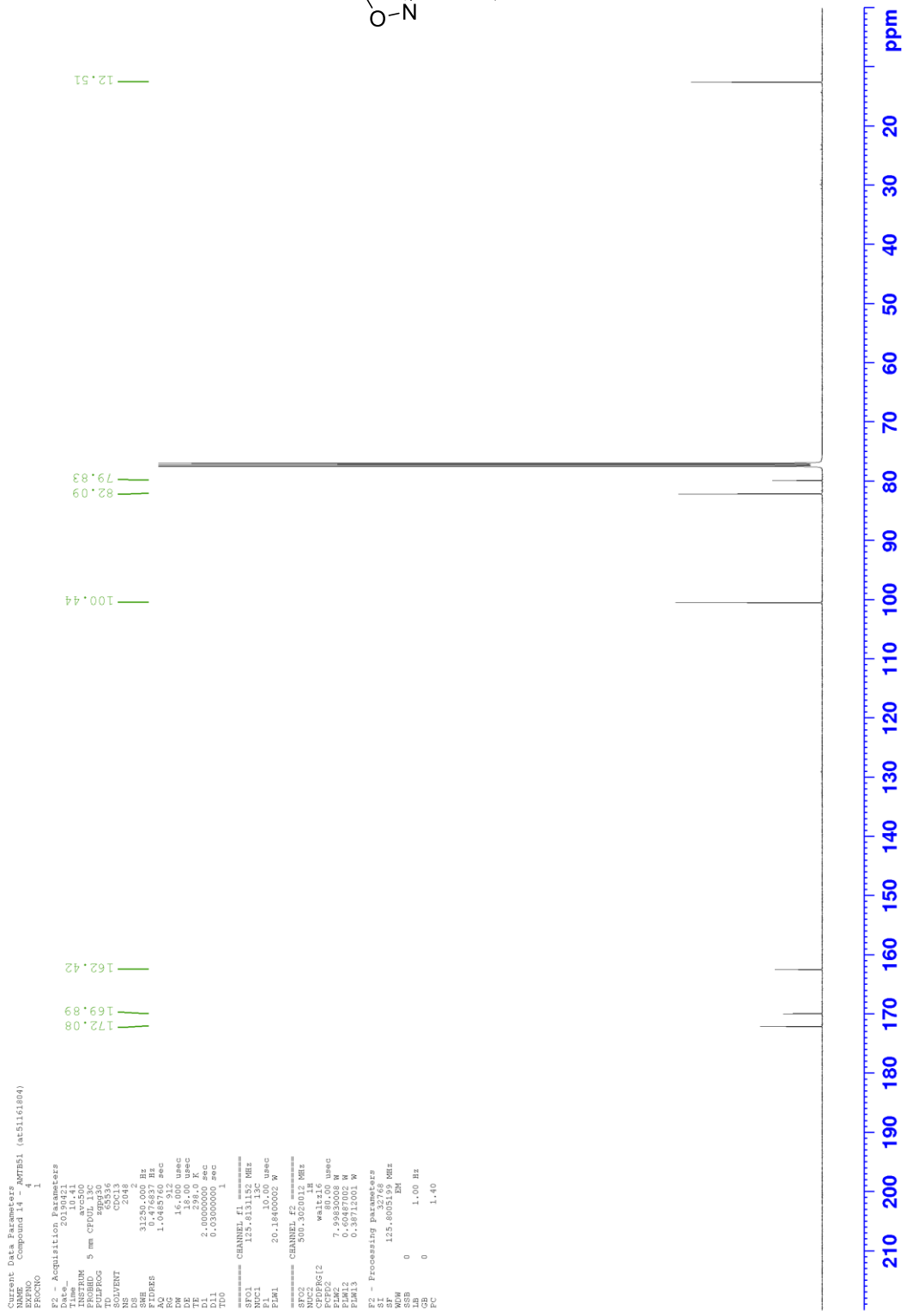
1^{i}-(5-Methylisoxazol-3-yl)prop-2^{i}-yn-1^{i}-one



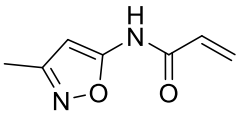


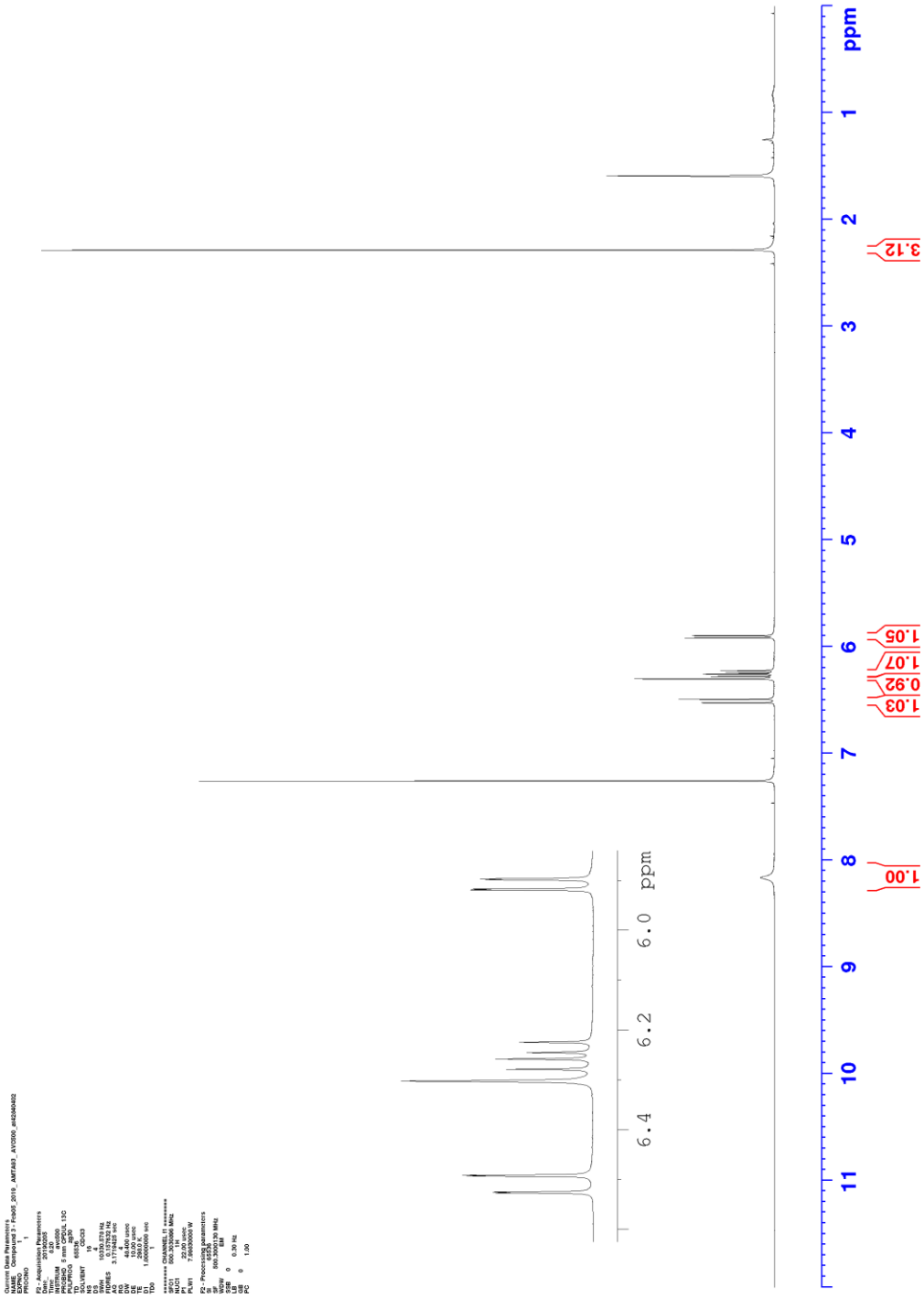
1^{i}-(5-Methylisoxazol-3-yl)prop-2^{i}-yn-1^{i}-one



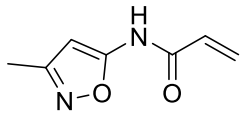


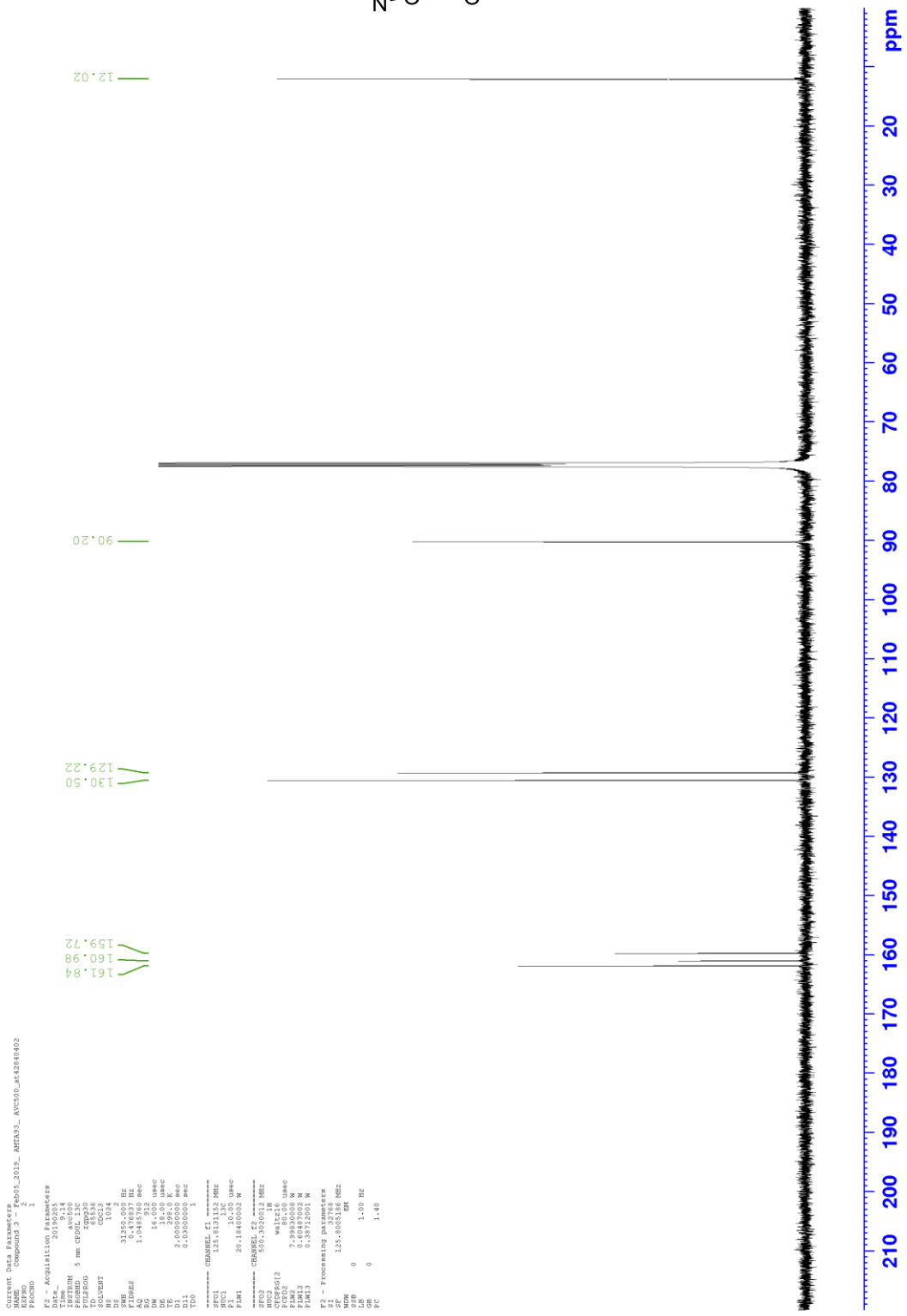
_N_-(3-Methylisoxazol-5-yl)prop-2^{ii}-enamide



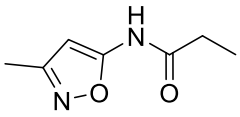


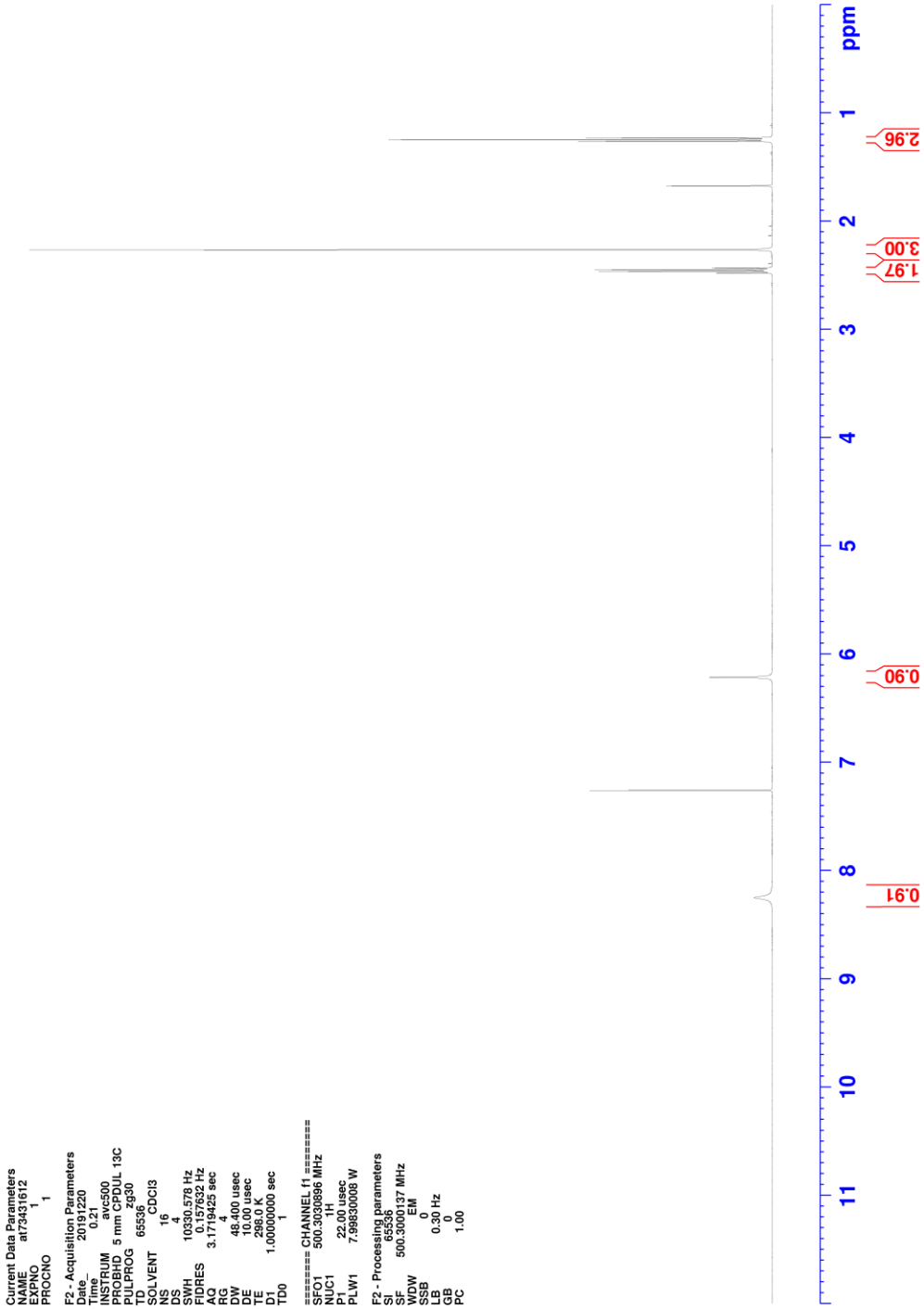
N-(3-Methylisoxazol-5-yl)prop-2[ii]-enamide





_N_-(3-Methyl-isoxazol-5-yl)propanamide

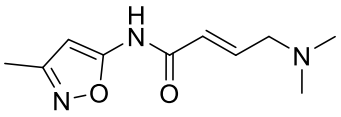


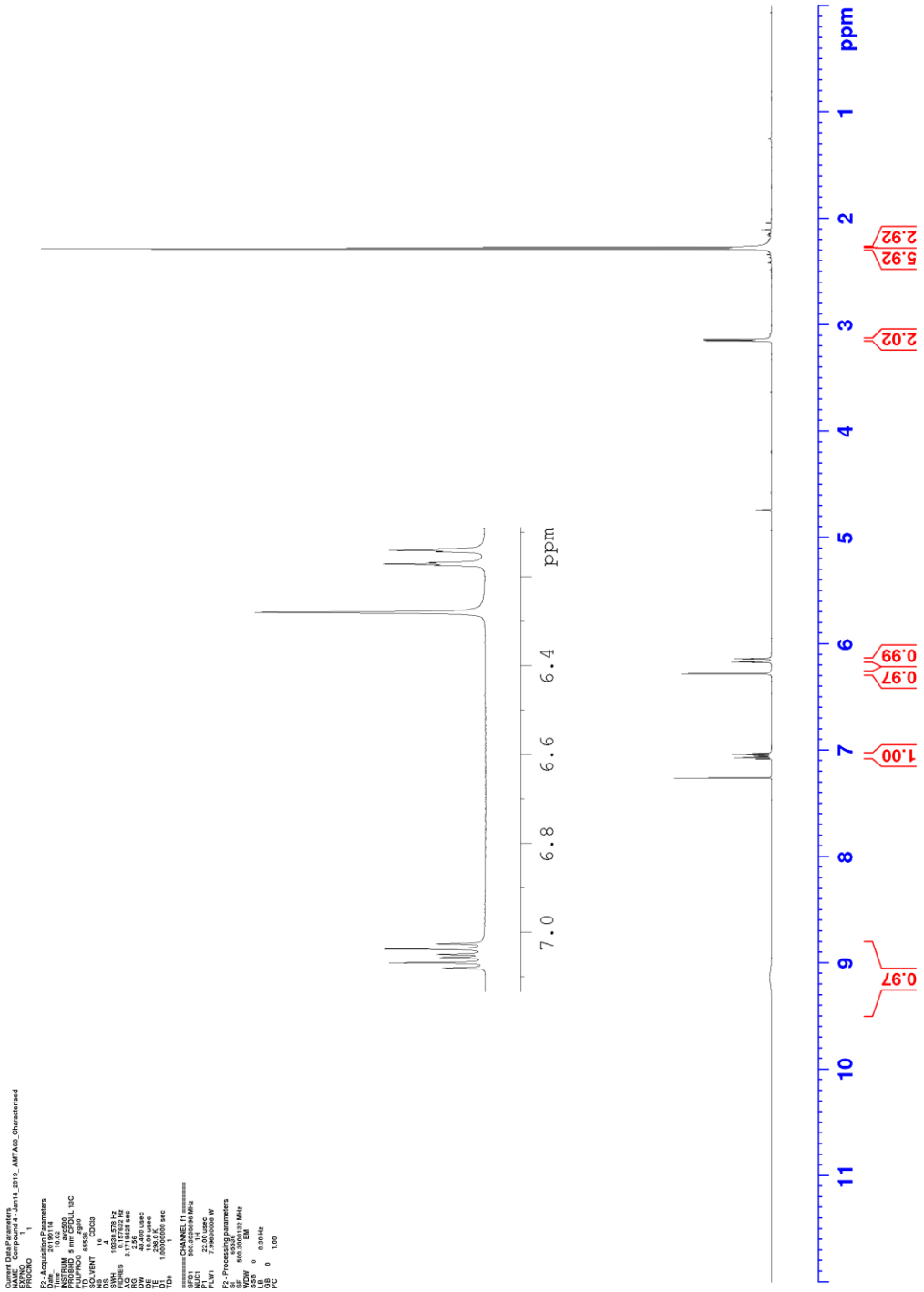


N-(3-Methyl-isoxazol-5-yl)propanamide

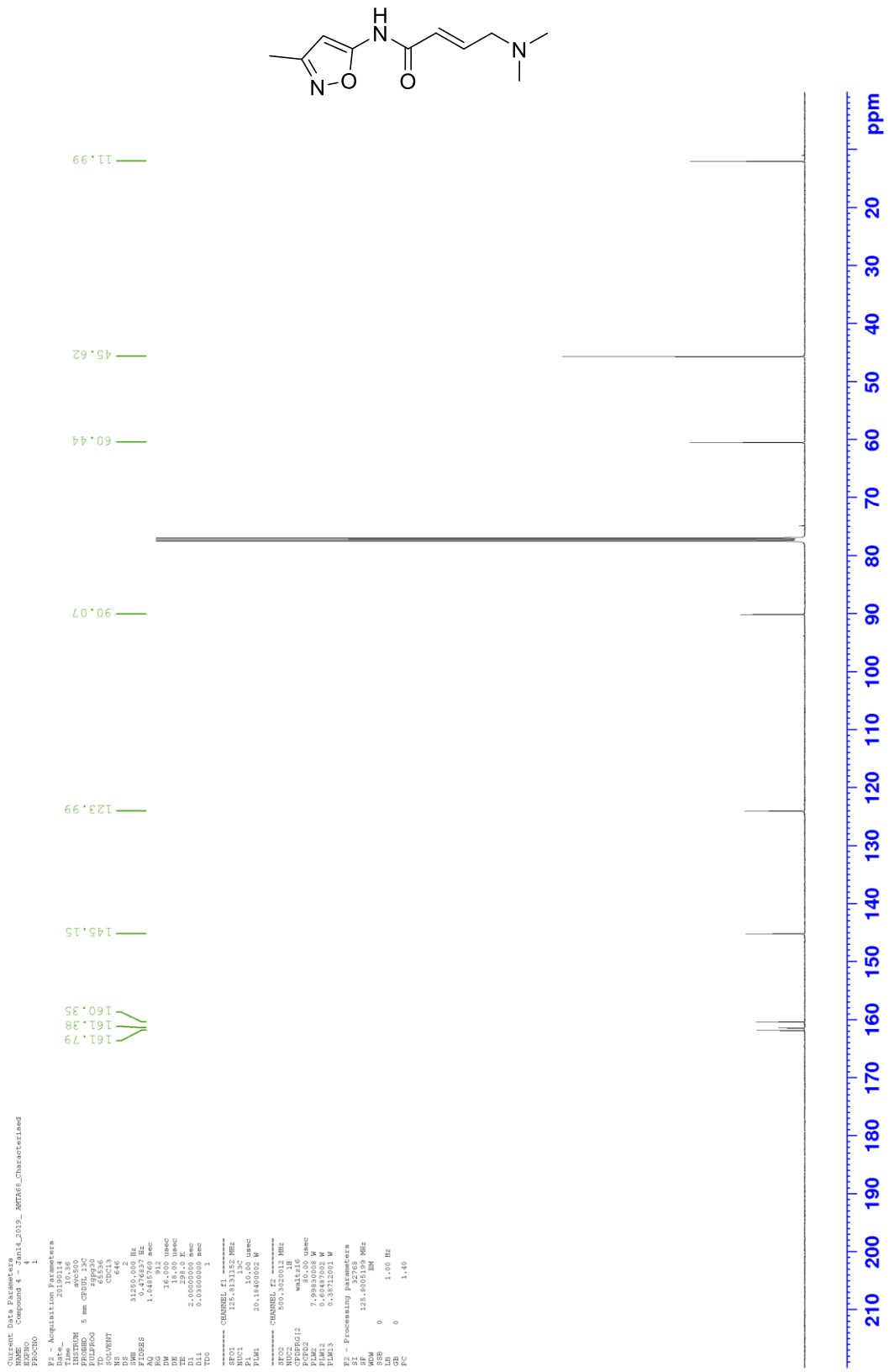


(*E*)-4[ii]-(Dimethylamino)-*N*-(3-methylisoxazol-5-yl)but-2[ii]-enamide

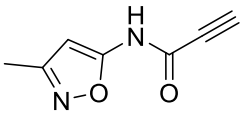


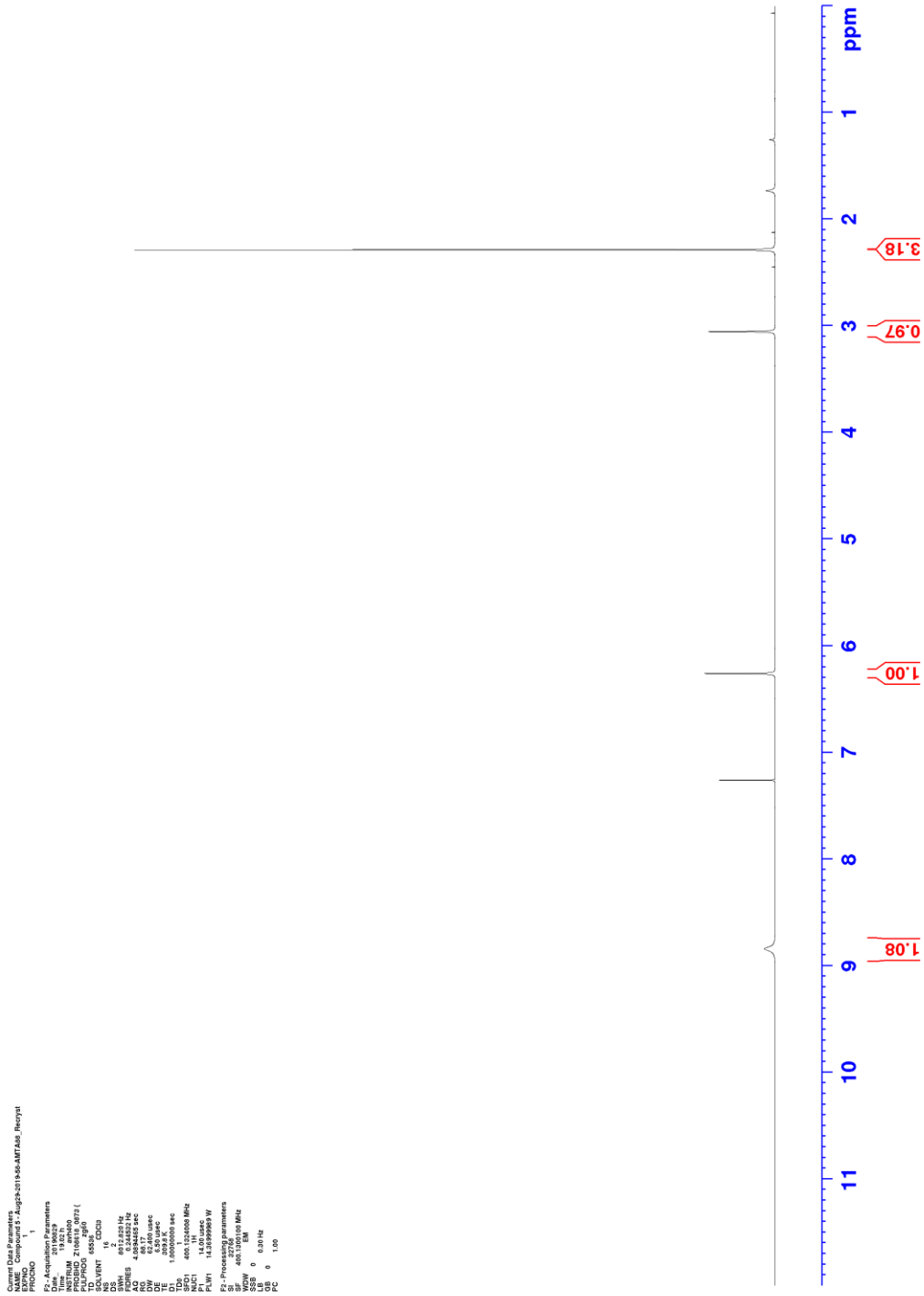


(*E*)-4[ii]-(Dimethylamino)-*N*-(3-methylisoxazol-5-yl)but-2[ii]-enamide

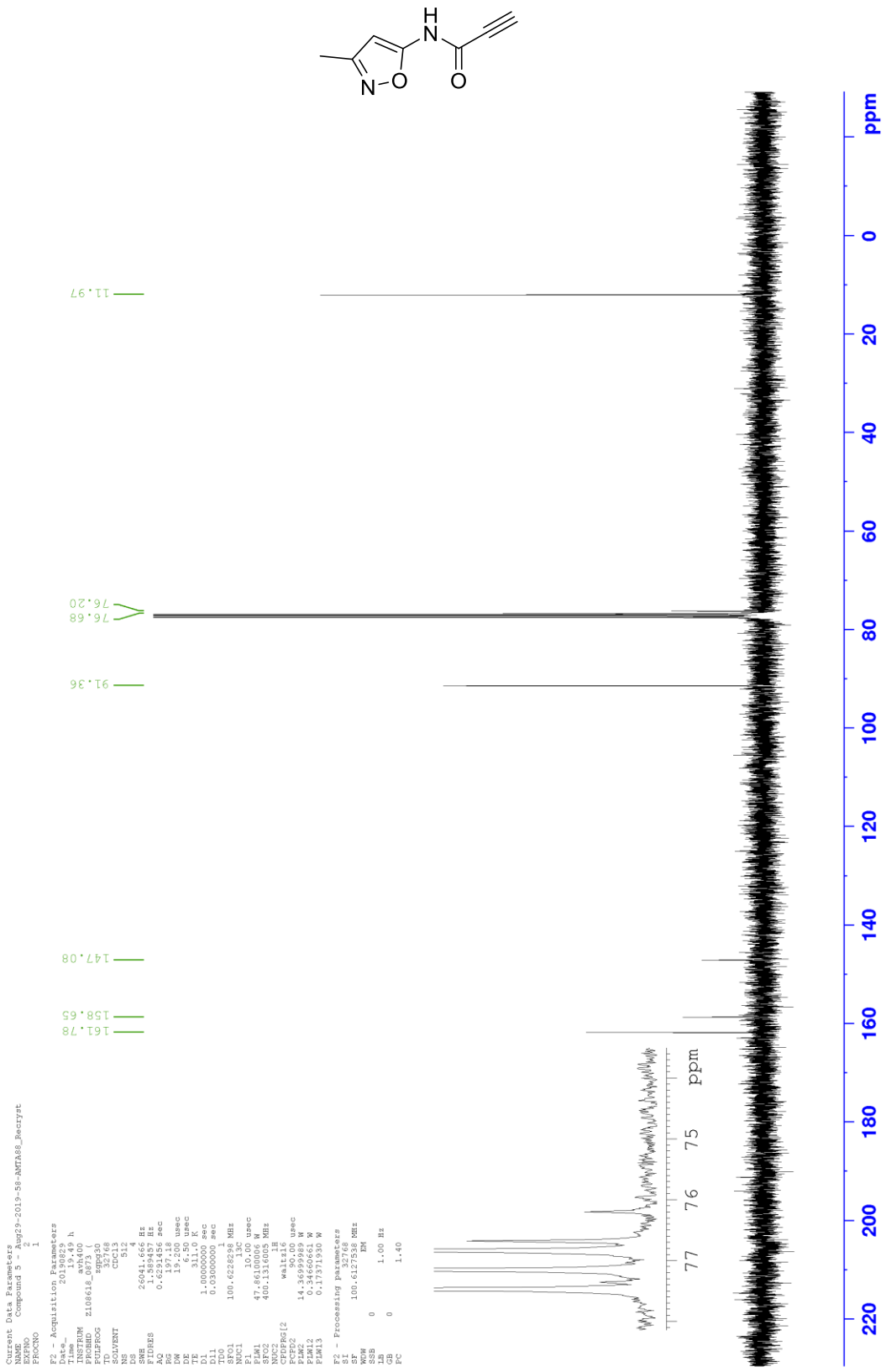


_N_-(3-Methylisoxazol-5-yl)prop-2[ii]-ynamide

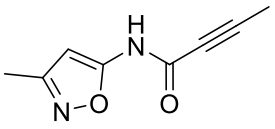


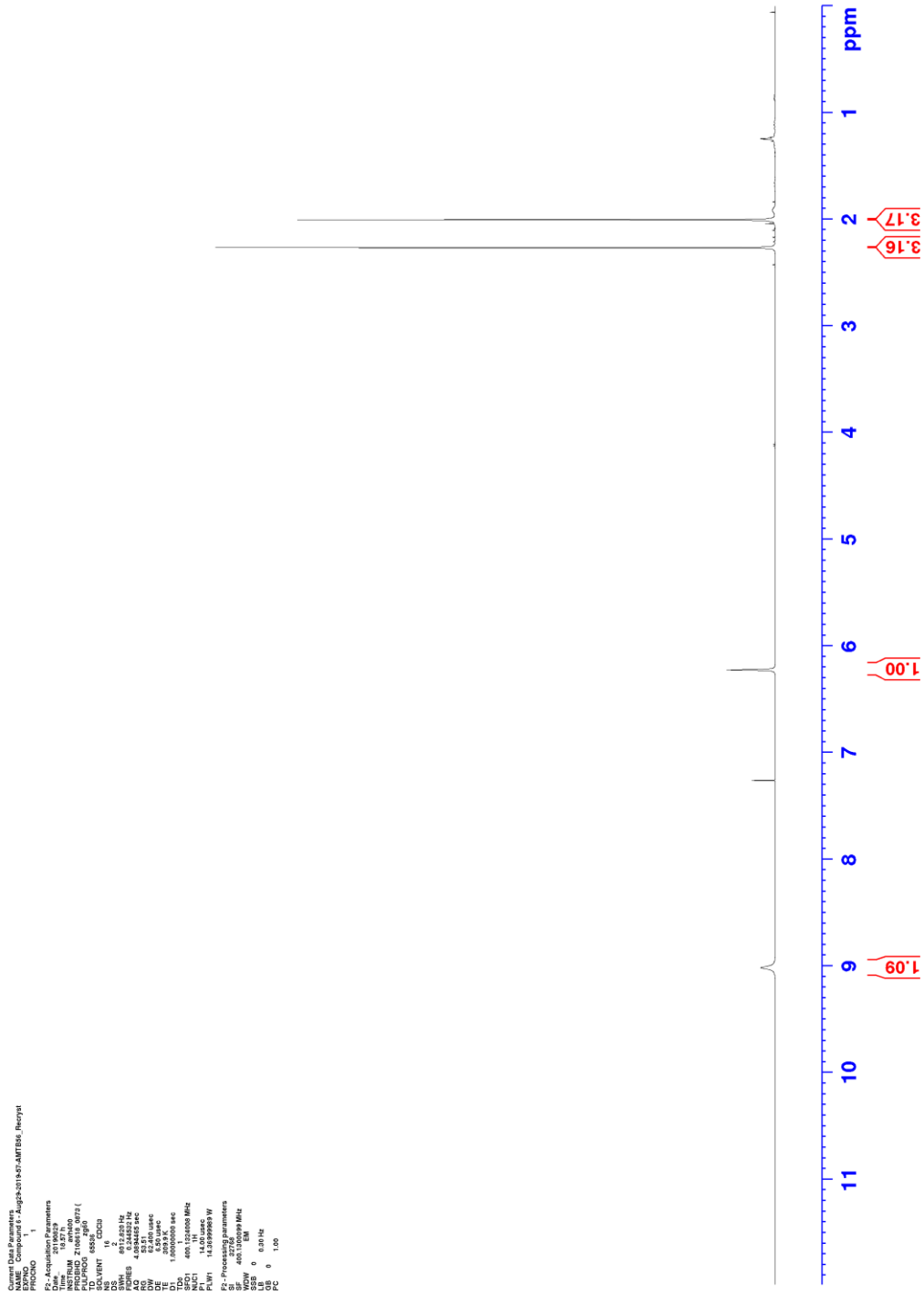


N-(3-Methylisoxazol-5-yl)prop-2[ii]-ynamide

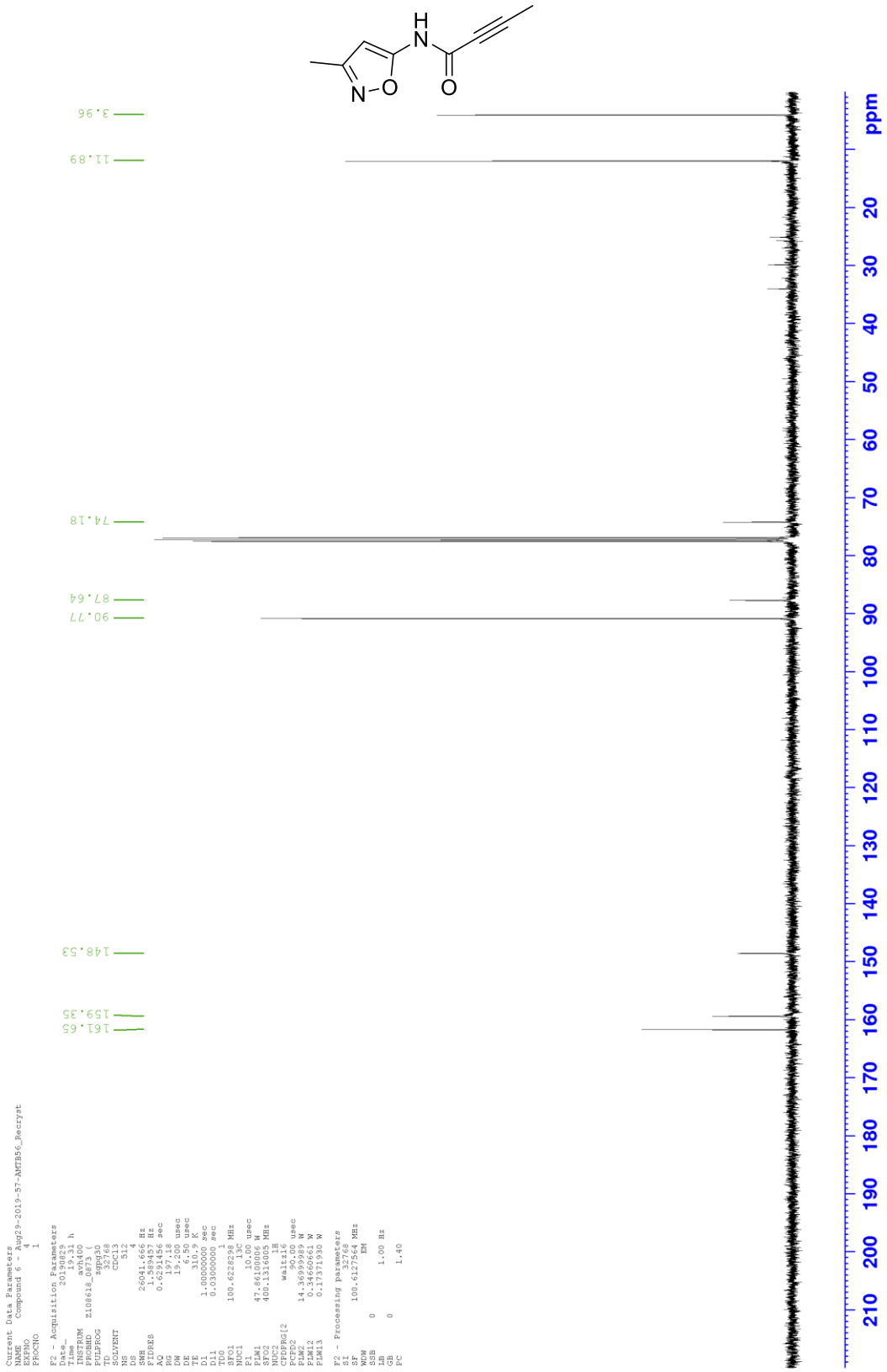


N-(3-Methylisoxazol-5-yl)but-2[ii]-ynamide

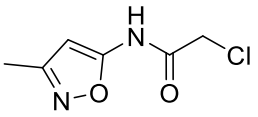


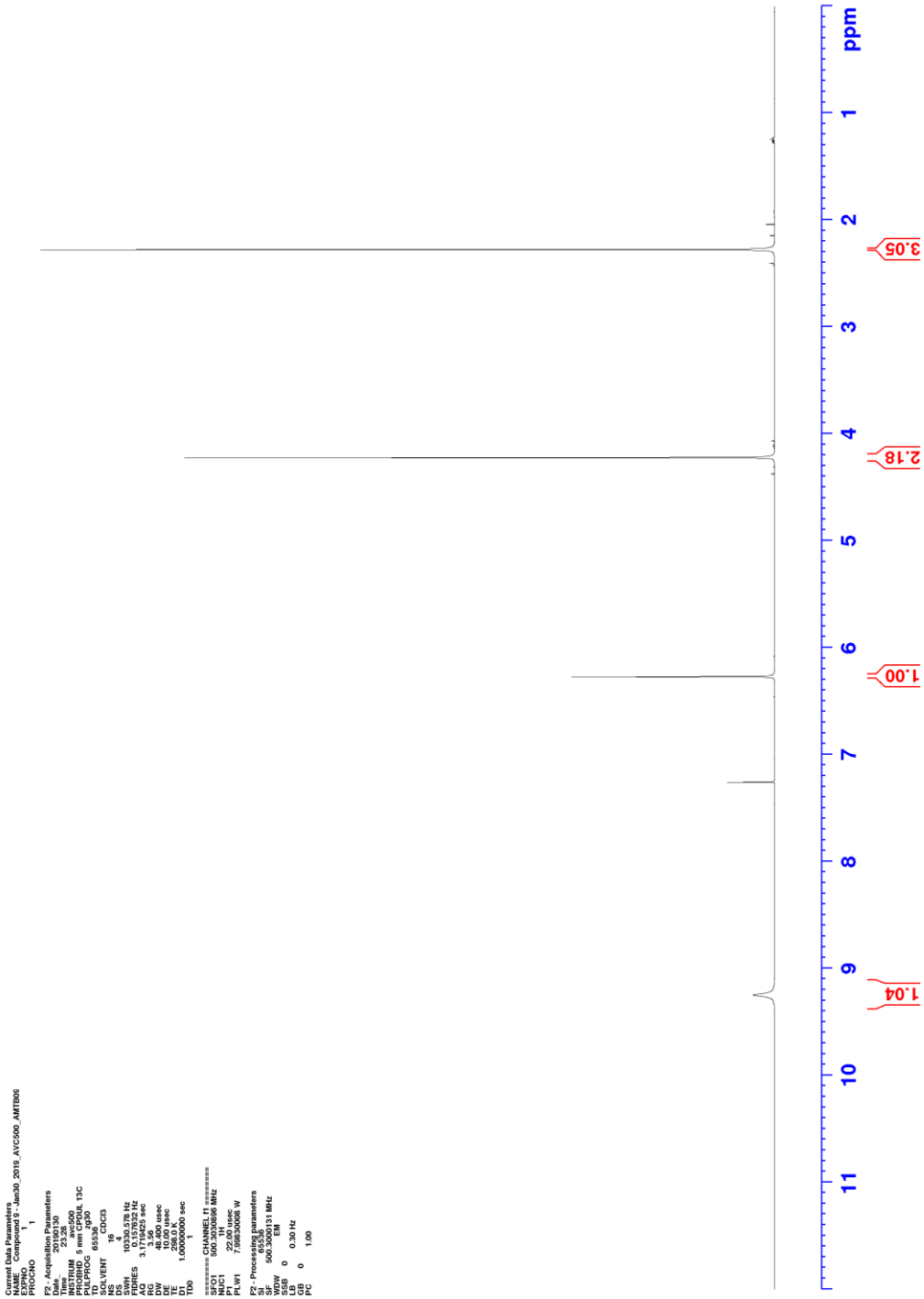


_N_-(3-Methylisoxazol-5-yl)but-2[ii]-ynamide

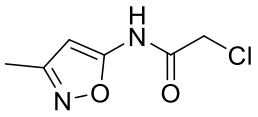


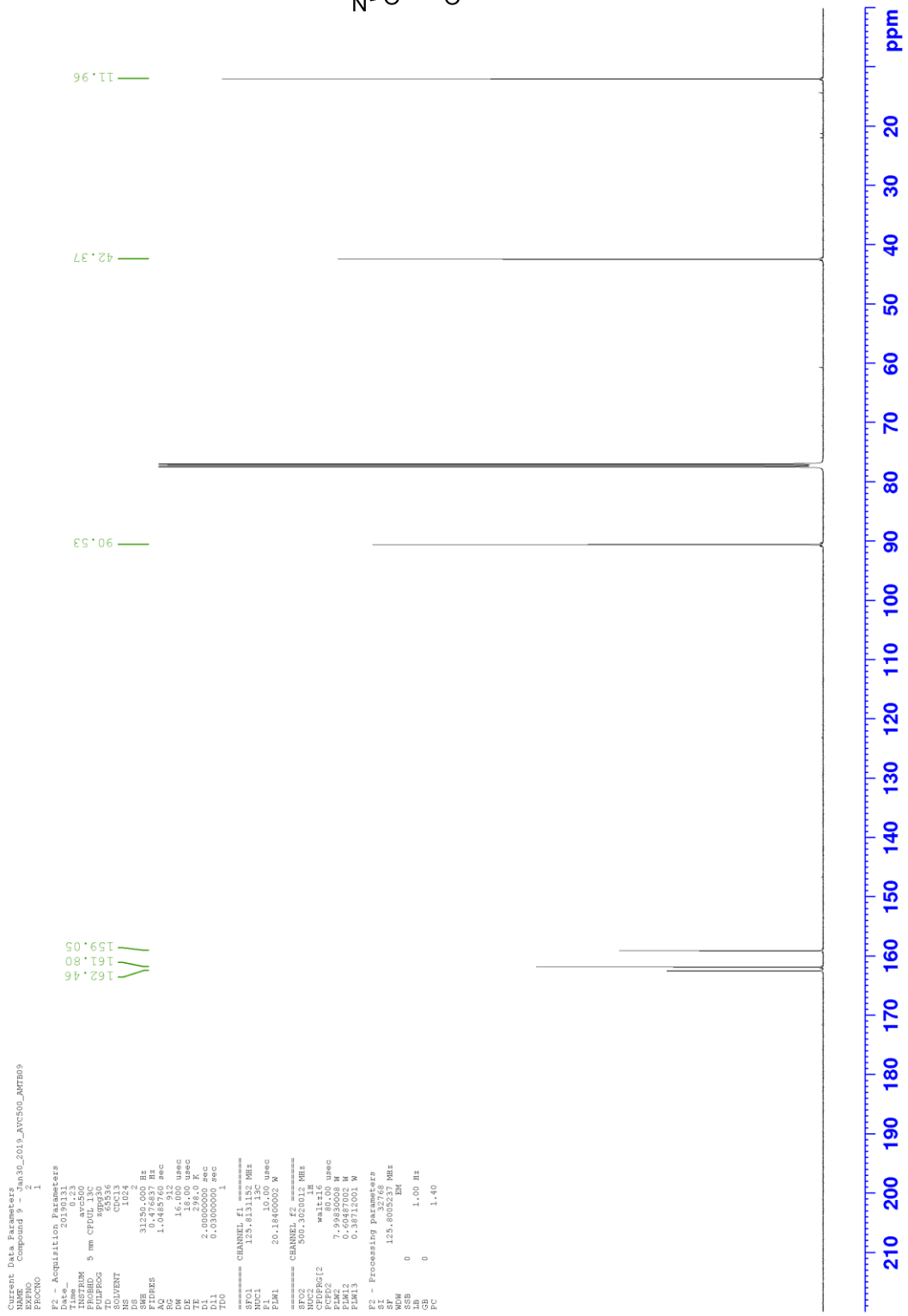
2[ii]-Chloro-*N*-(3-methylisoxazol-5-yl)acetamide



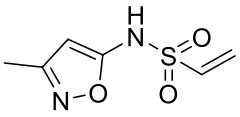


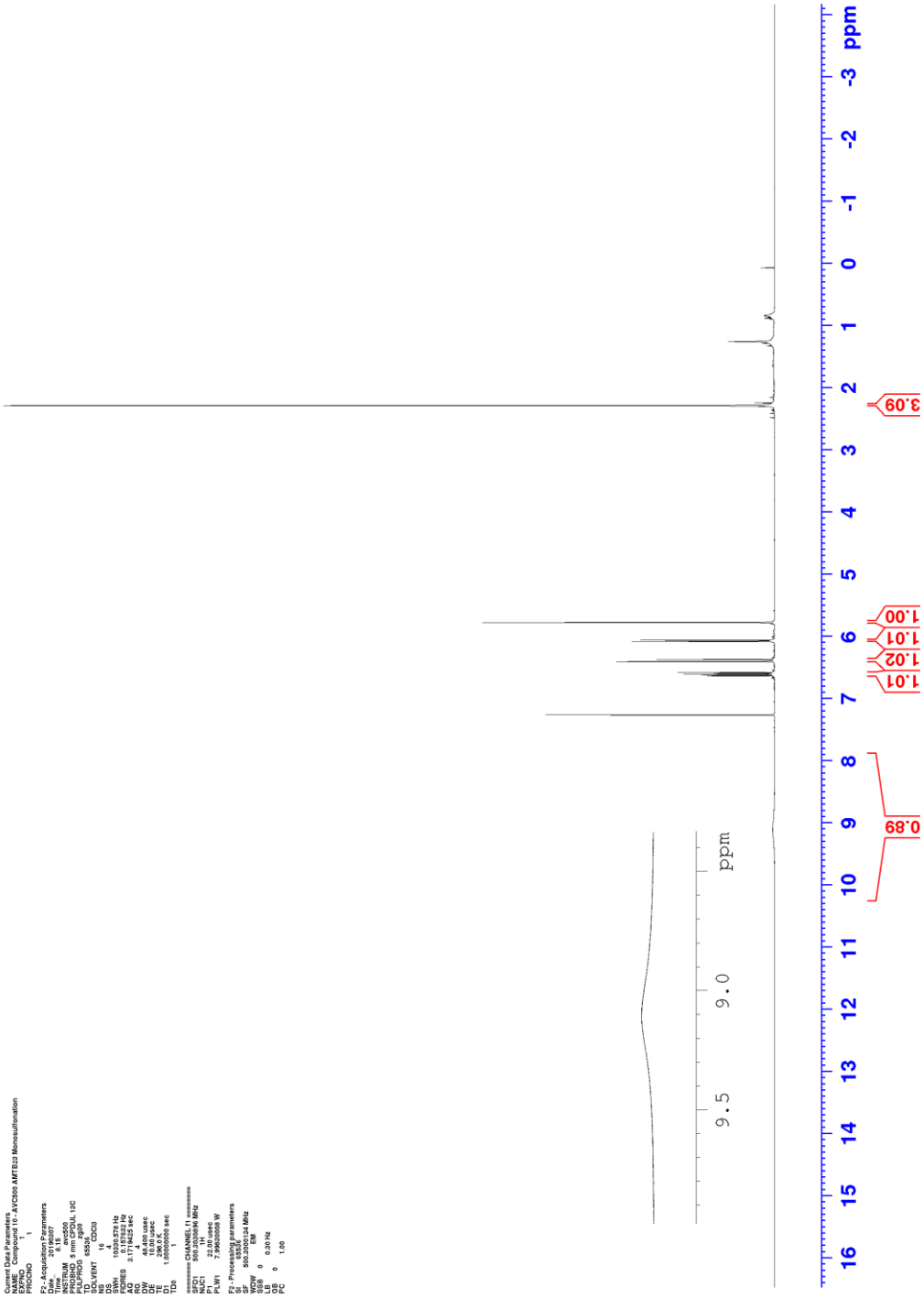
2[ii]-Chloro-*N*-(3-methylisoxazol-5-yl)acetamide



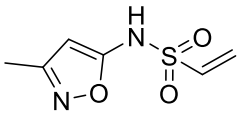


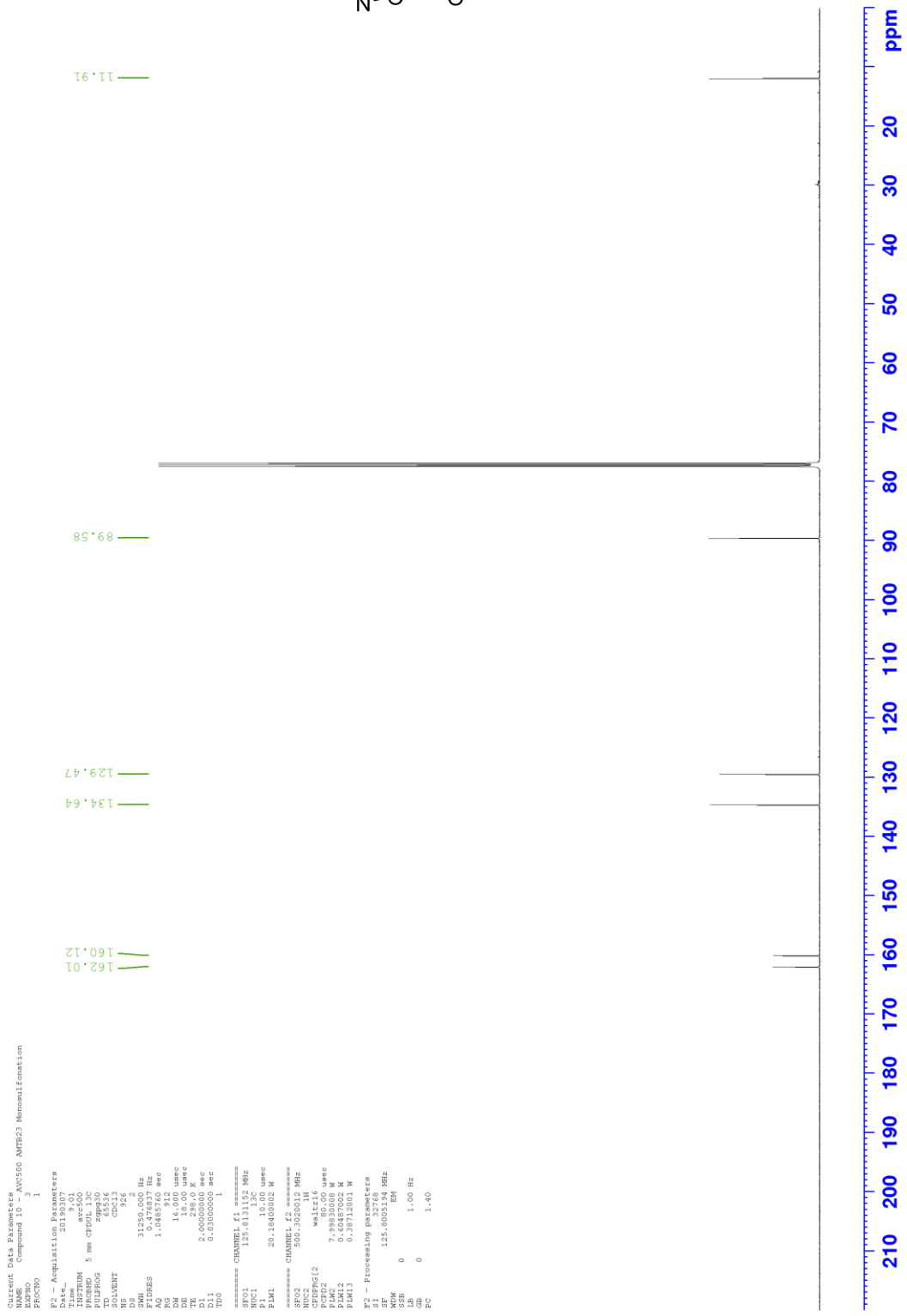
_N_-(3-Methylisoxazol-5-yl)ethenesulfonamide



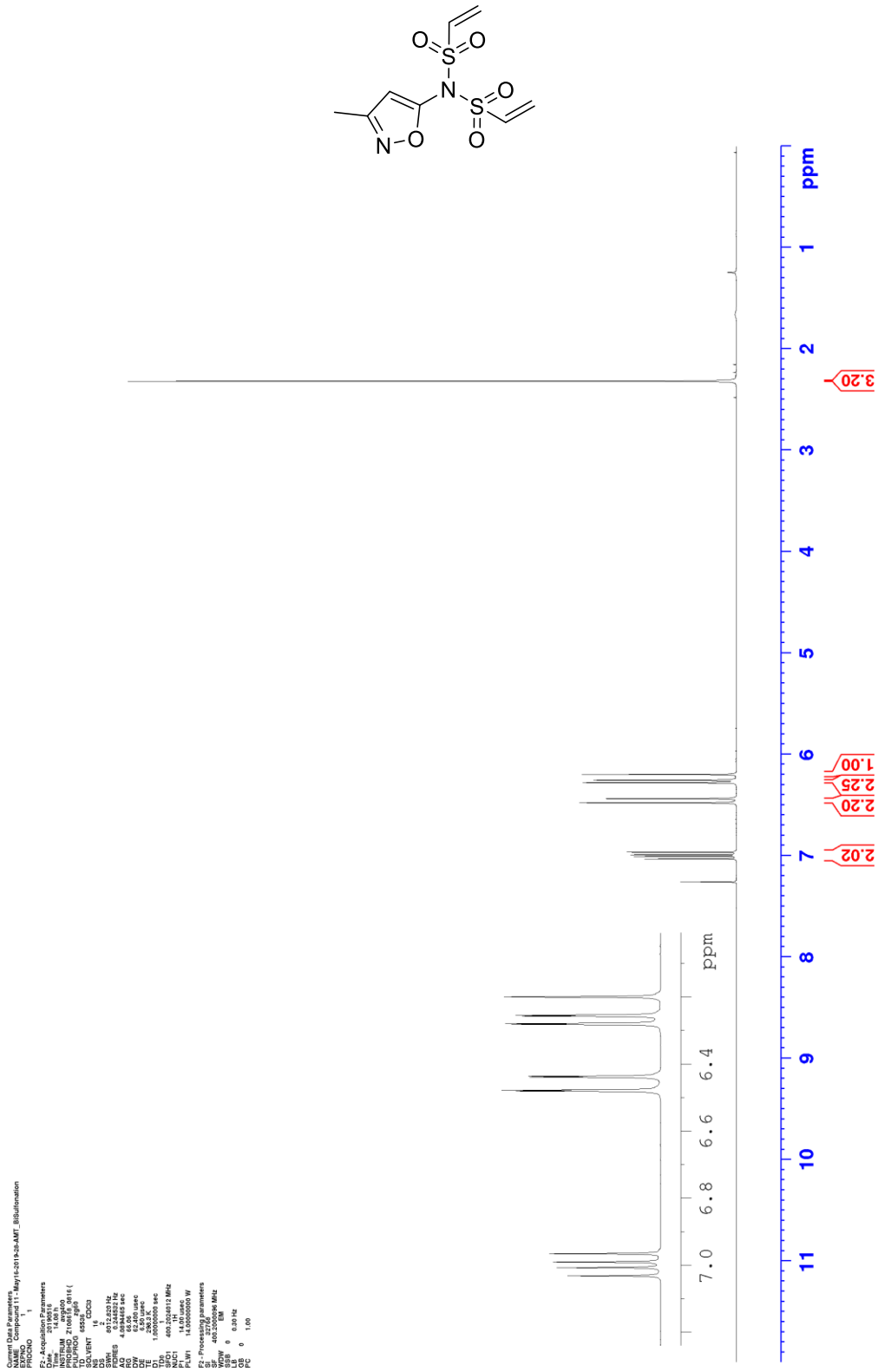


_N_-(3-Methylisoxazol-5-yl)ethenesulfonamide

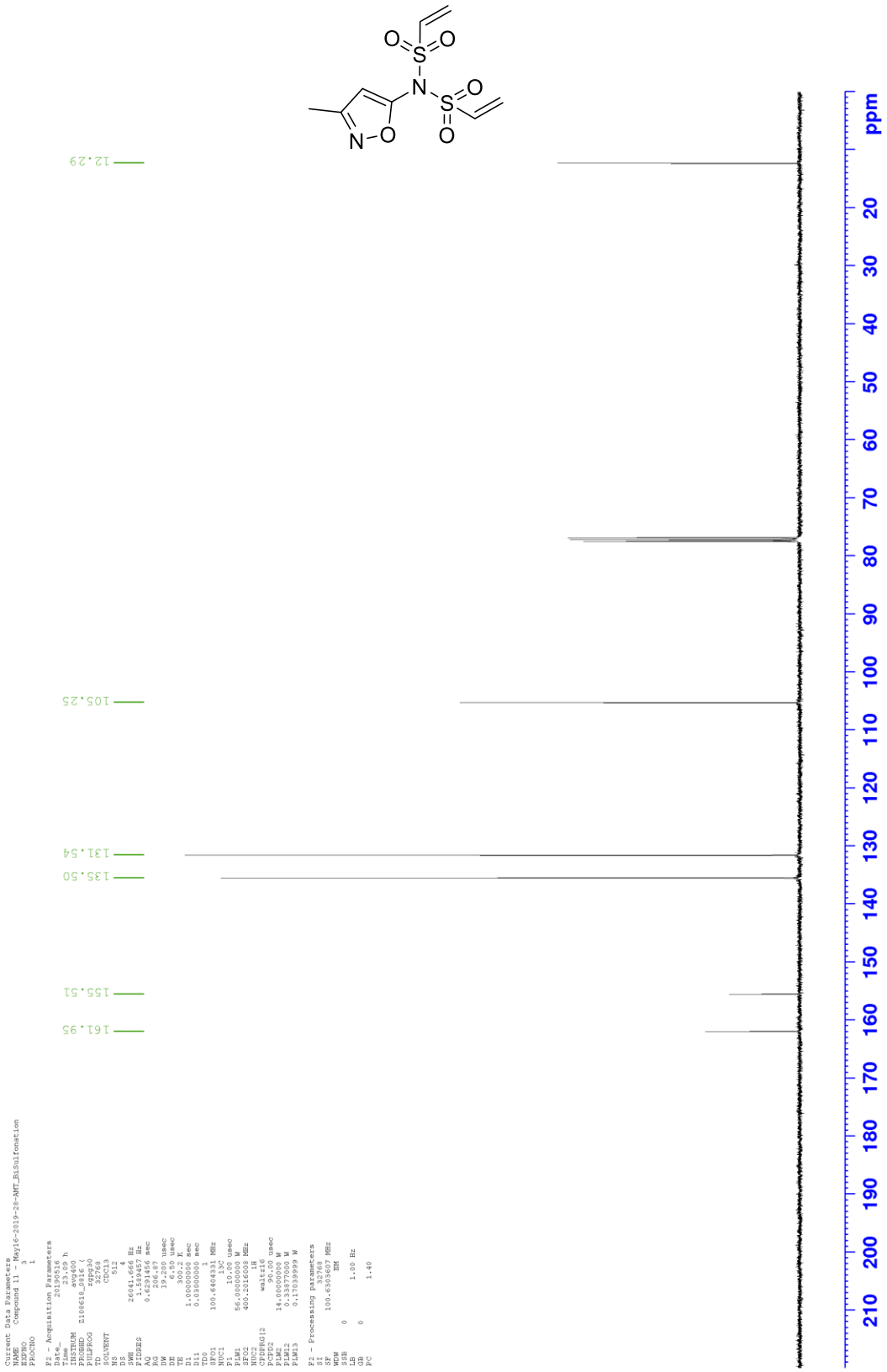




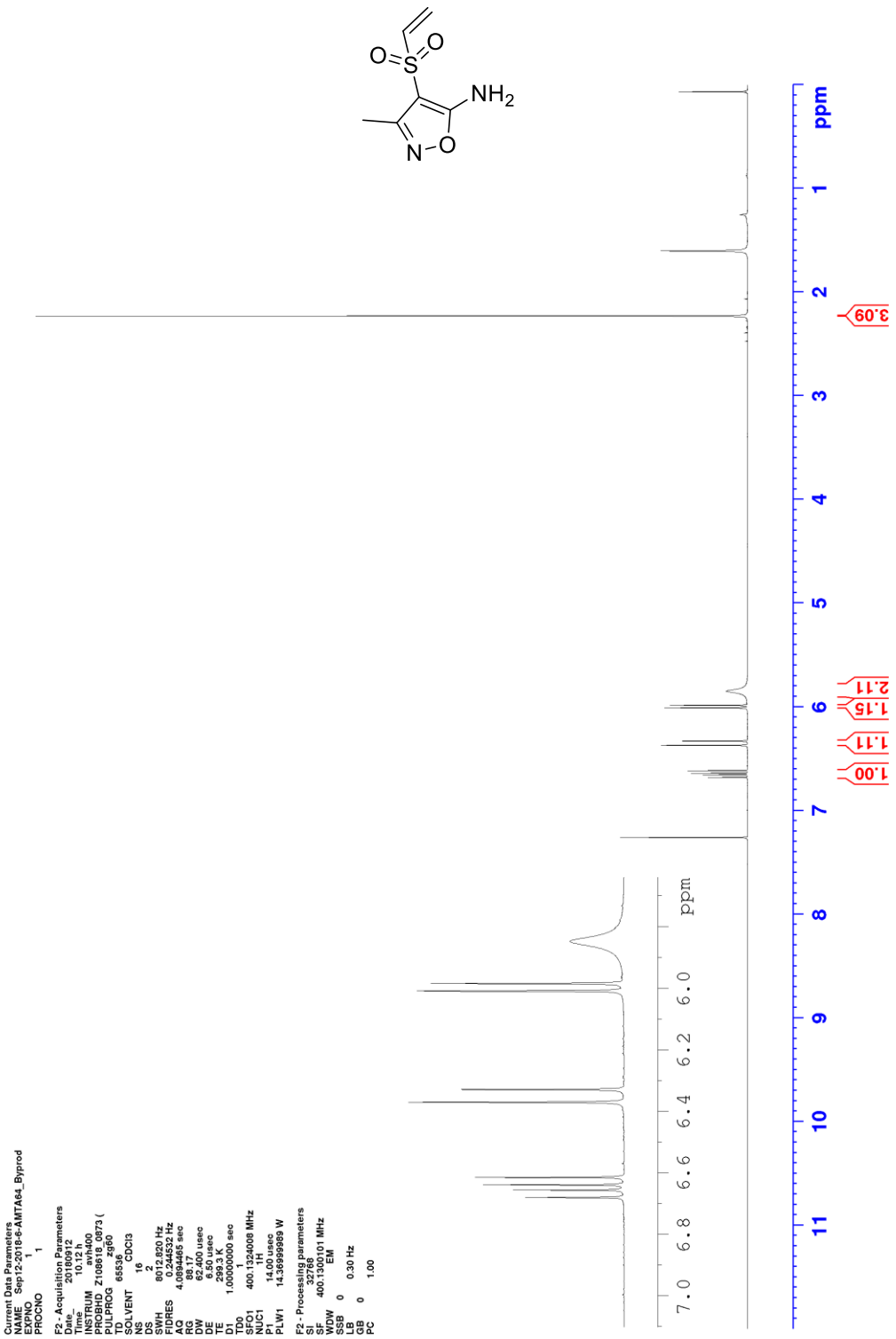
_N_-(Ethenesulfonyl)-_N_-(3-methylisoxazol-5-yl)ethenesulfonamide



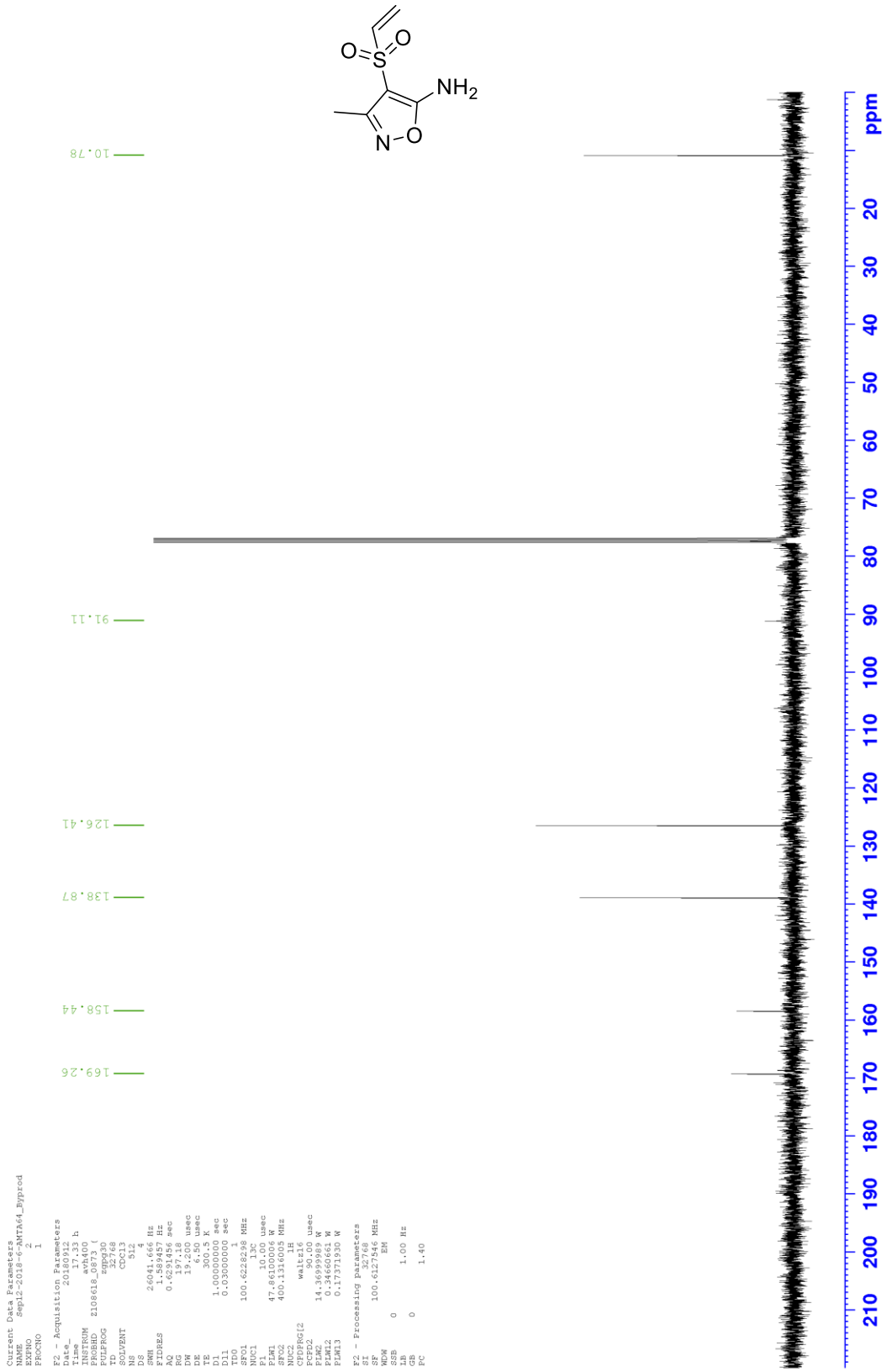
_N_-(Ethenesulfonyl)-_N_-(3-methylisoxazol-5-yl)ethenesulfonamide



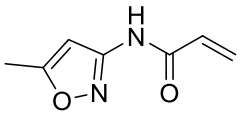
3-Methyl-4-(vinylsulfonyl)isoxazol-5-amine

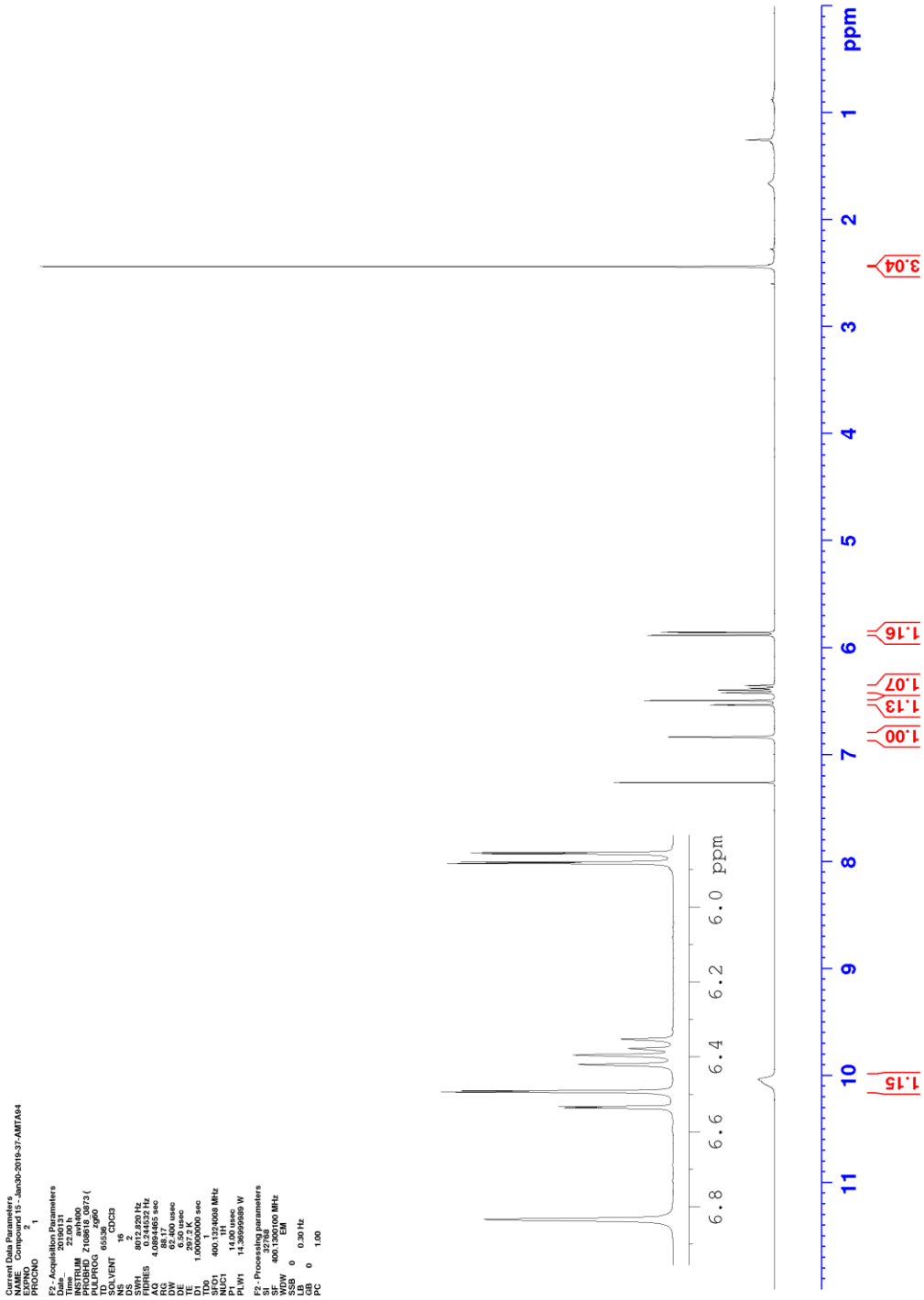


3-Methyl-4-(vinylsulfonyl)isoxazol-5-amine

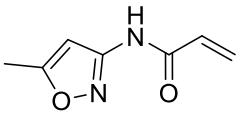


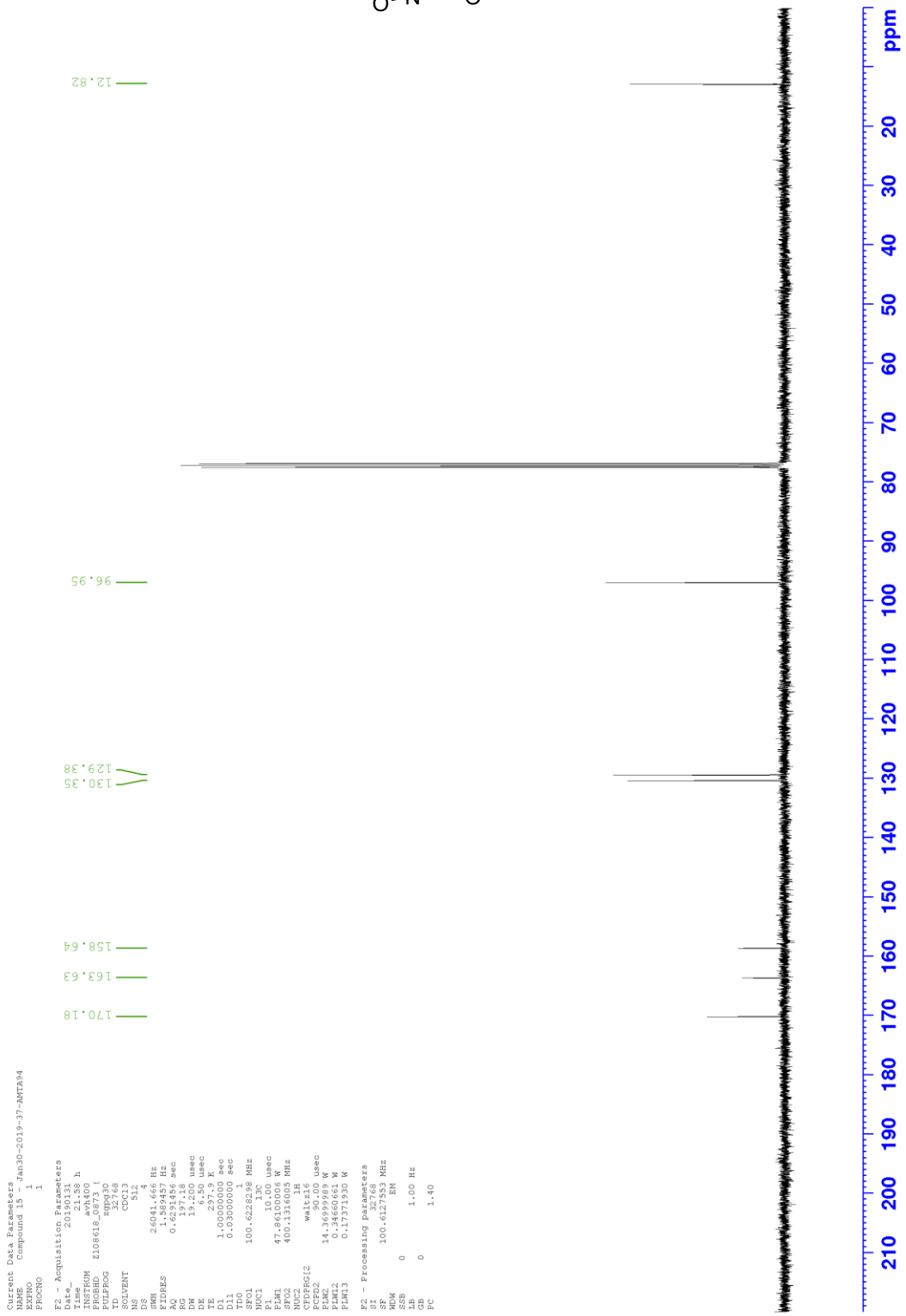
_N_-(5-Methylisoxazol-3-yl)prop-2^{ii}-enamide



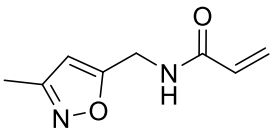


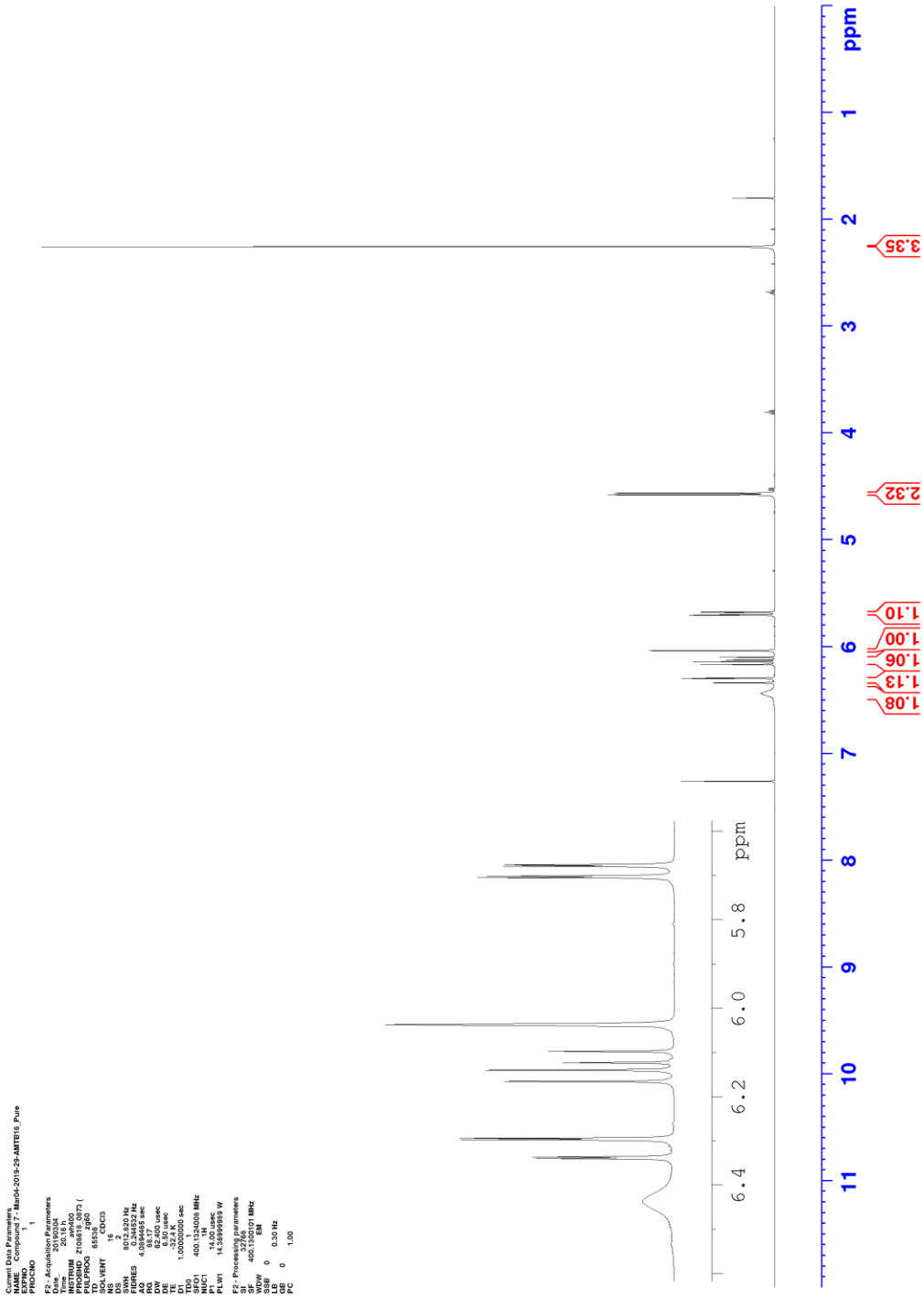
_N_-(5-Methylisoxazol-3-yl)prop-2^{ii}-enamide



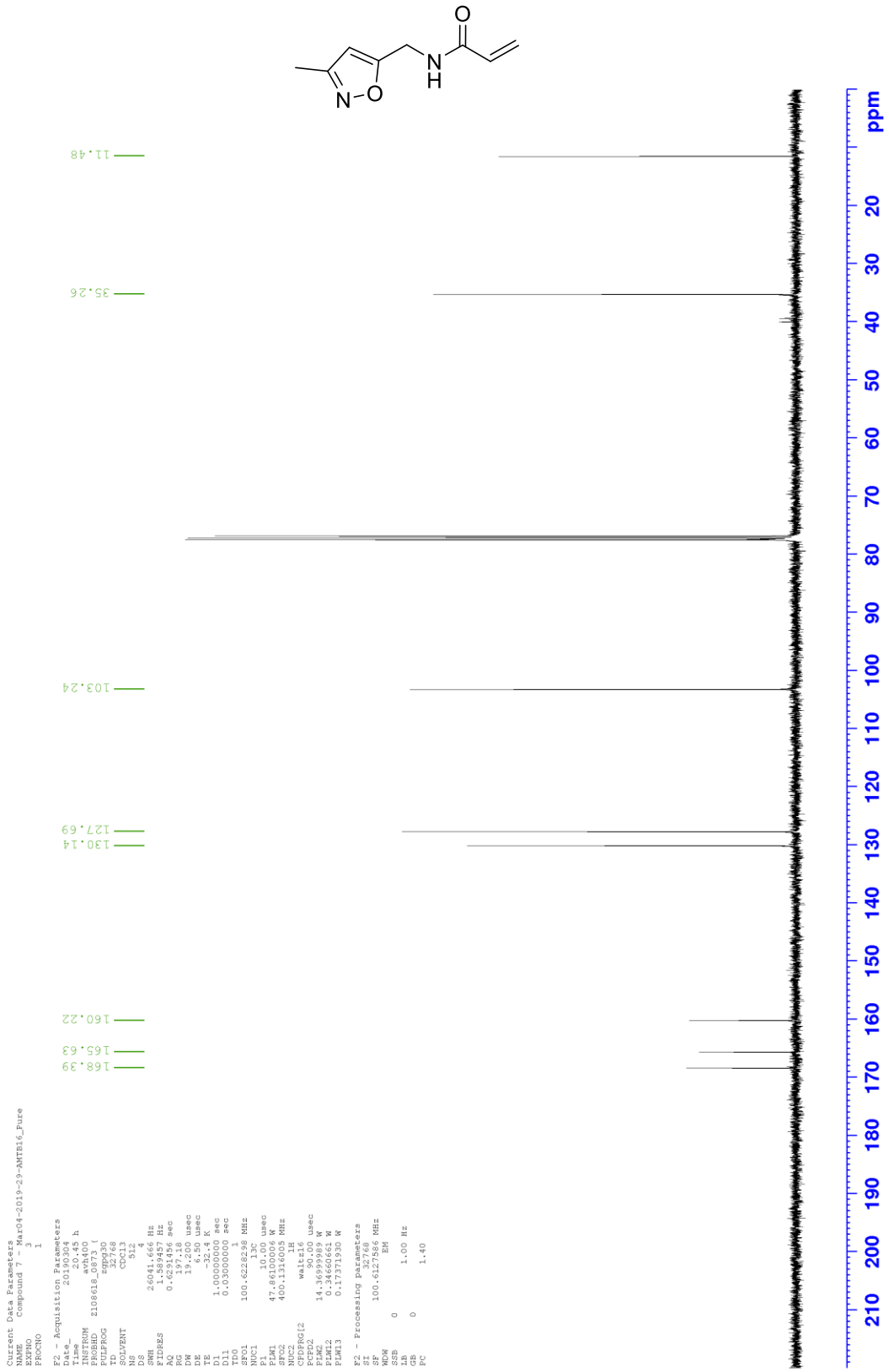


_N_-[(3-Methylisoxazol-5-yl)methyl]prop-2[ii]-enamide

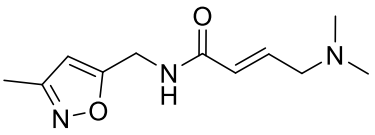


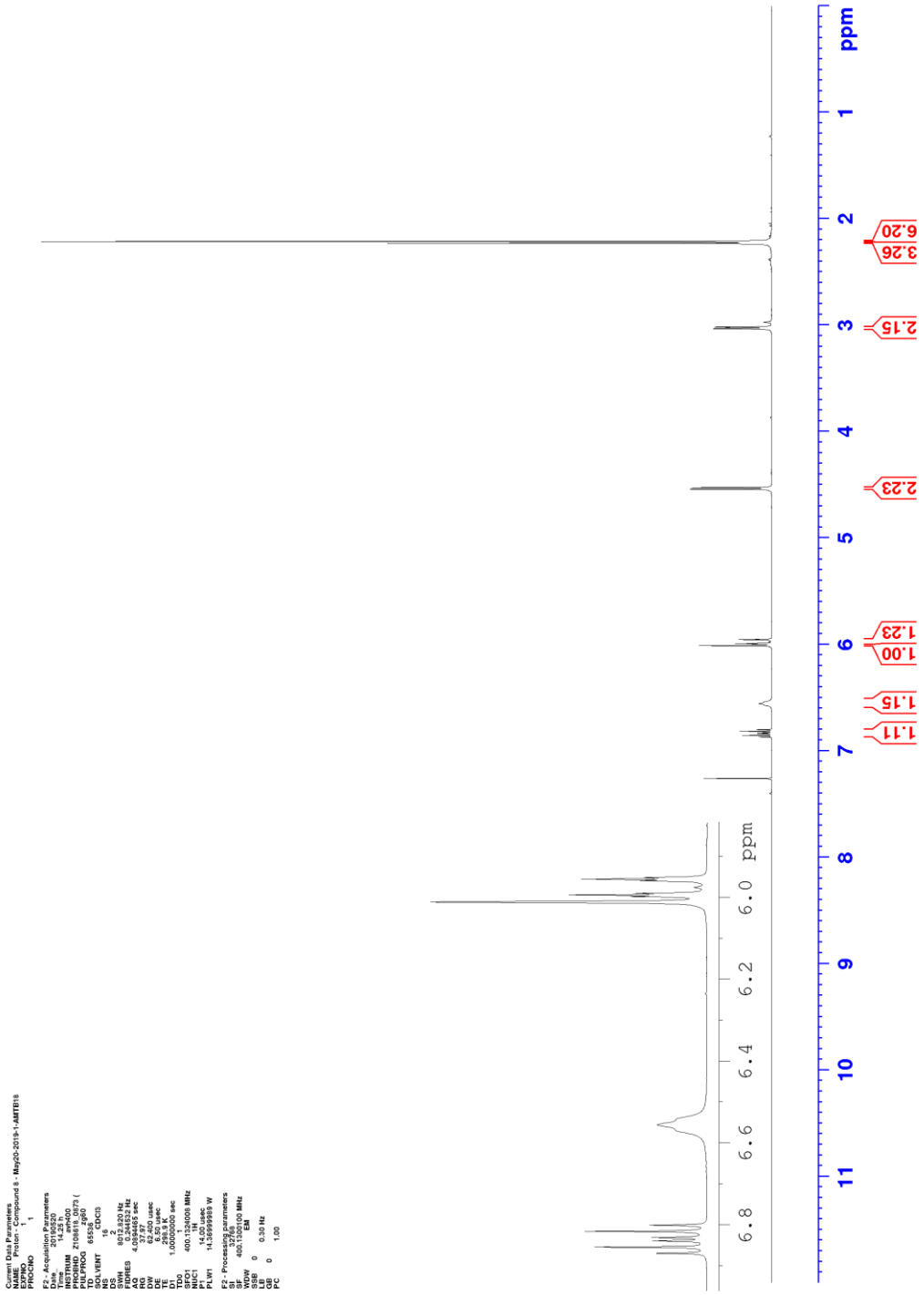


_N_-[(3-Methylisoxazol-5-yl)methyl]prop-2[ii]-enamide

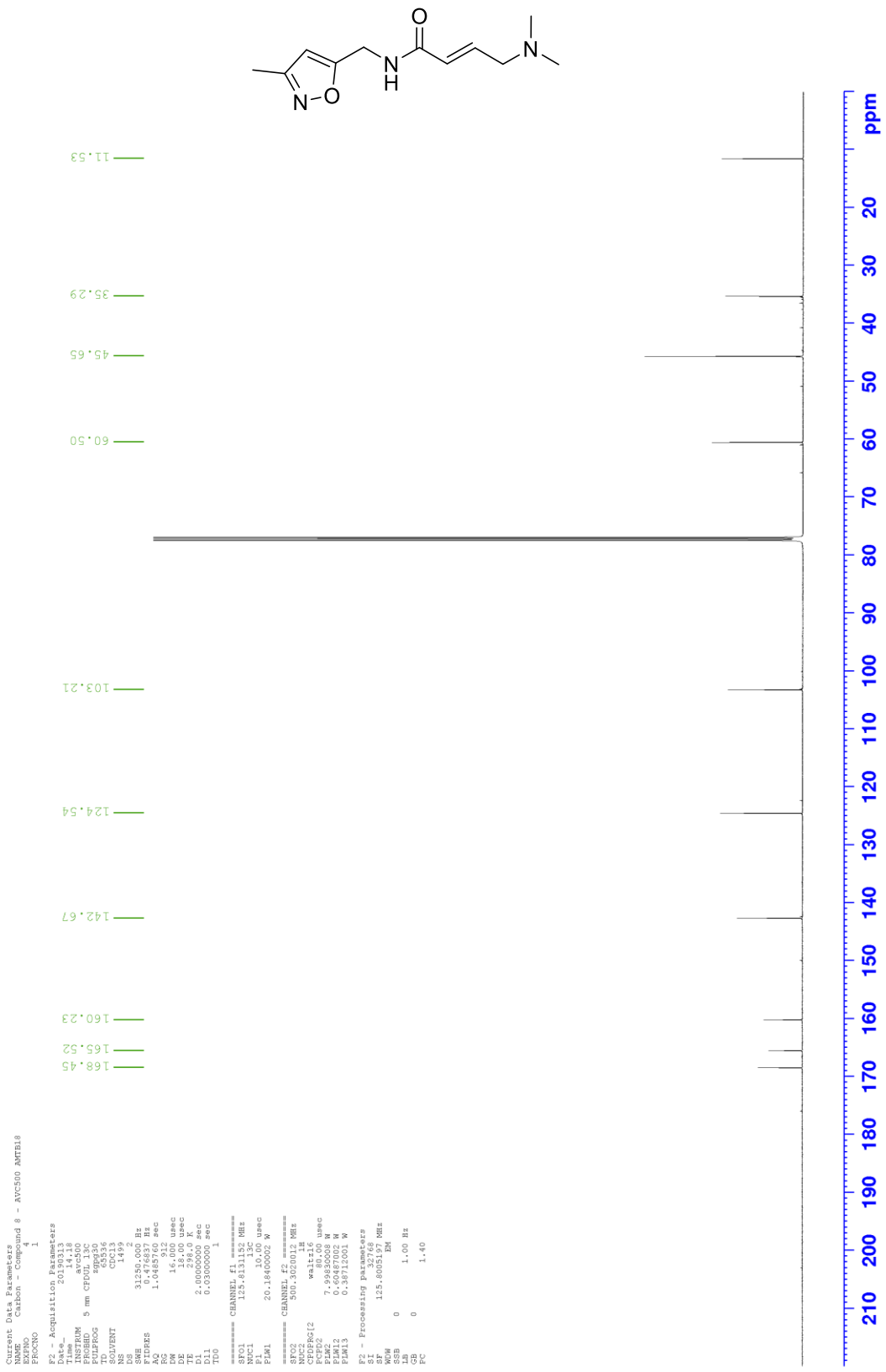


(*E*)-4[ii]-(Dimethylamino)-*N*-(3-methylisoxazol-5-yl)but-2[ii]-enamide

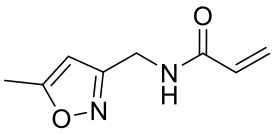


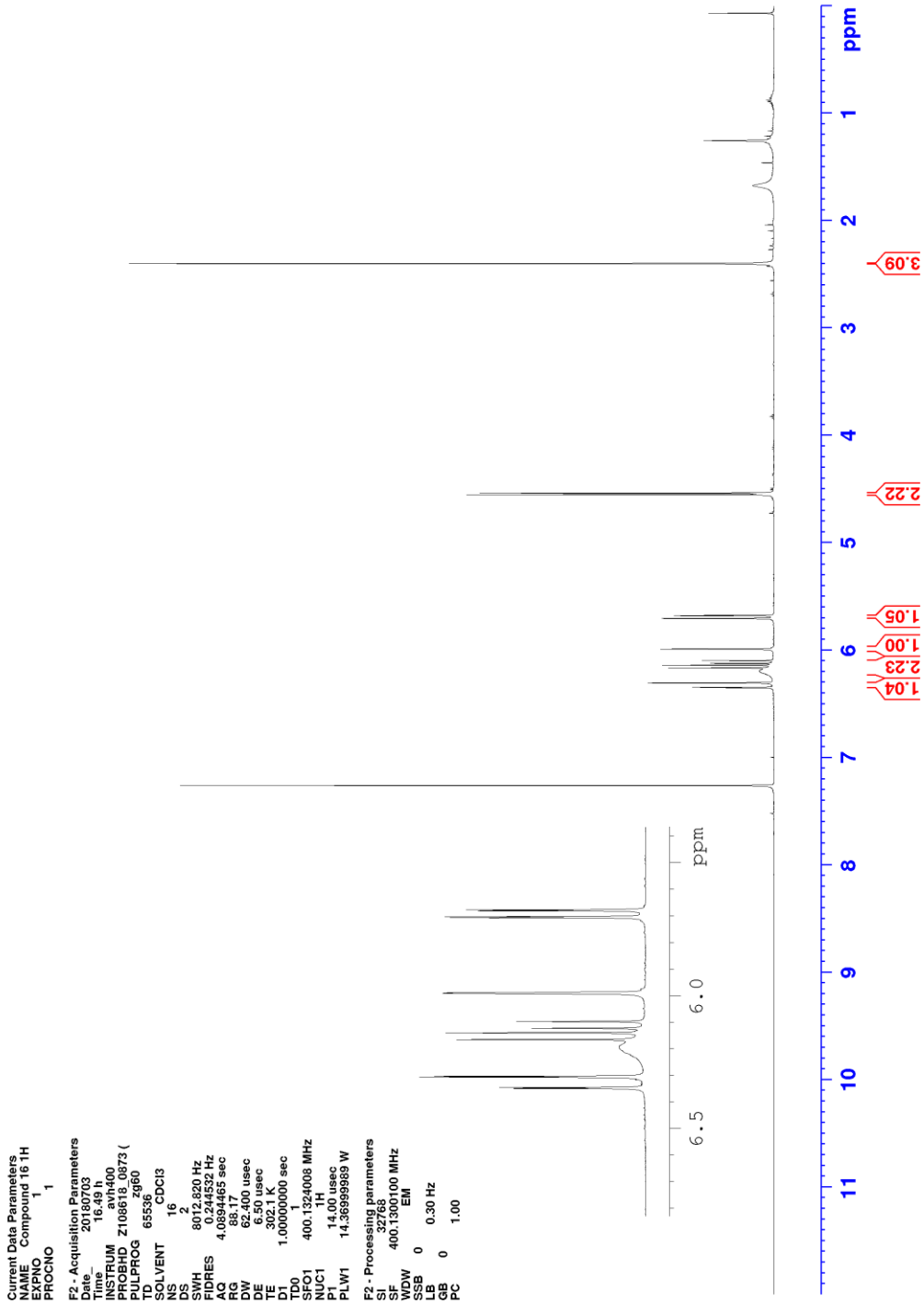


(*E*)-4[ii]-(Dimethylamino)-*N*-(3-methylisoxazol-5-yl)but-2[ii]-enamide

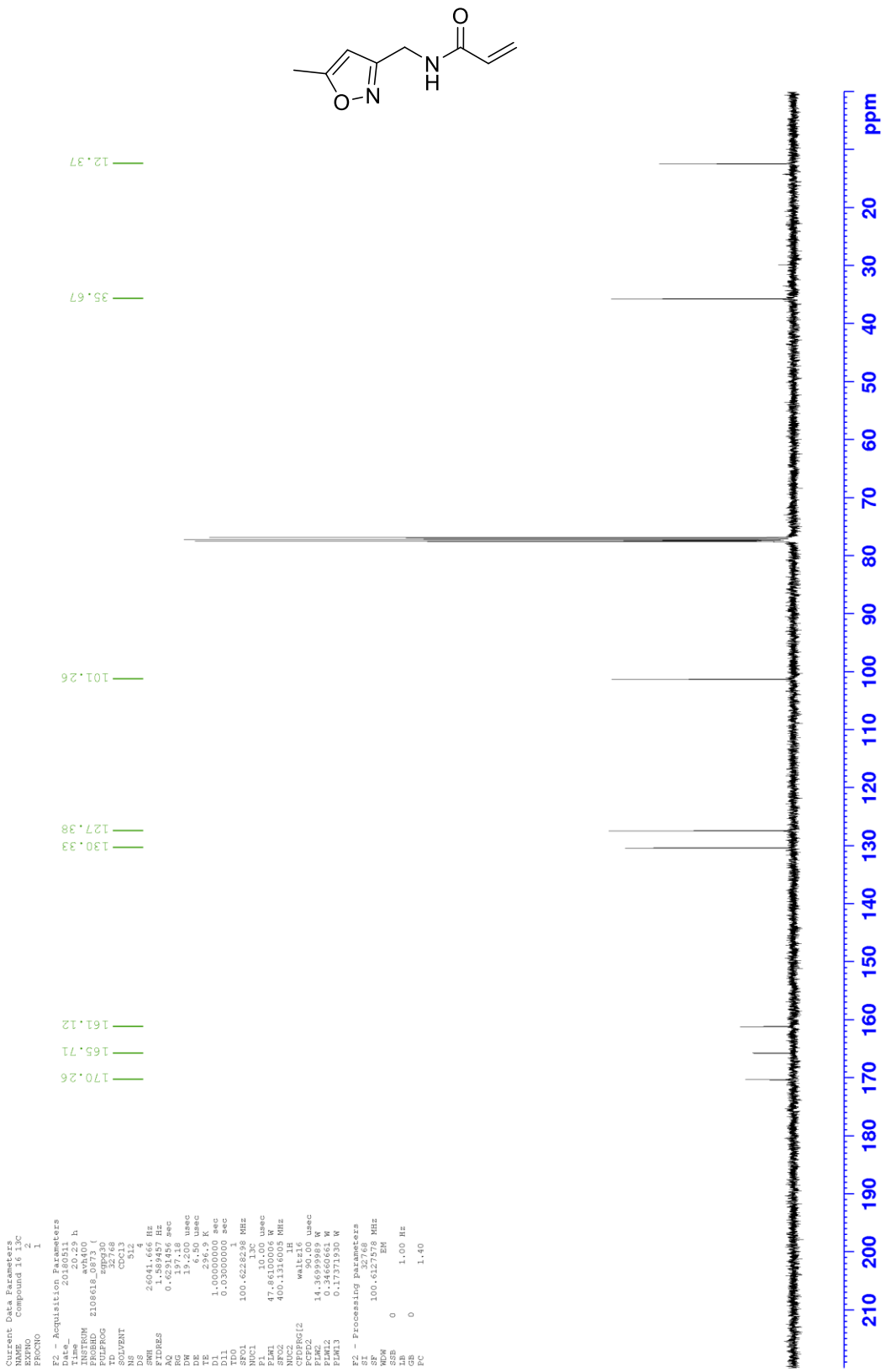


_N_-[(5-Methylisoxazol-3-yl)methyl]prop-2^{ii}-enamide



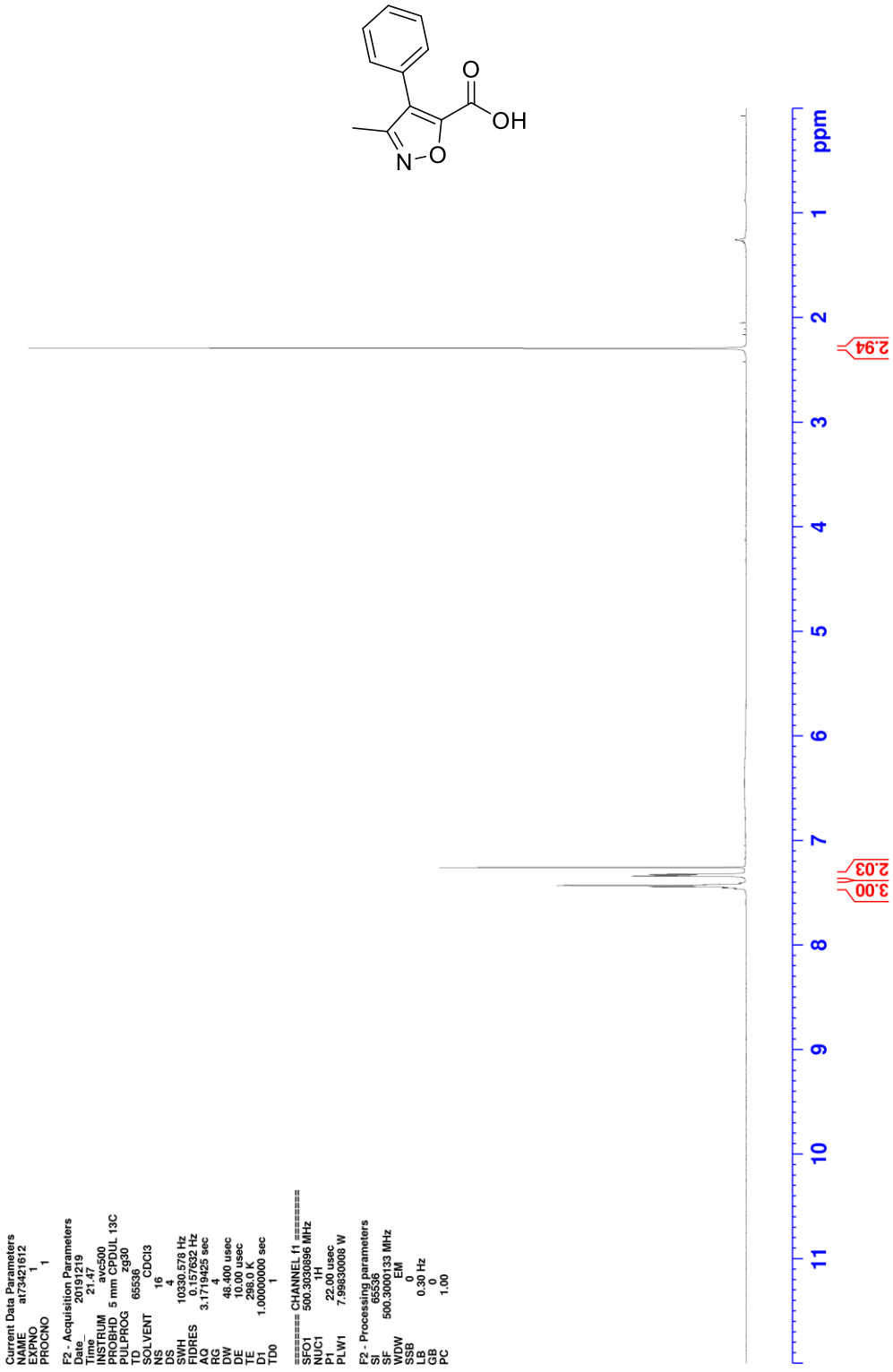


_N_-[(5-Methylisoxazol-3-yl)methyl]prop-2[ii]-enamide



Chapter 4 spectra

3-Methyl-4-phenylisoxazole-5-carboxylic acid



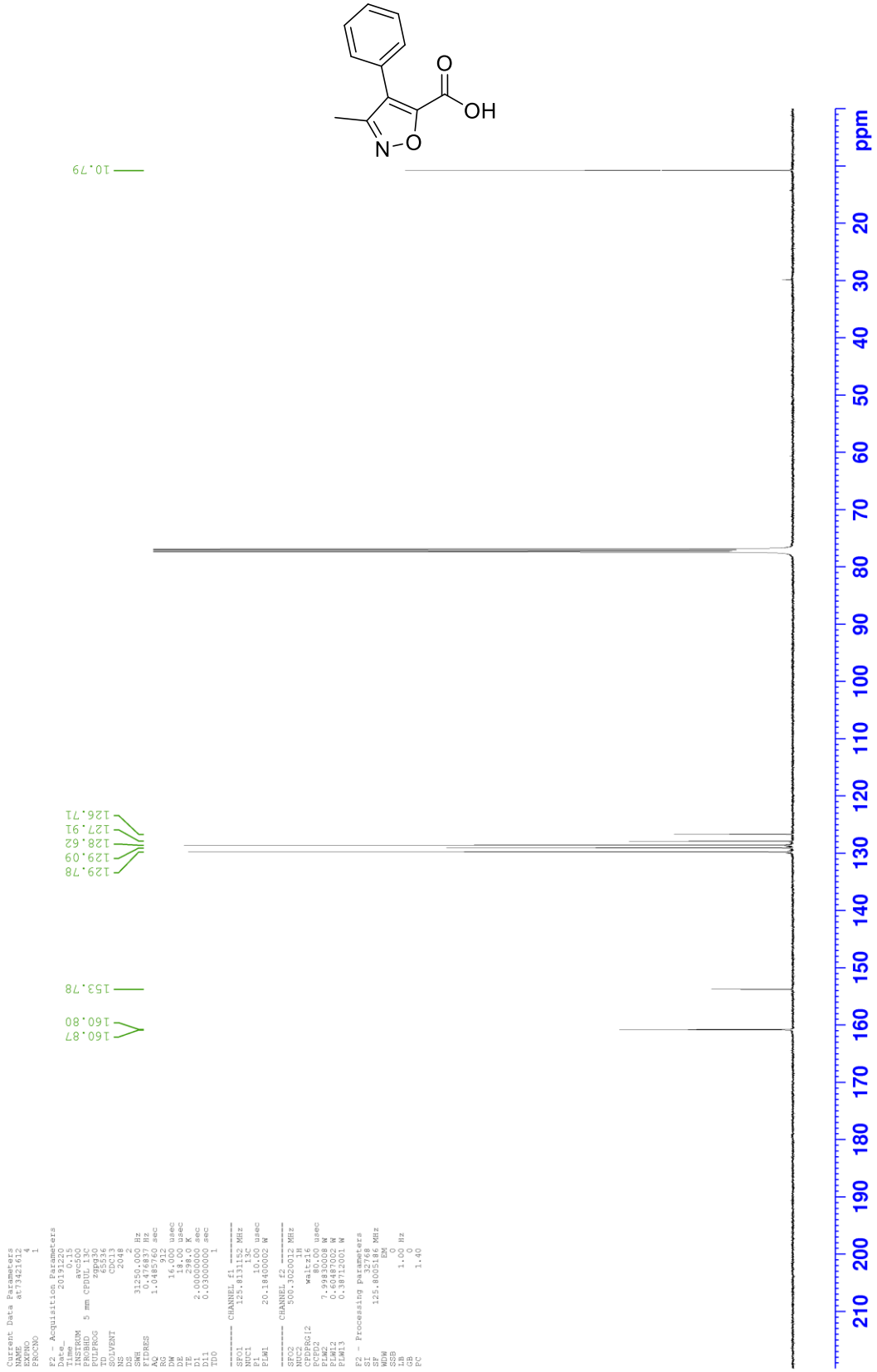
Current Data Parameters
NAME at73421612
EXPNO 1
PROCNO 1

F2 - Acquisition Parameters
Date_ 20191219
Time 21.47
INSTRUM avc500
PROBHD 5 mm CPDUL 13C
PULPROG zg30
TD 65536
SOLVENT CDCl3
NS 16
DS 4
SWH 10330.578 Hz
FIDRES 0.157632 Hz
AQ 3.17719425 sec
RG 4
DW 48.400 usec
DE 10.00 usec
TE 298.0 K
D1 1.00000000 sec
TD0 1

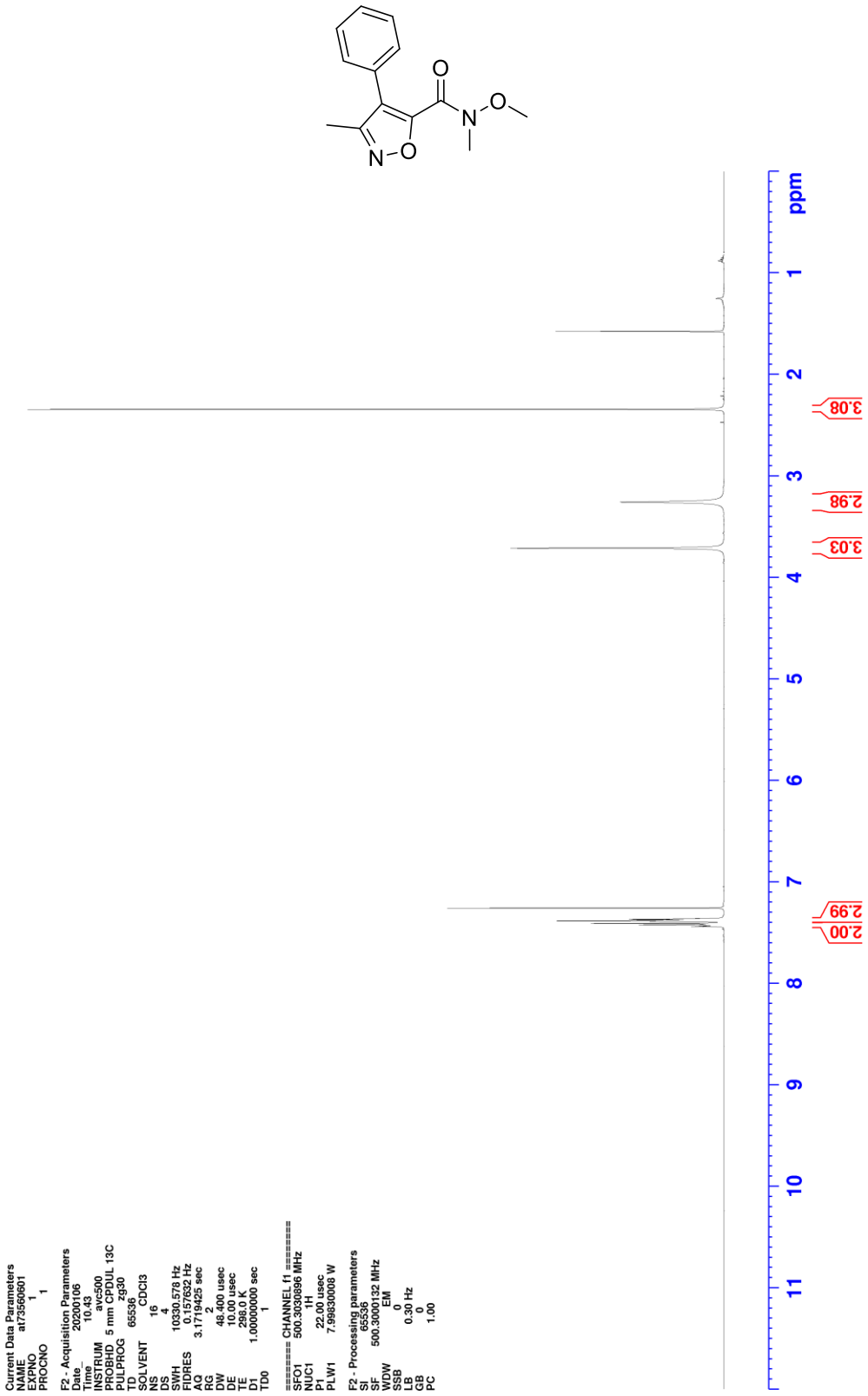
======== CHANNEL f1 ========
SFO1 500.3030896 MHz
NUC1 1H
P1 22.00 usec
PLW1 7.99830008 W

F2 - Processing parameters
SI 65536
SF 500.3000133 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

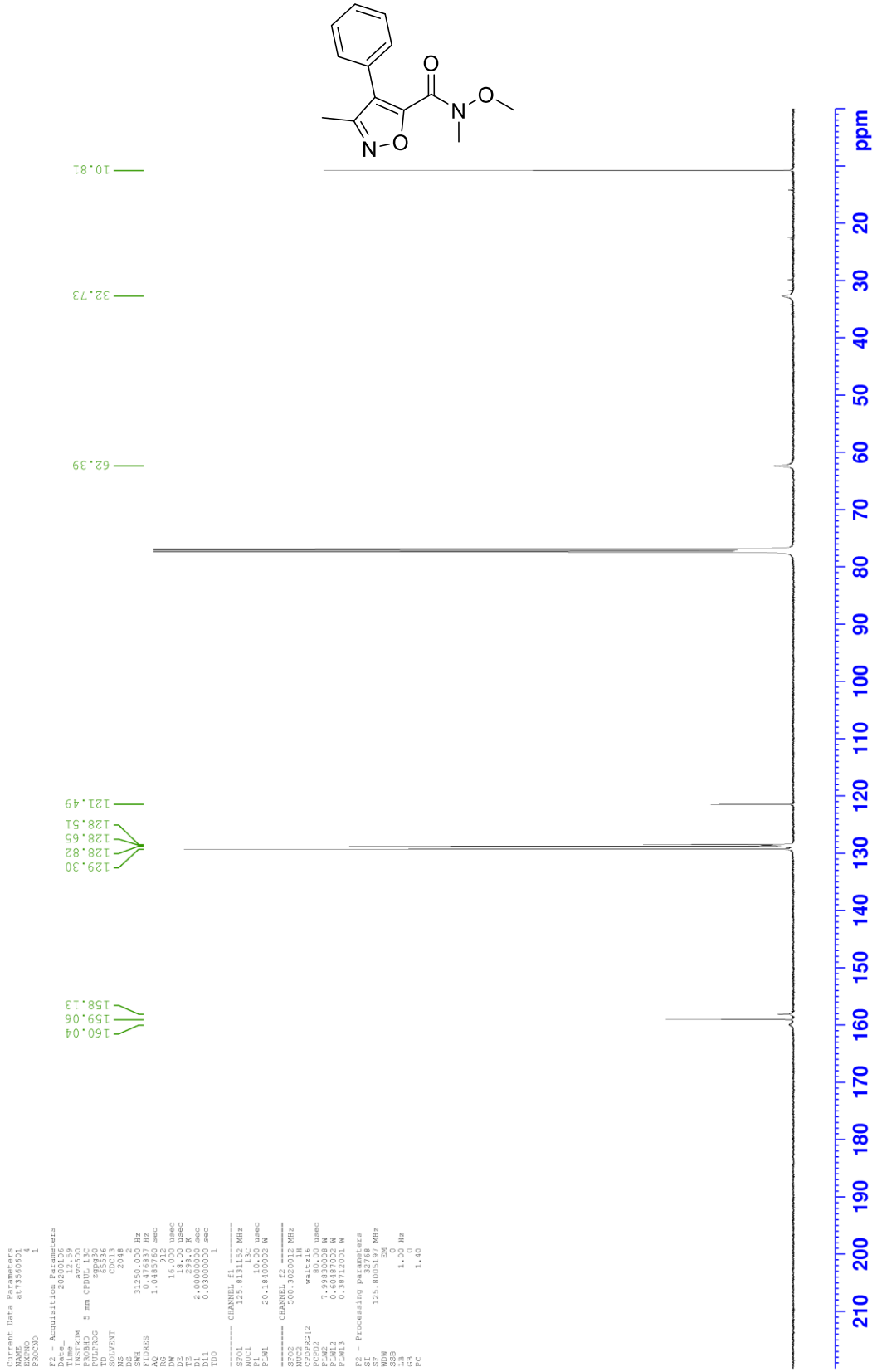
3-Methyl-4-phenylisoxazole-5-carboxylic acid



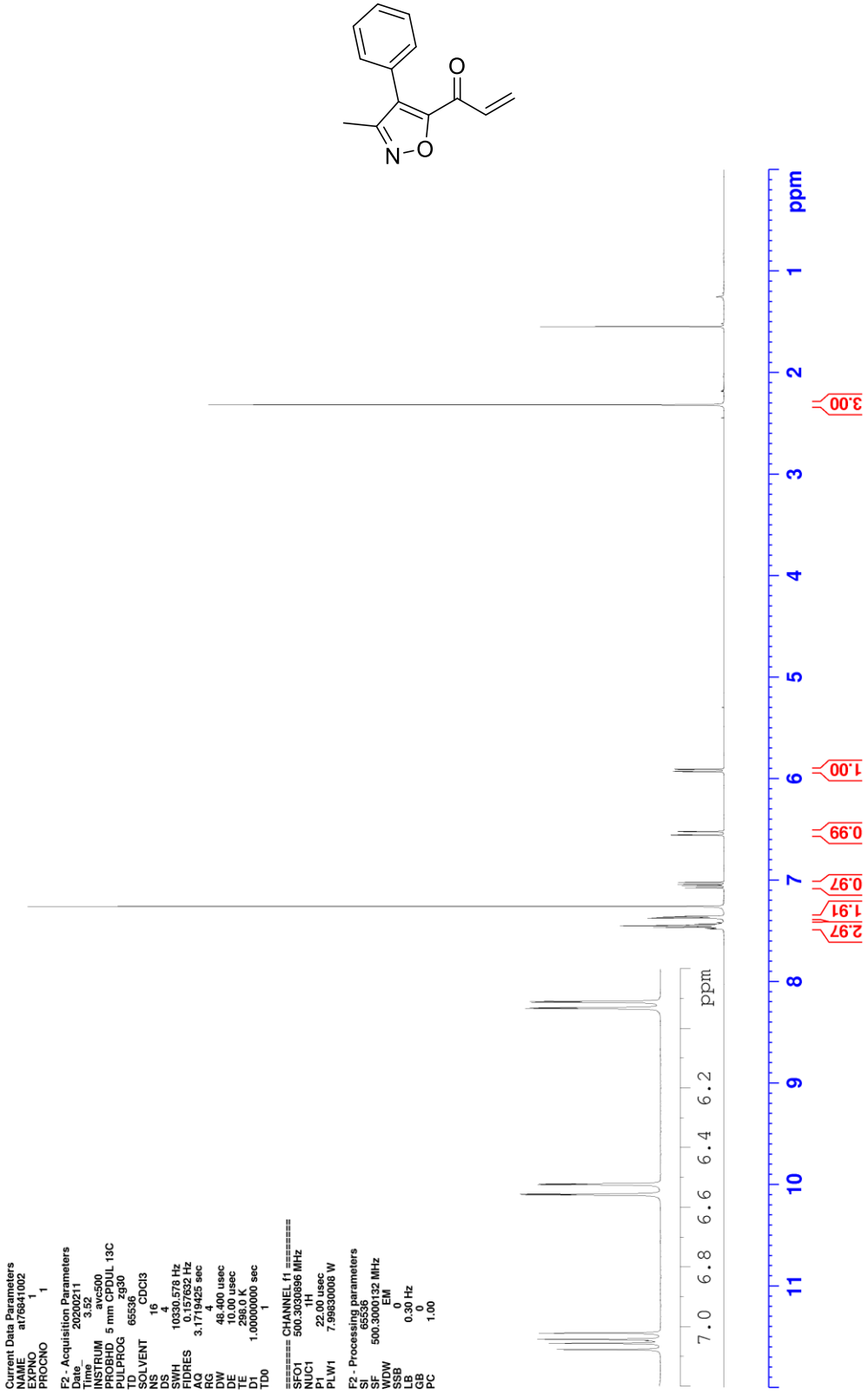
_N_-Methoxy-_N_,3-dimethyl-4-phenylisoxazole-5-carboxamide



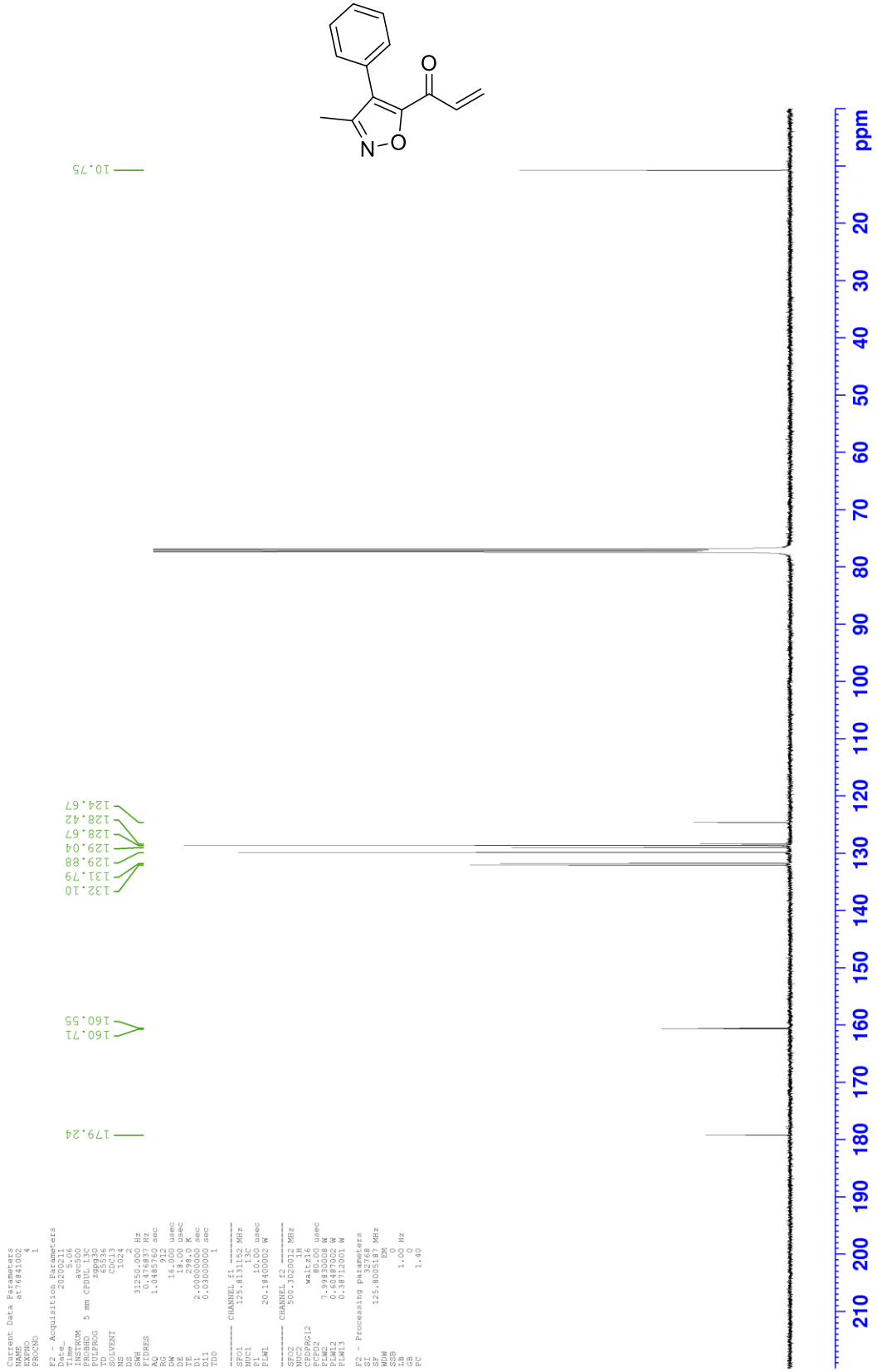
_N_-Methoxy-_N_,3-dimethyl-4-phenylisoxazole-5-carboxamide



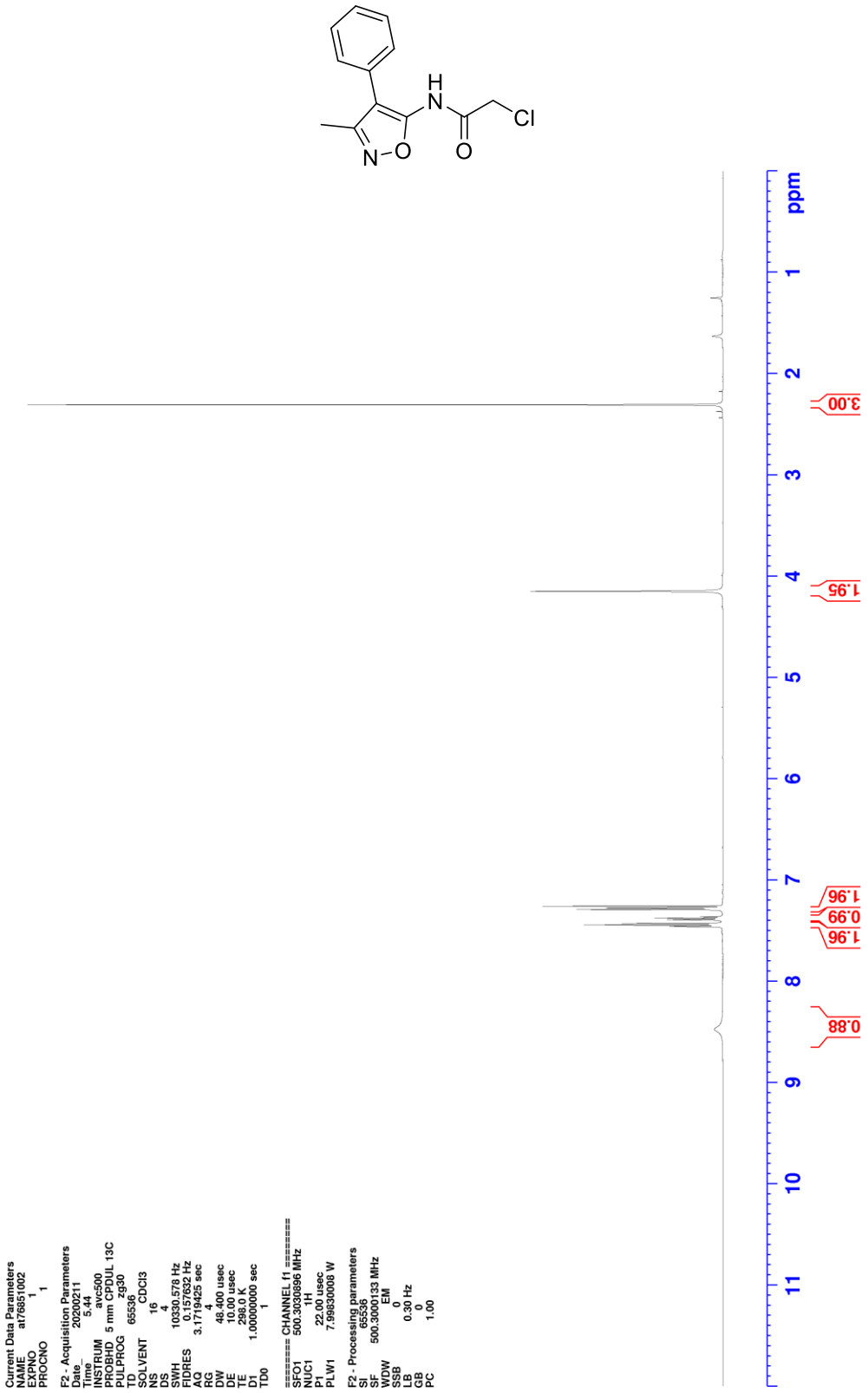
1[iii]-(3-Methyl-4-phenylisoxazol-5-yl)prop-2[iii]-en-1[iii]-one



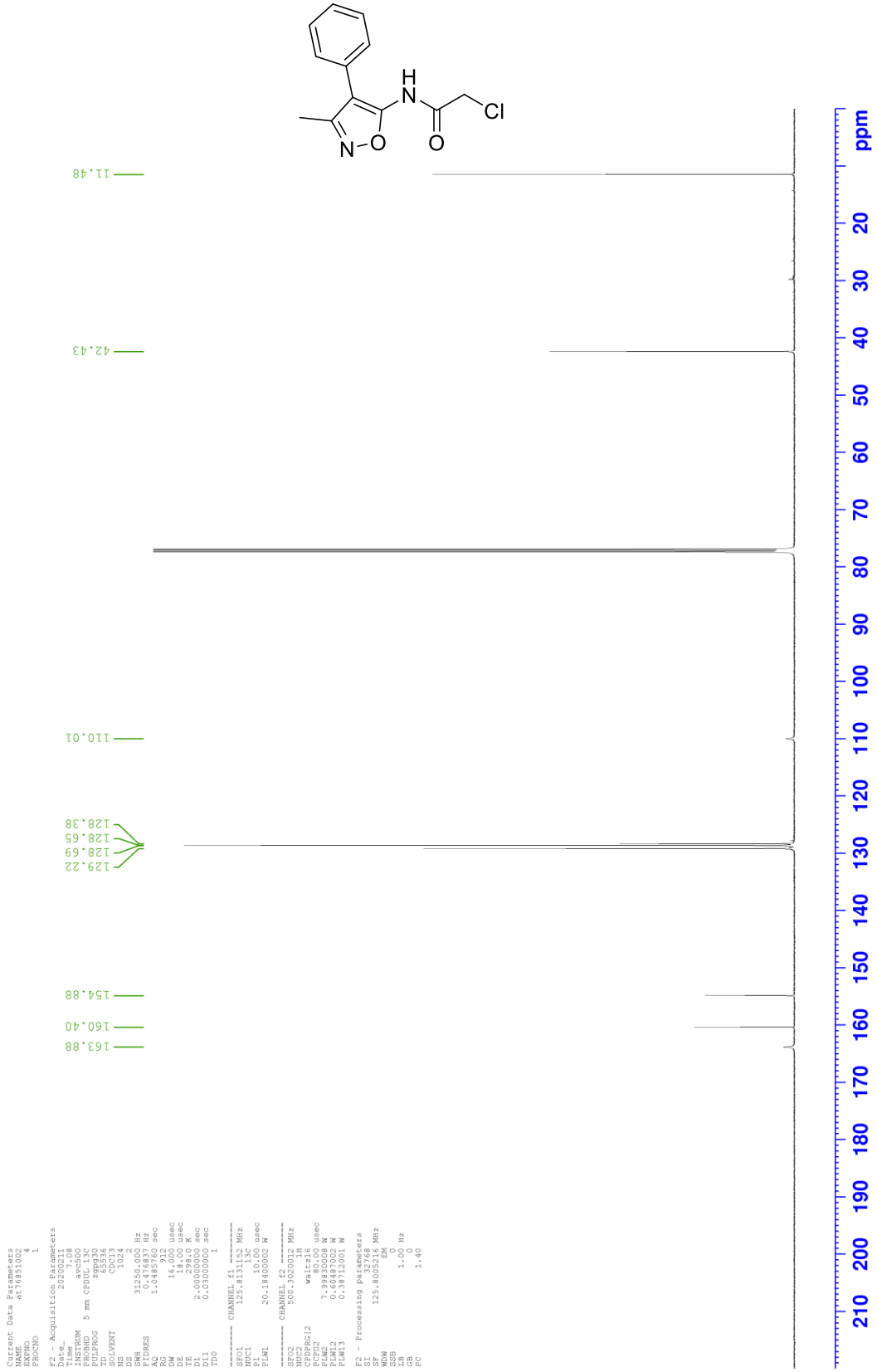
1[iii]-(3-Methyl-4-phenylisoxazol-5-yl)prop-2[iii]-en-1[iii]-one



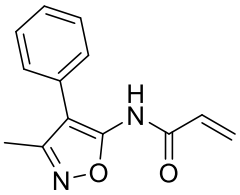
2[iii]-Chloro-*N*-(3-methyl-4-phenylisoxazol-5-yl)acetamide

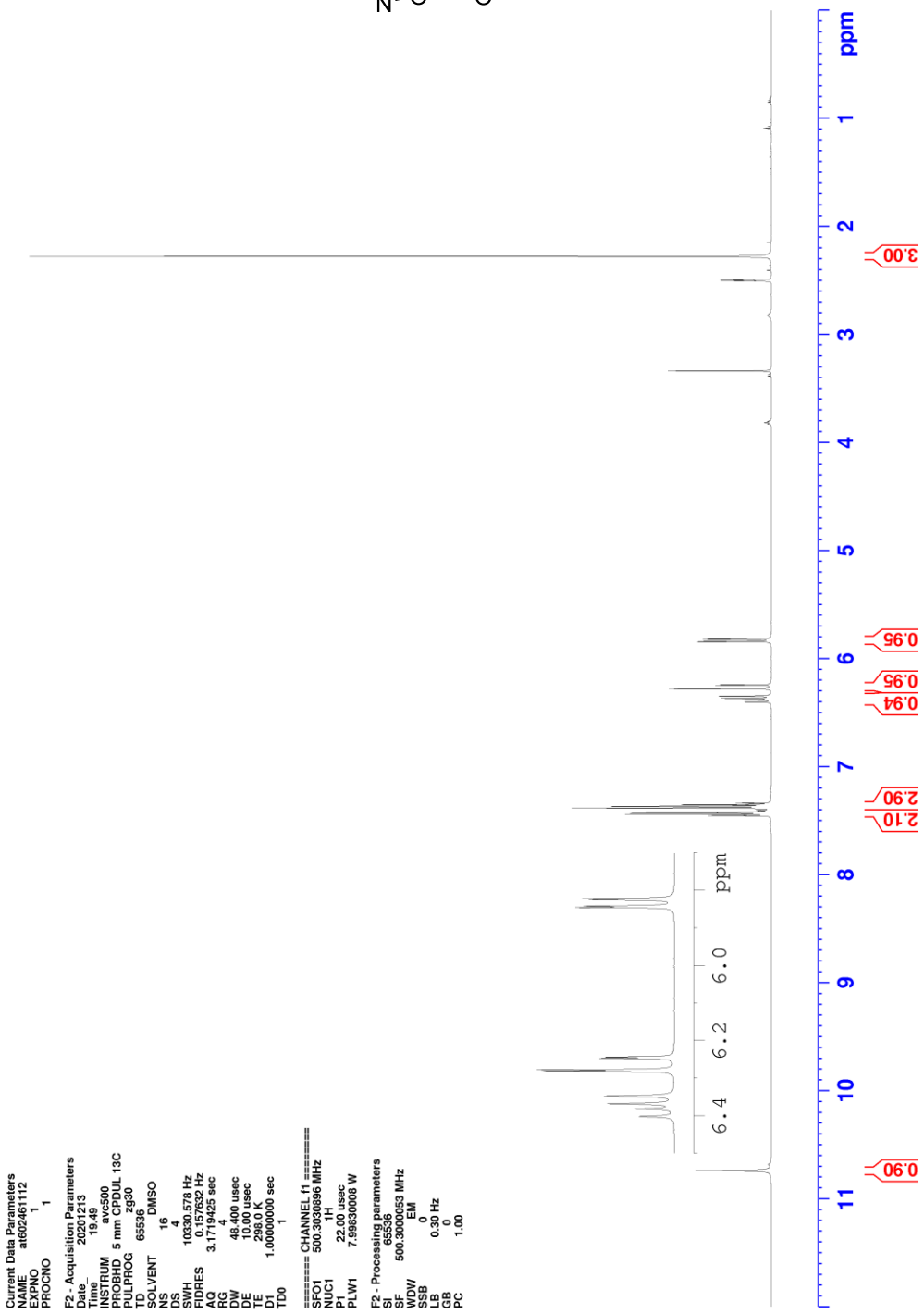


2[iii]-Chloro-*N*-(3-methyl-4-phenylisoxazol-5-yl)acetamide

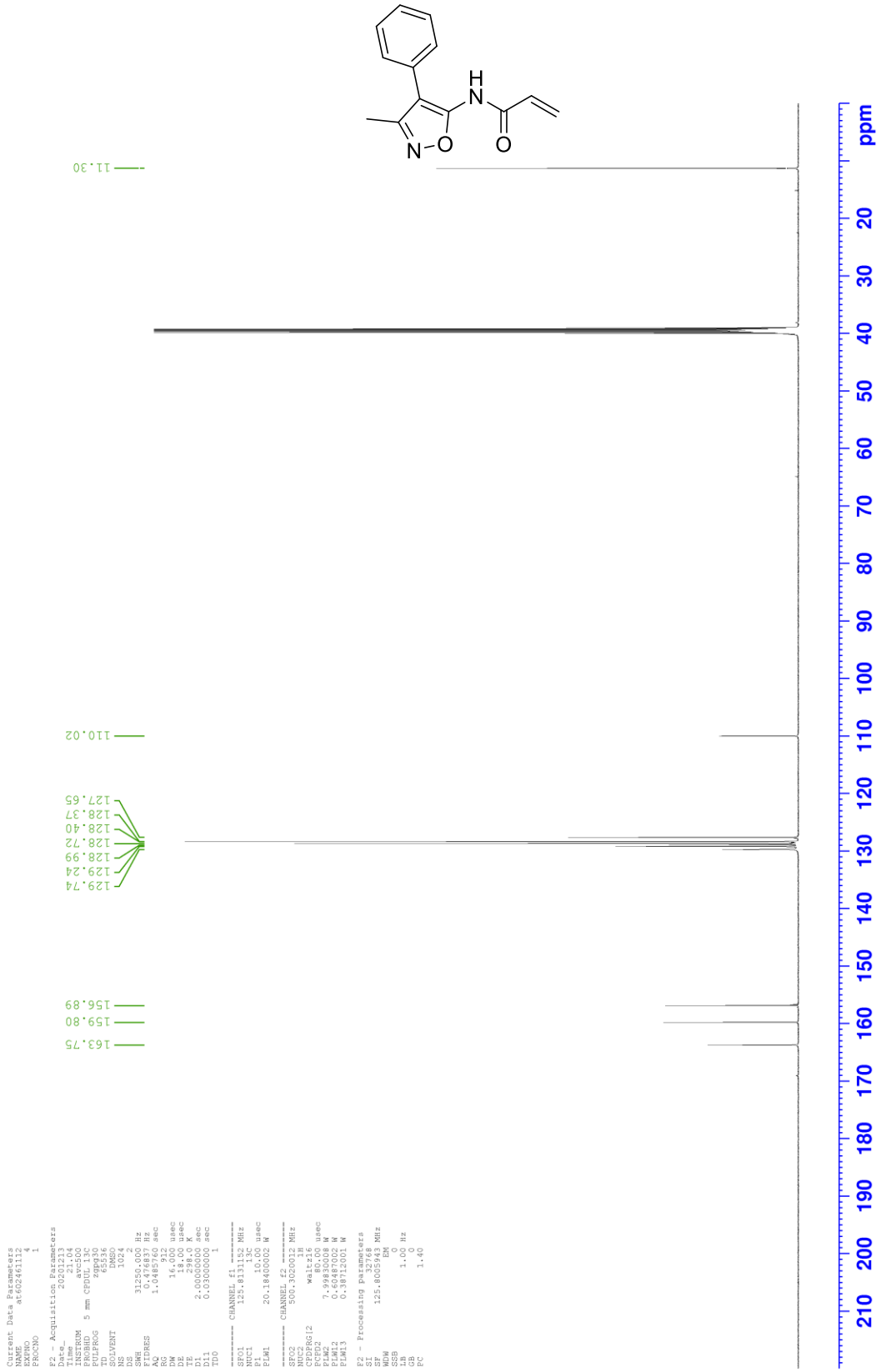


***N*-(3-methyl-4-phenylisoxazol-5-yl)prop-2[iii]-enamide**

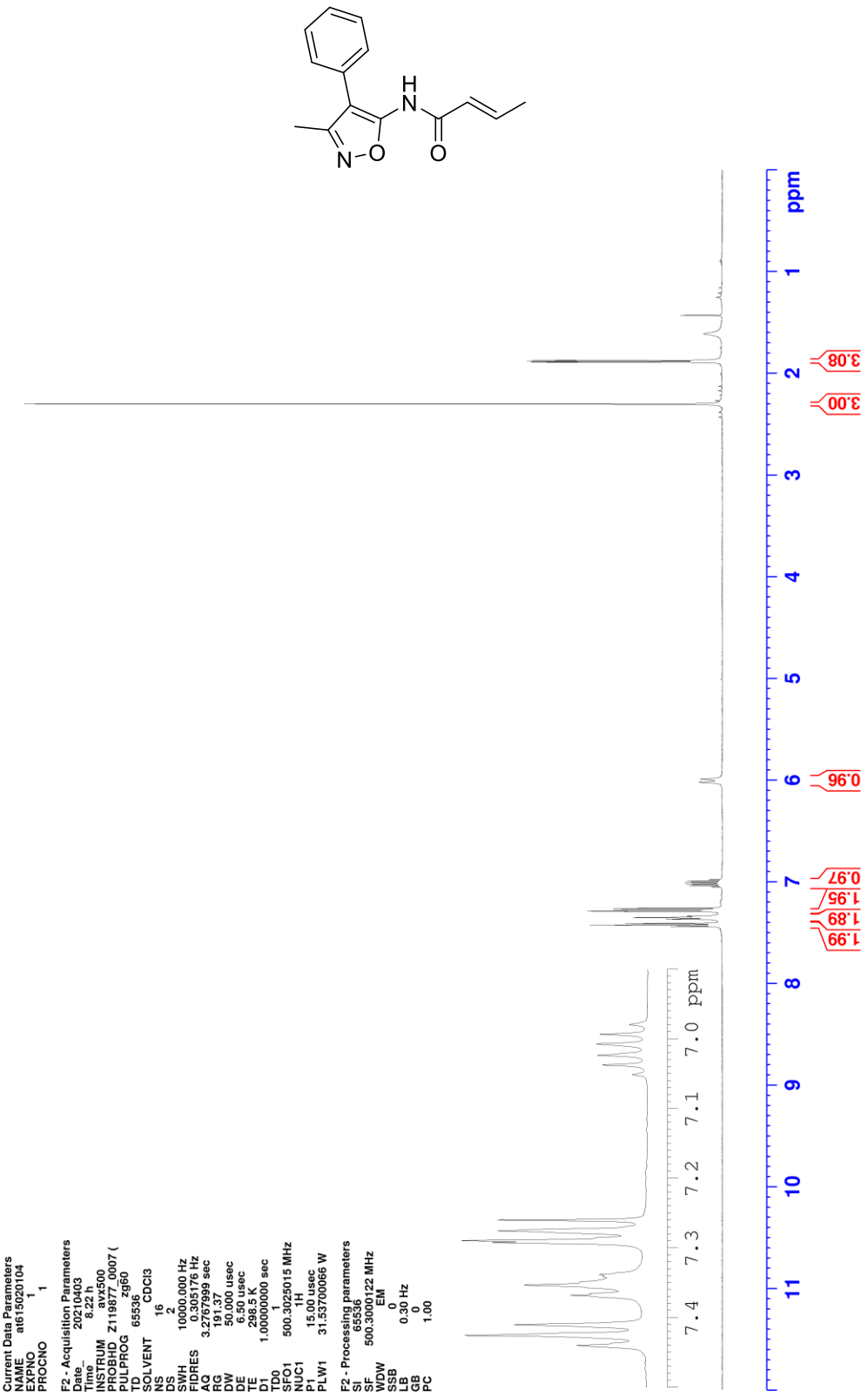




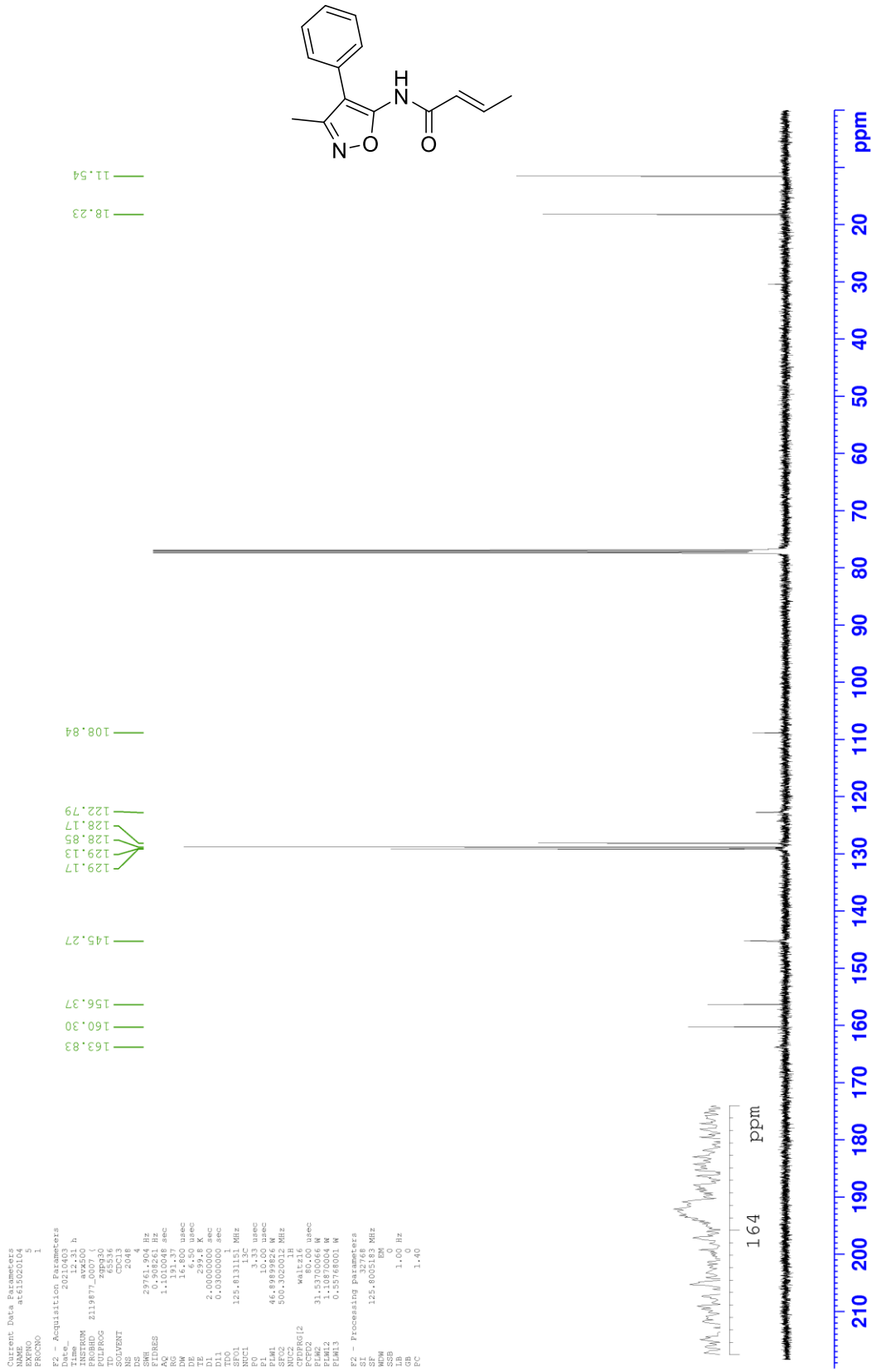
_N_-(3-methyl-4-phenylisoxazol-5-yl)prop-2[iii]-enamide



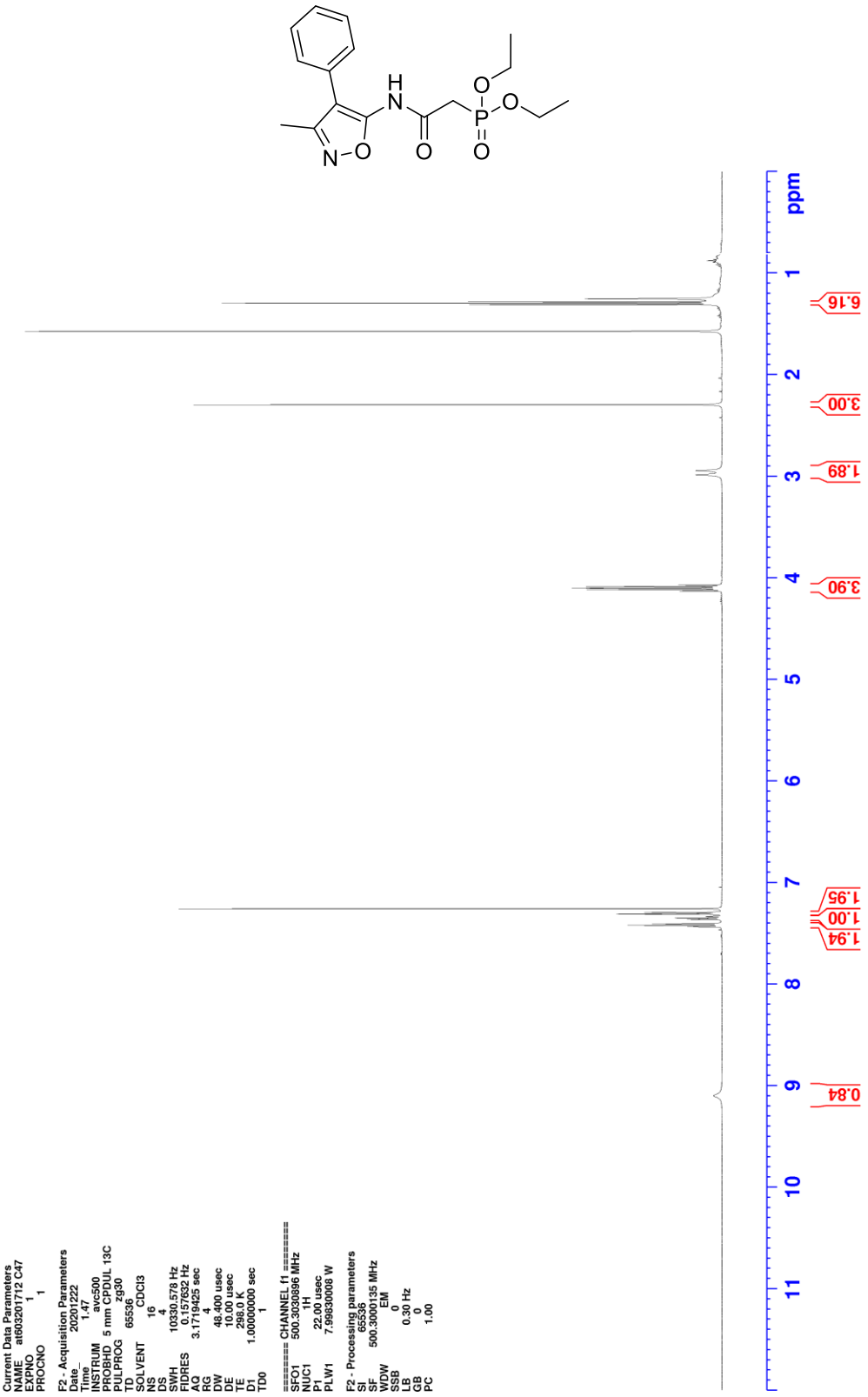
(*E*)-*N*-(3-methyl-4-phenylisoxazol-5-yl)but-2[iii]-enamide



(*E*)-*N*-(3-methyl-4-phenylisoxazol-5-yl)but-2[iii]-enamide



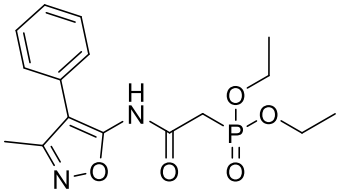
Diethyl {2[iii]-[(3-methyl-4-phenylisoxazol-5-yl)amino]-2[iii]-oxoethyl}phosphonate

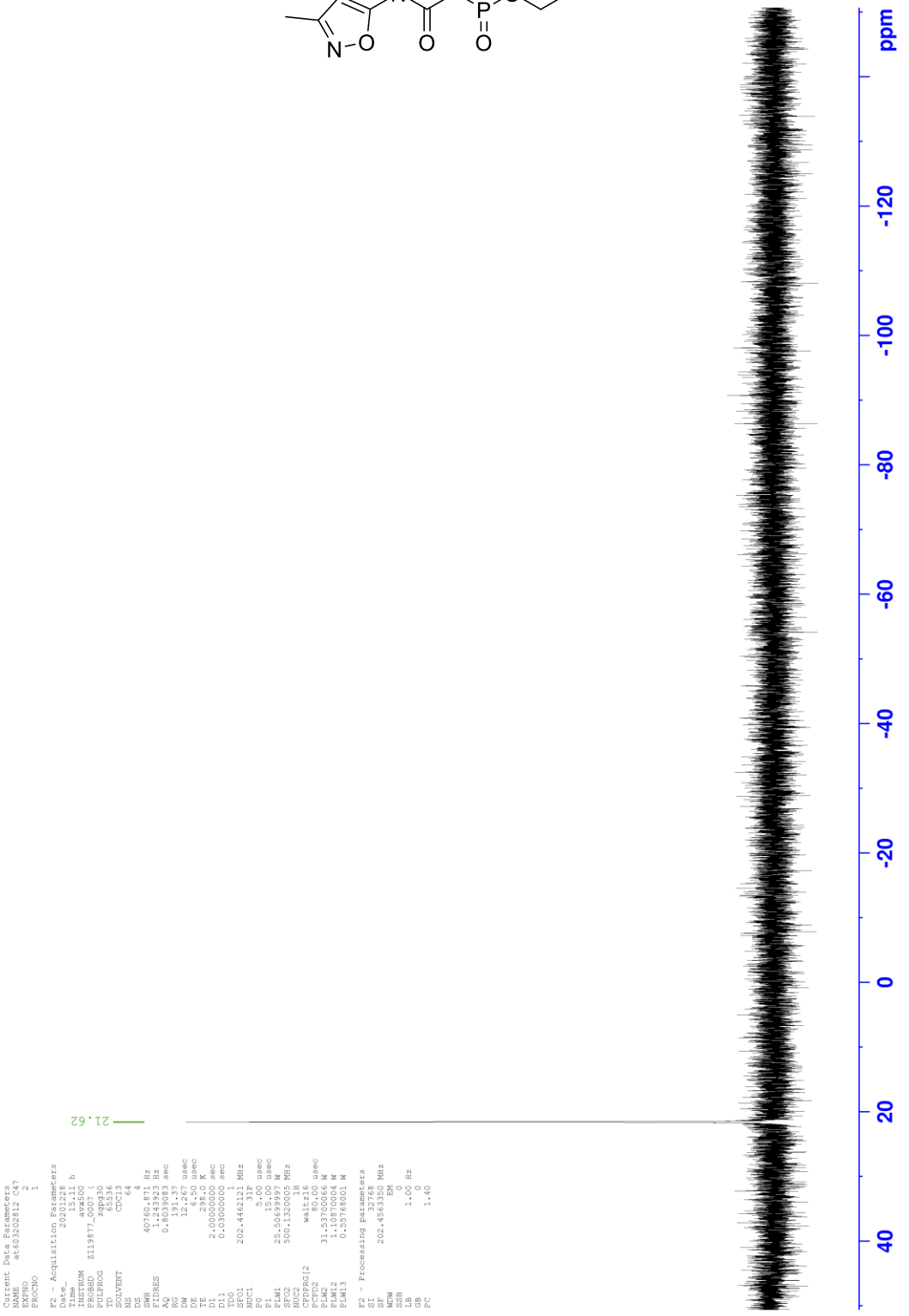


Diethyl {2^{iii}-[(3-methyl-4-phenylisoxazol-5-yl)amino]-2^{iii}-oxoethyl}phosphonate

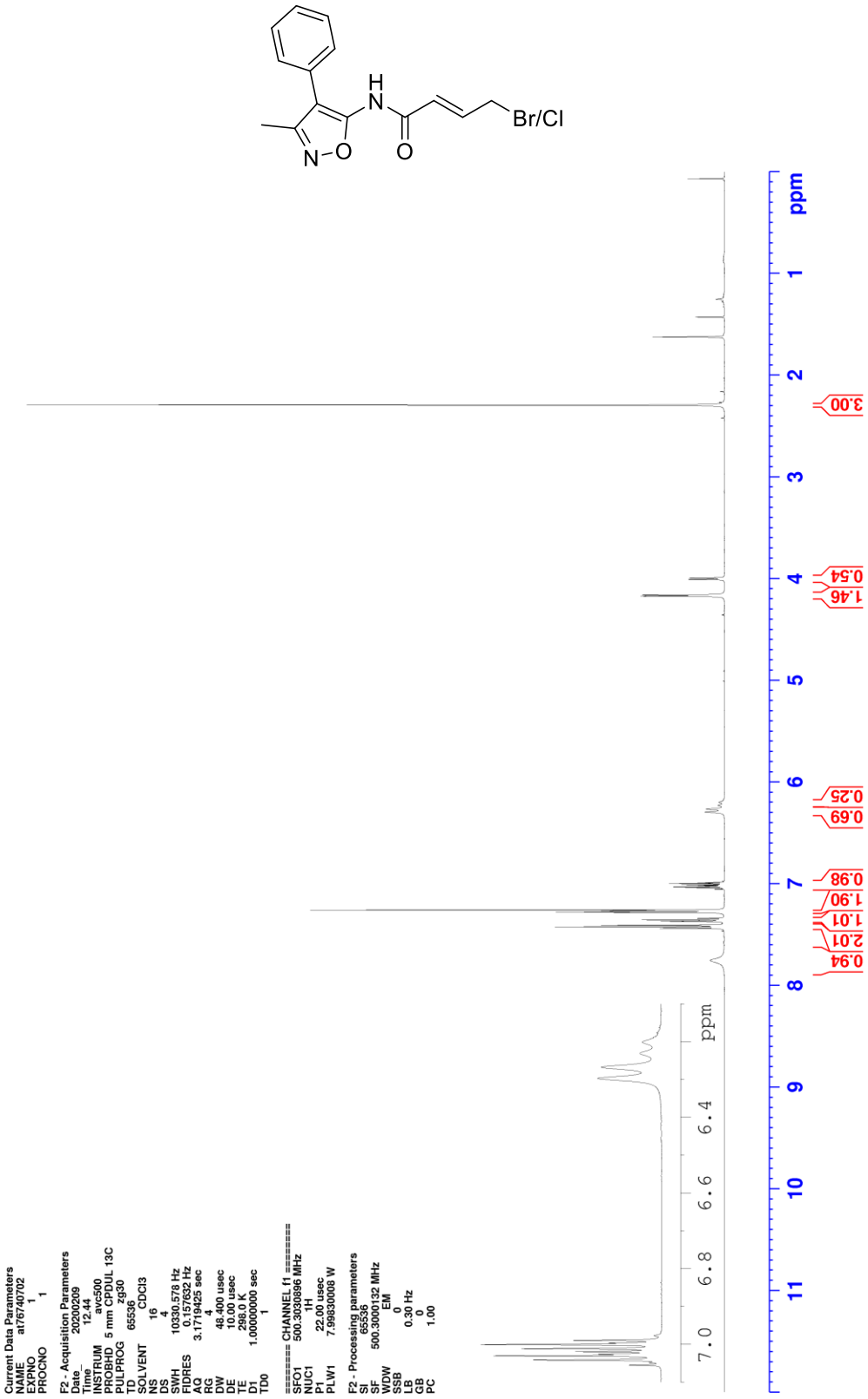


Diethyl {2^{iii}-[(3-methyl-4-phenylisoxazol-5-yl)amino]-2^{iii}-oxoethyl}phosphonate

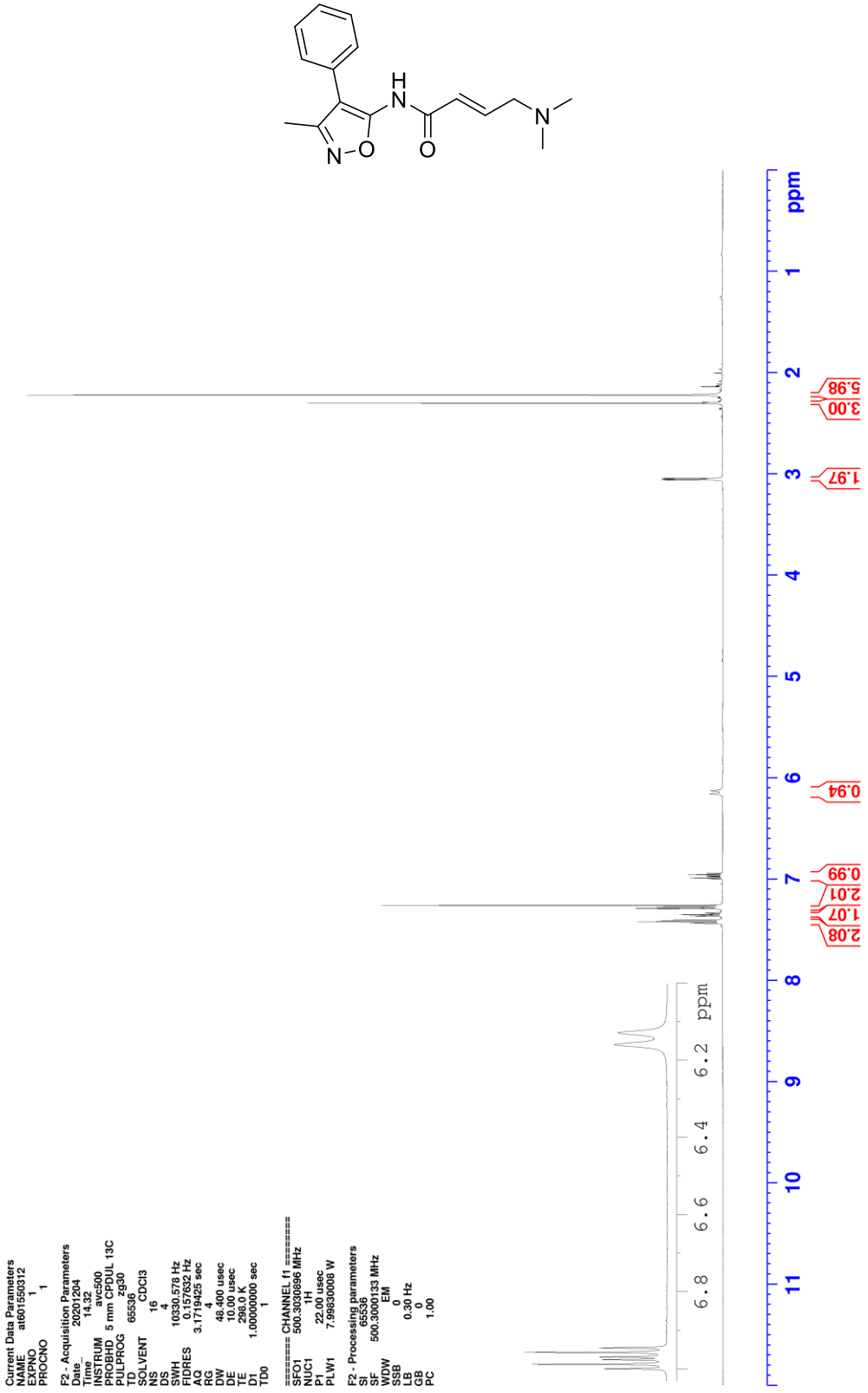




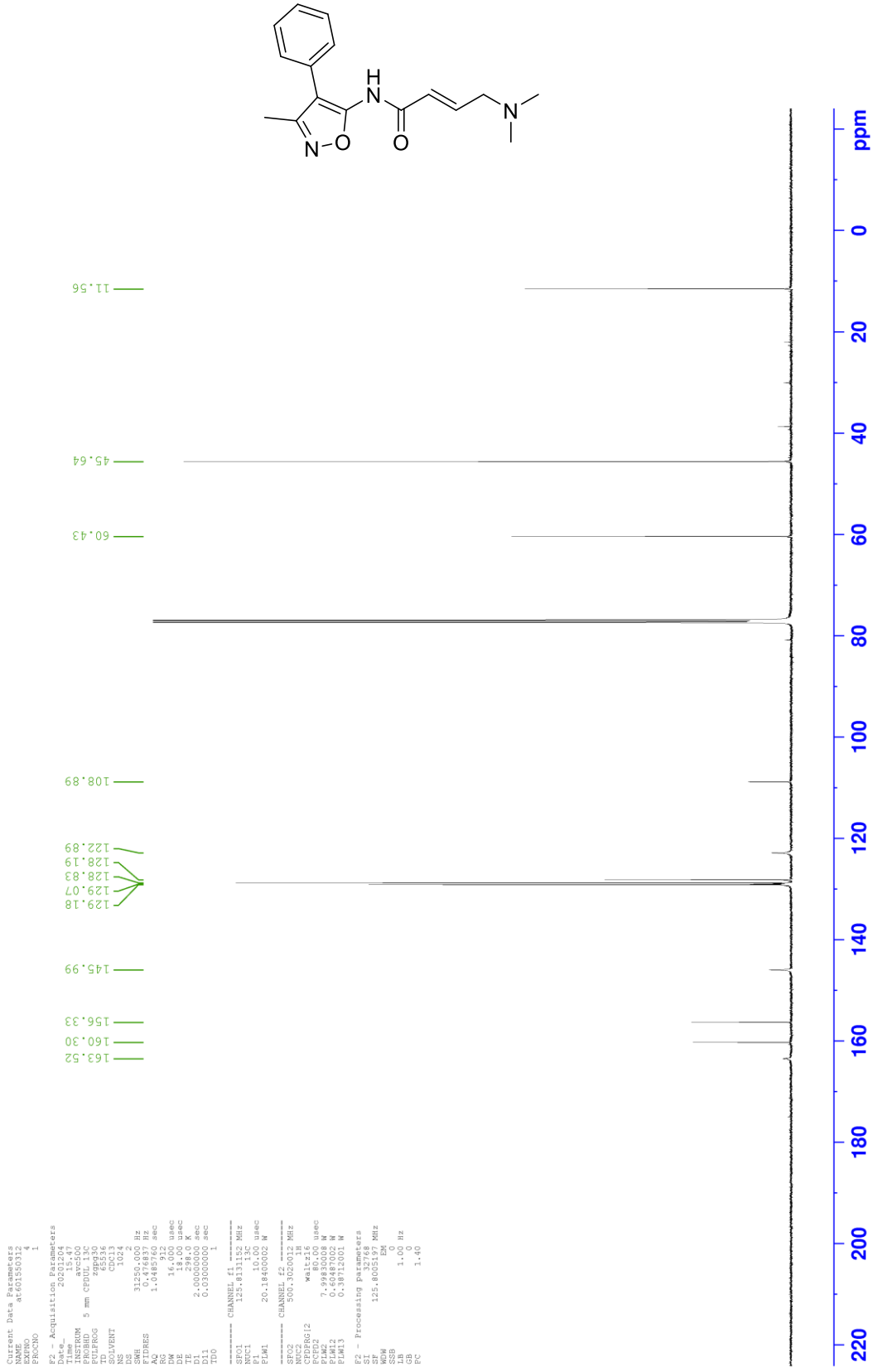
(*E*)-4[iii]-Bromo/Chloro-*N*-(3-methyl-4-phenylisoxazol-5-yl)but-2[iii]-enamide



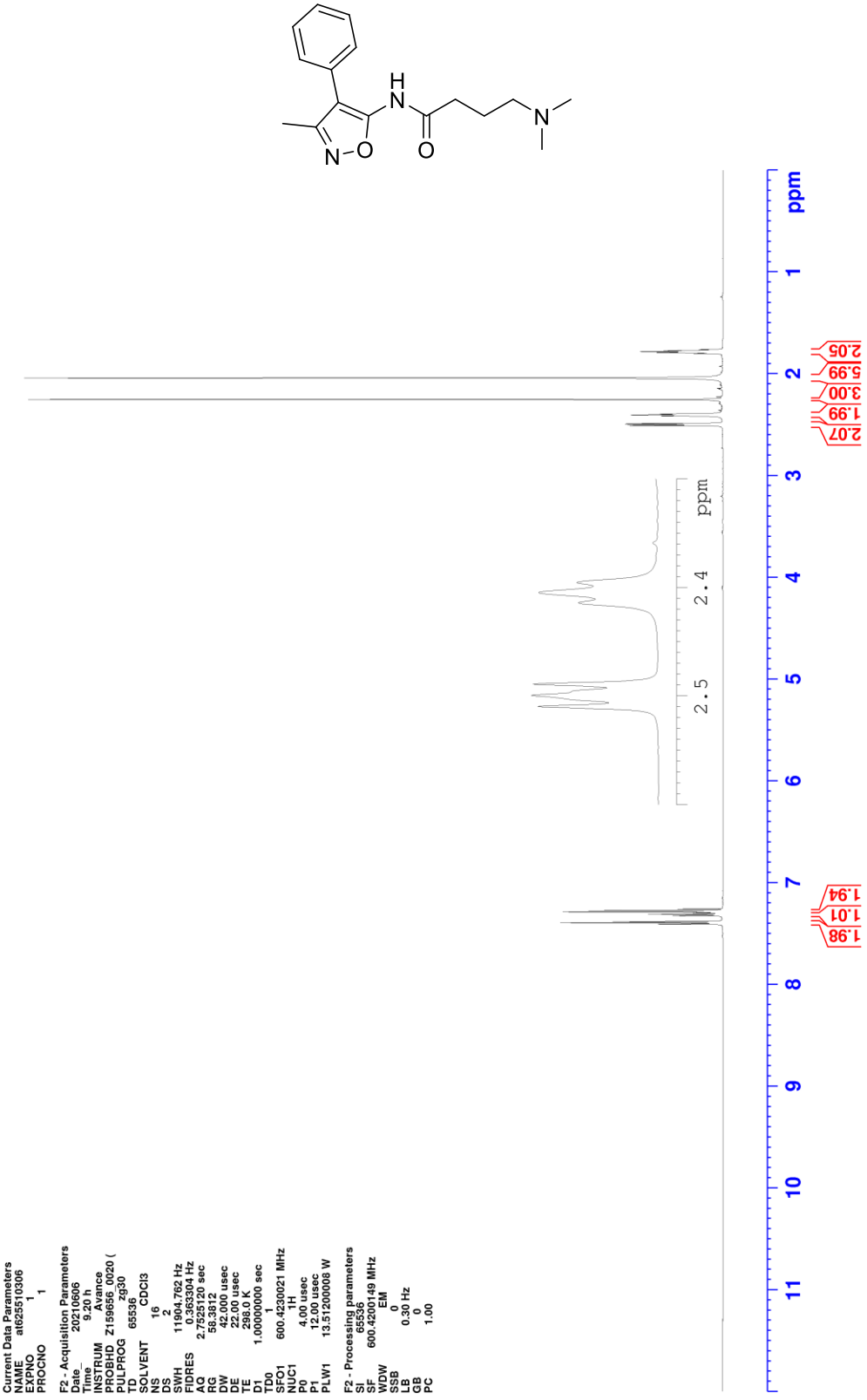
(*E*)-4[iii]-(Dimethylamino)-*N*-(3-methyl-4-phenylisoxazol-5-yl)but-2[iii]-enamide



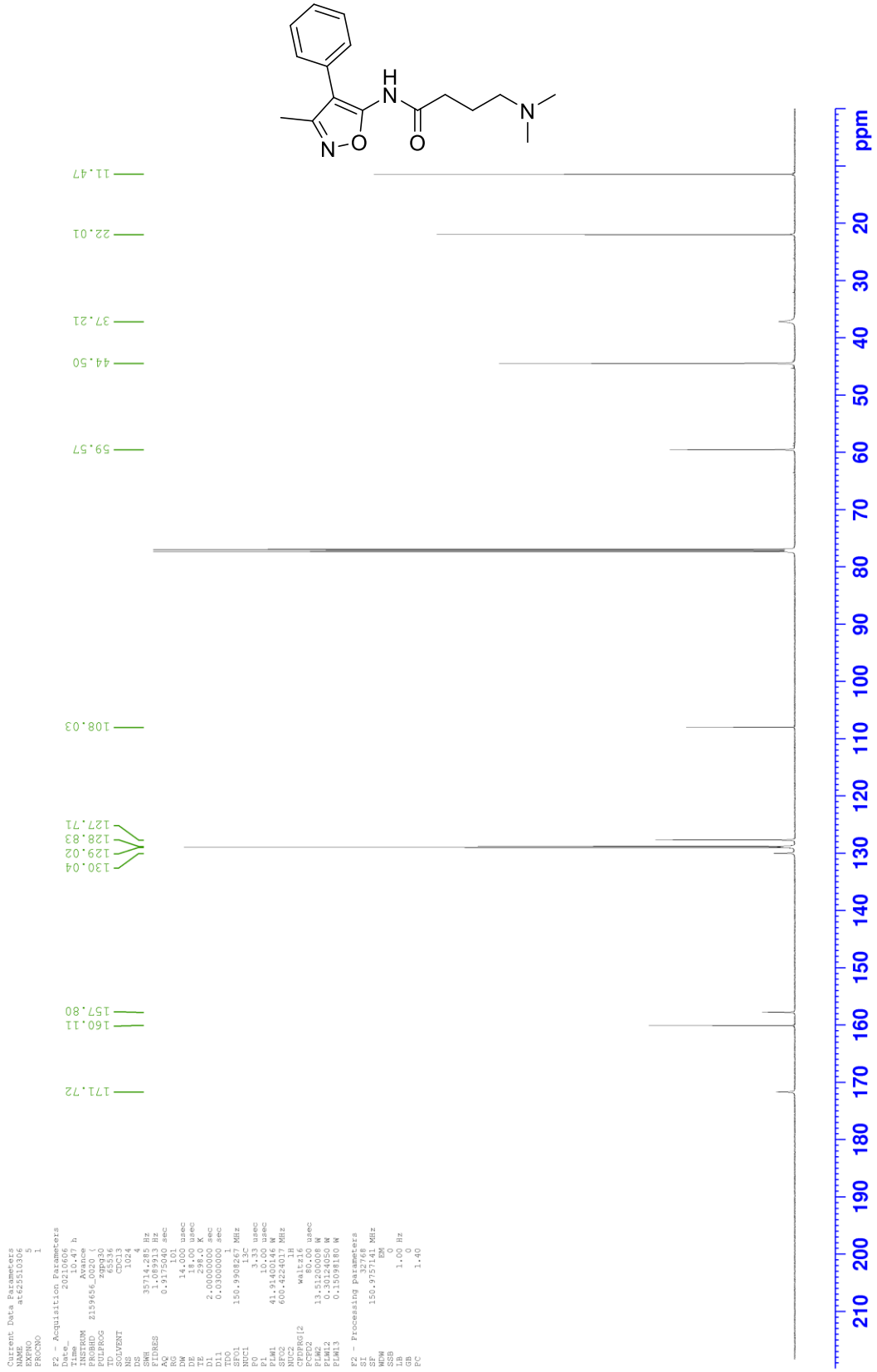
(*E*)-4[iii]-(Dimethylamino)-*N*-(3-methyl-4-phenylisoxazol-5-yl)but-2[iii]-enamide



4[iii]-(Dimethylamino)-*N*-(3-methyl-4-phenylisoxazol-5-yl)butanamide

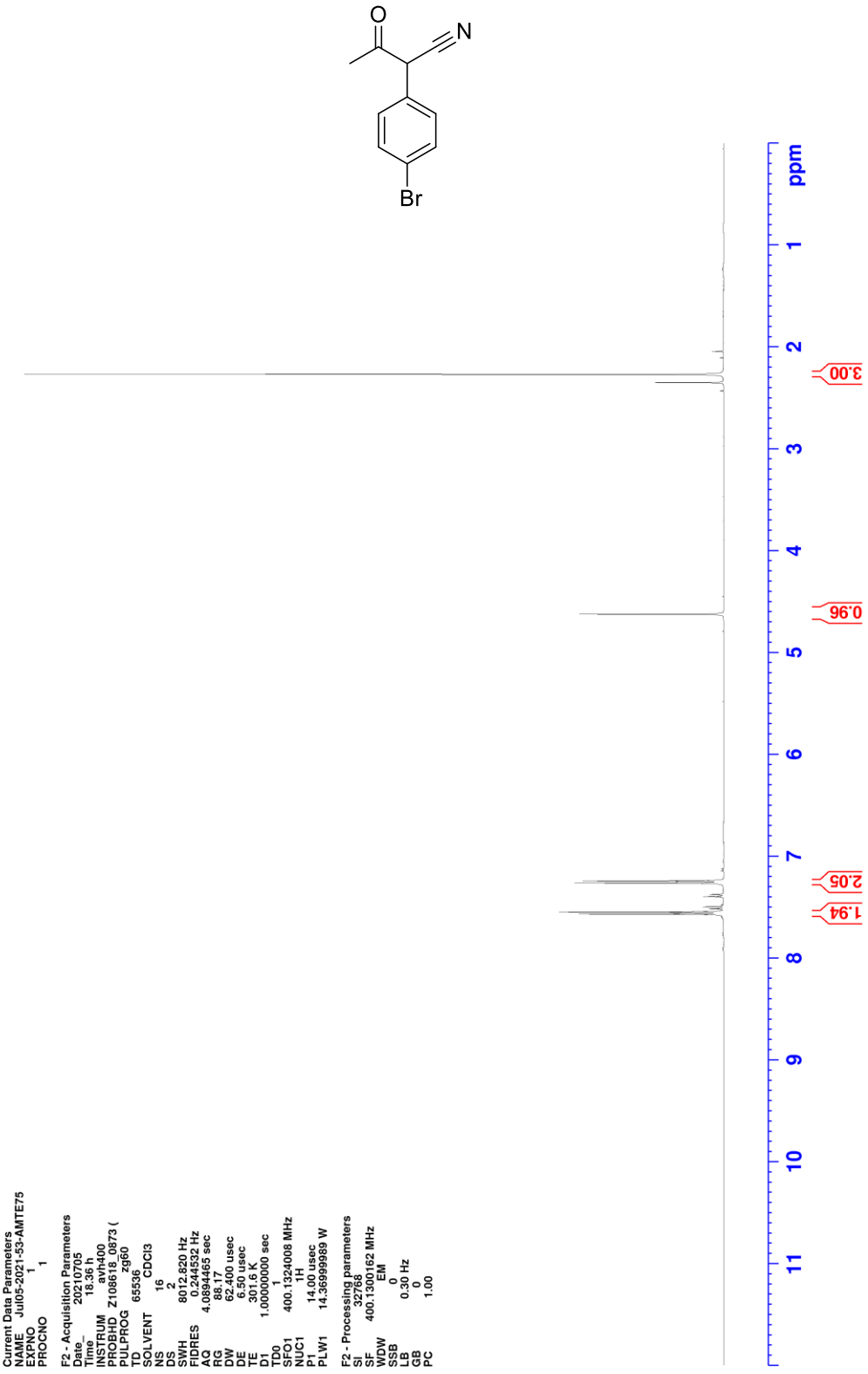


4[iii]-(Dimethylamino)-*N*-(3-methyl-4-phenylisoxazol-5-yl)butanamide

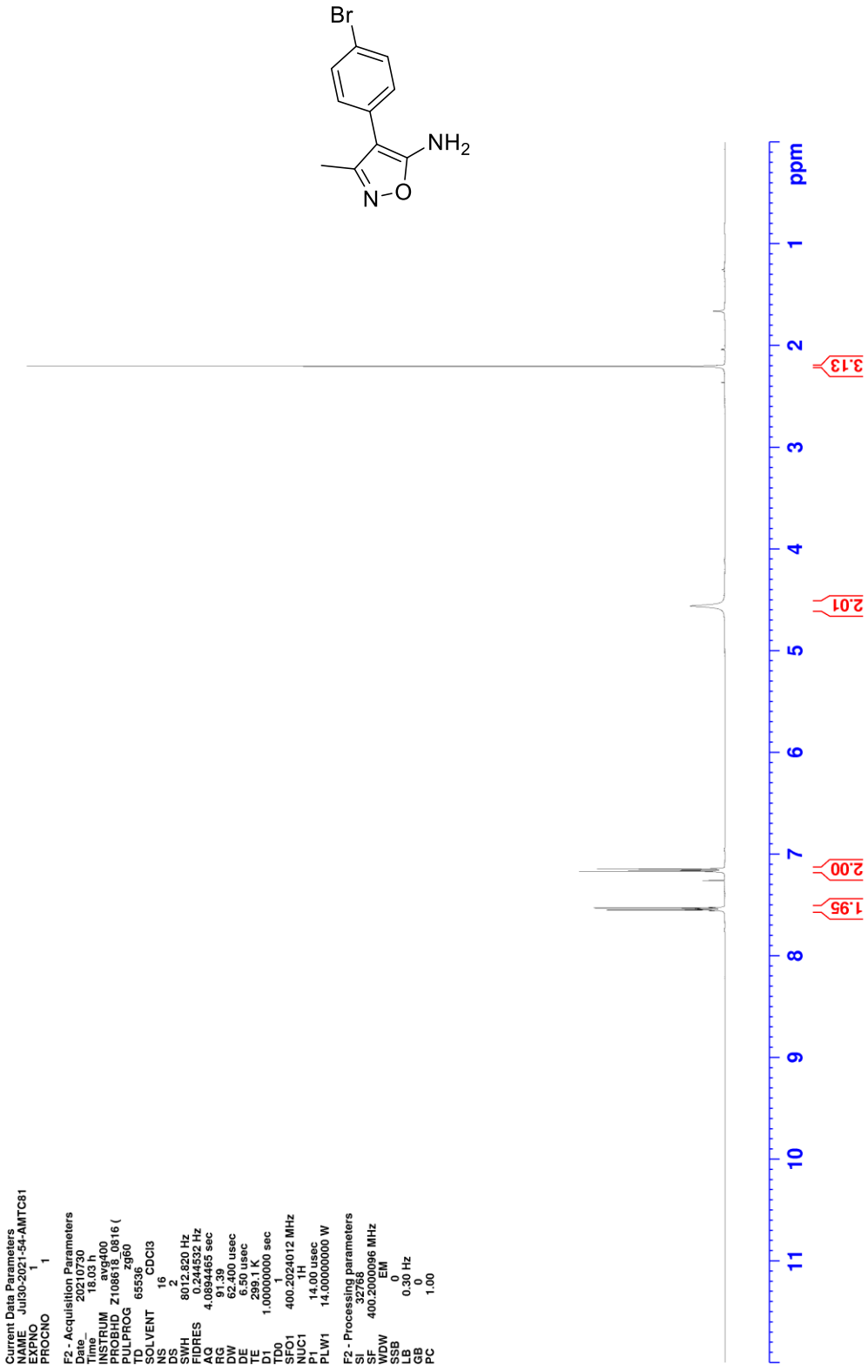


Chapter 5 spectra

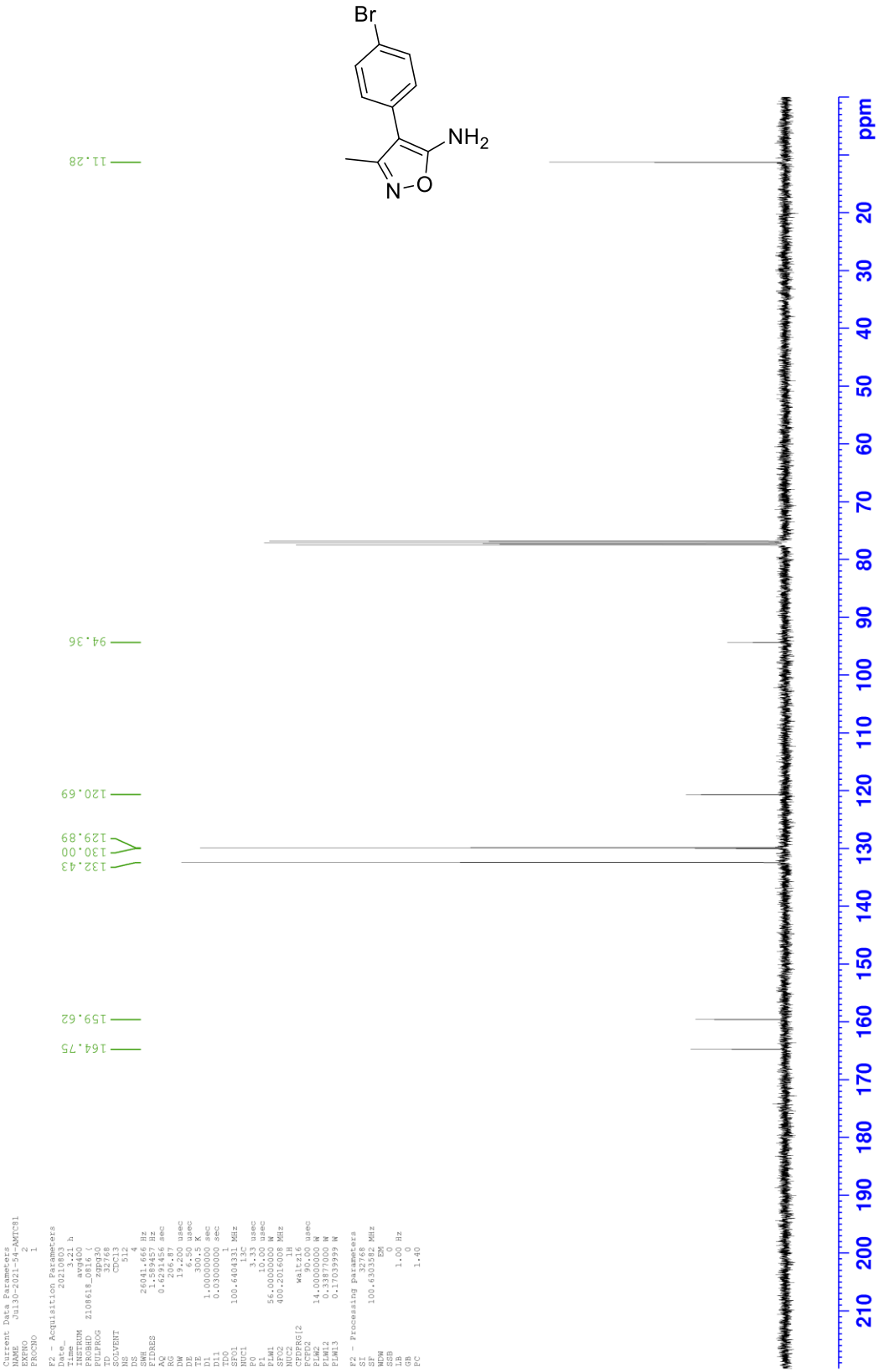
2-(4^{i}-Bromophenyl)-3-oxobutanenitrile



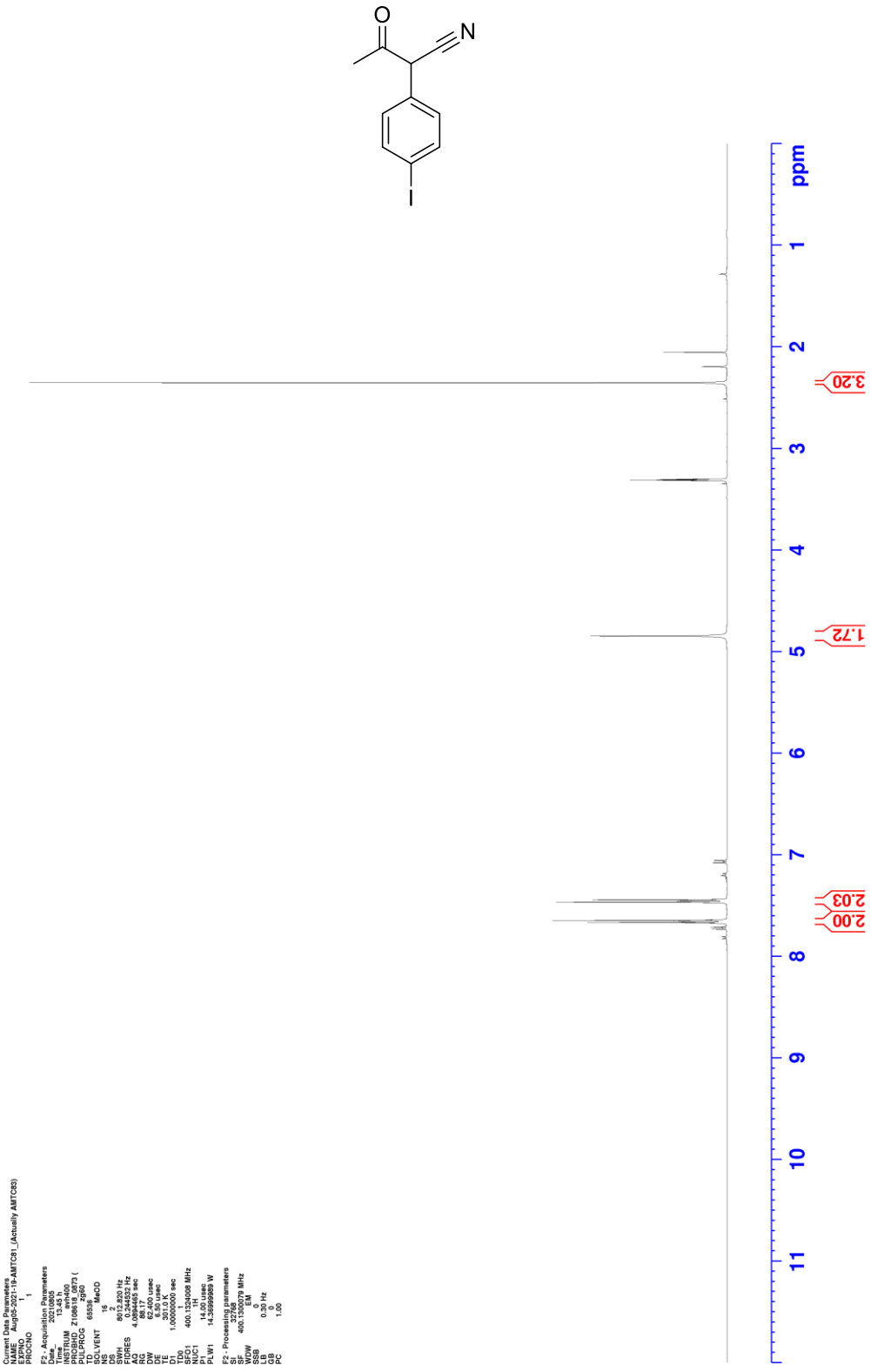
4-(4[ii]-Bromophenyl)-3-methylisooxazol-5-amine



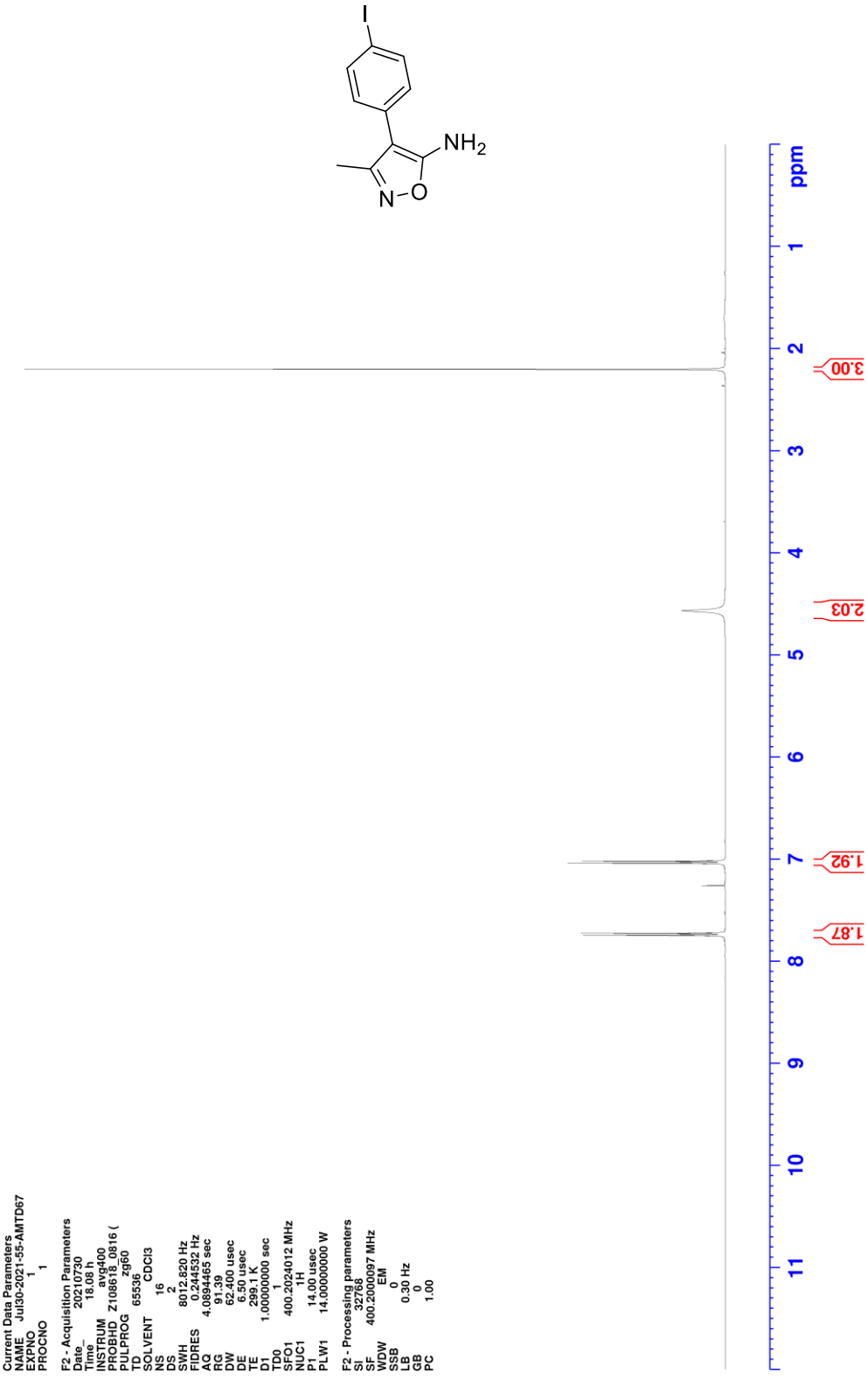
4-(4^{ii}-Bromophenyl)-3-methylisooxazol-5-amine



2-(4^{i}-Iodophenyl)-3-oxobutanenitrile



4-(4[ii]-Iodophenyl)-3-methylisoxazol-5-amine

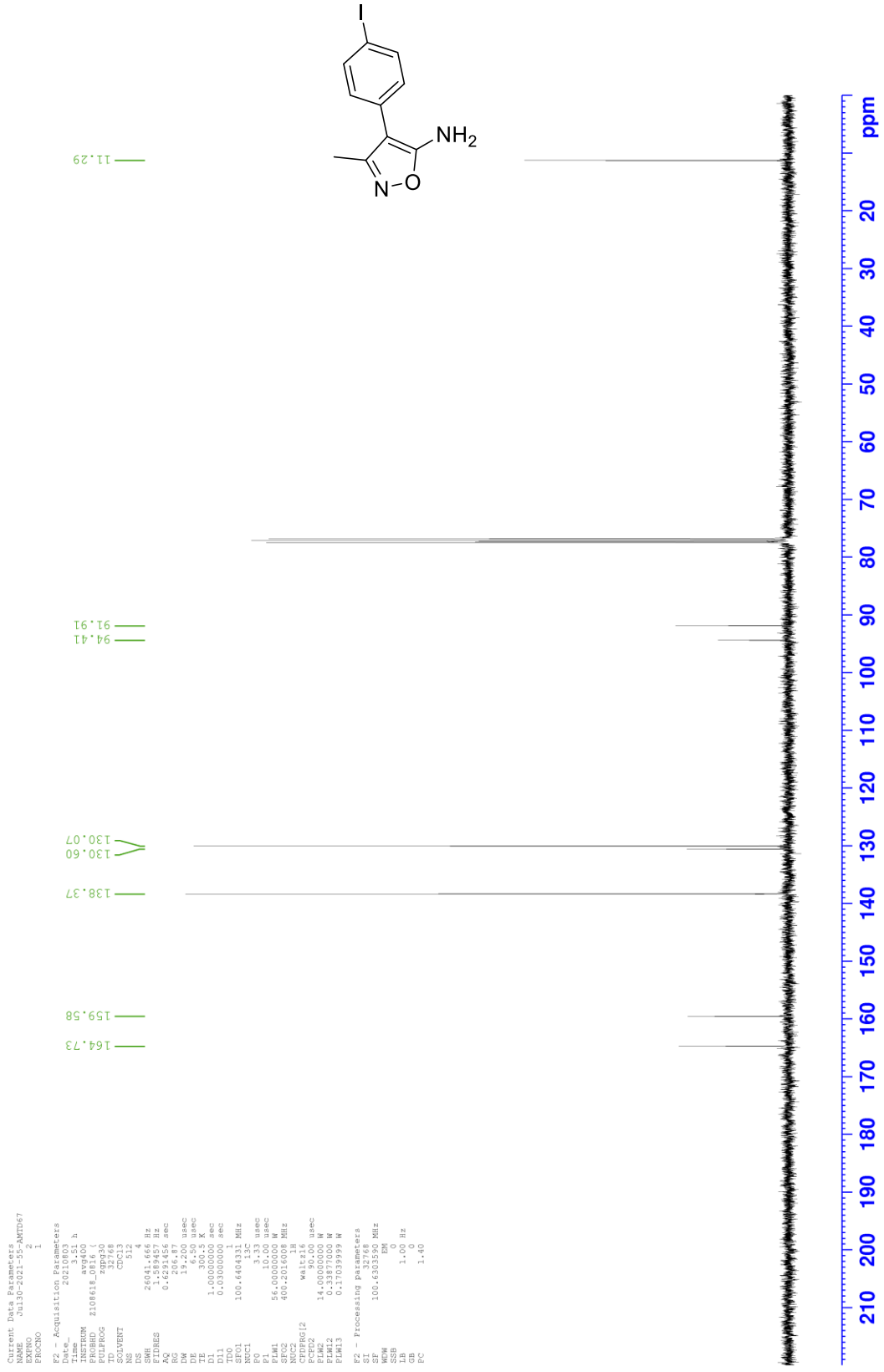


Current Data Parameters
NAME Jul30-2021-55-AMTD67
EXPNO 1
PROCNO 1

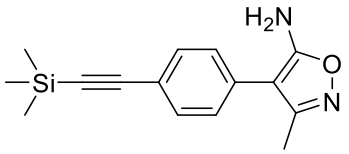
F2 - Acquisition Parameters
Date_ 20210730
Time 18.08 h
INSTRUM av400
PROBHD Z10813_0816 (
PULPROG zg60
TD 65536
SOLVENT CDCl3
NS 16
DS 2
SWH 8012.820 Hz
FIDRES 0.244532 Hz
AQ 4.0894465 sec
RG 91.39
DW 62.400 usec
DE 6.50 usec
TE 299.1 K
D1 1.00000000 sec
TD0 1
SFO1 400.2024012 MHz
NUC1 1H
P1 14.00 usec
PLW1 14.00000000 W

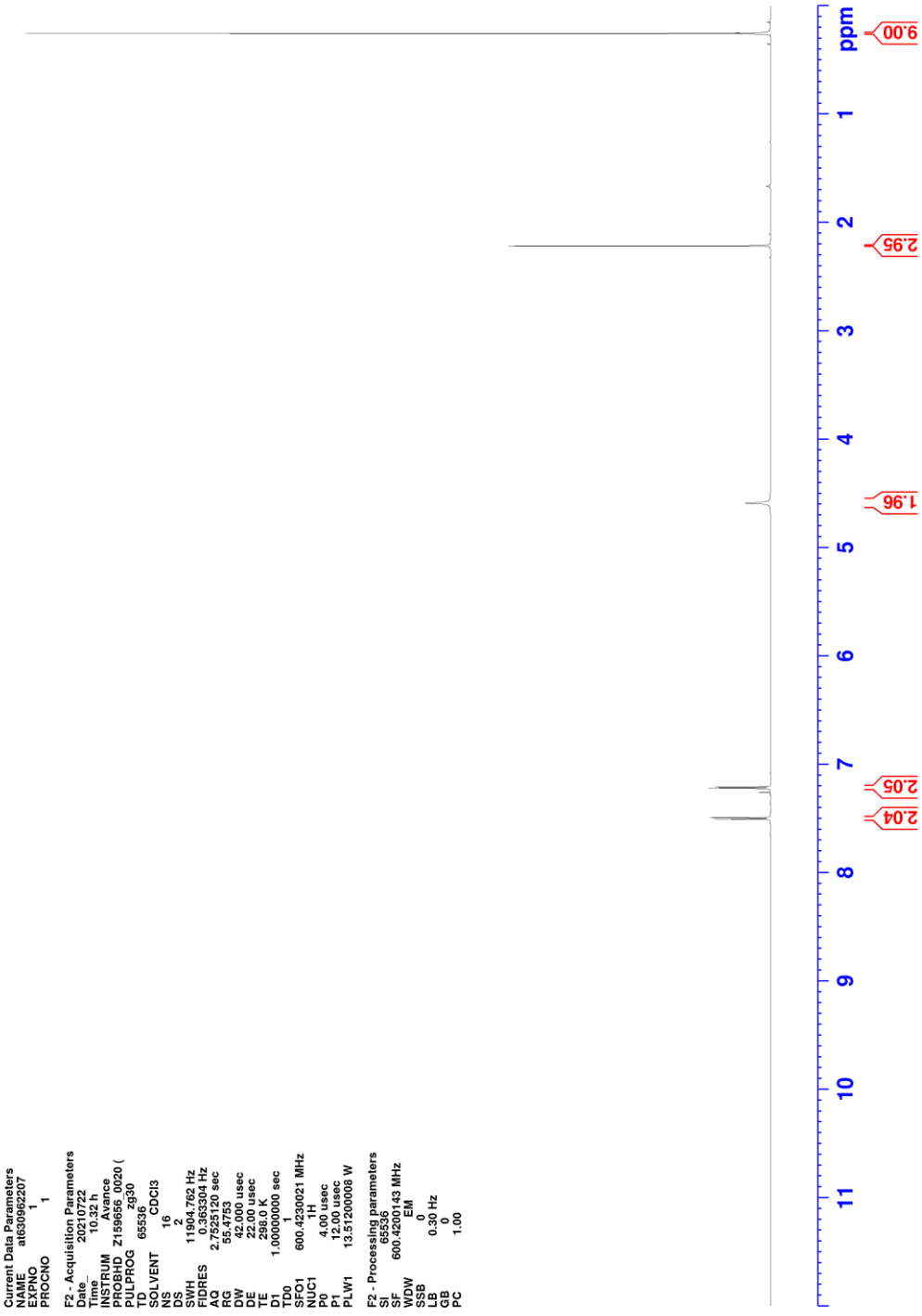
F2 - Processing parameters
SI 32768
SF 400.2000097 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

4-(4[ii]-Iodophenyl)-3-methylisoxazol-5-amine



3-Methyl-4-{4^{ii}-[(trimethylsilyl)ethynyl]phenyl}isoxazol-5-amine



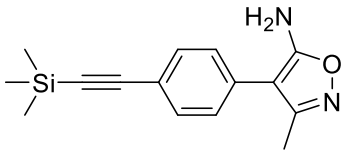


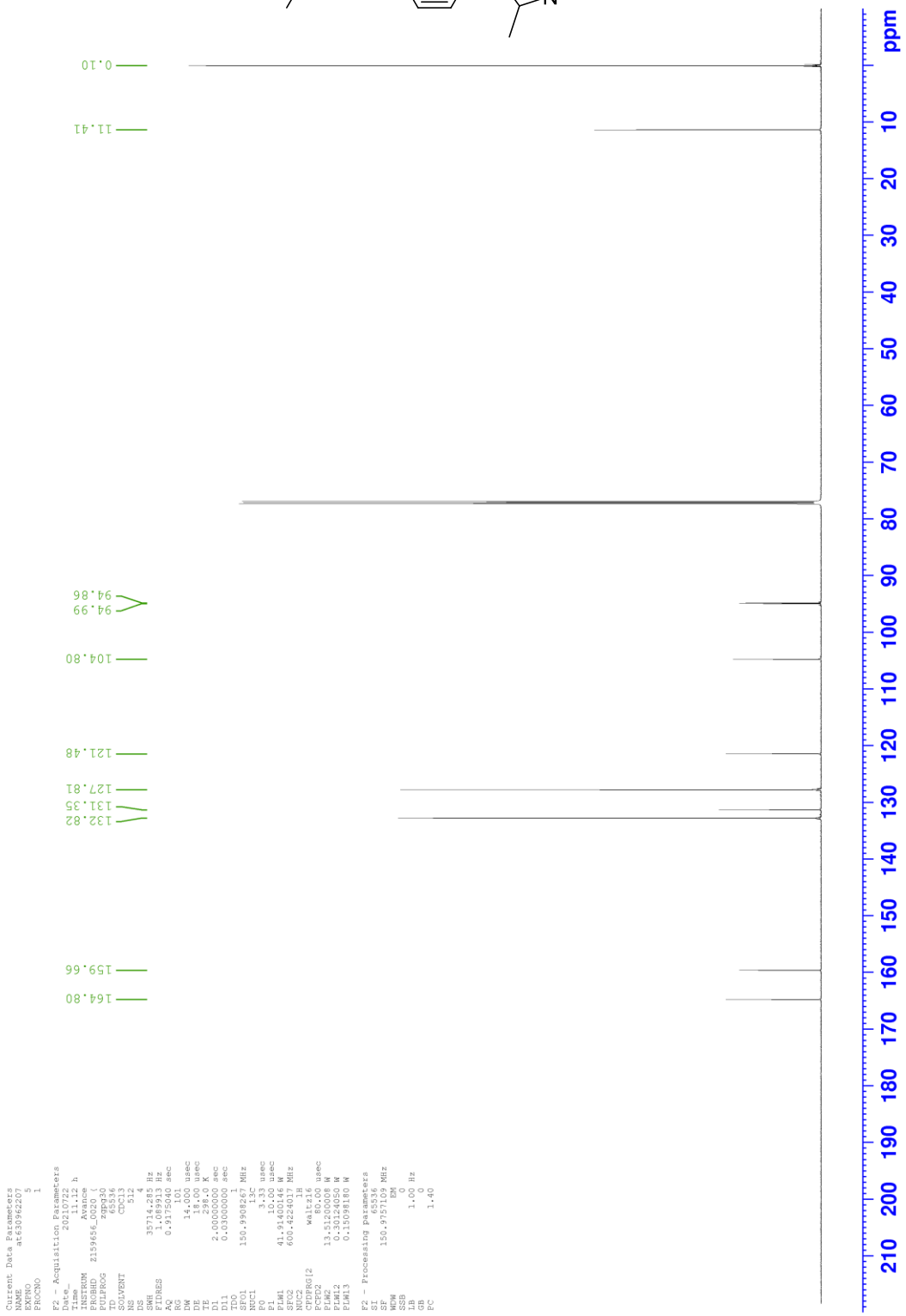

```
Current Data Parameters
NAME        ait630962207
EXPNO               1
PROCNO              1

F2 - Acquisition Parameters
Date_         20210722
Time          10.32 h
INSTRUM         Avance
PROBHD  Z159656_0020 (
PULPROG         zg30
TD             65536
SOLVENT        CDCl3
NS                16
DS                 2
SWH        11904.762 Hz
FIDRES      0.363304 Hz
AQ          2.77525120 sec
RG               55.4753
DW              42.000 usec
DE              22.00 usec
TE              298.0 K
D1          1.00000000 sec
TD0                1
SFO1    600.4230021 MHz
NUC1               1H
P0              4.00 usec
P1              12.00 usec
PLW1       13.51200008 W

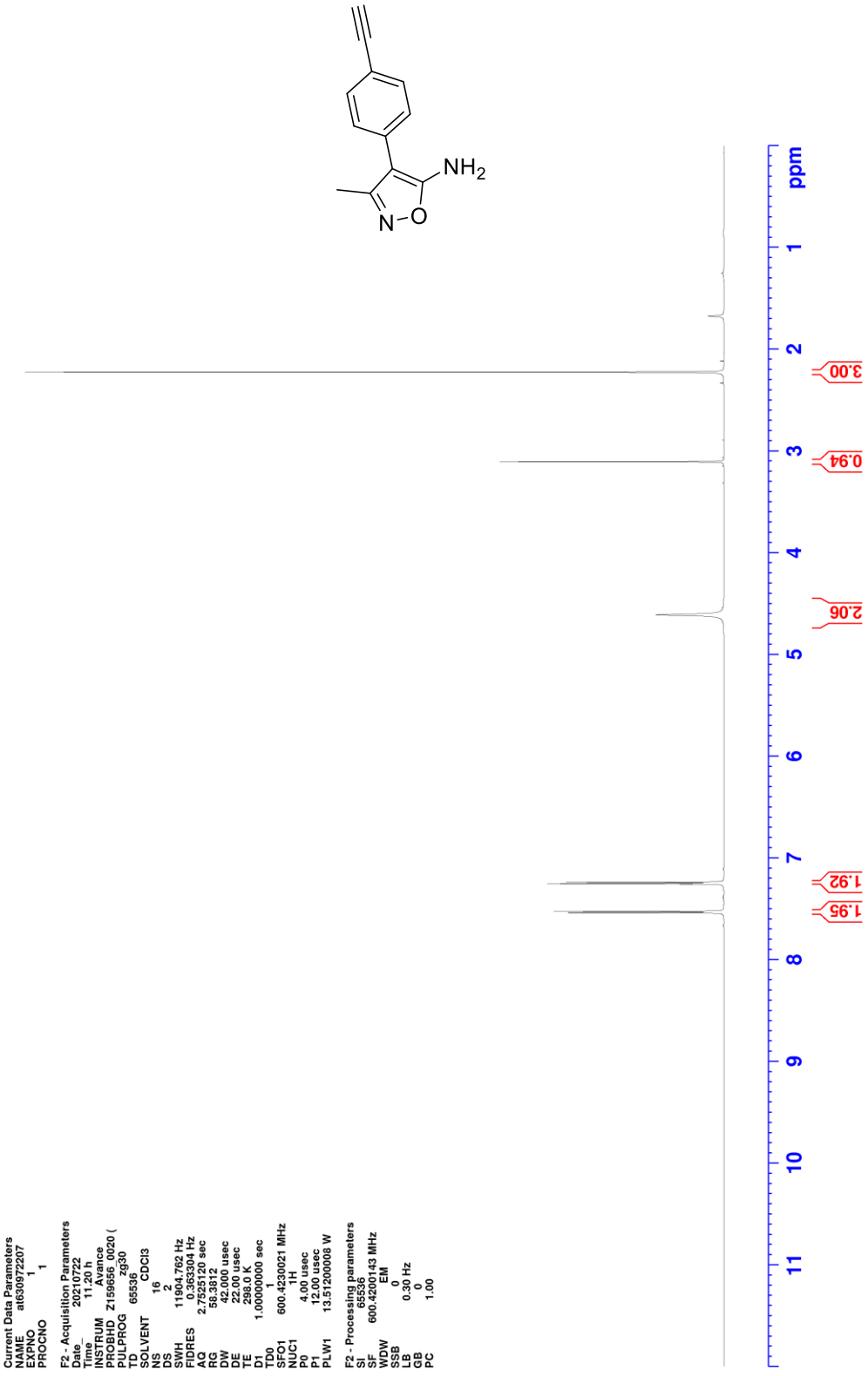
F2 - Processing parameters
SI             65536
SF      600.4200143 MHz
WDW               EM
SSB                0
LB              0.30 Hz
GB                 0
PC              1.00
```

3-Methyl-4-{4^{ii}-[(trimethylsilyl)ethynyl]phenyl}isoxazol-5-amine





4-(4[ii]-Ethynylphenyl)-3-methylisoxazol-5-amine

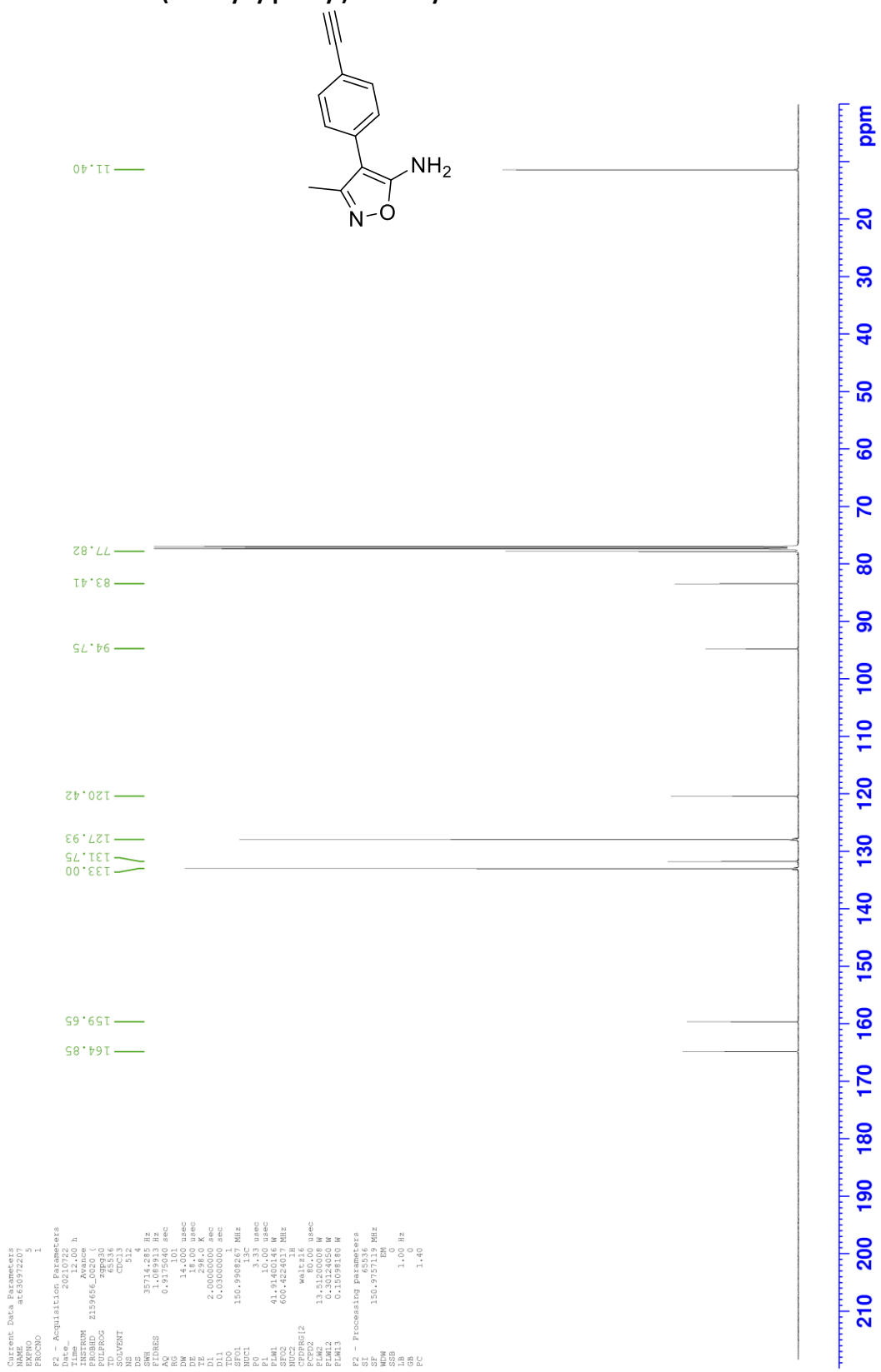


Current Data Parameters
NAME at630972207
EXPNO 1
PROCNO 1

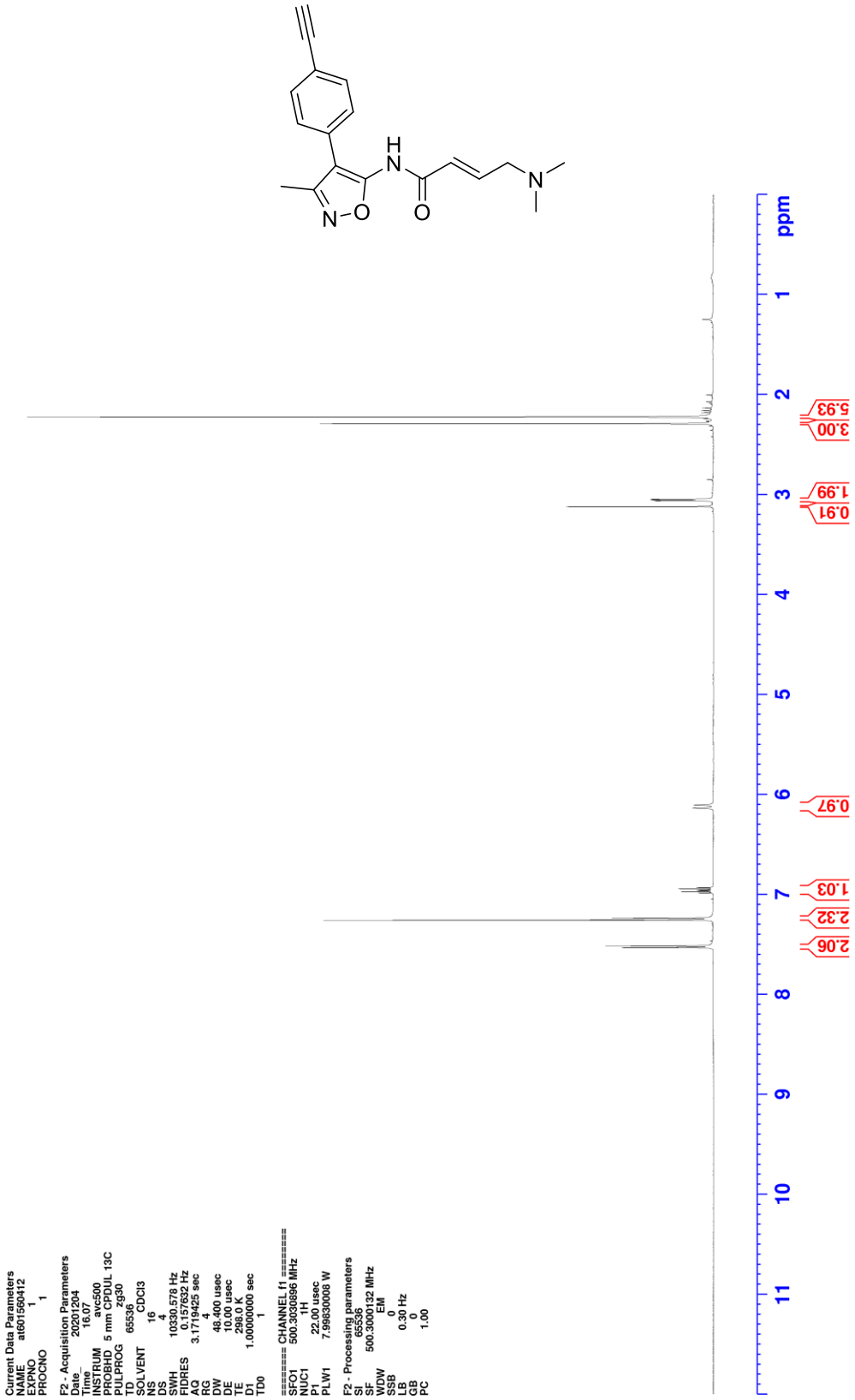
F2 - Acquisition Parameters
Date_ 20210722
Time_ 11.20 h
INSTRUM Avance
PROBHD Z159656_0020 (
PULPROG zg30
TD 65536
SOLVENT CDCl3
NS 16
DS 2
SWH 11904.762 Hz
FIDRES 0.363304 Hz
AQ 2.7525120 sec
RG 58.3812
DW 42.000 usec
DE 22.00 usec
TE 298.0 K
D1 1.00000000 sec
TD0 1
SFO1 600.4230021 MHz
NUC1 1H
P0 4.00 usec
P1 12.00 usec
PLW1 13.51200008 W

F2 - Processing parameters
SI 65536
SF 600.4200143 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

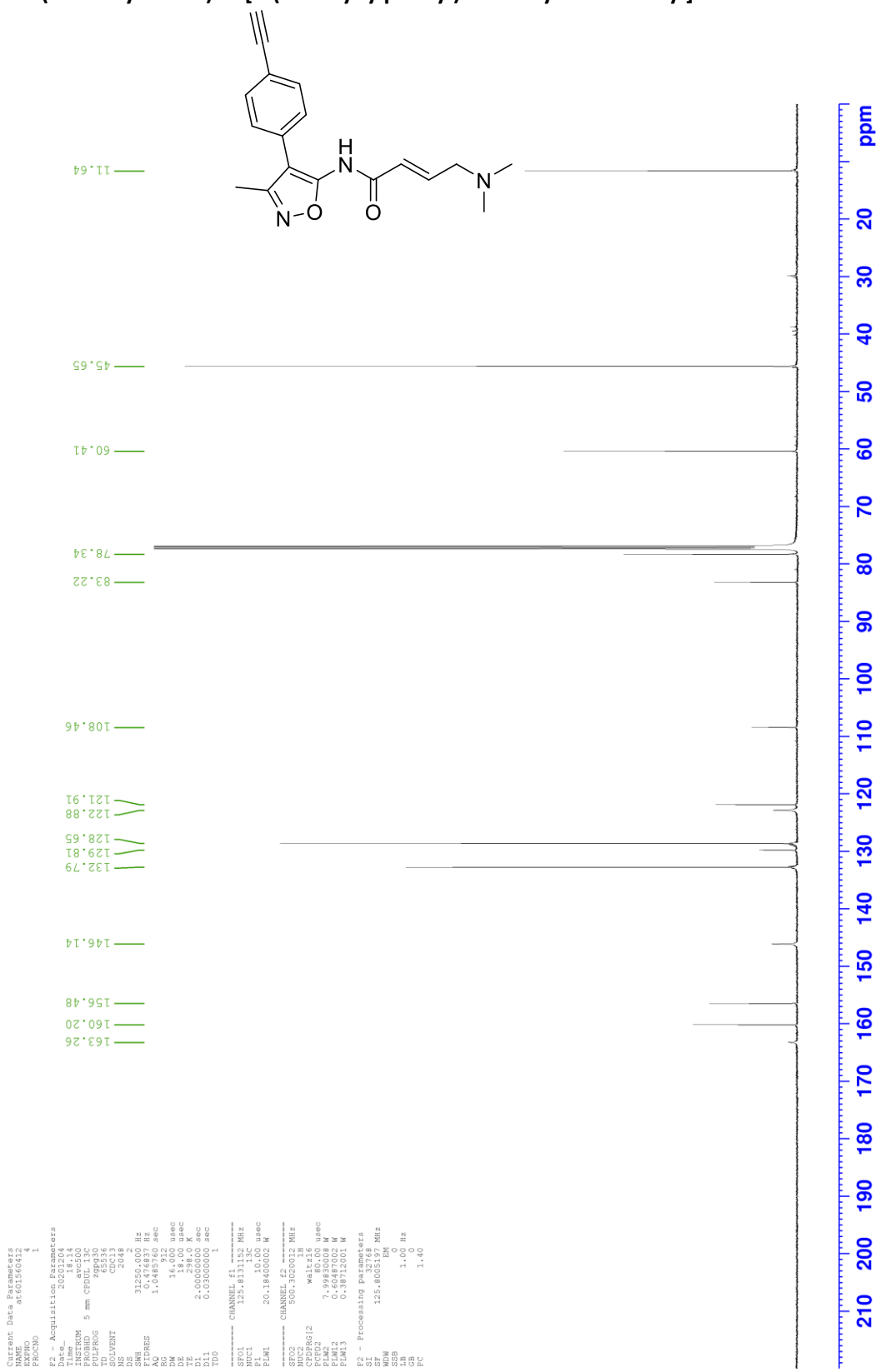
4-(4^{ii}-Ethynylphenyl)-3-methylisoxazol-5-amine



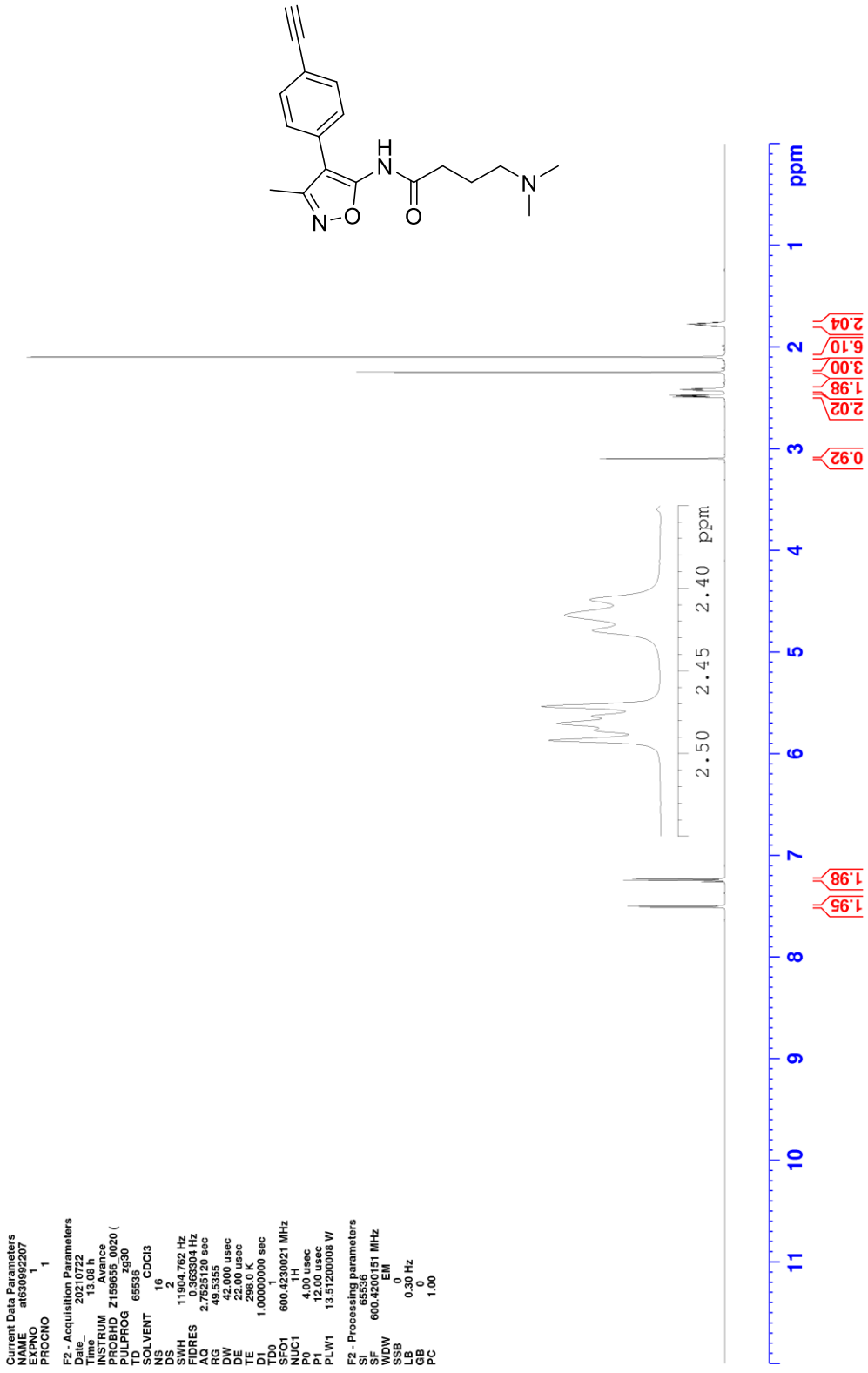
(*E*)-4[iii]-(Dimethylamino)-*N*-[4-(4[ii]-ethynylphenyl)-3-methylisoxazol-5-yl]but-2[iii]-enamide



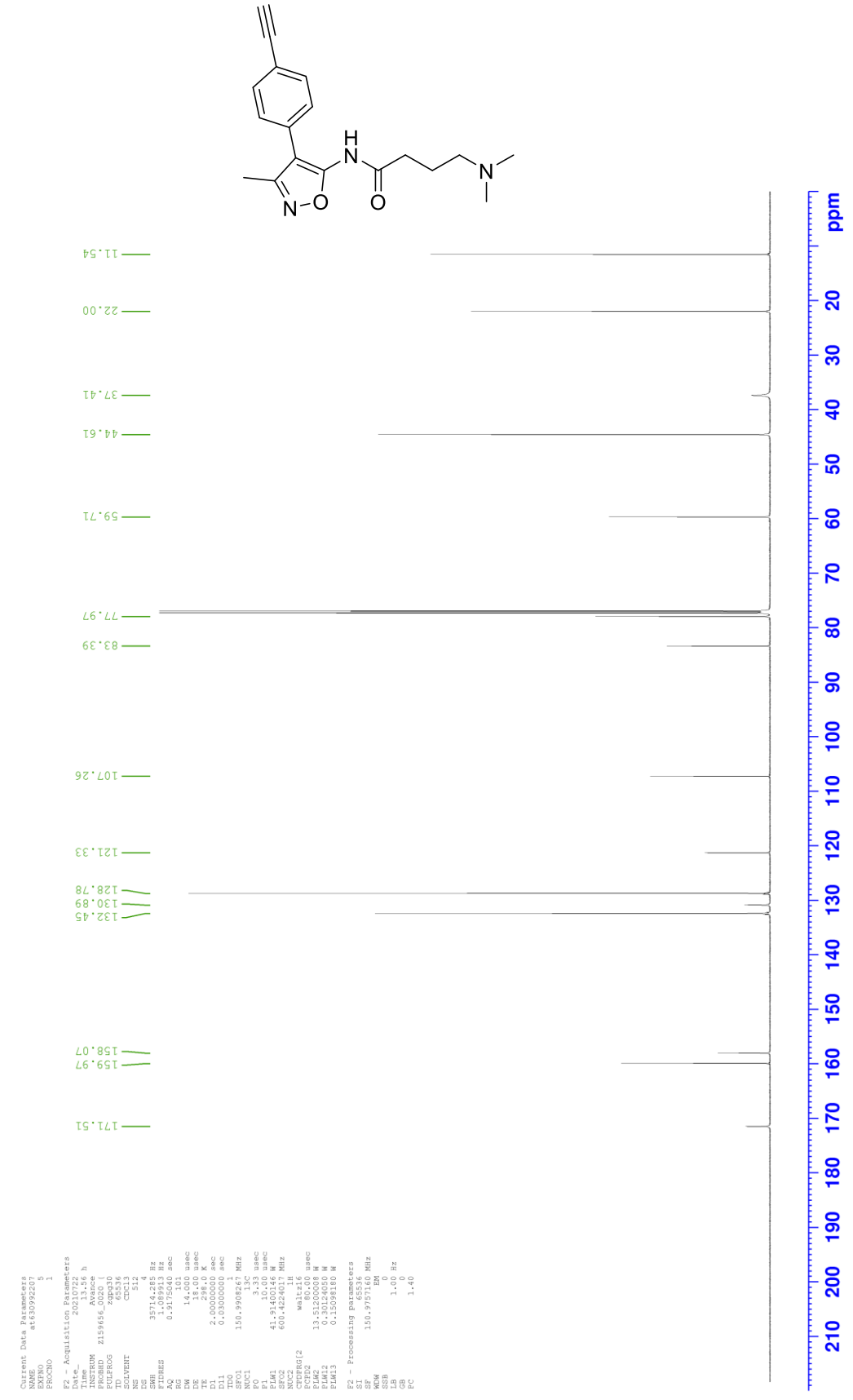
(*E*)-4[iii]-(Dimethylamino)-*N*-[4-(4[ii]-ethynylphenyl)-3-methylisoxazol-5-yl]but-2[iii]-enamide



4[iii]-(Dimethylamino)-*N*-[4-(4[ii]-ethynylphenyl)-3-methylisoxazol-5-yl]butanamide

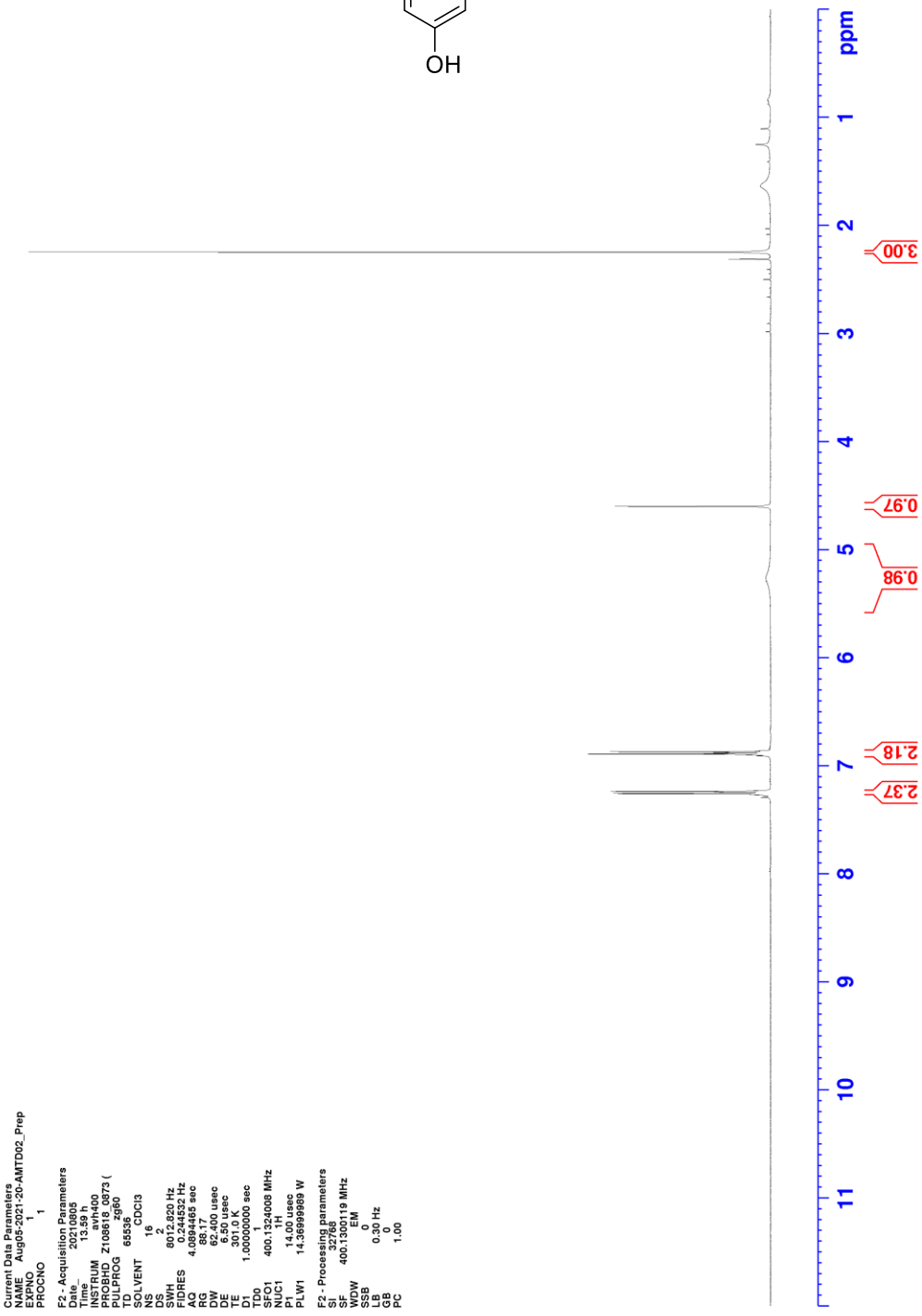


4^{iii}-(Dimethylamino)-*N*-[4-(4^{ii}-ethynylphenyl)-3-methylisoxazol-5-yl]butanamide

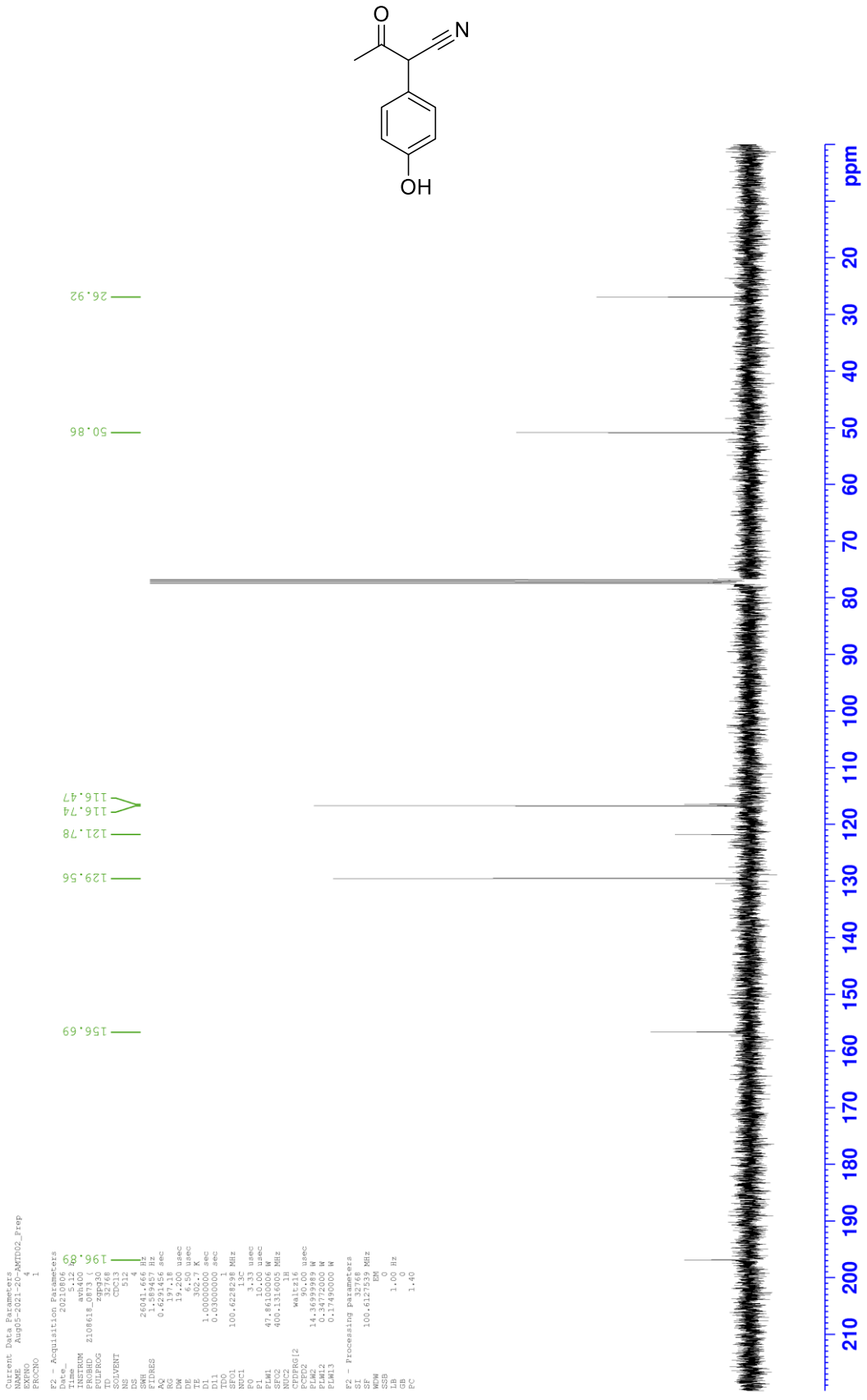


2-(4^{i}-Hydroxyphenyl)-3-oxobutanenitrile

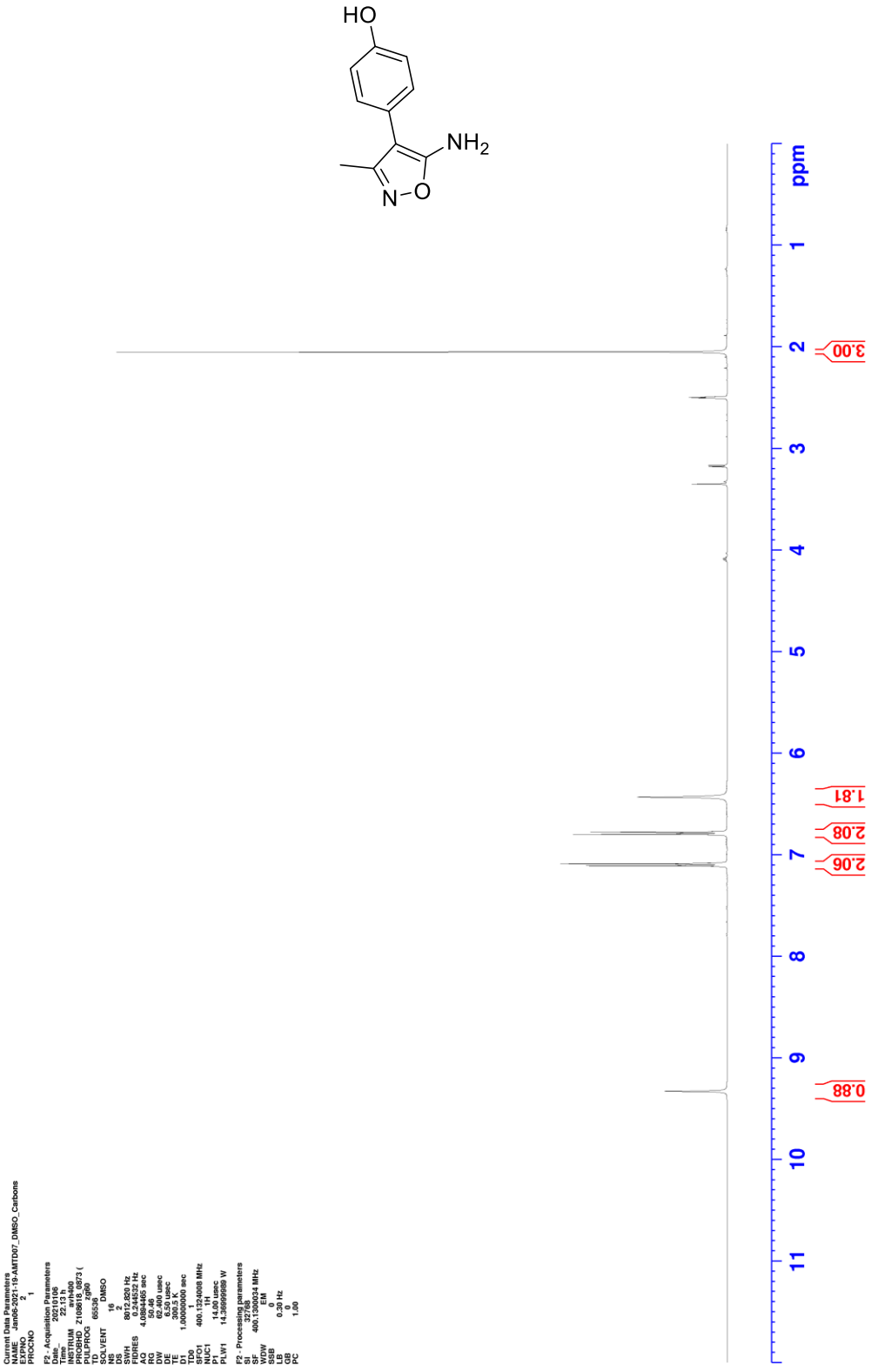




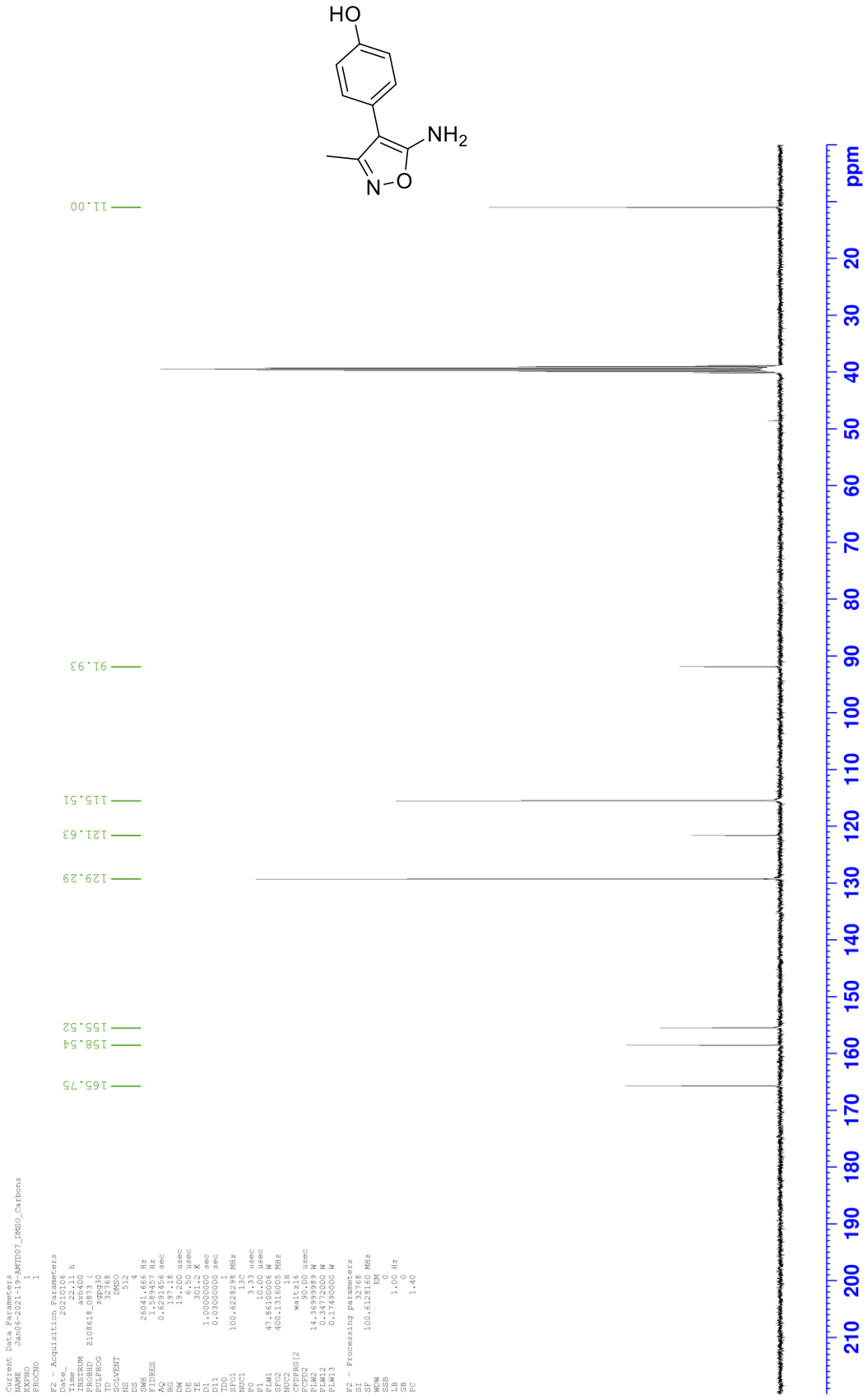
2-(4^{i}-Hydroxyphenyl)-3-oxobutanenitrile



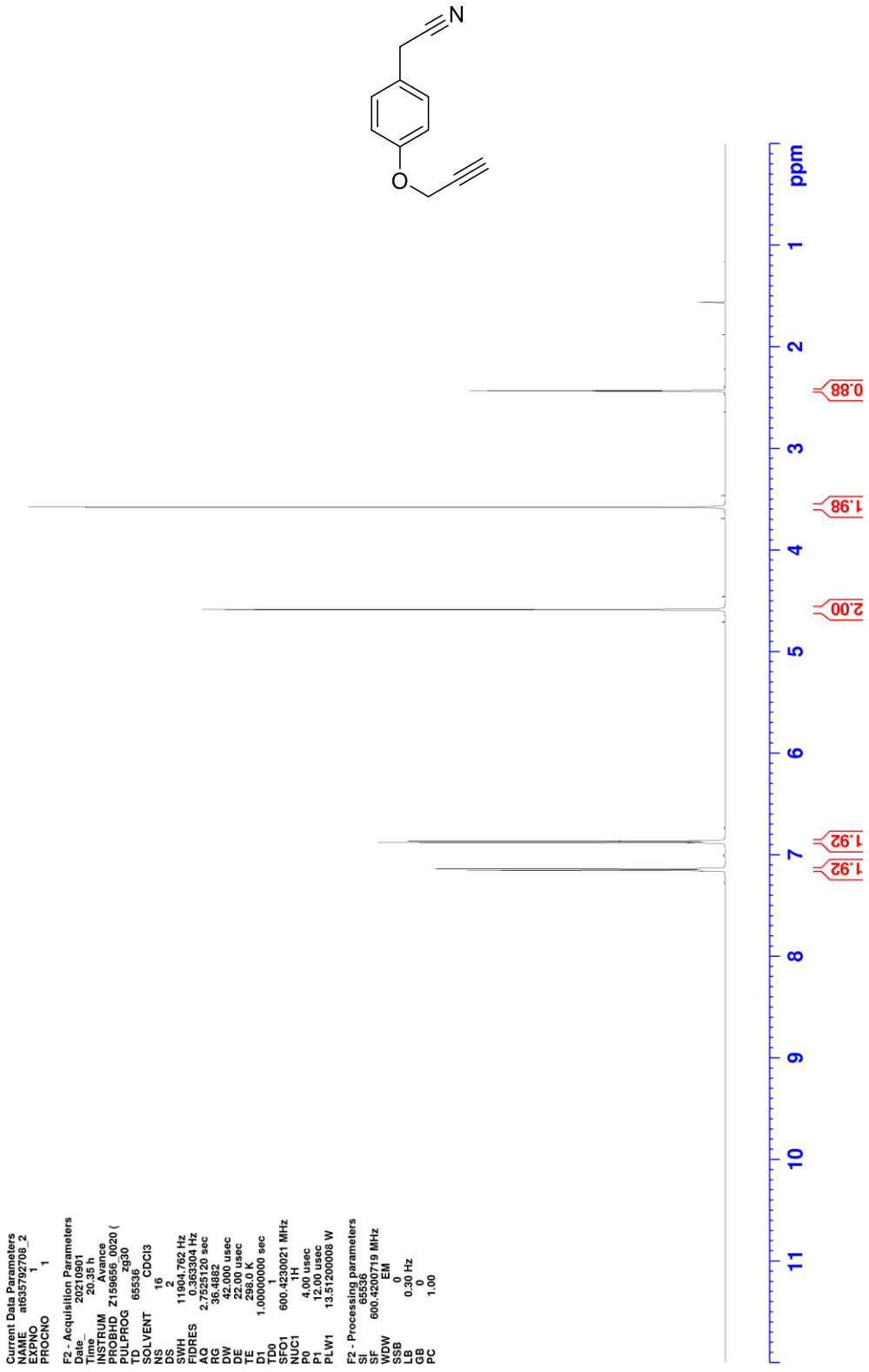
4[ii]-(5-Amino-3-methylisoxazol-4-yl)phenol



4[ii]-(5-Amino-3-methylisoxazol-4-yl)phenol



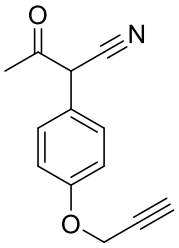
2-(4[i]-Prop-2[ii]-ynoxyphenyl)acetonitrile

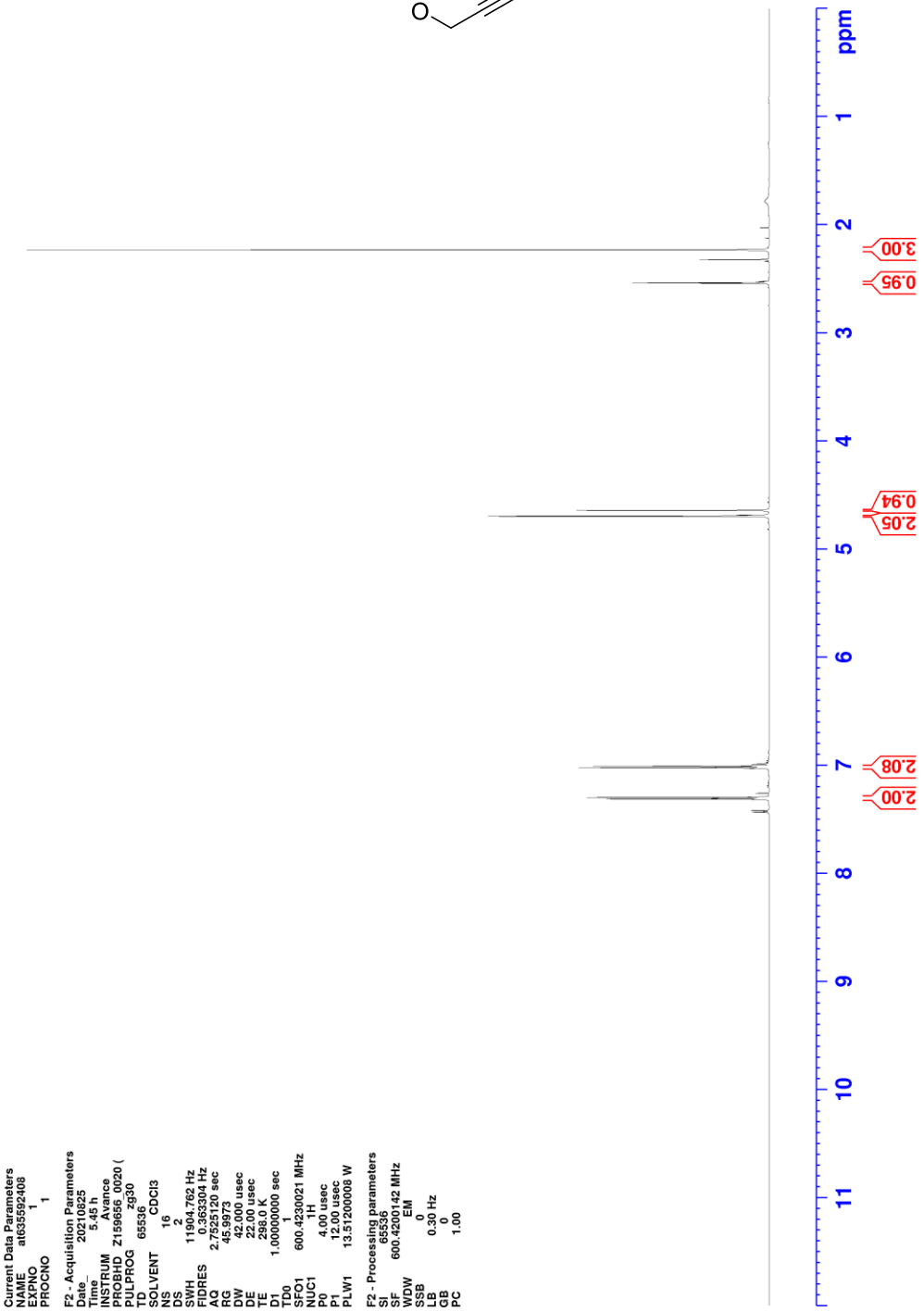


2-(4[i]-Prop-2[ii]-ynoxyphenyl)acetonitrile

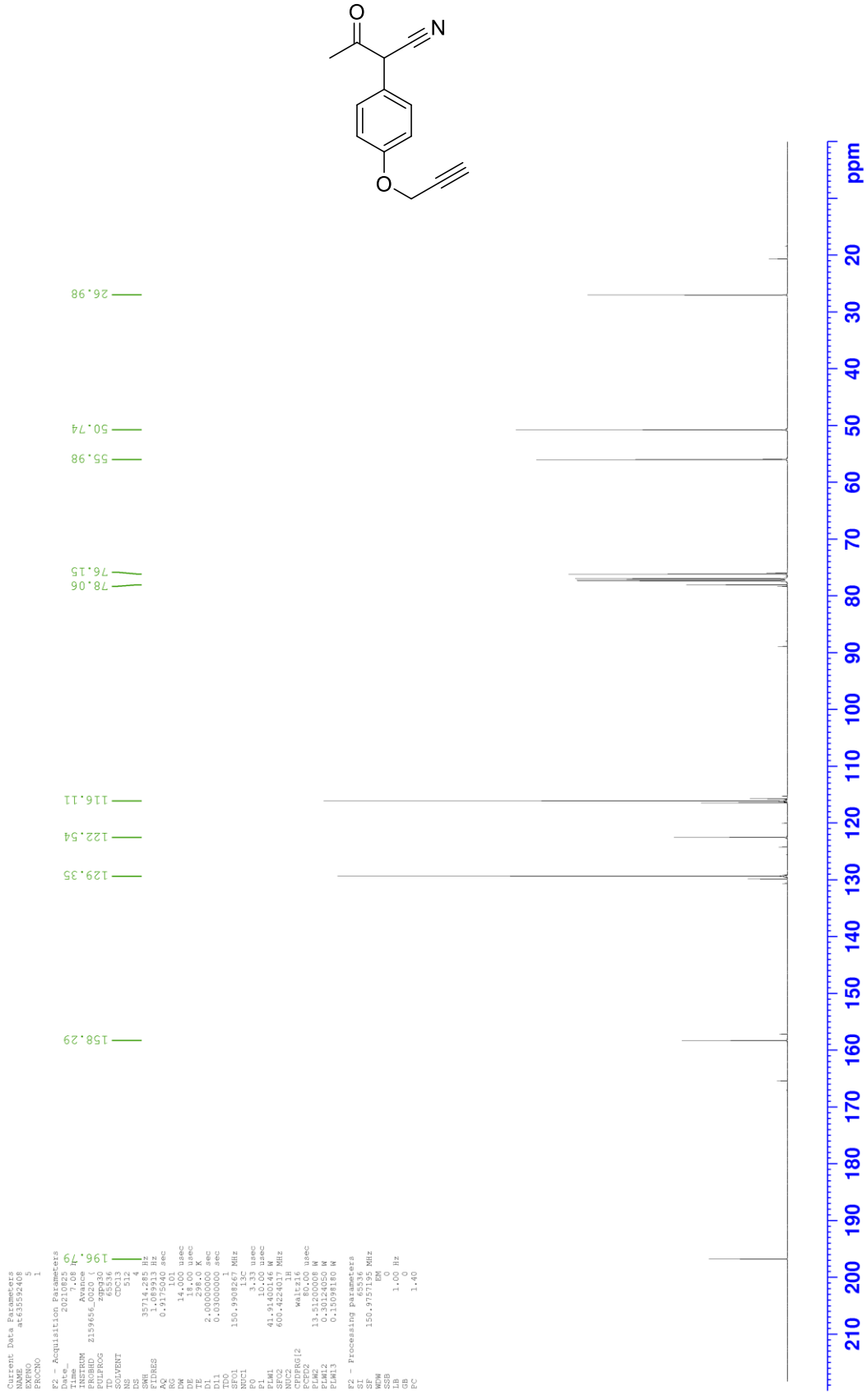


3-Oxo-2-(4[i]-prop-2[ii]-ynoxyphenyl)butanenitrile

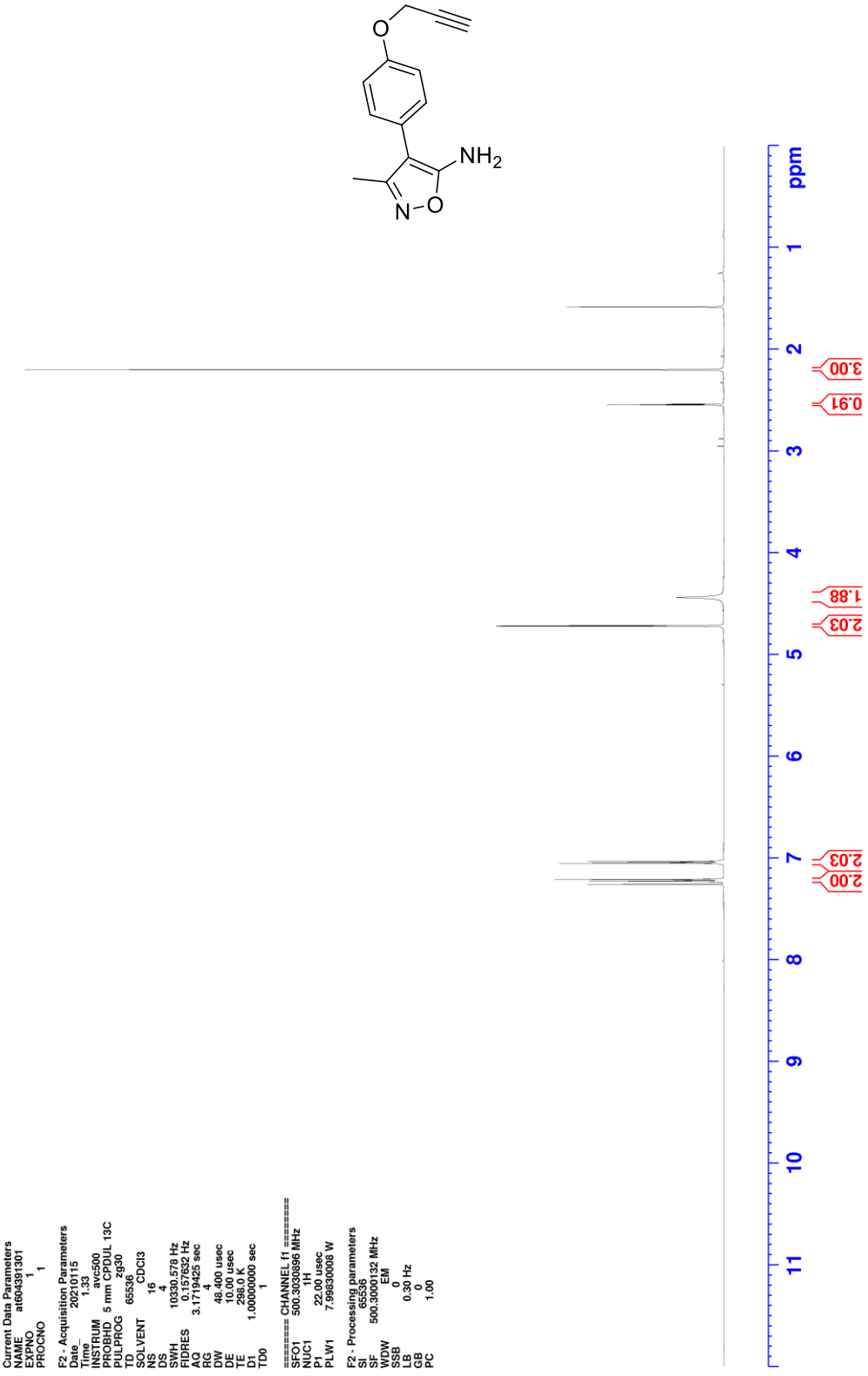




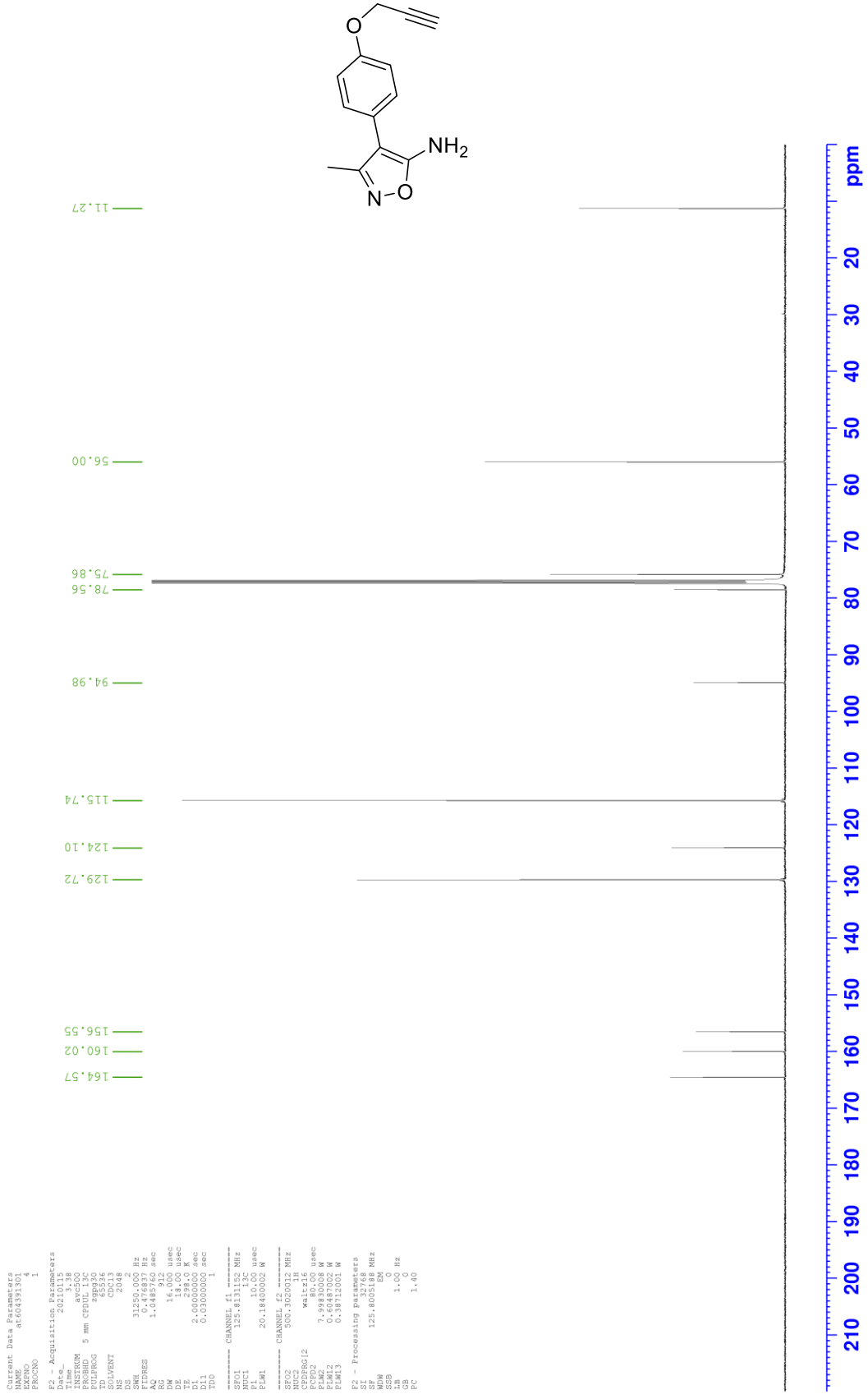
3-Oxo-2-(4[i]-prop-2[ii]-ynoxyphenyl)butanenitrile



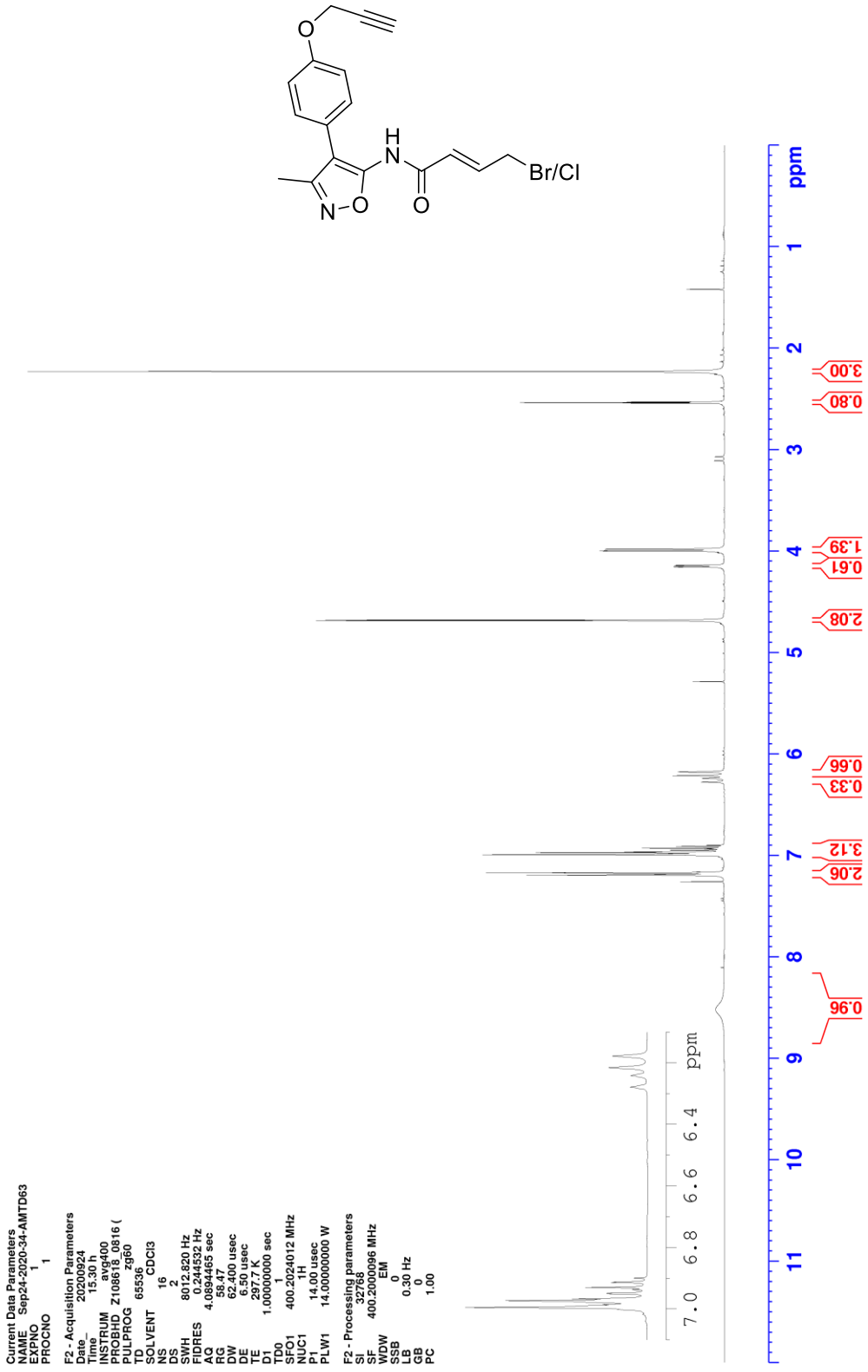
3-Methyl-4-(4[ii]-prop-2[iii]-ynoxyphenyl)-isoxazol-5-amine



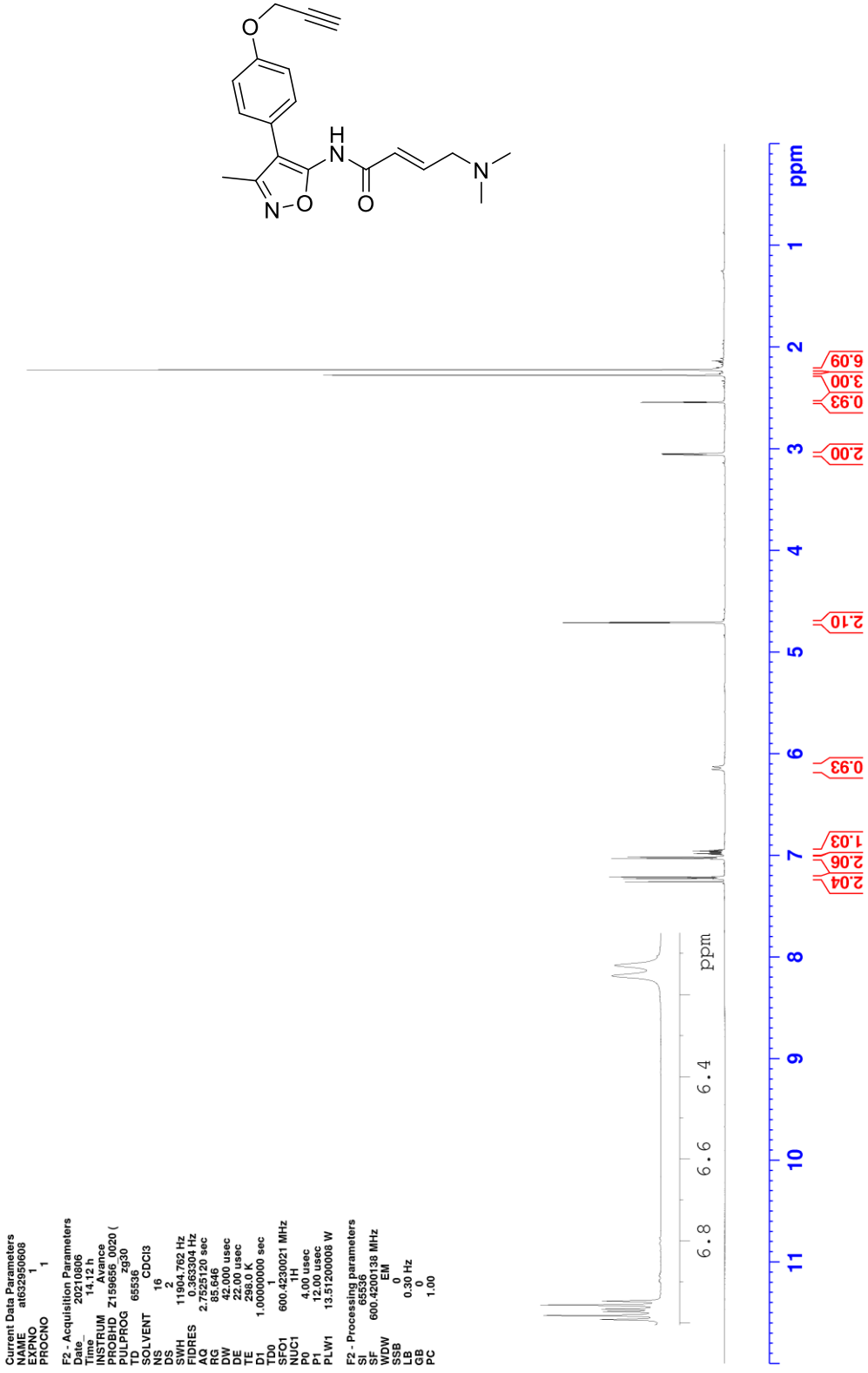
3-Methyl-4-(4[ii]-prop-2[iii]-ynoxyphenyl)-isoxazol-5-amine



(*E*)-4[iv]-Bromo/Chloro-*N*-[3-methyl-4-(4[ii]-prop-2[iii]-ynoxyphenyl)-isoxazol-5-yl]but-2[iv]-enamide



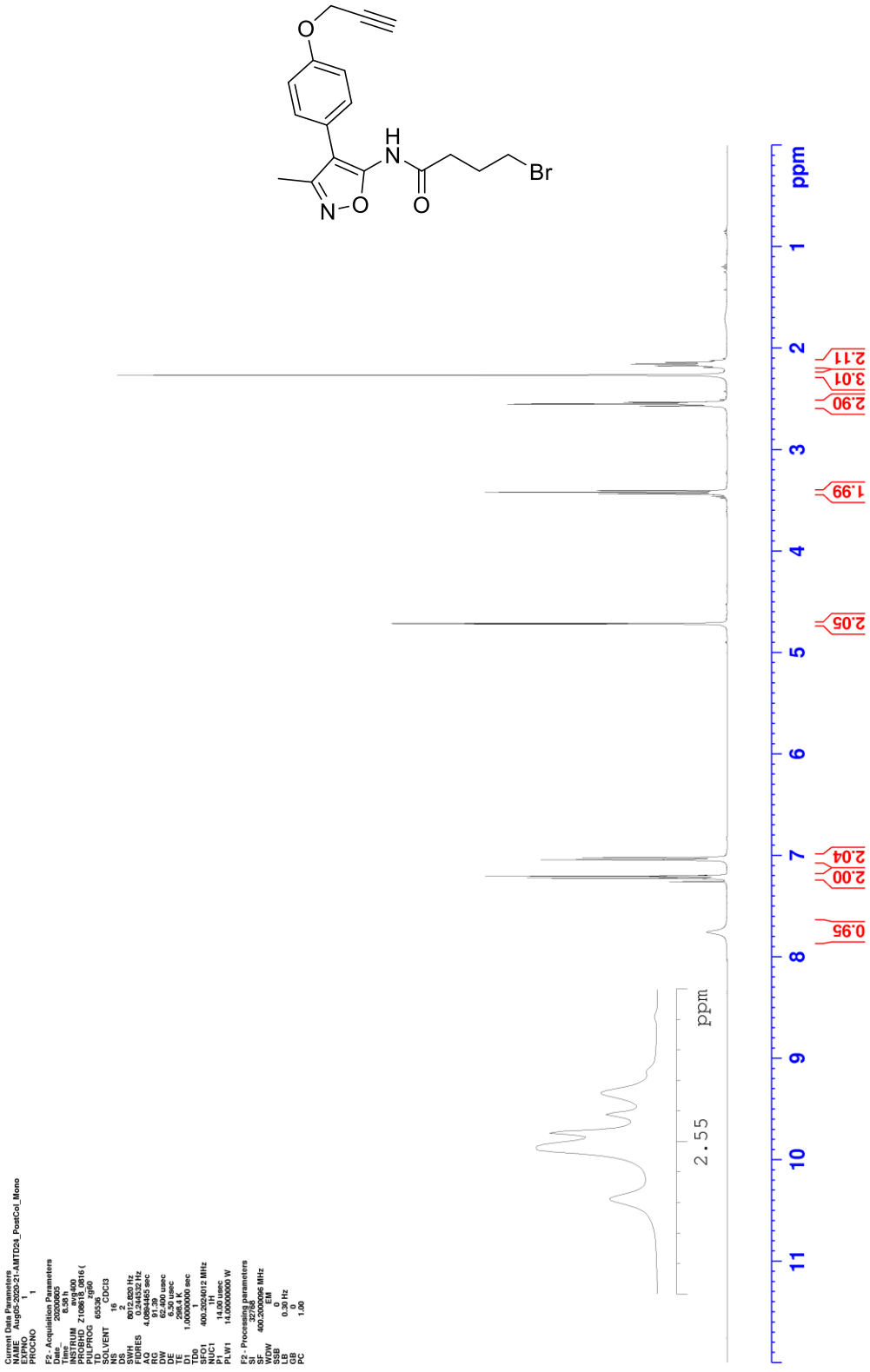
(*E*)-4^{iv}-(Dimethylamino)-*N*-[3-methyl-4-(4^{ii}-prop-2^{iii}-ynoxyphenyl)-isoxazol-5-yl]but-2^{iv}-enamide



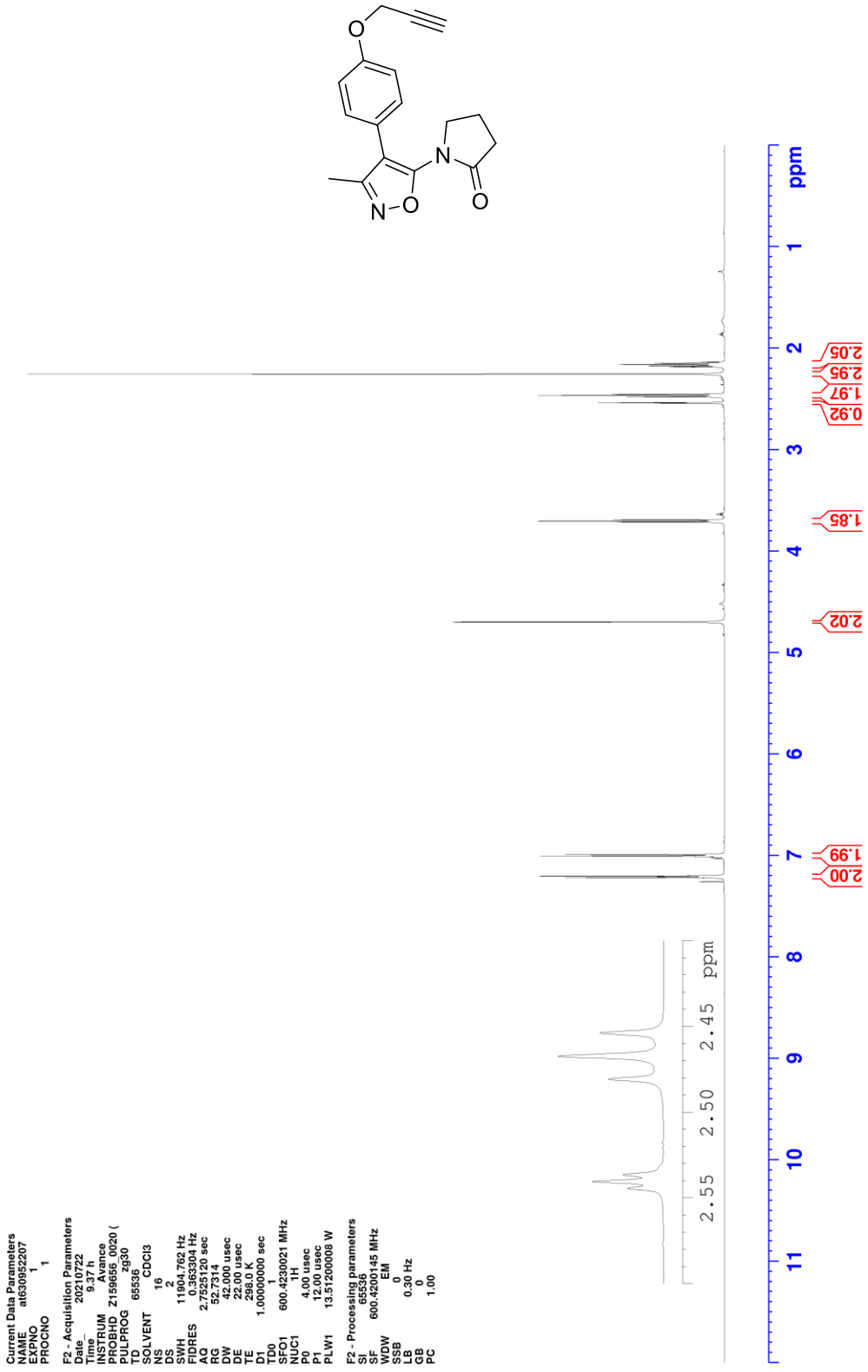
(*E*)-4^{iv}-(Dimethylamino)-*N*-[3-methyl-4-(4^{ii}-prop-2^{iii}-ynoxyphenyl)-isoxazol-5-yl]but-2^{iv}-enamide



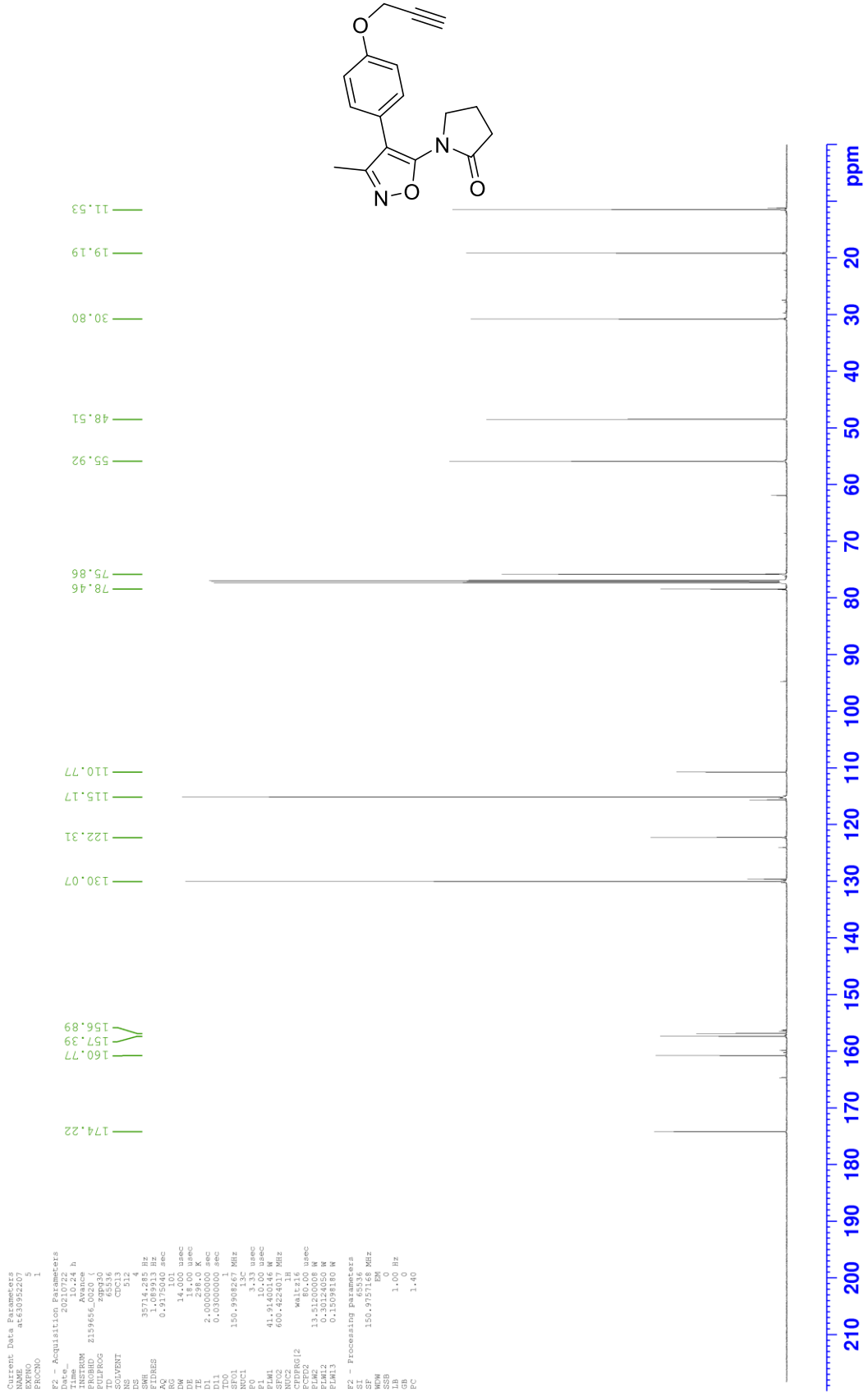
4[iv]-Bromo-*N*-[3-methyl-4-(4[ii]-prop-2[iii]-ynoxyphenyl)-isoxazol-5-yl]butanamide



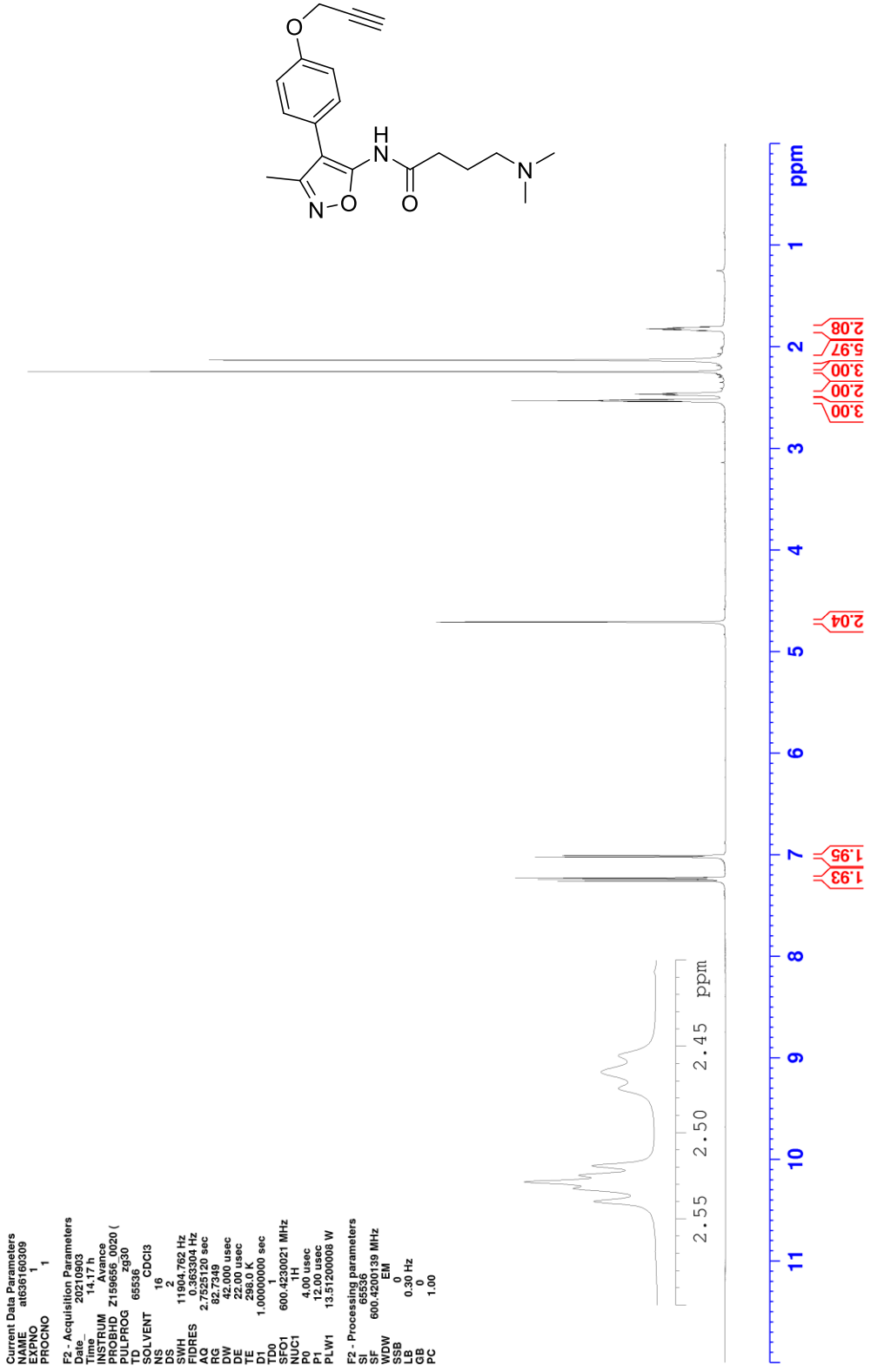
1[iv]-[3-Methyl-4-(4-prop-2[iii]-ynoxyphenyl)-isoxazol-5-yl]pyrrolidin-2[iv]-one



1[iv]-[3-Methyl-4-(4-prop-2[iii]-ynoxyphenyl)-isoxazol-5-yl]pyrrolidin-2[iv]-one



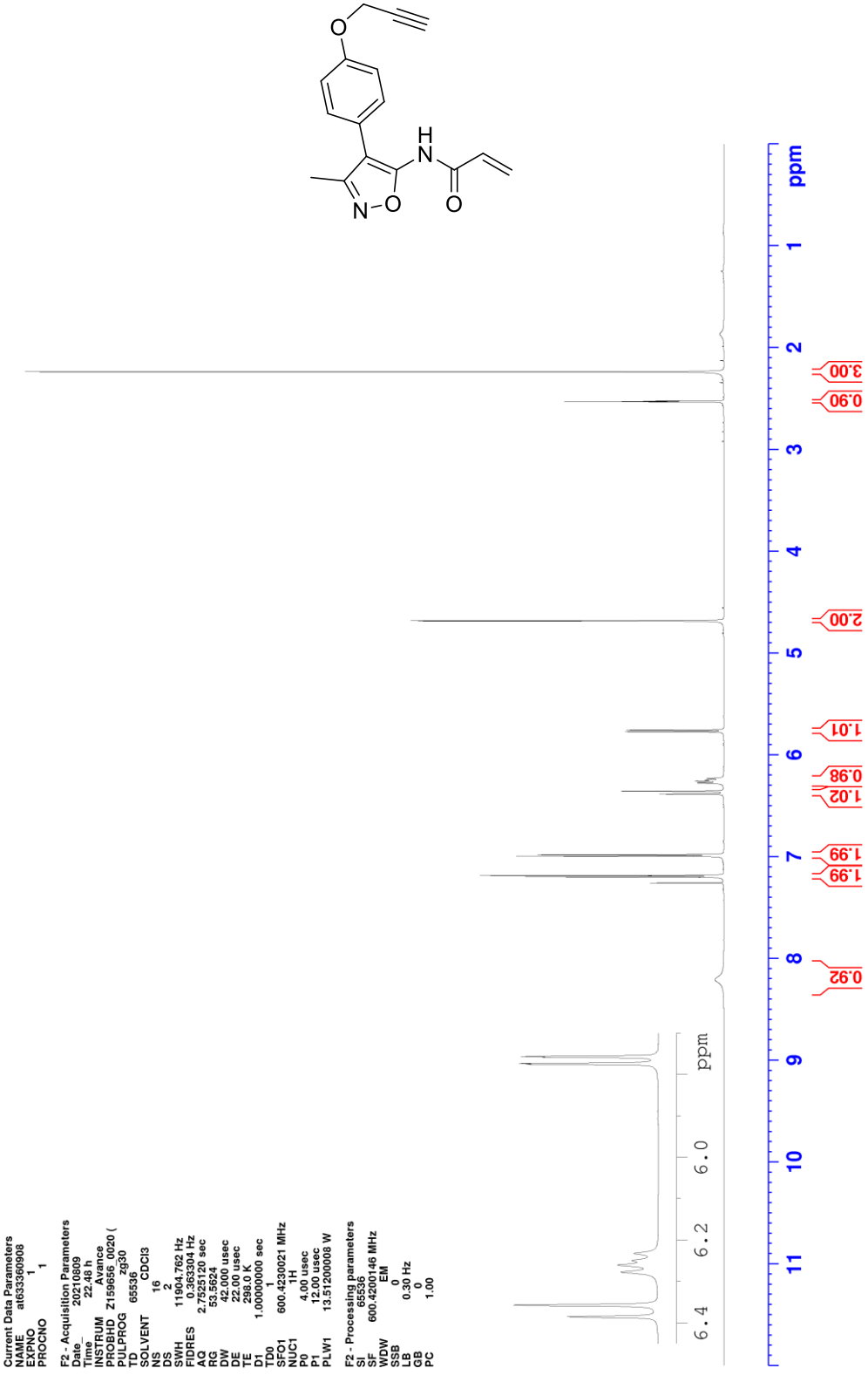
4[iv]-(Dimethylamino)-*N*-[3-methyl-4-(4[ii]-prop-2[iii]-ynoxyphenyl)-isoxazol-5-yl]butanamide



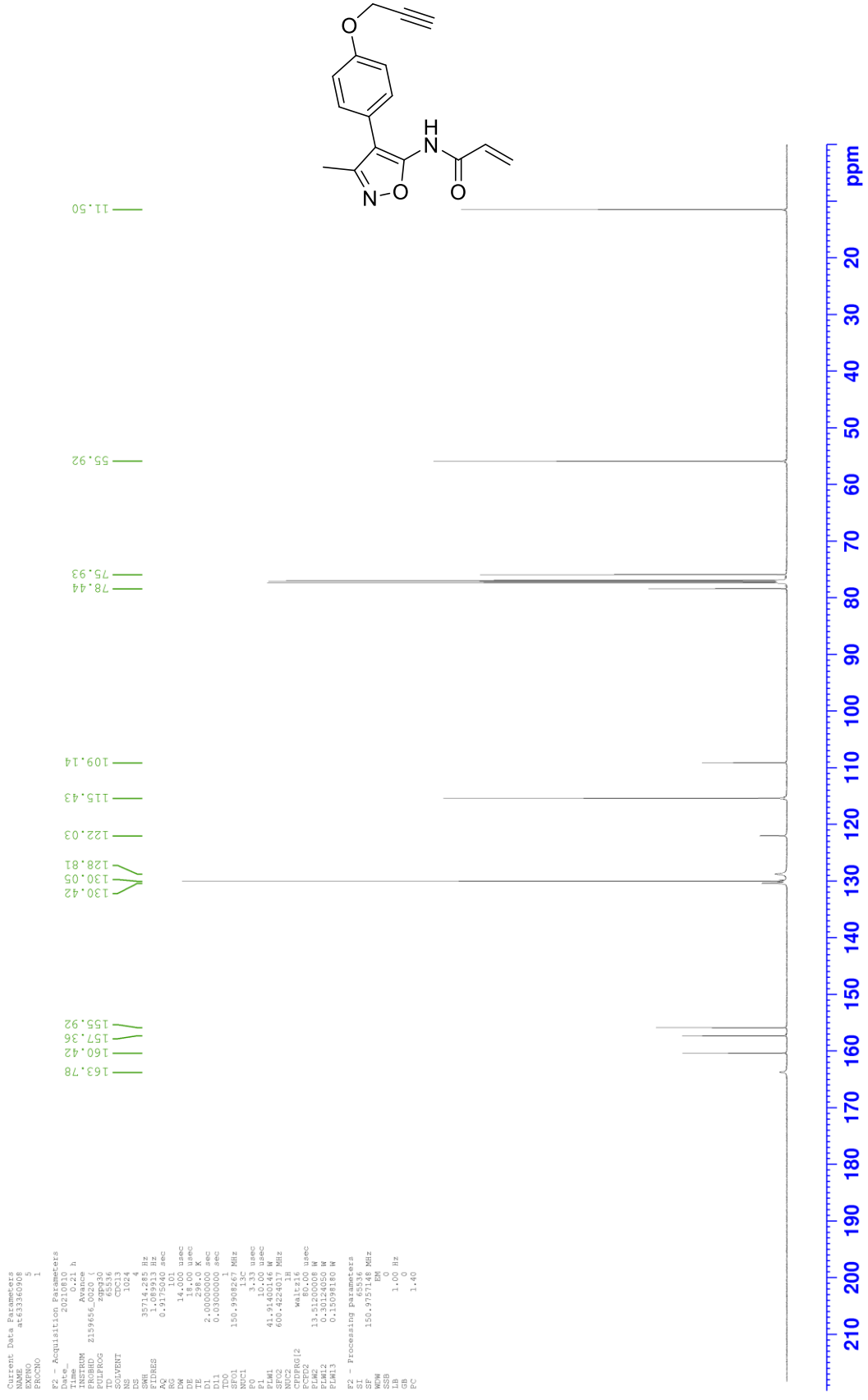
4[iv]-(Dimethylamino)-*N*-[3-methyl-4-(4[ii]-prop-2[iii]-ynoxyphenyl)-isoxazol-5-yl]butanamide



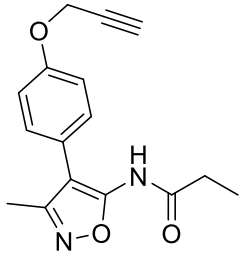
_N_-[3-Methyl-4-(4[ii]-prop-2[iii]-ynoxyphenyl)-isoxazol-5-yl]prop-2[iv]-enamide

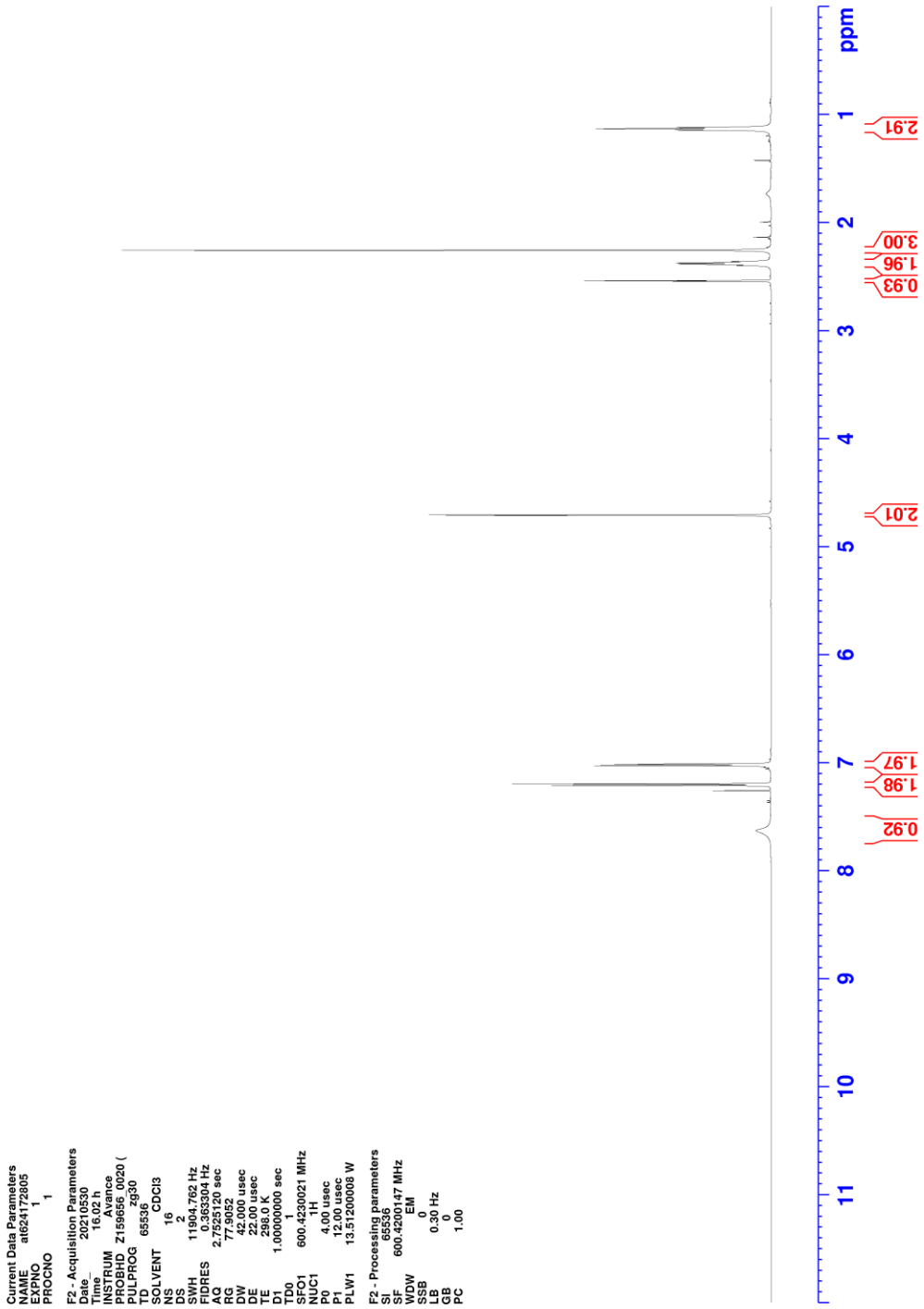


_N_-[3-Methyl-4-(4[ii]-prop-2[iii]-ynoxyphenyl)-isoxazol-5-yl]prop-2[iv]-enamide



***N*-[3-Methyl-4-(4[ii]-prop-2[iii]-ynoxyphenyl)-isoxazol-5-yl]propenamide**

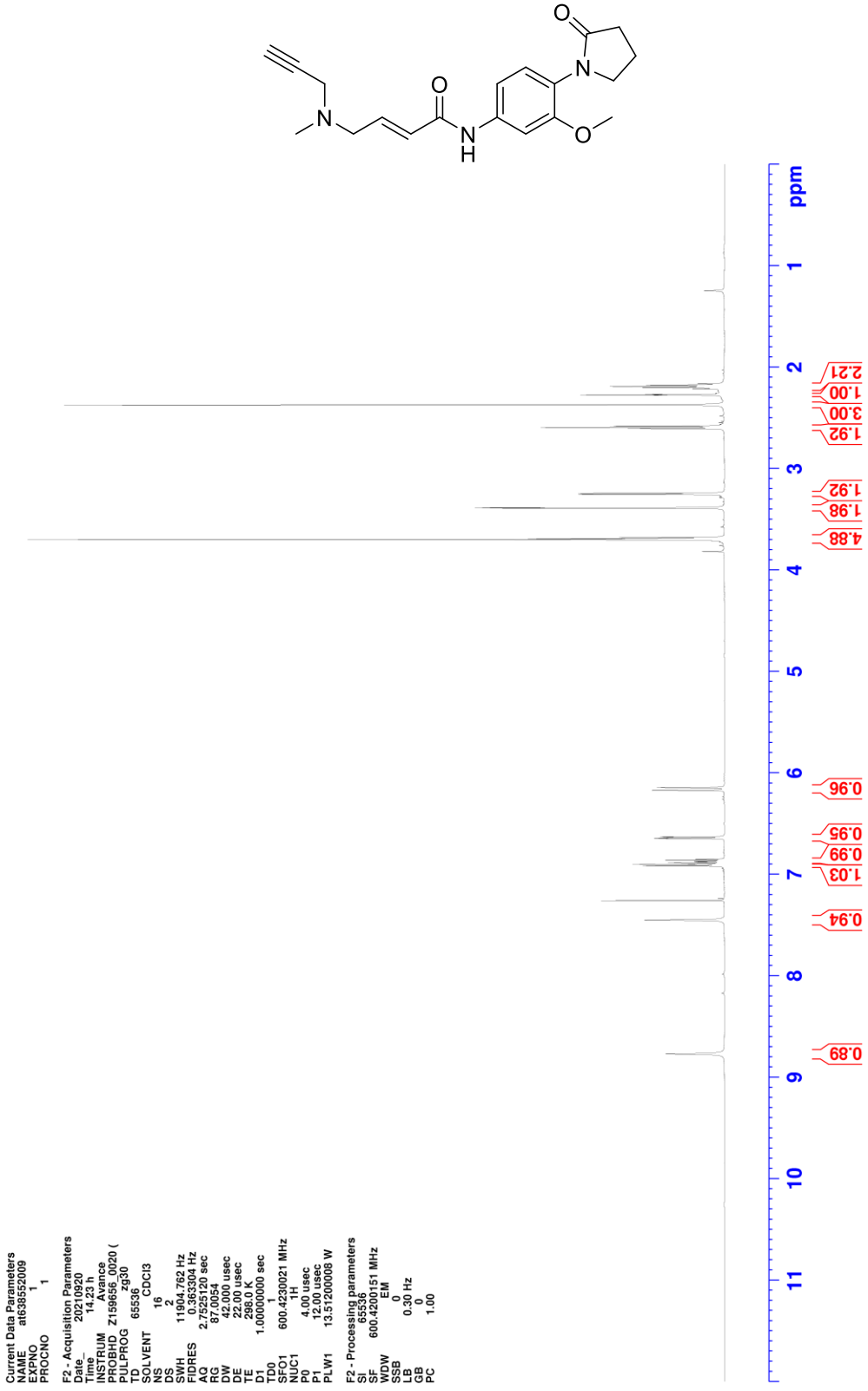




_N_-[3-Methyl-4-(4[ii]-prop-2[iii]-ynoxyphenyl)-isoxazol-5-yl]propenamide



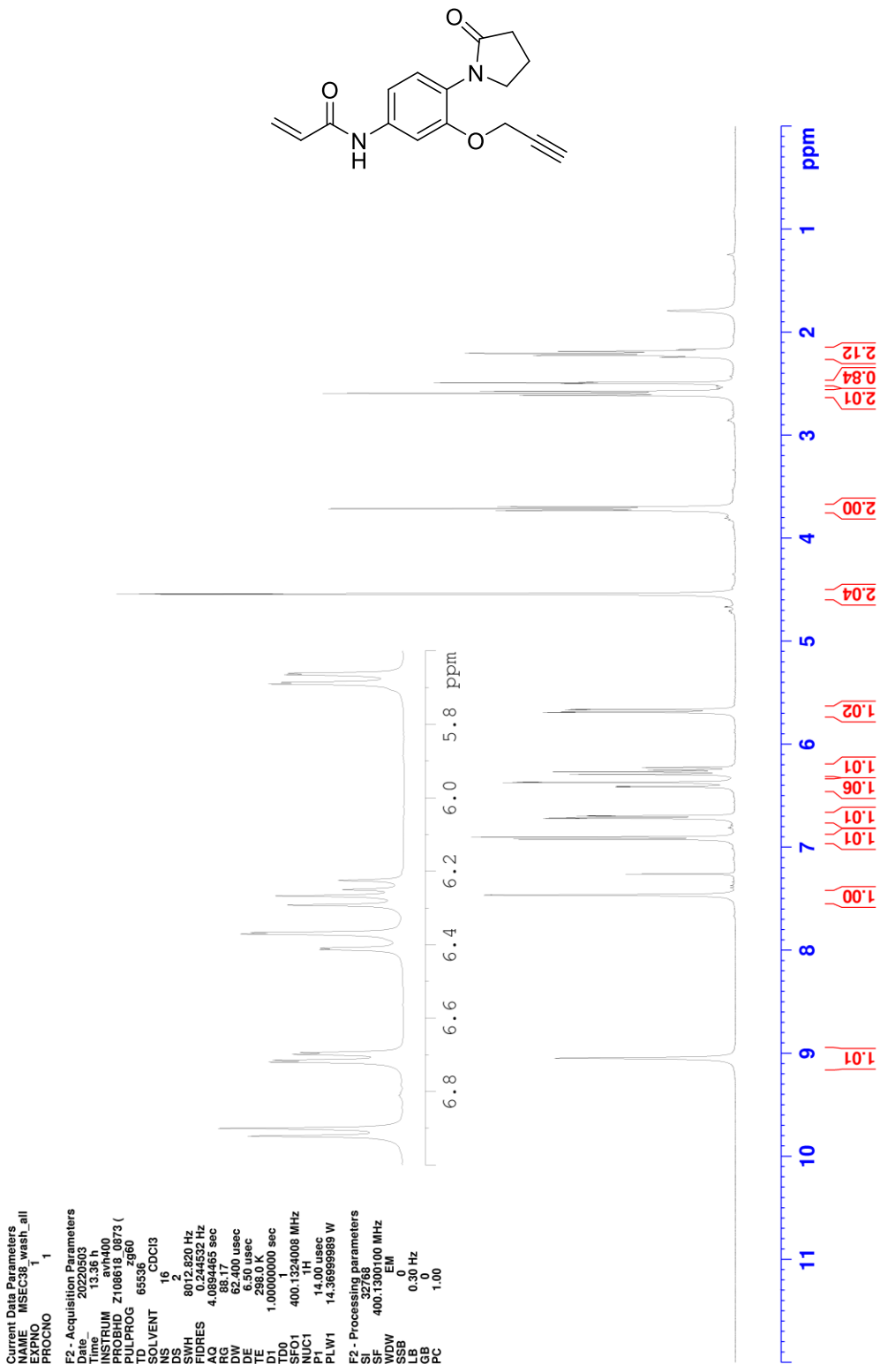
(*E*)-*N*-[3-Methoxy-4-(2^{v}-oxopyrrolidin-1^{v}-yl)phenyl]-4^{ii}-[methyl(prop-2^{iii}-ynyl)amino]but-2^{ii}-enamide



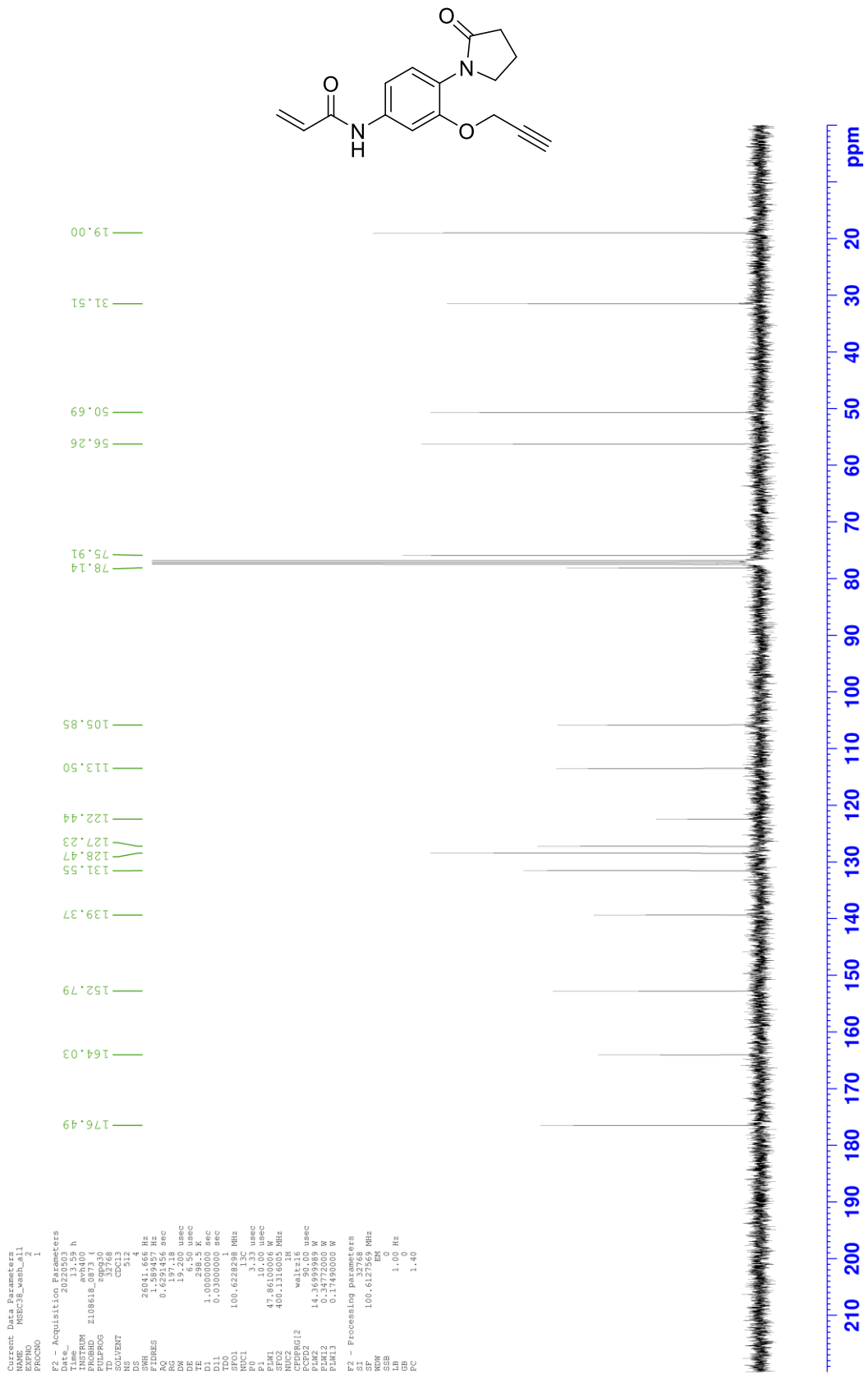
(*E*)-*N*-[3-Methoxy-4-(2^{v}-oxopyrrolidin-1^{v}-yl)phenyl]-4^{ii}-[methyl(prop-2^{iii}-ynyl)amino]but-2^{ii}-enamide



_N_-[4-(2-Oxopyrrolidin-1-yl)-3-prop-2-ynoxyphenyl]prop-2-enamide

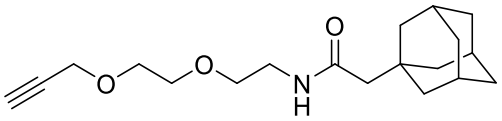


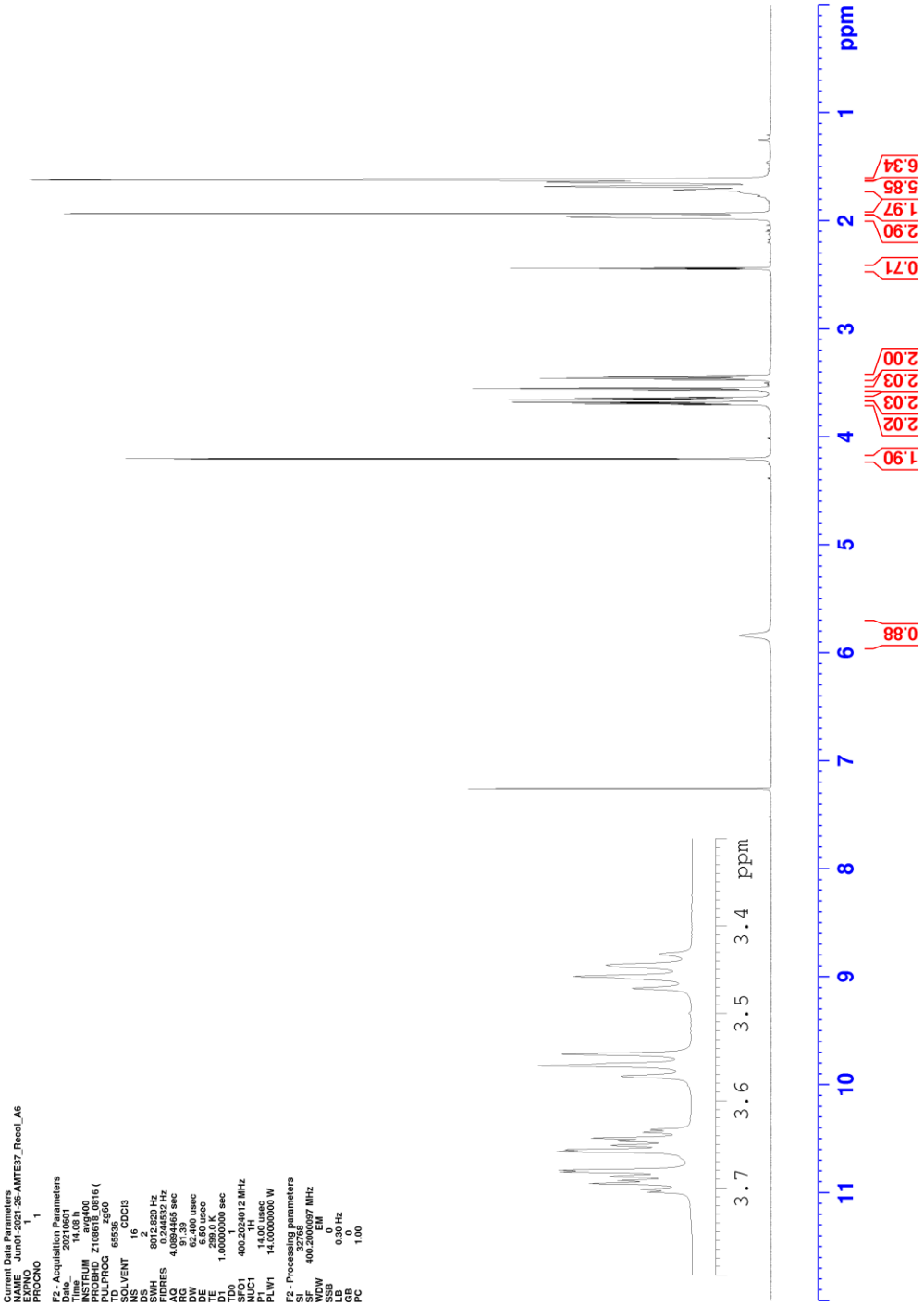
_N_-[4-(2-Oxopyrrolidin-1-yl)-3-prop-2-ynoxyphenyl]prop-2-enamide



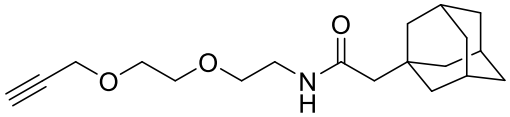
Chapter 6 spectra

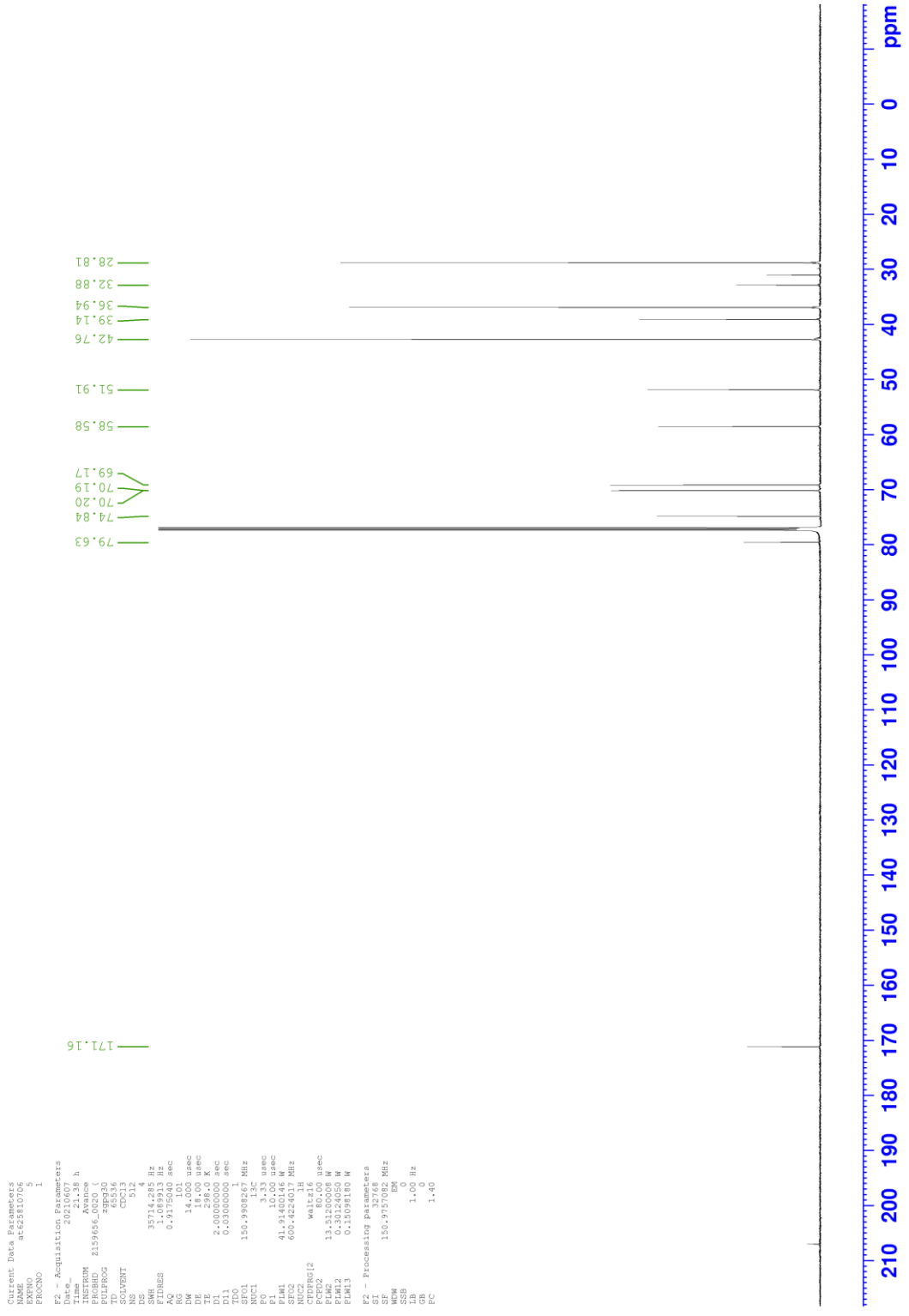
2^{iv}-(Adamantan-1^{v}-yl)-*N*-(2^{ii}-{2^{i}-[(prop-2-yn-1-yl)oxy]ethoxy}ethyl)acetamide



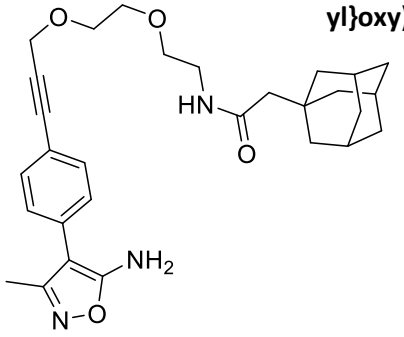


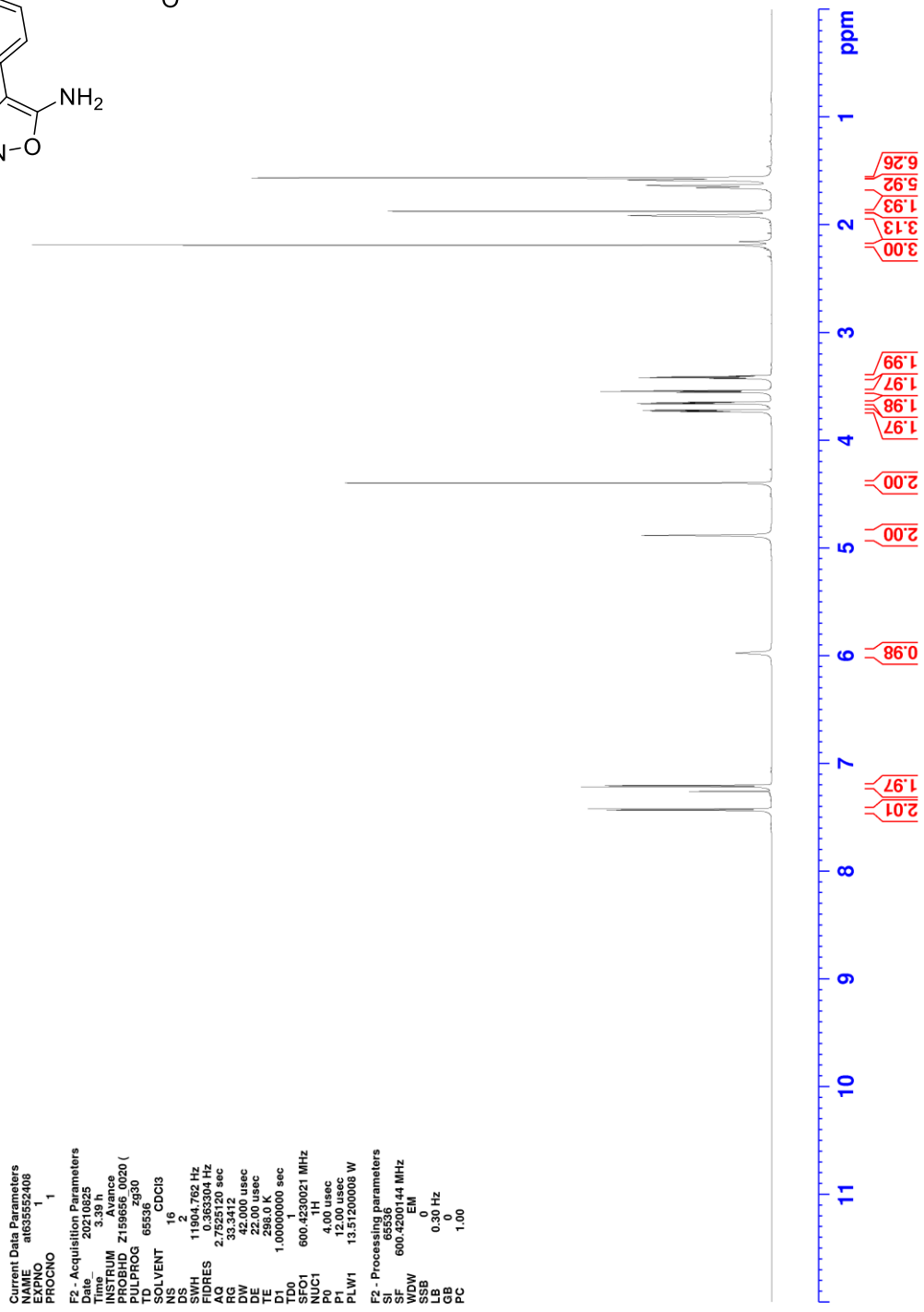
2[iv]-(Adamantan-1[v]-yl)-*N*-(2[ii]-{2[i]-[(prop-2-yn-1-yl)oxy]ethoxy}ethyl)acetamide





2[vii]-(Adamantan-1[viii]-yl)-*N*-{2[v]-[2[iv]-({3[iii]-[4[ii]-(5-amino-3-methylisoxazol-4-yl)phenyl]prop-2[iii]-yn-1[iii]-yl}oxy)ethoxy]ethyl}acetamide



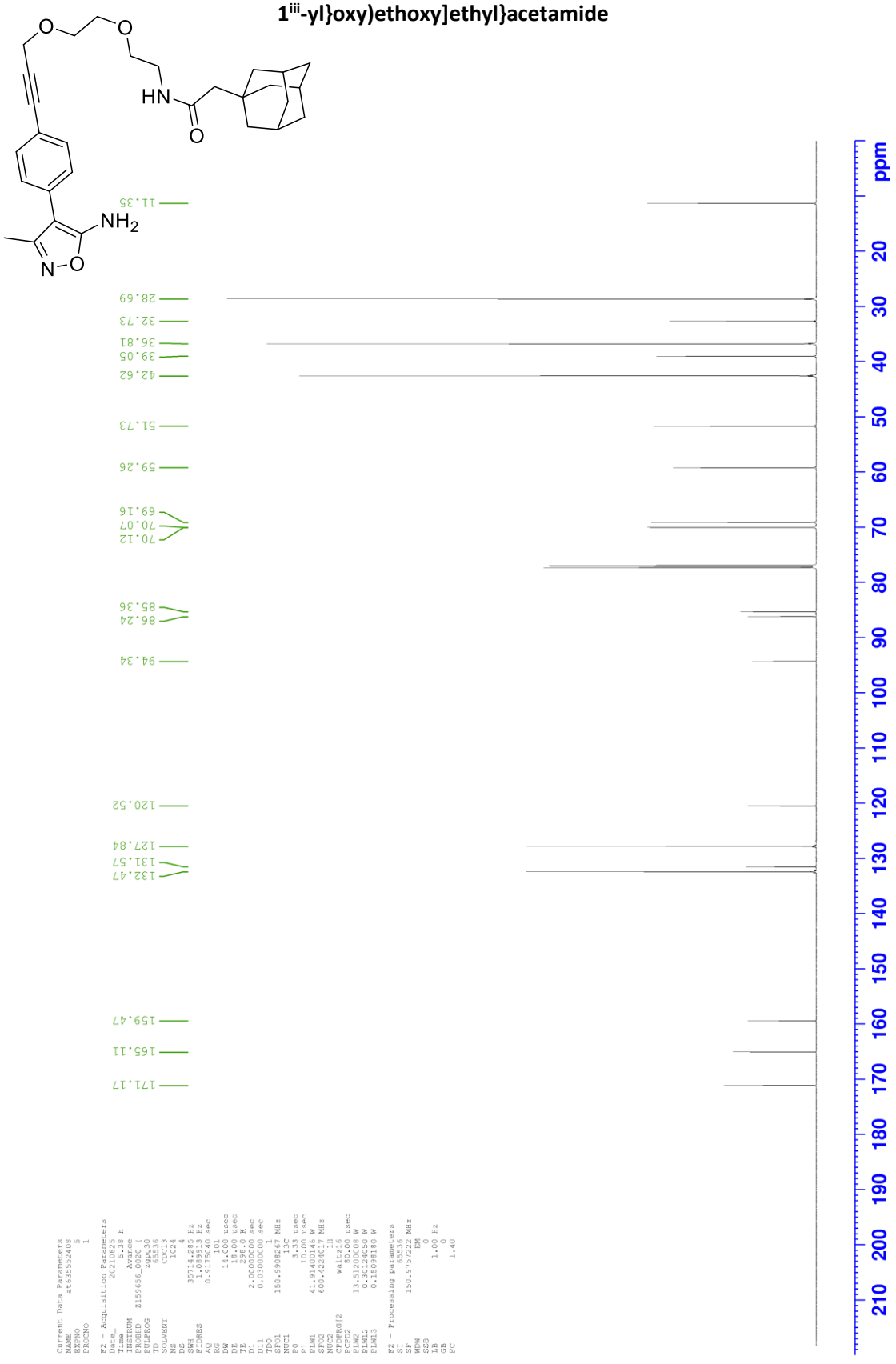


Current Data Parameters
NAME at635552408
EXPNO 1
PROCNO 1

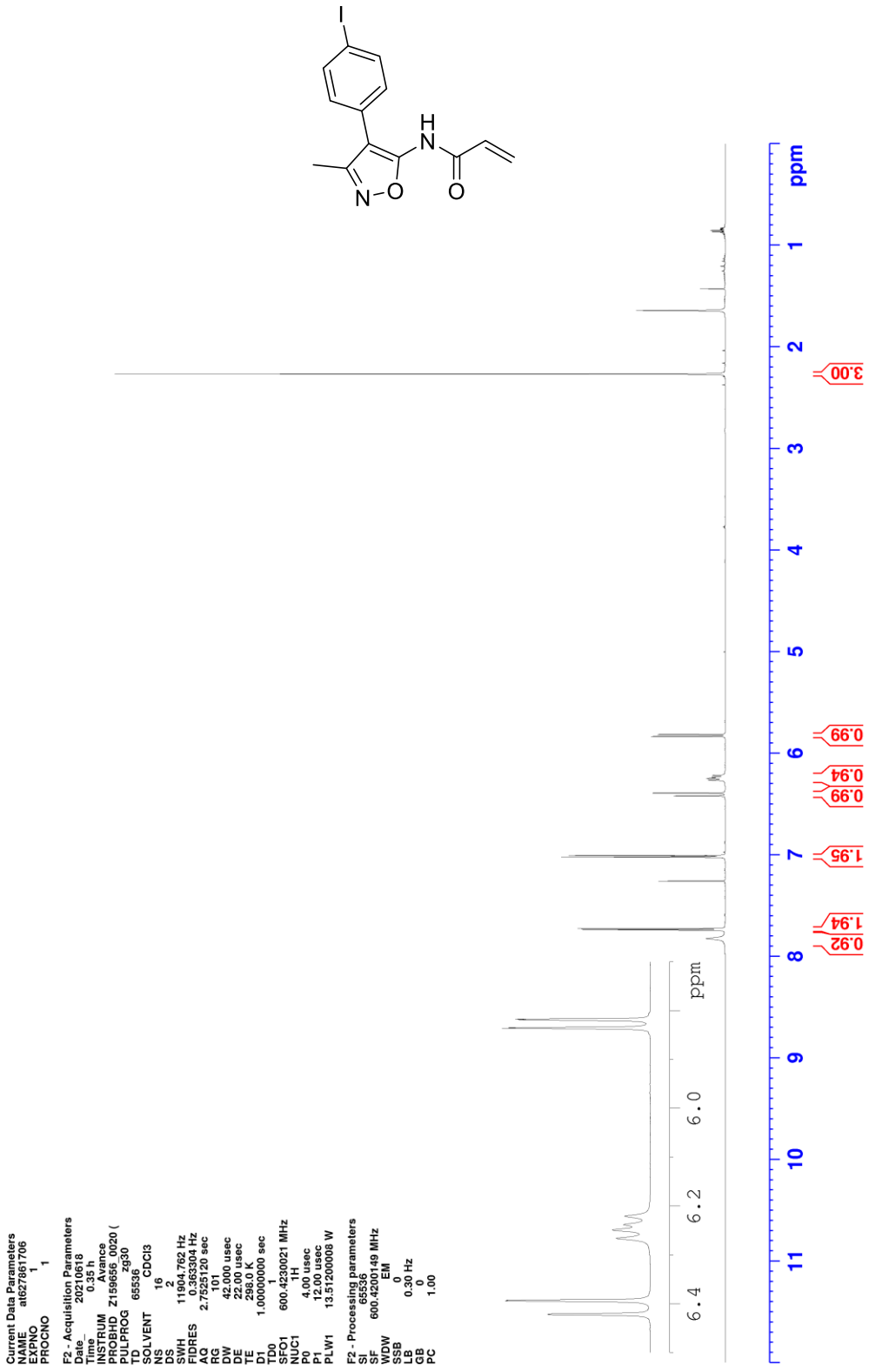
F2 - Acquisition Parameters
Date_ 20210825
Time_ 3.39 h
INSTRUM Avance
PROBHD Z159656_0020 (
PULPROG zg30
TD 65536
SOLVENT CDCl3
NS 16
DS 2
SWH 11904.762 Hz
FIDRES 0.363304 Hz
AQ 2.7525120 sec
RG 33.3412
DW 42.000 usec
DE 22.00 usec
TE 298.0 K
D1 1.00000000 sec
TD0 1
SFO1 600.4230021 MHz
NUC1 1H
P0 4.00 usec
P1 12.00 usec
PLW1 13.51200008 W

F2 - Processing parameters
SI 65536
SF 600.4200144 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

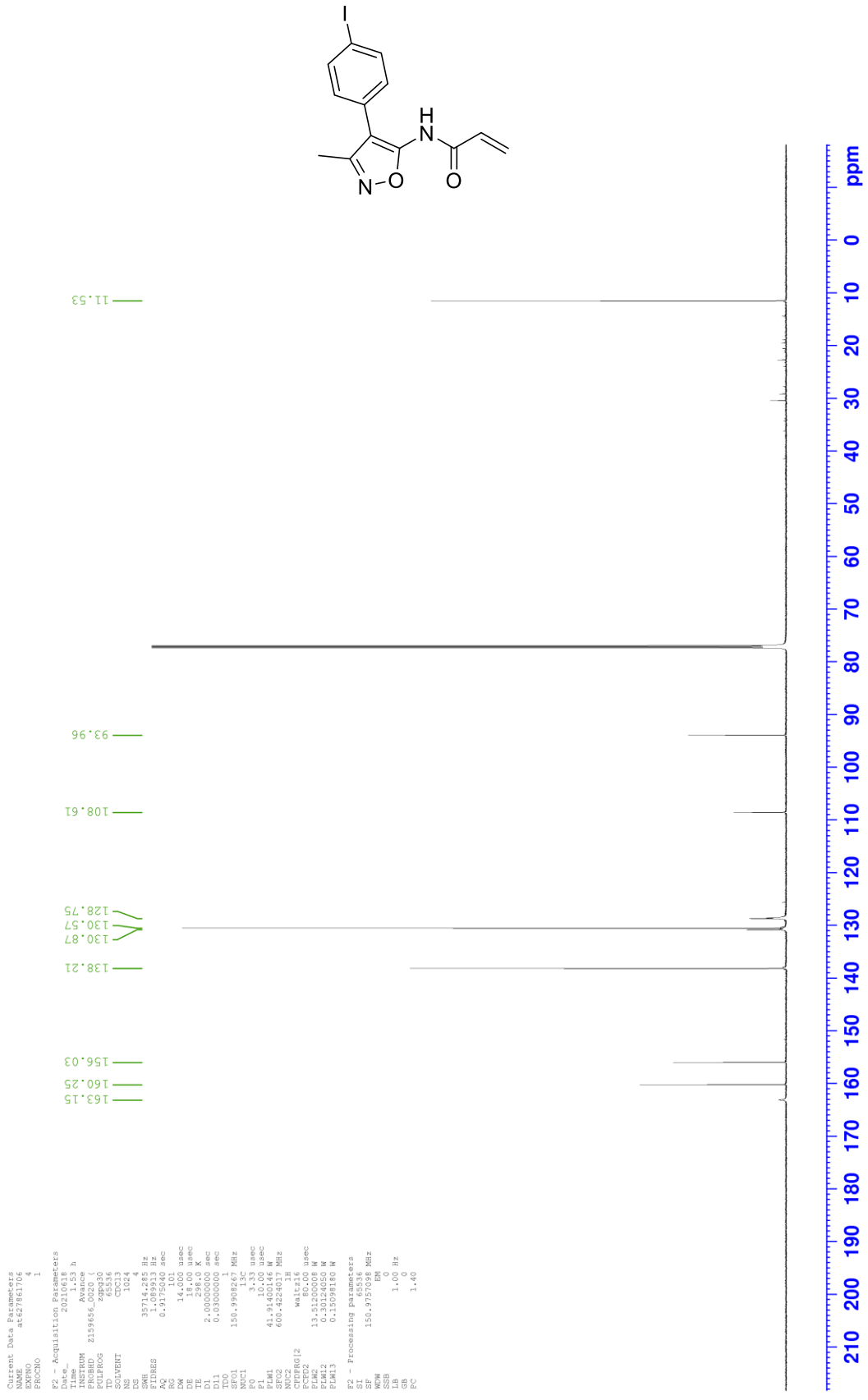
2^{vii}-(Adamantan-1^{viii}-yl)-*N*-{2^{v}-[2^{iv}-({3^{iii}-[4^{ii}-(5-amino-3-methylisoxazol-4-yl)phenyl]prop-2^{iii}-yn-1^{iii}-yl}oxy)ethoxy]ethyl}acetamide



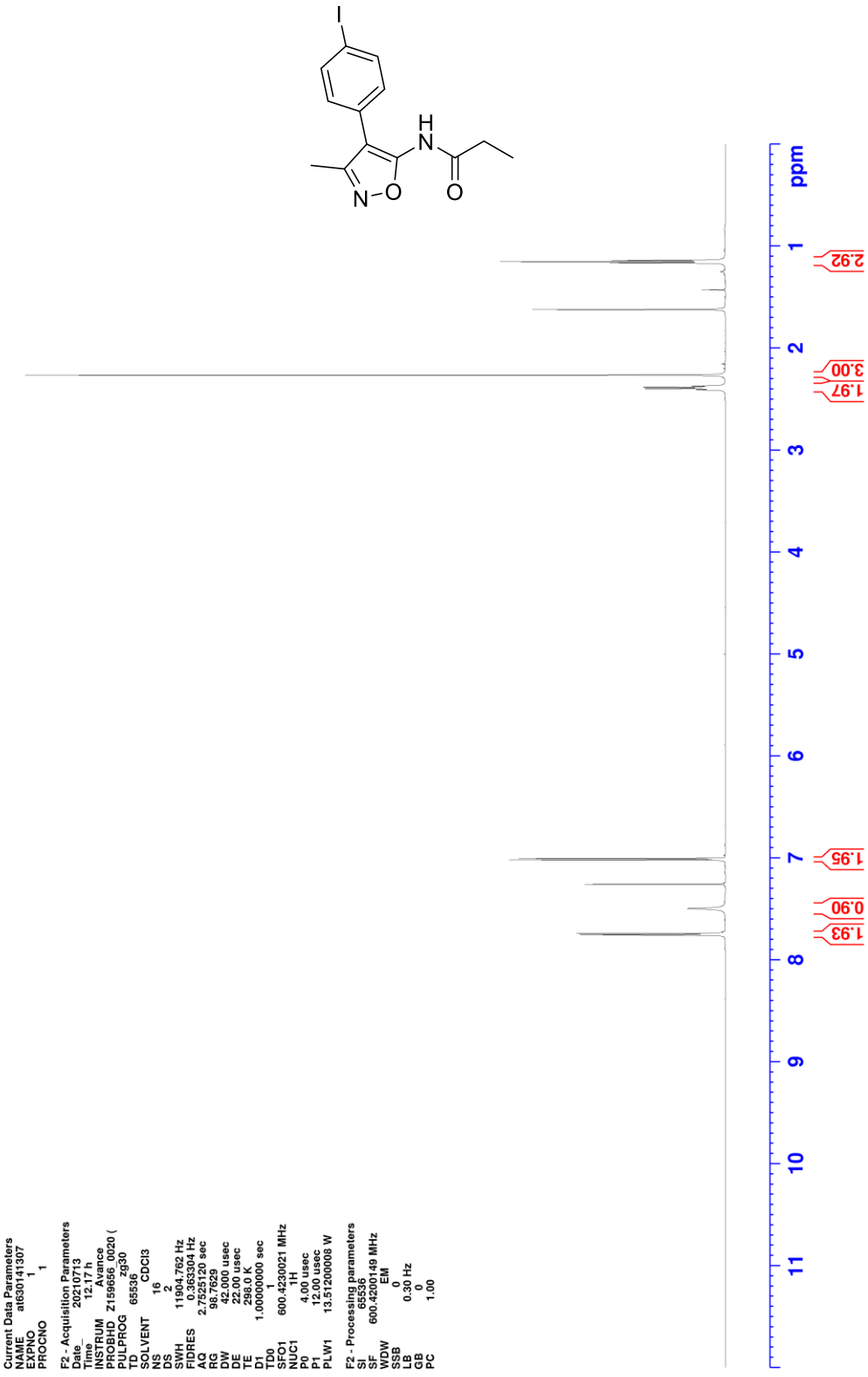
_N_-[4-(4[ii]-Iodophenyl)-3-methylisoxazol-5-yl]prop-2[iii]-enamide



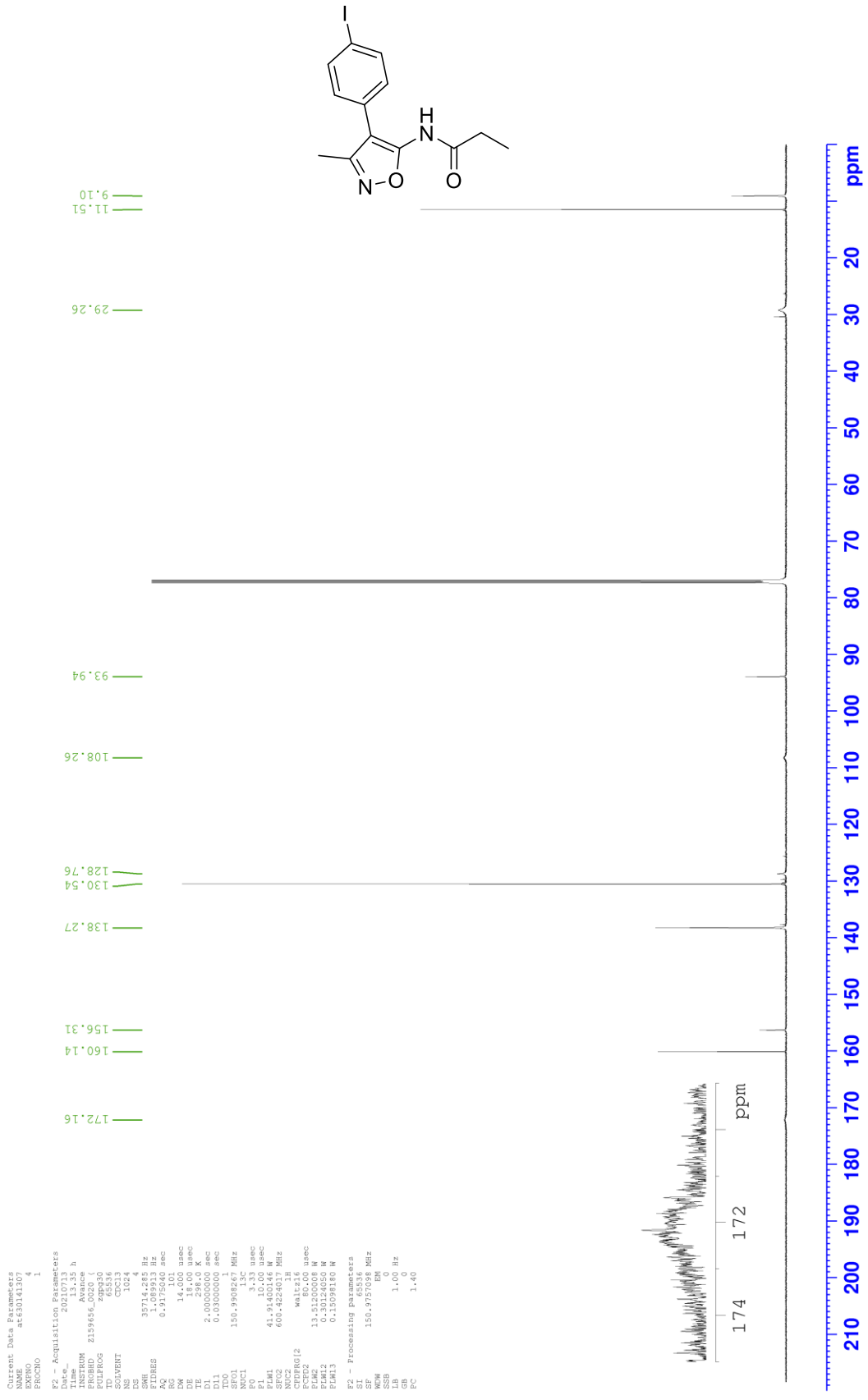
_N_-[4-(4[ii]-Iodophenyl)-3-methylisoxazol-5-yl]prop-2[iii]-enamide



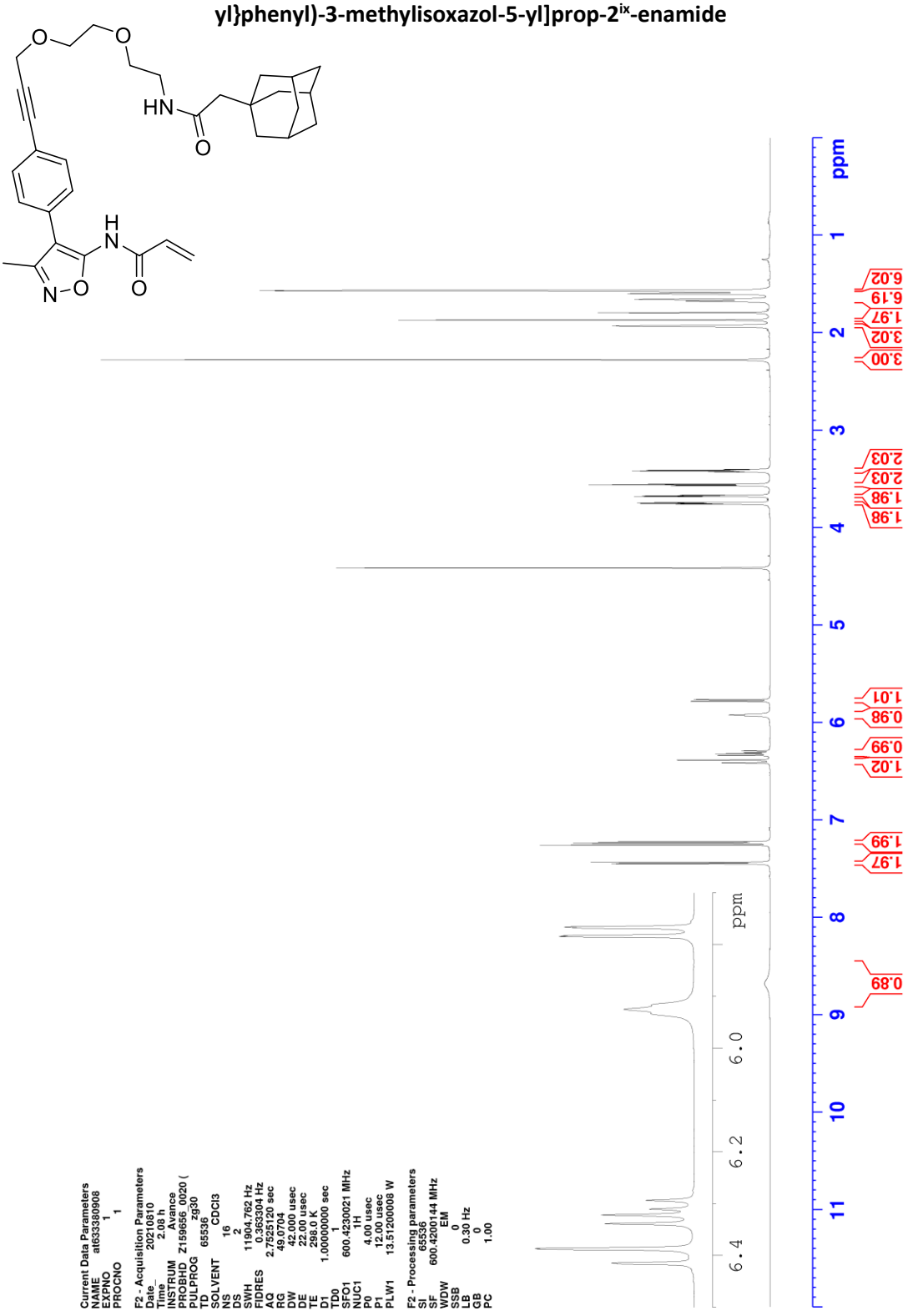
_N_-[4-(4[ii]-Iodophenyl)-3-methylisoxazol-5-yl]prop-2[iii]-enamide



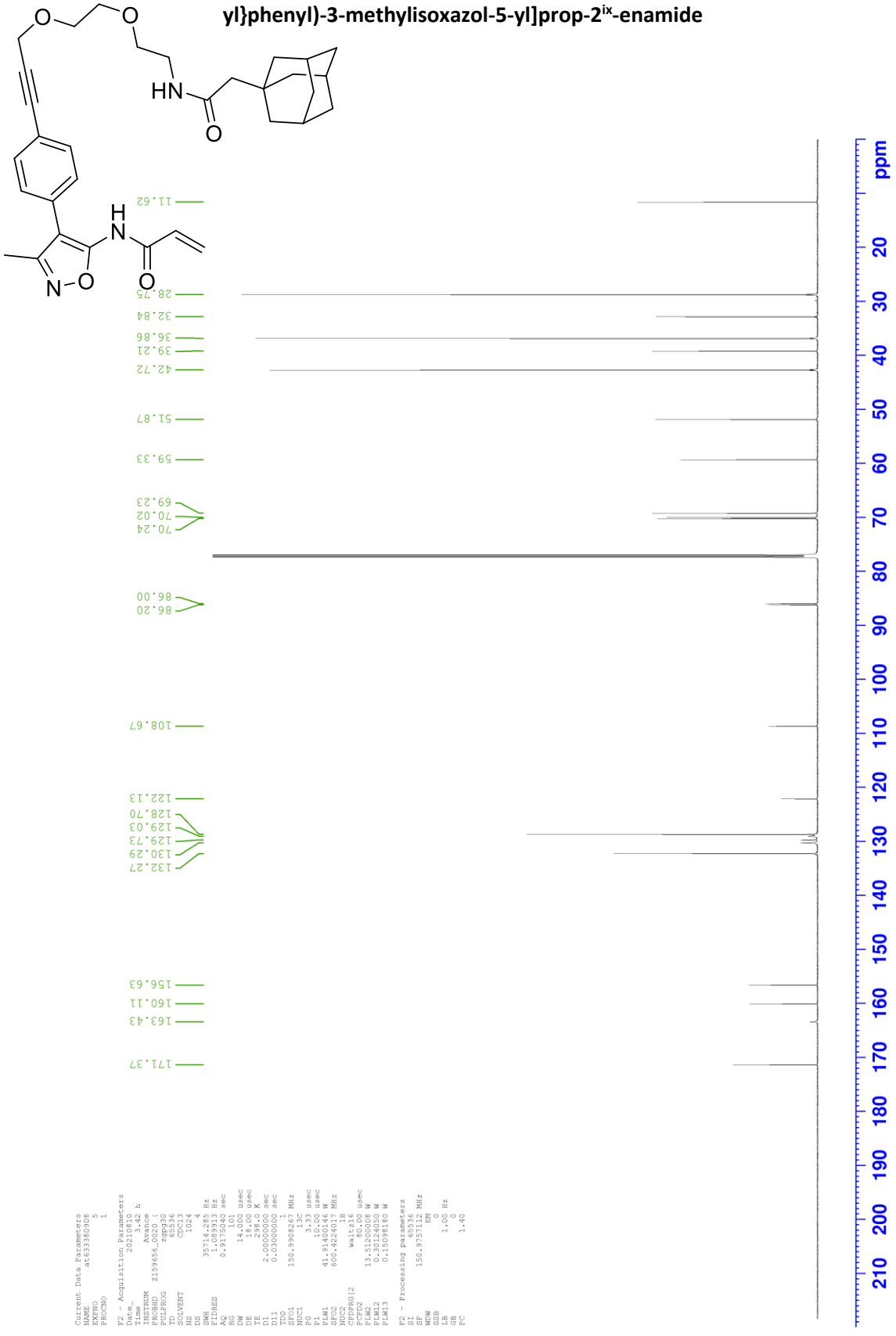
_N_-[4-(4[ii]-Iodophenyl)-3-methylisoxazol-5-yl]prop-2[iii]-enamide



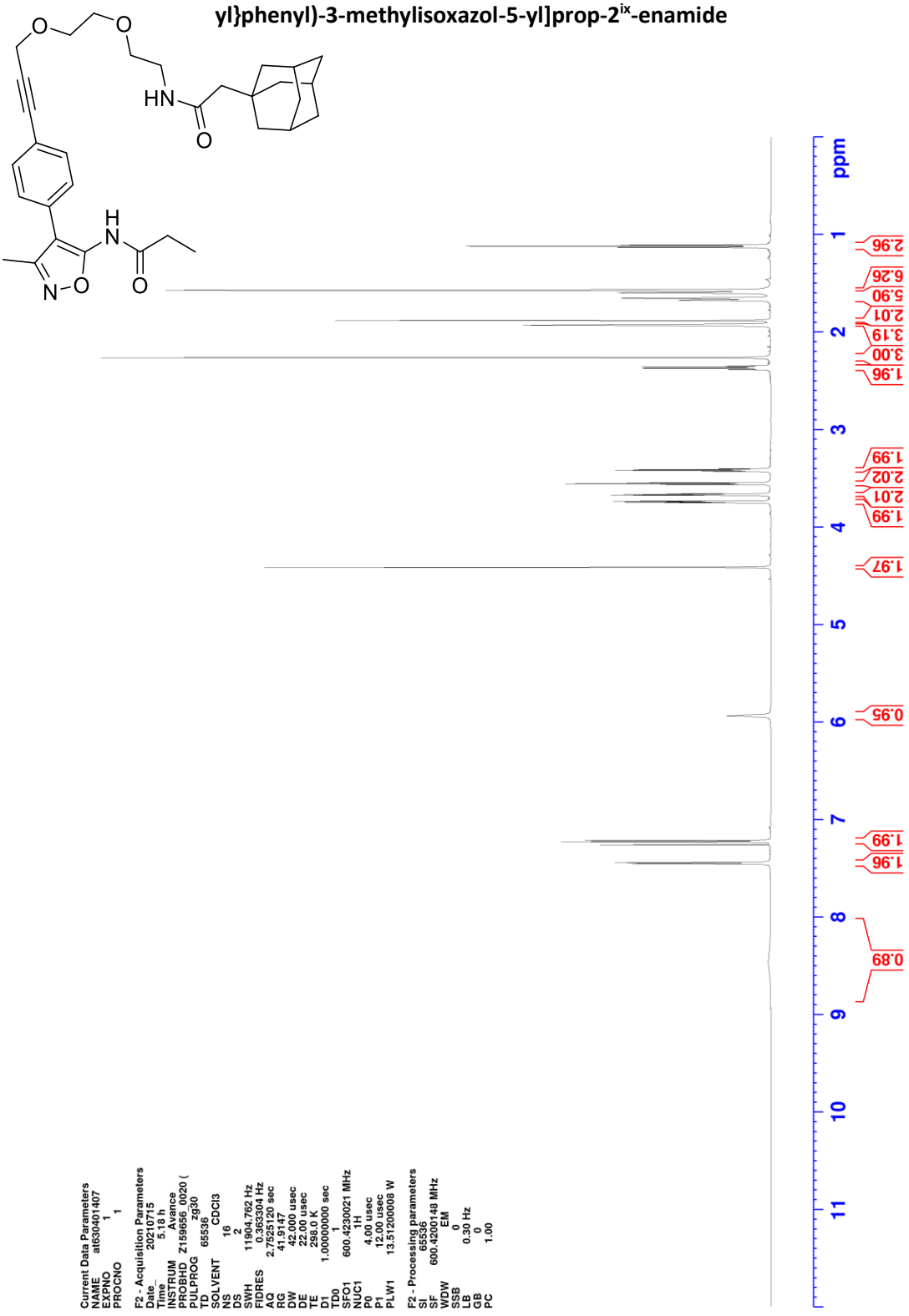
***N*-[4-(4[ii]-{3[iii]-[2[v]-(2[iv]-{2[vii][Adamantan-1[viii]-yl]acetamido}ethoxy)ethoxy]prop-1[iii]-yn-1[iii]-yl}phenyl)-3-methylisoxazol-5-yl]prop-2[ix]-enamide**



***N*-[4-(4[ii]-{3[iii]-[2[v]-(2[iv]-{2[vii][Adamantan-1[viii]-yl]acetamido}ethoxy)ethoxy]prop-1[iii]-yn-1[iii]-yl}phenyl)-3-methylisoxazol-5-yl]prop-2[ix]-enamide**



***N*-[4-(4^{ii}-{3^{iii}-[2^{v}-(2^{iv}-{2^{vii}[Adamantan-1^{viii}-yl]acetamido}ethoxy)ethoxy]prop-1^{iii}-yn-1^{iii}-yl}phenyl)-3-methylisoxazol-5-yl]prop-2^{ix}-enamide**

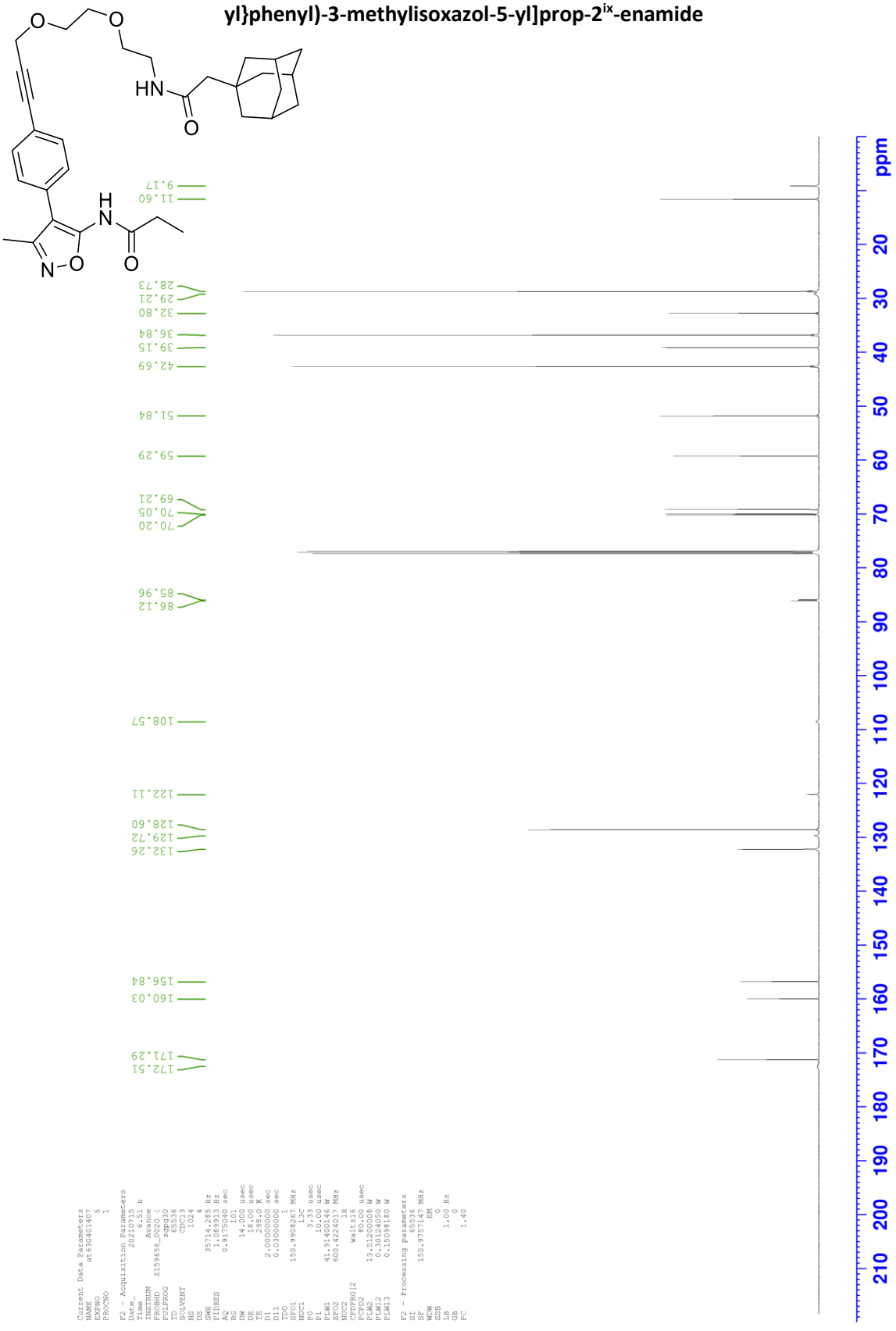


Current Data Parameters
NAME at630401407
EXPNO 1
PROCNO 1

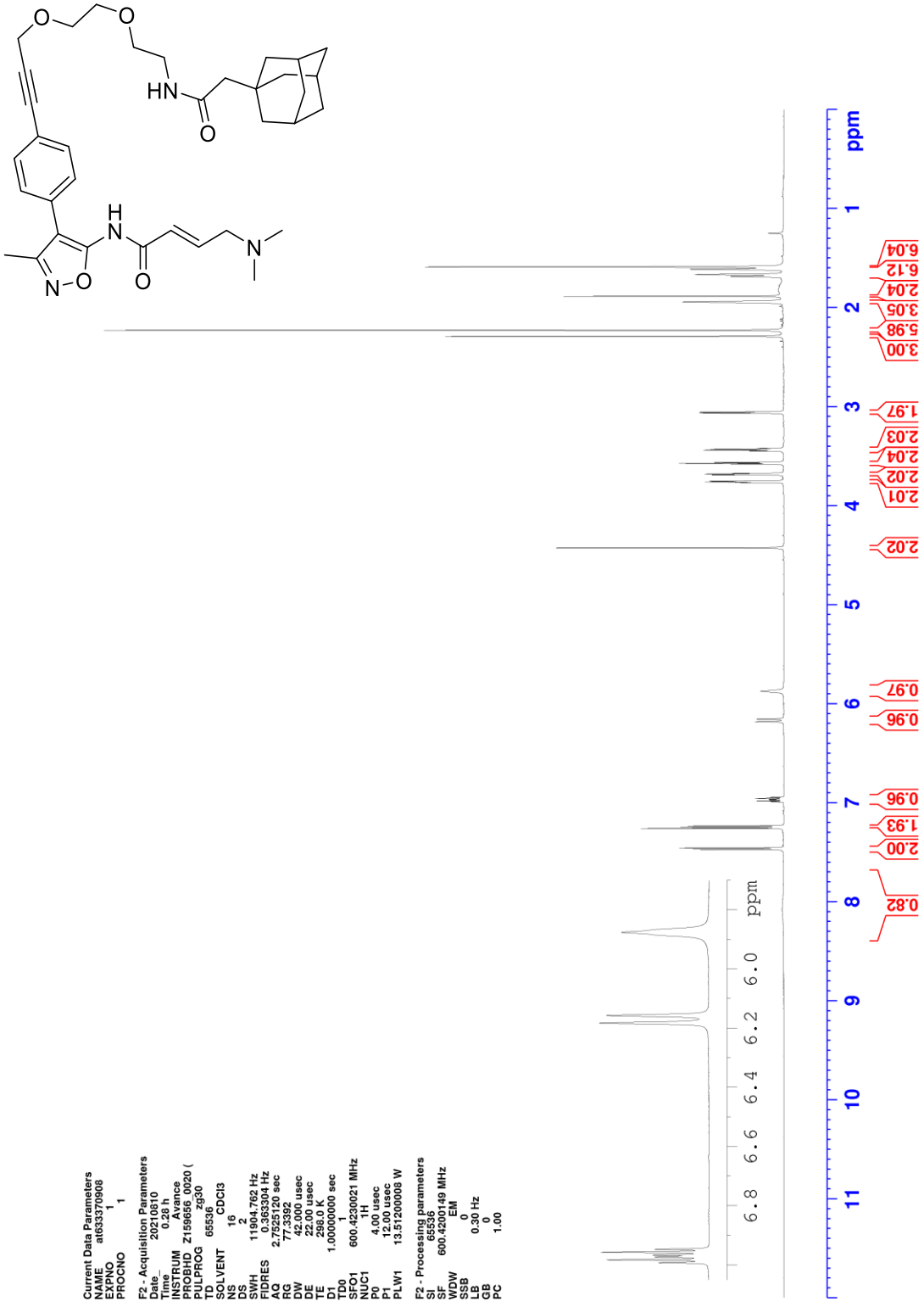
F2 - Acquisition Parameters
Date_ 20210715
Time 5.18 h
INSTRUM Avance
PROBHD Z159656_0020 (
PULPROG zg30
TD 65536
SOLVENT CDCl3
NS 16
DS 2
SWH 11904.762 Hz
FIDRES 0.363304 Hz
AQ 2.7525120 sec
RG 41.9147
DW 42.000 usec
DE 22.00 usec
TE 298.0 K
D1 1.00000000 sec
TD0 1
SFO1 600.4230021 MHz
NUC1 1H
P0 4.00 usec
P1 12.00 usec
PLW1 13.51200008 W

F2 - Processing parameters
SI 65536
SF 600.4200148 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

_N_-[4-(4[ii]-{3[iii]-[2[v]-(2[iv]-{2[vii][Adamantan-1[viii]-yl]acetamido}ethoxy)ethoxy]prop-1[iii]-yn-1[iii]-yl}phenyl)-3-methylisoxazol-5-yl]prop-2[ix]-enamide



(*E*)-*N*-[4-(4[ii]-{3[iii]-[2[v]-(2[iv]-{2[vii][Adamantan-1[viii]-yl]acetamido}ethoxy)ethoxy]prop-1[iii]-yn-1[iii]-yl}phenyl)-3-methylisoxazol-5-yl]-4[ix]-(dimethylamino)but-2[ix]-enamide

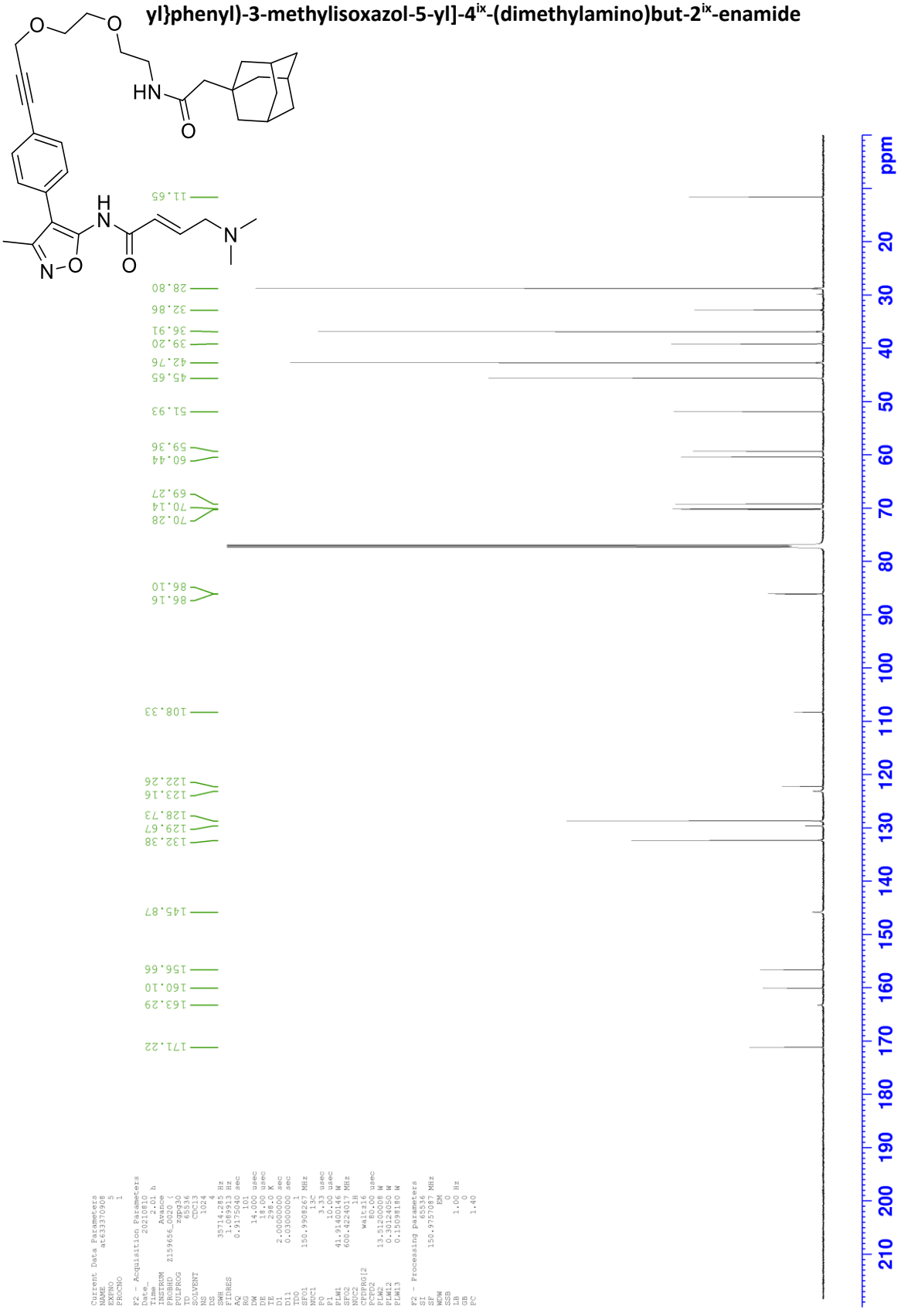


Current Data Parameters
NAME at633370908
EXPNO 1
PROCNO 1

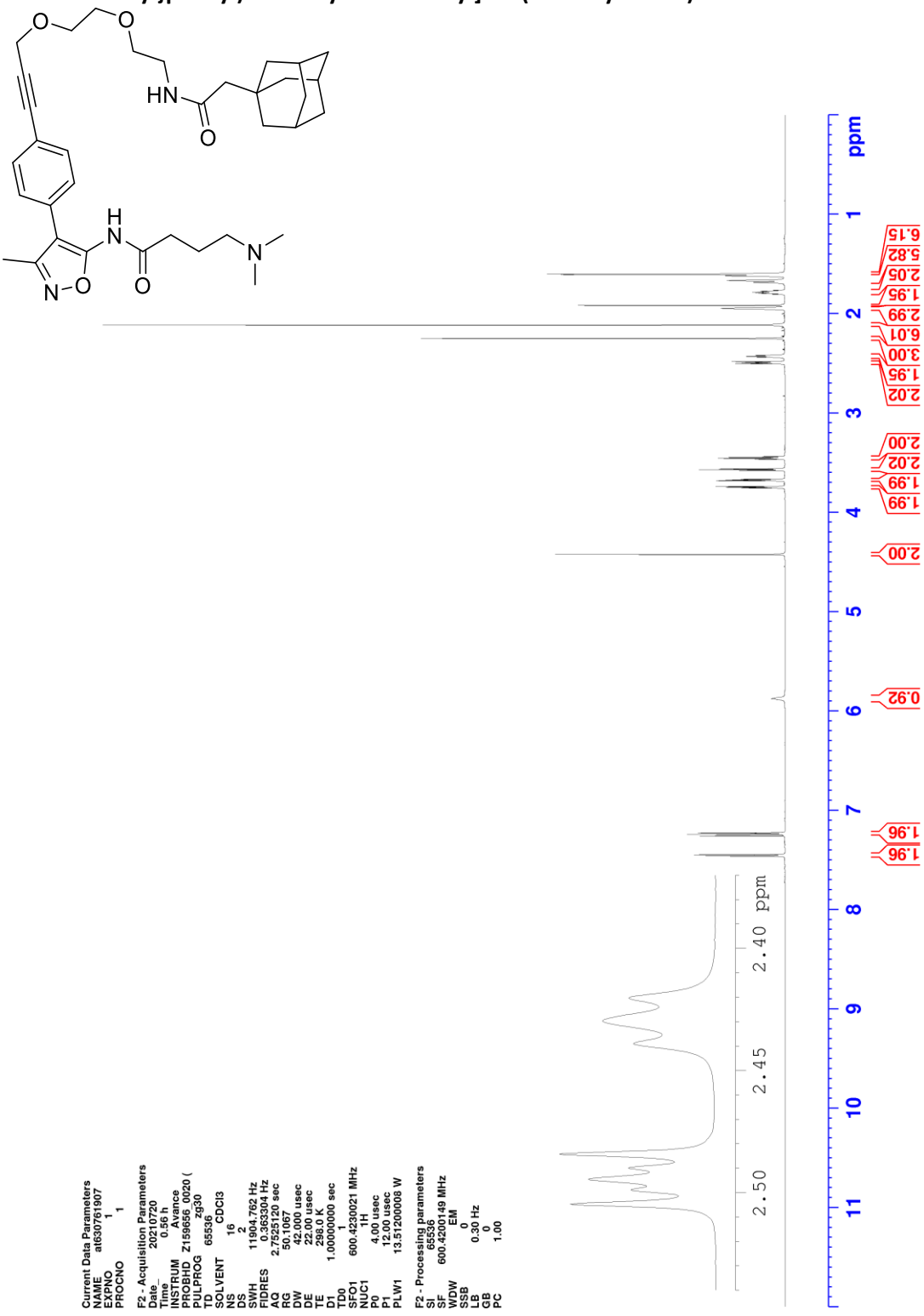
F2 - Acquisition Parameters
Date_ 20210810
Time_ 0.28 h
INSTRUM Avance
PROBHD Z159656_0020 (
PULPROG zg30
TD 65536
SOLVENT CDCl3
NS 16
DS 2
SWH 11904.762 Hz
FIDRES 0.363304 Hz
AQ 2.7525120 sec
RG 77.3392
DW 42.000 usec
DE 22.00 usec
TE 298.0 K
D1 1.00000000 sec
TD0 1
SFO1 600.4230021 MHz
NUC1 1H
P0 4.00 usec
P1 12.00 usec
PLW1 13.51200008 W

F2 - Processing parameters
SI 65536
SF 600.4200149 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

(*E*)-*N*-[4-(4[ii]-{3[iii]-[2[v]-(2[iv]-{2[vii][Adamantan-1[viii]-yl]acetamido}ethoxy)ethoxy]prop-1[iii]-yn-1[iii]-yl}phenyl)-3-methylisoxazol-5-yl]-4[ix]-(dimethylamino)but-2[ix]-enamide



***N*-[4-(4[ii]-{3[iii]-[2[v]-(2[iv]-{2[vii][Adamantan-1[viii]-yl]acetamido}ethoxy)ethoxy]prop-1[iii]-yn-1[iii]-yl}phenyl)-3-methylisoxazol-5-yl]-4[ix]-(dimethylamino)butanamide**



***N*-[4-(4[ii]-{3[iii]-[2[v]-(2[iv]-{2[vii][Adamantan-1[viii]-yl]acetamido}ethoxy)ethoxy]prop-1[iii]-yn-1[iii]-yl}phenyl)-3-methylisoxazol-5-yl]-4[ix]-(dimethylamino)butanamide**

