

# Deposit & Copying of Hardbound Dissertation Declaration



UNIVERSITY OF  
CAMBRIDGE

Board of Graduate Studies

Please note that you will also need to bind a copy of this Declaration into your final, hardbound print copy of your dissertation - this has to be the very first page of the hardbound dissertation. Do not include this form in the electronic version of your dissertation.

1	Surname (Family Name)	Forenames(s)	Title
2	Title of Dissertation as approved by the Degree Committee		

In accordance with the University Regulations in *Statutes and Ordinances* for the PhD, MSc and MLitt Degrees, I agree to deposit one or more copies of my dissertation and summary with the Secretary of the Board of Graduate Studies in a form or forms approved by the Board. The Secretary of the Board of Graduate Studies shall deposit a hardbound print copy of my dissertation and summary in the University Library under the following terms and conditions:

## 1. Dissertation Author Declaration

I am the author of this dissertation and hereby give the University the right to make my dissertation available in print form as described in 2. below.

My dissertation is my original work and a product of my own research endeavours and includes nothing which is the outcome of work done in collaboration with others except as declared in the Preface and specified in the text. I hereby assert my moral right to be identified as the author of the dissertation.

The deposit and dissemination of my dissertation by the University does not constitute a breach of any other agreement, publishing or otherwise, including any confidentiality or publication restriction provisions in sponsorship or collaboration agreements governing my research or work at the University or elsewhere.

## 2. Access to Dissertation

I confirm that I have selected an appropriate access and permission usage level for my dissertation, which is detailed on the Access Confirmation Form that I submitted with the electronic version of my dissertation. I understand that one print copy of my dissertation will be deposited in the University Library for archival and preservation purposes, and that, unless upon my application an embargo or restricted access has been granted for a specified period of time by my Supervisor or Degree Committee respectively prior to this deposit, the dissertation will be made available by the University Library for consultation by readers in accordance with University Library Regulations and copies of my dissertation may be provided to readers in accordance with applicable legislation, or on other access and permission usage terms detailed on my Access Confirmation Form and/or as may be marked on the dissertation.

## 3. Agreement to terms

By including this Declaration in my Hardbound Dissertation I confirm that any statements on this Declaration are correct and that I agree to the terms for access to my dissertation.

Synthesis of Structurally Diverse *N*-  
Substituted Quaternary Carbon  
Containing Small Molecules  
and Their Application as Novel  
Starting Points for  
Drug Discovery



**Sarah L. Kidd**

Trinity College

University of Cambridge

December 2018

Supervised by Professor David R. Spring

This dissertation is submitted for the degree of **Doctor of Philosophy**



# Synthesis of Structurally Diverse N-Substituted Quaternary Carbon Containing Small Molecules and Their Application as Novel Starting Points for Drug Discovery

Sarah L. Kidd

New bioactive small molecules are urgently required to seed both medicinal chemistry and chemical biology research. Within the two last decades, however, organic synthesis has been identified as a limiting factor in this regard owing to the historically uneven exploration of chemical space. Whilst fragment-based screening has proven a fruitful strategy for identifying novel bioactive small molecules, in a similar vein to high-throughput screening approaches, obstacles within this paradigm have been recognised. This relates to a lack of fragment novelty, the overrepresentation of  $sp^2$ -rich compounds and the lack of readily accessible exit vectors currently demonstrated by commercial screening collections. Thus, the development of novel screening libraries that address these deficiencies is therefore of significant importance. Moreover, innovative strategies are required to aid the somewhat challenging hit validation phase. In this context, the development of high-throughput platforms to accelerate the synthesis of derivatives could provide the opportunity to lower the economic impact of this process.

This thesis seeks to explore how diversity-oriented synthesis can be leveraged to tackle these issues. To populate new areas of fragment space, in the first chapter we utilise an  $\alpha,\alpha$ -disubstituted amino alcohol building block to generate 12 novel three-dimensional fragments containing an underrepresented *N*-substituted quaternary carbon moiety. These molecules ultimately contributed to a screening library of 40 compounds featuring this key motif. Importantly, throughout the library medicinally relevant functionalities and modifiable substituents were installed to enable biological recognition and three-dimensional fragment growth. Finally, cheminformatic analysis demonstrated the broad molecular shape diversity of the library and adherence of the library to commonly adopted guidelines within the field. In addition, the library was also shown to compare favourably with existing commercial collections, exhibiting a higher number of chiral centres, a lower fraction aromatic and an improved molecular shape distribution.

In the second chapter, we demonstrate the ability of this library to deliver novel hits against challenging biological targets and facilitate multidirectional fragment growth. To achieve this, state-of-the-art high-throughput X-ray crystallographic fragment screening was conducted, leading to the identification of six hits against three novel biological targets (CFI<sub>25</sub>, Activin A and Penicillin-binding protein 3). Utilising the modular nature of the synthetic route developed,

derivatives of each hit were readily generated interrogating several synthetic exit vectors, including the *N*-substituted quaternary centre. In turn, these analogues were successfully employed as a hit validation strategy, thus providing novel starting points for drug discovery or chemical probe development.

Finally, the third chapter describes investigations toward the coupling of rapid analogue construction with the XChem screening platform to provide a direct structural-binding readout *via* screening of crude reaction mixtures. In this section, we demonstrate the applicability of this library toward high-throughput derivative synthesis using solid-phase catalysis to generate 78 analogues in less than one week. In addition, these studies formed the basis of preliminary investigations toward the development of methodology to enable click chemistry to be implemented within the crude screening protocol.

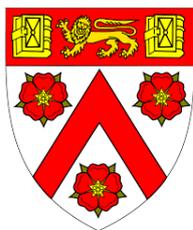
Together, these findings demonstrate the effective nature of structurally diverse *N*-substituted quaternary small molecules within a fragment-based screening context. Herein, we demonstrate libraries of this nature can provide novel three-dimensional scaffolds to seed screening campaigns and ultimately deliver novel starting points for drug discovery.

## Declaration

This dissertation is submitted in fulfilment of the requirements for the degree of Doctor of Philosophy. It describes work carried out in the Department of Chemistry, University of Cambridge, between October 2015 and December 2018 under the supervision of Prof. David Spring. Unless otherwise indicated, the research described is my own and not the product of collaboration. The work presented in this dissertation has not been submitted for any other degree. It does not exceed the prescribed word limit for the Physics and Chemistry Degree Committee.

Sarah Kidd

Trinity College, University of Cambridge





## Abstract

New bioactive small molecules are urgently required to seed both medicinal chemistry and chemical biology research. Within the two last decades, however, organic synthesis has been identified as a limiting factor in this regard owing to the historically uneven exploration of chemical space. Whilst fragment-based screening has proven a fruitful strategy for identifying novel bioactive small molecules, in a similar vein to high-throughput screening approaches, obstacles within this paradigm have been recognised. This relates to a lack of fragment novelty, the overrepresentation of sp<sup>2</sup>-rich compounds and the lack of readily accessible exit vectors currently demonstrated by commercial screening collections. Thus, the development of novel screening libraries that address these deficiencies is therefore of significant importance. Moreover, innovative strategies are required to aid the somewhat challenging hit validation phase. In this context, the development of high-throughput platforms to accelerate the synthesis of derivatives could provide the opportunity to lower the economic impact of this process.

This thesis seeks to explore how diversity-oriented synthesis can be leveraged to tackle these issues. To populate new areas of fragment space, in the first chapter we utilise an  $\alpha,\alpha$ -disubstituted amino alcohol building block to generate 12 novel three-dimensional fragments containing an underrepresented *N*-substituted quaternary carbon moiety. These molecules ultimately contributed to a screening library of 40 compounds featuring this key motif. Importantly, throughout the library medicinally relevant functionalities and modifiable substituents were installed to enable biological recognition and three-dimensional fragment growth. Finally, cheminformatic analysis demonstrated the broad molecular shape diversity of the library and adherence of the library to commonly adopted guidelines within the field. In addition, the library was also shown to compare favourably with existing commercial collections, exhibiting a higher number of chiral centres, a lower fraction aromatic and an improved molecular shape distribution.

In the second chapter, we demonstrate the ability of this library to deliver novel hits against challenging biological targets and facilitate multidirectional fragment growth. To achieve this, state-of-the-art high-throughput X-ray crystallographic fragment screening was conducted, leading to the identification of six hits against three novel biological targets (CFI<sub>25</sub>, Activin A and Penicillin-binding protein 3). Utilising the modular nature of the synthetic route developed, derivatives of each hit were readily generated interrogating several synthetic exit vectors, including the *N*-substituted quaternary centre. In turn, these analogues were successfully employed as a hit validation strategy, thus providing novel starting points for drug discovery or chemical probe development.

Finally, the third chapter describes investigations toward the coupling of rapid analogue construction with the XChem screening platform to provide a direct structural-binding readout *via* screening of crude reaction mixtures. In this section, we demonstrate the applicability of this library toward high-throughput derivative synthesis using solid-phase catalysis to generate 78 analogues in less than one week. In addition, these studies formed the basis of preliminary investigations toward the development of methodology to enable click chemistry to be implemented within the crude screening protocol.

Together, these findings demonstrate the effective nature of structurally diverse *N*-substituted quaternary small molecules within a fragment-based screening context. Herein, we demonstrate libraries of this nature can provide novel three-dimensional scaffolds to seed screening campaigns and ultimately deliver novel starting points for drug discovery.

## Acknowledgements

First and foremost, I would like to thank my supervisor Professor David Spring for providing me with the wonderful opportunity to be part of such an inspiring research group, for all his support and allowing me to pursue my research interests both in Cambridge and beyond. I would also like to thank my industrial supervisor Andrew Madin who has been a great source of helpful advice and suggestions to steer my project and to Professor Frank von Delft for enabling me to delve into the world of fragment-screening and XChem during my time within his group. I am also very grateful to the University/AstraZeneca studentship scheme for the financial support during my studies.

I would like to say a big thank-you to Dr. Natalia Mateu-Sanchis for her invaluable and inspiring mentorship during my PhD, in addition to conducting the research that provided the basis for this project. Beyond this, she has been a great source of encouragement and advice, both in the UK and now continued over Skype. Thanks also to Dr. Hannah Sore for all of the helpful project discussions during my studies and to both Till and Elaine for all of your contributions towards the work described. I would also like to thank Anthony for his guidance, helpful discussions and input on the fragment screening/OpenTrons work and Romain for teaching me the wizardry of XChem crystallography.

The Spring group has been a wonderful place to work, full of many great people (past and present) and I will cherish many memories from my time here. In particular, thanks to Steve for following me here from GSK, being an ice-cream enabler and always a listening ear. To Hannah for being my blue blanket buddy and keeping me sane whilst writing this thesis. Thanks also to Tommy for all of your advice and supporting my somewhat questionable music interests, I hope the Kidder playlist continues to bring you joy (apologies to the rest of you). Thanks also to the wonderful proof-readers, many already mentioned but also to Gabri and Kim. There are many more people who have not been named, but the daily tea breaks with all the cake have been great brain fuel!

Outside of the research group, I am honoured to have met such wonderful people. Huge thanks go to Emma for always being a source of gin-spiration and being such a truly brilliant friend, I am very excited for my bridesmaid duties. Thanks to Elinor, for all our twin antics, to Tom and Ivy for your visionary fleece trendsetting and to Ruth, Drakso and Nelson for filling my college time with laughter, food comas and many GIFs.

Finally, I would like to thank my family who have wholeheartedly supported me over the years. To Dad for inspiring me to pursue my academic studies, and to both him, Rachel and Sowsan for their support and foodie weekends. And most importantly to Mum, thanks for always being a beacon of encouragement, for our wonderful daily phone calls of 'drivel' and your concern

as to whether my molecules are 'behaving'. You have always believed in me and been there in times of need. Last, but definitely not least, to Kris for being such a wonderful partner in crime for the past nine years. I am so excited that our paths have finally aligned and very much looking forward to many more globetrotting adventures together and arguments about where to go for dinner.

## Abbreviations

$^1\text{H}$	proton
$^{\circ}\text{C}$	degrees centigrade
3-D	3-dimensional
Ac	acetyl
ACN	acetonitrile
ActIIb-EDC	extracellular domain of the type II activin receptor
Akt	protein kinase B, also known as PKB
$\text{Ap}_4\text{A}$	diadenosine tetraphosphate
APA	alternative polyadenylation
Aq	aqueous
Ar	aromatic
Aux	auxiliary
B/C/P	build/couple/pair
BACE-1	beta-secretase 1
BET	bromodomain and extra terminal
Bn	benzyl
Boc	<i>tert</i> -butyloxycarbonyl
$\text{Boc}_2\text{O}$	di- <i>tert</i> -butyl dicarbonate
BS1	binding site 1
BS2	binding site 2
BS3	binding site 3
Bu	butyl
CAN	cerium (IV) ammonium nitrate
CAS	Chemical Abstracts Service
$\text{CFI}_{25}$	Cleavage factor I 25kDa
$\text{CFI}_m$	Cleavage factor Im
CHK1	Checkpoint Kinase 1
CNS	central-nervous system
COD	cyclooctadiene
$\text{Cp}^*$	pentamethylcyclopentadiene
Cp	cyclopentadiene
CT	cyclotrimerisation
$\delta$	chemical shift
d	doublet
DAR	diels-Alder reaction
DBU	1,8-Diazabicyclo[5.4.0]undec-7-ene

DEAD	diethyl azodicarboxylate
DHP	dihydropyran
DIBALH	diisobutylaluminium hydride
DIPEA	diisopropyl ethylamine
DLS	Diamond Light Source
DMA	dimethylacetamide
DMEDA	<i>N,N</i> -dimethylethane-1,2-diamine
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
DOS	Diversity-Oriented Synthesis
<i>dr</i>	diastereomeric ratio
DSPL	Diamond and SGC Poised Library
IC <sub>50</sub>	half maximal inhibitory concentration
ELF	European Lead Factory
ELSD	evaporative light scattering detector
Eq(uiv).	equivalent(s)
ESI	electrospray ionisation
Et	ethyl
FBDD	fragment-based drug discovery
FBS	fragment-based screening
FDA	Food and Drug Administration
g	gram
GPCR	G-protein coupled receptor (1,3-Bis(2,4,6-trimethylphenyl)-2-
Grubbs II	imidazolidinylidene)dichloro(phenylmethylene)(tricyclohexylphosphine)ruth enium
HAC	heavy atom count
HBA	(number of) hydrogen-bond acceptors
HBD	(number of) hydrogen-bond donors
HRMS	high-resolution mass spectrometry
het	heterocycle
hrs	hours
HMBC	heteronuclear multiple-bond correlation (spectroscopy)
HSQC	heteronuclear single-quantum correlation (spectroscopy)
HTS	high-throughput screening
ID	identification
IR	infrared
J	coupling constant

JackiePhos	2-{Bis[3,5-bis(trifluoromethyl)phenyl]phosphino}-3,6-dimethoxy-2',4',6'-triisopropyl-1,1'-biphenyl
JackiePhos Pd G3	[(2-{Bis[3,5-bis(trifluoromethyl)phenyl]phosphine}-3,6-dimethoxy-2',4',6'-triisopropyl-1,1'-biphenyl)-2-(2'-amino-1,1'-biphenyl)] palladium(II) methanesulfonate
Kd	dissociation constant
Ki	inhibition constant
LCMS	liquid chromatography mass-spectrometry
LE	Ligand efficiency
logP	partition coefficient (lipophilicity)
M	metal/molar
m	milli/multiplet
m.p.	melting point
max	maximum
Me	methyl
MeCP2	methyl CpG-binding protein 2
mg	milligram
mL	millilitre
mmol	millimol
mins	minutes
MRSA	Methicillin-resistant <i>Staphylococcus aureus</i>
Ms	methane sulfonyl (mesyl)
mW	microwave
MW	molecular weight
NBE	new biologic entity
NHS	<i>N</i> -hydroxy succinimide
NME	new molecular entity
NMO	<i>N</i> -Methylmorpholine- <i>N</i> -Oxide
NMR	nuclear magnetic resonance
nOe	nuclear Overhauser effect
NOESY	nuclear Overhauser effect spectroscopy
nPR	Normalised principal moment of inertia ratio
NPs	natural products
NSQC	nitrogen-substituted quaternary carbon
<i>p</i>	<i>para</i>
PanDDA	pan-dataset density analysis
Pd/C	palladium on activated carbon
PG	protecting group

Ph	phenyl
PHIP2	pleckstrin homology domain-interacting protein
PKR	Pauson-Khand reaction
PMI	principal moment of inertia
PMP	<i>para</i> -methoxyphenyl
PPh <sub>3</sub>	triphenylphosphine
PPI	protein-protein interaction
ppm	parts per million
PSA	(topological) polar surface area
PSBH	partially saturated bicyclic heteroaromatics
q	quartet
R&D	research and development
r.t.	room temperature
RBC	rotatable bond count
RCEYM	ring closing enyne metathesis
RCM	ring closing metathesis
R <sub>f</sub>	retention factor
RO3	rule of three
RO5	rule of five
ROCK	RHO-associated protein kinase
s	singlet
S6K1	ribosomal protein S6 kinase 1
SAR	structure-activity relationship
SBR	structure-binding relationship
SGC	structural Genomics Consortium
SlogP	log octanol/water partition coefficient
SMEs	small/medium enterprises
SM	small molecule
SPR	surface-plasmon resonance
t	triplet
<i>t</i>	<i>tertiary</i>
TFA	trifluoroacetic acid
TGFβ	transforming growth factor β
THF	tetrahydrofuran
TLC	thin layer chromatography
TOS	Target-Oriented Synthesis
Ts	toluenesulfonyl
μmol	micromol

US	United States
USD	United State Dollar
W	watt
XCE	XChemExplorer



# Table of Contents

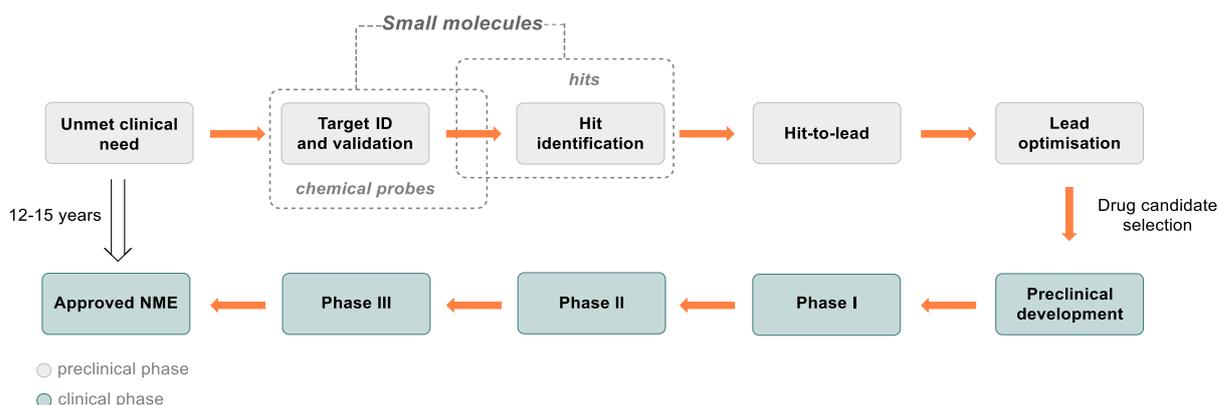
<b>1 Introduction</b> .....	1
<b>1.1 The role of small molecules within drug discovery</b> .....	1
<b>1.2 Sources of small molecules</b> .....	2
<b>1.3 Diversity-Oriented Synthesis</b> .....	4
<b>1.3.1. Diversity-Oriented Synthesis approaches: reagent, substrate and multi-component-based methods</b> .....	6
<b>1.3.2. The Build/Couple/Pair algorithm</b> .....	6
<b>1.3.3. The success of DOS</b> .....	7
<b>1.4 FBS</b> .....	8
<b>1.4.1 The success of FBS</b> .....	11
<b>1.4.2 Outstanding challenges</b> .....	12
<b>1.5 Thesis overview and aims</b> .....	15
<b>2 Results and discussion: DOS for the synthesis of <i>N</i>-substituted quaternary small molecules</b> .....	19
<b>2.1 Background</b> .....	19
<b>2.2 Project proposal</b> .....	23
<b>2.3 Library synthesis</b> .....	27
<b>2.3.1 Synthesis of the key amino alcohol building block I</b> .....	27
<b>2.3.2 Preparation of highly functionalised intermediates and further cyclisation: Couple and Pair phase</b> .....	29
<b>2.4 The complete DOS library</b> .....	50
<b>2.5 Cheminformatic assessment of the DOS library</b> .....	52
<b>2.5.1 PMI analysis</b> .....	52
<b>2.5.2 Physicochemical property analysis</b> .....	54
<b>3 Results and discussion: Investigations into the application of <i>N</i>-substituted quaternary carbon fragment libraries</b> .....	59
<b>3.1 Background</b> .....	59
<b>3.1.1 X-ray crystallography-based methods for fragment screening</b> .....	59
<b>3.1.2 The XChem screening platform</b> .....	59
<b>3.2 XChem screening of fragment-like DOS libraries</b> .....	63
<b>3.2.1 CFI<sub>25</sub></b> .....	63
<b>3.2.2 Follow-up compounds to NM450</b> .....	70
<b>3.2.3 Hit validation</b> .....	79
<b>3.2.4 Activin A</b> .....	83

3.2.5 Follow-up compounds to NM466 .....	85
3.2.6 Hit validation.....	89
3.2.7 On-going screening collaborations.....	90
<b>4 Results and discussion: Investigations into high-throughput methods of analogue production.....</b>	<b>91</b>
4.1 Background.....	91
4.2 DOS library application and toolbox expansion .....	96
4.2.1 High-throughput synthesis.....	97
4.2.2 Crude screening results .....	103
<b>5 Conclusions .....</b>	<b>105</b>
<b>6 Experimental .....</b>	<b>111</b>
6.1 General Experimental Details.....	111
6.2 <i>N</i> -substituted quaternary fragments .....	113
6.3 Synthesis of NM450 derivatives .....	141
6.4 Synthesis of NM466 derivatives .....	159
6.5 Synthesis of alkynes for high-throughput analogue synthesis .....	164
6.6 X-Ray crystallography and library screening.....	166
<b>7 References.....</b>	<b>169</b>
<b>8 Appendix .....</b>	<b>193</b>
8.1 The full NSQC library.....	193
8.2 Computational analysis .....	194
A) Principal Moment of Inertia (PMI) .....	194
B) Computational evaluation of physicochemical properties.....	203
8.3 Selected spectra .....	204
8.4 Publications .....	297

# 1 Introduction

## 1.1 The role of small molecules within drug discovery

The discovery of small molecules (SMs) with the ability to influence the function of biological macromolecules (e.g. proteins, DNA and polysaccharides) remains one of the greatest challenges within chemical biology.<sup>1</sup> Firstly, these entities are crucial to enable chemical genetics, in which SMs are used to probe and elucidate biological mechanisms, to develop our understanding of these complex systems.<sup>2-4</sup> Secondly, in a medicinal chemistry context where disease states are chemically modified,<sup>5</sup> SMs have continued to play a vital role over the decades. Indeed, their invaluable application as chemical probes has aided both target selection and validation for drug discovery purposes,<sup>6</sup> in addition to providing starting points for novel therapies (Figure 1.1). Crucially, whilst the development of new biologic entities (NBEs) such as antibody-based medicines have gained significant interest in recent years,<sup>7</sup> SMs continue to dominate pharmaceutical pipelines and the drug market.<sup>8</sup>



**Figure 1.1.** An overview of the drug discovery pipeline highlighting the essential role of SMs. Figure adapted from reference.<sup>9</sup> ID = identification, NME = new molecular entity.

However, despite the many technological and scientific advances within Research and Development (R&D) in the 1960s, a steady decline in the efficiency of drug discovery has been identified within recent years.<sup>8</sup> Indeed, since 1950 there has been a significant decrease in the number of New Molecular Entities (NMEs) approved per billion dollars spending (USD).<sup>10</sup> Strikingly in 2016 the FDA approved only 12 SM NMEs,<sup>11</sup> the lowest in almost five decades, despite an increase in the average cost to bring a drug to market reaching \$1.5bn in that timeframe.<sup>12</sup> Encouragingly in 2017 this increased to 34 SM-based NMEs,<sup>13</sup> however, staggeringly the costs are now estimated to be in the region of \$2-3bn.<sup>14</sup>

Whilst in part this decline can be attributed to more stringent safety regulations,<sup>14</sup> other factors such as drug attrition remain a major challenge. Over the last few years, the percentage of clinical candidates progressing from Phase I to clinical approval has steadily increased,

however, this figure is still alarmingly low in the region of 10-14%.<sup>15,16</sup> Previously the primary causes of attrition were attributed to poor bioavailability and pharmacokinetics, whereas failures are now understood to result from a lack of efficacy and safety issues, meaning more costlier late-stage failures.<sup>17-19</sup> Among the factors accredited to contribute to this is a lack of sufficient target validation, often linked to the disparity in the effects observed when targeting a given protein with a SM compared to those observed using genetic association studies and knockout models, for example *via* DNA mutations.<sup>20-23</sup> Thus, in order to draw robust disease hypotheses from these studies, there is an urgent need for new SM chemical probes with the ability to both inhibit and activate cellular pathways to complement these strategies.<sup>3,24</sup>

In addition to tackling drug candidate attrition, there remains a need for the discovery of new SMs capable of interacting with challenging and underexplored targets.<sup>25</sup> As the recent identification of SM inhibitors for the bromodomain and extra terminal (BET) family of bromodomains serves to demonstrate, these quests can indeed prove fruitful in enabling breakthrough chemical genetic research and inspiring the development of clinical candidates for novel oncology-based therapies.<sup>26-28</sup> Nevertheless, whilst the number of protein targets exploited for therapeutic purposes has more than doubled since 2006, G-protein coupled receptors (GPCRs), ion channels, kinases and nuclear receptors remain the most dominant target families, accounting for 70% of all SM drugs.<sup>29</sup> Moreover, a significant overlap between pharmaceutical portfolios has been identified, with over 58% of all targets and 46% of novel targets being pursued by more than one organisation in 2013.<sup>30</sup> Thus, a significant proportion of biological space remains underexploited for research and drug discovery purposes.

Nevertheless, a major hurdle in these pursuits is the availability or development of appropriate SMs. In cases where the biological target or native ligand are well characterised, structure-based or rational design strategies can be utilised.<sup>31</sup> In the absence of this information, however, this process can be somewhat challenging. An alternative strategy is the high-throughput screening (HTS) of SM libraries.<sup>32</sup> Typically, sources of compounds to populate these libraries comprise those obtained through either isolation of Nature's secondary metabolites (i.e. natural products) or chemical synthesis.<sup>33</sup> However, in this case the selection criteria of molecules to populate these collections becomes vastly complicated.<sup>34</sup>

## 1.2 Sources of small molecules

### *A) Nature as a source of molecular diversity*

Owing to the exquisite structural complexity and diversity exhibited by natural products (NPs), these SMs continue to provide inspiration for novel tools to perturb biological systems.<sup>35</sup> It is well established that these molecules interact with a wide range of targets and have delivered a rich source of drug discovery leads,<sup>36-38</sup> with around half of all clinical drugs deriving from NPs.<sup>39</sup> However, the isolation of these small molecules is non-trivial, and commonly purification

issues are faced.<sup>40</sup> As a result, these compounds are often screened as mixtures, which can complicate the identification of the active species.<sup>41</sup> As such, there has been a sustained interest within the synthetic and medicinal chemistry community in the artificial laboratory synthesis of such compounds and related analogues, which has aided our understanding and application of these SMs.<sup>42–44</sup> However, due to their structural complexity, often with multiple stereocenters and  $sp^3$  rich motifs, the synthesis of these structures remains incredibly challenging and time-consuming.

*B) Commercially available and pharmaceutical collections: Our exploration of chemical space*

Commercially available and pharmaceutical compound collections constitute an alternative source of bioactive molecules. The era of combinatorial ‘split-pool’ libraries within medicinal chemistry resulted in the vast population of small molecule libraries for HTS campaigns with millions of compounds.<sup>45,46</sup> As an inherent feature of their syntheses, however, these libraries commonly only feature appendage diversity around a limited number of molecular scaffolds.<sup>33,47,48</sup> Nature sees small molecules in a 3-dimensional (3-D) fashion as a binding surface, and as such complex biological targets will only interact with molecules featuring a complementary 3-D surface. Thus, it has been argued that the intrinsic shape diversity of a library can be attributed to its functional diversity; those with a limited number of core scaffolds are limited in the number of biological targets the library may interact with.<sup>5</sup>

Chemical space is astronomically vast with estimates in the region of  $10^{63}$  of all possible SMs.<sup>49</sup> Whilst generation of every permutation is physically impossible, the coverage of chemical space by synthetic small molecule collections thus far can be considered sub-optimal. Indeed, an enlightening analysis of the CAS registry by Lipkus *et al.* highlighted fifty-percent of known molecules explored only 0.25% of all known molecular scaffolds, whilst the most popular 5% of scaffolds were present in over 75% of compounds.<sup>50,51</sup> In addition to their methods of synthesis, this bias can be attributed to a historical preference of SMs fulfilling Lipinski’s Rule of Five (RO5) as a strategy to improve oral bioavailability and gauge ‘drug-likeness’.<sup>52–54</sup> Furthermore, recent analysis of large pharma company screening libraries suggested a substantial overlap and similarity of compounds within these collections.<sup>55</sup> This overall lack of diversity is alarming and a factor which has often been attributed to the limited success of these libraries when seeking to identify novel bioactive agents.<sup>56</sup> Moreover, this can lead to the assumption of a target as ‘undruggable’ due to the intractability when screening traditional compound libraries.<sup>57</sup>

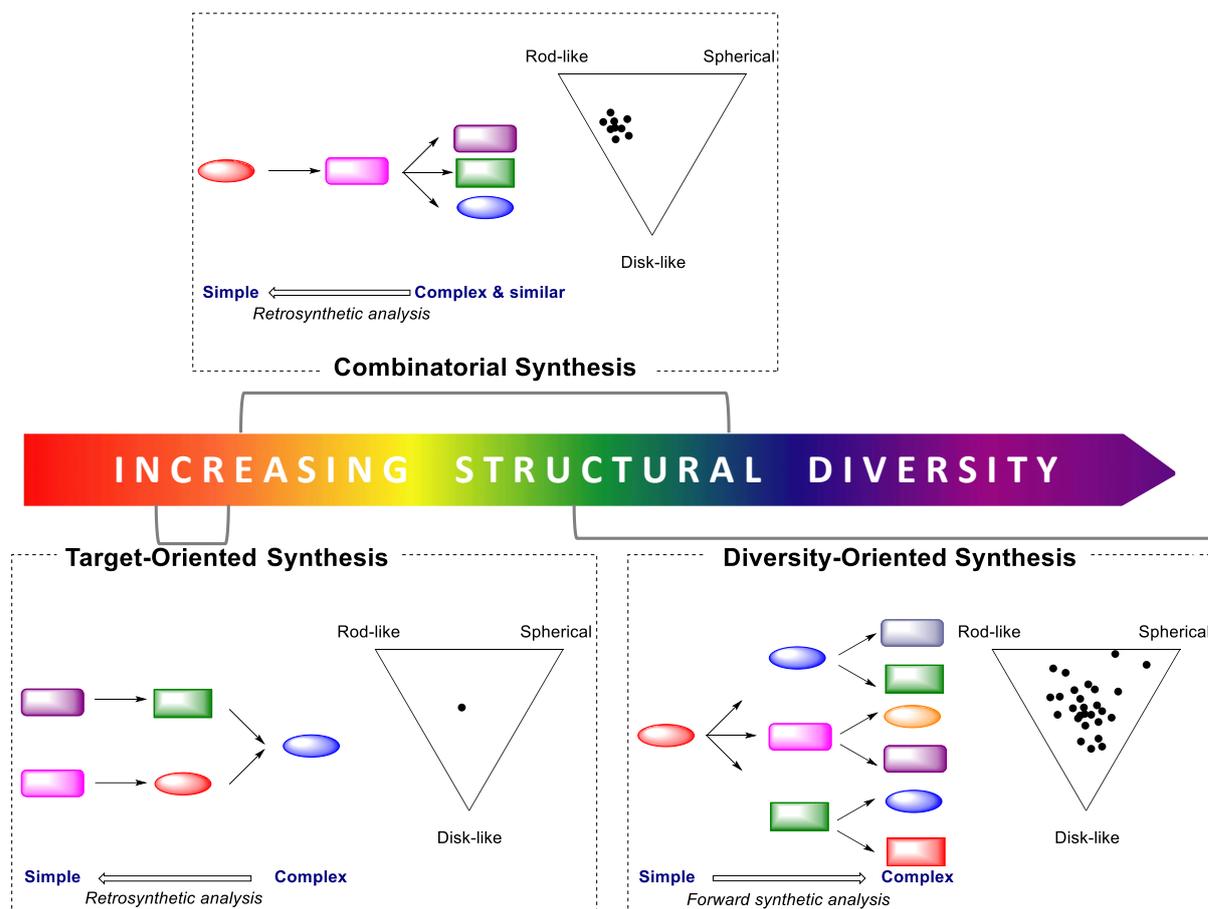
In addition to this structural bias, these libraries often contain a large percentage of scaffolds that are  $sp^2$  rich.<sup>48</sup> It has been noted that molecules featuring multiple aromatic groups have reduced solubility and bioavailability due to increased intermolecular interactions resulting from

$\pi$ -stacking.<sup>58</sup> Thus, the inherent physicochemical properties that stem from the molecular framework of these libraries can also be considered sub-optimal. Arguably, the ‘flat’ nature of these libraries also restricts the expanse of bioactive chemical space that we can explore. Studies undertaken by Schreiber *et al.* highlighted the importance of chirality, increased  $sp^3$  content within small molecules, and their ability to selectively bind to a panel of 100 different protein targets.<sup>59</sup> Molecules displaying such structural features (mostly those derived from Diversity-Oriented Synthesis and natural sources) positively correlated to an increased binding selectivity and frequency relative to commercially available sources. More importantly, an increase in the fraction of  $sp^3$  atoms has been shown to have a positive correlation with increased solubility, pharmacokinetic-profile and successful transition of molecules from discovery stage, through to clinical trials and FDA approval.<sup>60</sup>

In order to address the deficiencies of such screening collections, two strategies termed Diversity-oriented synthesis (DOS) and Fragment-based screening (FBS) emerged in the early years of this century offering alternative methodologies to develop novel chemical tools.<sup>53,61</sup> Often, molecules derived from DOS strategies resemble drug-like compounds with respect to their molecular mass and seek to populate novel and underexplored areas of chemical space;<sup>5,62</sup> in contrast, the FBS technique approaches this challenge by screening much smaller (<300 Da) molecules to increase the chemical space sampling. It is worth mentioning that in addition to these strategies the development of lead-oriented synthesis,<sup>63–66</sup> biology-oriented synthesis<sup>67–69</sup> and collaborations between industry, small/medium enterprises (SMEs) and academics such as the European Lead Factory (ELF)<sup>70</sup> in recent years has also provided tools for the generation of high-quality small molecule compound libraries.

### 1.3 Diversity-Oriented Synthesis

DOS emerged in the early 2000s<sup>71,72</sup> as a strategy to construct compounds that interrogate large regions of chemical space *via* the efficient and deliberate synthesis of different molecular scaffolds in a divergent process.<sup>73</sup> This involves the application of ‘forward’ synthetic analysis, in which common simple building blocks are transformed into a variety of structurally diverse products (Figure 1.2).<sup>5</sup> In this case, the most challenging but important concept is the generation of high levels of skeletal diversity within the library to facilitate increased chemical space sampling.<sup>74</sup> This opposes traditional target-oriented and combinatorial approaches, which follow linear and/or convergent sequences to generate a specific molecule or closely relating library occupying a defined area of chemical space.<sup>45,74</sup>



**Figure 1.2.** A comparison of target-oriented, combinatorial and diversity-oriented synthesis strategies for library generation and their respective molecular shape space distribution, depicted by principle moment of inertia plots. Adapted from references.<sup>9,40,75</sup>

The definition of a 'scaffold' and its diversity is a subjective concept, which can vary depending on the context of their use. Typically, however, the literature describes four types of molecular diversity that are commonly identified and include:

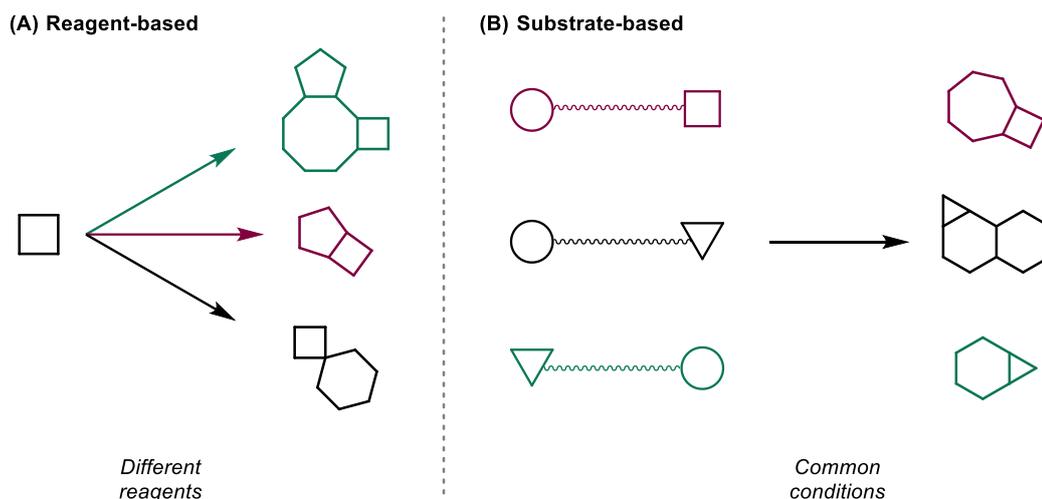
- 1) Skeletal diversity – different molecular architectures which result from a variation in ring structure/other rigidifying features, which determine the overall molecular shape.
- 2) Appendage diversity – variation in the building block and therefore substituents around a common scaffold.
- 3) Stereochemical diversity – variation in the spatial projection of different elements from the scaffold.
- 4) Functional group diversity – variation of the different functional groups present within a molecule that have the potential to interact with features present within the biological macromolecule in question.

Accordingly, a successful DOS library should incorporate all four elements to maximise the chemical space coverage.

### 1.3.1. Diversity-Oriented Synthesis approaches: reagent, substrate and multi-component-based methods

Commonly two methods to generate skeletal diversity are adopted within DOS.<sup>5,76</sup> The first involves a *reagent-based* approach whereby a common pluripotent building block is utilised in a branching pathway *via* the employment of varied reaction conditions and co-substrates that dictate the stereochemical and skeletal diversity that is produced. (Scheme 1.1, A).<sup>77</sup> To enable this, the building block usually features dense functionalisation of pluripotent handles to facilitate complexity generation.

The second strategy involves the employment of folding pathway, in which diverse starting materials featuring pre-encoded functionality are subjected to common reaction conditions to effect the scaffold formation (Scheme 1.1, B).<sup>78,79</sup> In this case, at least two appendages are encoded within the starting material, which in turn dictate the skeletal information when subjected to the common reaction conditions. Often, however, these approaches are not mutually exclusive and often modern DOS campaigns involve a combination of both strategies.<sup>5</sup>



**Scheme 1.1.** DOS can be achieved using several strategies to generate structural complexity and diversity. Adapted from reference.<sup>9</sup>

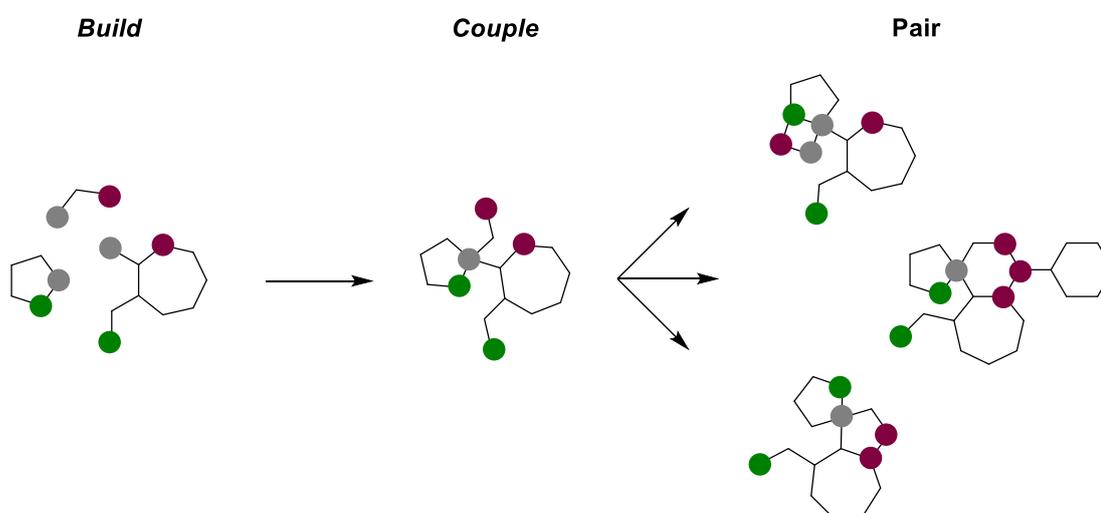
### 1.3.2. The Build/Couple/Pair algorithm

Independent of the DOS approach employed, a general method to describe the use of a systematic modular scheme was proposed by Nielsen and Schreiber, termed as the Build/Couple/Pair (B/C/P) algorithm (Scheme 1.2).<sup>80</sup> It was noted that DOS strategies to generate skeletally and stereochemically diverse products commonly adopt a three-stage process involving:

- 1) *Build Phase* – assembly of building blocks containing orthogonal functionalities, often including chiral motifs to introduce stereochemical diversity.

- 2) *Couple Phase* – intermolecular coupling reactions between the building blocks and other commercially available and/or readily synthesised molecules to obtain an array of molecules featuring additional functionality.
- 3) *Pair phase* – intramolecular coupling of different moieties within the precursors leading to the rigidified complex scaffolds. This stage is usually where the skeletal diversity is installed.

Recent advancements in the application of this strategy include the development of multi-dimensional approaches, whereby different linking motifs are utilised during the couple phase to increase the structural diversity of the respective library. This has been exemplified as an effective approach for the generation of novel macrocyclic SMs.<sup>81</sup>



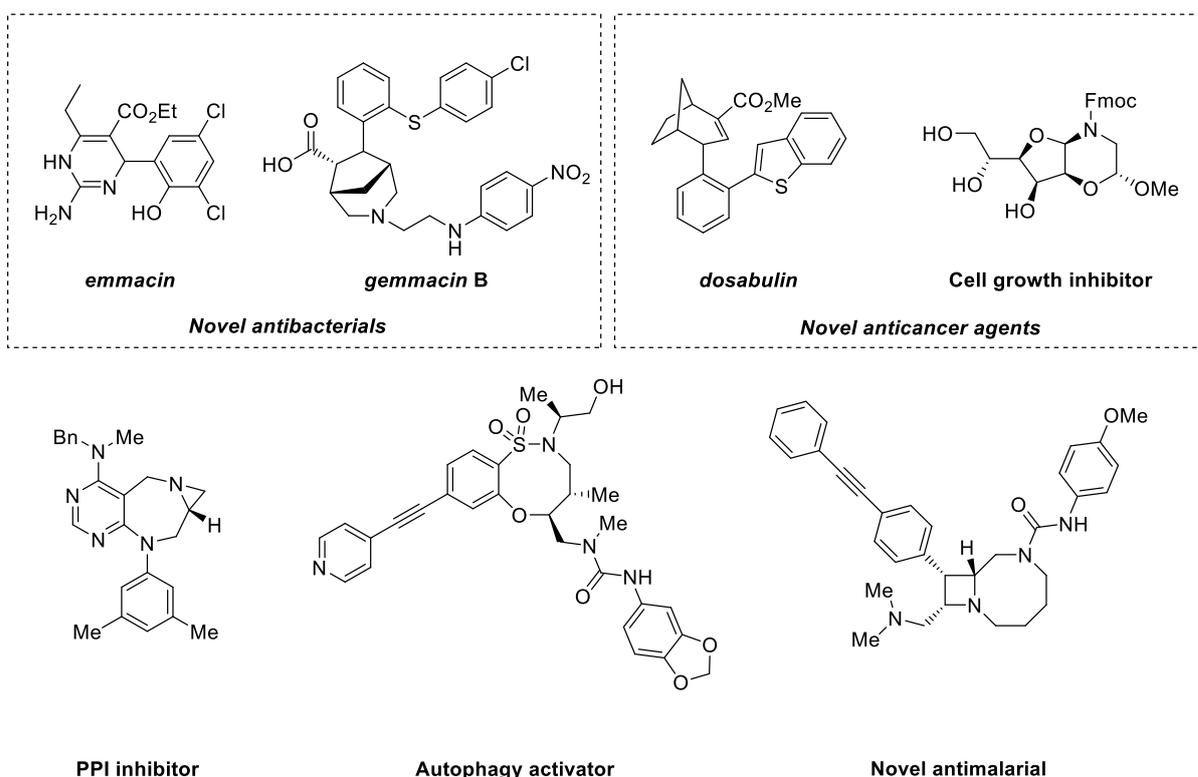
**Scheme 1.2.** The build/couple/pair strategy. Grey circles represent coupling functional group, green circles represent pairing functional group (non-polar) and purple circles represent pairing functional groups (polar). Adapted from reference.<sup>5</sup>

Typically, a DOS strategy is limited to five synthetic steps starting from the chosen building block in order to maximise efficiency, a concept which is key due to the large cost overheads associated with synthesis.<sup>82</sup> Thus, within a DOS approach it is desirable to utilise reactions that generate high levels of complexity and diversity. As such, strategies which feature multicomponent,<sup>83</sup> domino/tandem reactions,<sup>84,79</sup> and cycloadditions<sup>85,86</sup> have commonly been employed.

### 1.3.3. The success of DOS

Since its conception, DOS has had significant impact upon the biomedical community. Indeed, its vision of addressing screening collection deficiencies has instigated efforts towards library enrichment within pharmaceutical collections.<sup>61,87</sup> Due to the molecular size, compounds resulting from DOS are commonly utilised in high-throughput and phenotypic screening contexts,<sup>9</sup> vitally providing tools for the discovery of SMs in the absence of structural

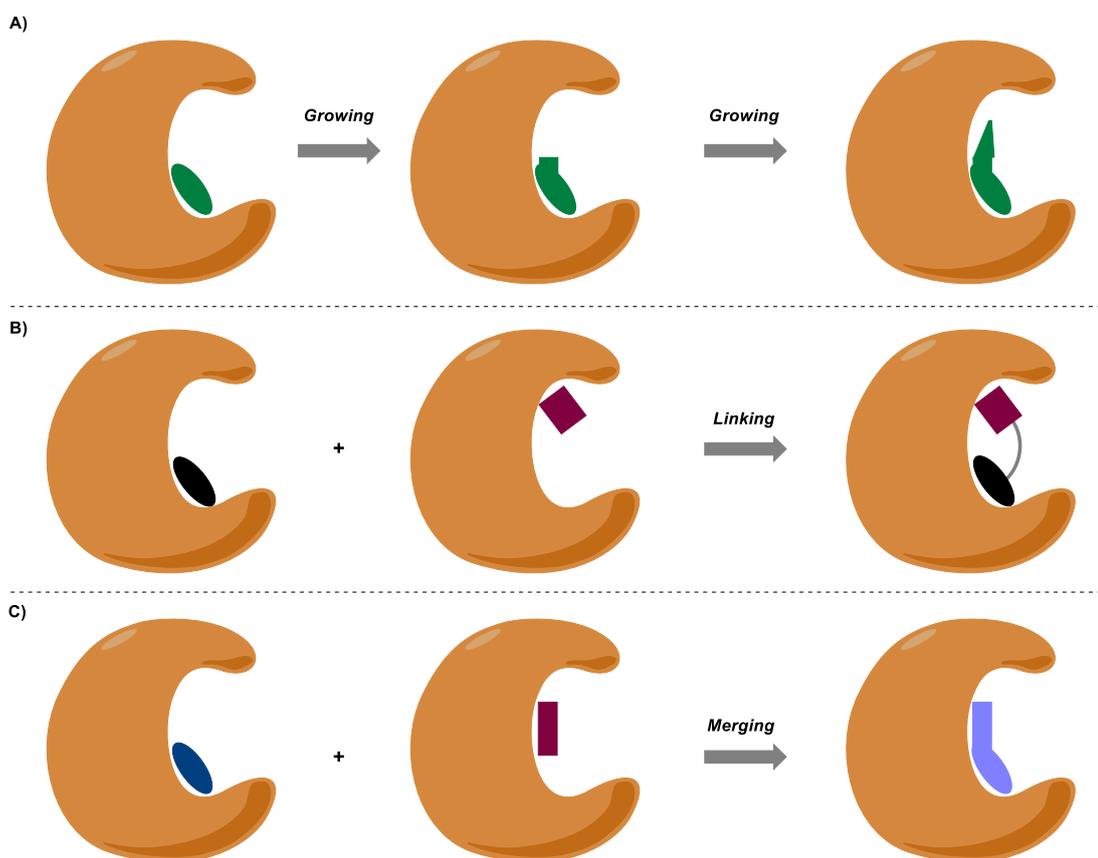
information or knowledge into the target responsible for eliciting a given phenotype.<sup>61</sup> Indeed, due to the diverse and 3-D nature of compounds created through DOS, these libraries have proven particularly successful in this regard. This includes the identification of compounds with a broad range of bioactivities, including new anticancer agents,<sup>88,89</sup> novel antimalarials,<sup>90,91</sup> and chemical probes such as modulators of autophagy<sup>92,93</sup> and  $\beta$ -cell apoptosis (Figure 1.3).<sup>94</sup> Importantly, molecules discovered through DOS have also been shown to interact with biological targets that have previously been difficult to modulate and are of increased biological relevance including protein-protein interactions (PPIs),<sup>95,96</sup> those with transcription factor activity,<sup>97</sup> drug resistant MRSA pathogens,<sup>98,99</sup> and those which activate membrane receptors (Figure 1.3).<sup>100</sup>



**Figure 1.3.** A selection of novel SM leads derived from DOS campaigns.

## 1.4 FBS

An alternative strategy for lead generation is the application of FBS. First described in 1996 by co-workers at Abbott,<sup>101</sup> this strategy relies on the screening of libraries of low molecular weight (<300 Da) compounds to identify weak ( $\mu$ M to mM affinity) ligands for a given target. In turn, these ‘fragment’ binders are grown,<sup>102</sup> linked<sup>103</sup> or merged<sup>104</sup> through an iterative process of design, synthesis and testing cycles to obtain potent lead compounds (Figure 1.4).



**Figure 1.4.** The three strategies commonly used in FBS to elaborate a hit. A) Fragment growth can be achieved utilising structural binding information to achieve potency through picking up additional interactions. B) If two fragments are bound within the pocket or adjacent pockets these two motifs can be chemically linked. C) If two fragments overlap within a binding pocket these entities can be merged to generate a more potent lead.

Whilst the basis of FBS appears counterintuitive for lead generation due to the weakly binding nature of the resulting hits, it is widely accepted that the screening of smaller molecules allows for a more efficient sampling of chemical space. Indeed, the number of possible 'fragment' chemical entities bearing 17 heavy atoms or fewer ( $10^{11}$ )<sup>105</sup> is considerably lower than the value possible HTS molecules with 30 heavy atoms or fewer ( $10^{63}$ ).<sup>49</sup> A direct consequence is that much smaller fragment libraries, commonly in the region of  $10^3$  members, are required to effectively sample 'fragment space'. In comparison, whilst HTS libraries are often in the order of  $10^6$  in magnitude, these can only fractionally cover the astronomical drug-sized space.<sup>106</sup> Crucially, due to this superior sampling significantly higher hit rates are generally observed with FBS campaigns compared to those conducted *via* HTS.<sup>107</sup>

This higher hit-rate can also be explained by considering the notion of *molecular complexity* and the probability of matching the binding site for a given target, a theory proposed in 2001 by Hann and co-workers.<sup>108</sup> According to this model, the larger and more complex a compound becomes, the greater the number of interactions between the molecule and a binding site. Thus, the window of opportunity of finding a hit becomes significantly narrower with more

complex molecules due to the reduced chance of complementarity. As a result, in addition to the enhanced chemical space sampling, it would be expected that smaller fragments with reduced complexity should also lead to a higher number of hits against a larger number of targets.<sup>109,110</sup>

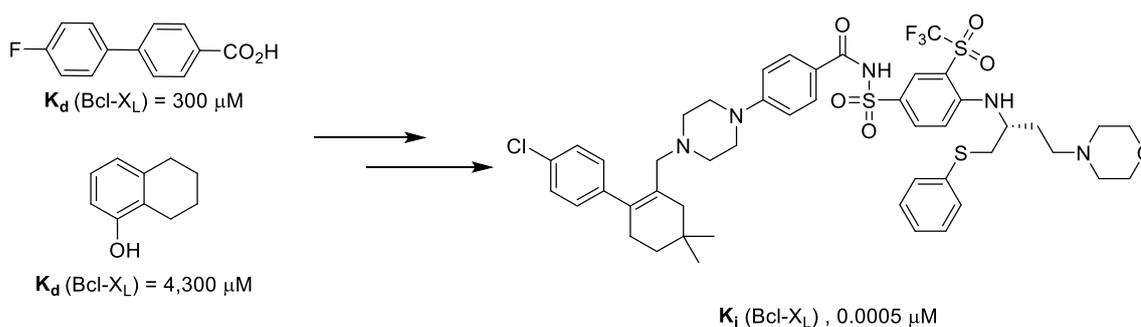
Another contributing factor toward the success of FBS, is the identification of efficient binders that form higher-quality interactions with their targets, despite their lower potency.<sup>111</sup> Often this factor is judged by the binding energy per atom, or ligand efficiency (LE).<sup>112</sup> Thus, fragment-sized molecules of higher or equal potency but with less atoms than their more drug-like counterparts can be considered more efficient binders.<sup>113</sup> Nevertheless, as a result of the reduced size and affinity, the challenge of hit identification is more problematic in FBS and biophysical methods of hit detection are often required.<sup>114,115</sup> Typically, this involves the use of X-ray crystallography, surface plasmon resonance (SPR) and NMR-based techniques due to their increased sensitivity.<sup>116,117</sup> Often, however, these results can provide ambiguity and commonly more than one technique is required to decipher true hits from artefacts, false positives and false negatives.<sup>118–120</sup>

One crucial consideration with FBS is the availability of structural information to leverage the power of structure-based design for hit elaboration.<sup>110,121,122</sup> Indeed, the utilisation of structural information has been shown to dramatically increase the success of obtaining potent inhibitors from 33% to 93%.<sup>123</sup> Nevertheless, the availability of crystal structures is not a trivial pursuit and the viability of this process often remains highly dependent on the nature of the target and the conformation of the protein in its crystal form. Thus, alternative approaches to guide hit optimisation have been explored, including NMR and computational modelling methods<sup>124</sup> and examples of successful campaigns devoid of structural information have emerged.<sup>125</sup>

In order to retain the advantages of FBS, the assembly of suitable screening libraries remains key. In 2003, a series of guidelines were proposed by co-workers at Astex termed the 'Rule of Three' (RO3, molecular weight <300 Da, clogP <3 and number of H-bond donors/acceptors <3), derived from analysis of the physicochemical properties of successful fragments.<sup>126</sup> Among these, the most important descriptor is considered to be molecular size, often also calculated by the number of non-hydrogen atoms (HAC), with fragments typically considered to contain 20 or less in order to maintain the efficiency of chemical space sampling.<sup>110</sup> Whilst recently smaller fragments have tended to be favoured (HAC < 16),<sup>115</sup> the utility of fragments at the upper limit of fragment space (~300 Da) has been demonstrated, for example in the development of inhibitors of challenging targets such as PPIs.<sup>127</sup> Finally, due to the low affinity of fragment binders, solubility is also a key parameter to consider in order to enable high-concentration screening.<sup>122</sup>

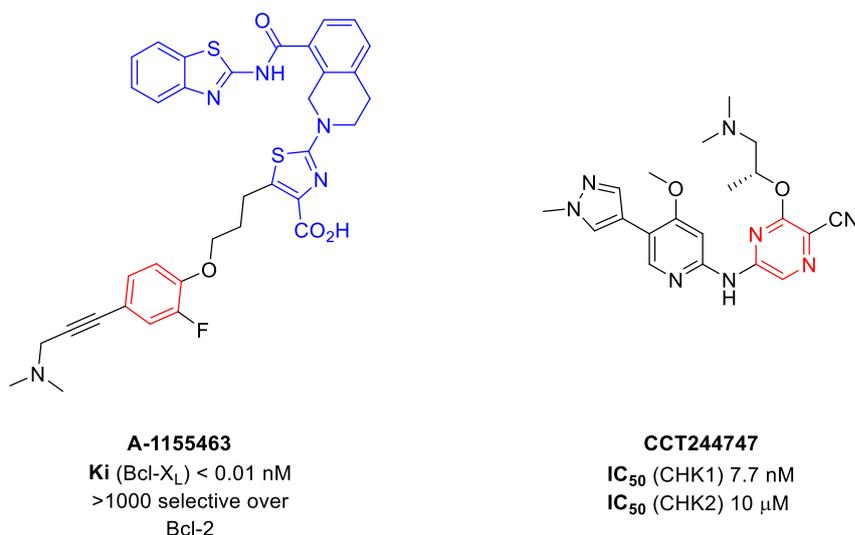
### 1.4.1 The success of FBS

Since its emergence over two-decades ago FBS has been widely adopted throughout industry and academia.<sup>128,129</sup> In this timeframe, this technique has proven particularly effective delivering two marketed drugs (Vemurafenib and Venetoclax)<sup>130,131</sup> and more than 30 clinical candidates.<sup>110,132</sup> Importantly, amongst these lie inhibitors of challenging targets, such as members of the Bcl-2 family and BACE-1,<sup>133</sup> which previously proved intractable with HTS approaches.<sup>111,134</sup> One such example is the development of a novel mode of action candidate, Navitoclax, a dual Bcl-2 and Bcl-X<sub>L</sub> inhibitor Navitoclax, which disrupts intracellular PPIs between other Bcl-2-family members (Figure 1.5).<sup>135</sup> At the other end of the size spectrum, owing to the smaller initial hits, FBS has also provided clinical candidates with reduced molecular weight and lipophilicity, characteristics which positively correlate with reduced attrition.<sup>136,137</sup> The ability of FBS to provide both large molecule (PPI inhibitors) and small molecules is one of the reasons why this approach has seen a dramatic development in the last decade.



**Figure 1.5.** The initial fragment hits and final candidate against Bcl-X<sub>L</sub> developed by co-workers at Abbott adapted from reference.<sup>111</sup>

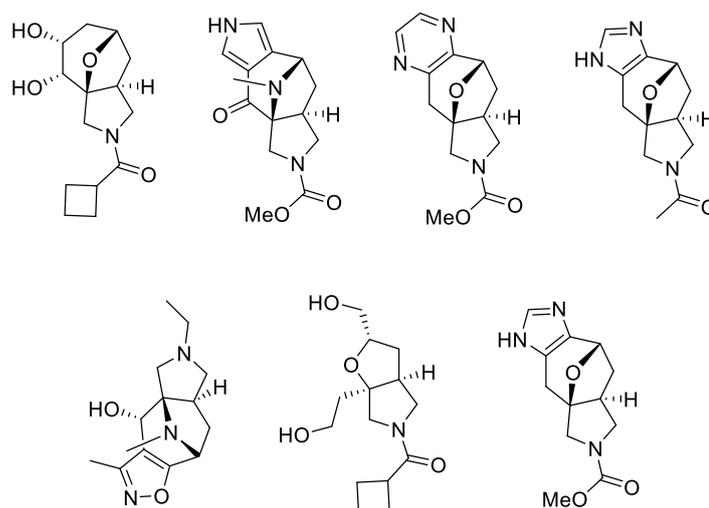
In the context of chemical probe discovery FBS approaches have also been highly successful, enabling the development of selective modulators for a wide range of targets.<sup>138–141</sup> Among the many examples, notably this includes the development of the Bcl-X<sub>L</sub>-selective chemical probe A-1155463<sup>142</sup> and the CHK1-selective compound CCT244747 (Figure 1.6).<sup>143</sup> In both cases, these targets had previously only explored in a dual-inhibitory manner, importantly providing tools to elucidate the associated biology of each protein.



**Figure 1.6.** Two examples of chemical probes developed by FBS methods. The two core components of A-1155643 were developed by combining HTS (portion in blue) and FBS (portion in red) strategies. The original fragment hit constituent of CCT244747 is highlighted in red.

### 1.4.2 Outstanding challenges

Despite these successes, in a similar vein to HTS screening collections, organic synthesis has been identified as a significant hurdle within FBS. More specifically, this relates to the lack of fragment space coverage by commercial collections, in addition to an overreliance on sp<sup>2</sup>-rich fragments of limited shape diversity.<sup>61,144,145</sup> Despite the potential for 3-D fragments to display lower hit rates when considering Hann's view of molecular complexity (*vide supra*), complementary tools to existing collections are required to interrogate challenging target classes with low tractability.<sup>144</sup> Recent work by Foley *et al.* indeed validated this theory through the screening of 52 3-D fragments against three epigenetic targets from two mechanistic target classes (BRD1, JMJD2 and ATAD2), resulting in the identification of 17 novel binders.<sup>146</sup> Importantly, this included seven hits against ATAD2, a novel oncology target previously deemed to have low druggability (Figure 1.7).<sup>147</sup> Nevertheless, whilst the growing demand for novel 3-D fragment architectures has received attention,<sup>148–150</sup> these scaffolds remain underrepresented.

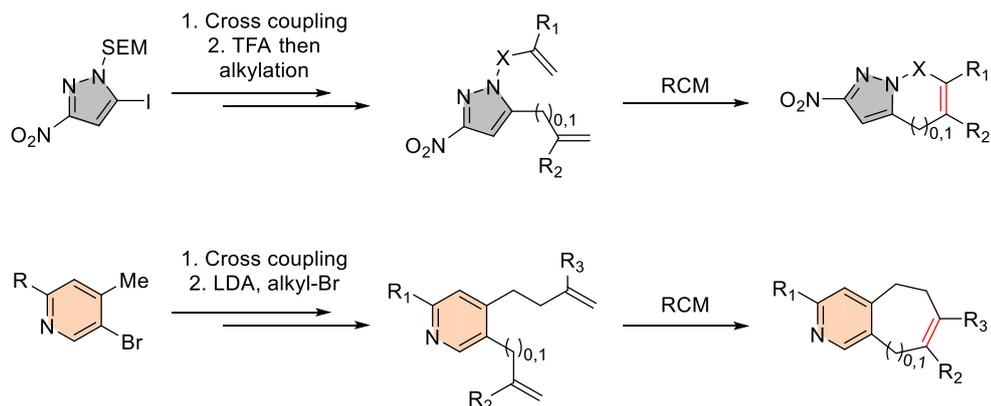


7 hits against ATAD2 from  
a library of 52 3-D Fragments

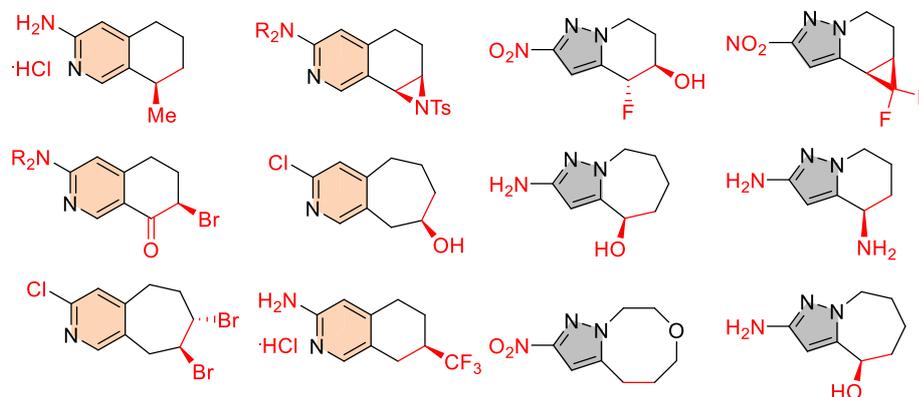
**Figure 1.7.** Screening of diverse 3-D fragments enabled the identification of novel binders against a challenging epigenetic target.

In addition to their utility against challenging targets, sp<sup>3</sup>-rich fragments also offer the prospect of increased numbers of fragment growth vectors. This has been noted as another significant deficiency of commercial libraries,<sup>115,121</sup> yet the incorporation of synthetic handles remains a crucial factor in enabling rapid fragment elaborations without the requirement of new synthetic route development. Thus, calls from within the scientific community for novel fragments featuring multi-directional exit vectors with synthetic tractability have been made. One example of research meeting this criteria was conducted within the Spring group by Twigg *et al.*<sup>151</sup> describing the development of a modular route to partially saturated bicyclic aromatic (PSBH) fragments (Figure 1.8). This work resulted in the formation of 15 fragments based on five scaffolds, varying in saturated ring size and heteroatom content in addition to the adjacent heterocycle. Furthermore, an alkene and either an amino, nitro or chloro handle were incorporated as secondary points for scaffold diversification.

A) Modular routes to PSBH fragments



B) Demonstration of the alkene exit vectors



**Figure 1.8.** The partially saturated bicyclic heteroaromatic library presented by Twigg *et al.* featuring enhanced  $sp^3$  character and synthetic handles for fragment elaboration. LDA = Lithium diisopropylamide, TFA = Trifluoroacetic acid, RCM = Ring closing metathesis.

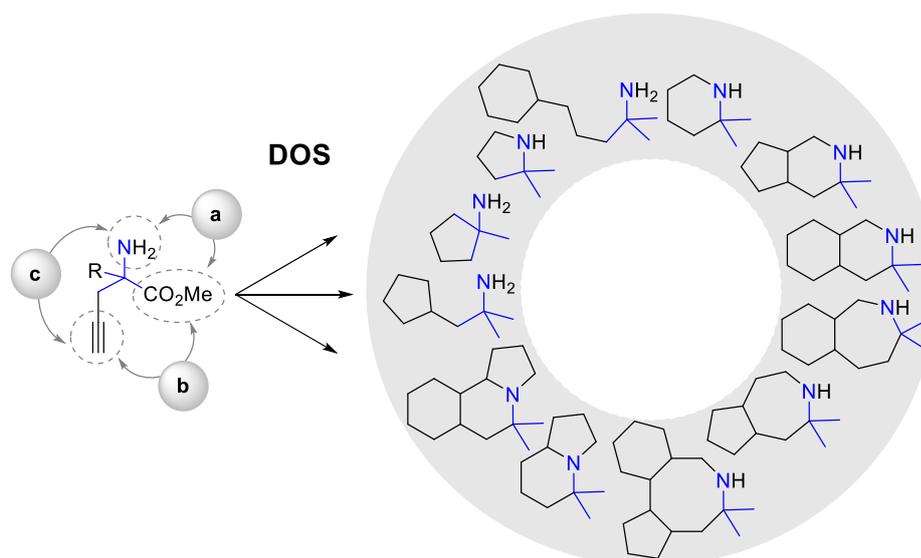
However, only a handful of additional examples seeking to tackle this issue have been published in recent years.<sup>152,153</sup> Thus, the development of methodologies seeking to generate many different scaffolds *via* a single strategy to rapidly populate fragment space whilst maintaining synthetic tractability would be highly beneficial. Moreover, whilst notable examples of strategies to access novel  $sp^3$ -rich fragments have been reported, the biological relevance of such libraries remains underexplored. Consequently, new research into the utility of such collections is required.

Another reason to generate fragments bearing readily functionalisable handles is that, due to the low affinity and sometimes ambiguous nature of initial hits resulting from X-ray screening campaigns, access to analogues is often required to probe the binding interactions.<sup>121</sup> Often this can be achieved using commercially available derivatives, however, this aspect can be more difficult with novel libraries. Moreover, within drug discovery the balance of compound value with resource investment is an important consideration.<sup>154</sup> Thus, the early stages of hit validation and elaboration not only presents a synthetic, but also an economic challenge.

## 1.5 Thesis overview and aims

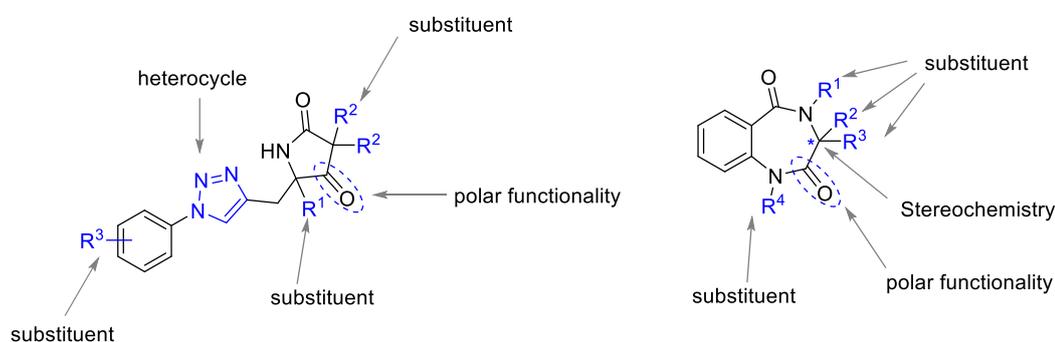
Traditionally, DOS and FBS have been regarded as competing technologies, but in recent years examples of the intersection of these two fields have blossomed within the literature.<sup>146,155–158</sup> Moreover, direct calls from leaders in the field have encouraged the development of fragment-focused DOS to address the aforementioned deficiencies with commercial screening libraries.<sup>121</sup> With these points in mind, this dissertation describes how DOS approaches can be leveraged to tackle these important outstanding challenges within the field.

**The design and synthesis of novel,  $sp^3$ -rich and diverse fragment libraries to complement existing screening collections.** As discussed in this chapter, there remains a need for the development of novel chemotypes to expand the FBS toolbox. Thus, Chapter Two describes the application of DOS methodology for the construction of a novel *N*-substituted quaternary carbon-containing screening library utilising  $\alpha,\alpha$ -disubstituted amino ester building blocks (Scheme 1.3). This resulted in the generation of a fragment collection based on 40 distinct scaffolds incorporating this important and currently underrepresented 3-D motif. In addition, modifiable groups were installed at a range of positions to enable multidirectional fragment growth. Computational analysis was conducted to demonstrate the high levels of diversity achieved, the importance of the quaternary centre and the applicability of this library toward FBS. Moreover, the collection featured a higher number of chiral centres and lower fraction aromatic compared to commercial collections.



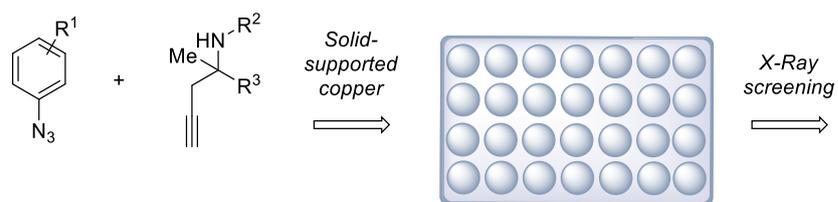
**Scheme 1.3.** DOS starting from  $\alpha,\alpha$ -disubstituted propargyl amino esters.

**Demonstration of the utility of DOS fragment libraries for enabling the discovery of chemical leads against challenging targets and structure-activity relationship exploration from multiple vectors around a ligand.** Since DOS libraries offer increased structural diversity, and therefore chemical space sampling, these collections should be well-suited to identify hits against challenging targets. Accordingly, Chapter Three describes the implementation of the quaternary DOS collection within multiple X-ray based fragment screening campaigns, illustrating the biological relevance through the identification of hits against several novel targets. Importantly, we also demonstrate the advantages of the inherently modular and flexible nature of DOS for enabling the construction of rationally designed analogues. This included the exploration of several 3-D fragment growth vectors, including the quaternary centre, to enable hit validation and probe the binding (Figure 1.9).



**Figure 1.9.** An example of the exit vectors and modifiable groups that were used as synthetic handles to derivatise the molecules in a 3-D manner.

**The development of novel methodologies for rapid analogue generation coupled to a structural-binding readout to accelerate the evolution of weak hits to potent lead compounds.** Due to the weakly binding nature of fragment hits, the evolution process of initial hits to achieve compounds with activity in solution assays remains a significant hurdle. Thus, the development of novel methodologies to accelerate this process would have significant utility for FBS. Accordingly, Chapter Four describes the development of a novel high-throughput strategy to enable rapid analogue construction *via* the formation of 1,2,3-triazoles utilising liquid handling robotics (Scheme 1.4). Firstly, we demonstrate the compatibility of the DOS chemistry both with this technology and solid-phase catalysis, enabling the formation of 78 analogues. Importantly, it was envisaged this would facilitate widespread adoption of the DOS library within a FBS setting. Secondly, these studies formed the basis of preliminary investigations toward the implementation of crude reaction mixtures within an X-ray screening platform to provide a direct structural-binding readout to enable rapid structure-binding relationship data.



Combinatorial robotic synthesis  
in 96 well format

**Scheme 1.4.** Using liquid handling robotics analogues were constructed in a high-throughput manner to investigate crude X-ray screening.

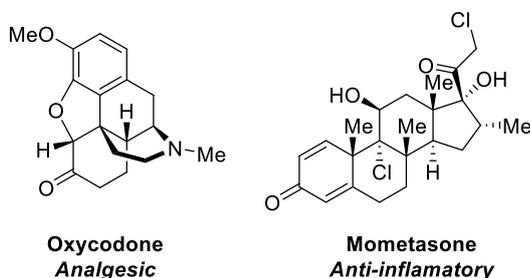


## 2 Results and discussion: DOS for the synthesis of *N*-substituted quaternary small molecules

### 2.1 Background

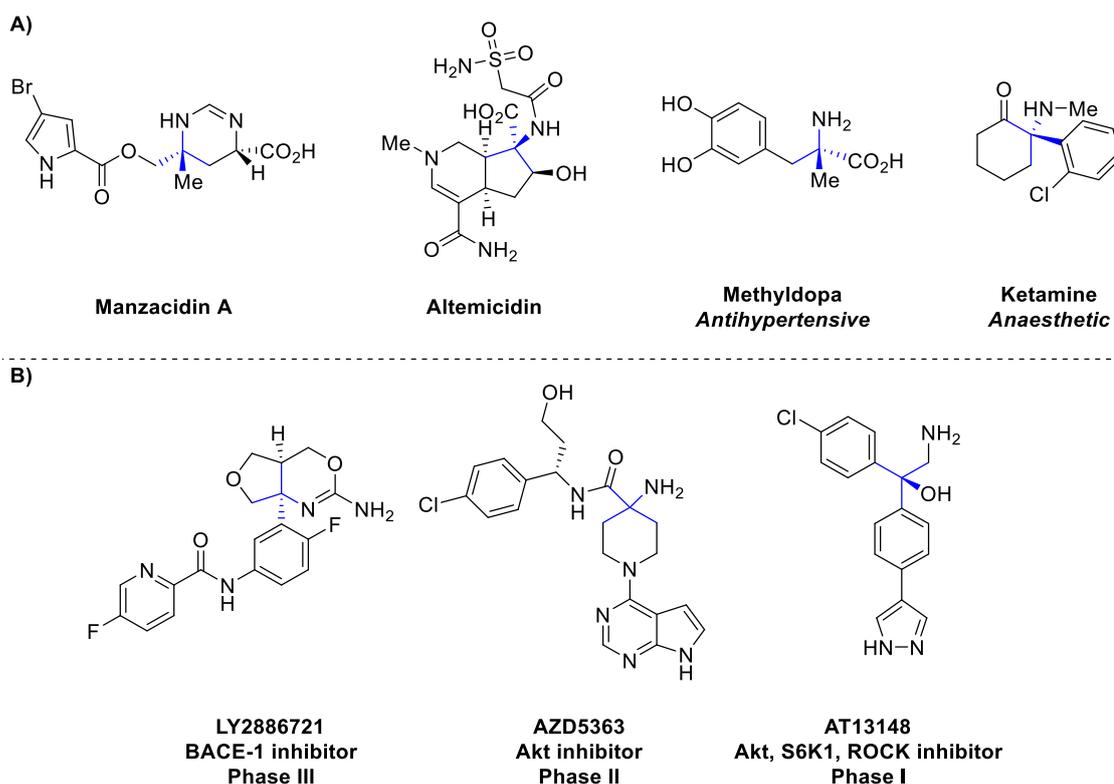
#### Strategies for the synthesis of *non-flat* fragment DOS libraries

Among the different methodologies described in the literature to generate complex small molecules, the introduction of restricted elements such as cyclic structures or quaternary  $sp^3$  carbons has received significant interest due to their increased 3-dimensionality and conformational restriction.<sup>159</sup> Certainly, the construction of quaternary carbon motifs has represented an important but nevertheless challenging quest for synthetic chemists in recent decades, forming the focus of many natural product and drug molecule endeavours.<sup>160,161</sup> The all-carbon variant of this motif represents an important structural feature within many marketed drugs,<sup>162</sup> such as Oxycodone and Mometasone (Figure 2.1). Nevertheless, as a result of the difficult nature of constructing this congested centre, naturally-derived precursors are commonly employed in their syntheses.<sup>159,163</sup> Thus, over recent years efforts toward the development of novel methodologies to engender all-carbon quaternary centres has surged.<sup>164–167</sup>



**Figure 2.1.** Quaternary carbon-containing small molecule drugs.

In addition to all-carbon quaternary centres, *N*-substituted quaternary carbons (NSQCs) represent an equally challenging but crucial variant of this motif.<sup>168–170</sup> Indeed, the conformational restriction imparted upon  $\alpha,\alpha$ -disubstituted amino acids has been demonstrated as an effective biomimetic strategy for modification of the properties of bioactive peptides. Thus, novel strategies concerning the synthesis of non-proteinogenic amino acids have received significant attention in recent years.<sup>171–173</sup> Moreover, the presence of NSQCs within natural products such as Manzacidin A, Altemicidin<sup>174,175</sup> and the essential medicines Methyldopa and Ketamine<sup>176</sup> (Figure 2.2) serves to highlight the relevance of this motif within a pharmaceutical context.



**Figure 2.2.** Examples of NSQCs in A) Natural products and essential medicines and B) FBS-derived clinical candidates.

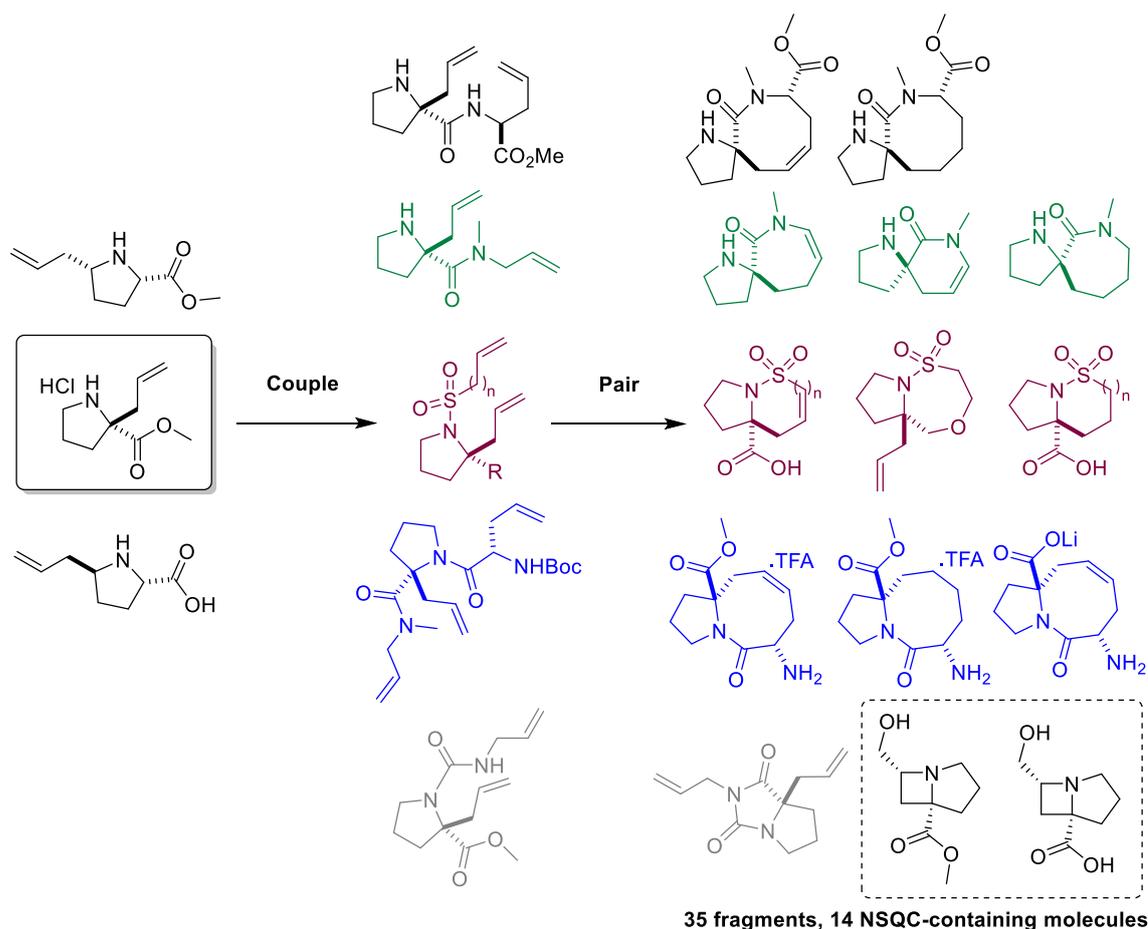
More specifically to the FBS paradigm NSQCs have been demonstrated as crucial elements within several FBS-derived clinical candidates (Figure 2.2), with seven out of the thirty-two in trials during 2016 containing this motif.<sup>110</sup> Five of these candidates, such as LY288682, relate to BACE-1 inhibition for Alzheimer's disease therapy. In this context, it has been shown that the incorporation of a NSQC within such molecules can be vital to facilitate the projection of substituents into different pockets within the protein to enable key binding interactions, in addition to improving compound stability.<sup>177–179</sup> Furthermore, examples such as the Akt inhibitors AZD5363 and AT13148 have additionally highlighted the potential of NSQC functionalities to mediate vital *H*-bond interactions, contributing to the Akt inhibitor pharmacophore.<sup>180–184</sup> Moreover, the inclusion of molecules such as Ketamine within the Astex 2012 Fragment Library<sup>145</sup> highlights the relevance of NSQC-containing sp<sup>3</sup>-rich compounds within fragment screening collections.

Despite their importance, however, due to the challenging nature of constructing this motif, NSQC-containing SMs remain underrepresented in screening collections. Furthermore, to date many examples of the incorporation of NSQCs within SMs have focused on Target-Oriented Synthesis (TOS) strategies, resulting in libraries of limited structural diversity.<sup>175,185</sup> With these points in mind, the generation of flexible synthetic routes toward diverse, fragment- and lead-like scaffolds that incorporate NSQCs are highly desirable. Thus, in recent years a number of

examples have emerged within the literature regarding the use of DOS methodology to construct screening libraries incorporating this structural feature.

### **A) DOS starting from allyl proline derivatives**

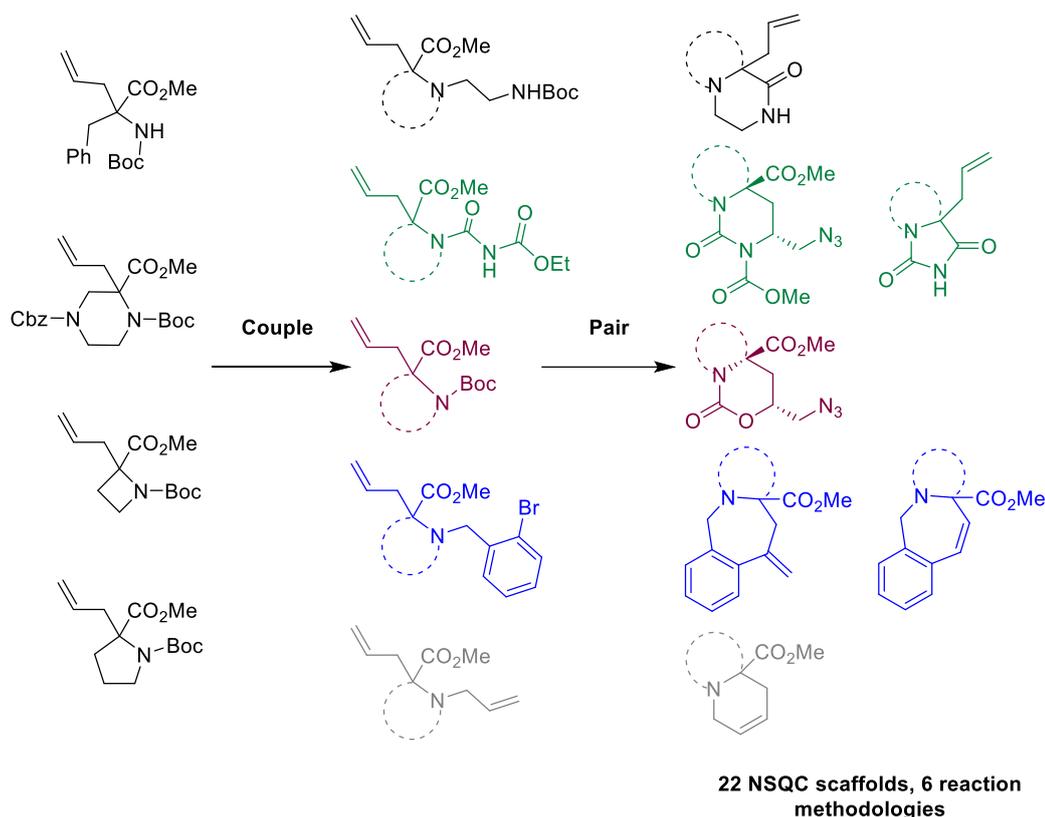
Allyl-containing cyclic amino esters building blocks represent one strategy utilised within DOS campaigns to generate NSQC-containing fragments. A seminal paper published by Young *et al.* in 2011 described the first application of DOS toward the synthesis of  $sp^3$ -rich fragments utilising this approach. In this research, a 35-member library was constructed utilising proline-derived cyclic building blocks in a B/C/P algorithm (Scheme 2.1).<sup>186</sup> A division of this work involved the employment of a quaternary allyl-proline substrate, which ultimately delivered a subset of the library comprising 14 NSQC-containing fragments. This approach involved the installation of a second allyl olefin *via* *N*- and ester- functionalisation to generate six intermediates, which were intramolecularly cyclised by RCM, hydantoin formation or oxo-Michael reactions to yield a series of fused and spiro bicyclic compounds. Furthermore, following a previously reported procedure,<sup>187</sup> two quaternary bicyclic azetidines were additionally synthesised from the same building block to enhance the scaffold diversity. Finally, members of the fragment library were additionally modified in a post-pair phase manner *via* functional group interconversion to modify the electronic properties and increase  $sp^3$  content. Importantly, the authors demonstrated the adherence of the full library to the RO3 principles and additional cheminformatic analysis confirmed 3-D nature and diversity of the compounds.<sup>186</sup>



**Scheme 2.1.** The 14 compounds containing a NSQC were included in the fragment library, 12 of which were synthesised from a single quaternary proline building block.

## B) DOS starting from $\alpha,\alpha$ -disubstituted amino acid-related scaffolds

The second example in this field was reported in 2015 by Foley *et al.* (Scheme 2.2) describing the use of four Boc-protected  $\alpha,\alpha$ -amino acid derived building blocks in a matrix-like B/C/P approach.<sup>188</sup> In this case, the four quaternary building blocks were diversified *via* *N*-alkylation or urea formation to produce five classes of trifunctional intermediates. Subsequent pairing of these functionalised substrates could then be achieved using six reaction methodologies to generate twenty-two NSQC-containing scaffolds. Notably, the resulting compounds featured medically relevant functionalities such as ureas, hydantoins and lactams, in addition to functional synthetic handles. Furthermore, the resulting library proved to possess a high degree of novelty, whilst subsequent virtual enumeration afforded a library of 1,110 lead-like 3-D compounds, with several examples meeting the RO3 criteria.

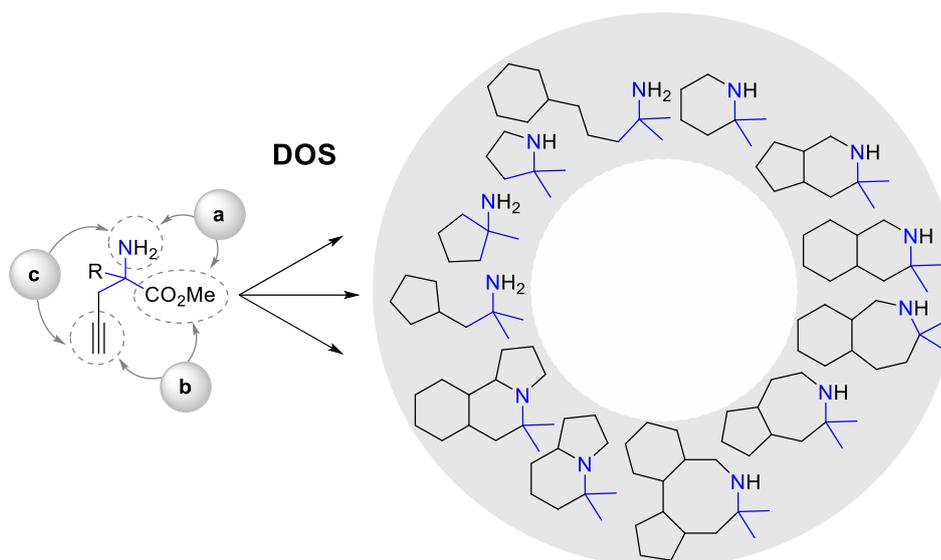


**Scheme 2.2.** Using four  $\alpha$ -allyl substituted amino acid derivatives 22 NSQC scaffolds were efficiently constructed in a matrix B/C/P fashion.

Nevertheless, in both instances the biological utility of these NSQC-containing fragments remains unexplored. Accordingly, we envisaged the generation of a novel NSQC library would provide the potential to the opportunity to demonstrate the biological relevance of 3-D fragments featuring this motif.

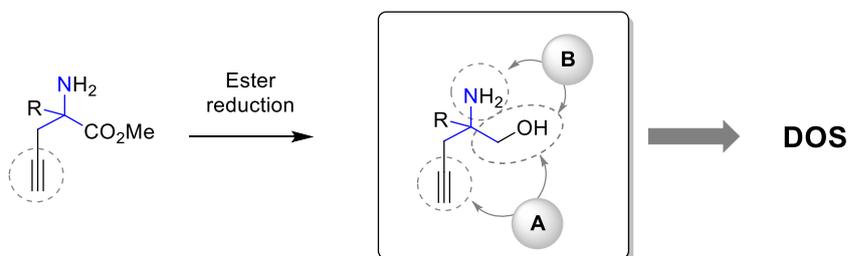
## 2.2 Project proposal

Inspired by the successful reports on the utility of cyclic allyl-based quaternary substrates for the synthesis of diverse  $sp^3$ -rich libraries, we hypothesised that acyclic propargyl-based  $\alpha,\alpha$ -disubstituted amino esters could serve as an orthogonal and novel approach for this purpose. Owing to the acyclic nature of these moieties, it was envisaged that the differential reactivity of the three synthetic handles (the alkyne, amino and ester functionalities) could be exploited to generate extensive scaffold diversity. Importantly, it was hoped that the development of modular DOS methodology in this context would be well-suited to address the aforementioned challenges within FBS. More specifically this related to the incorporation of multiple modifiable substituents for exit vector exploration, in addition to a versatile quaternary substituent for 3-D fragment growth. Preliminary studies in this regard were conducted in the Spring group by Dr. Natalia Mateu (Scheme 2.3).



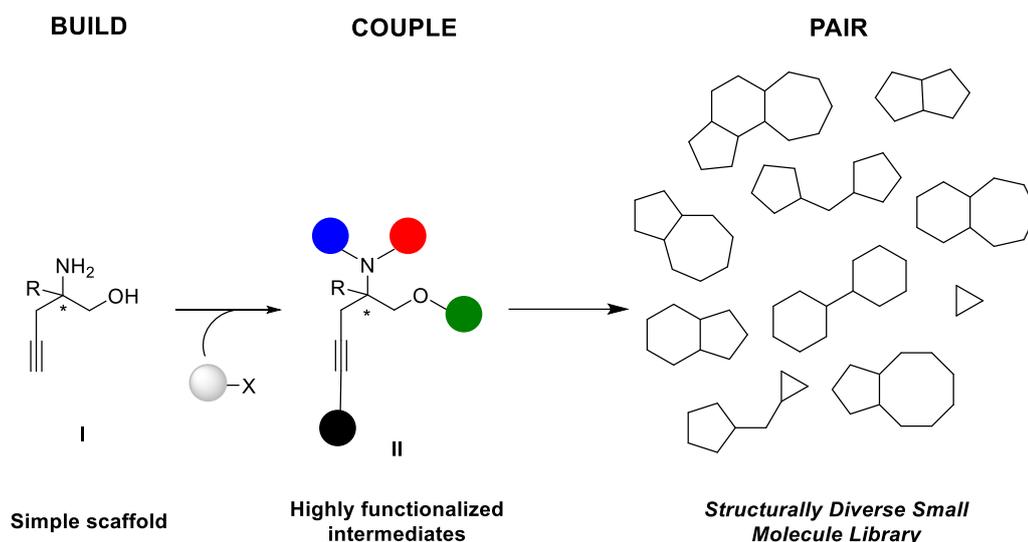
**Scheme 2.3.** DOS starting from  $\alpha,\alpha$ -disubstituted propargyl amino esters.

Stemming from this work, we hypothesised that an interesting complementary extension could involve utilising simple ester reduction to furnish a densely functionalised quaternary propargyl amino alcohol building block. Consequently, it was envisioned incorporation of this hydroxyl moiety would enable expansion of the scope of chemistries that could be employed within the DOS strategy (Scheme 2.4). In this instance, it was anticipated that pairing of the electrophilic alkyne (Scheme 2.4, pathway A) or the nucleophilic hydroxyl with a functionalised amino functionality (Scheme 2.4, pathway B) would enable access to 3-D oxygen-linked constrained fragments.



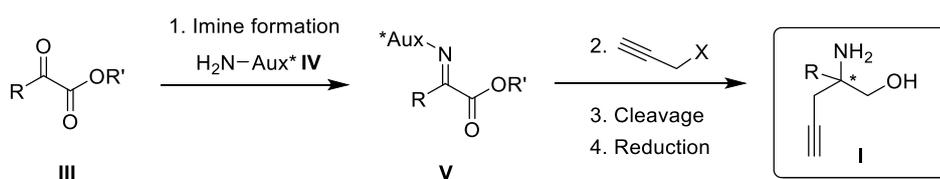
**Scheme 2.4.** The pluripotent nature of the quaternary propargylic amino alcohol in DOS.

With this concept in mind, a general strategy for the construction of the new library was proposed following a B/C/P format. In this manner, it was envisioned the starting amino alcohol building block(s) **I** (*build phase*) could be transformed into a variety of highly functionalised linear precursors **II** through varying both the linking motifs and substrates used (*couple phase*) (Scheme 2.5). Subsequent pairing of the resultant structures through intramolecular rigidification of the linear intermediates would afford multiple complex and diverse scaffolds in a straightforward manner (*pair phase*).



**Scheme 2.5.** DOS of a 3-D fragment library following a B/C/P approach.

Initially, a racemic strategy to generate **I** was proposed owing to the fact that fragment racemates are often preferred within FBS campaigns to enable the screening of both enantiomers.<sup>115</sup> It was envisaged that the first stage of the project would focus on the construction of building block **I** in a four step procedure starting from the  $\alpha$ -keto ester **III** (R = alkyl or aryl) as follows: (1) imine **V** formation; (2) regioselective nucleophilic addition of the propargylic chain; (3) protecting group cleavage and (4) reduction of the ester functionality (Scheme 2.6). Another important aspect of this project was the versatile nature of the proposed synthetic route, allowing for modification of the substituent at the quaternary stereocenter simply through selection of the desired ketoester **III**. Accordingly, this would enable the incorporation of a cohort of substituents at this position to generate additional diversity and facilitate further fragment growth.

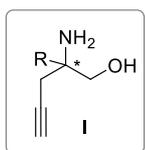


**Scheme 2.6.** The proposed route to building block **I**. Aux = auxiliary.

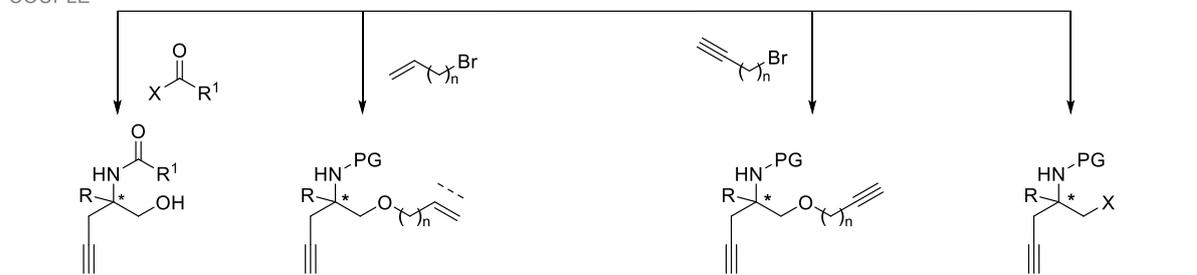
Inspired by the dense functionalities embedded within **I**, a divergent branching scheme for the DOS strategy was proposed (Scheme 2.7). This involved the generation of multiple linear precursors using a variety of transformations including ether synthesis, hydroxyl substitution and amine acylation. In the subsequent *pair phase*, it was anticipated that the employment of a wide range of complexity-generating and multi-component-coupling reactions would generate high levels of scaffold diversity across the library. This included ruthenium-catalysed intramolecular enyne metathesis (RCEYM), cobalt-catalysed [2+2+2] cyclotrimerisations, 1,3-dipolar cycloaddition reactions, and base-mediated cyclisations. Consequently, different

atomic skeletons such as morpholinones, benzoxepanes, aziridines and dihydropirans bearing a quaternary  $sp^3$  carbon would be constructed from a single common precursor.

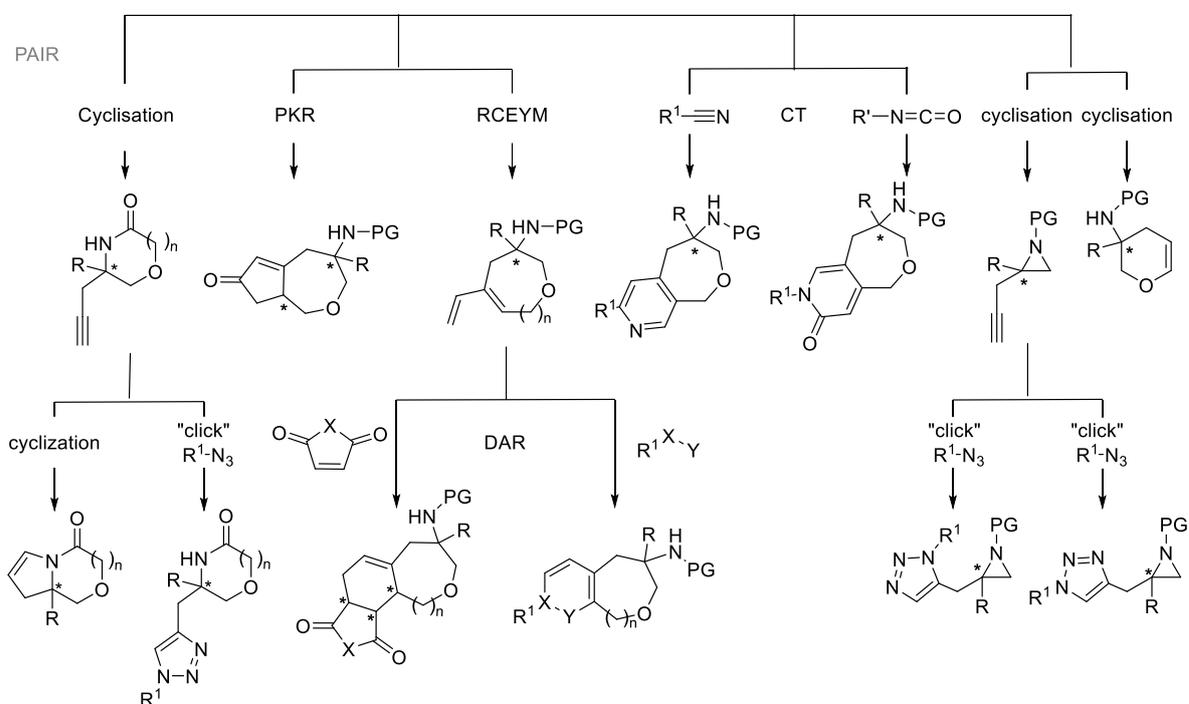
BUILD



COUPLE



PAIR

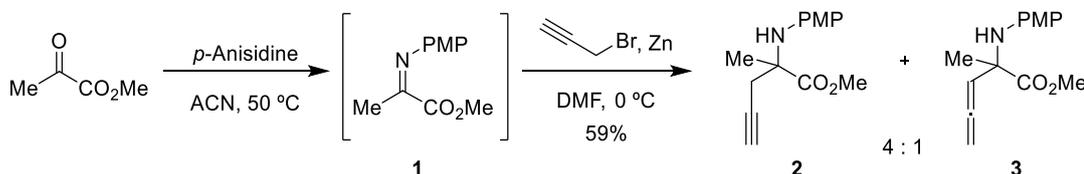


**Scheme 2.7.** The proposed DOS strategy starting from **I**. \* = stereocenter, PG = protecting group, PKR = Pauson Khand reaction, RCEYM = ring closing enyne metathesis, CT = [2+2+2] cyclootrimerization, DAR = Diels-Alder reaction.

## 2.3 Library synthesis

### 2.3.1 Synthesis of the key amino alcohol building block I

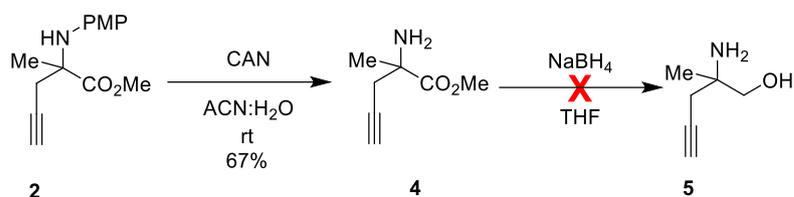
In line with the project proposal, initial investigations began with the racemic synthesis of the key amino alcohol intermediate **I** bearing a methyl substituent at the quaternary stereocenter. In this manner, construction commenced with the formation of imine **1** through condensation of the commercially available methyl pyruvate and *p*-anisidine (Scheme 2.8). However, despite previous reports on the formation of *p*-methoxy phenyl (PMP)-imino esters<sup>189–191</sup> we observed that overheating the reaction and purification of the intermediate imine **1** led to a considerable reduction in the yield. Instead, it was found that a simple work-up (*via* trituration and filtration) was sufficient to isolate the crude imine, which was then immediately subjected to the next reaction conditions without further purification.



**Scheme 2.8.** The synthetic route to amino ester **2**.

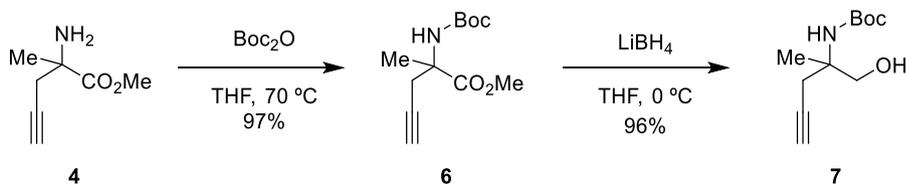
Treatment of **1** with propargyl bromide in the presence of Zn under Barbier-type conditions facilitated the regioselective formation of the key quaternary centre.<sup>192,193</sup> As a result of isomerisation<sup>194–196</sup> alkyne **2** was obtained together with allene **3** in a 4:1 ratio, respectively (determined by <sup>1</sup>H NMR), however both isomers were separable by column chromatography. The scalability of this synthetic procedure was demonstrated by the preparation of a 12 g mixture of **2** and **3** in a simple two-step operation. Thus, this route proved a viable procedure for the large-scale synthesis of **I** sufficient to enable the DOS investigations.

With **2** in hand, mild cerium(IV) ammonium nitrate (CAN) radical-mediated removal of the PMP group afforded the desired free amine **4** in good yield (Scheme 2.9). This amine could be isolated in good purity *via* an acidic/basic work-up, avoiding the need for potentially problematic column chromatography of the polar amino compound. Disappointingly, however, when a direct reduction was carried out upon ester **4** no desired amino alcohol **5** could be isolated (Scheme 2.9), and only a complex mixture was observed by <sup>1</sup>H NMR.



**Scheme 2.9.** CAN-mediated PMP deprotection and first attempt to obtain the key amino alcohol.

In order to access the desired alcohol functionality, it was therefore decided to first protect the amine with an alternative group and subsequently reduce the ester to furnish the desired hydroxyl building block (Scheme 2.10). Thus, amino ester **4** was heated with an excess of di-*tert*-butyl dicarbonate ( $\text{Boc}_2\text{O}$ ) in THF affording the *N*-Boc protected derivative **6** in high yield. This was subsequently subjected to a lithium-mediated ester reduction generating the desired starting quaternary amino alcohol **7** in excellent yield. It is worth mentioning that the use of an acid-labile protecting group was chosen to enable regioselective modifications to both *N*- and *O*- functionalities as required.



**Scheme 2.10.** The final synthetic steps to yield **6** and **7**.

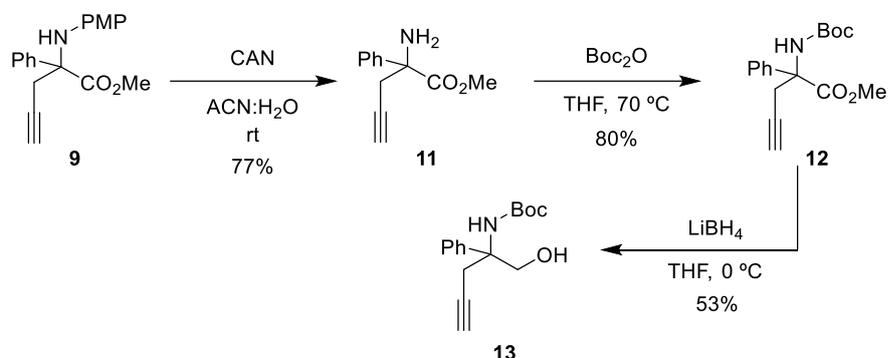
### Exemplification of the building block versatility

As a means to exemplify the versatile nature of the synthetic route to the starting building blocks, it was decided to generate an alternative substituent at the quaternary centre. In line with the project proposal, it was hoped that simply through variation of the ketoester used in the first step a substituent of choice could easily be installed at this position. To demonstrate this, applying the previously developed route, the phenyl variant of **7** was synthesised (Scheme 2.11). Pleasingly, upon reflux of commercially available methyl benzoylformate with *p*-anisidine and catalytic amounts of *p*-toluenesulfonic acid, imine **8** was isolated in an 88% yield. Subsequent subjection of **8** to the same Barbier reaction conditions used previously, achieved the nucleophilic propargyl addition. In this case, alkyne **9** and allene **10** were observed in a 6:1 ratio (determined by  $^1\text{H}$  NMR), again separable by column chromatography.



**Scheme 2.11.** Synthesis of a phenyl building block derivative.

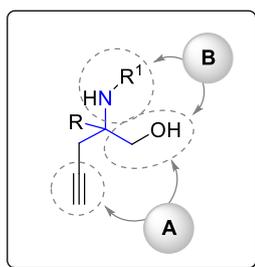
Accordingly, applying the previously developed synthetic sequence the amine was subsequently deprotected to give **11**, followed by *N*-Boc protection to afford **12** and ester reduction to afford the corresponding phenyl-substituted amino alcohol **13** in a moderate overall yield (Scheme 2.12).



**Scheme 2.12.** Synthesis of the key amino alcohol with an alternative quaternary substituent.

### 2.3.2 Preparation of highly functionalised intermediates and further cyclisation: Couple and Pair phase

With building blocks **7** and **13** in hand an exploration into the potential to form 3D diverse scaffolds from these substrates was undertaken. *This was achieved in two parts: (i) cyclisation of the hydroxyl and alkyne moieties (Figure 2.3, type A) or (ii) cyclisations between the nitrogen and hydroxyl moieties (Figure 2.3, type B). It was envisaged that during this, manipulation of the N- or O- moieties through the introduction of additional reactive functionalities would expand the scope, allowing the generation of multiple heterocyclic scaffolds.*

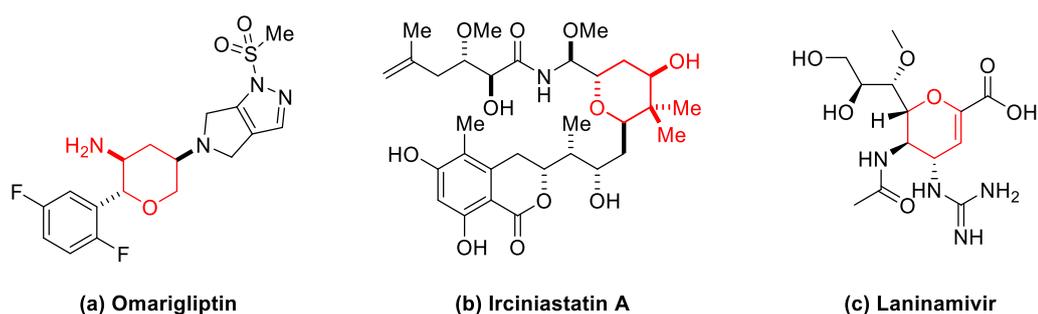


**Figure 2.3** The proposed strategy to explore the use of **I** in DOS.

### 2.3.2.1 Investigations into hydroxyl and alkyne pairing via pathway A

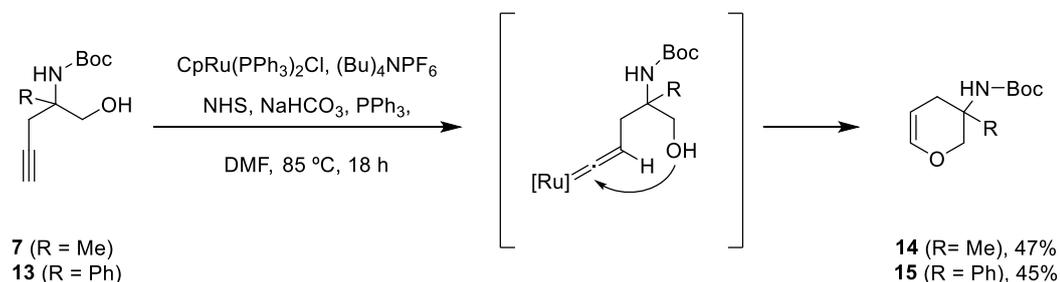
#### Direct intramolecular cyclisations

Substituted dihydropyrans (DHPs) and their fully saturated counterparts are found in many interesting natural products such as *Irciniastatin A*, with anti-tumour properties,<sup>197</sup> Omarigliptin, under development for the treatment of type 2 diabetes,<sup>198,199</sup> and the antiviral Laninamivir, under development for the treatment of Influenza A and B<sup>200</sup> (Figure 2.4). Given the presence of this characteristic motif in such structures, we envisioned that the generation of fragment-like, NSQC-bearing scaffolds *via* a versatile synthetic route would be of great interest.



**Figure 2.4.** DHP motifs within bioactive molecules.

Inspired by recent examples of the reactivity of ruthenium vinylidene complexes with N/O/S nucleophiles,<sup>201</sup> we envisioned a disubstituted DHP could be afforded *via* O-mediated nucleophilic attack at the alkyne upon exposure of **7** to transition metal catalysts<sup>201</sup> (Scheme 2.13). To this aim, following the procedure developed by Zacuto *et al.*, CpRu(PPh<sub>3</sub>)<sub>2</sub>Cl was used to initiate the cyclisation, affording DHP **14** in moderate yield (Scheme 2.13). Furthermore, to demonstrate the versatile introduction of different substituents at the quaternary carbon within the final small molecules, the corresponding phenyl derivative **15** was additionally synthesised, furnishing the scaffold in a similarly moderate yield.

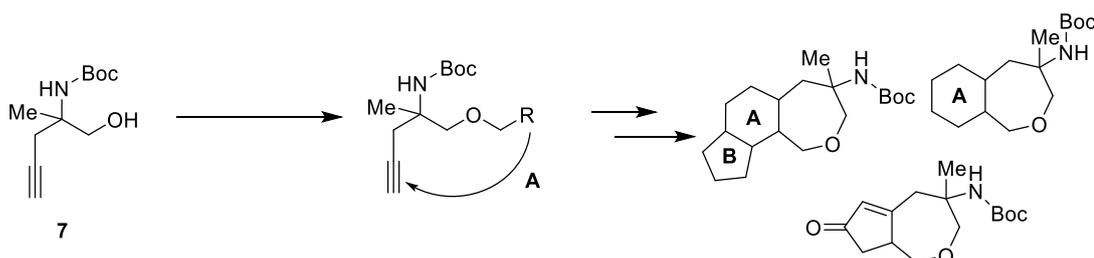


**Scheme 2.13.** The synthesis of DHP fragments with varied quaternary substituents. NHS = N-hydroxy succinimide.

## Hydroxyl alkylation strategies

Whilst oxepane-containing scaffolds are less well explored in drug discovery than their aza-counterparts,<sup>202</sup> this ring structure can be found embedded within several bioactive natural products,<sup>203</sup> whilst oxepanes flanked by two aromatic rings are commonly found within central nervous system (CNS)-active drugs, such as doxepin or asenapine.<sup>204</sup> As such, there remains a sustained interest in the development of novel methodologies to construct these motifs. Indeed, the privileged nature of this moiety has led to the development of novel methods to generate oxepane-based  $sp^3$ -rich libraries for the ELF<sup>205</sup> and as a chemical navigator within biology-oriented synthesis campaigns.<sup>206</sup>

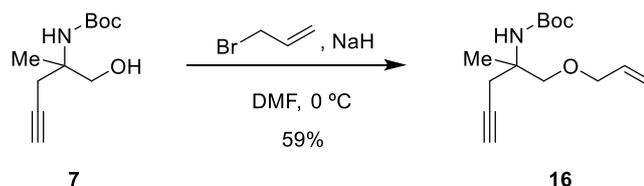
Inspired by these reports, we next considered the scope of introducing additional pendant functionalities to **7** via alkylation of the hydroxyl motif for generating oxepane-containing scaffolds. Accordingly, it was hoped the reactivity of these newly introduced functionalities could be harnessed to generate different polycyclic scaffolds in a chemoselective fashion. More specifically it was envisaged that the introduction of an additional unsaturation within the molecule would provide the opportunity to utilise different transition-metal based cyclisations to afford novel oxepane-fused scaffolds (Scheme 2.14).



**Scheme 2.14.** Cyclising the pendant groups and the alkyne following pathway **A**.

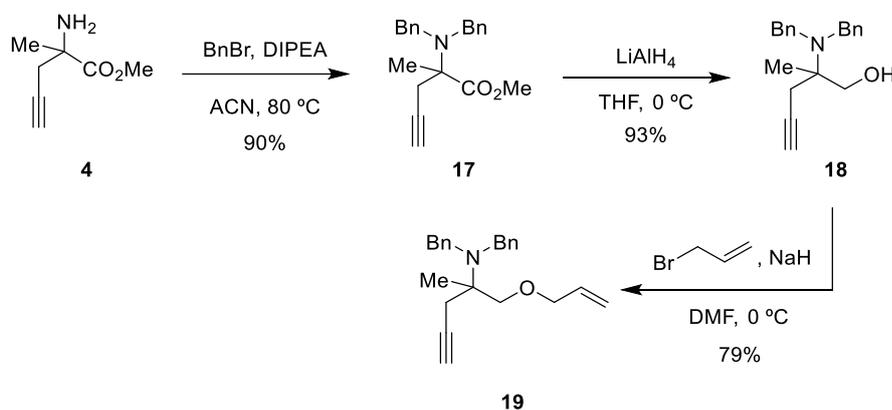
## Metathesis-Diels Alder cascade

With this proposal in mind, our investigations initially focused on the installation of an allyl group in the hope to afford enyne variant **16** (Scheme 2.15). Thus, with the *N*-Boc protected amino alcohol in hand, alkylation was investigated using allyl bromide and NaH as a base in DMF. Starting from **7**, only a moderate 57% yield of **16** was achieved, with crude <sup>1</sup>H NMR analysis suggesting secondary *N*-alkylation to be the origin of the low observed yield.



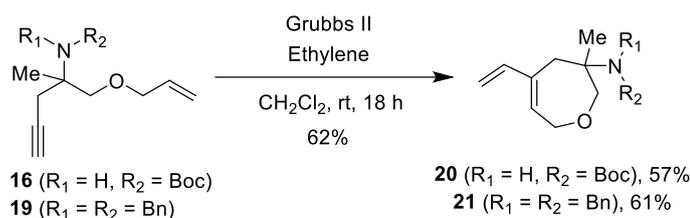
**Scheme 2.15.** Alkylation of *N*-protected amino alcohol **7**.

In order to mitigate this, an alternative *N*-protection strategy was explored. Thus, starting from the unprotected amino ester **4**, benzyl bromide and DIPEA were added in excess to yield the corresponding diprotected amino ester **17** in high yield (Scheme 2.16). Then, treatment of **17** with LiAlH<sub>4</sub> to mediate ester reduction afforded **18** in an excellent 93% yield. Pleasingly, applying the same ether formation of the same conditions to **18**, **19** was isolated in an improved 79% yield.



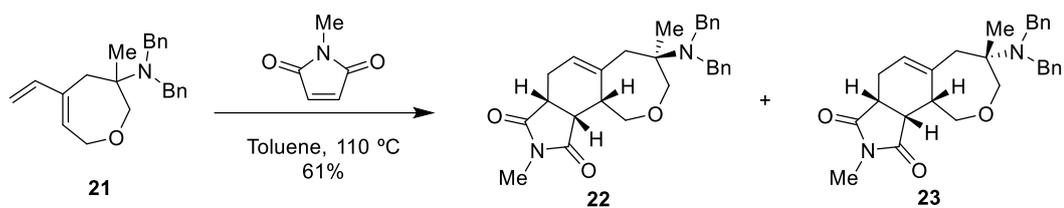
**Scheme 2.16.** Synthesis of benzyl protected amino ether **19**.

With enynes **16** and **19** in hand, it was envisaged that RCEYM under Mori conditions<sup>207,208</sup> would afford tetrahydrooxepines **20** or **21** (Scheme 2.17). Indeed, upon exposure of these intermediates to Grubbs second generation catalyst using a reduced reaction molarity (0.0255 mol dm<sup>-3</sup>) the intramolecular reaction was promoted, furnishing **20** and **21** in moderate yields. In this case, the yield of this procedure was attributed to the observed instability of the resulting diene products under TLC analysis, suggesting decomposition at room temperature after only 24 hours.



**Scheme 2.17.** RCEYM of **16** and **19** afforded dienes **20** and **21**.

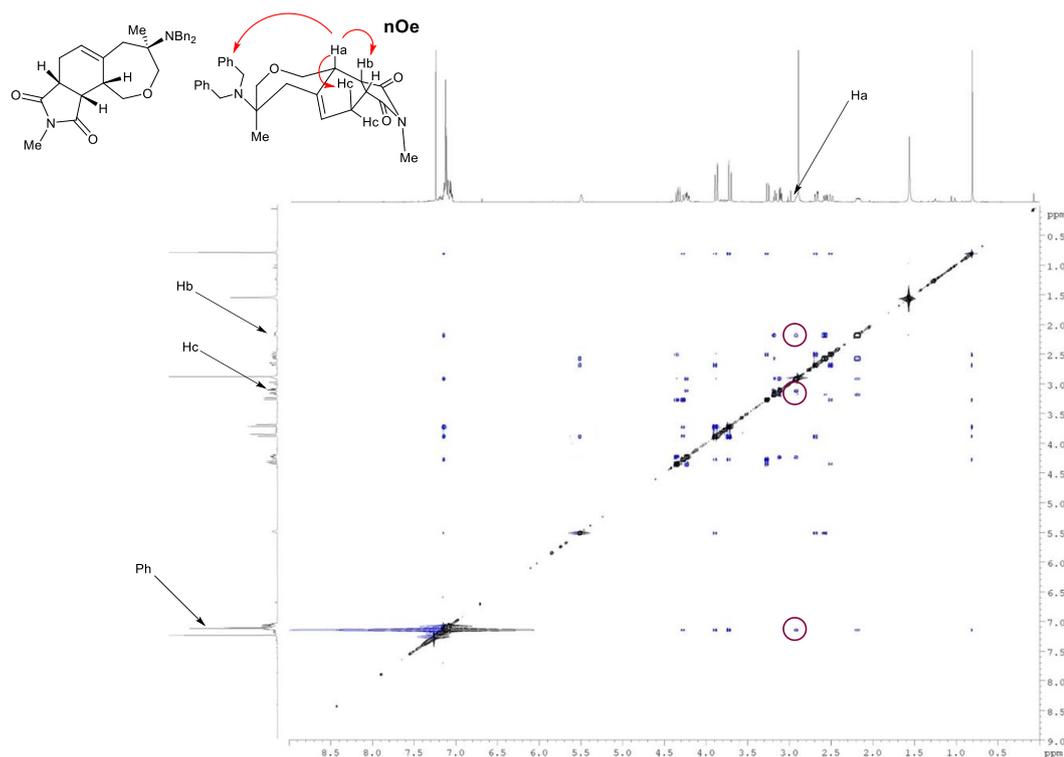
In a divergent manner, it was thought that dienes of this nature would serve as ideal substrates to participate within further transformations *via* a DAR using different dienophile components to create multiple diverse scaffolds.<sup>206,207</sup> To this aim, the [4+2] cycloadditions could be promoted through thermal activation of **2** in the presence of *N*-methylmaleimide yielding bicycles **22** and **23**, as a mixture of diastereoisomers (*dr ca.* 3:1, determined by <sup>1</sup>H NMR) which could be easily separated by flash column chromatography (Scheme 2.18). In the case of **20**, however, only a complex mixture was isolated.



*dr* 3:1

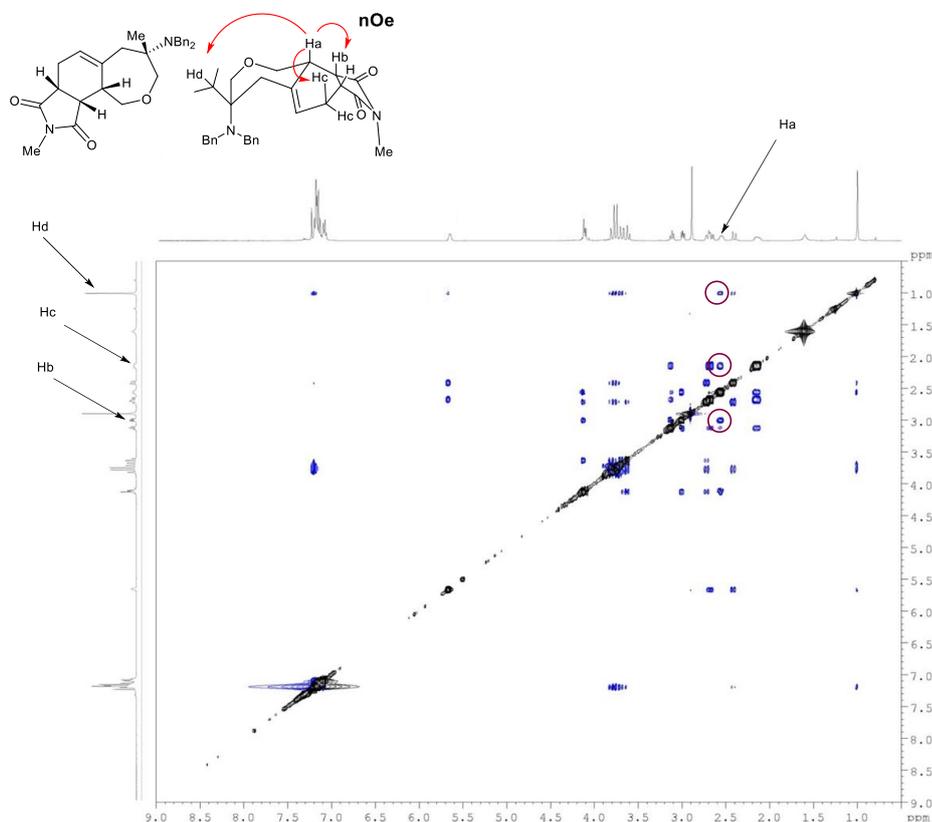
**Scheme 2.18.** DAR of 1,3-diene **21** with *N*-methylmaleimide to afford **22** and **23**.

The relative configuration between the quaternary carbon and the three new stereocenters formed after the DAR was deduced by nOe correlation analysis in both tricycles **22** and **23** (Figure 2.6). It was found that within the nOe spectrum of **22**, an interaction was present between the two protons of the newly formed stereocenters (Ha) and (Hb) indicating a *cis* relationship and the presence of the *endo* product. The relative orientation of these stereocenters to the remainder of the molecule could also be determined through the observed strong nOe interactions between the Ha proton and Hc proton within the cyclohexene ring and the axial phenyl protons of the benzyl group.



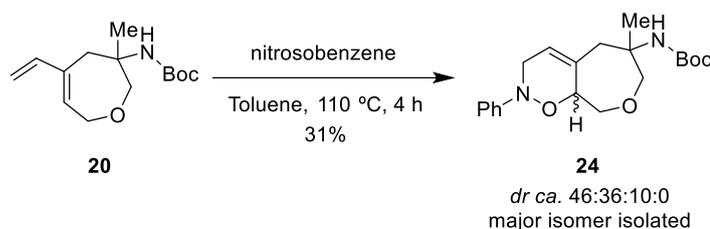
**Figure 2.5.** The NOESY 2-D spectra for diastereomer **22**.

Similarly, within the nOe spectra of **23** the *endo* product was again confirmed through the interaction between Ha with both Hb and Hc, whilst the presence of the alternative diastereomer was suggested through an additional interaction between Ha and the axial methyl protons Hd (Figure 2.6). This provided sufficient information to allow the determination of the relative stereochemistry of both polycyclic compounds **22** and **23**.



**Figure 2.6.** The NOESY 2-D spectra for diastereomer **23**.

With the success of this sequence in mind, it was thought to investigate alternative dienophiles which could be employed to generate additional novel scaffolds. Thus, in light of the previous evidence of the influence of the *N*-protection strategy on the ease of separation of diastereomers, both the diene products **20** and **21** were subjected to the DAR conditions with nitrosobenzene (Scheme 2.19). However, in the case of **21** only an inseparable mixture of the four possible regio- and stereoisomers was generated (ca. 4:2.2:2:1.8 ratio by crude  $^1\text{H}$  NMR analysis). The analogous reaction of the *N*-Boc diene **20** afforded a 4.6:3.6:1.0:0 distribution of regio and stereoisomers (determined by crude  $^1\text{H}$  NMR analysis). Pleasingly, in this case the major isomer from this mixture could be isolated in a 31% overall yield. Attempts to deprotect the amine under acidic conditions using trifluoroacetic acid to improve the separation of the remaining isomers proved in vain, with only a complex mixture obtained in this case and evidence of starting material decomposition detected by  $^1\text{H}$  NMR analysis.



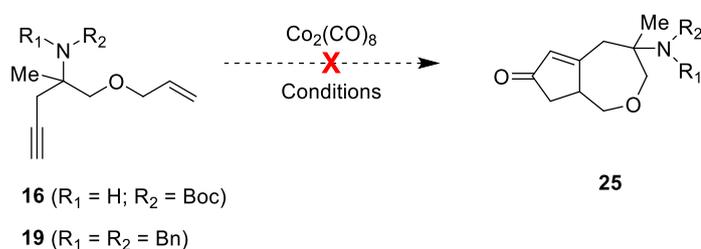
**Scheme 2.19.** Using nitrosobenzene as a dienophile, novel  $sp^3$  rich architectures could be constructed.

Efforts to determine the relative position of the substituents present within **24** were subsequently investigated. Unfortunately, nOe experiments proved inconclusive to definitively determine the diastereomer generated due to lack of both diastereomers for analysis and overlap of vital  $^1\text{H}$  NMR signals. Furthermore, no X-ray data could be obtained for the product due to the oily composition of the resulting compound.

### Cycloaddition pair phase of enyne **23**

With **22**, **23** and **24** in hand, alternative methodologies to pair the alkene and alkyne moieties within the branching intermediates **16** and **19** were then pursued. The Pauson Khand reaction (PKR) is a widely used method of forming fused cyclopentenone derivatives, mediated *via* Cobalt catalysis.<sup>209</sup> It was proposed that the alkene and alkyne moieties could be additionally paired through such a [2+2+1] cycloaddition to yield cyclopentenone-fused [5,7]- ring system **25** (Scheme 2.20).

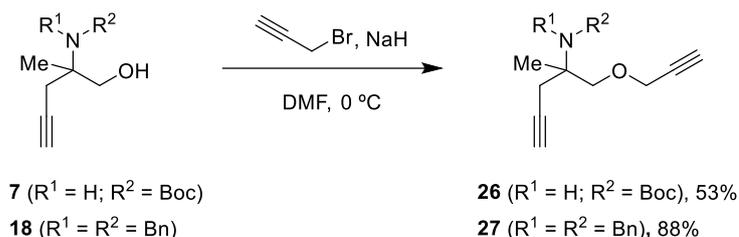
To this aim, our first efforts in the preparation of bicyclic system **25** were based on conditions reported using similar substrates.<sup>210</sup> This involved using stoichiometric amounts of the  $\text{Co}_2(\text{CO})_8$  catalyst along with 10 equivalents of the oxidant (NMO) to effect the transformation. When bis-benzyl protected intermediate **16** was subjected to these conditions, we were disappointed to observe only complete decomposition of the starting material and no desired product formation. Moreover, a similar result was obtained when subjecting *N*-Boc protected amine **19** to the same conditions. In order to maintain focus on investigating the scaffold diversity as opposed to detailed optimisations the formation of this scaffold was abandoned.



**Scheme 2.20.** Attempts to the synthesis bicyclic compound **25**.

## Alkylation of the hydroxyl with a second propargyl group

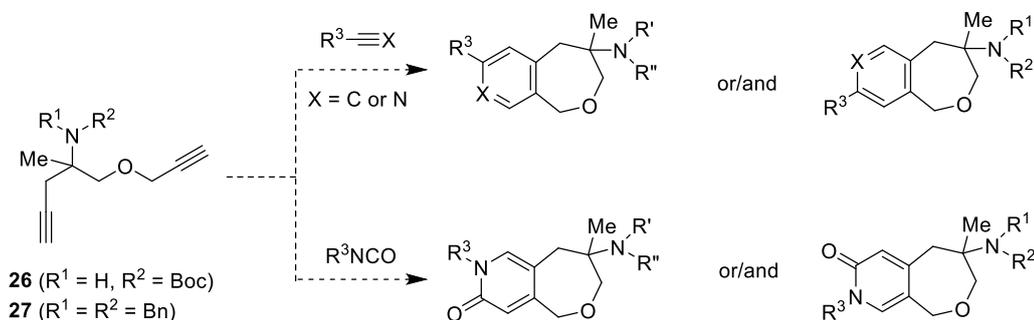
Inspired by the success of the metathesis cascade pair phase of enynes **16** and **19**, it was envisaged that alkylation of the hydroxyl with an alternative unsaturated moiety would provide the opportunity to form different complex molecules. Thus, alcohols **7** and **18** were subjected to the same ether-formation conditions instead using propargyl bromide to form **26** and **27** (Scheme 2.21). Analogous to the previous alkylations of these substrates, this reaction was found to be significantly more successful with the *N*-benzyl building block **18** delivering **27** in 88% yield, compared to only 53% of **26** from **7**.



**Scheme 2.21.** The formation of diynes **26** and **27**.

## [2+2+2] cyclotrimerisation pair phase

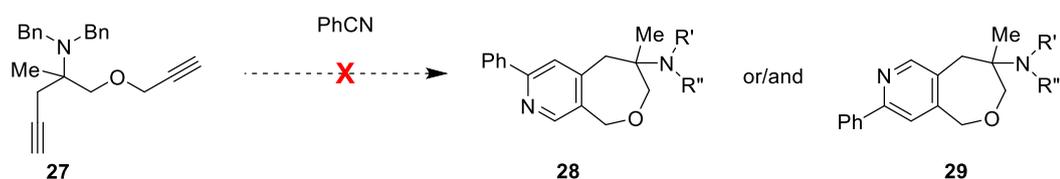
According to previous results on the formation of *N*-containing heterocycles *via* [2+2+2] cyclotrimerisations<sup>211</sup> it was envisaged that different [6,7]-bicyclic compounds could ultimately be synthesised starting from **26** and **27** (Scheme 2.22). Indeed, transition-metal catalysed [2+2+2] cyclotrimerisations have been extensively used to generate benzene derivatives due to the exceptional synthetic versatility of this methodology to construct complex molecules in an efficient manner.<sup>212,213</sup> Modern synthetic advancements in this reaction describe modifications to the procedure including the use of alternative variants of the third 'ene' such as nitriles and isocyanates to generate both pyridine and pyridinone moieties,<sup>214–216</sup> both of which are medically relevant motifs.<sup>202</sup>



**Scheme 2.22.** Scaffold diversification through [2+2+2] cyclotrimerisation starting from dialkyne **26** or **27**.

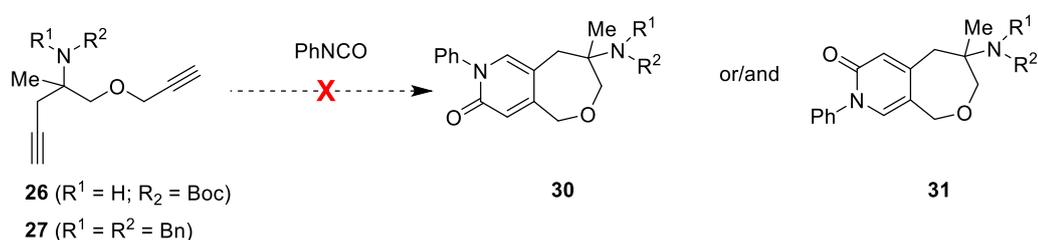
To this aim, intermediate **27** was subjected to a range of cyclisation conditions, employing benzonitrile as a third 'pair' component, with the objective of forming scaffolds **28** and **29**

(Scheme 2.23). Two different catalyst systems were trialled -  $\text{CpCo}(\text{CO})_2$  and Wilkinson's catalyst – however, in both cases only a complex reaction mixture was observed upon crude  $^1\text{H}$  NMR analysis of the resultant reaction mixture. Conditions reported by Nicolaus *et al.*<sup>217</sup> suggested the use of  $\text{PPh}_3$  may be beneficial in order to stabilise the metal in the catalytic cycle, thus increasing the reactivity of such processes when also used in conjunction with microwave heating. However, these conditions also proved unsuccessful with only starting material observed in both cases.



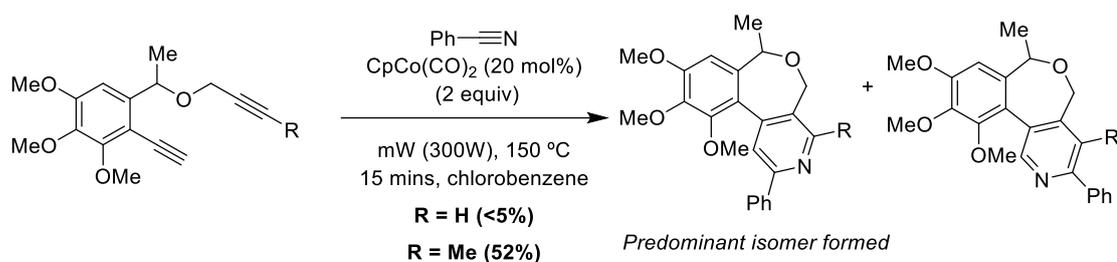
**Scheme 2.23.** [2+2+2] cyclotrimerisations were attempted using **23** to form oxepinopyridine scaffolds.

It was hypothesised that the lack of reactivity could be attributed to the nature of the third 'pair' component. Thus, using the same reaction conditions, phenyl isocyanate was employed, replacing benzonitrile, to generate bicyclic pyridinones **30** and **31** (Scheme 2.24). However, following this procedure once more only a complex mixture was observed when using both **26** and **27**.



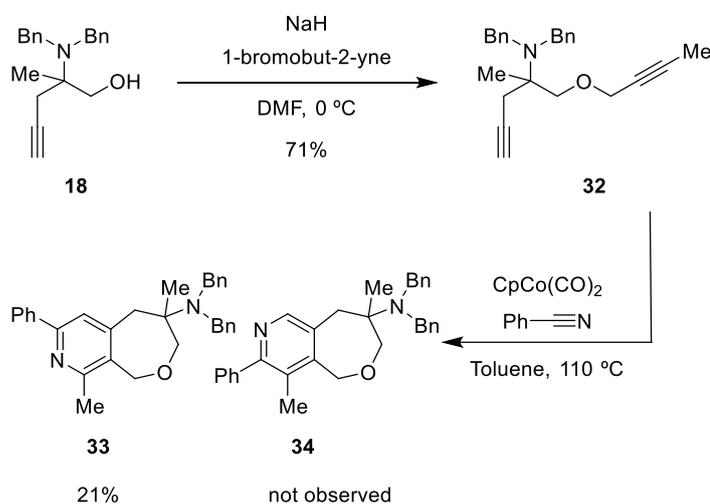
**Scheme 2.24.** Attempts toward the synthesis of pyridinones **30** and **31**.

Literature observations suggested that terminal alkynes can display reduced activity compared to substituted alkynes when they are exposed to [2+2+2] cyclotrimerisation reactions. Reports by Nicolaus *et al.*<sup>218</sup> demonstrated that when similar oxygen-linked substrates were subjected to cyclotrimerisation conditions, methyl-substituted alkynes could be utilised to form the aromatic products, where the terminal alkynes failed (Scheme 2.25). In addition to changing the *R* substituents, this methodology was shown to tolerate a wide range of different substituted nitriles to generate appendage diversity.



**Scheme 2.25.** [2+2+2] cyclotrimerisations using methyl-substituted alkynes.<sup>218</sup> mW = microwave.

Inspired by these reports, an alternative approach was proposed utilising substituted alkyne derivative within the [2+2+2] cyclotrimerization reaction. Thus, dibenzyl amino alcohol **18** was treated with 1-bromo-2-butyne to furnish internal alkyne **32** in good yield (Scheme 2.26). Pleasingly, upon subsequent subjection of **32** to Co catalysis and benzonitrile at 110 °C for 48 hours, a single regioisomer **33**, was isolated.

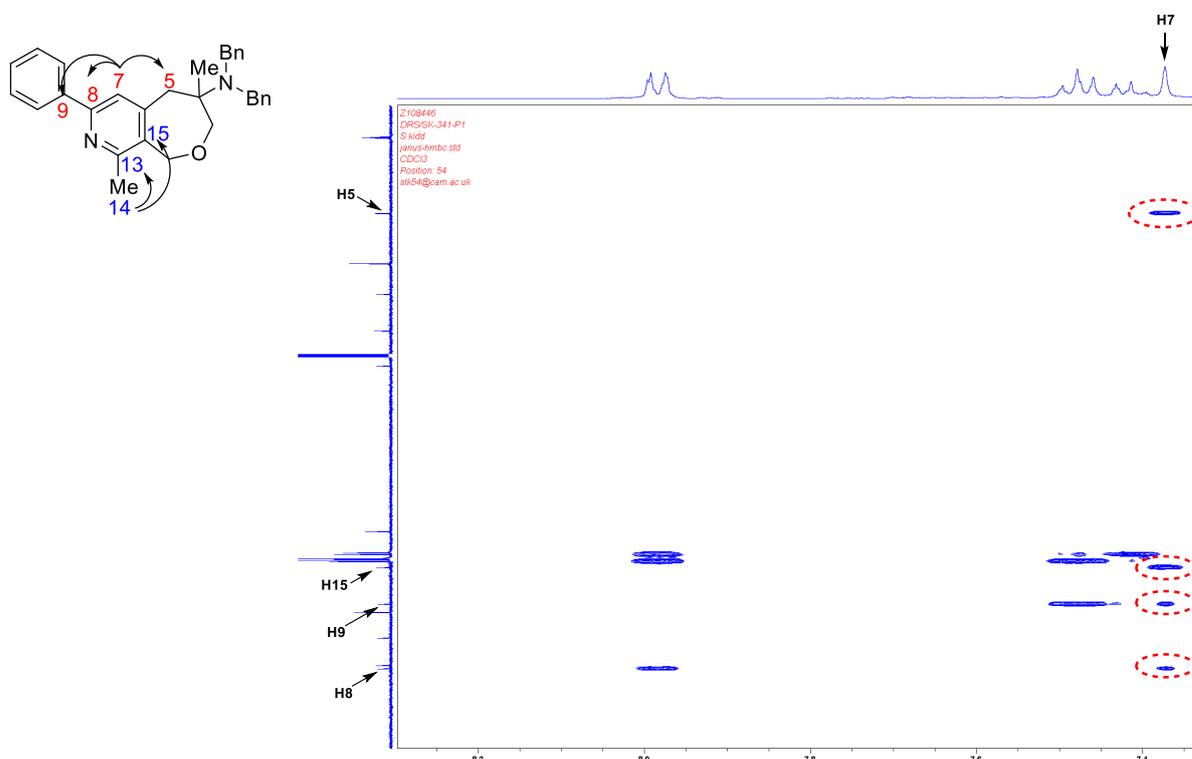


**Scheme 2.26.** An alternative approach for the [2+2+2] cyclotrimerisations yielding scaffold **33**.

The formation of a single regioisomer during [2+2+2]-cyclotrimerisations had previously been rationalised by Boñaga *et al.*<sup>219</sup> by considering the metallacyclopentadiene intermediate generated under the reaction conditions. In this manner, it was hypothesised that the increased steric clash between the phenyl and methyl substituents within intermediate of type **34** would be disfavoured. The poor yield of this reaction was attributed to decomposition of the starting material as a result of the forcing conditions required to promote the reaction. Nevertheless, considering the challenging nature of generating 7-membered rings compared to their 6- or 5-membered counterparts, in addition to the comparably low yields reported in the literature,<sup>218,219</sup> this result was still pleasing. Moreover, sufficient quantities of the compound were obtained to enable biological testing and it was envisioned that optimisation of this procedure could be achieved in the future, if required.

The regiochemistry of **33** was established through means of HMBC 2-D NMR spectral analysis whereby coupling was observed between the single proton within the pyridine moiety (H<sub>7</sub>) to

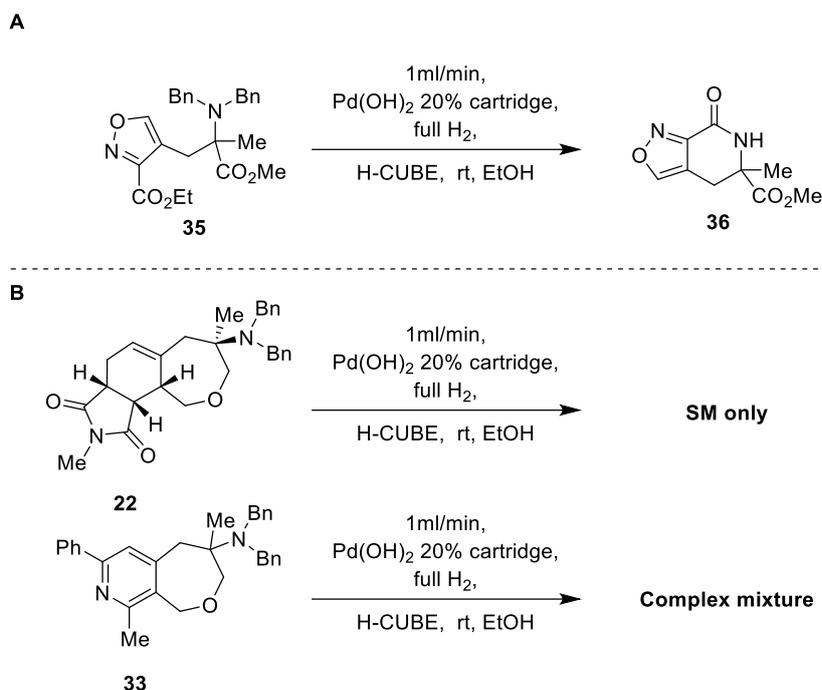
those within the pendant phenyl ring ( $H_8$  and  $H_9$ ), in addition to the methylene adjacent to the quaternary centre ( $H_5$ ) (Figure 2.7). Furthermore, only coupling between the methyl protons ( $H_{14}$ ) to the quaternary carbons ( $H_{13}$  and  $H_{15}$ ) was observed. This evidence was further corroborated by the upfield shift of both the  $^1H$  peak of  $H_7$  and  $^{13}C$  peaks for  $C_7$ , compared to the down field shift of  $C_{13}$  due to the differences in proximity from the deshielding nitrogen nucleus.



**Figure 2.7.** 2-D HMBC analysis enabled determination of the regiochemistry of **33**.

### Amine deprotection

With **22**, **23** and **33** oxepane scaffolds in hand, it was thought necessary to investigate the potential of scaffold deprotection to reveal the free primary amine to exemplify its use as a fragment exit vector. Previous efforts to remove the benzyl group from **35** to generate **36** conducted by Dr. Natalia Mateu concluded that hydrogenolysis *via* the use of H-Cube apparatus at room temperature could be the optimum conditions capable to affect this transformation (Scheme 2.27, A).

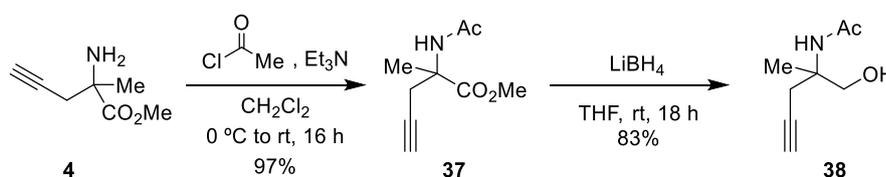


**Scheme 2.27.** Previous success in *N*-benzyl removal was observed by Dr Mateu. However, this route proved unsuccessful for **22** and **33**.

Thus, both **22** and **33** were subjected to identical conditions (Scheme 2.27, B), however, no evidence of benzyl removal was observed with only starting material or a complex mixture observed in both cases. Moreover, it was hypothesised that alternative forcing conditions could result in undesirable loss of the alkene moiety within **22**. Considering these findings and the difficulties experienced in removal of the *N*-Boc group from **24**, an alternative strategy to enable modification at this position was devised.

### Alternative *N*-Acyl substituents

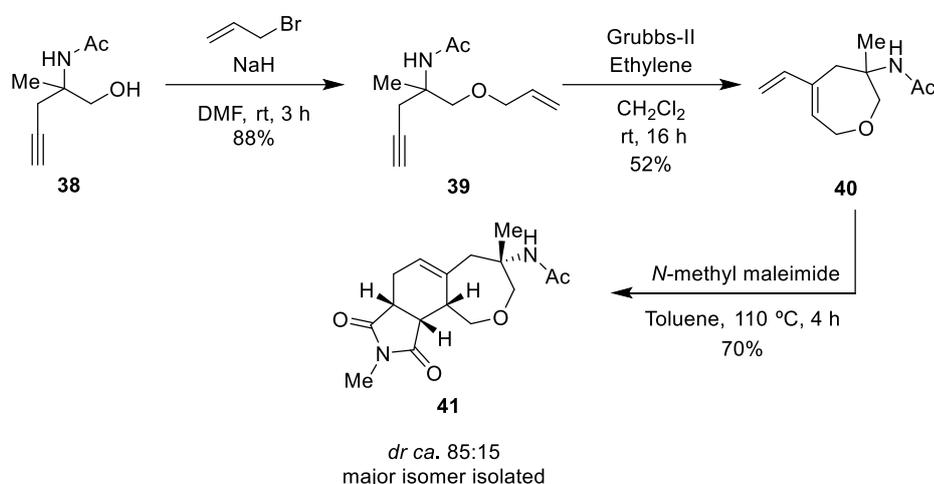
To this aim, it was instead thought an alternative non-removable *N*-substituent, such as an acetamide, could be incorporated at this position. It was envisaged that in turn this functionality could be exploited as a fragment growth vector, if required upon the identification of a hit by means of biological testing. Accordingly, the *N*-acetamide variant of ester **37** was synthesised, followed by reduction to yield the key alcohol building block **38** (Scheme 2.28).



**Scheme 2.28.** Acylation of the primary amine **4** followed by ester reduction to afford **38**.

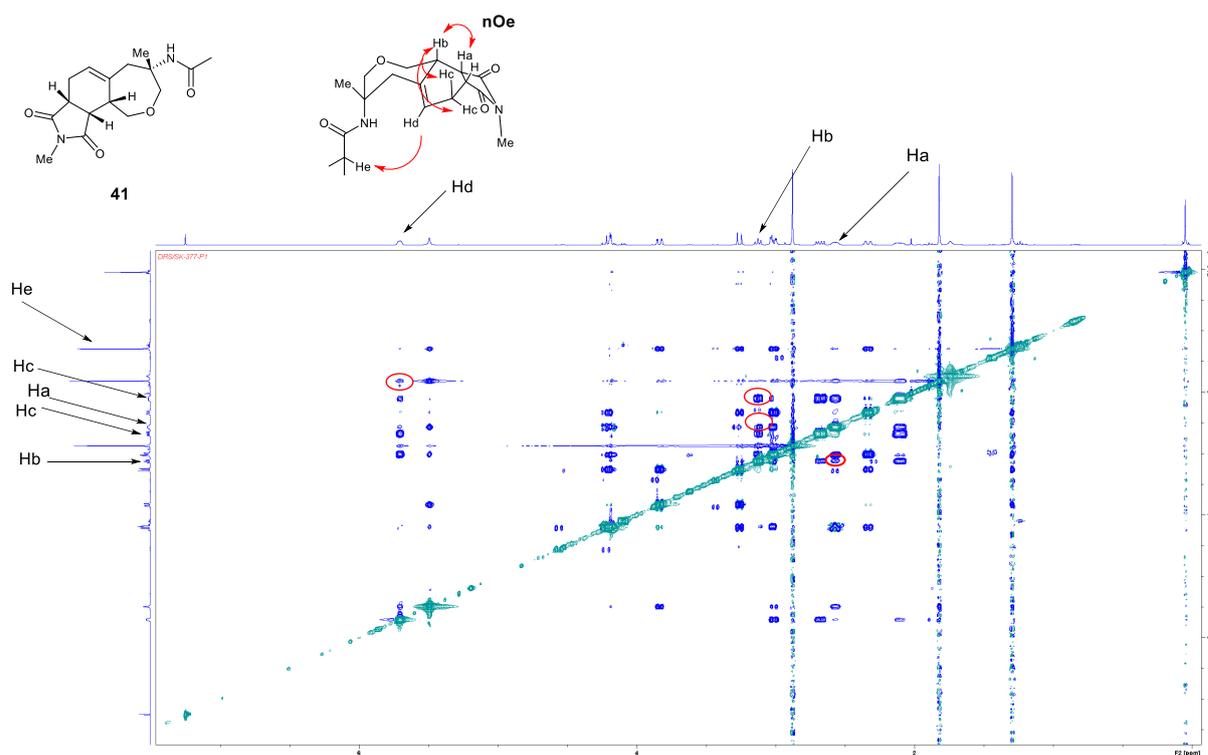
## Metathesis cascade strategy

With **38** in hand, the identical alkylation-metathesis-Diels-Alder synthetic sequence developed previously was then explored (Scheme 2.29). Pleasingly, alkylation with allyl bromide proceeded smoothly, affording enyne intermediate **39** in good yield. Subsequent treatment of **39** to Grubbs-II provided **40** in moderate yield, followed by reaction with *N*-methyl maleimide to generate scaffold **41** in an 85:15 ratio of *endo* product diastereomers (determined by crude <sup>1</sup>H NMR analysis). In this case, whilst both isomers were observed, due to the scale of the reaction and difficulties in separation of the diastereomers during purification, insufficient quantities of the minor diastereomer were isolated for full characterisation. Importantly, the major diastereomer was isolated in sufficient quantities for biological testing.



Scheme 2.29. The metathesis-Diels Alder sequence using **38** to afford **41**.

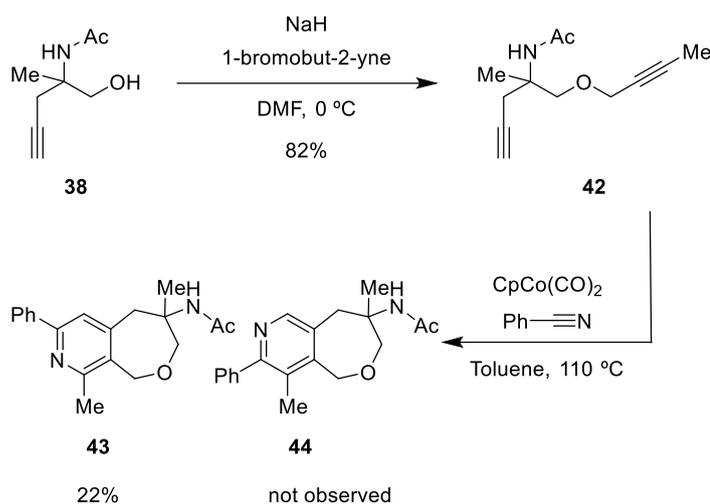
In an analogous fashion to **22** and **23**, nOe studies indeed proved fruitful in the determination of the relative positions of the substituents (Figure 2.8). It was found that within the nOe spectrum of **41**, interactions were observed between the two protons of the newly formed stereocenters (Ha) and (Hb) indicating a *cis* relationship and between Ha and Hc, indicating the presence of the *endo* product. The relative orientation of these stereocenters to the remainder of the molecule could also be determined through the observed nOe interactions between the alkene proton Hd proton to acetamide methyl protons He.



**Figure 2.8.** The NOESY 2-D spectra for diastereomer **41**.

### **[2+2+2] cyclotrimerisation strategy**

Satisfied by the success of the *N*-Acetamide strategy within the above route, the corresponding [2+2+2] methodology was hence investigated. To this end, **38** was alkylated with 1-bromobut-2-yne, furnishing the internal alkyne **42** in good yield (Scheme 2.30). Subsequent subjection of this intermediate to the same Co-mediated [2+2+2] conditions, utilising benzonitrile as the third alkyne component, yielded fused pyridine scaffold **43**. In a similar fashion, only one of the two possible regioisomers of the reaction was observed and isolated, albeit in a comparably poor yield.



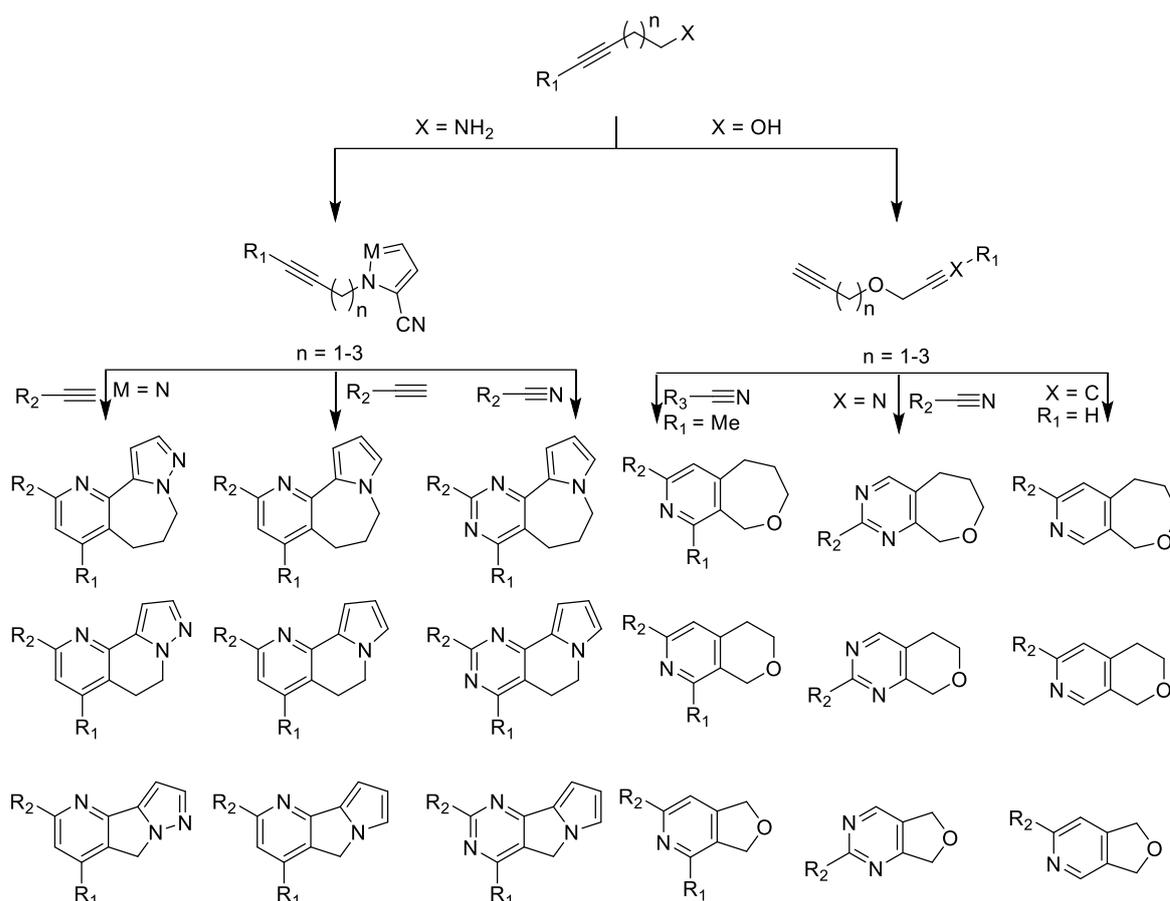
**Scheme 2.30.** The same [2+2+2] methodology was applied to **42** forming the fused oxepine **43**.

Analogous to the dibenzyl analogue **33**, the regiochemistry of **43** was established through means of HMBC 2-D NMR spectral analysis, where the same evidence was utilised to determine the regiochemistry.

### Methodology expansion

Due to the novelty of the fused-oxepane scaffolds generated using this route, it was envisioned that versatility of this methodology could be exploited to construct alternative challenging architectures in an efficient manner. Indeed, partially saturated heteroaromatic motifs have been of recent interest<sup>151</sup> and remain important architectures to develop modular synthetic routes toward. These initial findings therefore formed the basis of a M.Sci project undertaken by Ben Mackworth (Spring Group) investigating the expansion of this methodology (Scheme 2.31).

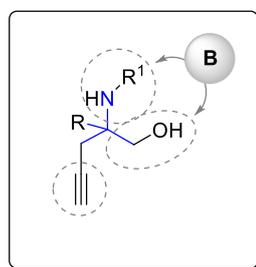
To simplify the synthesis, it was proposed that the quaternary centre could be removed from the starting substrate, allowing the key intermediate to be constructed from commercially available materials in only one step. It was anticipated that modification of the linker length would enable the formation of five, six and seven membered fused cycles, whilst the use of both unsubstituted and substituted alkynes within the starting material would result in the incorporation of both mono- and di-substituted pyridine rings. Furthermore, installation of a nitrile within the building block would allow the construction of pyrimidine scaffolds. Research in this project is on-going within the Spring group, continued by Edward Ou (Spring Group) as part of his PhD studies and will be published in due course.



**Scheme 2.31.** The proposed extension of the [2+2+2] methodology for the formation of multiple novel partially saturated architectures.

### 2.3.2.2 Investigations into N- and O- pairing via pathway B

Satisfied by the scope of pathway **A**, we next sought to explore pairings between the hydroxyl and alkyne motifs within **I**, following pathway **B** (Figure 2.9). It was envisioned this could be achieved in a similar manner to the exploration of pathway **A** - either through direct cyclisation or attachment of a pendant pairing functionality to generate additional scaffold diversity.

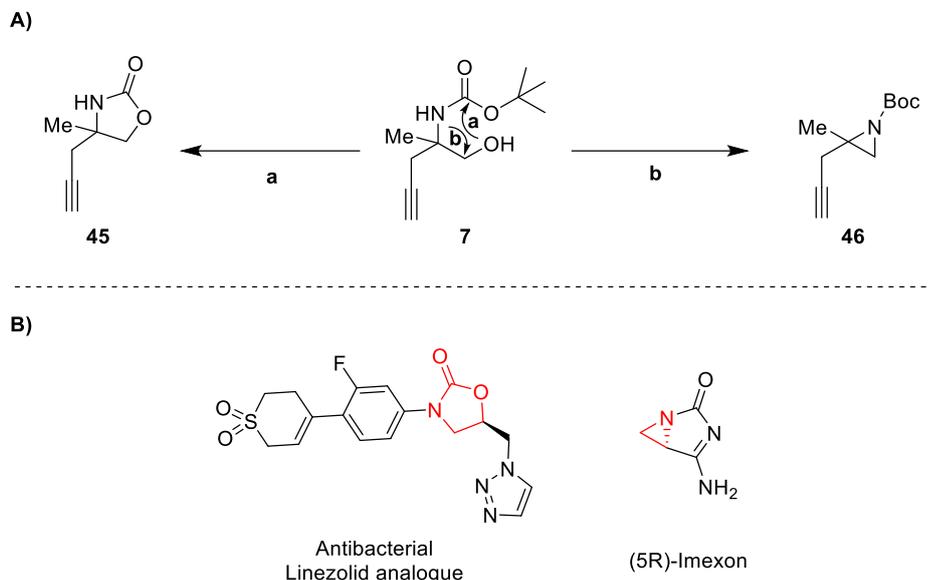


**Figure 2.9.** Subsequent investigations into pathway **B** were explored.

### Direct intramolecular cyclisations

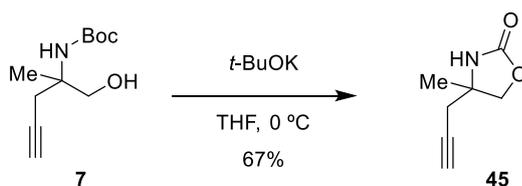
With the presence of both nucleophilic (hydroxyl and nitrogen) and electrophilic (C-X and carbonyl) moieties within the *N*-Boc amino alcohol **7**, it was envisaged that these functionalities could be exploited to form quaternary oxazolidinone **45**, or aziridine **46** motifs through two

different intramolecular reactions (Scheme 2.32, A). Both oxazolidinone and aziridine motifs represent important medicinal chemistry functionalities, derivatives of which can be found within both bioactive molecules reported in the literature, (Scheme 2.32, B).<sup>220–223</sup>



**Scheme 2.32.** A) quaternary oxazolidinone and aziridine formation through intramolecular reactions. B) examples of bioactive molecules containing these motifs.

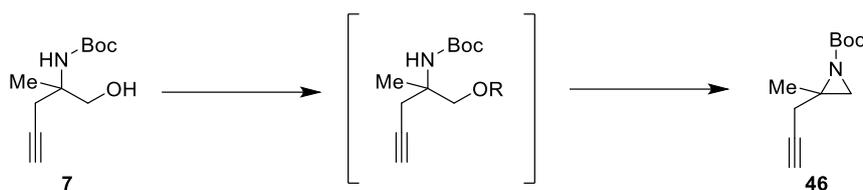
It was proposed that by considering the Boc-protecting group as a coupling partner, oxazolidinone **45** could be furnished starting from **7** through the nucleophilic attack of the hydroxyl to the carbamate moiety (Scheme 2.33). Thus, upon treatment with *t*-BuOK, the cyclisation between the nucleophilic hydroxyl group and the carbonyl was promoted, affording constrained carbamate **45** in a moderate 67% yield.



**Scheme 2.33.** Formation of the quaternary oxazolidinone **45**.

Alternatively, it was hoped that conversion of the hydroxyl within **7** to a suitable leaving group would promote *N*-substitution at the  $\alpha$ -carbonyl centre, forming **46**. Thus, initial work sought to investigate suitable methods to construct the aziridine moiety (Table 2.1). Firstly, a two-step procedure was trialled in which **7** was first transformed to a mesylate intermediate and subsequently subjected to *t*-BuOK to mediate the cyclisation (Table 2.1, Entry 1). Intramolecular-Mitsunobu reaction<sup>224,225</sup> (Table 2.1, Entry 2) *via* a phosphonium intermediate was also trialled, however in both cases only a poor yield of **46** was achieved. A literature search highlighted a simple alternative approach using tosyl chloride and potassium hydroxide in a one-pot process for the synthesis of aziridines.<sup>226</sup> Gratifyingly, applying this methodology

to quaternary substrate **7**, aziridine **46** was obtained in a pleasing 78% yield (Table 2.1, Entry 3).

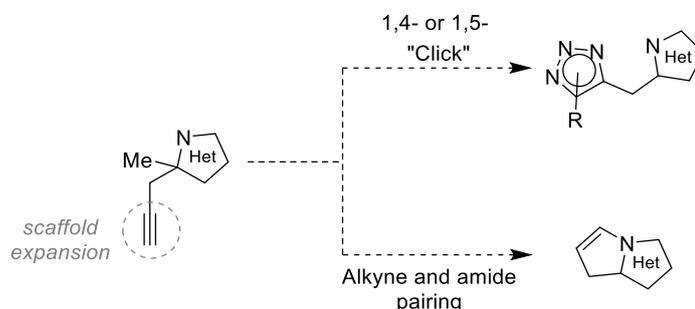


Entry	Conditions	R	Yield
	i. MsCl, Triethylamine, CH <sub>2</sub> Cl <sub>2</sub> , 0 °C then 40 °C,		
1	4 h	Ms	13%
	ii. <i>t</i> -BuOK THF, r.t., 20 h		
2	PPh <sub>3</sub> , DEAD, THF, r.t. then 50 °C, 5 h	PPh <sub>3</sub>	39%
3	TsCl, KOH, Et <sub>2</sub> O, 40 °C, 20 h	Ts	78%

**Table 2.1.** Three different conditions were trialled in attempts to form aziridine **46**. DEAD = diethyl azodicarboxylate.

### Further scaffold expansion *via* alkyne derivatisation

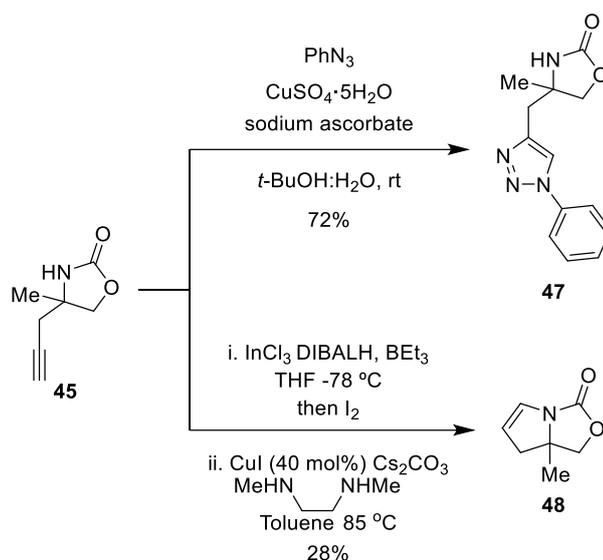
The pluripotent nature of the terminal alkyne functional handle present within both **45** and **46** prompted the hypothesis that this motif could be utilised in post-pair derivatisations to generate additional biologically relevant nitrogen heterocycles (Scheme 2.34). Examples of such transformations include the 1,3-dipolar cycloaddition between **45** and a suitable azide to form either a 1,4- or 1,5- triazole in a regioselective manner or alkyne-amide pairing. Such 5,5-linked scaffolds are analogous to those found in antibacterial analogues of Linezolid<sup>221,227</sup> and in fact those found in BACE-1 inhibitors,<sup>228</sup> demonstrating the biological relevance of such architectures.



**Scheme 2.34.** Other potential post pair reactions using the alkyne motif to generate novel restricted small molecules. Het = heterocycle.

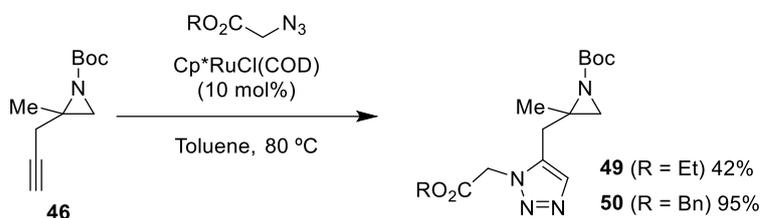
To this aim, upon exposure of **45** to azidobenzene in the presence of CuSO<sub>4</sub>·H<sub>2</sub>O and sodium ascorbate, 1,4- triazole **47** was furnished in a 75% yield (Scheme 2.35). Alternatively, utilising

methodology developed by Nicolai *et al.*<sup>229</sup> the fused 5,5-oxazolidinone **48** could also be generated. This was achieved through a two-step procedure in which the terminal alkyne was first converted to a vinyl iodide *via* radical-mediated *in situ* formation of  $\text{HIInCl}_2$ . Subsequent subjection of this intermediate to conditions developed by Buchwald<sup>230</sup> to initiate amide vinylation afforded **48** albeit in low yield. Importantly, within both the scaffolds derived from **45**, fragment growth vectors could be envisaged through modification of the azido substituent used in the triazole formation, or alternatively through modification of the alkene moiety within **48**.



**Scheme 2.35.** Modification of **45** via diversification of the alkyne moiety to give **47** and **48**. DIBALH = Diisobutylaluminium hydride.

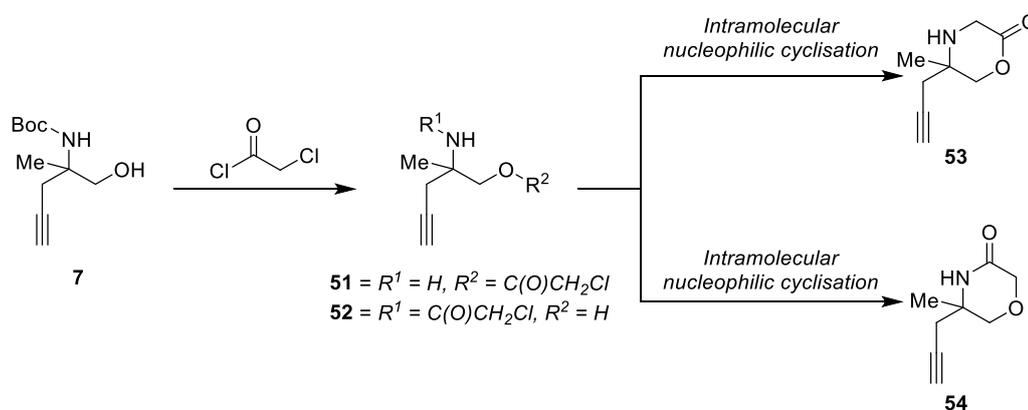
With exemplifications of **45** in hand, we next sought to investigate 1,5-triazole formation, instead using the alkyne within **46** (Scheme 2.36). Indeed, heating **46** with  $\text{Cp}^*\text{Ru}(\text{COD})\text{Cl}$ <sup>231</sup> and ethyl 2-azidoacetate afforded the substituted aziridine **49**, albeit in a modest 42% yield. To exemplify the versatile nature of introducing alternative functionalities at this position benzyl 2-azidoacetate was employed under the same reaction conditions affording **50** in an excellent 95% yield.



**Scheme 2.36.** Transformation of aziridine **46** into 1,5-triazoles **49** and **50** via ruthenium-catalysed cycloadditions. Benzyl 2-azidoacetate previously synthesised by Dr. Natalia Mateu.

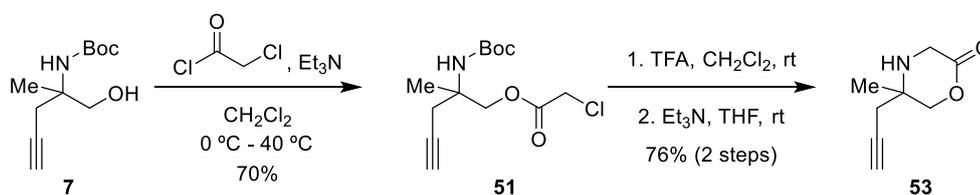
## Acylation strategies for the formation of morpholinone motifs using pathway A

In addition to direct cyclisations that could be inflicted upon **7** we were also interested in exploring the incorporation of additional reactive functionality to the substrate following a traditional *build/couple/pair* DOS approach. The presence of both amine and hydroxyl nucleophiles within **7** prompted the idea that acylation of either the nitrogen or hydroxyl with chloroacetyl chloride would allow the formation of two distinct morpholinone motifs in a regioselective manner. This could be achieved via either (i) *O*-acylation followed by *N*-mediated cyclisation to give **53** or *vice versa* (ii) *N*-deprotection and acylation followed by *O*-mediated cyclisation to give **54** (Scheme 2.37).



**Scheme 2.37.** The proposed *N*- or *O*-acylation, followed by ring closure to form morpholinones **53** and **54**.

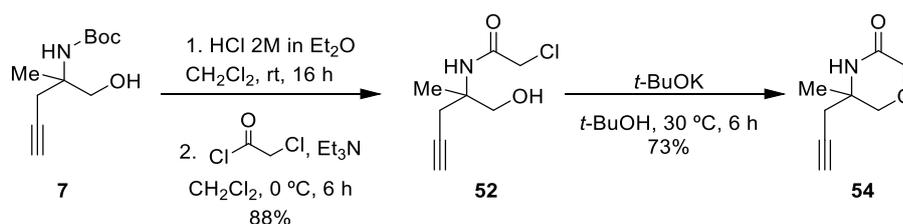
Accordingly, subjection of **7** to 1.2 equivalents of chloroacetyl chloride, in the presence of triethylamine afforded **51** in a good yield (Scheme 2.38). Then, TFA-mediated deprotection of the Boc-protecting group, followed by the exposure of the resulting TFA salt to an excess of triethylamine promoted the ring closure to afford the morpholin-2-one **53** in 76% yield over two steps.



**Scheme 2.38.** Synthesis of morpholin-2-one **53**.

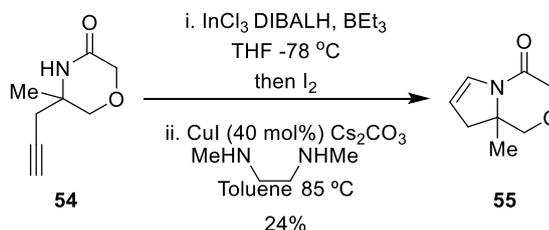
Similarly, the alternative regioisomer of **54** could also be furnished in a two-step fashion (Scheme 2.39). This was achieved through firstly Boc group deprotection under acidic conditions, revealing the free amine *in situ*, followed by *N*-acylation *via* the addition of chloroacetyl chloride and excess base. Following the isolation of **52**, the resultant intermediate

could then be cyclised through the addition of *t*-BuOK as a base, promoting intramolecular substitution to afford the desired morpholin-3-one **54** in a good overall yield.



**Scheme 2.39.** The synthesis of morpholin-3-one **54**.

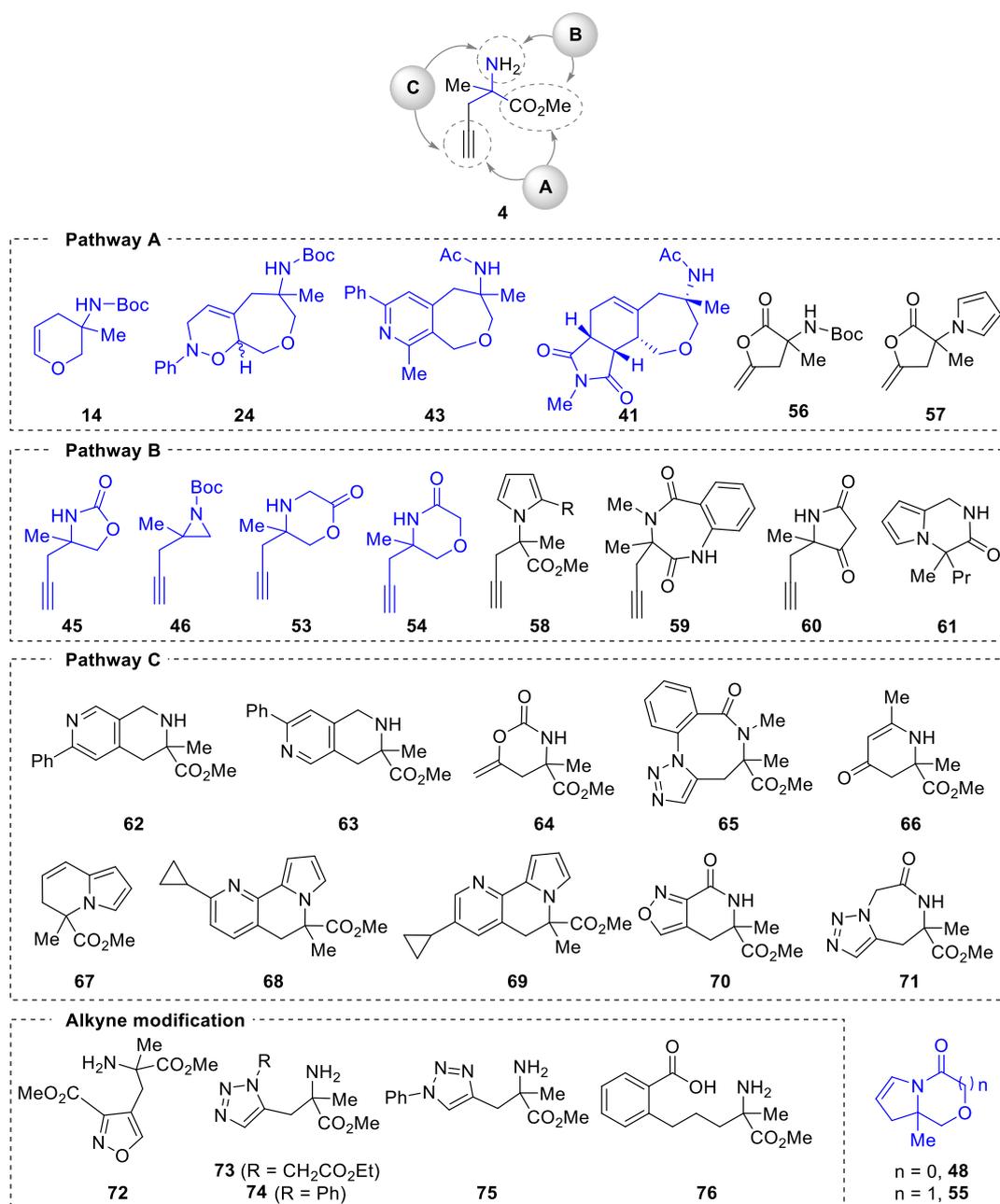
Furthermore, analogous to our previous examples (see Scheme 2.35), it was envisaged that the terminal alkyne handle within these scaffolds would provide the opportunity for further rigidification and/or scaffold expansion to furnish additional medicinally relevant heterocycles (Scheme 2.40). Thus, it was proposed that intramolecular cyclisation of the alkyne with the preinstalled amide functionality within **54** could be used to generate fused scaffold **55**. Indeed, using the same reaction conditions as for the formation of **48**, the 6,5-bicyclic scaffold **55** was isolated albeit in a low yield. Unfortunately, however, attempts to modify the alkyne within **53** and **54** using Click chemistry proved unsuccessful using both Ru- and Cu- catalysed methods.



**Scheme 2.40.** Scaffold expansion of **54** to give the constrained 6,5-heterocycle **55**.

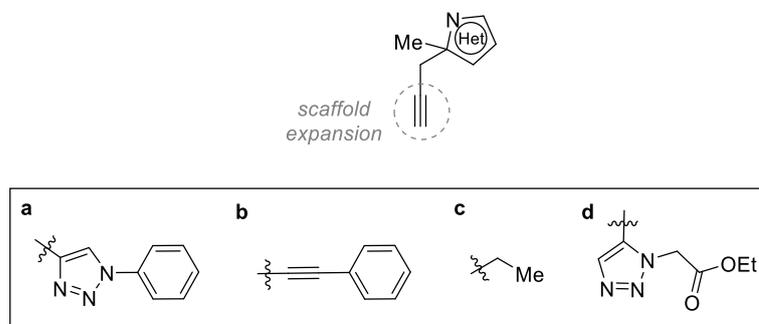
## 2.4 The complete DOS library

In addition to the synthetic efforts described herein, research carried out by Dr. Natalia Mateu (Spring Group) resulted in the synthesis of a further 28 compounds *via* investigations into utility of the amino ester **4** directly within a DOS strategy (data not shown).<sup>232</sup> In total, 31 structurally diverse compounds were constructed (Figure 2.10). The resultant combined library featured a broad range of heterocyclic scaffolds, polar and medically relevant motifs and the key *N*-substituted quaternary centre throughout. Importantly, the efficiency of the DOS approach was maintained since no more than five (and an average of only three) synthetic steps were required to access each scaffold starting from the amino ester **4**.



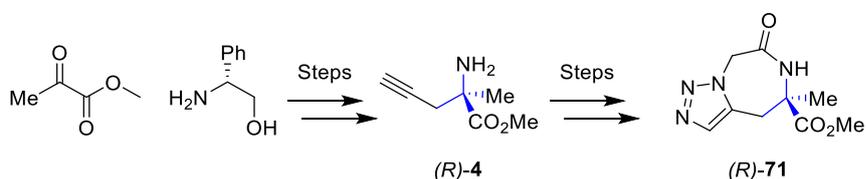
**Figure 2.10.** The full library of compounds synthesised from amino ester **4**. Compounds in black synthesised by Dr. Mateu. The synthesis of compounds in blue is described in this thesis.

Furthermore, through these joint efforts, the utility of the alkyne within several scaffolds was exemplified *via* diversification of five scaffolds to give nine additional compounds. This included either a Cu-mediated 1,4- substituted triazole formation, Sonogashira reaction to produce substituted alkynes, Pd-mediated reduction to an alkyl chain, or Ru-promoted regioselective 1,5- substituted triazole formation (Figure 2.11, see appendix for structures).



**Figure 2.11.** Exemplification of additional viable alkyne modifications.

Importantly, whilst the library was synthesised in a racemic manner for screening purposes, a flexible asymmetric route to access both (*R*)- and (*S*)- enantiomer of the amino ester was also developed by Dr. Mateu (Scheme 2.41). In this manner, upon the identification of a screening hit the desired enantiomer of a given library member could be synthesised, demonstrated through the synthesis of (*R*)-**71** from (*R*)-**4**.<sup>232</sup>



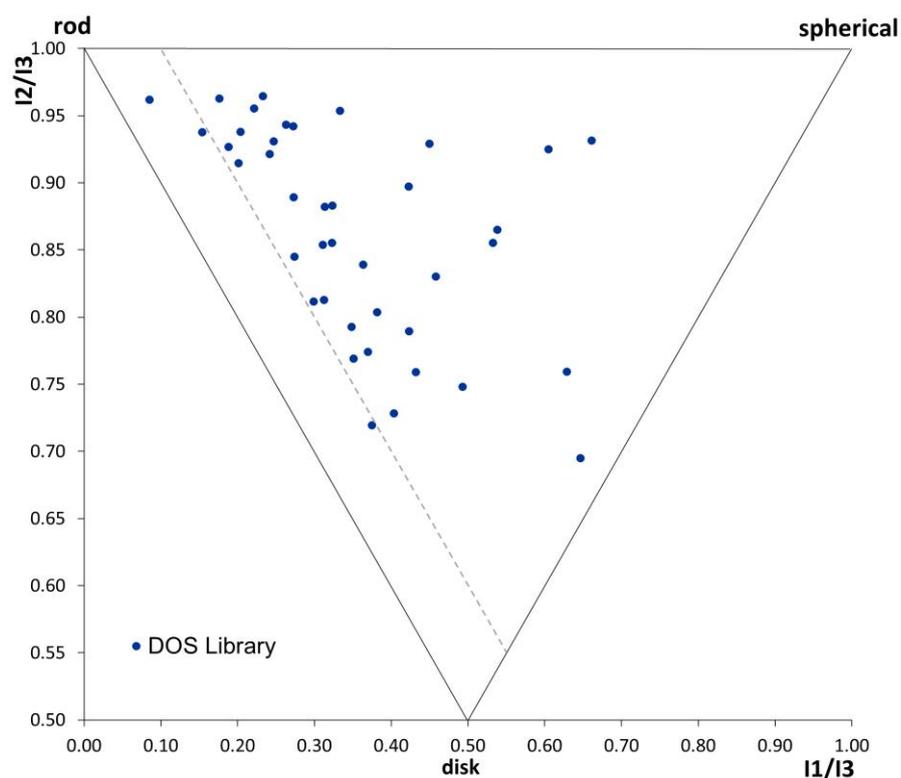
**Scheme 2.41.** The asymmetric route to (*R*)-**4** developed by Dr Mateu and demonstration of the synthesis of a corresponding library member.

## 2.5 Cheminformatic assessment of the DOS library

### 2.5.1 PMI analysis

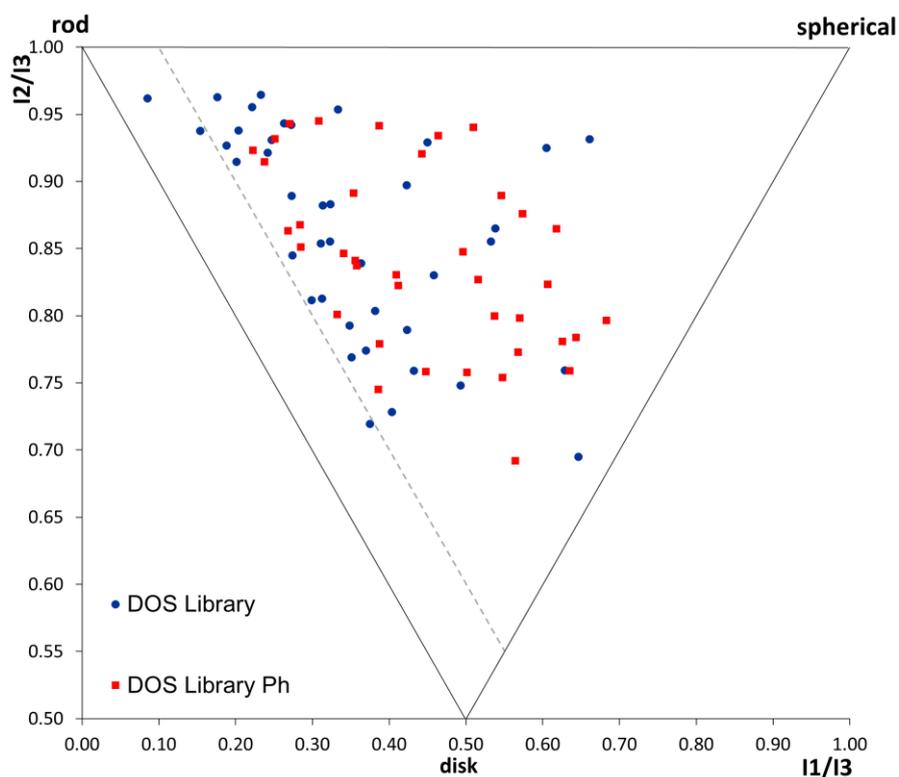
With the DOS library in hand, focus was subsequently placed on investigating computational methods of assessing the diversity of the library. PMI analysis is commonly adopted to investigate the 3-dimensional molecular shape diversity of a given compound collection.<sup>233</sup> This technique requires the computation of the normalised ratios of principal moment of inertia ( $npr1/npr2$ ) for each molecule, which are plotted on a two-dimensional triangular graph. The three vertices of the graph represent the three extreme shapes of rod, disk and sphere and therefore allow assessment of the 3-dimensionality of a given collection.

PMI analysis of the DOS library was conducted using the Molecular Operating Environment software package 2012.10 from the Chemical Computing Group (see Appendix for details). This data highlighted the compounds access a broad region of the molecular shape space with rod to spherical-like features and few disk-like compounds (Figure 2.12). Furthermore, when considering the boundaries of  $sp^2$ -rich 'flatland' as compounds with  $npr1 + npr2 < 1.1$ , as proposed by Morley *et al.*,<sup>144</sup> pleasingly only 7% of the DOS library appeared within this region further highlighting the 3-D nature of the molecules.



**Figure 2.12.** PMI plot analysis of the DOS library (blue circles). Dashed line represents  $npr1 + npr2 = 1.1$  and the boundary of flatland.<sup>144</sup>

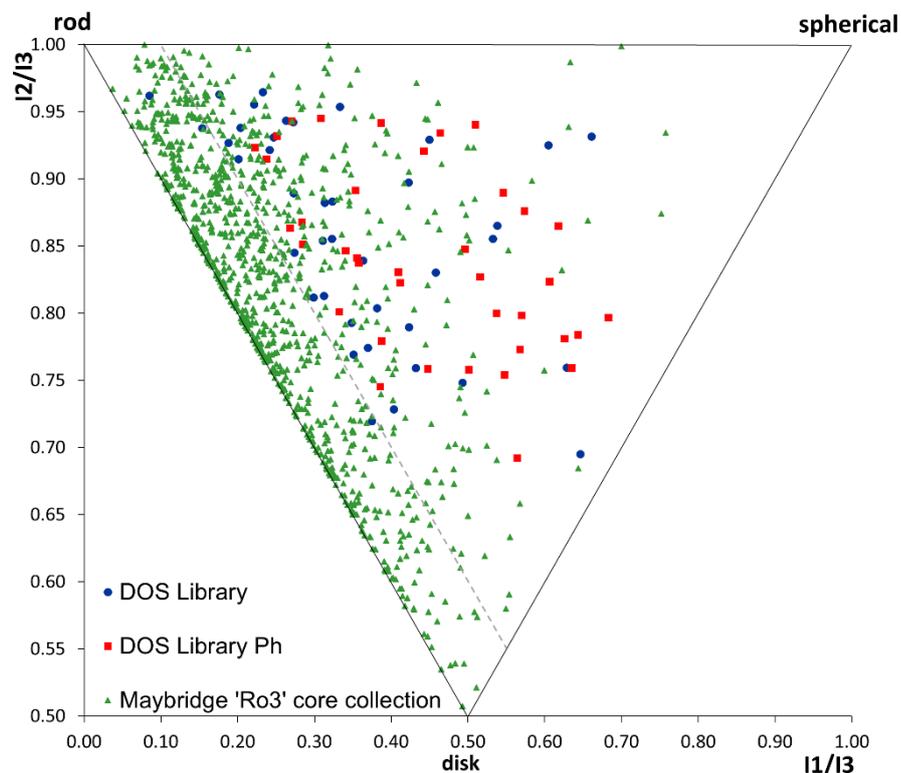
In addition to computing values for the synthesised DOS library, it was thought that virtual enumeration could be utilised to ‘construct’ an analogous collection whereby the phenyl quaternary centre had been installed throughout. Subsequent PMI analysis and comparison with the methyl collection would allow conclusions to be drawn on the effect of changing the quaternary centre on the molecular shape diversity. Accordingly, the fragments were enumerated, and as expected upon modification of the quaternary centre, PMI analysis indeed demonstrated that the molecular shape distribution was in fact altered (Figure 2.13). In this case, the Ph DOS library tended toward the right hand more spherical region of the plot, whilst importantly now 0% of the fragments appeared within flatland. Since molecular shape can be considered an essential factor in determining protein binding, it could therefore be hypothesised that the installation of the modifiable quaternary centre throughout the DOS library could serve as an important tool to modifying molecular shape and therefore the binding of a compound upon hit identification.



**Figure 2.13.** PMI plot comparing the DOS library and virtual Ph DOS library.

Finally, comparisons were then made between the commercially available Maybridge ‘Rule of Three Diversity Core Collection’ fragment set comprising 1000 compounds and the two DOS libraries (Figure 2.14). Staggeringly, using the same analysis it was found that 72% of the commercial library lay within the flatland region of the plot, demonstrating the overrepresentation of flat structures within commercial fragment collections. Moreover, upon visual inspection of the Maybridge library, only two compounds were found to contain a *N*-substituted quaternary centre, exemplifying the underrepresentation of this motif within

screening collections. These results further highlight the success of the DOS library in terms of both molecular shape diversity and the incorporation of this key quaternary motif within novel screening libraries.



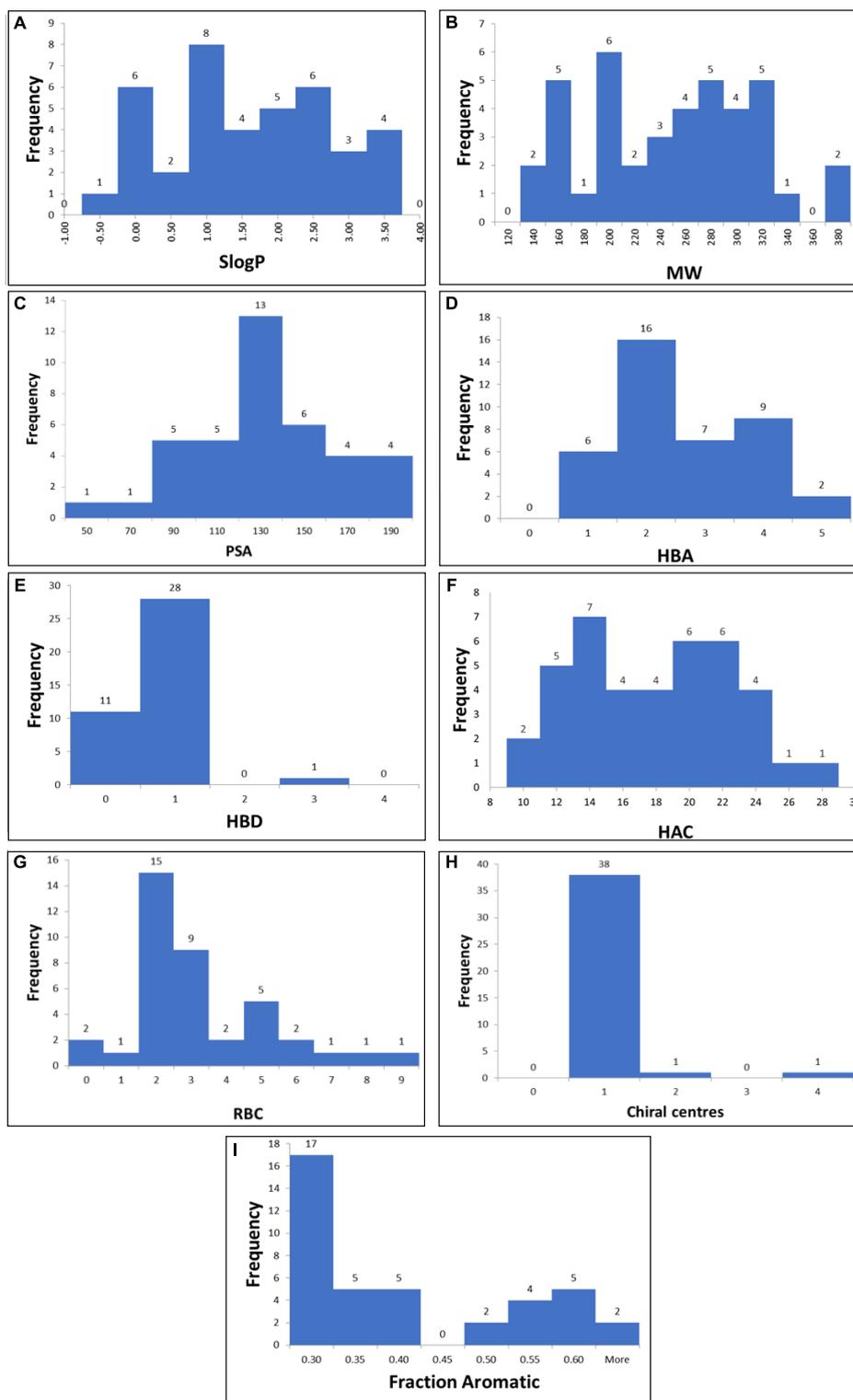
**Figure 2.14.** PMI plot analysis of the DOS library (blue circles), Ph DOS library (red squares) and the Maybridge commercial fragment set (green triangles).

## 2.5.2 Physicochemical property analysis

With the chemical space constraints of FBS in mind, computation of the physicochemical properties relating to commonly adopted guidelines within the field<sup>115,121,234,235</sup> were considered. Accordingly, the compounds were analysed for the following properties: SlogP (lipophilicity), molecular weight (MW), polar surface area (PSA), number of hydrogen-bond acceptors (HBA), number of hydrogen-bond donors (HBD), HAC, number of rotatable bonds (RBC), number of chiral centres and fraction aromatic (the number of aromatic atoms expressed as a fraction of the total number of heavy atoms). The distribution of the data is displayed in a series of histograms in Figure 2.15, A-I.

When considering the core RO3 guidelines (MW < 300, clogP < 3, HBD/HBA < 3), pleasingly, only 10% of the library had a marginally higher SlogP than three (Figure 2.15, A) only eight compounds had a MW greater than 300 (Figure 2.15, B), and all members were in-line with HBD rule (Figure 2.15, E). When considering the number of HBA, however, a significant number (28%) had more than the upper limit of three HBA (Figure 2.15, D). Despite these

findings, overall using these principles as a guideline the library showed acceptable properties for use within FBS.



**Figure 2.15.** Histograms showing the distribution of predicted physicochemical properties for compounds in the DOS library.

However, when considering the supplementary rules of  $PSA < 60$  and  $RBC < 3$  (Figure 2.15, C and G), some notable discrepancies were found. Namely, these included a PSA of  $> 60$  for

all but one of the library members, and 30% of the library with a RBC > 3. Whilst not fatal to the application of the library within a FBS setting, when considering upon hit identification, these properties should be considered to improve the oral bioavailability<sup>236</sup> and/or the blood-brain barrier penetration<sup>237</sup> of the resulting compounds during the fragment elaboration process.

To enable comparison of the DOS library with those calculated for the Maybridge and Chembridge<sup>151</sup> commercial collections, the mean values of the physicochemical properties were subsequently calculated (Table 2.2). In general, the DOS library compared near equally with both commercial collections. In addition to the PSA and RBC, the mean HAC lay marginally outside the Astex guidelines. Nevertheless, importantly an increased number of chiral centres and decreased fraction aromatic were observed. Overall, these results were pleasing since the primary aim of the project was to incorporate an increased fraction of sp<sup>3</sup> atoms and a NSQC motif within a fragment library to address deficiencies within commercially available libraries, a hypothesis which was also validated through these comparisons.

Property [a]	Guidelines			DOS	Maybridge 'RO3'	Chembridge
	RO3	Astex	collection			
<b>SlogP</b>	<3	0-2	1.37	1.92	1.31	
<b>MW</b>	<300	140-230	237	182	222	
<b>PSA</b>	<60	-	120.7	104.7	53.9	
<b>HBA</b>	<3	<3	2.63	1.83	1.81	
<b>HBD</b>	<3	<3	0.78	1.01	1.04	
<b>HAC</b>	-	10-16	16.8	12.5	15.5	
<b>RBC</b>	<3	<3	3.18	1.99	3.20	
<b>Chiral centres</b>	-	0-1.5	1.10	0.14	0.27	
<b>Fraction Ar</b>	-	-	0.29	0.52	0.42	

**Table 2.2.** Mean values of the DOS fragment library compared to the Maybridge RO3 core collection and Chembridge commercially available fragment libraries. Green = within guidelines, yellow = values at the extremes of the guidelines, red = values outside the guidelines.

Importantly, in a similar vein to the application of Lipinski's RO5, several discussions within the literature have emphasised the application of RO3 as indeed only a guideline rather than a definitive boundary.<sup>235</sup> Certainly, as screens conducted by Köster *et al.*<sup>238</sup> have demonstrated, non-compliant libraries can still prove fruitful in FBS campaigns, providing high hit-rates and enabling the identification of additional chemo-type diverse hits that would otherwise have been missed through stringent application of the RO3.

Furthermore, as discussed in the Practical Fragments Blog<sup>239</sup> and reinforced in a paper published by Hann,<sup>240</sup> it is often considered that *molecular obesity* should be the major consideration for both fragment-based drug discovery (FBDD) and traditional drug discovery programs. In this case, the importance of MW and lipophilicity across a fragment library could

be considered the pivotal constraints for shaping the starting points of these campaigns. Finally, as discussed previously, larger fragments can be useful when targeting challenging biological targets such as PPIs.<sup>127</sup> Consequently, the application of the DOS library in question can therefore be further justified and it is envisioned this library will serve as an excellent complementary collection to those already available. Accordingly, in line with the second objective, we next sought to investigate the biological utility of these fragments and demonstrate the ease of analogue construction using DOS methodology, discussed in the following chapter.



# 3 Results and discussion: Investigations into the application of *N*-substituted quaternary carbon fragment libraries

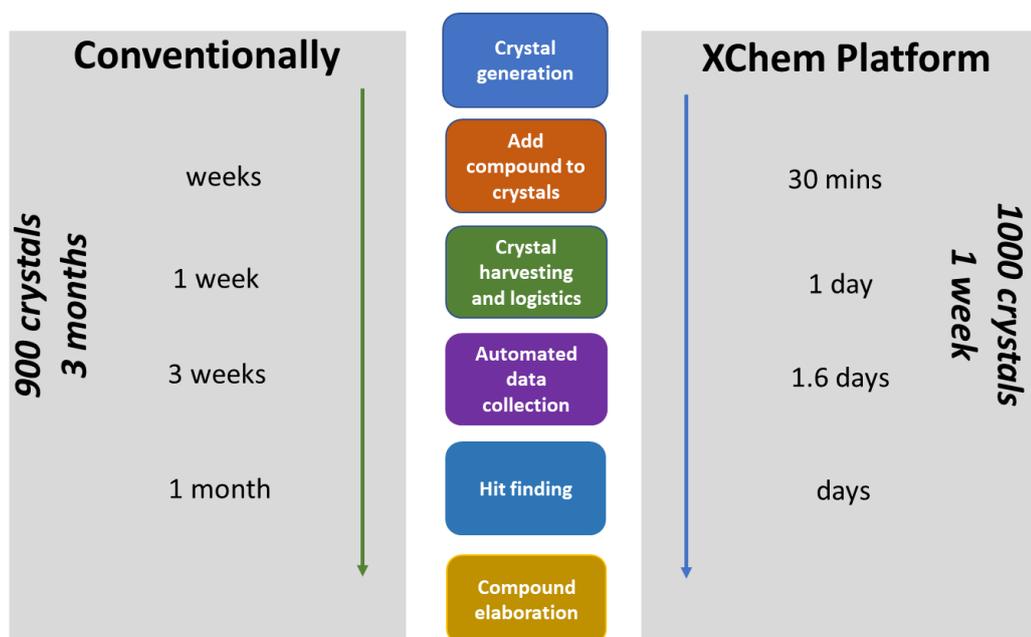
## 3.1 Background

### 3.1.1 X-ray crystallography-based methods for fragment screening

Among the many different methods utilised to conduct fragment screening, X-ray crystallography remains a well-established technique within the field of FBS in both academic and industrial settings. This can be attributed to the powerful nature of this technique to simultaneously acquire both hit identification and structural information for a given molecule bound to a target of interest.<sup>110</sup> In the early phases of FBS, X-ray crystallographic methods of screening proved fruitful<sup>134,241</sup> and the technique was used extensively, resulting in many successful medicinal chemistry programs.<sup>133,136</sup> Over time, however, presumably due to the low-throughput nature of this technique, faster and higher-throughput technologies such as SPR or NMR were often employed to conduct primary fragment screens or utilised as pre-filters to identify smaller numbers of compounds to be screened *via* X-ray methods.<sup>115,120,242,243</sup> Nevertheless, although effective it has been found that these biophysical methods can result in a large proportion of false negative results, vitally highlighting that fragment hits can evade detection by these methods when compared to X-ray based fragment screens.<sup>244</sup> Thus, the powerful ability of X-ray crystallography to identify fragments across a range of binding affinities, including very weak binders should not be overlooked.

### 3.1.2 The XChem screening platform

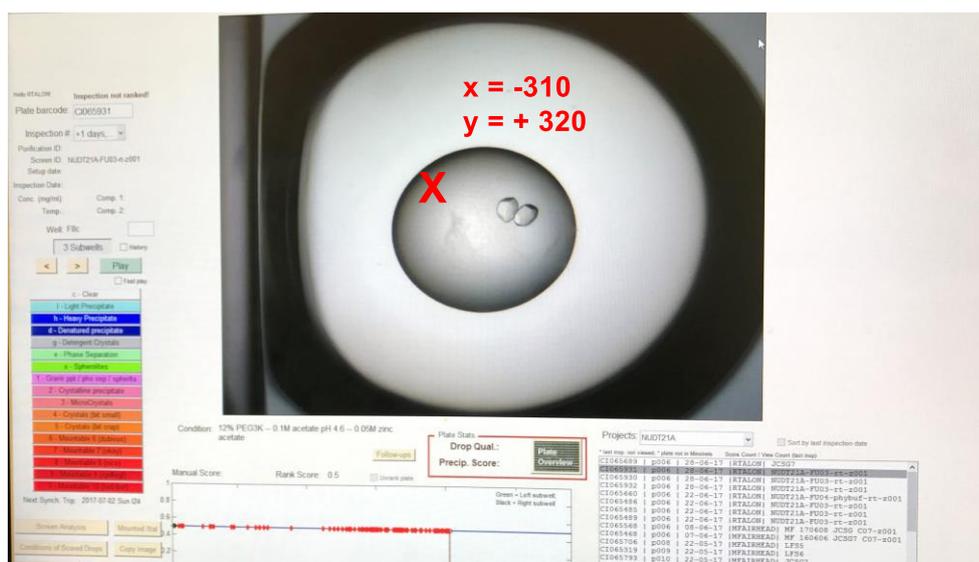
Recent developments in the foundational technologies within this screening paradigm such as third generation synchrotrons and the advancement of high-throughput technologies has however reinstated this methodology as a feasible and well-used tool for primary FBS.<sup>115,245-247</sup> Over the last few years the von Delft group have pioneered the development of a cutting-edge high-throughput fragment-screening platform, XChem, which enables rapid screening of thousands of compounds within only a 1-2-week period *via* X-ray crystallographic methods (Figure 3.1).<sup>248</sup> Based at the i04-1 beamline at the Diamond Light Source synchrotron (DLS) in Oxford, importantly this platform provides the opportunity to conduct screens in a singleton format, as opposed to cocktails (solutions containing 2-10 fragments).<sup>241,249</sup> The XChem strategy of expediting the fragment screening protocol relies on the implementation of several innovative technologies within four key stages of the traditional workflow (Figure 3.1).



**Figure 3.1.** A comparison between conventional and XChem methods for fragment screening.

### Adding compounds to crystals

Using Formulatrix® plate imagers, images of the crystal plates can be viewed over 1-, 4-, 7-day and weekly time intervals to monitor the crystal growth in each well. Through subsequent application of *TexRank* algorithm developed by Tsing Ng *et al.*<sup>250</sup> the resulting images are ranked according to their likelihood of containing crystals for manual inspection. During this process, the desired crystal(s) and precise location within the well for the soaking event to occur are then chosen, denoted by x-y coordinates given from the centre of the well (Figure 3.2).



**Figure 3.2.** Clicking a specific location on the crystal image in *TexRank* records the x-y coordinates for the crystal soaking event.

An Echo® 550 (Labcyte, Inc., Sunnyvale, California, USA) acoustic liquid handler is used for dispensing compounds in singleton format onto the required crystal wells, allowing for compound diffusion into the protein binding sites through the solvent channels within the crystal lattice.<sup>251</sup> Aside from significant speed advantages (the soaking of 1000 crystals within only 10 minutes), the accuracy of this machinery results in the precise shooting of 2.5 nL aliquots of compound solution to specific locations within a crystal well. As a result, compound solutions can be accurately delivered to a specific site (pre-defined using *TexRank*) as to avoid crystal breakage.

### **Crystal harvesting**

Whilst the crystal harvesting process is conducted manually, the employment of the *Shifter* (robotic x-y staging linked to a microscope)<sup>252</sup> significantly increases the output of this process. Utilising robot controlled mechanical navigation of the crystal plate under the microscope, the user is automatically sequentially directed to the pre-selected crystal wells required for harvesting where each crystal can be extracted using Mitegen loops. The implementation of this technology has been particularly effective, enabling the harvesting of typically >100 crystals per hour, even by an untrained user. Upon cryogenic cooling, these samples can then be stored in pucks within dewars until data collection is required.

### **Data collection**

Near continuous data collection has been achieved within the XChem platform through the implementation of a BART robot, facilitating the automated collection of 700 crystals within a 24-hour period, allowing for fully unattended beamtime.<sup>253</sup>

### **Hit finding**

The analysis of X-Ray diffraction data to generate crystal structures is non-trivial, requiring the co-operation of several software packages to complete each stage of the process. In response to this challenge, Krojer *et al.*<sup>254</sup> developed the XChemExplorer (XCE) tool to provide a graphical workflow to enable the seamless transition between the tools required to process, solve and refine protein-ligand structures. This interface retrieves information stored within the corresponding SQLite database, *SoakDB*, which collates the data collected during the previous four stages for a given crystal. Following the workflow, diffraction data can be retrieved from auto-processing pipelines using the *Datasets* tab, initial maps and ligand constraints are generated using DIMPLE, AceDRG and RDkit in the *Maps & Restraints* tab, whilst the *Hit Identification* and *Refinement* tabs are utilised to identify binding fragments and generate the corresponding crystal structures through the employment of the Pan-dataset density analysis (*PanDDA*) algorithm and *Coot* (Figure 3.3).

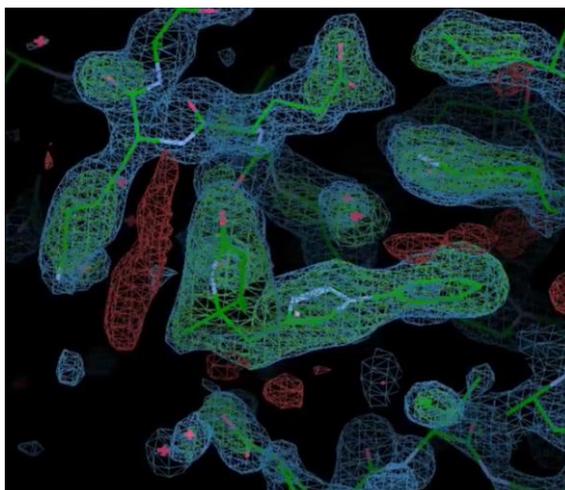
The screenshot shows the XChemExplorer application window. At the top, there is a menu bar with 'File', 'Datasource', 'Preferences', 'Deposition', 'Proasis', and 'Help'. Below the menu bar, there are tabs for 'Overview', 'Datasets', 'Maps', 'PANDDA', 'Refinement', 'Deposition', and 'Settings'. The main area displays a table with the following columns: Sample ID, Compound ID, Smiles, Visit, Resolution [Min<(sig(I))> = 1.5], Refinement Rfree, Data Collection Date, Puck, PuckPosition, and Ligand Confidence. The table contains 18 rows of data, numbered 60 to 77. At the bottom of the window, there is a control panel with four tabs: 'Update Tables From Datasource', 'Datasets', 'Maps & Restraints', 'Hit Identification', and 'Refinement'. Each tab has a 'Run' button and a 'Status' indicator. The 'Datasets' tab is currently active, showing 'Get New Results from Autoprocessing'. The 'Maps & Restraints' tab shows 'Run DIMPLE on selected MTZ files'. The 'Hit Identification' tab shows 'pandda.analyse'. The 'Refinement' tab shows 'Open COOT'. The status bar at the very bottom indicates 'idle'.

Sample ID	Compound ID	Smiles	Visit	Resolution [Min<(sig(I))> = 1.5]	Refinement Rfree	Data Collection Date	Puck	PuckPosition	Ligand Confidence
60	NUDT21A-x0282	NM424	NC(C)(C)(OC)...	1.7	None	07/07/2017 10:36	DLS347	3	None
61	NUDT21A-x0283	NM442	C=C(C)...	None	None	07/07/2017 10:37	DLS347	4	None
62	NUDT21A-x0284	NM413	C=C(O)...	1.73	None	07/07/2017 10:40	DLS347	5	None
63	NUDT21A-x0285	NM448	NC(C)(O)...	None	None	07/07/2017 10:40	DLS347	6	None
64	NUDT21A-x0286	NM144	NC(CC1)...	2.1	None	07/07/2017 10:43	DLS347	7	None
65	NUDT21A-x0287	NM474	NC(C)(O)...	1.82	None	07/07/2017 10:45	DLS347	8	None
66	NUDT21A-x0288	NM145	O=C(C)...	2.01	None	07/07/2017 10:47	DLS347	9	None
67	NUDT21A-x0289	NM393_2	O=C1C...	1.78	None	07/07/2017 10:53	DLS347	10	None
68	NUDT21A-x0290	NM395_1	O=C(N)...	None	None	07/07/2017 10:50	DLS347	11	None
69	NUDT21A-x0291	NM309_1	O=C1N...	2.52	None	07/07/2017 10:52	DLS347	12	None
70	NUDT21A-x0292	NM341	CC1(C)...	2.01	None	07/07/2017 10:55	DLS347	13	None
71	NUDT21A-x0293	NM454_2	O=C(N)...	2.08	None	07/07/2017 10:56	DLS347	14	None
72	NUDT21A-x0294	NM270_2	O=C(O)...	None	None	07/07/2017 10:58	DLS347	15	None
73	NUDT21A-x0295	NM464_1	CC(CC1)...	2.14	None	07/07/2017 11:00	DLS347	16	None
74	NUDT21A-x0296	SK342	O=C1C...	2.14	None	07/07/2017 11:02	CP50...	1	None
75	NUDT21A-x0297	SK351	O=C1C...	1.81	0.24607	07/07/2017 11:04	CP50...	2	None
76	NUDT21A-x0298	SK343	CC1(CO)...	1.94	None	07/07/2017 11:05	CP50...	3	None
77	NUDT21A-x0299	SK173	CC(C)...	None	None	07/07/2017 11:07	CP50...	4	None

**Figure 3.3.** The XCE workflow interface comprising of four tabs: Datasets, Maps & Restraints, Hit Identification and Refinement.

## PanDDA

The detection of electron density resulting from weakly binding fragments is vastly complicated since it is often challenging to identify such changed states due to the presence of fractionally present molecular states within the ground state of a protein crystal form. To overcome this challenge the *PanDDA* algorithm was developed by Pearce *et al.* in 2017.<sup>255</sup> This tool utilises the statistical analysis of density distributions within ~30 hit-free crystals to build an accurate ground state model, which can then be subtracted from the bound state data to elegantly reveal changes in electron density (Figure 3.4). As a result, the detection of these changed states, *i.e.* ligand binding, is vastly simplified. This provides an additional advantage when implemented within the XChem platform, since only the changed state events identified by the algorithm are required for manual inspection and ligand modelling/refinement. Thus, the application of this tool within the XCE interface greatly streamlines the process of hit identification within from the extreme numbers of datasets (>500) collected in such an experiment.



**Figure 3.4.** PanDDA map highlighting changed states of electron density (green and red) compared to the ground state (blue), simplifying the identification of a bound fragment hit.

### **XChem success**

Since its development in 2015, the XChem Platform has facilitated the screening of over 60 protein targets and enabled the identification of >1000 fragment hits, with users spanning both academia and industry. Moreover, this system has proven successful against a variety of protein targets, including those related to epigenetic gene regulation,<sup>248,256</sup> oncology<sup>257</sup> and neurodegenerative disease.<sup>258</sup> Thus, through the ability to identify particularly weak fragment binders and leverage structure-based design, this technology provides a powerful strategy to tackle the FBS of challenging targets.

In the early stages, commonly commercially available fragment collections such as Maybridge (see Chapter 2) were employed within this platform. Whilst these proved fruitful, due to the deficiencies of such libraries in terms of 3-dimensionality, novelty and fragment growth vectors, the XChem team had sought to supplement these collections.

## **3.2 XChem screening of fragment-like DOS libraries**

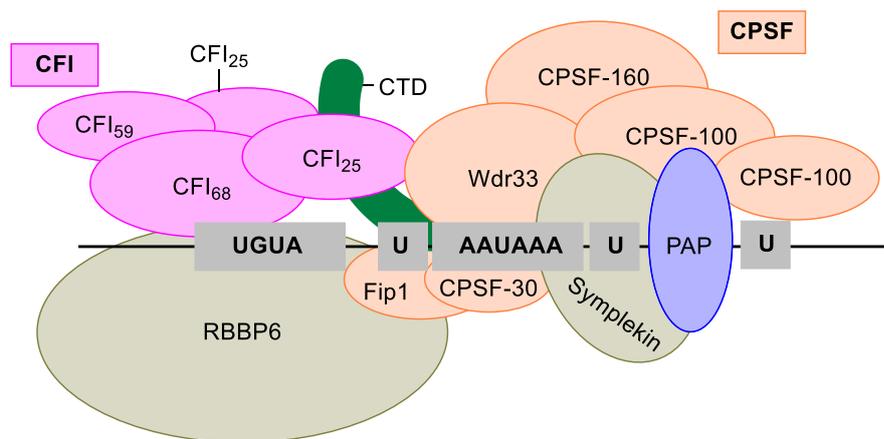
In 2017, a collaboration was established between the Spring group and XChem to enable the routine screening of 3-D fragments derived from DOS. Thus, considering the excellent properties of the NSQC compounds described in Chapter 2, these fragments were the first to be incorporated into the pipeline. This library was created using 250-500 mM stock solutions of the compounds in  $d_6$ -DMSO in line with the XChem standard protocols. To initially validate the utility of this library within this platform, a proof-of-concept fragment screen was conducted.

### **3.2.1 CFI<sub>25</sub>**

Cleavage factor 25kDa (CFI<sub>25</sub>, also known as CPSF5 or NUDT21), a promising oncology target currently under research by the Huber and von Delft groups in the Structural Genomics

Consortium, Oxford. CFI<sub>25</sub> is a member of the Nudix superfamily of hydrolases, characterised by the presence of the characteristic Nudix box generated by a conserved 23-residue sequence, which is in turn incorporated into a distinctive  $\alpha/\beta/\alpha$  fold motif. These so called 'house-keeping' proteins function to catalyse the hydrolysis of a diverse range of Nucleoside diphosphates linked to other moieties, X, Nudix substrates including nucleoside triphosphates, capped RNA and dinucleotide coenzymes. Importantly, however, CFI<sub>25</sub> is devoid of hydrolase activity due to the absence of two key glutamate residues required for catalysis, instead replaced by Leu124 and Ile128.<sup>259,260</sup>

The NUDT21 gene encodes the CFI<sub>25</sub> protein, which is a sub-unit of the pre-mRNA cleavage factor Im (CFIm) component of polyadenylation machinery (Figure 3.5).<sup>261</sup> In its entirety, CFIm consists of two CFI<sub>25</sub> units in a heterotetramer with either CFI<sub>59</sub> or CFI<sub>68</sub>. Importantly, it is thought that within this complex the CFI<sub>25</sub> component is responsible for UGUA RNA motif recognition and binding, whilst the remaining CFI<sub>59</sub> and CFI<sub>68</sub> domains facilitate RNA looping and enhance the RNA binding due to the presence of an RNA recognition motif.<sup>262</sup> Around 19 other complexes and proteins are involved in the remainder of the polyadenylation machinery.<sup>263</sup>



**Figure 3.5.** A subsection of the metazoan polyadenylation machinery composed of ~20 proteins and complexes. The CFI machinery (in pink) is formed by two units of CFI<sub>25</sub> (responsible for UGUA recognition) and either CFI<sub>59</sub> or CFI<sub>68</sub>. Adapted from reference.<sup>263</sup>

The crystal structures of the free and diadenosine tetraphosphate (Ap<sub>4</sub>A, a native Nudix hydrolase substrate<sup>264</sup>) complex of CFI<sub>25</sub> were published in 2008 by Coseno *et al.*<sup>260</sup> More recently, Yang *et al.*<sup>265</sup> similarly disclosed the crystal structure of CFI<sub>25</sub> in complex with UUGUAU RNA sequence. In both instances, interaction of these substrates with Arg63 was identified as crucial for binding between with the phosphate component of Ap<sub>4</sub>A or the U3 element of the RNA sequence. Moreover, the significance of this residue was further demonstrated through mutation of this residue to Serine, which was shown to almost entirely eradicate RNA binding.<sup>265</sup>

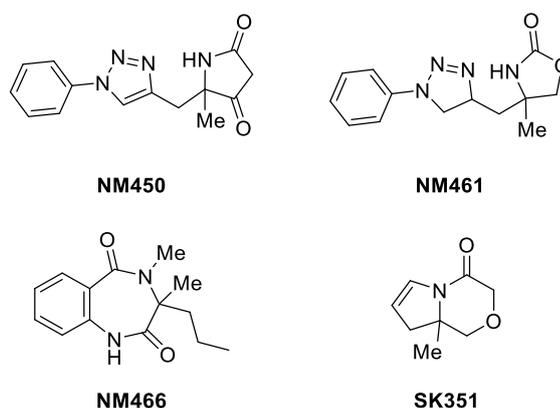
Importantly, studies have demonstrated the role of the CFI<sub>m</sub> machinery in the alternative polyadenylation (APA) of mRNA, an important mechanism involved in gene regulation.<sup>263</sup> Indeed, the CFI<sub>m</sub> complex has been shown to play a key role in determining the size of the 3' untranslated region of mRNA leading to different mRNA isoforms by mediating APA.<sup>266</sup> More specifically to the significance of CFI<sub>25</sub>, RNA interference knockdown of CFI<sub>25</sub> has revealed a significant number of genes with shortened 3' untranslated regions, which Masamha *et al.*<sup>267</sup> were able to demonstrate results in the increased expression of several oncogenes, ultimately affecting both tumour growth and tumour size in glioblastoma cells.

In addition to its potential role in cancer, studies by Gennarino *et al.*<sup>268</sup> recently demonstrated the relationship between CFI<sub>25</sub> and the post-translational gene regulation of the mRNA encoding methyl CpG-binding protein 2 (MeCP2), an important protein involved in brain function and the development of neuropsychiatric disease. Through the study of lymphoblastoid cells of patients exhibiting copy number variants of the NUDT21 gene and neuropsychiatric syndromes, significantly, differential levels of the MeCP2 protein were identified. Moreover, siRNA-mediated knockdown of NUDT21 resulted in adjustment of these levels to those of the wild-type or healthy individuals. Ultimately, the authors were able to attribute these findings to regulation in the expression of the short or long isoforms of the related mRNA by CFI<sub>25</sub>.<sup>268</sup>

Despite these pivotal findings, however, to the best of our knowledge there are currently no literature small molecule chemical probes for CFI<sub>25</sub>. As a result, the identification of novel fragment hits for this purpose represented an exciting avenue of research to enable further investigations into the associated biology.

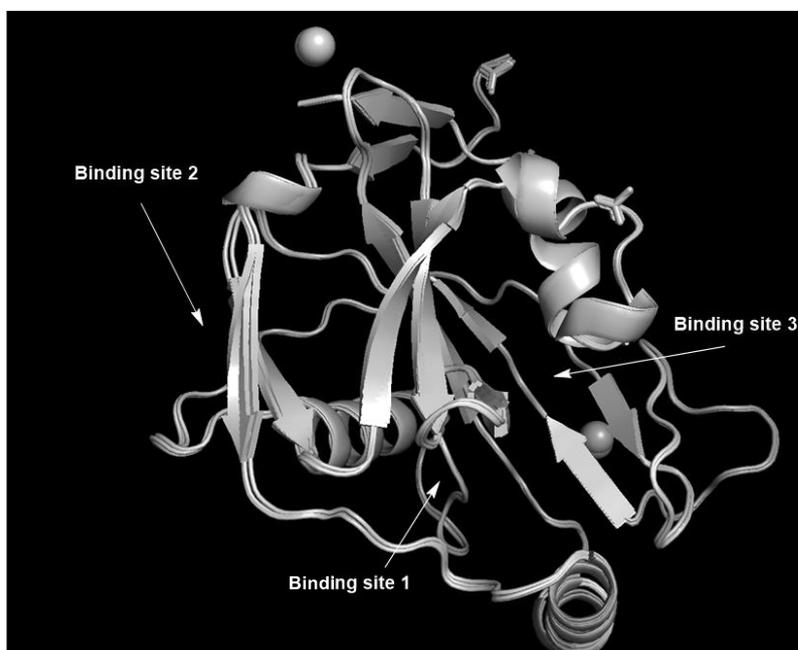
### **X-Ray screening**

Due to the robust and replicable nature of CFI<sub>25</sub> crystals, this protein was deemed amenable to fragment screening *via* the XChem platform in collaboration with the Structural Genomics Consortium. Thus, in July 2017 a fragment screening campaign was conducted against CFI<sub>25</sub> utilising the NSQC library. Upon analysis of the resulting PanDDA event maps, four X-Ray hits were identified (Figure 3.6). Importantly, as a result of the diverse nature of the DOS library these hits related to several different chemotypes, highlighting the potential of this collection to deliver hits of varied molecular architecture.



**Figure 3.6.** The four hits identified during the XChem screen against CFI<sub>25</sub>.

Superimposing the resulting crystal structures of these hits revealed three different fragment binding sites within the protein (Figure 3.7).

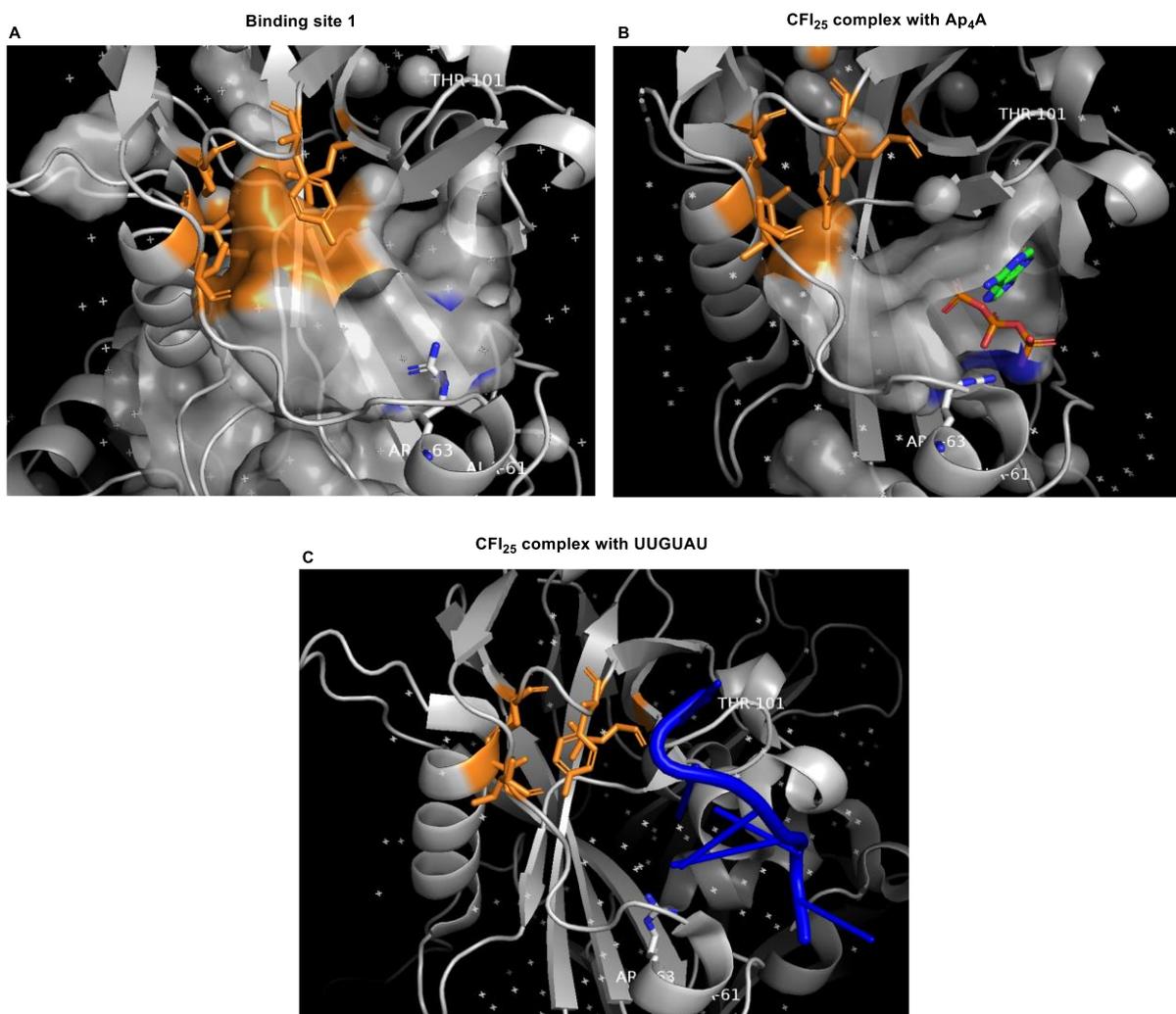


**Figure 3.7.** An overlay of the crystal structures containing hits highlighted three different fragment binding sites.

### Binding site 1 hits

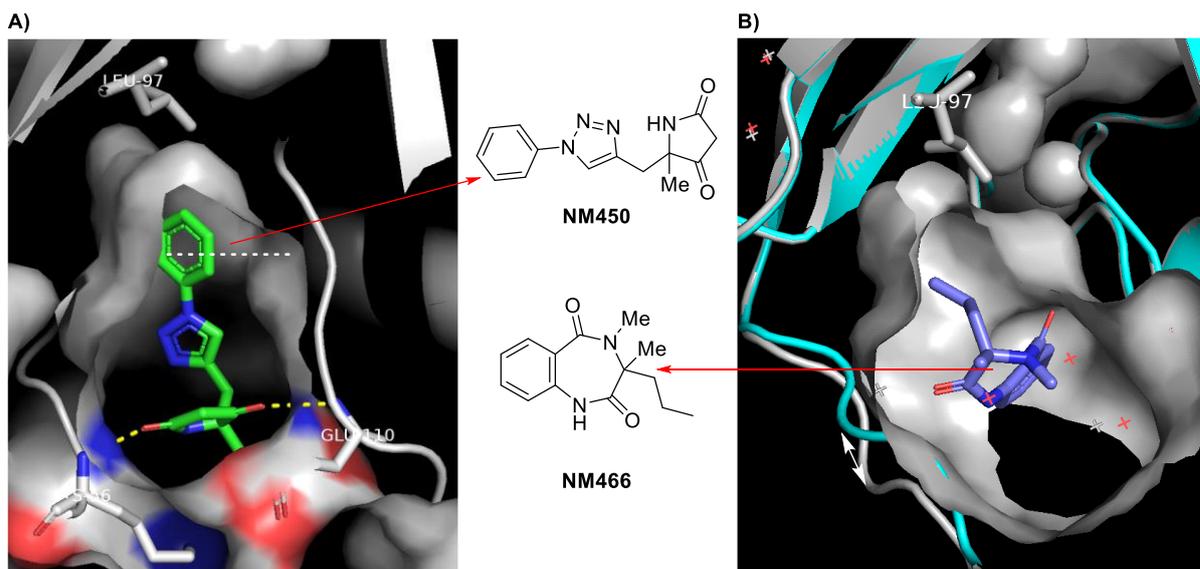
Binding site 1 (BS1) was found to be the most significant with both **NM450** and **NM466**, and within a single hydrophobic pocket formed by five residues; Leu53, Leu97, Ile128, Pro188 and Tyr191 (Figure 3.8, A). Of significant importance is the location of BS1 within Nudix box region of the protein (residues 109-131)<sup>260</sup> highlighted by the role of Ile28 in the formation of the pocket. Contextualisation of this binding site through comparison to the binding site for the Ap<sub>4</sub>A<sup>260</sup> and the UUGUAU RNA sequence<sup>265</sup> between Thr101 and Ala61 revealed BS1 to be in close proximity, located at the back of the open cavity where these moieties bind (Figure 3.8, B and C). Interestingly, comparison of the Ap<sub>4</sub>A, UUGUAU-bound structures with those

of the DOS fragments revealed a substantial shift in the location of Tyr191 in the latter case, resulting in an expanded lipophilic binding pocket.



**Figure 3.8.** A) The hydrophobic pocket BS1 formed by Leu53, Leu 97, Ile128, Pro188 and Tyr191 (orange sticks) in relation to the natural substrate binding location marked by Thr101 and Ala61 (white labels) and Arg63 (blue-grey sticks). B) Reported crystal structure of CFI<sub>25</sub> in complex with Ap<sub>4</sub>A,<sup>260</sup> (green and red sticks) with the residues responsible for BS1 highlighted in orange. C) Reported crystal structure of CFI<sub>25</sub> in complex with UUGUAU,<sup>265</sup> with the residues responsible for BS1 highlighted in orange.

Significantly, it was found that upon binding of **NM450**, this hydrophobic pocket was further extended in an open conformation, mediated through additional movement of Leu97 to accommodate the phenyl ring (Figure 3.9, dashed white line). It is worth mentioning that whilst the electron density for the phenyl and triazole components proved to be well defined, the pyrrolidinone portion was ambiguous. In this regard, determination of the enantiomer preference of the protein was not possible. Modelling of both enantiomers, however, revealed interactions could be made in both instances and thus it was hypothesised that the pyrrolidinone portion positively contributed toward the binding.

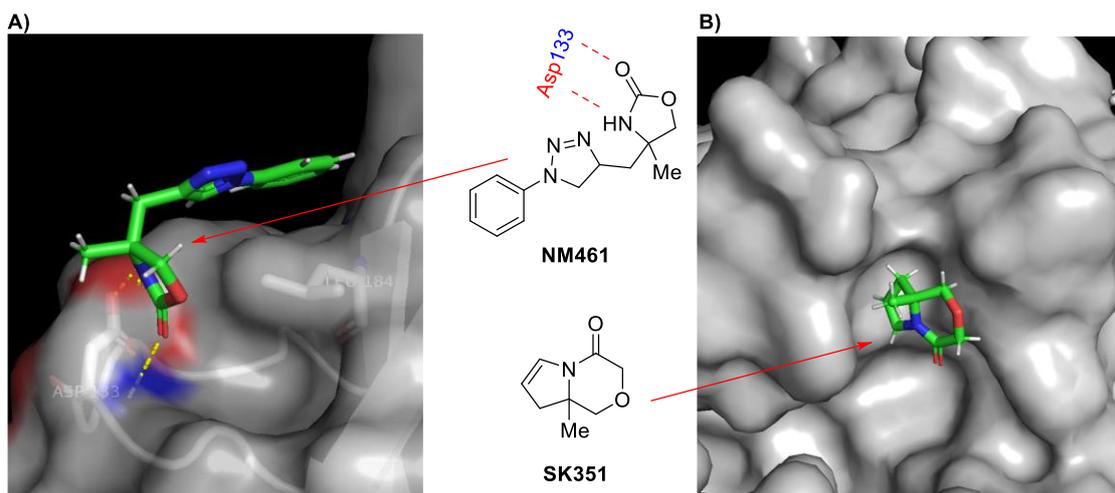


**Figure 3.9.** Crystal structures of A) **NM450** (green sticks) and B) **NM466** (purple sticks) within the main pocket of  $CFI_{25}$ . In (A) the dashed white line demonstrates the closed conformation of the lipophilic pocket whilst in (B) the protein backbone for **NM450** is shown in cyan and the white arrow highlights the significant loop shift between the two fragments binding. Dashed yellow lines represent polar interactions with the protein.

The crystal structure of **NM466** within  $CFI_{25}$ , revealed the fragment to sit in an alternative binding mode to **NM450**. Here, the key hydrophobic pocket was found in the closed conformation, with the aromatic component instead binding in the position of the triazole ring of **NM450** (Figure 3.9, B). This observation was corroborated through comparison of the corresponding binding pockets and backbone location, highlighting a significant loop shift between that observed in the binding of **NM450** (highlighted in cyan) to **NM466**.

### Binding sites 2 and 3

The second two binding sites, Binding site 2 (BS2) and Binding site 3 (BS3), represented two shallow grooves on the surface of the protein. During the screen, one fragment was identified as binding within each site (Figure 3.10). The binding of **NM461** in BS2 was unexpected due to the similarity in structure to **NM450**, differing only through the quaternary heterocycle present. This result was not presumed to be an anomaly due to the replication of this binding event within the duplicate of run the screen. Instead, two potential hydrogen bonds between the Asp133 residue and oxazolidinone were believed to anchor **NM461**, whilst the aromatic portion bound tightly over a hydrophobic shelf created by Leu184.

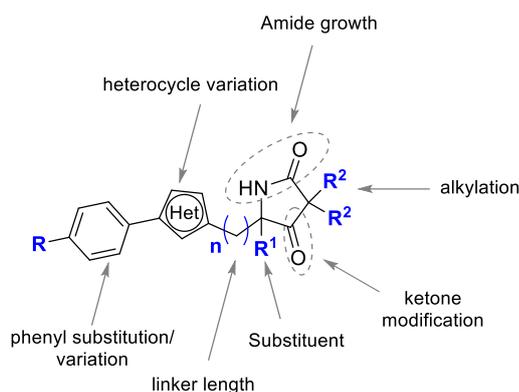


**Figure 3.10.** Additional binders were identified to bind in two further sites within  $CFI_{25}$  A) **NM461** within BS2 and B) **SK351** within BS3.

In the case of **SK351**, however, no polar contacts could be predicted between the fragment and the protein. In view of this and the location along the crystal contacts of the protein, this hit was deemed to be insignificant.<sup>269</sup>

Importantly, these exciting results not only provided the opportunity to highlight the advantage of the DOS-derived libraries for analogue formation, but additionally served as initial starting points for lead-generation. Thus, two objectives were envisaged: (1) rapid multidirectional derivatisation of the hits to showcase the modular chemistry and explore the binding modes, followed by (2) screening of further analogues to provide hit validation.

To address the first objective, it was anticipated that the amenability of the DOS chemistry toward multi-directional vector growth could be demonstrated *via* derivatisation to almost every functionality within **NM450** (Figure 3.11). Specifically, in line with the structural data, it was thought these investigations could include modification of the benzene ring through substitution, variation in the bridging heterocycle, modification of the quaternary substituent and derivatisation of the pyrolidinone heterocycle *via* amide growth,  $\alpha$ -alkylation and ketone modification. Moreover, it was also important to demonstrate the potential to access varied derivatives in a rapid fashion, owing to the highly modular DOS strategy.



**Figure 3.11.** Potential structural modifications to **NM450**.

The second objective of hit validation, however, was more challenging. Whilst X-Ray crystallography can be used to identify very weakly binding compounds, as discussed often issues can arise in the detection or validation of these compounds using orthogonal biophysical methods.<sup>244</sup> Moreover, due to the novel and challenging nature of the protein in question, the development of these techniques was non-trivial. Initial attempts within the Huber group to characterise the eight hits using biophysical methods (SPR and an AlphaScreen) proved unsuccessful. Whilst, several of the fragments could in fact be detected (data not shown), inconsistencies in the results achieved with both techniques suggested further optimisation would be required to produce reliable and consistent results. Investigations into this and the application of additional techniques for this purpose is on-going within the Huber group.

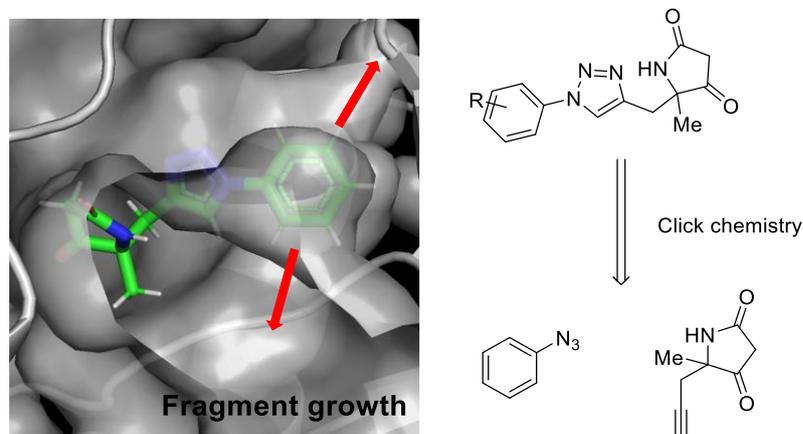
As an alternative initial hit validation strategy, the XChem platform had previously found success in validating initial X-Ray hits through the screening of related analogues. Following this procedure, a second round of synthesis was generally first conducted to generate small libraries of derivatives. Through screening of these libraries false positives could then be identified and true positives validated upon binding of related analogues, simultaneously providing structural binding information for subsequent medicinal chemistry explorations. Thus, the derivatives formed for the first objective would additionally serve as useful tools in this context.

To exemplify this, analogues of **NM450** were initially pursued due to the interesting binding mode observed. It is worth mentioning, that due to the lack of conclusive data regarding the stereochemistry of the bound fragment hit, a racemic synthesis of the corresponding analogues was initially pursued.

### 3.2.2 Follow-up compounds to **NM450**

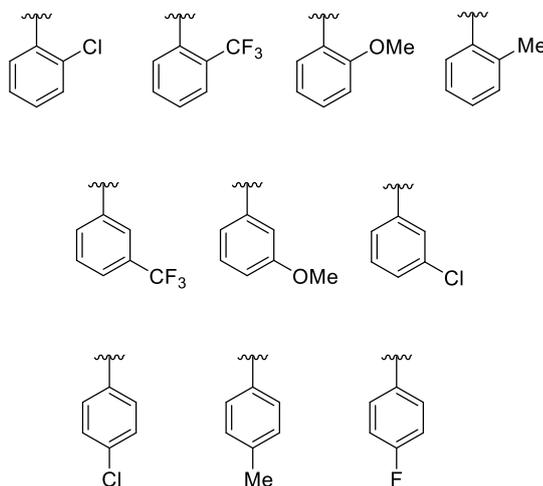
Following the proposed synthetic plans (*vide supra*), firstly, it was envisioned that initial access to derivatives bearing aromatic substituent variation would serve as simple analogues that could be used to confirm the binding, whilst exploring the scope of several growth vectors. Due

to the presence of this moiety within an open pocket, it was anticipated that a variety of substituent vectors would be viable (Figure 3.12). Moreover, utilising the modularity of the DOS route, it was hoped that analogues of this nature could be easily generated through disconnection of the triazole ring to its constituent aromatic azide and heterocyclic alkyne components.



**Figure 3.12.** The crystal structure of **NM450** suggested substitution of the aromatic ring to be a viable at a range of positions. Disconnection at the triazole ring would enable rapid analogue synthesis.

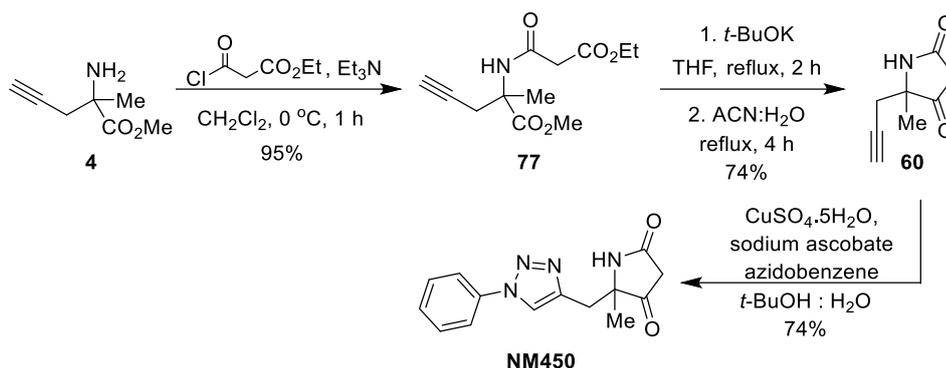
Computational modelling predicted both electron-withdrawing and electron-donating groups in *ortho/meta/para*-positions could be tolerated. Thus, ten commercially available aromatic azides were selected incorporating functionalities at different *ortho/meta/para* positions, in addition to variation in the size and polarity of the functionality (Figure 3.13).



**Figure 3.13.** The substituted analogues chosen for synthesis.

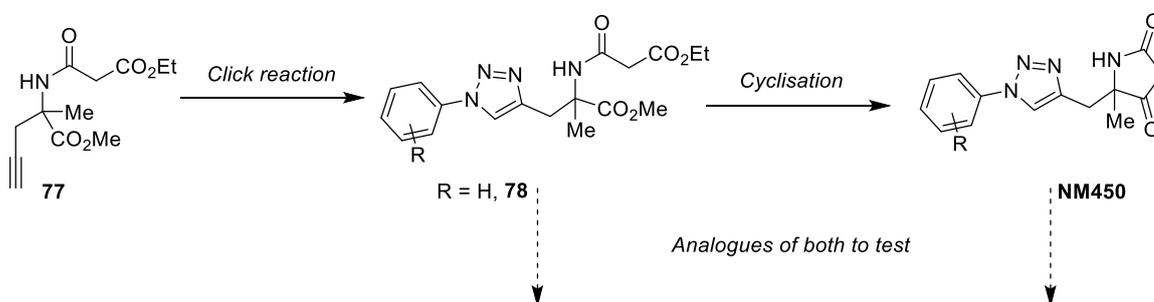
The original synthesis to **NM450** involved acylation of amino ester **4** to give **77**, followed by Dieckmann condensation and decarboxylation to yield alkyne **60**. In turn, subjecting the terminal alkyne within **60** to Cu-mediated regioselective 1,4-triazole formation using azidobenzene under standard click conditions afforded **NM450** (Scheme 3.1).<sup>232</sup> However, this

click reaction proved to not be reproducible upon variation of the aryl-azide and as a result an alternative synthetic route was designed.



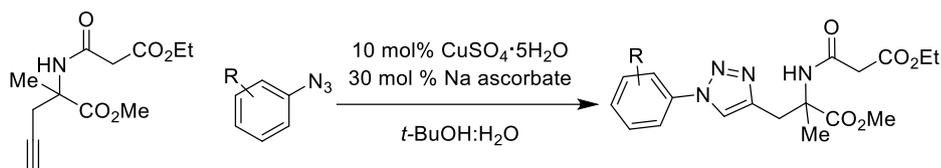
**Scheme 3.1.** Original synthesis to **NM450** developed by Dr. N Mateu.

It was hypothesised that alternatively, the aromatic substituents could be introduced at the previous step *via* click reaction of the linear acylated moiety **77**. Subsequent cyclisation could then be achieved following the same procedure as previously established, to isolate the desired analogues of **NM450** (Scheme 3.2). Furthermore, it was hoped that generation of the substituted intermediates of type **78**, in addition to the fully cyclised compounds would be advantageous for biological evaluation to provide additional SBR datapoints. Thus, a gram-scale synthesis of **77** was successfully conducted following the pre-established procedure.



**Scheme 3.2.** The proposed route to afford multiple analogues for screening purposes.

As an initial proof of concept, the click reaction of 1-azido-2-(trifluoromethyl)benzene with **77** was trialled (Table 3.1, Entry 1). It is worth mentioning that a 1:0.9 ratio of alkyne to azide was selected to mitigate isolation of potentially explosive unreacted azide substrate. Accordingly, **77** was subjected to standard Copper click reaction conditions, affording **79** in a good 69% yield. Satisfied by this result, expansion of this methodology to the remaining nine azides was subsequently conducted.

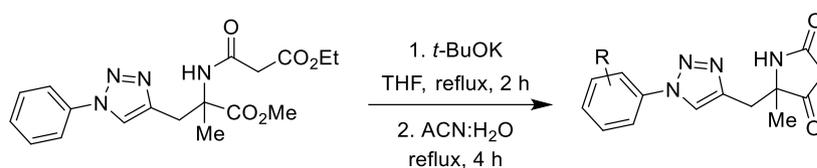


Entry	Equiv. of alkyne	Equiv. of azide	R	Yield	Product
1	1	0.9	<i>o</i> -CF <sub>3</sub>	70%	<b>79</b>
2	1	0.9	<i>o</i> -Cl	69%*	<b>80</b>
3	1	0.9	<i>o</i> -OMe	65%	<b>81</b>
4	1	0.9	<i>o</i> -Me	12%*	<b>82</b>
5	1	0.9	<i>m</i> -OMe	69%	<b>83</b>
6	1	0.9	<i>m</i> -CF <sub>3</sub>	84%	<b>84</b>
7	1	0.9	<i>m</i> -Cl	68%	<b>85</b>
8	1	0.9	<i>p</i> -F	83%	<b>86</b>
9	1	0.9	<i>p</i> -Cl	74%	<b>87</b>
10	1	0.9	<i>p</i> -Me	66%	<b>88</b>

**Table 3.1.** The Cu-mediated click reaction of **77** with ten substituted aromatic azides proceeded smoothly, except in the case of entry 4, generating **79** - **88**. \* = second batch of CuSO<sub>4</sub>·H<sub>2</sub>O added.

Incorporation of *ortho*-chloride and methoxy substituents proceeded smoothly yielding **80** and **81**, however a significantly lower yield was obtained using 2-azidotoluene (Table 3.1, entries 2-4). It could be hypothesised that this could have been a result of the diminished electron withdrawing effect induced by the methyl substituent relative to the -Cl and -OMe groups, reducing the nucleophilicity of the corresponding *ortho* azido functionality. Attempts to improve the yield *via* the addition of a further 20 mol% of Cu and increased reaction time proved in vein. Gratifyingly, however, the remaining six transformations all proceeded smoothly affording the *meta* and *para* substituted products **83** – **88** in moderate yields (Table 3.1, entries 5-10). Concurrently, investigations into the compatibility of this reaction using a flow reactor were initiated due to the potential to provide both a safer and faster route to access substituted triazoles of type **78**.<sup>270</sup> Preliminary efforts in this regard, however, had limited success.

Next, compounds **79** – **88** were taken forward for cyclisation with the aim of affording the desired pyrrolidinone analogues **89** – **97** (Table 3.2, entries 1-10). Applying the standard Dieckmann condensation conditions followed by thermal decarboxylation indeed returned analogues **89** – **97** in good yields. Due to the lack of material isolated during the formation of **82**, no cyclisation was attempted and instead only **82** was retained for biological evaluation (Table 3.2, entry 4).

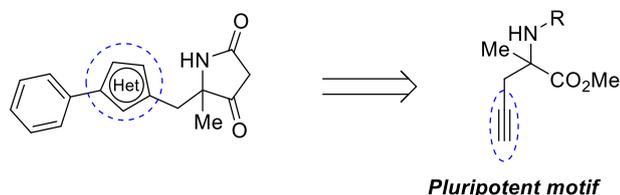


Entry	Starting material	R	Cyclisation	Product
1	79	<i>o</i> -CF <sub>3</sub>	64%	89
2	80	<i>o</i> -Cl	51%	90
3	81	<i>o</i> -OMe	62%	91
4	82	<i>o</i> -Me	NA	NA
5	83	<i>m</i> -OMe	78%	92
6	84	<i>m</i> -CF <sub>3</sub>	74%	93
7	85	<i>m</i> -Cl	66%	94
8	86	<i>p</i> -F	73%	95
9	87	<i>p</i> -Cl	62%	96
10	88	<i>p</i> -Me	70%	97

**Table 3.2.** Cyclisation of substrates **79** – **88** under standard Dieckmann conditions followed by decarboxylation to yield **89** - **97**.

## Heterocycle analogues

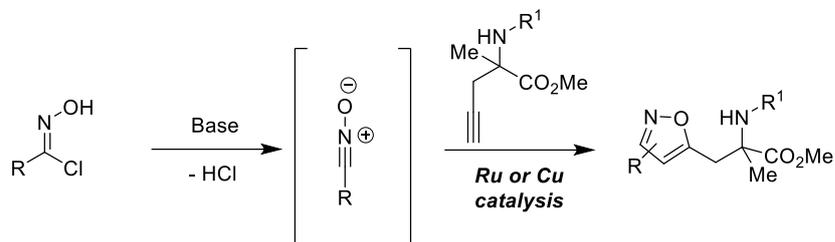
In addition to producing analogues of **NM450** bearing variation in the aromatic substituents, it was envisioned that alternative bridging heterocycles could additionally be introduced, showcasing the versatility of the DOS route to introduce alternative moieties at this position. Indeed, variation in the position of heteroatoms within an aromatic cycle in such a manner is a commonly adopted technique within medicinal chemistry pursuits. Accordingly, it was expected this could be achieved utilising the pluripotent nature of the terminal alkyne present within the key building block to generate alternative heterocycles in a modular approach (Scheme 3.3).



**Scheme 3.3.** The alkyne motif could be used as a pluripotent handle to generate alternative bridging heterocycles.

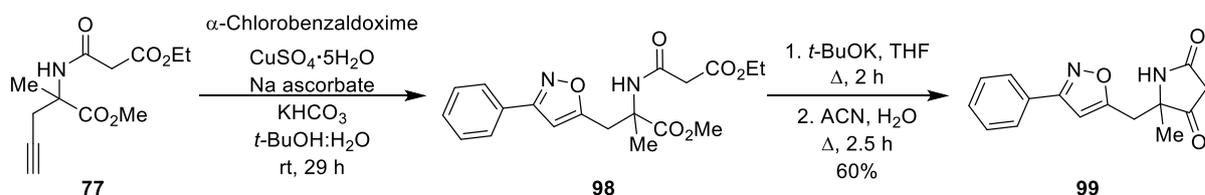
In addition to the reactivity of alkynes with azides to form triazoles, chlorooximes have been used in a similar fashion, whereby under basic conditions a reactive nitrile oxide 1,3 dipole can be generated (Scheme 3.4).<sup>271</sup> In turn, reaction of this species with a suitable alkyne under

either Ruthenium or Copper catalysis can produce 1,5- or 1,4- substituted isoxazoles, respectively, in a regioselective manner.<sup>272,273</sup> In relation to the binding of **NM450** within CF1<sub>25</sub> pocket, investigation into the binding mode of the fragment *via* the synthesis of both isoxazole isomers proved an interesting avenue of research.



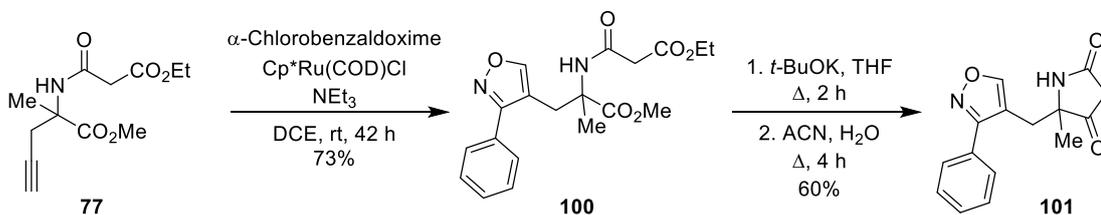
**Scheme 3.4.** The formation of a reactive nitrile oxide to generate 1,4- or 1,5- isoxazoles regioselectively.

To this aim, alkyne **77** was subjected to Copper-mediated cycloaddition conditions<sup>274</sup> to furnish the 1,4- substituted isoxazole **98** (Scheme 3.5). Following the pre-established procedure, this could then be cyclised to generate the analogue **99**.



**Scheme 3.5.** 1,4- regio selective isoxazole formation, yielding **99**.

As previously mentioned, it was hypothesised that the alternative 1,5-regioisomer of **101** could also serve as an interesting SBR data point. Consequently, following conditions developed by Grecian *et al.*,<sup>273</sup> reaction of **77** under Cp<sup>\*</sup>Ru(COD)Cl catalysis with  $\alpha$ -chlorobenzaldoxime and base, 1,5-triazole **100** was isolated in good yield (Scheme 3.6). Finally, following the pre-established cyclisation conditions **101** was afforded in good yield.

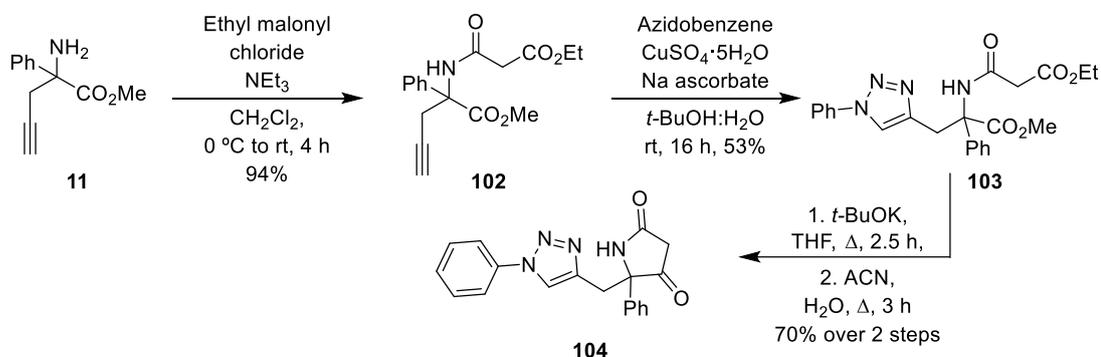


**Scheme 3.6.** Ru-mediated formation of the 1,5-isoxazole regioisomer **100**.

## Quaternary centre modifications

With the crystal structure of **NM450** in mind, it was hypothesised that the utility of the quaternary centre growth vector within **NM450** could likewise be demonstrated. This functionality appeared to sit within a larger, more open pocket within the binding site (Figure 3.9), and it was therefore envisioned that much larger substituents could be tolerated at this

position. With the phenyl amino ester already in hand (see Chapter 2), it was proposed that this could be utilised to generate the corresponding phenyl variant of **NM450**. Thus, amino ester was acylated using the pre-established conditions, affording **102** in an excellent 94% yield (Scheme 3.7). Subsequent reaction under the Cu-catalysed click conditions then afforded **103**, whilst finally condensation followed by decarboxylation, delivered analogue **104**.

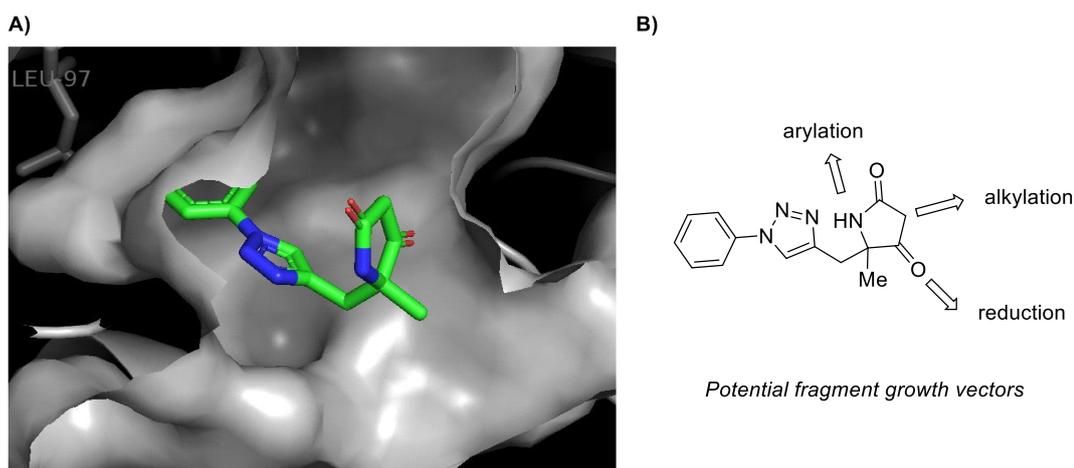


**Scheme 3.7.** Synthesis of **104**, an analogue of **NM450** bearing an alternative quaternary substituent.

Since the prior aim of this research was to investigate the rapid synthesis of **NM450** analogues for hit validation, further derivatives of this nature were postponed until validation of **NM450** binding had been acquired. It was anticipated, however, that following the demonstration of the viability of this procedure through the formation of **104**, alternative functionalities could be indeed introduced following this route if required.

### Investigations into late-stage diversifications of **NM450**

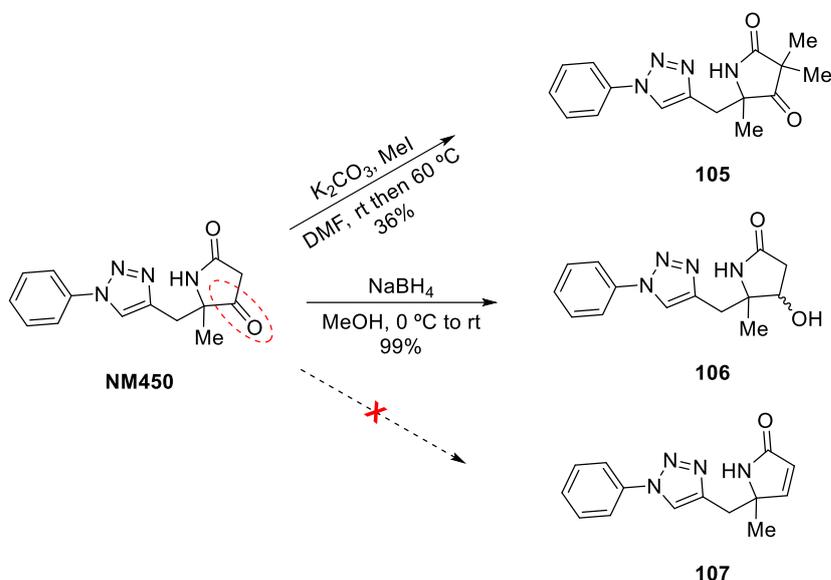
Finally, hypothesis-driven proposals of additional more 3-D modifications to the molecule based around the pyrrolidinone core were considered. Whilst the aromatic portion of the molecule was buried within the hydrophobic pocket, the pyrrolidinone core was found to be situated within a much larger channel providing scope for expansion across multiple growth vectors (Figure 3.14). Accordingly, it was envisaged that a range of transformations could be explored including amide derivatisation, ketone modification and growth from the enolisable  $\alpha$ -keto position to enable fragment growth. Importantly, to maintain synthetic efficiency we proposed these modifications could be installed using late-stage modifications to the **NM450** scaffold.



**Figure 3.14.** A) X-ray crystal structure of **NM450** bound within **CFI<sub>25</sub>**, showing the pyrrolidinone within an open pocket. B) Potential fragment growth vectors of interest.

### Ketone modifications

With the pluripotent nature of the ketone functionality within **NM450** in mind, we hypothesised that several additional derivatives could be easily accessed *via* modification of this moiety. It was envisaged that this could be achieved following three different strategies: (1)  $\alpha$ -alkylation, (2) ketone reduction or (3) ketone elimination (Scheme 3.8).

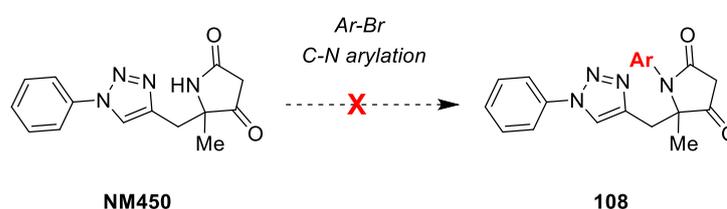


**Scheme 3.8.** Investigating methods to modify the ketone functionality within **NM450**.

Encouragingly upon treatment of **NM450** with MeI under basic conditions the dimethylated species **105** was isolated, albeit in low yield (Scheme 3.8). Furthermore, subjecting of **NM450** to sodium borohydride mediating reduction of the ketone functionality afforded a mixture of diastereomers of the resulting alcohol product **106** in almost quantitative yield. However, attempts to remove the ketone *via* mesylation followed by DBU-prompted elimination yielded only starting material with no evidence of the desired product **107** observed. It is envisaged

that on-going investigations into the applicability of differences bases could prove fruitful in this regard.

We next considered extension of the molecule from the amide vector through the application of various amide arylation conditions to furnish **108** (Scheme 3.9). The JackiePhos ligand was developed in response to challenges associated with the *N*-arylation of secondary amides, substrates which often prove challenging with more conventional Buchwald-type ligands. During ligand development it was established that the specific substitution of the JackiePhos ligand increased the electrophilicity, whilst reduced the steric demands of the Palladium intermediate, encouraging the otherwise rate-limiting transmetallation step.<sup>275</sup> Thus, this seemed an ideal starting point to facilitate the arylation of the secondary amide within **NM450**. Accordingly, investigations began using the commercially available JackiePhos Pd G3 pre-catalyst system.

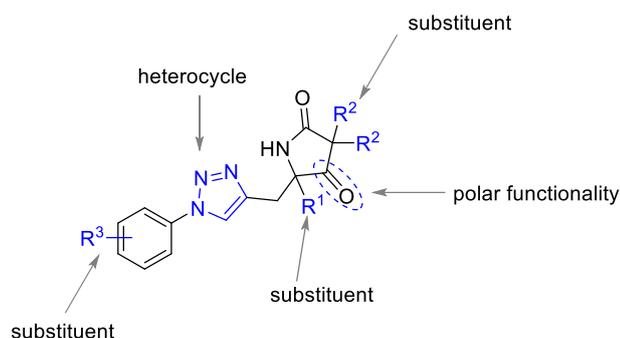


**Scheme 3.9.** Attempts to modify the amide nitrogen proved unsuccessful.

Discouragingly, however, it was found that alteration to the equivalents and nature of the aryl bromide in addition to the temperature and duration of microwave heating proved unsuccessful, with only trace amounts of product observed in all cases. In light of the failures experienced using the JackiePhos Pd Buchwald conditions, alternative Cu-catalysed Ullmann-type coupling conditions were considered. Consequently, **NM450** was subjected to CuI and iodobenzene in the presence of DMEDA as a ligand under thermal heating conditions. Disappointingly, in this case only a complex mixture was obtained with no evidence of **108** formation observed by <sup>1</sup>H NMR.

In a final attempt to install functionality at this position, it was anticipated that **NM450** may be amenable to alkylation methods to introduce further functionality. Thus, **NM450** was treated with benzyl bromide under basic conditions. Encouragingly, complete consumption of the starting material was observed in this case. However, <sup>1</sup>H NMR analysis of the resulting compound suggested alkylation had instead taken place at the  $\alpha$ -keto position, with some evidence of di-alkylation present. With these findings in mind, it was hypothesised a more viable route to modify the amide nitrogen position would instead involve installation of functionality at an earlier and less sterically hindered stage of the synthesis, for example prior to cyclisation of **78** or triazole formation from **77**.

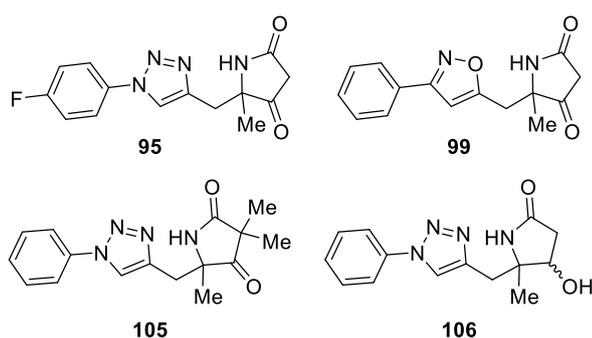
Considering the first objective, the highly modular DOS strategy successfully enabled the rapid synthesis of 26 **NM450** derivatives. Building on previous research in the field of DOS fragments for hit evolution,<sup>153</sup> in this example we have demonstrated how this can be achieved in a multidirectional fashion through leveraging the inherent modularity, the NSQC motif and sp<sup>3</sup> carbons (Figure 3.15). Thus, in addition to the physicochemical properties (see Chapter 2), we have highlighted the potential of the NSQC library to address key outstanding requirements within the field of FBDD.



**Figure 3.15.** These investigations resulted in 26 derivatives exploring a number of fragment growth vectors.

### 3.2.3 Hit validation

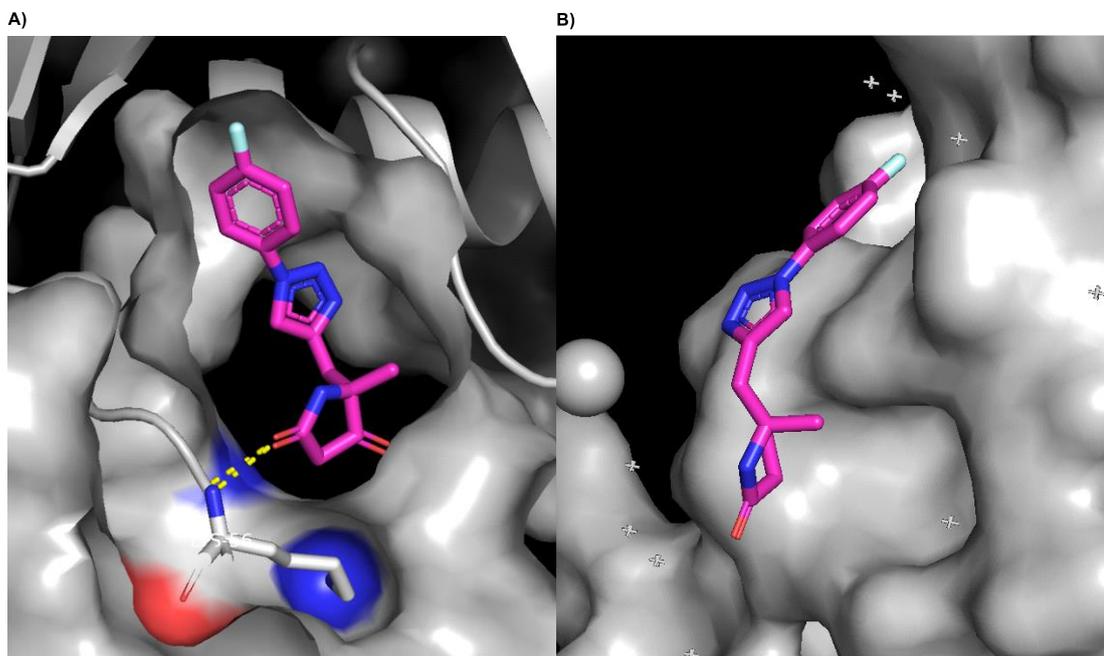
With the library of 26 analogues in hand, a second round of X-Ray screening was subsequently conducted to address the second objective. In a similar fashion, this was completed at the XChem screening facility at DLS, and all compounds were screened in duplicate. Gratifyingly, upon analysis of the resulting data using PanDDA four hits were successfully identified (Figure 3.16).



**Figure 3.16.** The four hits identified from the follow-up X-Ray screen.

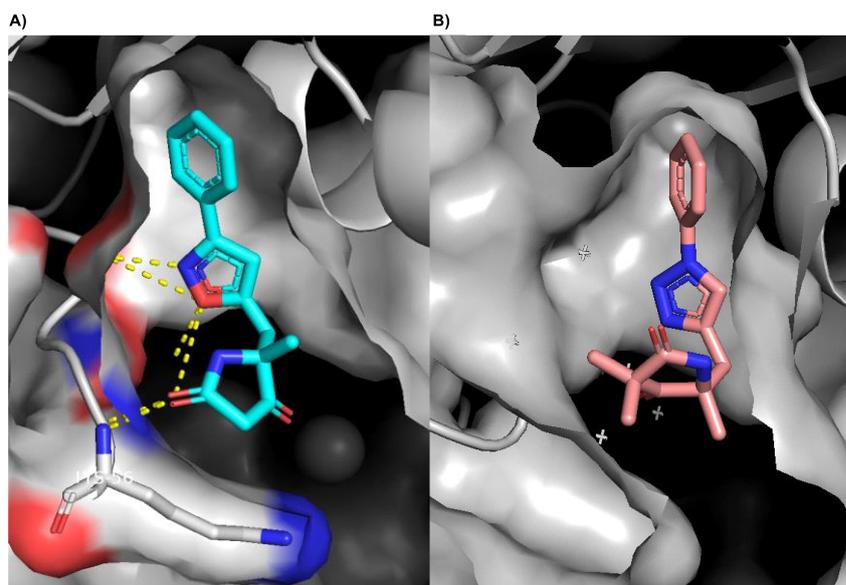
Evaluation of the structural X-Ray data generated revealed that in relation to the ten substituted aromatic analogues screened, only the *para*-fluoro analogue **95** was in fact tolerated (Figure 3.17, A). This data revealed the aromatic portion of the molecule to bind in a similar fashion to **NM450** within BS1, with the amide carbonyl interacting with Lys56 within the protein back bone. Intriguingly, however, a second portion of electron density was identified for the molecule within

BS2 with the same binding mode as **NM461** (Figure 3.17, B). It was hypothesised, however, that this could be a crystallographic artefact, potentially binding between the crystal contacts. This theory was corroborated by the fact no polar interactions could be identified between the molecule and protein.



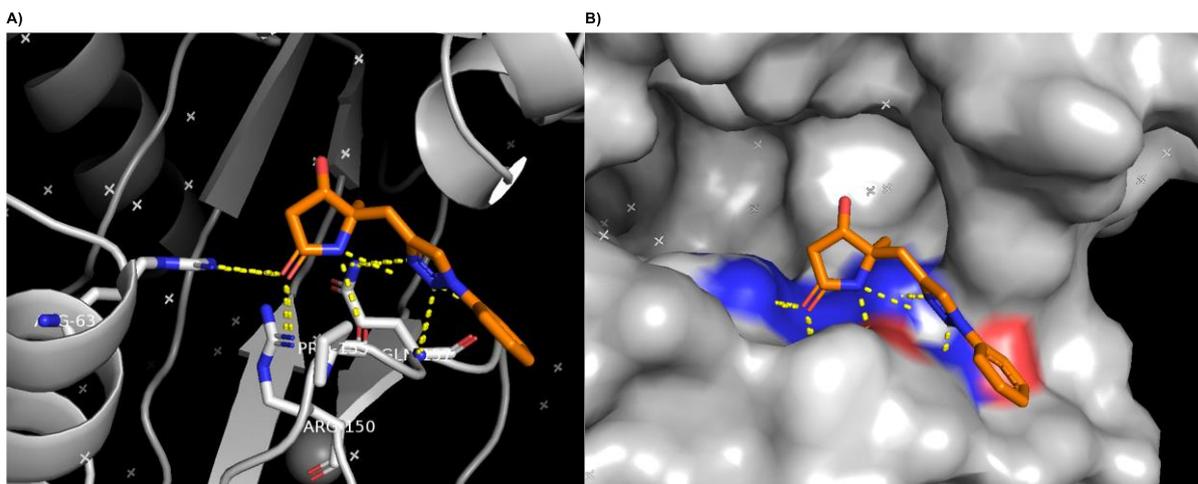
**Figure 3.17.** The crystal structure of **95** within A) BS1 and B) BS2.

The binding of **99** revealed that the alternative isoxazole bridging heterocycle could also be tolerated. Importantly, selectivity for the 1,4-regioisomer could be inferred from these results since no binding of the 1,5 isomer **101** was identified. Inspection of the structural data revealed **99** bound in the same manner as **NM450** within BS1 (Figure 3.18, A). In this instance polar interactions could be predicted between the amide carbonyl and the isoxazole heteroatoms within the CFI<sub>25</sub> backbone. Similarly, **101** also exhibited this binding mode (Figure 3.18, B). Interestingly, in this case the *gem*-dimethyl substituents and NSQC were oriented toward different channels within the protein, suggesting these could be utilised as two alternative 3-D growth vectors from the molecule.



**Figure 3.18.** The crystal structures of A) **99** and B) **105** within BS1 of CFI<sub>25</sub>.

Finally, the most significant and unexpected result was the data for the hydroxyl variant **106**. Upon analysis of the crystal structure it was found that this compound bound within the substrate channel for both ApA4 and the UUGUAU motif (*vide supra*, Figure 3.8). Moreover, importantly this predicted a polar interaction between the amide carbonyl and the key Arg63 residue responsible for ApA4 and UUGUAU binding and Arg150. Moreover, further interactions could be identified between the amide nitrogen toward Arg63 and the triazole with Gln157 (Figure 3.19, A and B). It is worth mentioning, however, as with **NM450** and all bound derivatives once more the aromatic region proved to be much more defined, whilst the electron density for the quaternary heterocycle was ambiguous. Thus, whilst these interactions could be hypothesised, screening of the single diastereomer and enantiomer variants of **106** would provide vital information into the true binding preference and spatial orientation of the heterocycle.



**Figure 3.19.** The crystal structure of **106** bound within the substrate channel of CFI<sub>25</sub> with the predicted hydrogen bonds and important residues highlighted in A) and the substrate binding channel surface highlighted in B).

In view of these results, the initial **NM450** hit was successfully validated through the identification of the binding of additional analogues, meeting the second objective proposed. Consequently, **NM450** therefore serves as a novel lead for the development of the first small molecule modulators of CFI<sub>25</sub>. Thus, as this example demonstrates, DOS fragment libraries can be effective in delivering new hits against challenging targets. It is envisioned that future efforts will involve hit evolution in line with the results described utilising the chemistry developed and crystal structures presented herein. In particular, one important objective would involve the synthesis of the single enantiomer variants of these compounds to determine the enantiomer preference of the protein to aid the application of structure-based design for fragment growth. Moreover, as discussed previously, methods to measure the binding affinity of these fragments using biophysical techniques are currently under development within the Huber group. In due course, once established, it is expected these this data will enable prioritisation of subsequent fragment evolution and guide the development of novel leads for this important biological target.

Importantly, as a result of its success, the library was subsequently made openly available to external XChem collaborators enabling the routine screening of the library.

### 3.2.4 Activin A

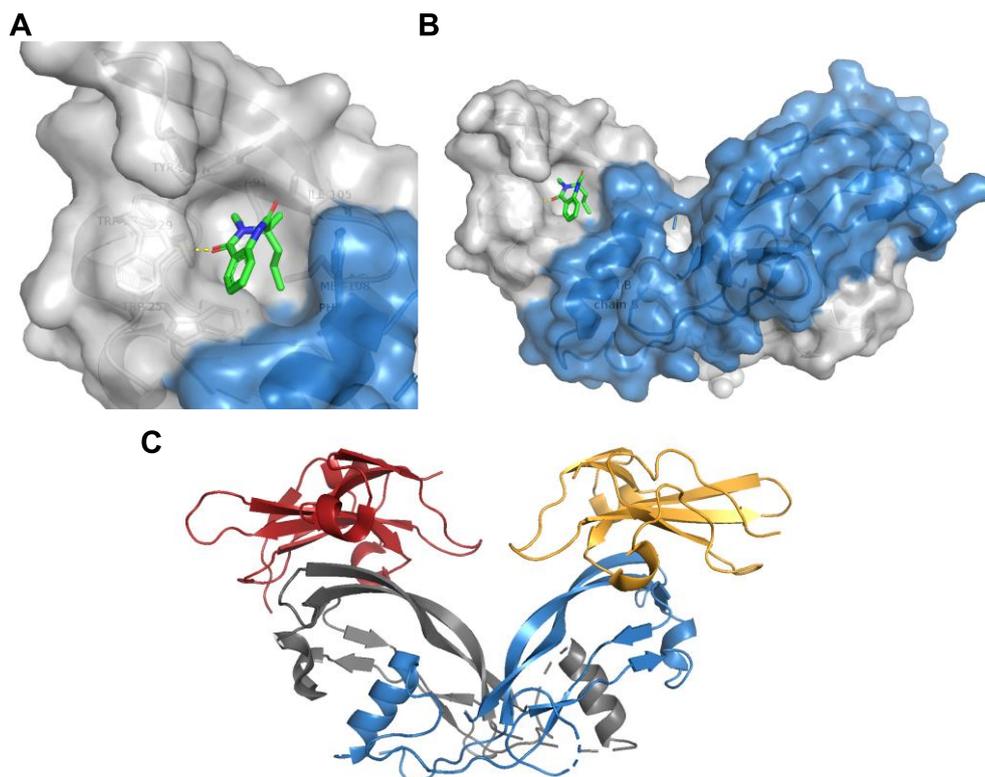
The first collaborator to utilise the NSQC library was Joseph McLoughlin from the Hyvönen group during a screen seeking to identify novel fragment binders of activin A. Activins are members of the transforming growth factor  $\beta$  (TGF- $\beta$ ) superfamily of growth factors, which play an essential roles in homeostasis and development and have been studied for many years.<sup>276</sup> Research has shown activins mediate an intramolecular signalling cascade *via* binding to the extracellular domains of transmembrane serine/threonine kinases known as Type I or Type II receptors.<sup>277,278</sup> In turn, this conduces the phosphorylation of type I receptors by the kinase domain of the type II receptors. This then enables the subsequent phosphorylation of receptor-Smad proteins, which results in their translocation to the nucleus where they are involved in the regulation of target gene expression.<sup>279</sup> Importantly, in this context, binding of the type II receptors has been shown to be crucial for type I receptor binding and therefore vital to generate the first step of this signalling pathway.

Several studies have associated the role of activin A signalling with the regulation of embryogenesis, stem cell differentiation and wound healing, among other processes.<sup>280–282</sup> Moreover, dysregulation of activin A signalling or expression has been linked to human disease such inflammatory conditions and cancer.<sup>280,283,284</sup> Indeed, recently STM434, a fusion protein of the type II receptor extracellular domain with a human Fc domain, entered clinical trials for the treatment of ovarian cancer and other solid tumours, suggesting inhibition of this protein can be exploited for therapeutic benefit.<sup>285</sup> However, negative off-target effects were observed, hypothesised to result from bone morphogenic protein 9 inhibition.<sup>286</sup>

Nevertheless, despite the potential of this target, to the best of our knowledge no small molecule modulators of this protein exist to enable further investigations into the associated biology. Following their recent success in determining the crystal structure of the pro-domain of activin A,<sup>276</sup> the Hyvönen group became interested in pursuing the development of novel activin SM modulators. This included investigating the potential to develop novel small molecule inhibitors of this protein *via* high-throughput fragment screening methods. This provided the possibility to explore an orthogonal approach to the development of stapled peptide-based libraries against Activin A, also under investigation within the both the Hyvönen and Spring research groups.

From these efforts, a single hit from the NSQC library was identified, **NM466**. Inspection of the crystal structure of this hit revealed the fragment bound within a lipophilic pocket lined by Trp25, Trp28, Ile29, Tyr93, Met91, Ile105 and Met108 of chain A with Phe58 of chain B (Figure 3.20, A). Importantly, it was believed the fragment was anchored within this site *via* hydrogen bond interaction between the carbonyl of the tertiary amide within **NM466** toward Trp28 from

chain A of the protein backbone, whilst the propyl chain bound tightly over a shelf generated by Phe58 of chain B.

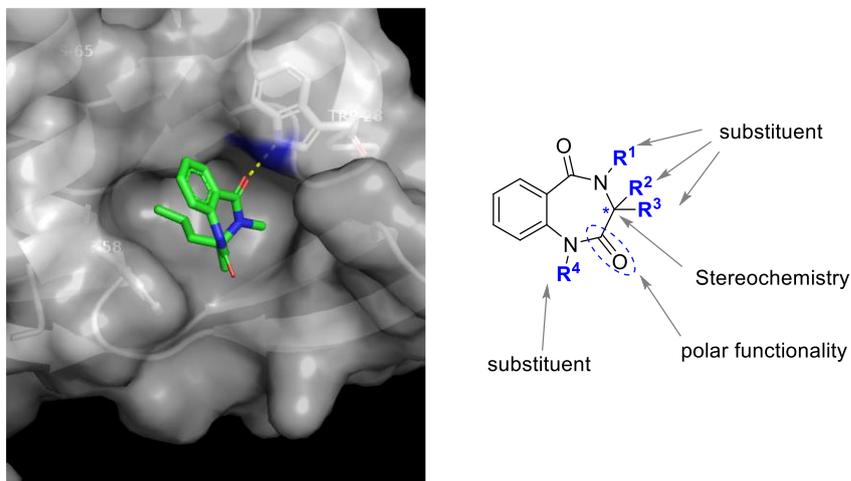


**Figure 3.20.** A) The fragment binding site of **NM466** within activin A B) An overview of the **NM466** binding site in relation to chain A (depicted in grey) and chain B (depicted in blue) of activin A C) The crystal structure of activin A (Chain A depicted in grey and Chain B depicted in blue) bound to the extracellular domains of the type II receptor (ActRIIb-ECD) (depicted in red and orange).<sup>287</sup>

Moreover, comparison of the fragment binding site (Figure 3.20, B) to the reported crystal structure of activin A bound to the extracellular domains of the type II receptor (ActRIIb-ECD)<sup>287</sup> (Figure 3.20, C) highlighted the **NM466** to be bound in close proximity to one such domain. It could therefore be hypothesised that conformational changes induced through binding of **NM466** could enable modulation of the activin A signalling cascade.

Analogous to the CFI<sub>25</sub> project, these results once more provided the opportunity to showcase the DOS chemistry, in addition to potentially providing novel starting points for lead generation. Importantly, during modelling and refinement of the crystal structures it was found that only (*R*)- variant of **NM466** was capable of fitting the electron density, suggesting the preferential or sole binding of this enantiomer. As a result, the first priority was the synthesis of both the (*R*)- and (*S*)- enantiomers to confirm this hypothesis utilising the modular chemistry. Secondly, once more devoid of biophysical hit validation methods due to the challenging nature of developing such techniques for novel targets, it was anticipated the synthesis of **NM466** analogues would address this objective.

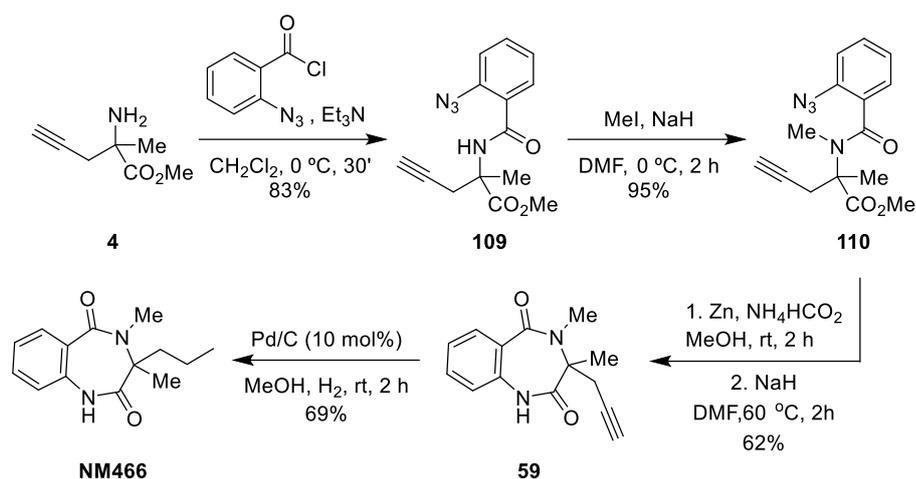
With the crystal structure and modular DOS route in mind, several analogues were envisioned for this purpose (Figure 3.21). More specifically, it was proposed that benzodiazepine core of the molecule provided the opportunity for several points of derivatisation, including simple *N*-alkylation, aromatic substitution and alkyne chain modifications. Thus, a synthetic strategy to enable the generation of such analogues was devised.



**Figure 3.21.** The crystal structure of **NM466** bound to *Activin A* and the proposed positions of derivatisation.

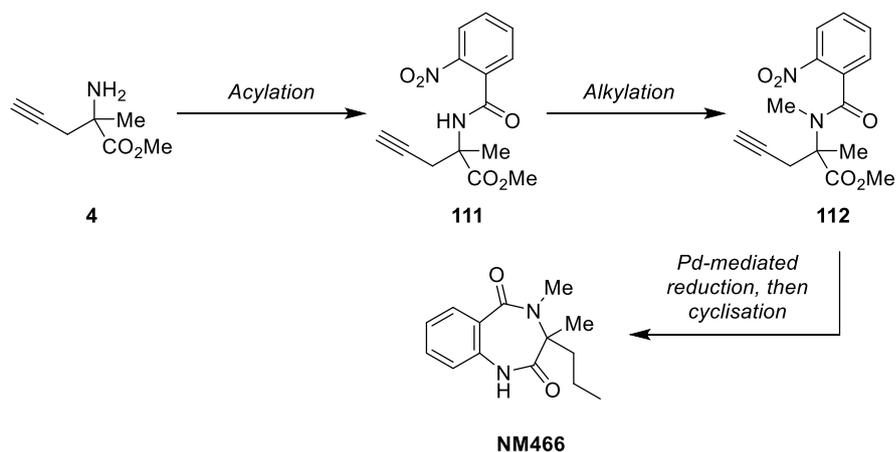
### 3.2.5 Follow-up compounds to **NM466**

The original synthetic route to **NM466** developed by Dr. Natalia Mateu is outlined in Scheme 3.10.<sup>232</sup> This straightforward four-step procedure involved acylation of **4** with 2-azidobenzoyl chloride to give **109**, followed by methylation to give **110**. Subsequent Zn-mediated azide reduction followed by subjection of the crude intermediate to basic conditions promoted the intramolecular cyclisation to deliver the benzodiazepine **59** in good yield. Finally, Pd-mediated reduction was then employed to generate the desired propyl chain from the alkyne moiety to afford **NM466**.



**Scheme 3.10.** The original route to **NM466** developed by Dr. Natalia Mateu starting from amino ester **4**.

Initially, during library construction this route had been designed with the vision of accessing additional scaffolds through derivatisation of the intact alkyne within **59**. However, with respect to the activin A project, this motif was not required. Alternatively, it was envisioned that direct analogues of **NM466** could instead be accessed *via* a nitro variant of **110**. Accordingly, a condensed three-step synthetic route from the key amino ester **4** was proposed (Scheme 3.11). It was anticipated that this alternative route could employ commercially available 2-nitrobenzoyl chloride derivatives in place of 2-azidobenzoyl functionalities. In turn, it was hoped that in this instance Pd-mediated reduction could be instead utilised to simultaneously affect the reduction of both alkyne and nitro moieties in a single step, prior to cyclisation.

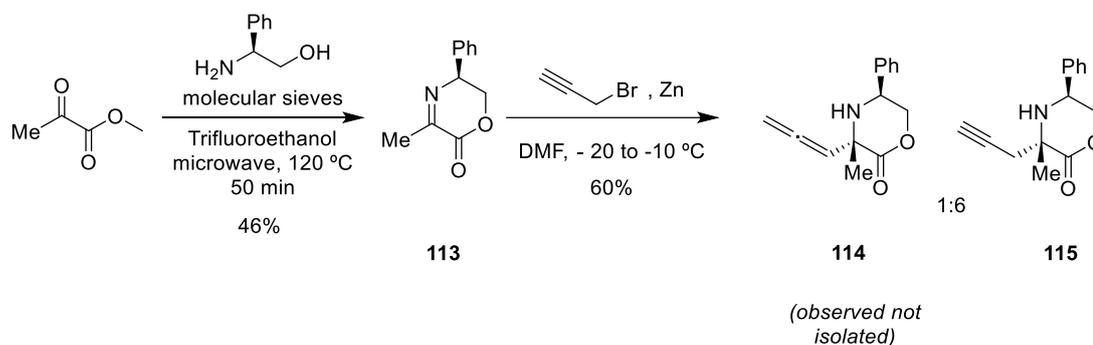


**Scheme 3.11.** The proposed shorter three-step route to access **NM466** analogues.

Following this procedure, firstly utilising the (*R*)- or (*S*)- variant of **4**, the corresponding single enantiomers of **NM466** could be constructed. In turn, to access derivatives it was envisioned that acylation could be used to introduce aromatic core variation using commercially available materials, alkylation could be used to explore amide substitution, whilst modifications to the benzodiazepine core could be introduced using late-stage functionalisation once the scaffold

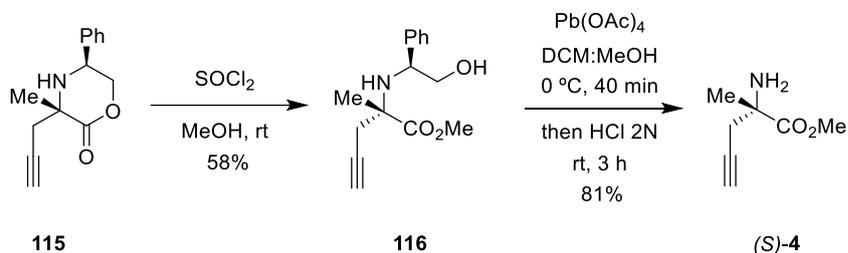
was constructed. Finally, modification of the amine **4** used in the first step could allow for alternative quaternary substituents to be introduced.

Accordingly, initial efforts sought to generate the (*S*)-amino ester **4** following the synthetic procedure established by Dr. Natalia Mateu (Scheme 3.12).<sup>232</sup> Synthesis began with the formation of the cyclic imine **114** *via* microwave-assisted condensation of (*S*)-phenyl glycinol and methyl pyruvate. In turn, this intermediate was subjected to zinc-mediated Barbier conditions to install the quaternary substituent, with the stereochemistry controlled *via* rear-face attack due to top-face steric hindrance induced by the phenyl substituent. In a similar fashion to the racemic synthesis of **4** (see Chapter 2), both the alkene and allene products were detected by crude <sup>1</sup>H NMR analysis. It was found that the temperature of the reaction heavily influenced the observed ratio of these two species, with reactions carried out at -20 °C offering the best ratio of 6:1 alkyne to allene (determined by <sup>1</sup>H NMR). Fortunately, the major isomer could be separated using column chromatography.



**Scheme 3.12.** Phenyl glycinol was employed as a cheap chiral auxiliary in the asymmetric route to assess optically pure (*S*)-**115**.

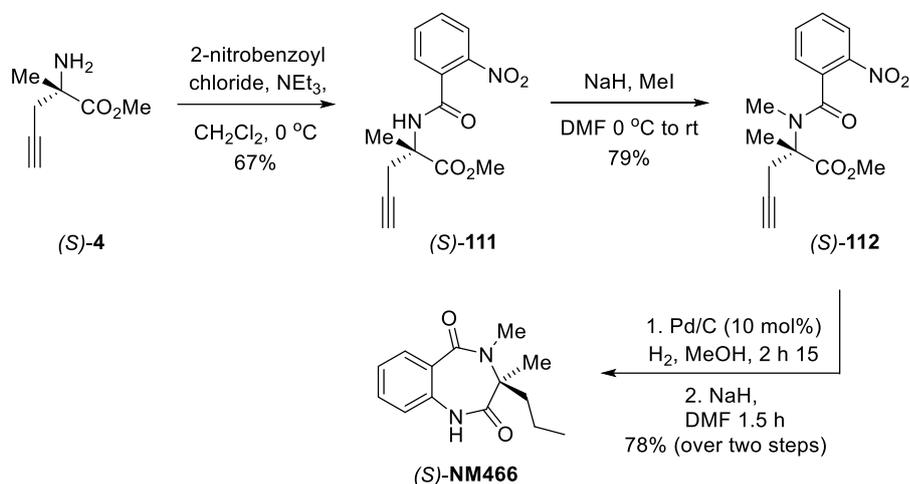
Following this, thionyl chloride-mediated ring opening of **115** was utilised to afford **116** (Scheme 3.13). Finally, auxiliary cleavage using lead (IV) acetate delivered the enantiopure (*S*)-**4** amino ester in a good 81% yield.



**Scheme 3.13.** Ring cleavage and auxiliary removal afforded the enantiopure amino ester (*S*)-**4**.

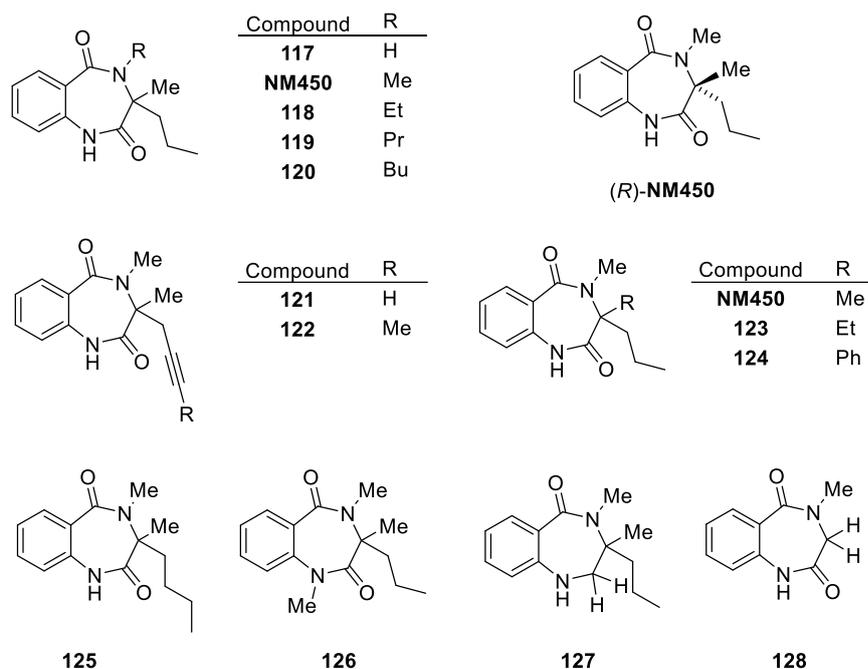
With (*S*)-**4** in hand, the shortened synthetic route to **NM466** was subsequently investigated. Accordingly, (*S*)-**4** was treated with 2-nitrobenzoyl chloride and triethylamine, yielding the acylated intermediate (*S*)-**111** (Scheme 3.14). Following this, methylation of this intermediate proceeded smoothly, affording (*S*)-**111** in a 79% yield. Finally, subjection of (*S*)-**112** to 10 mol%

of Pd/C under a hydrogen atmosphere promoted the simultaneous reduction of both nitro and alkyne functionalities. It was found that in an analogous fashion to the original Zn-mediated procedure, filtration of the reaction through Celite proved to be a sufficient work-up and upon treatment of the crude intermediate to sodium hydride in DMF, the desired product (**S**)-**NM466** was isolated in a good 78% yield (over two steps).



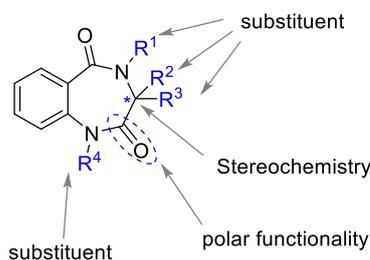
**Scheme 3.14.** Using (**S**)-**4** the corresponding (**S**)-**NM466** was rapidly constructed.

In light of these findings, the synthesis of the remaining (*R*)-enantiomer and further analogues of **NM466** formed the basis of an M.Sci project undertaken by Till Reinhardt within the Spring group. This project proved to be particularly successful leading to the synthesis of an additional 15 analogues for screening (Figure 3.22). This included sequential growing of the fragment from the alkylated amide position (**117** - **120**), variation in the alkyne and quaternary substituents (**121** - **124**), in addition to modification of the secondary amide (**126** & **127**) and complete removal of the quaternary centre (**128**). This was achieved using computational modelling to allow compound prioritisation.



**Figure 3.22.** The 15 analogues successfully synthesised by Till Reinhardt to validate and probe the binding of **NM466** to Activin A.

Thus, once more, the utility of the NSQC library and DOS chemistry in a fragment screening setting has been demonstrated. Additionally, through these efforts we also exemplified the flexibility of the DOS strategy through the generation of a condensed three-step synthetic route to access **NM466**, enabling the rapid generation of derivatives. By exploiting the modularity of this methodology, both the (*R*)- and (*S*)- enantiomers could be easily accessed in efforts to validate our original findings relating to the preferential binding of (*R*)-**NM466**. Moreover, in line with the crystal data, several vectors of the molecule have been explored (Figure 3.23). Importantly, this included the generation of a third ethyl quaternary substituent, highlighting the potential of this motif to be utilised as a 3-D fragment growth vector.



**Figure 3.23.** An overview of the vectors from **NM466** explored during these efforts.

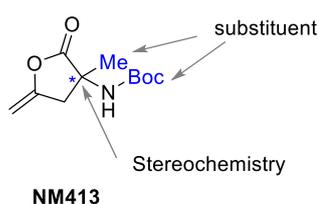
### 3.2.6 Hit validation

Preliminary X-Ray screening data conducted by Joseph McLoughlin of **117**, **118** and **124** successfully validated the initial binding of **NM466** through the additional binding of **117** and **115**. The screening of the remaining 12 analogues is yet to be completed. Thus, these **NM466**

therefore represents a novel starting point for the development of new chemical leads. On-going work in this regard is focused on the screening of the remaining analogues in order to confirm the viable vectors for fragment growth. Moreover, to guide the hit evolution, the development of NMR-based and biophysical methods to determine binding affinity is also underway within the Hyvönen group. It is envisioned that with these techniques in hand, further analogues could be rapidly accessed to develop the binding affinity and/or potency as required.

### 3.2.7 On-going screening collaborations

In addition to the examples described herein, the NSQC library has subsequently been utilised in other XChem screenings. Once more, this library has proven fruitful in this context and provided novel fragment leads. This includes the hit **NM413** against penicillin binding protein 3 (PBP-3), in collaboration with the Dowson group at The University of Warwick. On-going follow-up chemistry efforts in this regard are under investigation by Elaine Fowler and will be published in due course. It is worth mentioning, that in a similar fashion to the above examples in this context the DOS strategy has facilitated exploration of several fragment growth vectors (Figure 3.24), including modification to the quaternary substituent.



**Figure 3.24.** On-going work conducted by Elaine Fowler exploring the 3-D growth of **NM413**.

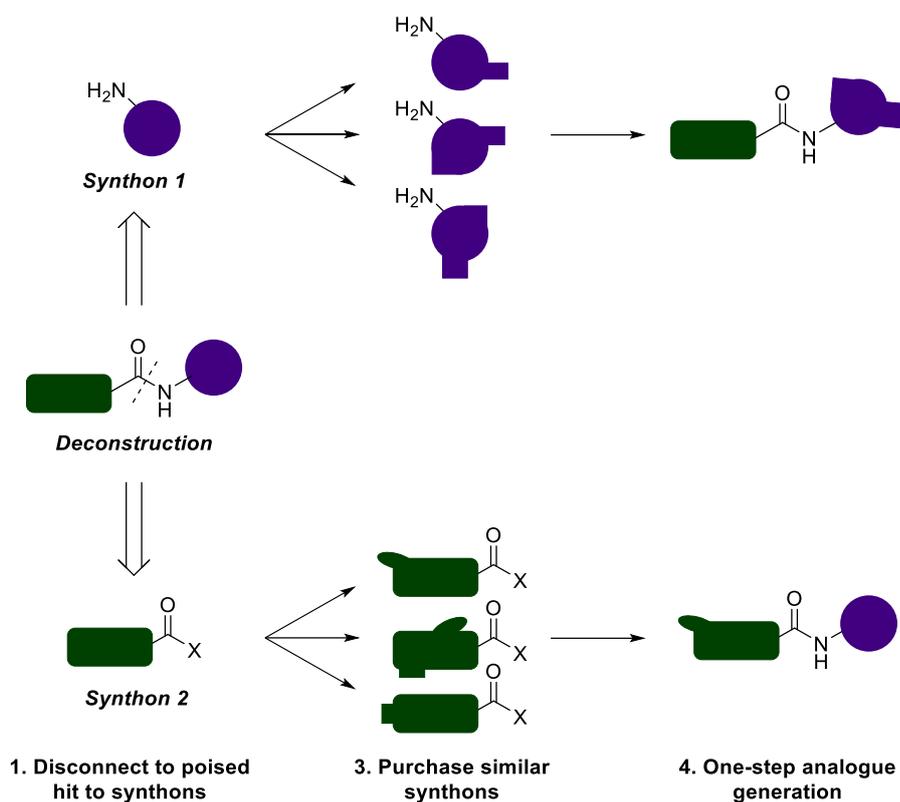
As these examples demonstrate, the NSQC library has proven effective in delivering novel leads against challenging biological targets, owing to the broad range of scaffold diversity. Additionally, as an inherent feature of its synthesis access to derivatives can be achieved in a straightforward fashion without the requirement of novel synthetic route development. Moreover, as the latter two cases exemplify, the NSQC-motif can provide an additional 3-D exit vector, with both alkyl and aryl substituents described. Thus, overall this work demonstrates that this library serves as an excellent complementary collection to commercial fragments, and considering the success described, it is envisioned that these molecules will prove fruitful in many future FBS campaigns. Future work will focus on building upon these preliminary results with the vision of obtaining potent compounds against all three targets.

## 4 Results and discussion: Investigations into high-throughput methods of analogue production

### 4.1 Background

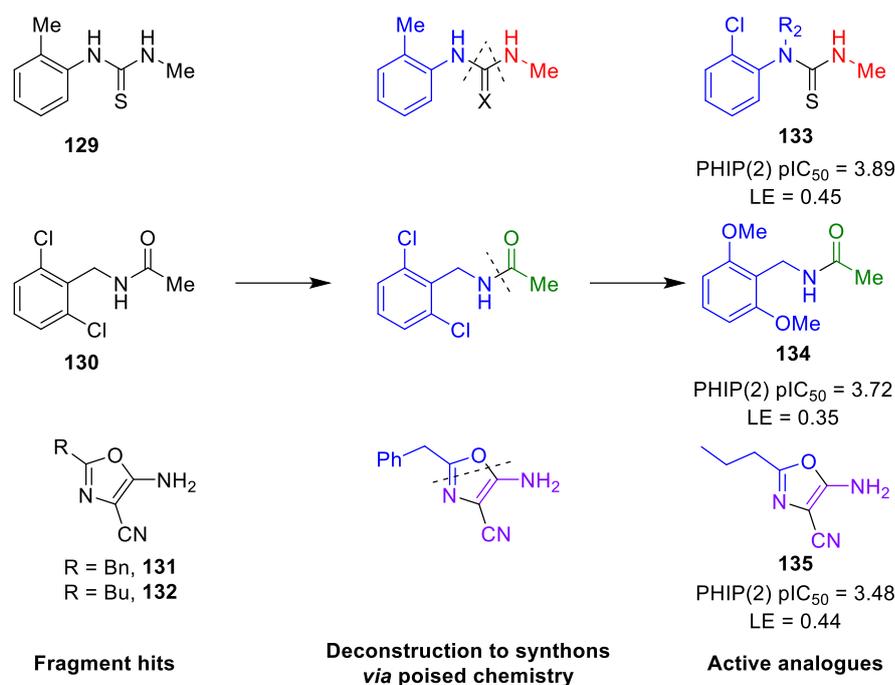
One significant challenge within the FBS screening paradigm lies with the subsequent hit validation and evolution process to attain compounds with measurable potency in solution assays. Indeed, due to the low affinity ( $\mu\text{M}$  to  $\text{mM}$ ) and sometimes ambiguous nature of initial hits resulting from these efforts, this process is dramatically complicated.<sup>110,115</sup> Often, the first stage of this process can be achieved in an 'off-the-shelf' approach using commercially available fragment analogues (termed SAR-by Catalogue).<sup>288,289</sup> However, in some cases this method can be more challenging, for example when screening novel compound libraries. Thus, the availability of appropriate compounds for this purpose in a time- and cost- efficient manner is crucial.

In view of these challenges, the Diamond and SGC Poised Fragment Library (DSPL) was developed by Cox *et al.*<sup>248</sup> with the aim of enabling the rapid synthesis of compound analogues to facilitate hit validation and optimisation. This entailed the generation of a 'poised' strategy relying on the utility of 23 robust and reliable reactions commonly adopted within medicinal chemistry<sup>290</sup> and used to form heterocycles<sup>291</sup> to enable reliable disconnection of fragments into readily available synthons (Figure 4.1). Accordingly, upon the identification of fragment hits, the corresponding synthons could be readily enumerated using commercial collections and analogues readily synthesised in one synthetic step. Using these principles, a screening library of 779 compounds with these properties was constructed, termed the *DSPL*.



**Figure 4.1.** Generation of analogues using the DSPL library. Adapted from reference.<sup>248</sup>

In a proof-of-concept study, the authors demonstrated the utility of this poised strategy and the *DPSL* using a case study screening campaign against PHIP(2), an atypical bromodomain.<sup>248</sup> Following an initial X-ray screen four fragment hits (**129** – **132**) within the key acetyl-lysine binding region were identified, corresponding to three different ‘poised’ chemotypes of thioureas, amides and oxazoles (Figure 4.2). Through exploitation of the poised chemistry handles, the subsequent construction of 64 analogues could be achieved in a relatively short time period, which ultimately resulted in the identification of novel micromolar inhibitors. This included the highly ligand efficient compounds **133**, **134** and **135**.



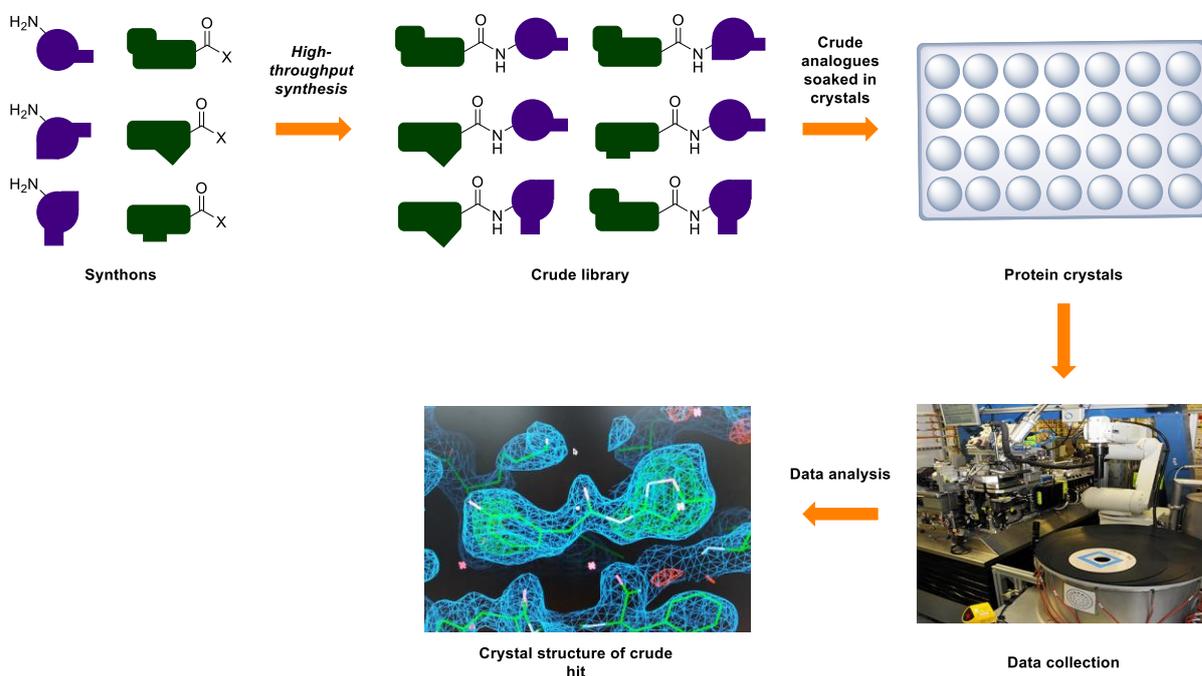
**Figure 4.2.** The poised approach case study against PHIP(2). Three hit chemotypes were rapidly enumerated leading to micromolar affinity analogues detected by an AlphaScreen assay.

Within the XChem setting (see Chapter 3) the *DPSL* has proven particularly successful, enabling the identification of several initial hits and subsequent analogue syntheses across several therapeutic targets.<sup>248,257</sup> Building on these promising results, the second generation library, Diamond, SGC and iNext Poised (*DSi-P*) library was developed and is now routinely proposed to XChem users. To aid library uptake, the *DSi-P* library was recently made available from the chemical supplier *Enamine*.<sup>292</sup> However, a challenge in the application of the library remained in instances where researchers sought to access compounds outside the initial library scope. As a result, the XChem team sought to enhance this collection with additional 3-D libraries, including those containing restricted elements such as NSQCs, that allow rapid multi-directional structure-activity relationship exploration. In this regard, the NSQC library (see Chapter 2 and 3) seemed an ideal library to compliment the *DPSL* due to its excellent physicochemical properties and successful applications toward challenging targets.

One challenge arising from the application of both libraries concerned the availability of bespoke analogues. It is generally recognised that within the medicinal chemistry paradigm organic synthesis remains a significant rate-limiting step due to the large significant cost and time overheads associated with this phase of a project.<sup>154</sup> In this regard, machine-assisted synthesis has demonstrated significant merit within recent years, and calls toward the demonstration of reaction compatibility with automation technologies have been made.<sup>154</sup> Thus, within recent years the application of automatable technologies towards reaction discovery,<sup>293–295</sup> reaction optimisation<sup>296,297</sup> and the miniaturised synthesis of high-value medicinal compounds<sup>298–300</sup> has transformed the landscape of synthetic chemistry. Certainly,

the ability of such platforms to accelerate the timescale and increase the efficiency of such procedures cannot be rivalled, importantly providing one solution to overcoming the aforementioned barriers within drug discovery. More recently, literature examples have also emerged demonstrating the powerful integration of high-throughput methods of synthesis with bioactive screens for SAR exploration<sup>301,302</sup> enabling a direct bioactivity readout for the resulting compounds.

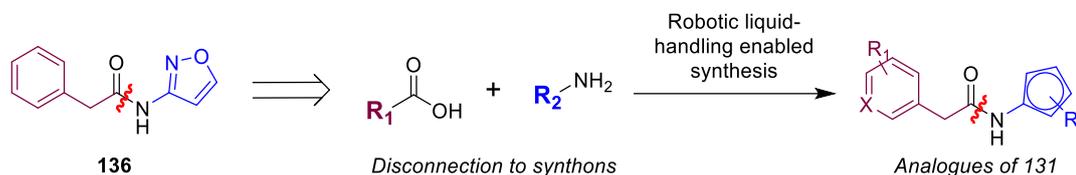
Inspired by these reports, unpublished work undertaken by researchers at the XChem facility sought to investigate the possibility of merging the novel XChem screening platform with high-throughput strategies of analogue synthesis (Figure 4.3). More specifically, it was hoped that in this instance the bottleneck of reaction purification<sup>303</sup> could be mitigated by instead utilising the exquisite nature of protein binding specificity, observed for example in crystal soaking methods during fragment screening, to exclusively bind to the reaction component of choice. In this case, it was envisioned that crude reaction mixtures could be directly employed in crystal soaking events to identify follow-up analogue hits, letting the protein crystal filter out reagents and side-products. Moreover, this would serve as an orthogonal method to existing technologies such as activity-directed synthesis,<sup>304,305</sup> nanoscale synthesis and affinity ranking<sup>302</sup> and crude SPR,<sup>306,307</sup> enabling the simultaneous rapid acquirement of structural binding information of the resultant compounds.



**Figure 4.3.** An overview of the high-throughput crude workflow.

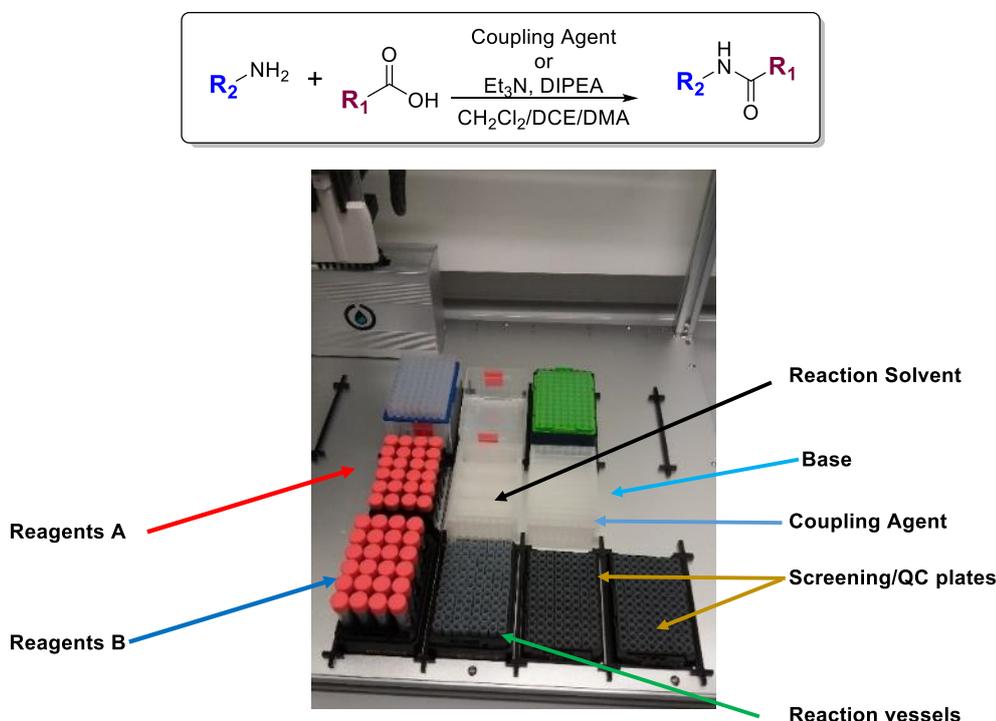
To initially investigate this proposal, it was envisioned that by taking advantage of the inherent poised nature of the *DSPL*, liquid handling robotics could be utilised to rapidly generate analogues in a high-throughput manner. To achieve this, the Opentrons robot was selected due to the ability to generate customised protocols using Python and economic benefits

including low maintenance, outlay and consumable costs.<sup>308</sup> As a proof-of-concept study, investigations carried out by Dr. Anthony Aimon sought to re-synthesise a series of previously confirmed analogues of **136** via amide coupling reactions utilising the Opentrons in a 96-well format (Figure 4.4).<sup>309</sup>



**Figure 4.4.** Disconnection of **136** via the poised synthetic handle to give the corresponding synthons for variation.

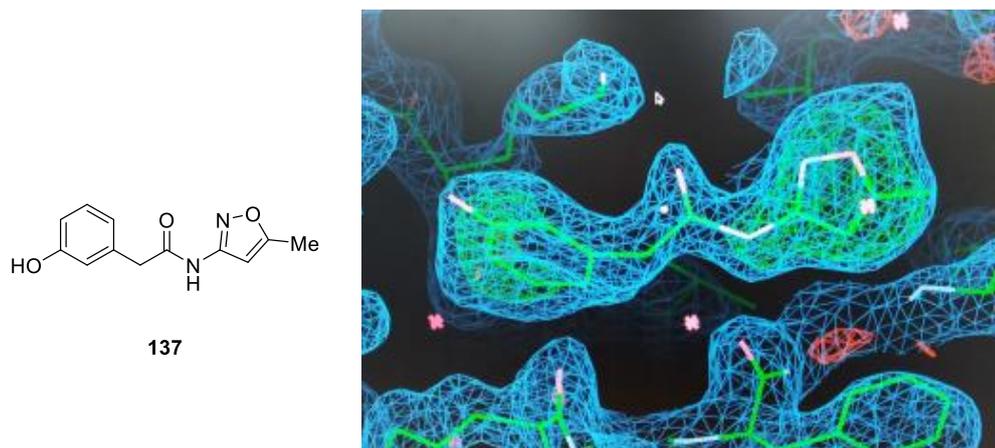
Utilising stock solutions of amines (Reagents A) and either carboxylic acids or acids chlorides (Reagents B), in addition to troughs of base, coupling agent and solvent, the replication of 176 compounds was successfully achieved (Figure 4.5). It was found that, if required, the resulting reaction mixture could be subjected to a surrogate aqueous work up via sequential pipette injections and aspirations, allowing for the removal of aqueous soluble reaction by-products. Additionally, liquid transfers of the crude mixtures to screening and quality-control plates in bespoke format (96/384/1536 wells) or solvent (Acetonitrile, DCM, DMSO) were facilitated, although this required utilisation of an evaporation system, e.g Genevac.



**Figure 4.5.** The proof-of-concept amide formation reaction to generate a small library of analogues and robotic set-up. Work carried out by Dr. Anthony Aimon, XChem.

In turn, reaction progression was assessed by means of LCMS analysis and upon completion the crude reactions were subsequently directly employed in crystal soaking experiments

against the target of interest following the XChem workflow (see Chapter 3). Gratifyingly, upon screening of the crude reaction mixtures identical hit compounds were identified from both the purified and crude compounds (for example **137**) confirming the validity of this process (Figure 4.6). To provide open-access, the initial Python code for this procedure was subsequently generalised and a repository created by Dr. Anthony Bradley on the open-source platform GitHub.<sup>310</sup>



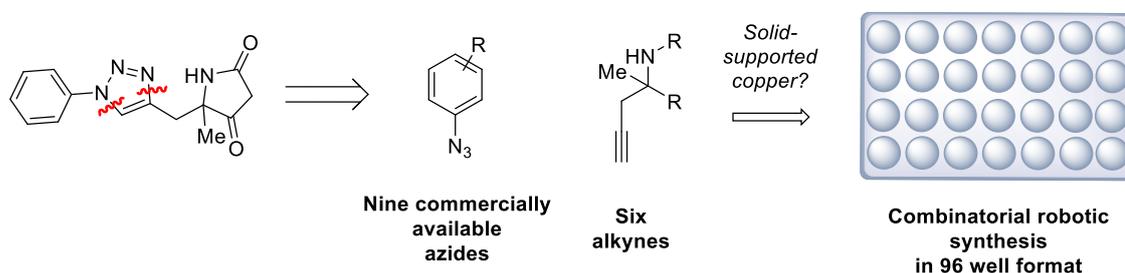
**Figure 4.6.** Compound **137** and the electron density map of the crude compound within the pocket of the desired protein. Work carried out by Dr. Anthony Aimon, XChem.

## 4.2 DOS library application and toolbox expansion

Whilst the application of the DPSL within the high-throughput crude platform had proven successful, two tasks remained outstanding: (1) demonstration of the compatibility of the technology with more 3-D fragment libraries and (2) expansion of the crude toolbox to show compatibility with alternative reaction conditions. In line with these aims, we hypothesised that the inherently modular nature of the quaternary DOS strategy could be exploited to this address these objectives. To achieve this, it was envisioned that a library of second-generation of **NM450** follow-up compounds (see Chapter Three) could be rapidly synthesised utilising this technology. It was hoped this would showcase the potential of the DOS library, whilst enabling exploration into the application of a second reaction methodology.

In view of the presence of the key bridging triazole within six DOS library members, including the fragment hit **NM450**, we hypothesised that the generality and robust nature of triazole-forming click chemistry could be harnessed for this purpose. The classical Cu(I)-catalysed or strain-promoted Huisgen 1,3-dipolar cycloaddition between an alkyne and azide to generate 1,2,3-triazoles constitutes an important click chemistry variant owing to its characteristic features of high selectivity, reliability and bio-compatibility.<sup>311,312</sup> Indeed, the utility of this reaction within the drug discovery paradigm has been previously demonstrated by applications in bioconjugation including peptide stapling,<sup>313–315</sup> DNA modification<sup>316,317</sup> and cell surface labelling.<sup>318</sup> In view of the merits of this reaction,<sup>318</sup> utilising the **NM450** hit in a proof-of-concept

study, it was proposed that following a combinatorial approach, libraries of analogues exhibiting variation in both the aromatic substitution and quaternary heterocycle could be generated in an automated fashion (Figure 4.7).

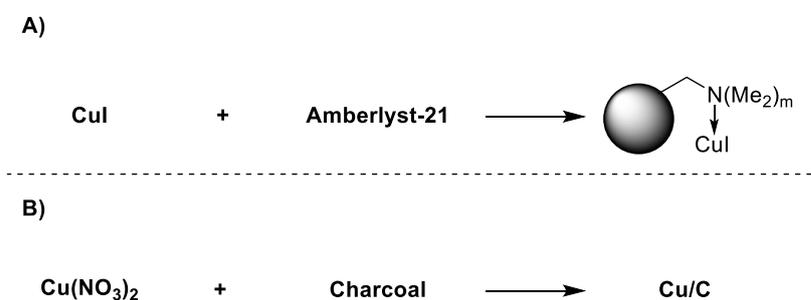


**Figure 4.7.** The combinatorial-type approach proposed for use in the crude high-throughput synthesis.

#### 4.2.1 High-throughput synthesis

To investigate this proposal, initially six quaternary alkyne substrates were chosen to be combinatorially reacted with nine commercially available aromatic azide substrates to generate 54 corresponding analogues of **NM450**. However, a major concern was the imperative use of copper agents in the reactions, which we hypothesised may result in a reduction in the diffraction quality and resolution of the crystals.<sup>319</sup> To minimize this, it was decided to investigate the use of solid-supported copper sources, in addition using the conventional reaction conditions.

The first approach, developed by Girard *et al.*,<sup>320</sup> entailed the non-covalent attachment of CuI to an Amberlyst-21 polymer resin *via* chelation to the amine-functionality within the polymer (Figure 4.8). Importantly, due to the use of CuI within this procedure, the active Cu(I) species is generated within the reaction mixture, therefore removing the requirement of reducing agents to promote the reaction. Furthermore, the reported reaction scope of this procedure appeared sufficiently broad and thus seemed a promising option.

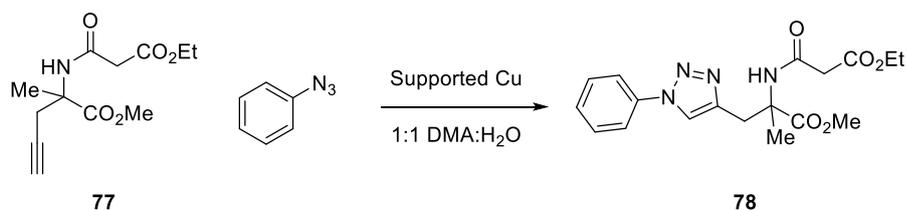


**Figure 4.8.** A) Attachment of CuI to polymer resin *via* chelation.<sup>320</sup> B) The formation of nanoparticle Cu/C reported by Lipshutz *et al.*<sup>321</sup>

The second strategy, developed by Lipshutz *et al.*,<sup>321</sup> involved impregnation of activated wood charcoal particles with copper nitrate to afford nanoparticles of Cu/C (Figure 4.9). In a similar

fashion, the requirement for a reducing agent was also eliminated due to the presence of both CuO and Cu<sub>2</sub>O within the matrix. Moreover, the reaction was shown to proceed with or without the addition of base, although the addition of one equivalent of NEt<sub>3</sub> was shown to be beneficial. If successful, it was envisaged that the removal of such nanoparticles could be mediated *via* the incorporation of a reaction mixture filtration step.

Consequently, both catalysts were synthesised following the reported procedures and an investigation into their ability to transform **77** to **78** using azido benzene was initiated. When **77** was treated with Cu/C and azidobenzene, disappointingly no evidence of **78** was observed with only starting material present (Table 4.1, Entry 1). Initial subjection of **77** to the Amberlyst21-Cu system and azidobenzene at room temperature proved sluggish, however, upon heating of the reaction to 40 °C full conversion was observed and a good 75% yield of **78** was obtained (Table 4.1, Entry 2). Finally, for a better compatibility with the XChem screening platform, the traditional solution phase CuSO<sub>4</sub> procedure was modified by switching from *t*-butanol to DMA successfully (Table 4.1, Entry 3).

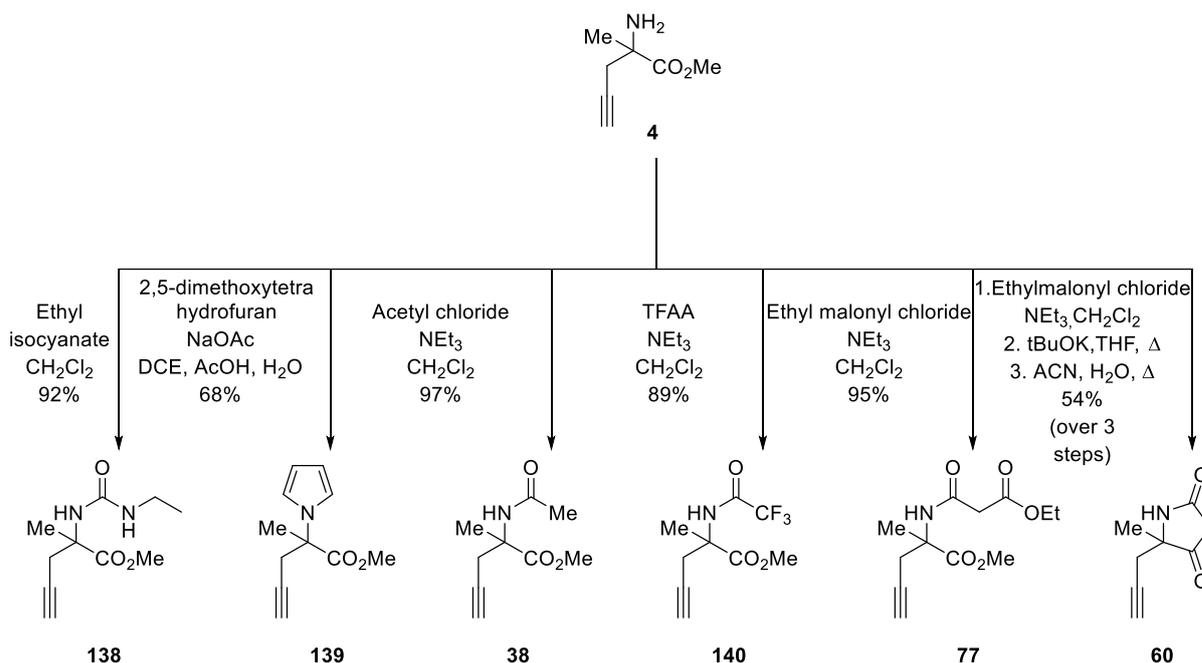


Entry	Copper source	Conditions	Result
1	Cu/C	DMA:H <sub>2</sub> O rt then 40 °C 18 hr	Only SM by TLC
2	Amberlyst21-Cu	DMA:H <sub>2</sub> O rt then 40 °C 18 hr	Full conversion, 75%
3	CuSO <sub>4</sub>	Sodium ascorbate (0.3 equiv) DMA:H <sub>2</sub> O	Full conversion, 58%

**Table 4.1.** Investigations of reaction conditions for automated synthesis.

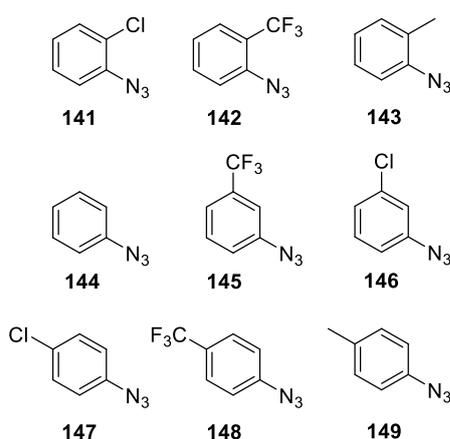
Considering the identification of two alternative binding sites for **NM450** and **NM461** within CFI<sub>25</sub> (see Chapter 3), it was envisaged that variation of the quaternary heterocycle would provide valuable information into the binding preference of such molecules. With this theory in mind, four biologically relevant motifs that could be obtained in one-step transformations from **4** were selected for synthesis (Scheme 4.1). In addition, alkynes **77** and **60** were also included to serve as controls, delivering the corresponding triazole **NM450** and **78** previously synthesised. Hence, **4** was treated with ethyl isocyanate yielding urea **138**, whilst the corresponding pyrrole was achieved *via* application of Clauson-Kass conditions,<sup>322,323</sup> proceeding smoothly to yield **139**. Next, acylation of **4** with acetyl chloride or trifluoroacetic anhydride afforded the *N*-acetamide variants **140** and **141**. Finally, acylation with ethyl malonyl

chloride afforded **77** and subsequently, half of this material could then be cyclised *via* a Dieckmann condensation and decarboxylation to deliver the pyrrolidinone **60**.



**Scheme 4.1.** Six alkynes were readily synthesised *via* N-modifications from amine **4**.

With the alkynes in hand, nine commercially available azides (**142 – 150**) were selected matching those used in the conventional synthesis (Figure 4.9, see Chapter 3) and varying in substitution and functionality around the aromatic core. It should be noted that the screening of the analogues described in Chapter 3 was conducted concurrently to these investigations and as such the SBR data obtained from the follow-up library screening could not be utilised to inform azide selection at this stage.

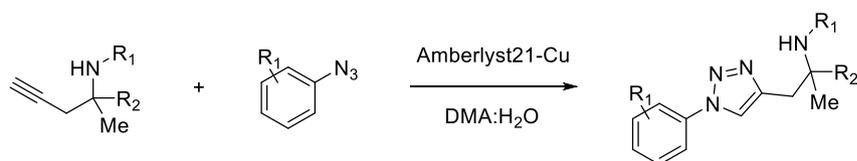


**Figure 4.9.** Commercially available azides used in this initial study.

With these substrates in hand, investigations into the utility of this methodology within the high-throughput synthesis platform were initiated. The reactions were carried out on a 0.05 mmol scale using 0.8M stock solutions of each alkyne in DMA. Due to the instability of the azide

functionality within the aromatic substrates, these materials were commercially provided in 0.5M solutions in *tert*-butyl methyl ether. As a result, the reactions were conducted in DMA, which was later removed *via* Genevac evaporation and subsequently dissolved in DMSO to enable X-ray screening. In an analogous fashion to the conventional triazole synthesis (*vide supra*) a 1:0.9 ratio of alkyne to azide was once more utilised to avoid the isolation of unreacted azide substrate. Agitation and gentle heating to 37 °C was achieved through employment of an orbital shaker incubator. After 24 hours the reaction plates were cooled to room temperature, the solvents were exchanged, and the reaction progression was analysed by LCMS.

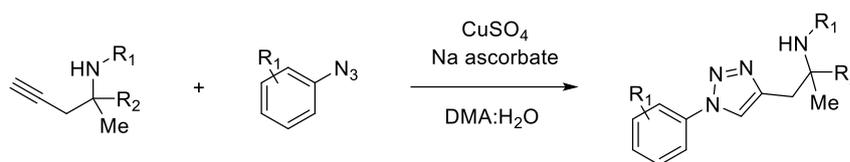
Firstly, considering the Amberlyst21-Cu system, 33 of the 45 reactions that were successfully analysed showed greater than 70% conversion (assessed through ELSD chromatograms) (Table 4.2). For nine reaction wells instrumental measurement errors resulted in poor determination of the reaction conversion, however pleasingly the desired product mass was still observed on three occasions. Interestingly, the same trend of reduced reactivity of the 1-azido-2-methylbenzene alkyne (see Chapter 3) was observed with poor conversion to the desired products in all but one case.



Azide	Alkyne					
	77	60	138	139	38	140
141			Mass found			
142				No signal		
143						
144						
145						
146						
147						Mass not found
148			Mass found		Mass found	
149	Mass not found	Mass not found			Mass not found	Mass not found

**Table 4.2.** Reactions conducted under Amberlyst21-Cu catalyst conditions: Approx. 4 mg Amberlyst21-Cu, DMA:H<sub>2</sub>O. Dark green = excellent conversion (100%), Pale green = good conversion (70 - 90%), Yellow = moderate conversion (50 - 60%), Orange = low conversion (20 - 40%), Red = poor conversion (< 20%). Grey = detection error no conversion could be determined.

The corresponding reactions employing the copper in solution protocol were analysed (Table 4.3). It should be noted that in this case, due to insufficient quantities of alkyne **140**, only the reactions of alkynes **38**, **60**, **77**, **138** and **139** were investigated. Gratifyingly, LCMS analysis revealed a similar trend with the majority of the reactions (30 out of 39 successfully analysed wells) displaying greater than 70% conversion. Once more instrumental error was observed in six cases whereby the conversion could not be assessed, however in all cases the desired product masses were still in fact detected.



Azide	Alkyne				
	77	60	138	139	38
141	Pale green	Yellow	Pale green	Pale green	Pale green
142	Dark green	Pale green	Pale green	Pale green	Dark green
143	Red	Red	Red	Red	Red
144	Pale green	Orange	Pale green	Pale green	Pale green
145	Dark green	Dark green	Dark green	Pale green	Dark green
146	Pale green	Pale green	Dark green	Pale green	Dark green
147	Pale green	Mass found	Pale green	Mass found	Pale green
148	Pale green	Mass found	Dark green	Yellow	Mass found
149	Orange	Mass found	Pale green	Pale green	Mass found

**Table 4.3.** Reactions conducted under standard conditions: 10 mol% CuSO<sub>4</sub>·5H<sub>2</sub>O, 30 mol% Sodium ascorbate, DMA:H<sub>2</sub>O. Dark green = excellent conversion (100%), Pale green = good conversion (70 - 90%), Yellow = moderate conversion (50 - 60%), Orange = low conversion (20 - 40%), Red = poor conversion (< 20%). Grey = detection error no conversion could be determined.

These results therefore validate the first objective of these investigations, demonstrating the applicability of the quaternary DOS library to a high-throughput synthesis setting. Firstly, this enabled the rapid formation of 78 analogues within one week, but secondly exemplified the compatibility of the library with solid-phase catalysis, potentially unlocking further applications of the screening library in a high-throughput context. Significantly, the modular nature of the synthetic strategy used to construct the library was leveraged to generate several alkyne synthons. Whilst only the nitrogen vector was explored in this instance, it is envisaged that other functionalities could be additionally enumerated (following the approaches described in Chapter 2) to generate more varied synthons. Thus, it could be hypothesised that multi-

directional growth vectors of **NM450** could be rapidly explored in high-throughput fashion using this approach.

#### 4.2.2 Crude screening results

To investigate the second objective of crude toolbox expansion, the 96 crude mixtures were then screened against CFI<sub>25</sub> using the XChem platform in duplicate. However, in all cases, upon analysis of the diffraction data using the PanDDA algorithm revealed no electron density relating to bound molecules was identified. Alarming, even the crude compound corresponding to the original hit **NM450**, which served as a control for this procedure was not detected by this analysis.

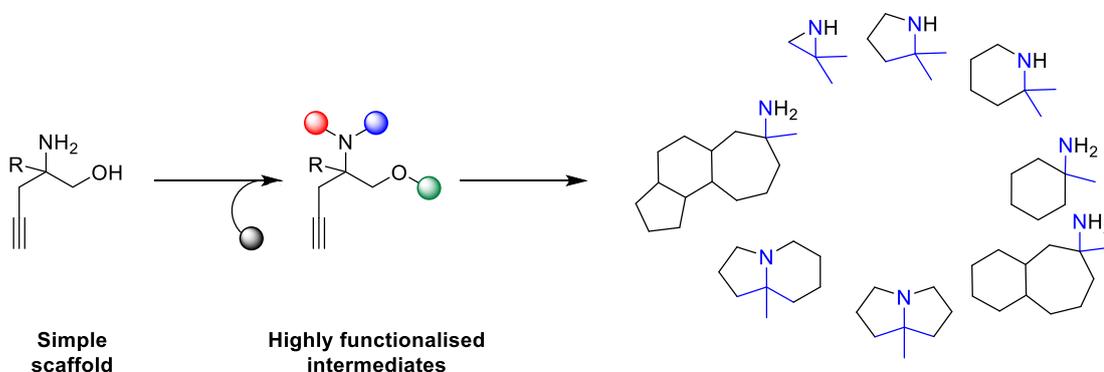
Significantly, a large proportion of the crystals exhibited poor to no diffraction, conflicting the characteristic robust nature of this crystal system previously observed during the initial CFI<sub>25</sub> screening campaign. An average of 41% of the crystals subjected to solid-supported crudes, and 47% of the crystals subjected to the solution-phase crudes produced data which was inadequate for incorporation into the PanDDA algorithm. This suggested upstream processing issues as a result of the poor crystal diffraction and resolution as a result of copper exposure. With respect to the Amberlyst-21 system it was proposed that metal-leaching could have occurred,<sup>324</sup> resulting in a higher concentration of copper within the crude reaction mixture, therefore explaining the similarity in results for both catalyst systems. With these points in mind, it is envisioned that future work on this project will involve the employment of metal scavenging reagents with the aim of reducing the in-solution copper levels. The application of these materials e.g. Quadrapure has previously been demonstrated, enabling the reduction to single digit parts per million.<sup>325–328</sup> Overall, as these investigations have demonstrated, whilst the quaternary DOS library is amenable to a high-throughput synthesis platform, further reaction optimisation is required to enable triazole formation to be compatible with the crude screening protocol.



## 5 Conclusions

An outstanding requirement within the field of organic synthesis is the development of novel strategies to access new SMs to seed both drug discovery and the development of new chemical probes. In a similar fashion to HTS approaches of identifying bioactive SMs, within the FBS paradigm organic synthesis has been identified as a limiting factor. Thus, in recent years calls for the development of new fragment libraries that feature increased diversity, 3-dimensionality and multiple modifiable vectors have been made. Moreover, due to the challenges associated with detecting particularly weak fragment binders, the development of novel strategies to accelerate analogue formation could provide the opportunity to lower the economic impact of this process.

DOS is a strategy that aims to generate varied molecular architectures in a modular and rapid fashion and is therefore well-suited to populate new areas of fragment space. Accordingly, in Chapter one, we described the design and synthesis of a 3-D fragment collection to address the aforementioned issues in fragment diversity by utilising this methodology. More specifically, this included the incorporation of a *N*-substituted quaternary stereocenter motif as a strategy for introducing 3-D character and conformational restriction throughout the library. This was achieved by harnessing the differential reactivity nature of  $\alpha,\alpha$ -disubstituted amino ester and alcohol substrates to generate quaternary-carbon containing SMs (Scheme 5.1). Initially, in the *build phase*, we developed robust and scalable synthetic routes to three building blocks with varied *N*-substitutions to enable the DOS investigations. Following this, systematic explorations into the *couple* and *pair* phases were conducted, resulting in the synthesis of 12  $sp^3$ -rich heterocyclic scaffolds based on eight molecular frameworks. This included the formation of three-, five-, six- and seven-membered constrained cycles, in addition to a series of novel oxepane-fused motifs. Furthermore, modifiable substituents were installed at a range of positions to enable fragment elaboration in a straightforward manner.

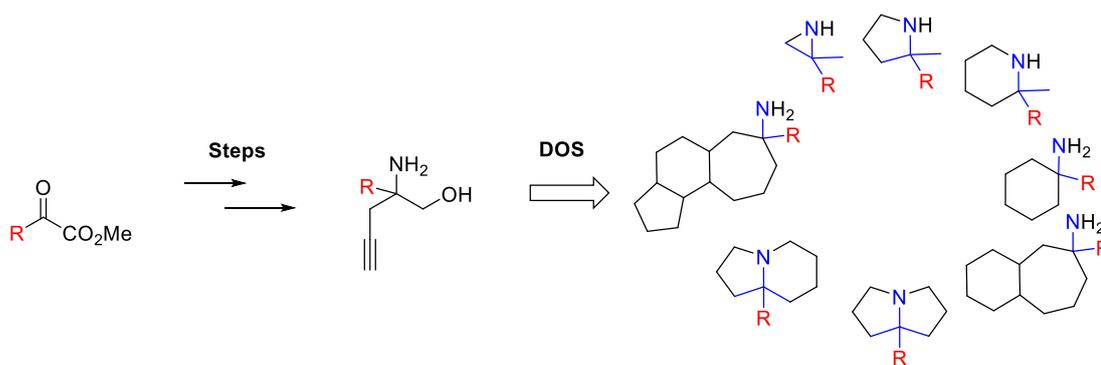


**Scheme 5.1.** The overall DOS strategy using an  $\alpha,\alpha$ -disubstituted amino alcohol building block.

Ultimately, these compounds were incorporated into a quaternary screening collection based on 40 distinct scaffolds incorporating medicinally-relevant motifs. The efficiency of these combined efforts was exemplified since no more than five synthetic steps were employed to generate each library member. Moreover, as an inherent feature of the building block motif, the quaternary centre was installed throughout the library, introducing a 3-D exit vector within the final molecules. The synthetic utility of this quaternary motif was demonstrated through the generation of the phenyl variant of the amino alcohol building block and subsequent incorporation of this motif into a final scaffold within the library.

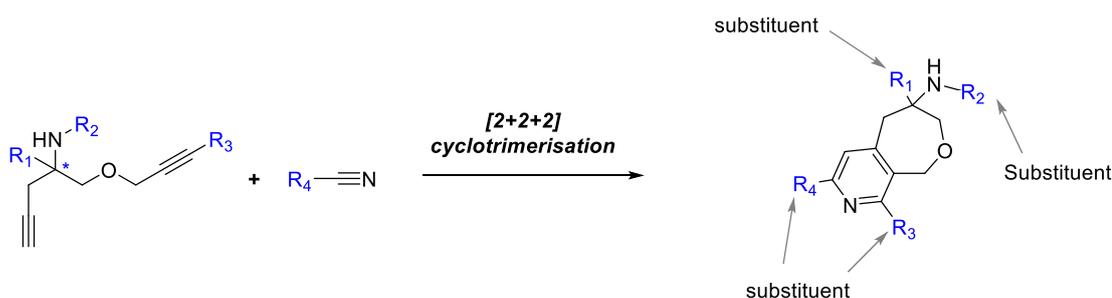
Subsequent computational analysis of this collection revealed the library accessed a broad range of molecular shape space, whilst the quaternary motif could be utilised to modify this distribution, illustrated *via* PMI plots. Thus, overall, we have shown that DOS can be used to synthesise a library of structurally diverse fragments featuring restricted elements to enhance their overall 3-D shape. Moreover, prediction of the physicochemical properties of the library demonstrated the amenability of this library to a FBS setting through RO3 guideline adherence. Finally, comparison of the DOS library to existing commercial collections revealed a lower fraction of aromatic atoms and an increased number of chiral centres was achieved in this work.

In Chapter one, we sought to illustrate the broad skeletal and stereochemical diversity that could be achieved using the DOS approach however, the functional group and appendage diversity elements remain to be extensively explored. It is envisioned that computational virtual enumeration could be utilised to further expand the scope of the work presented. Firstly, this could be addressed through expansion of the quaternary motif through the incorporation different substituents at this position. As demonstrated through the synthesis of the phenyl derivative, this motif can simply be modified through variation of the commercially available keto ester utilised within building block construction. Thus, following this methodology appendage diversity could be introduced among the scaffolds presented (Scheme 5.2). Moreover, whilst the library was synthesised in a racemic fashion, owing to the demonstration of the asymmetric synthetic route, the corresponding enantiopure (*R*)- and (*S*)- variants of each compound could additionally be virtually generated.



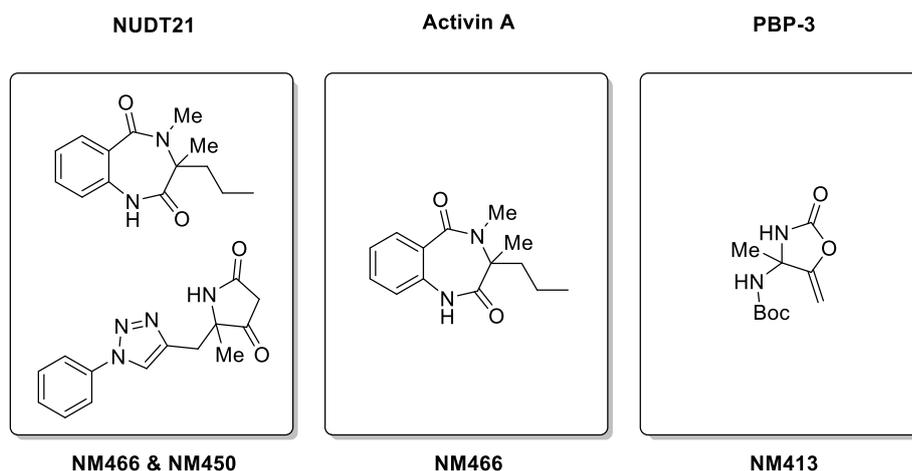
**Scheme 5.2.** Virtual library enumeration through variation of the quaternary substituent.

Alternatively, library enumeration could be investigated on an individual basis depending on the library member in question. In this instance, taking advantage of the modular synthetic strategy for constructing the quaternary precursors and the expanse of potential coupling partners employed at each stage, further appendage and functional group diversity could be achieved (Scheme 5.3). This would involve expansion of the functionalities previously demonstrated as viable positions for modification as disclosed in Chapter 2.



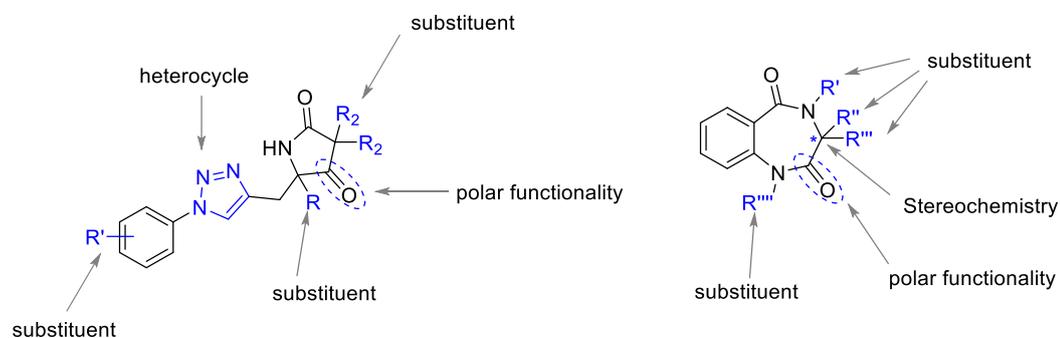
**Scheme 5.3.** Virtual library expansion through R group variation.

In the second chapter, the implementation of the DOS-derived library within a high-throughput X-ray crystallographic-based screening platform was described. Indeed, the biological relevance and structural diversity of the collection was highlighted through the identification of several novel binders against three challenging targets from distinct protein families (Figure 5.1). Remarkably, in the case of CFI<sub>25</sub> these results represent the first examples of SM binders for this important target. These results therefore present an exciting opportunity to develop novel SM probes to enable further elucidation of its role within disease settings.



**Figure 5.1.** XChem screens revealed hits against targets from three different protein families.

Subsequently, we initiated synthetic efforts toward hit analogues with the goal of hit validation. In this work, we successfully demonstrated the advantages of modular DOS-derived synthetic routes for enabling the rapid generation several derivatives of both hits (Figure 5.2). Importantly, this work highlighted the potential for fragment growth from a variety of different exit vectors as a result of utilising DOS chemistry, answering important calls from within the field. This included the exemplification of several viable synthetic strategies, enabling modification of almost every position around **NM450** and **NM466** to explore the binding mode. It is envisaged that upon biophysical determination of the binding affinity these explorations will prove beneficial to inform subsequent fragment growth.



**Figure 5.2.** A summary of the synthetic efforts toward the generation of hit analogues demonstrating the feasibility of fragment growth from multiple vectors.

In turn, these analogues were utilised to validate the initial crystallography results and therefore represent novel starting points for drug discovery or chemical probe development. Thus, following these initial results, subsequent investigations could be envisioned focusing on the development of these hits to generate potent and selective hit compounds. Indeed, on-going work in this area is focused on the exploration of appropriate biophysical methods to enable binding affinity measurement. Following this, pragmatic decisions on fragment growth could subsequently be made, building upon the preliminary follow-up investigations described

herein. In this context, the crystal structures generated in this work will provide valuable structural information to aid this process.

Finally, in the third chapter we described initial developments toward novel methodology for high-throughput analogue production and crude reaction mixture screening in collaboration with XChem. Organic synthesis remains one of the most time-consuming and cost-intensive segments within drug discovery and the emergence of robot-automated methods of synthesis and removing the necessity for purification represents one strategy often utilised to tackle this hurdle. In this work, we successfully demonstrated the applicability of the quaternary DOS library with to both a high-throughput synthesis setting and solid-supported catalysis, through the formation of 78 1,2,3-triazoles. Thus, these proof-of-concept studies highlight the potential for DOS library enumeration and rapid derivative production to enable widespread adoption of the library. Whilst utility of this reaction in a crude screening context remain to be validated, we believe this initial work will provide a solid foundation for on-going optimisation studies toward reaching this goal. Future investigations should seek to explore the application of metal chelating resins in the hope to avoid crystal degradation.

As this work serves to demonstrate, DOS methodology can be leveraged to tackle outstanding challenges within the field of FBS. Moreover, the establishment of translational collaborations can be particularly fruitful in enabling the identification of novel starting points for drug discovery stemming from such libraries. It is envisioned that the continued partnership of academic and industrial groups would be highly beneficial in this regard for future fragment-based projects.



## 6 Experimental

### 6.1 General Experimental Details

Proton nuclear magnetic resonance ( $^1\text{H}$  NMR) were recorded at ambient temperature on a Bruker DPX-400 (400 MHz), Bruker Advance 400 QNP (400 MHz,) and Bruker Avance 500 Cryo Ultrashield (500 MHz). Tetramethylsilane was used as an internal standard.  $^1\text{H}$  NMR chemical shifts ( $\delta_{\text{H}}$ ) are reported in parts per million (ppm), to the nearest 0.01 ppm and are referenced to the residual non-deuterated solvent peak ( $\text{CDCl}_3$ : 7.26,  $\text{CD}_3\text{OD}$ : 3.31,  $\text{D}_2\text{O}$ : 4.79). Coupling constants ( $J$ ) are reported in Hertz (Hz) to the nearest 0.1 Hz. Data are reported as follows: chemical shift, integration, multiplicity (s = singlet; d = doublet; t = triplet; q = quartet; qn = quintet; sep = septet; m = multiplet; or as a combination of these), coupling constant(s) and assignment. Carbon NMR ( $^{13}\text{C}$  NMR) were recorded at ambient temperature on a Bruker DPX-400 (400 MHz), Bruker Advance 400 QNP (400 MHz,) and Bruker Avance 500 Cryo Ultrashield (500 MHz). Chemical shifts ( $\delta_{\text{C}}$ ) are quoted in ppm, to the nearest 0.1 ppm, and are referenced to the residual non-deuterated solvent peak ( $\text{CDCl}_3$ : 77.16,  $\text{CD}_3\text{OD}$ : 49.00).

$^1\text{H}$  NMR and  $^{13}\text{C}$  NMR spectra assignments were supported by DEPT-135, COSY (2D,  $^1\text{H}$ - $^1\text{H}$  correlations), HSQC (2D, one bond  $^1\text{H}$ - $^{13}\text{C}$  correlations), HMBC (2D, multi-bonds  $^1\text{H}$ - $^{13}\text{C}$  correlations) where appropriate. The numbering of molecules used for  $^{13}\text{C}$  and  $^1\text{H}$  NMR assignments does not conform to IUPAC standards. High resolution mass spectrometry (HRMS) measurements were recorded with a Micromass Q-TOF mass spectrometer or a Waters LCT Premier Time of Flight mass spectrometer using Electrospray ionisation (ESI) techniques. Mass values are reported within the  $\pm 5$  ppm error limit.

Thin layer chromatography (TLC) was performed using pre-coated Merck glass backed silica gel 60 F<sub>254</sub> plates and visualised by quenching of UV fluorescence ( $\lambda_{\text{max}} = 254$  nm) or by staining with potassium permanganate. Retention factors ( $R_f$ ) are quoted to 0.01. Flash column chromatography was carried out Merck 9385 Kieselgel 60  $\text{SiO}_2$  (230-400 mesh) under a positive pressure of nitrogen unless otherwise stated. Visualisation was achieved *via* ultraviolet light (254 nm) and chemical staining with  $\text{KMnO}_4$  as appropriate.

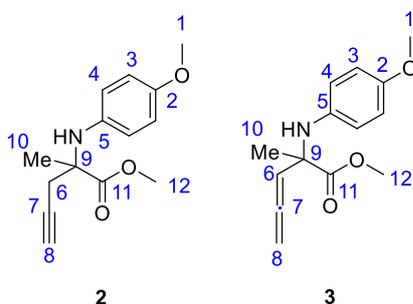
Liquid chromatography mass spectroscopy (LCMS) was carried out using a Waters ACQUITY H-Class UPLC with an ESCi Multi-Mode Ionisation Waters SQ Detector 2 spectrometer using MassLynx 4.1 software; EI refers to the electrospray ionisation technique; LC system: solvent A: 2 mM  $\text{NH}_4\text{OAc}$  in  $\text{H}_2\text{O}/\text{MeCN}$  (95:5); solvent B: MeCN; solvent C: 2% formic acid; column: ACQUITY UPLC<sup>®</sup> CSH C18 (2.1 mm  $\times$  50 mm, 1.7  $\mu\text{m}$ , 130  $\text{\AA}$ ) at 40  $^\circ\text{C}$ ; gradient: 5 – 95 % B with constant 5 % C over 1 minute at flow rate of 0.6 mL/minute; detector: PDA e $\lambda$  Detector 220 – 800 nm, interval 1.2 nm. Melting points (m.p.) were obtained using a Büchi Melting Point B-545 melting

point apparatus and are uncorrected. Infrared (IR) spectra were recorded neat on a Perkin-Elmer Spectrum One spectrometer using an ATR sampling accessory either as solids or films in CH<sub>2</sub>Cl<sub>2</sub>. Selected absorptions ( $\nu_{\max}$ ) are reported in wavenumbers (cm<sup>-1</sup>) with the following abbreviations: w, weak; m, medium; s, strong; br, broad.

Ethyl acetate, methanol, dichloromethane, acetonitrile and toluene were distilled from calcium hydride. Diethyl ether was distilled from a mixture of lithium aluminium hydride and calcium hydride. Tetrahydrofuran was dried using Na wire and distilled from a mixture of lithium aluminium hydride and calcium hydride with triphenylmethane as an indicator. Petroleum ether was distilled before use and refers to the fraction between 40-60 °C. All other solvents and reagents were obtained from commercial suppliers and used without further purification. Reactions were carried out under a stream of dry nitrogen using oven-dried glassware unless otherwise stated. Room temperature (r.t.) refers to ambient temperature. All temperatures below 0 °C were maintained using an acetone-cardice bath. Temperatures of 0 °C were maintained using an ice-water bath.

## 6.2 *N*-substituted quaternary fragments

### 6.2.1 Synthesis of methyl 2-((4-methoxyphenyl)amino)-2-methylpent-4-ynoate (**2**) and methyl 2-((4-methoxyphenyl)amino)-2-methylpenta-3,4-dienoate (**3**)



To a solution of *p*-anisidine (4.40 g, 35.7 mmol) in acetonitrile (ACN) (40.0 mL) was added methyl pyruvate (4.70 mL, 52.0 mmol) and the reaction heated to 50 °C for 80 minutes. Upon completion, the reaction was concentrated *in vacuo* and the crude imine was redissolved in a mixture of petroleum ether:CH<sub>2</sub>Cl<sub>2</sub> (10:1). The resulting precipitate was removed by filtration to afford an orange/brown oil. The crude imine (7.10 g, 34.3 mmol) was dissolved in DMF (172 mL) and cooled to 0 °C before propargyl bromide (4.87 mL, 45.2 mmol) was added, followed by activated Zn\* powder (3.36 g, 51.7 mmol). The reaction was warmed to r.t. and then heated to 60 °C for 1 hour. Upon completion, the reaction was cooled to 0 °C and quenched with NH<sub>4</sub>Cl saturated aqueous solution. The product was extracted with EtOAc (3 x 60 mL) and washed with brine (4 x 50 mL). The combined organic extracts were dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 4:1) to give the products; **2** (4.16 g, 16.8 mmol, 47 %) and **3** (1.04 g, 4.20 mmol, 12 %) as orange/brown oils.

\*Zn was stirred with 2N HCl (aq) for 10-15 minutes and then washed with H<sub>2</sub>O, EtOH, Et<sub>2</sub>O and rigorously dried before use.

Data for **2**:

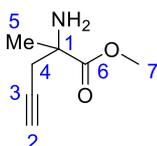
$R_f$  = 0.43 (petroleum ether/EtOAc, 4:1); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 6.77 – 6.70 (4H, m, H3/4), 3.75 (3H, s, H12), 3.74 (3H, s, H1), 2.79 (1H, dd,  $J$  = 16.9, 2.7 Hz, H6a), 2.66 (1H, dd,  $J$  = 16.8, 2.7 Hz, H6b), 2.09 (1H, t,  $J$  = 2.7 Hz, H8), 1.55 (3H, s, H10); <sup>13</sup>C (101 MHz, CDCl<sub>3</sub>) δ 175.1 (C11), 154.7 (C2), 137.9 (C5), 121.4 (C4), 114.5 (C3), 79.7 (C7), 71.9 (C8), 61.3 (C9), 55.6 (C1), 52.7 (C12), 28.0 (C6), 24.0 (C10); **HRMS** (ESI) calcd for [C<sub>14</sub>H<sub>18</sub>NO<sub>3</sub>]<sup>+</sup>: 248.1287, found 248.1275; **IR**  $\nu_{max}$  = 3283 (m, C≡C-H), 2952 (w, C-H), 2120 (w, C≡C), 1727 (s, C=O), 1509 (s, C=C), 1108 (s, Ar-O-C).

These characterisation data are in accordance with that previously reported in the literature.<sup>193</sup>

Data for **3**:

$R_f = 0.50$  (petroleum ether/EtOAc, 4:1);  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  6.76 – 6.72 (2H, m, H4), 6.68 – 6.64 (2H, m, H3), 5.55 (1H, t,  $J = 6.7$  Hz, H6), 4.96 – 4.87 (2H, m, H8), 3.74 (3H, s, H1), 3.73 (3H, s, H12), 1.60 (3H, s, H10);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  207.5 (C7), 175.0 (C11), 153.6 (C2), 138.7 (C5), 118.9 (C4), 114.5 (C3), 95.3 (C6), 79.2 (C8), 60.6 (C9), 55.7 (C1), 52.9 (C12), 24.5 (C10); **HRMS** (ESI) calcd for  $[\text{C}_{14}\text{H}_{18}\text{NO}_3]^+$ : 248.1287, found 248.1275; **IR**  $\nu_{\text{max}} = 2952$  (w, C-H), 1956 (w, C=C=C), 1728 (s, C=O), 1510 (s, C=C), 1236 (s, Ar-O-C).

### 6.2.2 Synthesis of methyl 2-amino-2-methylpent-4-ynoate (**4**)

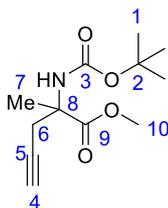


To an ice-cooled solution of alkyne **2** (7.10 g, 28.8 mmol) in ACN (68.0 mL) was added a solution of cerium ammonium nitrate (CAN) (31.4 g, 57.4 mmol) in  $\text{H}_2\text{O}$  (68.0 mL), dropwise over 30 minutes. The mixture was allowed to warm to r.t. and, upon completion, 2N HCl (120 mL) was added and the reaction stirred for 20 minutes. The solution was washed with EtOAc (3 x 80 mL) and the aqueous layer then basified to pH 10 using  $\text{Na}_2\text{CO}_3$ . The resulting aqueous emulsion was then extracted with  $\text{CH}_2\text{Cl}_2$  (2 x 150 mL) and the organic extracts combined, dried over  $\text{Na}_2\text{SO}_4$ , filtered and concentrated *in vacuo* to give **4** (2.80 g, 19.8 mmol, 67%) as a brown oil.

$R_f = 0.21$  (petroleum ether/EtOAc, 1:1);  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  3.72 (3H, s, H7), 2.64 (1H, dd,  $J = 16.5, 2.6$  Hz, H4a), 2.44 (1H, dd,  $J = 16.5, 2.6$  Hz, H4b), 2.04 (1H, t,  $J = 2.6$  Hz, H2), 1.86 (2H, br s,  $\text{NH}_2$ ), 1.36 (3H, s, H5);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  176.6 (C6), 79.8 (C3), 71.5 (C2), 57.5 (C1), 52.6 (C7), 31.0 (C4), 26.0 (C5); **HRMS** (ESI) calcd for  $[\text{C}_7\text{H}_{12}\text{NO}_2]^+$ : 142.0863, found 142.0865; **IR**  $\nu_{\text{max}} = 3293$  (w,  $\text{C}\equiv\text{C-H}$ ), 2956 (C-H), 2114 (w,  $\text{C}\equiv\text{C}$ ), 1730 (s, C=O), 1598 (w,  $\text{NH}_2$ ), 1208 (s, C-O-C).

These characterisation data are in accordance with that previously reported in the literature.<sup>329</sup>

### 6.2.3 Synthesis of methyl 2-((*tert*-butoxycarbonyl)amino)-2-methylpent-4-ynoate (6)

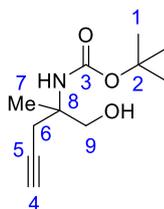


To a solution of **4** (1.00 g, 7.09 mmol) in THF (24 mL) was added  $\text{Boc}_2\text{O}$  (2.32 g, 10.6 mmol) and the reaction heated to 70 °C for 18 hours in a sealed tube. The reaction was concentrated *in vacuo* and purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 3:1) to give **6** as a white solid (1.66 g, 6.88 mmol, 97%).

$R_f$  = 0.42 (petroleum ether/EtOAc, 3:1); **m.p.** 76.4 – 78.1 °C;  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  5.24 (1H, br s, NH), 3.76 (3H, s, H10), 2.95 – 2.86 (2H, m, H6), 2.02 (1H, t,  $J$  = 2.6 Hz, H4), 1.55 (3H, s, H7), 1.44 (9H, s, H1);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  173.7 (C9), 154.5 (C3), 80.2 (C2), 79.5 (C5), 71.3 (C4), 58.4 (C8), 52.9 (C10), 28.4 (C1), 27.1 (C6), 23.4 (C7); **HRMS** (ESI) calcd for  $[\text{C}_{12}\text{H}_{20}\text{NO}_4]^+$ : 242.1392, found 242.1382; **IR**  $\nu_{\text{max}}$  = 3355 (w,  $\text{C}\equiv\text{C}-\text{H}$ ) 2997 (w, C-H), 2123 (w,  $\text{C}\equiv\text{C}$ ), 1733 (s, C=O), 1695 (s, C=O).

These characterisation data are in accordance with that previously reported in the literature.<sup>330</sup>

### 6.2.4 Synthesis of *tert*-butyl (1-hydroxy-2-methylpent-4-yn-2-yl)carbamate (7)

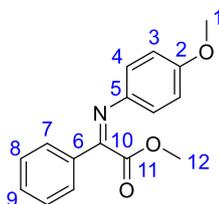


To an ice-cooled solution of **6** (1.62 g, 6.70 mmol) in THF (30 mL) was added  $\text{LiBH}_4$  (438 mg, 20.0 mmol) and the reaction stirred for 18 hours. Then, the reaction mixture was quenched with  $\text{H}_2\text{O}$  and extracted with EtOAc (3 x 40 mL). The combined organic extracts were dried over  $\text{Na}_2\text{SO}_4$ , filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc 3:1) to give **7** as a white solid (1.36 g, 6.38 mmol, 96%).

$R_f$  = 0.35 (petroleum ether/EtOAc, 4:1); **m.p.** 50.6 – 52.1 °C;  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  4.88 (1H, br s, OH), 3.68 (2H, s, H9), 2.73 (1H, dd,  $J$  = 16.8, 2.6 Hz, H6a), 2.45 (1H, dd,  $J$  = 16.9, 2.7 Hz, H6b), 2.05 (1H, t,  $J$  = 2.7 Hz, H4), 1.43 (9H, s, H1), 1.29 (3H, s, H7);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  156.2 (C3), 80.3 (C2/C5), 80.3 (C2/C5), 71.3 (C4), 69.2 (C9), 56.1 (C8), 28.5 (C1),

26.8 (C6), 22.7 (C7); **HRMS** (ESI) calcd for  $[C_{11}H_{19}NO_3Na]^+$ : 236.1257, found 236.1247; **IR**  $\nu_{max}$  = 3675 (b, OH), 3273 (W, C≡C-H), 2973 (s, C-H), 2110 (w, C≡C), 1667 (s, C=O).

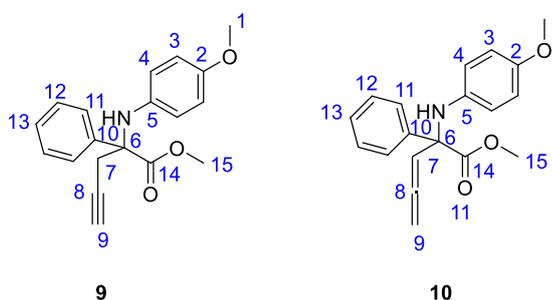
### 6.2.5 Synthesis of methyl (Z)-2-((4-methoxyphenyl)imino)-2-phenylacetate (8)



A solution of methyl benzyl formate (250 mg, 1.52 mmol), *p*-anisidine (197 mg, 1.60 mmol) and *p*-TsOH (14.4 mg, 0.0760 mmol) in toluene (3 mL) was heated to reflux with a dean-stark for 190 minutes. The reaction was cooled to r.t. and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 3:1) to give **8** (361 mg, 1.34 mmol, 88%) as a yellow solid.

$R_f$  = 0.55 (petroleum ether/EtOAc, 3:1); **m.p.** 87.3 – 88.8 °C;  $^1H$  NMR (400 MHz,  $CDCl_3$ )  $\delta$  7.87 – 7.84 (2H, m, H7), 7.53 – 7.43 (3H, m, H8 & H9), 6.99 – 6.95 (2H, m, H4), 6.90 – 6.87 (2H, m, H3), 3.81 (3H, s, H1), 3.70 (3H, s, H12);  $^{13}C$  (101 MHz,  $CDCl_3$ )  $\delta$  166.2 (C11), 159.3 (C10), 157.5 (C2), 143.3 (C5), 134.3 (C6), 131.7 (C9), 128.8 (C8), 127.9 (C7), 121.3 (C4), 114.3 (C3), 55.6 (C1), 52.1 (C12); **HRMS** (ESI) calcd for  $[C_{16}H_{16}NO_3]^+$ : 270.1133, found 270.1130; **IR**  $\nu_{max}$  = 2950 (w, C-H), 1733 (s, C=O), 1618 (m, C=N), 1576 (m, C=C), 1498 (s, C=C), 1226 (s, Ar-O-C).

### 6.2.6 Synthesis of methyl 2-((4-methoxyphenyl)amino)-2-phenylpent-4-ynoate (9) and methyl 2-((4-methoxyphenyl)amino)-2-phenylpenta-3,4-dienoate (10)



To an ice-cooled solution of **8** (620 mg, 2.30 mmol) and propargyl bromide (0.484 mL, 3.57 mmol) in DMF (13.6 mL) was added  $Zn^*$  (301 mg, 4.60 mmol) and the reaction warmed to r.t. over 1 hour and then heated to 60 °C for 75 minutes. The reaction was quenched with  $NH_4Cl$  saturated aqueous solution and extracted with EtOAc (3 x 30 mL). The combined organic extracts were washed with brine, dried over  $Na_2SO_4$ , filtered and concentrated *in vacuo*. The

crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 85:15) to give **9** (457 mg, 1.47 mmol, 64%) and **10** (76.2 mg, 0.246 mmol, 11%) both as orange oils.

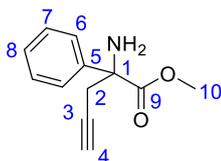
Data for **9**:

$R_f$  = 0.48 (petroleum ether/EtOAc, 85:15);  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.58 – 7.55 (2H, m, H11), 7.40 – 7.31 (3H, m, H12 & H13), 6.65 – 6.61 (2H, m, H3), 6.42 – 6.38 (2H, m, H4), 4.97 (1H, br s, NH), 3.72 (3H, s, H15), 3.69 (3H, s, H1), 3.39 (1H, dd,  $J$  = 16.4, 2.6 Hz, H7a), 3.21 (1H, dd,  $J$  = 16.4, 2.6 Hz, H7b), 2.02 (1H, t,  $J$  = 2.6 Hz, H9);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  173.1 (C14), 153.3 (C2), 139.8 (C10), 137.9 (C5), 128.9 (C12), 128.2 (C13), 127.0 (C11), 118.8 (C4), 114.5 (C3), 79.4 (C8), 71.9 (C9), 66.9 (C6), 55.6 (C1), 53.4 (C15), 25.04 (C7); **HRMS** (ESI) calcd for  $[\text{C}_{19}\text{H}_{20}\text{NO}_3]^+$ : 310.1443, found 310.1430; **IR**  $\nu_{\text{max}}$  = 3404 (w, C $\equiv$ C-H), 2952 (w, C-H), 2168 (w, C $\equiv$ C), 1732 (s, C=O), 1511 (s, C=C), 1239 (s, Ar-O-R).

Data for **10**:

$R_f$  = 0.58 (petroleum ether/EtOAc, 85:15);  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.64 – 7.61 (2H, m, H11), 7.38 – 7.28 (3H, m, H12 & H13), 6.65 – 6.61 (2H, m, H3), 6.42 – 6.38 (2H, m, H4), 5.97 (1H, t,  $J$  = 6.6 Hz, H7), 5.10 (1H, br s, NH), 4.88 – 4.78 (2H, m, H9), 3.69 (3H, s, H15) 3.69 (3H, s, H1);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  208.2 (C8), 173.0 (C14), 152.4 (C2), 140.1 (C10), 138.5 (C5), 128.7 (C12), 128.0 (C13), 127.6 (C11), 117.0 (C4), 114.2 (C3), 92.5 (C7), 78.8 (C9), 66.9 (C6), 55.6 (C1), 53.1 (C15); **HRMS** (ESI) calcd for  $[\text{C}_{19}\text{H}_{20}\text{NO}_3]^+$ : 310.1443, found 310.1434; **IR**  $\lambda_{\text{max}}$  = 2951 (w, C-H), 1735 (s, C=O), 1509 (C=C), 1236 (Ar-O-C).

### 6.2.7 Synthesis of methyl 2-amino-2-phenylpent-4-ynoate (**11**)

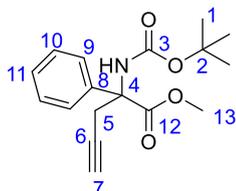


To an ice-cooled solution of **9** (217 mg, 0.702 mmol) in ACN (6 mL) was added a solution of CAN (1.53, 2.80 mmol) in  $\text{H}_2\text{O}$  (6.5 mL) dropwise. The mixture was allowed to warm to r.t. and then stirred for 23 hours. 2N HCl (aq) (6 mL) was added and the reaction stirred for 30 minutes. Upon completion, the reaction was extracted with EtOAc (3 x 20 mL) and the aqueous layer basified with  $\text{Na}_2\text{CO}_3$  to pH 10. The resulting cream suspension was extracted with  $\text{CH}_2\text{Cl}_2$  (3 x 30 mL) and the combined organic extracts were dried over  $\text{Na}_2\text{SO}_4$ , filtered and concentrated *in vacuo* to give **11** (110 mg, 0.542 mmol, 77%) as a brown oil without further purification.

$R_f$  = 0.47 (petroleum ether/EtOAc, 3:2);  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.53 – 7.50 (2H, m, H6), 7.38 – 7.33 (2H, m, H7) 7.32 – 7.28 (1H, m, H8), 3.75 (3H, s, H10), 3.14 (1H, dd,  $J$  = 16.5, 3.2 Hz, H2a), 2.77 (1H, dd,  $J$  = 16.5, 2.5 Hz, H2b), 2.18 (2H, br s,  $\text{NH}_2$ ), 2.06 (1H, t,  $J$  = 2.6 Hz,

H4);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  174.9 (C9), 141.5 (C5), 128.7 (C7), 128.1 (C6), 125.4 (C8), 80.0 (C3), 71.6 (C4), 63.5 (C1), 53.0 (C10), 31.1 (C2); **HRMS** (ESI) calcd for  $[\text{C}_{12}\text{H}_{14}\text{NO}_2]^+$ : 204.1021, found 204.1025; **IR**  $\nu_{\text{max}}$  = 3288 (m,  $\text{C}\equiv\text{C}-\text{H}$ ), 2953 (w, C-H), 1730 (s,  $\text{C}=\text{O}$ ), 1600 (m,  $\text{C}=\text{C}$ ).

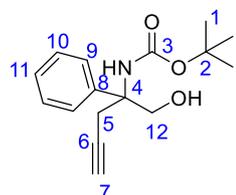
### 6.2.8 Synthesis of methyl 2-((*tert*-butoxycarbonyl)amino)-2-phenylpent-4-ynoate (12)



A solution of **11** (91.0 mg, 0.448 mmol) and  $\text{Boc}_2\text{O}$  (146 mg, 0.672 mmol) in THF (4 mL) was heated to 75 °C for 18 hours in a sealed tube. Further  $\text{Boc}_2\text{O}$  (48.0 mg, 0.219 mmol) was added and the reaction stirred at 80 °C for a further 6 hours then cooled to r.t. and stirred for 18 hours. The reaction was concentrated *in vacuo* and purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 85:15) to give **12** (108 mg, 0.356 mmol, 80%) as a pale-yellow solid.

$R_f$  = 0.32 (petroleum ether/EtOAc, 9:1); **m.p.** 100.3 – 101.9 °C;  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.44 (2H, d,  $J$  = 7.5 Hz, H9), 7.38 – 7.34 (2H, m, H10), 7.32 – 7.38 (1H, m, H11), 6.03 (1H, br s, NH), 3.71 (3H, s, H13), 3.62 – 3.43 (2H, m, H5), 1.99 (1H, m, H7), 1.41 (9H, br s, H1);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  171.9 (C12), 154.1 (C3), 138.5 (C8), 128.8 (C10), 128.3 (C11), 126.0 (C9), 85.1 (C2), 79.6 (C6), 71.3 (C7), 64.5 (C4), 53.5 (C13), 28.4 (C1), 25.3 (C5); **HRMS** (ESI) calcd for  $[\text{C}_{17}\text{H}_{21}\text{NO}_4\text{Na}]^+$ : 326.1368, found 326.1365; **IR**  $\lambda_{\text{max}}$  = 3321 (w,  $\text{C}\equiv\text{C}-\text{H}$ ), 2976 (w, C-H), 2012 (w,  $\text{C}\equiv\text{C}$ ), 1736 (s,  $\text{C}=\text{O}$ ), 1712 (s,  $\text{C}=\text{O}$ ), 1487 (s,  $\text{C}=\text{C}$ ), 1434 (C=C).

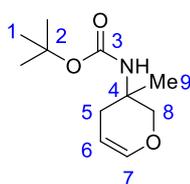
### 6.2.9 Synthesis of *tert*-butyl (1-hydroxy-2-phenylpent-4-yn-2-yl)carbamate (13)



To an ice-cooled solution of **12** (55.0 mg, 0.181 mmol) in THF (2 mL) was added  $\text{LiBH}_4$  (13.0 mg, 0.594 mmol) and the reaction warmed to r.t before heating to 60 °C for 3 hours. The reaction was cooled to r.t. and quenched with  $\text{H}_2\text{O}$ . The product was extracted with EtOAc (3 x 20 mL), washed with brine, dried over  $\text{Na}_2\text{SO}_4$ , filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 85:15) to give **13** (26.5 mg, 0.0962 mmol, 53%) as a white solid.

$R_f$  = 0.10 (petroleum ether/EtOAc, 9:1); **m.p.** 119.1 – 120.5 °C;  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.42 – 7.35 (4H, m, H9 & H10), 7.31 – 7.27 (1H, m, H11), 5.45 (1H, br s, NH), 4.04 (1H, m, H12a), 3.92 – 3.81 (1H, m, H12b & OH), 3.06 (1H, dd,  $J$  = 16.9, 2.6 Hz, H5a), 2.95 (1H, dd,  $J$  = 16.9, 2.8 Hz, H5b), 2.04 (1H, br s, H7), 1.45 (9H, br s, H1);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  156.1 (C3), 141.2 (C8), 128.7 (C10), 127.7 (C11), 126.0 (C9), 80.5 (C2), 79.9 (C6), 72.2 (C7), 69.7 (C12), 62.1 (C4), 28.4 (C1), 27.3 (C5); **HRMS** (ESI) calcd for  $[\text{C}_{16}\text{H}_{22}\text{NO}_3]^+$ : 276.1600, found 276.1595; **IR**  $\nu_{\text{max}}$  = 3306 (O-H), 3076 (m, C=C-H), 2046 (w, C≡C), 1682 (s, C=O), 1557 (s, C=C), 1582 (s, C=C).

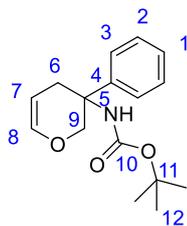
### 6.2.10 Synthesis of *tert*-butyl (3-methyl-3,4-dihydro-2*H*-pyran-3-yl)carbamate (**14**)



A solution of **7** (19.0 mg, 0.0890 mmol), *N*-hydroxysuccinimide (5.00 mg, 0.0445 mmol), tetrabutylammonium hexafluorophosphate (17.3 mg, 0.0445 mmol) and  $\text{NaHCO}_3$  (3.80 mg, 0.00270 mmol) in DMF (1 mL) was degassed with an argon purge for 15 minutes.  $\text{CpRu}(\text{PPh}_3)_2\text{Cl}$  (6.00 mg, 8.26  $\mu\text{mol}$ ) and  $\text{PPh}_3$  (2.10 mg, 8.00  $\mu\text{mol}$ ) were added and the reaction heated to 80 °C for 56 hours in a sealed tube. The reaction was cooled to r.t. and filtered through celite, washing with EtOAc. The filtrate was washed with brine, dried over  $\text{Na}_2\text{SO}_4$ , filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc 3:2) to give of **14** (9.10 mg, 0.427 mmol, 47%) as a colourless oil.

$R_f$  = 0.39 (petroleum ether/EtOAc 19:1);  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  6.36 (1H, dt,  $J$  = 6.1, 1.8 Hz, H7), 4.67 – 4.63 (2H, m, H6 & NH), 4.04 (1H, d,  $J$  = 10.2 Hz, H8a), 3.56 (1H, d,  $J$  = 11.0 Hz, H8b), 2.28 (1H, d,  $J$  = 16.9 Hz, H5a), 1.97 (1H, d,  $J$  = 17.2 Hz, H5b), 1.43 (9H, s, H1), 1.37 (3H, s, H9);  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  155.0 (C3), 143.3 (C7), 98.8 (C6), 79.3 (C2), 71.0 (C8), 48.7 (C4), 32.6 (C5), 28.6 (C1), 22.9 (C9); **HRMS** (ESI) calcd for  $[\text{C}_{11}\text{H}_{19}\text{NO}_3\text{Na}]^+$ : 236.1257, found 236.1250; **IR**  $\nu_{\text{max}}$  = 2975 (m, C-H), 1715 (s, C=O), 1680 (m, C=C), 1136 (s, C-O).

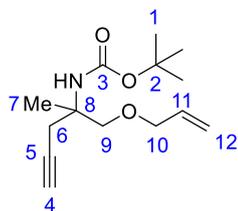
### 6.2.11 Synthesis of *tert*-butyl (3-phenyl-3,4-dihydro-2*H*-pyran-3-yl)carbamate (15)



To a solution of **13** (20.0 mg, 0.0726 mmol) in DMF (0.75 mL) was added *N*-hydroxy succinimide (4.20 mg, 0.0363 mmol), tetrabutylammonium hexafluorophosphate (3.60 mg, 9.29  $\mu$ mol) and NaHCO<sub>3</sub> (3.00 mg, 35.7  $\mu$ mol). The solution was degassed with argon for 15 minutes before CpRu(PPh<sub>3</sub>)<sub>2</sub>Cl (5.30 mg, 0.00730 mmol) and PPh<sub>3</sub> (4.00 mg, 0.00145 mmol) were added. The reaction was heated to 80 °C for 18 hours in a sealed tube. Upon completion, the reaction was cooled to r.t. and diluted with water. The aqueous solution was extracted with CH<sub>2</sub>Cl<sub>2</sub> (3 x 15 mL) and the combined organic extracts were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated *in vacuo*. The resulting oil was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 4:1) to give **15** (9.00 mg, 0.0329 mmol, 45%) as a yellow/brown oil.

$R_f$  = 0.44 (petroleum ether/EtOAc, 19:1); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  7.38 – 7.33 (4H, m, H2 & H3), 7.28 – 7.23 (1H, m, H1), 6.46 (1H, dt,  $J$  = 6.0, 1.7 Hz, H8), 5.09 (1H, s, NH), 4.84 – 4.80 (1H, m, H7), 4.26 (1H, d,  $J$  = 10.1 Hz, H9a), 3.88 (1H, d,  $J$  = 10.9 Hz, H9b), 2.64 – 2.44 (2H, m, H6), 1.37 (9H, s, H12); <sup>13</sup>C (101 MHz, CDCl<sub>3</sub>)  $\delta$  154.5 (C10), 143.5 (C8), 141.3 (C4), 128.6 (C2), 127.3 (C1), 125.4 (C3), 99.3 (C7), 79.6 (C11), 70.6 (C9), 54.0 (C5), 32.3 (C6), 28.4 (C12); HRMS (ESI) calcd for [C<sub>16</sub>H<sub>22</sub>NO<sub>3</sub>]<sup>+</sup> 276.1607, found 276.1600; IR  $\nu_{\max}$  = 2975 (w, C-H), 1698 (s, C=O), 1655 (m, C=C), 1494 (m, C=C), 1164 (C-O-C).

### 6.2.12 Synthesis of *tert*-butyl (1-(allyloxy)-2-methylpent-4-yn-2-yl)carbamate (16)

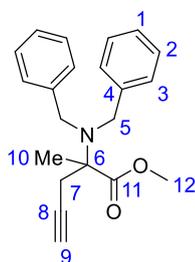


To an ice-cooled solution of **7** (40.0 mg, 0.188 mmol) in DMF (2 mL) was added NaH (8.20 mg, 0.206 mmol) and the reaction stirred for 15 minutes before allyl bromide (45.5  $\mu$ L, 0.376 mmol) was added and the reaction stirred for 90 minutes at r.t. Upon completion the reaction was diluted with H<sub>2</sub>O and extracted with EtOAc (2 x 30 mL). The combined organic extracts were dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated *in vacuo*. The crude product was purified

by flash column chromatography (silica gel; petroleum ether/EtOAc, 3:2) to give **16** (28.0 mg, 0.111 mmol, 59%) as a colourless oil.

$R_f$  = 0.54 (petroleum ether/EtOAc, 9:1);  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  5.94 – 5.84 (1H, m, H11), 5.27 (1H, m, H12a), 5.18 (1H, m, H12b), 4.84 (1H, br s, NH), 4.01 (2H, m, H10), 3.58 (1H, d,  $J$  = 9.1 Hz, H9a), 3.42 (1H, d,  $J$  = 9.1 Hz, H9b) 2.78 (1H, dd,  $J$  = 16.3, 2.3 Hz, H6a), 2.57 (1H, dd,  $J$  = 16.8, 2.7 Hz, H6b), 1.98 (1H, t,  $J$  = 2.7 Hz, H4), 1.43 (9H, s, H1), 1.39 (3H, s, H7);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  154.9 (C3), 134.7 (C11), 117.2 (C12), 80.8 (C5), 79.3 (C2), 74.1 (C9), 72.4 (C10), 70.5 (C4), 54.7 (C8), 28.5 (C1), 26.5 (C6), 21.6 (C7); **HRMS** (ESI) calcd for  $[\text{C}_{14}\text{H}_{23}\text{NO}_3\text{Na}]^+$ : 276.1576, found 276.1570; **IR**  $\nu_{\text{max}}$  = 3305 (w,  $\text{C}\equiv\text{C-H}$ ), 2977 (w, C-H), 1715 (s, C=O), 1496 (s, C=C), 1164 (C-O-C).

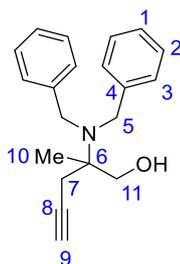
### 6.2.13 Synthesis of methyl 2-(dibenzylamino)-2-methylpent-4-ynoate (**17**)



A solution of **4** (400 mg, 2.83 mmol), DIPEA (2.47 mL, 14.1 mmol) and benzyl bromide (1.69 mL, 14.1 mmol) in ACN (18 mL) was heated to 80 °C for 75 hours. The reaction was cooled to r.t., concentrated *in vacuo* and the resulting gum diluted with  $\text{H}_2\text{O}$  and extracted with EtOAc (3 x 50 mL). The combined organic extracts were washed with brine, dried over  $\text{Na}_2\text{SO}_4$ , filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 85:15) to give **17** (256 mg, 2.56 mmol, 90%) as an off-white solid.

$R_f$  = 0.54 (petroleum ether/EtOAc, 9:1); **m.p.** 65.2 – 66.7 °C;  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.32 (4H, d,  $J$  = 7.4 Hz, H3), 7.20 (4H, t,  $J$  = 7.4 Hz, H2), 7.12 (2H, t,  $J$  = 7.3 Hz, H1), 3.85 (4H, s, H5), 3.72 (3H, s, H12), 2.77 (1H, dd,  $J$  = 16.8, 2.6 Hz, H7a), 2.69 (1H, dd,  $J$  = 16.8, 2.6 Hz, H7b), 2.09 (1H, t,  $J$  = 2.8 Hz, H9), 1.50 (3H, s, H10);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  174.4 (C11), 141.1 (C4), 128.5 (C3), 128.1 (C2), 126.7 (C1), 81.0 (C8), 71.3 (C9), 66.7 (C6), 54.9 (C5), 51.8 (C12), 28.2 (C7), 22.4 (C10); **HRMS** (ESI) Calcd for  $[\text{C}_{21}\text{H}_{24}\text{NO}_2]^+$  322.1807, found 322.1818; **IR**  $\nu_{\text{max}}$  = 3297 (w,  $\text{C}\equiv\text{C-H}$ ), 2947 (w, C-H), 2119 (w,  $\text{C}\equiv\text{C}$ ), 1725 (s, C=O), 1602 (m, C=C), 1205 (C-O-C).

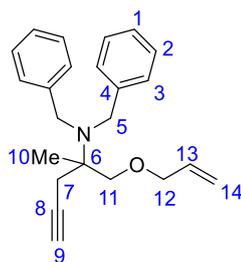
### 6.2.14 Synthesis of 2-(dibenzylamino)-2-methylpent-4-yn-1-ol (**18**)



To an ice-cooled solution of **17** (581 mg, 1.81 mmol) in THF (14 mL) was added a suspension of LiAlH<sub>4</sub> (137 mg, 3.62 mmol) in THF (4 mL) and the reaction stirred for 70 minutes whilst warming to r.t.. The reaction was quenched with 10% NaOH (aq) and extracted with EtOAc (3 x 40 mL). The combined organic extracts were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated *in vacuo* to give **18** (496 mg, 1.69 mmol, 93%) as a pale pink solid.

**R<sub>f</sub>** = 0.35 (petroleum ether/EtOAc, 85:15); **m.p.** 66.2 – 67.9 °C; **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.25 – 7.14 (10H, m, H1, H2 & H3), 3.86 (2H, d, *J* = 14.5 Hz, H5a), 3.79 (2H, d, *J* = 14.5 Hz, H5b), 3.69 (1H, d, *J* = 11.2 Hz, H11a), 3.53 (1H, d, *J* = 11.2 Hz, H11b), 2.63 (1H, br s, OH), 2.47 (1H, dd, *J* = 16.6, 2.6 Hz, H7a), 2.37 (1H, dd, *J* = 16.6, 2.7 Hz, H7b), 2.07 (1H, t, *J* = 2.7 Hz, H9), 1.26 (3H, s, H10); **<sup>13</sup>C** (101 MHz, CDCl<sub>3</sub>) δ 141.1 (C4), 128.5 (C2/3), 128.5 (C2/3), 127.0 (C1), 81.5 (C8), 71.3 (C9), 66.1 (C11), 62.3 (C6), 53.5 (C5), 25.1 (C7), 19.6 (C10); **HRMS** (ESI) calcd for [C<sub>20</sub>H<sub>24</sub>NO]<sup>+</sup> 294.1858, found 294.1849; **IR** *v*<sub>max</sub> = 3426 (b, O-H), 3277 (s, C≡C-H), 1598 (m, C=C), 1492 (C=C), 1450 (C=C).

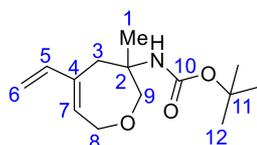
### 6.2.15 Synthesis of 1-(allyloxy)-*N,N*-dibenzyl-2-methylpent-4-yn-2-amine (**19**)



To an ice-cooled solution of NaH (54.0 mg, 1.36 mmol) and allyl bromide (177 uL, 2.05 mmol) in DMF (3 mL) was added a solution of **18** (190 mg, 0.648 mmol) in DMF (3.8 mL) dropwise. The reaction was stirred for 2 hours at r.t., quenched with NH<sub>4</sub>Cl saturated aqueous solution and extracted with CH<sub>2</sub>Cl<sub>2</sub> (2 x 20 mL). The combined organic extracts were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 9:1) to give **19** (161 mg, 0.485 mmol, 75%) as a colourless oil.

$R_f = 0.38$  (petroleum ether/EtOAc, 20:1);  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.32 (4H, d,  $J = 7.3$  Hz, H3), 7.18 (4H, t,  $J = 7.4$  Hz, H2), 7.10 (2H, t,  $J = 7.2$  Hz, H1), 5.95 – 5.85 (1H, m, H13), 5.28 (1H, dd,  $J = 6.3$  Hz, H14a), 5.17 (1H, dd,  $J = 4.0$ , H14b), 3.91 – 3.90 (6H, m H5 & H12), 3.52 (1H, dd,  $J = 9.6$  Hz, H11a), 3.46 (1H, dd,  $J = 9.6$  Hz, H11b), 2.60 (1H, dd,  $J = 16.8, 2.6$  Hz, H7a), 2.53 (1H, dd,  $J = 16.9, 2.6$  Hz, H7b), 2.07 (1H, t,  $J = 2.6$  Hz, H9), 1.26 (3H, s, H10);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  142.2 (C4), 135.1 (C13), 128.5 (C3), 127.9 (C2), 126.4 (C1), 116.5, (C14), 82.6 (C8), 75.0 (C11), 72.3 (C12), 70.5 (C9), 61.0 (C6), 54.3 (C5), 26.4 (C7), 21.2 (C10); **HRMS** (ESI) calcd for  $[\text{C}_{23}\text{H}_{28}\text{NO}]^+$ : 334.2171, found 334.2175; **IR**  $\nu_{\text{max}} = 3301$  (m,  $\text{C}\equiv\text{C-H}$ ), 2922 (m, C-H), 2120 (w,  $\text{C}\equiv\text{C}$ ), 1602 (m,  $\text{C}=\text{C}$ ), 1493 (s,  $\text{C}=\text{C}$ ), 1453 (s,  $\text{C}=\text{C}$ ), 1090 (C-O-C).

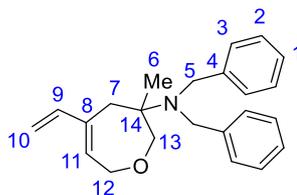
### 6.2.16 Synthesis of *tert*-butyl (3-methyl-5-vinyl-2,3,4,7-tetrahydrooxepin-3-yl)carbamate (**20**)



A solution of **16** (150 mg, 0.592 mmol) in  $\text{CH}_2\text{Cl}_2$  (10 mL) was degassed with an argon purge. Grubbs-II (50.0 mg, 0.0592 mmol) was added the reaction was degassed with ethylene and stirred under this atmosphere for 18 hours. Then, the reaction was concentrated *in vacuo* and purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 19:1) to give **20** (60.0 mg, 0.237 mmol, 40%) as a colourless oil.

$R_f = 0.30$  (petroleum ether/EtOAc, 8:1);  $^1\text{H}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  6.32 (1H, dd,  $J = 17.4, 10.9$  Hz, H5), 5.87 (1H, t,  $J = 5.5$  Hz, H7), 5.34 (1H, d,  $J = 17.4$  Hz, H6t), 5.05 (1H, d,  $J = 10.9$  Hz, H6c), 4.89 (1H, br s, NH), 4.29 (1H, dd,  $J = 15.0, 6.1$  Hz, H8a), 4.12 (1H, dd,  $J = 15.2, 4.9$  Hz, H8b), 3.91 (1H, d,  $J = 12.3$  Hz, H9a), 3.50 (1H, d,  $J = 12.3$  Hz, H9b), 3.19 (1H, d,  $J = 14.0$  Hz, H3a), 2.39 (1H, d,  $J = 15.0$  Hz, H3b), 1.39 (9H, s, H12), 1.37 (3H, s, H1);  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  154.9 (C10), 140.0 (C5), 139.5 (C4), 129.9 (C7), 113.8 (C6), 81.6 (C9), 78.9 (C11), 69.1 (C8), 52.8 (C2), 35.5 (C3), 28.4 (C12), 23.4 (C1), **HRMS** (ESI) calcd for  $[\text{C}_{14}\text{H}_{23}\text{NO}_3\text{Na}]^+$ : 276.1576, found 276.1585. **IR**  $\nu_{\text{max}} = 3423$  (w, N-H), 2975 (w, C-H), 1715 (s,  $\text{C}=\text{O}$ ), 1494 (m,  $\text{C}=\text{C}$ ), 1168 (s, C-O-C).

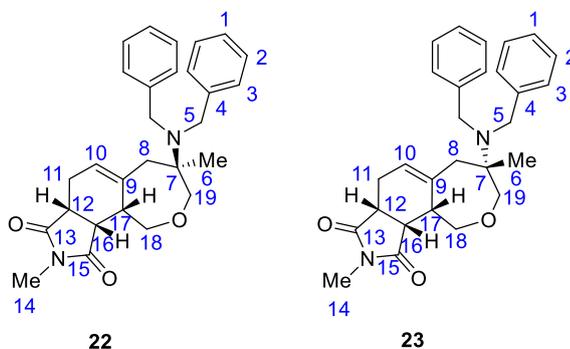
### 6.2.17 Synthesis of *N,N*-dibenzyl-3-methyl-5-vinyl-2,3,4,7-tetrahydrooxepin-3-amine (21)



To a solution of **19** (85.0 mg, 0.255 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (9 mL) previously degassed with an argon purge, was added Grubbs II catalyst (22.0 mg, 0.0259 mmol) and the reaction stirred under an atmosphere of ethylene for 18 hours. Upon completion, the reaction was concentrated *in vacuo* and the crude product purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 19:1) to give **21** (52.0 mg, 0.156 mmol, 61%) as a colourless oil.

*R*<sub>f</sub> = 0.29 (petroleum ether/EtOAc, 19:1); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.26 (4H, d, *J* = 7.0 Hz, H3), 7.20 – 7.16 (4H, m, H2), 7.12 – 7.08 (2H, m, H1), 6.36 (1H, dd, *J* = 17.3, 10.8 Hz, H9), 5.52 – 5.49 (1H, m, H11), 5.22 (1H, d, *J* = 17.4 Hz, H10t), 5.00 (1H, d, *J* = 10.7 Hz, H10c), 4.49 (1H, d, *J* = 17.5 Hz, H12a), 4.22 (1H, d, *J* = 17.6, 3.5 Hz, H12b), 4.05 (1H, d, *J* = 12.9 Hz, H13a), 3.92 (2H, d, *J* = 15.0 Hz, H5a), 3.85 (2H, d, *J* = 15.0 Hz, H5b), 3.46 (1H, d, *J* = 12.8 Hz, H13b), 2.97 (1H, d, *J* = 14.1 Hz, H7a), 2.44 (1H, d, *J* = 13.7 Hz, H7b) 1.04 (3H, s, H6); <sup>13</sup>C (101 MHz, CDCl<sub>3</sub>) δ 142.1 (C4), 140.8 (C9), 137.2 (C8), 129.4 (C11), 128.5 (C3), 128.0 (C2), 126.4 (C1), 111.5 (C10), 78.6 (C13), 71.6 (C12), 64.0 (C14), 54.5 (C5), 29.3 (C7), 23.9 (C6); HRMS (ESI) calcd for [C<sub>23</sub>H<sub>28</sub>NO]<sup>+</sup>: 334.2171, found 334.2175; IR *v*<sub>max</sub> = 2969 (w, C-H), 1604 (m, C=C), 1493 (m, C=C), 1453 (m, C=C), 1110 (C-O-C).

### 6.2.18 Synthesis of (4*R*)-4-(dibenzylamino)-4,9-dimethyl-1,3,4,5,7,7a,10a,10b-octahydro-8*H*-oxepino[3,4-*e*]isoindole-8,10(9*H*)-dione (22) and (4*S*)-4-(dibenzylamino)-4,9-dimethyl-1,3,4,5,7,7a,10a,10b-octahydro-8*H*-oxepino[3,4-*e*]isoindole-8,10(9*H*)-dione (23)



A solution of **21** (25.0 mg, 0.0755 mmol) and *N*-methyl maleimide (16.8 mg, 0.151 mmol) in toluene (0.75 mL) was heated to 110 °C for 1 hour 15 minutes. The reaction was concentrated

*in vacuo* and the crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 3:2) to give **22** (15.6 mg, 0.0351 mmol, 46%) and **23** (4.90 mg, 0.0110 mmol, 15%) both as colourless oils.

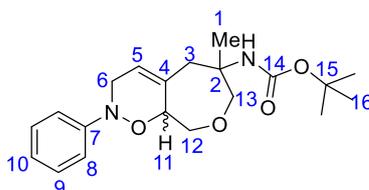
Data for **22**:

$R_f = 0.22$  (1:4, EtOAc/PE 40-60);  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.22 – 7.06 (10H, m, H1, H2 & H3), 5.52 – 5.51 (1H, m, H10), 4.37 – 4.33 (1H, m, H18a), 4.29 – 4.26 (1H, m, H19b), 4.23 (1H, dd,  $J = 12.9, 3.8$  Hz, H18a), 3.89 (2H, d,  $J = 14.7$  Hz, H5a) 3.72 (2H, d,  $J = 14.7$  Hz, H5b), 3.27 (1H, d,  $J = 13.2$  Hz, H19b), 3.20 – 3.16 (1H, m, H12), 3.11 (1H, dd,  $J = 8.7, 5.4$  Hz, H16), 2.93 – 2.91 (1H, m, H17), 2.90 (3H, s, H14), 2.69 (1H, dd,  $J = 15.1, 2.9$  Hz, H8a), 2.50 – 2.55 (1H, m, H11a), 2.50 (1H, d,  $J = 15.2$  Hz, H8b), 2.21 – 2.16 (1H, m, H11b), 0.81 (3H, s, H6);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  180.1 (C13) 178.5 (C15), 141.3 (C9), 140.9 (C4), 129.1 (C3), 127.8 (C2), 126.4 (C1), 123.3 (C10), 80.7 (C19), 72.6 (C18), 61.9 (C7), 52.4 (C5), 46.7 (C8), 43.0 (C16), 41.8 (C17), 40.7 (C12), 25.1 (C11), 25.0 (C14), 18.9 (C6); **HRMS** (ESI) calcd for  $[\text{C}_{28}\text{H}_{33}\text{N}_2\text{O}_3]^+$ : 445.2486, found 445.2474; **IR**  $\nu_{\text{max}} = 2938$  (m, C-H), 1690 (s, C=O), 1493 (m, C=C), 1131 (s, C-O-C).

Data for **23**:

$R_f = 0.10$  (petroleum ether/EtOAc, 4:1);  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.22 – 7.15 (8H, m, H2 & H3), 7.13 – 7.09 (2H, m, H1), 5.69 – 5.67 (1H, m, H10), 4.16 – 4.10 (2H, m, H18) 3.82 (2H, d,  $J = 14.9$  Hz, H5a), 3.75 (2H, d,  $J = 14.9$  Hz, H5b), 3.70 (1H, dd,  $J = 11.8, 2.1$  Hz, H19a), 3.64 (1H, d,  $J = 11.8$  Hz, H19b), 3.14 (1H, m, H12), 3.01 (1H, dd,  $J = 8.7, 5.2$  Hz, H16), 2.91 (3H, s, H14), 2.74 – 2.71 (1H, m, H8a) 2.70 – 2.67 (1H, m, H11a), 2.58 - 2.56 (1H, m, H17), 2.42 (1H, dd,  $J = 13.7, 2.0$  Hz, H8b), 2.18 – 2.13 (1H, m, H11b), 1.02 (3H, s, H6);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  180.0 (C13) 178.3 (C15), 141.7 (C4), 141.3 (C9), 128.3 (C3), 128.1 (C2), 126.6 (C1), 123.2 (C10), 81.0 (C19), 72.1 (C18), 61.6 (C7), 53.9 (C5), 42.7 (C16), 41.8 (C8), 41.7 (C17), 41.1 (C12), 25.2 (C11), 25.0 (C14), 22.2 (C6); **HRMS** (ESI) calcd for  $[\text{C}_{28}\text{H}_{33}\text{N}_2\text{O}_3]^+$ : 445.2486, found 445.2474; **IR**  $\nu_{\text{max}} = 3025$  (w, C $\equiv$ C-H), 2920 (w, C-H), 1693 (s, C=O), 1601 (w, C=C), 1124 (m, C-O-C).

### 6.2.19 Synthesis of *tert*-butyl (6-methyl-2-phenyl-3,5,6,7,9,9a-hexahydro-2H-oxepino[4,3-*e*][1,2]oxazin-6-yl)carbamate (**24**)



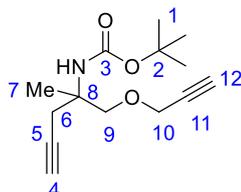
To a solution of **20** (60.0 mg, 0.237 mmol) in toluene (2.5 mL) was added nitrosobenzene (51.0 mg, 0.476 mmol) and the reaction heated to 110 °C for 4 hours. Then, the reaction was

concentrated under reduced pressure to give a mixture of diastereomers (dr ca. 46:36:10:0 determined by  $^1\text{H}$  NMR). The major isomer was isolated by flash column chromatography (silica gel; petroleum ether/EtOAc 22:3) to give **24** (26.0 mg, 0.0721 mmol, 31%) as a yellow oil.

Data for isolated diastereomer **24**:

$R_f$  = 0.15 (petroleum ether/EtOAc 22:3);  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.30 – 7.25 (2H, m, H9), 7.05 – 7.02 (2H, m, H8), 6.96 – 9.92 (1H, t,  $J$  = 7.3 Hz, H10), 5.71 (1H, d,  $J$  = 2.2 Hz, H5), 4.77 (1H, br s, NH), 4.55 – 4.51 (1H, m, H6a), 4.31 – 4.26 (2H, m, H6b & H11), 4.02 – 3.97 (2H, m, H12a & H13a), 3.63 (1H, dd,  $J$  = 11.4, 9.5 Hz, H12b), 3.28 (1H, d,  $J$  = 12.7 Hz, H13b), 2.74 – 2.65 (2H, m, H3), 1.44 (9H, s, H16), 1.30 (3H, s, H1);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  154.7 (C14), 147.6 (C7), 133.1 (C4), 129.0 (C9), 123.2 (C5), 121.8 (C10), 115.7 (C8), 80.9 (C15), 78.2 (C13), 72.0 (C12), 67.1 (C6), 61.8 (C11), 55.1 (C2), 45.7 (C3), 28.4 (C16), 22.3 (C1); HRMS (ESI) calcd for  $[\text{C}_{20}\text{H}_{28}\text{N}_2\text{O}_4\text{Na}]^+$ : 383.1941, found 383.1931. IR  $\nu_{\text{max}}$  = 3360 (w, N-H), 2977 (w, C-H), 2934 (w, C-H), 1711 (s, C=O), 1596 (m, C=C), 1491 (s, C=C).

### 6.2.20 Synthesis of *tert*-butyl (2-methyl-1-(prop-2-yn-1-yloxy)pent-4-yn-2-yl)carbamate (**26**)

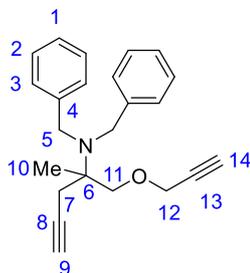


To an ice-cooled solution of alcohol **7** (40.0 mg, 0.188 mmol) in DMF (2 mL) was added NaH (8.20 mg, 0.206 mmol) and the reaction stirred for 15 minutes before propargyl bromide (44.7  $\mu\text{L}$ , 0.376 mmol) was added and the reaction stirred for 2 hours at r.t. Further NaH (1.00 mg, 0.0250 mmol) was added and the reaction stirred for additional 2 hours. Upon completion the reaction was diluted with  $\text{H}_2\text{O}$  and extracted with EtOAc (2 x 30 mL). The combined organic extracts was dried over  $\text{Na}_2\text{SO}_4$ , filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 3:2) to give the **26** (25.0 mg, 0.0995 mmol, 52%) as a colourless oil.

$R_f$  = 0.48 (petroleum ether/EtOAc, 9:1);  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  4.81 (1H, br s, NH), 4.21 – 4.13 (2H, m, H10), 3.67 (1H, d,  $J$  = 9.0 Hz, H9a), 3.55 (1H, d,  $J$  = 9.0 Hz, H9b), 2.79 (1H, dd,  $J$  = 16.7, 2.0 Hz, H6a), 2.56 (1H, dd,  $J$  = 16.7, 2.6 Hz, H6b), 2.44 (1H, t,  $J$  = 2.4 Hz, H12), 1.09 (1H, t,  $J$  = 2.7 Hz, H4), 1.43 (9H, s, H1) 1.39 (3H, s, H7);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  154.8 (C3), 80.6 (C5), 79.6 (C11), 79.4 (C2), 74.8 (C12), 73.9 (C9), 70.8 (C4), 58.7 (C10), 54.6 (C8), 28.5

(C1), 26.5 (C6), 21.6 (C7); **HRMS** (ESI) calcd for  $[C_{14}H_{22}NO_3]^+$ : 252.1600, found 252.1603; **IR**  $\nu_{\max}$  = 3291 (w, C≡C-H), 2977 (w, C-H), 2122 (w, C≡C), 1697 (s, C=O).

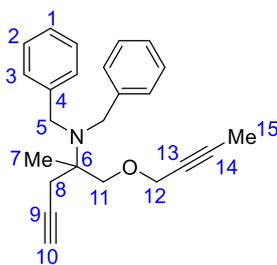
### 6.2.21 Synthesis of *N,N*-dibenzyl-2-methyl-1-(prop-2-yn-1-yloxy)pent-4-yn-2-amine (27)



To an ice-cooled solution of propargyl bromide (220  $\mu$ L, 2.05 mmol) in DMF (7 mL) was added NaH (60.0 mg, 1.50 mmol), followed by a solution of **18** (200 mg, 0.682 mmol) in DMF (1 mL) and the reaction warmed to r.t. over 2 hours. The reaction was quenched with  $NH_4Cl$  saturated aqueous solution and the product extracted with  $CH_2Cl_2$  (3 x 30 mL). The combined organic extracts were concentrated, washed with brine (2 x 20 mL), dried over  $Na_2SO_4$ , filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 4:1) to give **27** (200 mg, 0.604 mmol, 88%) as an orange oil.

$R_f$  = 0.64 (petroleum ether/EtOAc, 9:1);  $^1H$  NMR (400 MHz,  $CDCl_3$ )  $\delta$  7.33 – 7.31 (4H, m, H3), 7.21 – 7.17 (4H, m, H2), 7.13 – 7.09 (2H, m, Hz, H1), 4.08 (2H, d,  $J$  = 2.3 Hz, H12), 3.91 (4H, s, H5), 3.63 (1H, d,  $J$  = 9.4 Hz, H11a), 3.58 (1H, d,  $J$  = 9.4 Hz, H11b), 2.59 (1H, dd,  $J$  = 16.9, 2.7 Hz, H7a), 2.52 (1H, dd,  $J$  = 16.9, 2.7 Hz, H7b), 2.41 (1H, t,  $J$  = 2.5 Hz, H14), 2.08 (1H, t,  $J$  = 2.8 Hz, H9), 1.26 (s, 3H, H10);  $^{13}C$  (101 MHz,  $CDCl_3$ )  $\delta$  142.1 (C4), 128.5 (C3), 127.9 (C2), 126.4 (C1), 82.3 (C8), 80.0 (C13), 74.7 (C11), 74.4 (C14), 70.7 (C9), 60.8 (C6), 58.6 (C12), 54.3 (C5), 26.4 (C7), 21.2 (C10); **HRMS** (ESI) calcd for  $[C_{23}H_{26}NO]^+$ : 332.2009, found 332.2002; **IR**  $\nu_{\max}$  = 3294 (m, C≡C-H), 2949 (m, C-H), 2120 (w, C≡C), 1494 (m, C=C), 1453 (s, C=C), 1114 (s, C-O-C).

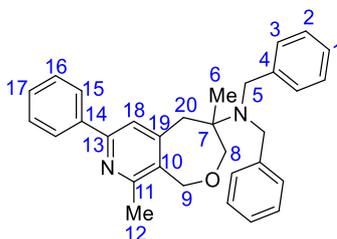
### 6.2.22 Synthesis of *N,N*-dibenzyl-1-(but-2-yn-1-yloxy)-2-methylpent-4-yn-2-amine (32)



A solution of NaH (16.0 mg, 0.405 mmol) and 1-bromo-2-butyne (48.0 uL, 0.552 mmol) in DMF (0.8 mL) was cooled to 0 °C before a solution of **18** (54.0 mg, 0.184 mmol) in DMF (1 mL) was added dropwise. The reaction was warmed to r.t., stirring for 2.5 hours. The reaction was quenched with NH<sub>4</sub>Cl saturated aqueous solution and extracted with CH<sub>2</sub>Cl<sub>2</sub> (3 x 20 mL). The combined organic extracts were dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 85:15) to give **32** (43.5 mg, 0.126 mmol, 71%) as a colourless oil.

$R_f$  = 0.61 (petroleum ether/EtOAc, 19:1); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.31 (4H, d,  $J$  = 7.2 Hz, H3), 7.18 (4H, t,  $J$  = 4.4 Hz, H2), 7.11 – 7.08 (2H, m, H1), 4.04 (2H, m, H12), 3.91 (4H, s, H5), 3.60 (1H, d,  $J$  = 9.5 Hz, H11a), 3.54 (1H, d,  $J$  = 9.5 Hz, H11b), 2.60 – 2.49 (2H, m, H8), 2.06 (1H, t,  $J$  = 2.7 Hz, H10), 1.84 (3H, t,  $J$  = 2.3 Hz, H15), 1.25 (3H, s, H7); <sup>13</sup>C (101 MHz, CDCl<sub>3</sub>) δ 142.2 (C4), 128.5 (C3), 127.9 (C2), 126.3 (C1), 82.5 (C9), 82.3 (C13), 75.5 (C14), 74.5 (C11), 70.6 (C10), 60.9 (C6), 59.1 (C12), 54.3 (C5), 26.5 (C8), 21.3 (C7), 3.8 (C15); HRMS (ESI) calcd for [C<sub>24</sub>H<sub>28</sub>NO]<sup>+</sup>: 346.2171, found 346.2182; IR  $\nu_{max}$  = 3281 (s, C≡C-H), 2199 (w, C≡C), 1600 (w, C=C), 1492 (s, C=C) 1089 (s, C-O-C).

### 6.2.23 Synthesis of *N,N*-dibenzyl-4,9-dimethyl-7-phenyl-1,3,4,5-tetrahydrooxepino[3,4-*c*]pyridin-4-amine (**34**)

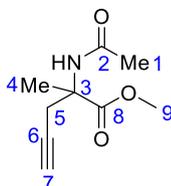


CpCo(CO)<sub>2</sub> (96.0 uL, 0.540 mmol) and benzonitrile (104 uL, 1.54 mmol) and) were added to a solution of **32** (80.0 mg, 0.231 mmol) in toluene (3.8 mL) previously degassed with an argon purge, and then heated to 110 °C for 48 hours in a sealed tube. The reaction was cooled to r.t. and the solvents removed *in vacuo*. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 9:1) **34** (22.0 mg, 0.0490 mmol, 21%) as a colourless oil.

$R_f$  = 0.28 (petroleum ether/EtOAc, 9:1); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.97 – 7.95 (2H, m, H15), 7.47 – 7.43 (2H, m, H16), 7.39 (1H, m, H17), 7.35 (1H, s, H18), 7.23 (4H, m, H3), 7.20 – 7.16 (4H, m, H2), 7.13 – 7.09 (2H, m, H1), 4.92 (1H, d,  $J$  = 14.8 Hz, H9a), 4.78 (1H, d,  $J$  = 14.9 Hz, H9b), 4.10 (1H, d,  $J$  = 12.7 Hz, H8a), 3.96 (2H, d,  $J$  = 14.8 Hz, H5a), 3.87 (2H, d,  $J$  = 14.9 Hz, H5b), 3.64 (1H, d,  $J$  = 12.5 Hz, H8b), 3.55 (1H, d,  $J$  = 13.6 Hz, H20a), 2.89 (1H, d,  $J$  = 13.4 Hz, H20b), 2.57 (3H, s, H12), 1.03 (3H, s, H6); <sup>13</sup>C (101 MHz, CDCl<sub>3</sub>) δ 155.6 (C13) 154.8 (C11) 148.0 (C19), 141.6 (C4), 139.5 (C14), 130.3 (C10), 128.8 (C16/C17), 128.8 (C16/C17), 128.4

(C3), 128.1 (C2), 127.0 (C15), 126.6 (C1), 121.3 (C18), 79.8 (C8), 70.9 (C9), 61.8 (C7), 54.1 (C5), 41.5 (C20), 22.6 (C12), 22.4 (C6); **HRMS** (ESI) calcd for  $[C_{31}H_{33}N_2O]^+$ : 449.2593, found 449.2581; **IR**  $\nu_{\max}$  = 3061 (w, C-H), 2848 (s, C-H), 1701 (m, C=N, Ar), 1589 (s, C=C), 1558 (s, C=C) 1114 (s, C-O-C).

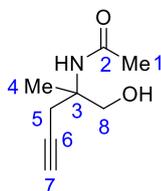
#### 6.2.24 Synthesis of methyl 2-acetamido-2-methylpent-4-ynoate (**37**)



To a solution of **4** (257 mg, 1.82 mmol) in  $CH_2Cl_2$  (18 mL) at 0 °C was added triethylamine (0.380 mL, 2.70 mmol) and acetyl chloride (0.193 mL, 2.70 mmol) and the reaction warmed to r.t. and stirred for 6 hours. Upon completion,  $NH_4Cl$  saturated aqueous solution was added and the reaction extracted with EtOAc (3 x 50 mL). The combined organic extracts were washed with brine, dried over  $Na_2SO_4$  and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 2:3) to give **37** (322 mg, 1.76 mmol, 97%) as a white solid.

$R_f$  = 0.19 (petroleum ether/EtOAc, 7:3); **m.p.** 98.5 – 99.4 °C;  $^1H$  NMR (400MHz,  $CDCl_3$ )  $\delta$  6.22 (1H, bs, NH), 3.77 (3H, s, H9), 3.05 (1H, dd,  $J$  = 16.9, 2.6 Hz, H5a), 2.92 (1H, dd,  $J$  = 16.9, 2.6 Hz, H5b), 2.00 (3H, s, H1), 1.99 (1H, t,  $J$  = 2.6 Hz, H7), 1.59 (3H, s, H4);  $^{13}C$  NMR (101MHz,  $CDCl_3$ )  $\delta$  173.6 (C8), 169.7 (C2), 79.7 (C6), 71.1 (C7), 58.8 (C3), 53.1 (C9), 26.4 (C5), 23.7 (C1), 23.0 (C4); **HRMS (ESI)** calcd for  $[C_9H_{14}NO_3]^+$ : 184.0968, found 184.0968; **IR**  $\nu_{\max}$  = 3385 (m, N-H), 3233 (m,  $C\equiv C-H$ ), 2179 (w,  $C\equiv C$ ), 1722 (s, C=O ester), 1663 (s, C=O amide), 1249 (s, C-O-C).

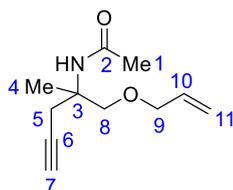
#### 6.2.25 Synthesis of *N*-(1-hydroxy-2-methylpent-4-yn-2-yl)acetamide (**38**)



To a solution of **37** (810 mg, 4.42 mmol) in THF (44 mL) at 0 °C was  $LiBH_4$  (192 mg, 8.81 mmol) and the reaction stirred for 16 hours. Upon completion, the reaction was quenched with water and extracted with EtOAc (3 x 50 mL). The combined organic extracts were washed with brine, dried over  $Na_2SO_4$  and concentrated *in vacuo* to give **38** (568 mg, 3.66 mmol, 83%) as a white solid.

$R_f = 0.22$  (petroleum ether/EtOAc, 1:4); **m.p.** 107.6 – 108.8 °C;  **$^1\text{H NMR}$**  (400MHz,  $\text{CDCl}_3$ )  $\delta$  5.74 (1H, bs, NH), 4.68 (1H, t,  $J = 5.9$  Hz, OH), 3.71 – 3.63 (2H, m, H8), 2.80 (1H, dd,  $J = 16.9, 2.6$  Hz, H5a), 2.43 (1H, dd,  $J = 16.9, 2.6$  Hz, H5b), 2.07 (1H, t,  $J = 2.7$  Hz, H7), 2.02 (3H, s, H1), 1.33 (3H, s, H4);  **$^{13}\text{C NMR}$**  (101MHz,  $\text{CDCl}_3$ )  $\delta$  171.5 (C2), 80.1 (C6), 71.6 (C7), 69.2 (C8), 57.8 (C3), 26.8 (C5), 24.2 (C1), 22.8 (C4); **HRMS (ESI)** calcd for  $[\text{C}_8\text{H}_{14}\text{NO}_2]^+$ : 156.1019, found 156.1019. **IR**  $\nu_{\text{max}} = 3308$  (m,  $\text{C}\equiv\text{C-H}$ ), 3246 (br m, O-H), 2876 (m, C-H), 1642 (s,  $\text{C}=\text{O}$ ), 1378 (s, C-O), 1052 (s, C-O).

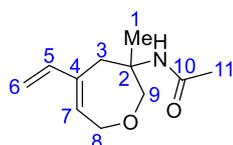
### 6.2.26 Synthesis of *N*-(1-(allyloxy)-2-methylpent-4-yn-2-yl)acetamide (**39**)



To a solution of **38** (250 mg, 1.61 mmol) in DMF (17.8 mL) at 0 °C was added NaH (70.0 mg, 1.77 mmol) and the reaction stirred for 15 minutes before allyl bromide (153  $\mu\text{L}$ , 1.77 mmol) was added. The reaction was stirred for 3 hours before  $\text{NH}_4\text{Cl}$  saturated aqueous solution was added and the reaction mixture was extracted with EtOAc (3 x 20 mL). The organic extracts were combined, washed with brine and dried over  $\text{Na}_2\text{SO}_4$ . The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 3:2) to give **39** (276 mg, 1.41 mmol, 88%) as a colourless oil.

$R_f = 0.25$  (petroleum ether/EtOAc, 3:2);  **$^1\text{H NMR}$**  (400MHz,  $\text{CDCl}_3$ )  $\delta$  5.94 – 5.84 (1H, m, H10), 5.67 (1H, bs, NH), 5.29 – 5.17 (1H, m, H11), 4.01 (2H, d,  $J = 5.6$  Hz, H9), 3.63 (1H, d,  $J = 9.3, \text{H8a}$ ), 3.44 (1H, d,  $J = 9.2$  Hz, H8b), 2.89 (1H, dd,  $J = 16.7, 2.7$  Hz, H5a), 2.62 (1H, dd,  $J = 16.7, 2.7$  Hz, H5b), 1.98 (1H, t,  $J = 2.7$  Hz, H7), 1.94 (3H, s, H1), 1.43 (3H, s, H4);  **$^{13}\text{C NMR}$**  (101MHz,  $\text{CDCl}_3$ )  $\delta$  170.2 (C2), 134.6 (C10), 117.4 (C11), 80.7 (C6), 73.9 (C8), 72.4 (C9), 70.7 (C7), 55.7 (C3), 26.1 (C5), 24.5 (C1), 21.3 (C4); **HRMS (ESI)** calcd for  $[\text{C}_{11}\text{H}_{18}\text{NO}_2]^+$ : 196.1329, found 196.1332; **IR**  $\nu_{\text{max}} = 3281$  (m,  $\text{C}\equiv\text{C-H}$ ), 3079 (w, C-H), 2929 (w, C-H), 2351 (w,  $\text{C}\equiv\text{C}$ ), 1653 (s,  $\text{C}=\text{C}$ ), 1543 s, ( $\text{C}=\text{C}$ ).

### 6.2.27 Synthesis of *N*-(3-methyl-5-vinyl-2,3,4,7-tetrahydrooxepin-3-yl)acetamide (**40**)

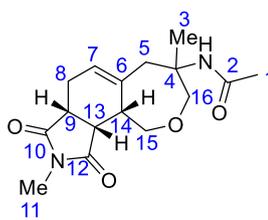


A solution of **39** (240 mg, 1.18 mmol) in  $\text{CH}_2\text{Cl}_2$  (22.8 mL) was degassed with an argon purge for 15 minutes. Grubbs-II (100 mg, 0.118 mmol) was added and the reaction was degassed

with ethylene and stirred under this atmosphere for 18 hours. Then, the reaction was concentrated *in vacuo* and purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 1:1) to give **40** (120 mg, 0.614 mmol, 52%).

$R_f$  = 0.22 (petroleum ether/EtOAc, 2:3);  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  6.31 (1H, dd,  $J$  = 17.4, 10.8 Hz, H5), 5.87 (1H, t,  $J$  = 5.2 Hz, H7), 5.66 (1H, br s, NH), 5.33 (1H, d,  $J$  = 17.4 Hz, H6t), 5.07 (1H, d,  $J$  = 10.8 Hz, H6c), 4.30 (1H, dd,  $J$  = 15.2, 6.2 Hz, H8a), 4.13 (1H, dd,  $J$  = 15.2, 4.9 Hz, H8b), 3.97 (1H, d,  $J$  = 12.3 Hz, H9a), 3.50 (1H, d,  $J$  = 12.4 Hz, H9b), 3.28 (1H, d,  $J$  = 15.0 Hz, H3a), 2.41 (1H, d,  $J$  = 15.0 Hz, H3b), 1.87 (3H, s, H11), 1.44 (3H, s, H1);  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  170.2 (C10), 140.0 (C5), 139.5 (C4), 129.7 (C7), 114.1 (C6), 81.4 (C9), 69.1 (C8), 53.8 (C2), 35.6 (C3), 24.3 (C11), 23.2 (C1); **HRMS** (ESI) calcd for  $[\text{C}_{11}\text{H}_{17}\text{NO}_2\text{Na}]^+$ : 218.1152, found 218.1146; **IR**  $\nu_{\text{max}}$  = 3315 (m, N-H), 2935 (m, C-H), 1655 (s, C=O), 1650 (s, C=C), 1541 (s, C=C), 1128 (m, C-O-C).

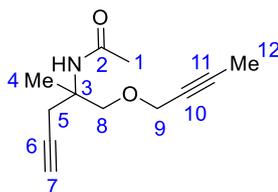
### 6.2.28 Synthesis of *N*-(4,9-dimethyl-8,10-dioxo-3,4,5,7,7a,8,9,10,10a,10b-decahydro-1*H*-oxepino[3,4-*e*]isoindol-4-yl)acetamide (**41**)



To a solution of **40** (45.0 mg, 0.230 mmol) in toluene (2.3 mL) was added *N*-methylmaleimide (51.0 mg, 4.62 mmol) and the reaction heated to 110 °C for 5 hours. Then, the reaction was cooled to r.t. and concentrated *in vacuo* to give a mixture of diastereomers (dr = ca. 85:15, determined by  $^1\text{H}$  NMR). The major diastereomer was isolated by flash column chromatography (silica gel, EtOAc) to give **41** (49.0 mg, 0.160 mmol, 70%) as a white solid.

$R_f$  = 0.12 (EtOAc); **m.p.** 134 – 137 °C;  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  5.73 – 5.71 (1H, m, H7), 5.50 (1H, s, NH), 4.26 – 4.16 (2H, m, H15), 3.85 (1H, dd,  $J$  = 12.5, 2.2 Hz, H16a), 3.27 (1H, d,  $J$  = 12.5 Hz, H16b), 3.16 – 3.12 (1H, m, H14), 3.05 – 3.01 (2H, m, H5a & H9), 2.89 (3H, s, H11), 2.69 (1H, ddd,  $J$  = 15.3, 7.3, 0.93 Hz, H8a), 2.60 – 2.56 (1H, m, H13), 2.35 (1H, d,  $J$  = 14.4 Hz, H5b), 2.16 – 2.08 (1H, m, H8b), 1.83 (3H, s, H1), 1.31 (3H, s, H3);  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  179.5 (C10), 178.0 (C12), 169.8 (C2), 139.6 (C6), 125.4 (C7), 81.4 (C16), 72.6 (C15), 56.7 (C4), 43.8 (C5), 42.7 (C9), 41.5 (C13), 40.6 (C14), 25.0 (C8), 24.9 (C11), 24.0 (C1), 22.0 (C3); **HRMS** (ESI) calcd for  $[\text{C}_{16}\text{H}_{23}\text{N}_2\text{O}_4]^+$ : 307.1658, found 307.1663; **IR**  $\nu_{\text{max}}$  = 3329 (w, N-H), 2930 (w, C-H), 1693 (s, C=O), 1664 (s, C=O), 1535 (m, C=C).

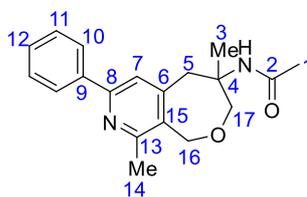
### 6.2.29 Synthesis of *N*-(1-(but-2-yn-1-yloxy)-2-methylpent-4-yn-2-yl)acetamide (42)



A solution of **38** (150 mg, 0.965 mmol) in DMF (9.6 mL) was cooled to 0 °C before NaH (42.0 mg, 1.06 mmol) was added. The reaction was stirred for 20 minutes before 1-bromo-2-butyne (93.0 uL, 1.06 mmol) was added and the reaction was stirred for 3 hours. Upon completion NH<sub>4</sub>Cl saturated aqueous solution was added and the reaction mixture was extracted with EtOAc (3 x 20 mL). The combined organic extracts were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 3:2) to give **42** (164 mg, 0.791 mmol, 82%) as a colourless oil.

$R_f$  = 0.22 (petroleum ether/EtOAc, 3:2); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 5.67 (1H, bs, NH), 4.13 (2H, m, H9), 3.89 (1H, d, *J* = 9.2 Hz, H8a), 3.51 (1H, d, *J* = 9.2 Hz, H8b), 2.91 (1H, dd, *J* = 16.8, 2.7 Hz, H5a), 2.61 (1H, dd, *J* = 16.7, 2.6 Hz, H5b), 1.99 (1H, t, *J* = 2.7 Hz, H7), 1.94 (3H, s, H1), 1.86 (3H, t, *J* = 2.3 Hz, H12), 1.44 (3H, s, H4); <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 170.2 (C2), 83.0 (C6), 80.7 (C10) 75.0 (C11) 73.6 (C8), 70.7 (C7), 59.4 (C9), 55.6 (C3), 26.0 (C5) 24.5 (C1), 21.2 (C4) 3.8 (C12); HRMS (ESI) calcd for [C<sub>12</sub>H<sub>17</sub>NO<sub>2</sub>Na]<sup>+</sup>: 230.1152, found 230.1146; IR  $\nu_{\max}$  = 3292 (m, C≡C-H) 2922 (w, C-H), 2364 (w, C≡C-R), 2242 (w, C≡C), 1650 (s, C=O), 1090 (s, C-O).

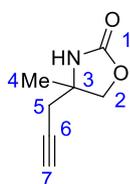
### 6.2.30 Synthesis of *N*-(4,9-dimethyl-7-phenyl-1,3,4,5-tetrahydrooxepino[3,4-*c*]pyridin-4-yl)acetamide (43)



CpCo(CO)<sub>2</sub> (96.0 uL, 0.540 mmol) and benzonitrile (104 uL, 1.54 mmol) and) were added to a solution of **42** (80.0 mg, 0.386 mmol) in toluene (3.8 mL) previously degassed with an argon purge, and the heated to 110 °C for 48 hours in a sealed tube. Then, the reaction was cooled to r.t. and the solvents removed *in vacuo*. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 7:3) **43** (16.0 mg, 0.0515 mmol, 22%, based on 39% recovered unreacted starting material) as a white solid.

$R_f = 0.12$  (petroleum ether/EtOAc, 1:1); **m.p.** 147 – 151 °C;  $^1\text{H NMR}$  (500 MHz,  $\text{CDCl}_3$ )  $\delta$  7.98 – 7.96 (2H, m, H10), 7.47 – 7.42 (3H, m, H7 & H11), 7.38 (1H, m, H12), 5.49 (1H, br s, NH), 5.13 (1H, d,  $J = 14.0$  Hz, H16a), 4.55 (1H, d,  $J = 14.0$  Hz, H16b), 3.98 (1H, dd,  $J = 12.3, 2.0$  Hz, H17a), 3.79 (1H, dd,  $J = 14.3, 2.0$  Hz, H5a), 3.64 (1H, d,  $J = 12.3$  Hz, H17b), 3.04 (1H, d,  $J = 14.2$  Hz, H5b), 2.67 (3H, s, H14), 1.65 (3H, s, H1), 1.46 (3H, s, H3);  $^{13}\text{C NMR}$  (126 MHz,  $\text{CDCl}_3$ )  $\delta$  170.2 (C2), 156.2 (C8), 154.8 (C13), 148.1 (C6), 139.1 (C9), 130.4 (C15), 128.9 (C12), 128.7 (C11), 126.9 (C10), 121.0 (C7), 83.0 (C17), 69.8 (C16), 54.0 (C4), 43.6 (C5), 23.9 (C1), 22.9 (C14), 22.8 (C3); **HRMS** (ESI) calcd for  $[\text{C}_{19}\text{H}_{23}\text{N}_2\text{O}_2]^+$ : 311.1760, found 311.1764. **IR**  $\nu_{\text{max}}$  = 3295 (m, N-H), 3063 (w, C-H), 2919 (m, C-H), 1649 (s, C=O), 1591 (m, C=C), 1539 (s, C=C), 1441 (s, C=C), 1087 (s, C-O-C).

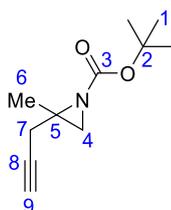
### 6.2.31 Synthesis of 4-methyl-4-(prop-2-yn-1-yl)oxazolidin-2-one (45)



To an ice-cooled stirred solution of **7** (33.0 mg, 0.155 mmol) in THF (1.60 mL) was added *t*-BuOK (35.0 mg, 0.312 mmol) and the reaction was allowed to warm to r.t. for 6 hours. The reaction was diluted with  $\text{NH}_4\text{Cl}$  saturated aqueous solution and extracted with  $\text{CH}_2\text{Cl}_2$  (3 x 20 mL). The combined organic extracts were dried over  $\text{Na}_2\text{SO}_4$ , filtered and concentrated *in vacuo* to give **45** (14.7 mg, 0.106 mmol, 67 %) as a white solid without further purification.

$R_f = 0.26$  (petroleum ether/EtOAc, 1:1); **m.p.** 83.2 – 85.1 °C;  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  5.92 (1H, br s, NH), 4.28 (1H, d,  $J = 8.7$  Hz, H2a), 4.10 (1H, d,  $J = 8.7$  Hz, H2b), 2.53 – 2.44 (2H, m, H5), 2.09 (1H, t,  $J = 2.6$  Hz, H7), 1.46 (3H, s, H4);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  158.7 (C1), 78.7 (C6), 75.0 (C2), 72.0 (C7), 57.2 (C3), 31.1 (C5), 25.2 (C4); **HRMS** (ESI) calcd for  $[\text{C}_7\text{H}_{10}\text{NO}_2]^+$ : 140.0706, found 140.0708; **IR**  $\nu_{\text{max}}$  = 3235 (s,  $\text{C}\equiv\text{C-H}$ ), 2891 (w, C-H), 2030 (w,  $\text{C}\equiv\text{C}$ ), 1729 (s, C=O).

### 6.2.32 Synthesis of *tert*-butyl 2-methyl-2-(prop-2-yn-1-yl)aziridine-1-carboxylate (46)

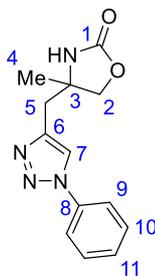


To a solution of **7** (500 mg, 2.34 mmol) in  $\text{Et}_2\text{O}$  (40 mL), TsCl (536 mg, 2.81 mmol) and powdered KOH (789 mg, 14.1 mmol) were added and the reaction refluxed for 18 hrs. Further

KOH (263 mg, 4.69 mmol) was added and the reaction stirred for additional 3 hours. The reaction was cooled to r.t and diluted with water. The aqueous solution was extracted with Et<sub>2</sub>O (2 x 60 mL) and the combined organic extracts dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 4:1) to give the **46** (357 mg, 1.83 mmol, 78%) as a colourless oil.

**R<sub>f</sub>** = 0.55 (petroleum ether/EtOAc, 9:1); **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 2.54 (1H, dd, *J* = 17.1, 2.6 Hz, H7a), 2.35 (1H, dd, *J* = 17.1, 2.7 Hz, H7b), 2.23 (1H, s, H4a), 2.10 (1H, s, H4b), 2.06 (1H, t, *J* = 2.8 Hz, H9), 1.46 (9H, s, H1), 1.37 (3H, s, H6); **<sup>13</sup>C** (101 MHz, CDCl<sub>3</sub>) δ 160.7 (C3), 81.2 (C2), 79.8 (C8), 71.2 (C9), 41.3 (C5), 36.5 (C4), 28.1 (C1), 27.2 (C7), 19.8 (C6); **HRMS** (ESI) calcd for [C<sub>11</sub>H<sub>18</sub>NO<sub>2</sub>]<sup>+</sup>: 196.1332, found 196.1333; **IR** *v*<sub>max</sub> = 3290 (m, C≡C-H), 2975 (w, C-H), 2160 (m, C≡C), 1711 (s, C=O), 1159 (m, C-O).

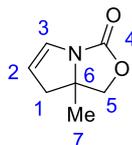
### 6.2.33 Synthesis of 4-methyl-4-((1-phenyl-1*H*-1,2,3-triazol-4-yl)methyl)oxazolidin-2-one (**47**)



To a solution of **45** (20.0 mg, 0.140 mmol) in H<sub>2</sub>O (0.5 mL) and *t*-BuOH (0.5 mL) was added PhN<sub>3</sub> (0.5M in *tert*-butyl methyl ether) (0.431 mL, 0.215 mmol) followed by CuSO<sub>4</sub> (3.60 mg, 0.0420 mmol). The reaction was stirred at r.t. for 24 hours. Upon completion the reaction was concentrated *in vacuo*. The sample was taken up in CH<sub>2</sub>Cl<sub>2</sub> (10 mL) and washed with 5% NH<sub>4</sub>OH aqueous solution (10 mL), dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel; EtOAc) to give **47** (26.0 mg, 0.100 mmol, 72%) as a white solid.

**R<sub>f</sub>** = 0.12 (petroleum ether/EtOAc, 1:1); **m.p.** 147.3 – 146.7 °C; **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.99 (1H, s, H7), 7.71 – 7.67 (2H, m, H9), 7.48 – 7.44 (2H, m, H10), 7.40 – 7.36 (1H, m, H11), 6.97 (1H, s, NH), 4.40 (1H, d, *J* = 8.6 Hz, H2a), 4.11 (1H, d, *J* = 8.6 Hz, H2b), 3.10 – 3.02 (2H, m, H5), 1.41 (3H, s, H4); **<sup>13</sup>C** (101 MHz, CDCl<sub>3</sub>) δ 159.3 (C1), 142.9 (C6), 136.9 (C8), 129.7 (C10), 128.7 (C11), 121.1 (C7), 120.4 (C9), 75.2 (C2), 57.7 (C3), 36.3 (C5), 25.7 (C4); **HRMS** (ESI) calcd for [C<sub>13</sub>H<sub>15</sub>N<sub>4</sub>O<sub>2</sub>]<sup>+</sup>: 259.1190, found 259.1196; **IR** *v*<sub>max</sub> = 2918 (m, C-H), 1737 (s, C=O), 1607 (m, C=C) 1453 (m, C=C), 1045 (C-O-C).

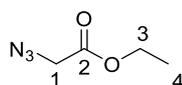
### 6.2.34 Synthesis of 7a-methyl-7,7a-dihydro-1H,3H-pyrrolo[1,2-c]oxazol-3-one (48)



InCl<sub>3</sub> (62.0 mg, 0.280 mmol) was introduced into 5 mL flask and heated with a heat gun (150 °C) under vacuum for 2 minutes. After being allowed to cool to r.t., THF (1 mL) was added under Argon. The mixture was stirred at r.t. for 10 minutes and cooled to -78 °C. DIBAL-H (1.00 M in hexane, 0.279 mL, 0.279 mmol) was added dropwise and the mixture was stirred at -78 °C for 40 minutes. Then, **45** (25.0 mg, 0.179 mmol) was then added, followed by BEt<sub>3</sub> (1.0 M in THF, 0.0900 mL, 0.0900 mmol) and the mixture was stirred at -78 °C for 3 hours. A solution of iodine (274 mg, 1.08 mmol) in THF (0.5 mL) was then added. After 40 minutes, the mixture was poured onto a NaHCO<sub>3</sub> saturated aqueous solution (5 mL). Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> was added under stirring until complete decoloration and the aqueous layer was extracted with EtOAc (5 x 10 mL). The combined organic extracts were dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated *in vacuo*. The crude product was dissolved in toluene (1.8 mL) and CuI (15.0 mg, 0.0788 mmol), Cs<sub>2</sub>CO<sub>3</sub> (70.0 mg, 0.214 mmol) and *N,N'*-dimethylethane-1,2-diamine (17.0 μL, 0.158 mmol) were added. The mixture was stirred at 85 °C for 2 hours and cooled to r.t., diluted with DCM (5 mL) and the organics washed with water (5 mL). The aqueous layer was extracted with DCM (3 x 5 mL) and the combined organic extracts were dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 2:1) to yield **48** (7.00 mg, 0.0503 mmol, 28%) as a colourless oil.

R<sub>f</sub> = 0.14 (petroleum ether/EtOAc, 85:15); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 6.38 – 6.35 (1H, m, H3), 5.23 – 5.21 (1H, m, H2), 4.28 (1H, d, *J* = 8.6 Hz, H5a) 4.20 (1H, d, *J* = 8.6 Hz, H5b), 2.71 (1H, dt, *J* = 16.6, 2.6 Hz, H1a), 2.37 (1H, qd, *J* = 11.5, 2.8 Hz, H1b), 1.39 (3H, s, H7); <sup>13</sup>C (101 MHz, CDCl<sub>3</sub>) δ 157.2 (C4), 127.2 (C1), 112.7 (C2), 77.4 (C5), 65.0 (C6), 43.0 (C3), 26.8 (C7); HRMS (ESI) [C<sub>7</sub>H<sub>10</sub>NO<sub>2</sub>]<sup>+</sup>: 140.0706, found 140.0712; IR ν<sub>max</sub> = 3133 (m, C-H), 1757 (s, C=O), 1594 (m, C=C), 1501 (s, C=C), 1041 (s, C-O-C) 762 (s, C-H, cis).

### 6.2.35 Synthesis of ethyl 2-azidoacetate



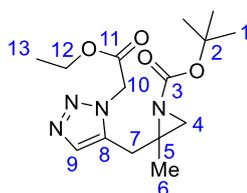
To a solution of ethyl bromoacetate (300 mg, 1.79 mmol) in acetone (7.5 mL) and H<sub>2</sub>O (2.5 mL), was added sodium azide (233 mg, 3.60 mmol) and the reaction stirred for 18 hours at r.t. The reaction was diluted with brine and extracted with EtOAc (2 x 20 mL). The combined

organic extracts were dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated *in vacuo* to give ethyl 2-azidoacetate (220 mg, 1.70 mmol, 95%) as a colourless oil without further purification.

$R_f$  = 0.80 (petroleum ether/EtOAc, 4:1); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 4.24 (2H, q,  $J$  = 7.2 Hz, H3), 3.84 (2H, s, H1), 1.29 (3H, t,  $J$  = 7.1 Hz, H4); <sup>13</sup>C (101 MHz, CDCl<sub>3</sub>) δ 168.4 (C2), 61.9 (C3), 50.4 (C1), 14.2 (C4); IR  $\nu_{\max}$  = 2986 (m, C-H), 2103 (s, N=N), 1741 (s, C=O), 1193 (C-O-C).

These characterisation data are in accordance with that previously reported in the literature.<sup>331</sup>

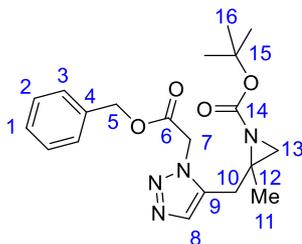
### 6.2.36 Synthesis of *tert*-butyl 2-((1-(2-ethoxy-2-oxoethyl)-1*H*-1,2,3-triazol-5-yl)methyl)-2-methylaziridine-1-carboxylate (**49**)



A solution of **48** (50.0 mg, 0.256 mmol) and ethyl 2-azidoacetate (49.6 mg, 0.384 mmol) in toluene (3 mL) was degassed with argon before Cp\*Ru(COD)Cl (3.90 mg, 0.0102 mmol) was added. The reaction was stirred at r.t. for 18 hours and then heated to 60 °C for 3 hours. Further Cp\*Ru(COD)Cl (7.80 mg, 0.0204 mmol) was added and the reaction stirred at the same temperature for 24 hours. Upon completion the reaction was concentrated *in vacuo* and the crude product purified by flash column chromatography (silica gel, petroleum ether/EtOAc, 3:2) to give the **49** (35.0 mg, 0.108 mmol, 42%) as a brown oil.

$R_f$  = 0.12 (petroleum ether/EtOAc, 1:1); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.61 (1H, s, H9), 5.41 (1H, d,  $J$  = 17.7 Hz, H10a), 5.26 (1H, d,  $J$  = 17.8 Hz, H10b) 4.26 – 4.19 (2H, m, H12), 3.01 (1H, d,  $J$  = 15.5 Hz, H7a), 2.73 (1H, d,  $J$  = 15.5 Hz, H7b), 2.17, (1H, s, H4a), 2.01 (1H, s, H4b), 1.41 (9H, s, H1), 1.28 (3H, t,  $J$  = 7.3 Hz, H13), 1.24 (3H, s, H6); <sup>13</sup>C (101 MHz, CDCl<sub>3</sub>) δ 166.9 (C11), 160.3 (C3), 134.1 (C9), 133.7 (C8), 81.8 (C2), 62.3 (C12), 49.5 (C10), 41.4 (C5), 36.1 (C4), 31.2 (C7), 28.1 (C1), 19.1 (C6), 14.2 (C13); HRMS (ESI) calcd for [C<sub>15</sub>H<sub>25</sub>N<sub>4</sub>O<sub>4</sub>]<sup>+</sup>: 325.1876, found 325.1871; IR  $\nu_{\max}$  = 2978 (w, C-H), 1749 (s, C=O), 1712 (s, C=O), 1553 (w, C=C), 1457 (m, C=C).

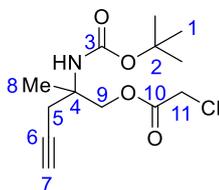
### 6.2.37 Synthesis of *tert*-butyl 2-(((1-(2-(benzyloxy)-2-oxoethyl)-1*H*-1,2,3-triazol-5-yl)methyl)-2-methylaziridine-1-carboxylate (**50**)



A solution of **48** (31.0 mg, 0.154 mmol) and benzyl 2-azidoacetate (44.2 mg, 0.231 mmol) in toluene (1.00 mL) was degassed with argon before Cp\**Ru*(COD)Cl (5.00 mg, 0.0130 mmol) was added and the reaction heated to 80 °C for 90 minutes. The reaction was filtered through Celite washing with EtOAc and the filtrate concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, gradient 1:1 to 0:1) to give the **50** (57.0 mg, 0.147 mmol, 95%) as a white solid.

$R_f$  = 0.34 (petroleum ether/EtOAc, 2:3); **m.p.** 92.1 – 93.0 °C;  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.62 (1H, s, H8), 7.38 – 7.31 (5H, m, H1/2/3), 5.49 (1H, d,  $J$  = 17.9 Hz, H7a), 5.32 (1H, d,  $J$  = 18.0 Hz, H7b), 5.24 – 5.17 (2H, m, H5), 2.99 (1H, d,  $J$  = 15.5 Hz, H10a), 2.68 (1H, d,  $J$  = 15.6 Hz, H10b), 2.15 (1H, s, H13a), 1.97 (1H, s, H13b), 1.40 (9H, s, H16), 1.21 (3H, s, H11);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  166.8 (C6), 160.3 (C14), 134.8 (C4), 134.2 (C8), 133.7 (C9), 128.9 (C2), 128.9 (C1), 128.6 (C3), 81.9 (C15), 67.9 (C5), 49.6 (C7), 41.5 (C12), 36.1 (C13), 31.2 (C10), 28.1 (C16), 19.0 (C11); **HRMS** (ESI) calcd for  $[\text{C}_{20}\text{H}_{27}\text{N}_4\text{O}_4]^+$ : 387.2024, found 387.2032; **IR**  $\nu_{\text{max}}$  = 2979 (w, C-H), 1748 (s, C=O), 1704 (s, C=O), 1553 (w, C=C), 1500 (w, C=C).

### 6.2.38 Synthesis of 2-((*tert*-butoxycarbonyl)amino)-2-methylpent-4-yn-1-yl 2-chloroacetate (**51**)

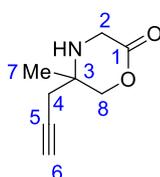


To an ice-cooled solution of **7** (84.0 mg, 0.394 mmol) and trimethylamine (65.0  $\mu\text{L}$ , 0.473 mmol) in  $\text{CH}_2\text{Cl}_2$  (3 mL) was added chloroacetyl chloride (37.6  $\mu\text{L}$ , 0.473 mmol) and the reaction warmed to r.t. over 20 minutes before heating to 40 °C for 7 hours. Further chloroacetyl chloride (9.40  $\mu\text{L}$ , 0.118 mmol) was added and the reaction heated for a further 1.5 hours. Upon completion, the reaction was diluted with  $\text{H}_2\text{O}$  and extracted with  $\text{CH}_2\text{Cl}_2$  (3 x 20 mL). The combined organic extracts were dried over  $\text{Na}_2\text{SO}_4$ , filtered and concentrated *in vacuo*. The

crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 3:2) to give **51** (80.3 mg, 0.278 mmol, 70%) as a colourless oil.

$R_f$  = 0.14 (petroleum ether/EtOAc, 85:15);  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  4.68 (1H, br s, NH), 4.41 (1H, d,  $J$  = 11.0 Hz H9a), 4.32 (1H, d,  $J$  = 11.0 Hz H9b), 4.08 (2H, s, H11), 2.74 (1H, dd,  $J$  = 16.7, 2.1 Hz, H5a), 2.54 (1H, dd,  $J$  = 16.7, 2.6 Hz, H5b), 2.04 (1H, t,  $J$  = 2.7 Hz, H7), 1.41 (9H, s, H1), 1.38 (3H, s, H8);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  167.0 (C10), 154.4 (C3), 79.9 (C2), 79.4 (C6), 71.7 (C7), 68.8 (C9), 54.0 (C4), 40.8 (C11), 28.4 (C1), 26.9 (C5), 22.0 (C8); **HRMS (ESI)** calcd for  $[\text{C}_{13}\text{H}_{21}\text{NO}_4\text{Cl}]^+$ : 290.1159, found 290.1144; **IR**  $\nu_{\text{max}}$  = 3295 (w,  $\text{C}\equiv\text{C-H}$ ), 2979 (w, C-H), 2002 (w,  $\text{C}\equiv\text{C}$ ), 1755 (m,  $\text{C}=\text{O}$ ), 1702 (s,  $\text{C}=\text{O}$ ), 1243 (m, C-Cl).

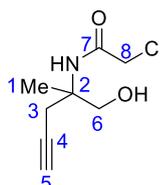
### 6.2.39 Synthesis of 5-methyl-5-(prop-2-yn-1-yl)morpholin-2-one (**53**)



To a solution of **51** (24.0 mg, 0.0830 mmol) in  $\text{CH}_2\text{Cl}_2$  (0.9 mL) was added TFA (0.85 mL) and the reaction stirred at r.t. for 2 hours 30 minutes. The reaction was concentrated *in vacuo* to afford a colourless gum. To a solution of this salt (40.0 mg, 0.138 mmol) in THF (1.30 mL) was added triethylamine (77.0  $\mu\text{L}$ , 0.553 mmol) and the reaction stirred for 6 hours. The reaction was diluted with water and extracted with  $\text{CH}_2\text{Cl}_2$  (3 x 20 mL). The combined organic extracts were washed with  $\text{NH}_4\text{Cl}$  saturated aqueous solution, dried over  $\text{Na}_2\text{SO}_4$ , filtered and concentrated *in vacuo* to give **53** (16.0 mg, 0.105 mmol, 76%) as a colourless oil.

$R_f$  = 0.41 (petroleum ether/EtOAc, 1:1);  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  6.91 (1H, br s, NH), 4.03 (2H, s, H2), 3.73 (2H, s, H8), 2.78 (1H, dd,  $J$  = 17.0, 2.8 Hz, H4a), 2.48 (1H, dd,  $J$  = 16.9, 2.7 Hz, H4b), 2.11 (1H, t,  $J$  = 2.8 Hz, H6), 1.40 (3H, s, H7);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  166.7 (C1), 79.5 (C5), 72.1 (C6), 68.3 (C8), 58.0 (C3), 43.0 (C2), 26.9 (C4), 22.2 (C7); **HRMS (ESI)** calcd for  $[\text{C}_8\text{H}_{12}\text{NO}_2]^+$ : 154.0863, found 154.0862; **IR**  $\nu_{\text{max}}$  = 3305 (w,  $\text{C}\equiv\text{C-H}$ ), 3057 (w, C-H), 1666 ( $\text{C}=\text{O}$ ), 1264 (s, C-O).

### 6.2.40 Synthesis of 2-chloro-N-(1-hydroxy-2-methylpent-4-yn-2-yl)acetamide (**52**)

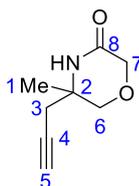


To a solution of **7** (140 mg, 0.656 mmol) in  $\text{CH}_2\text{Cl}_2$  (1 mL) was added HCl (2M in  $\text{Et}_2\text{O}$ ) (4.28 mL, 0.856 mmol) and the reaction stirred for 22 hours. Then, the reaction was concentrated *in*

*vacuo* to give the intermediate salt. A solution of the intermediate (93.0 mg) in CH<sub>2</sub>Cl<sub>2</sub> (6.5 mL) was cooled to 0 °C before triethylamine (0.260 mL, 1.97 mmol) was added. Chloroacetyl chloride (59.0 μL, 0.744 mmol) was added drop-wise and the reaction was warmed to r.t. and stirred for 5 hours. Then, NH<sub>4</sub>Cl saturated aqueous solution (30 mL) was added and extracted with EtOAc (3 x 10 mL). The combined organics were dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc 7:3) to give **52** (102 mg, 0.538 mmol, 88%) as a colourless oil.

**R<sub>f</sub>** = 0.21 (petroleum ether/EtOAc 3:2); **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 6.91 (1H, br s, NH or OH), 4.04 (2H, s, H8), 3.95 (1H, s, NH or OH), 3.73 (2H, s, H6), 2.78 (1H, dd, *J* = 16.9, 2.7 Hz, H3a), 2.48 (1H, dd, *J* = 16.9, 2.6 Hz, H3b), 2.11 (1H, t, *J* = 2.7 Hz, H5), 1.40 (3H, s, H1); **<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>) δ 166.7 (C7), 79.4 (C4), 72.1 (C5), 68.4 (C6), 58.0 (C2), 43.0 (C8), 26.9 (C3), 22.2 (C1); **HRMS** (ESI) calcd for [C<sub>8</sub>H<sub>13</sub>NO<sub>2</sub><sup>35</sup>Cl]<sup>+</sup>: 190.0629, found 190.0629. **IR** ν<sub>max</sub> = 3294 (w, C≡C-H), 2937 (w, C-H), 1660 (s, C=O), 1262 (m, C-OH), 1057 (s, C-OH).

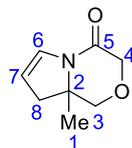
#### 6.2.41 Synthesis of 5-methyl-5-(prop-2-yn-1-yl)morpholin-3-one (**54**)



To a solution of **52** (102 mg, 0.540 mmol) in *t*-BuOH (5.4 mL) was added *t*-BuOK (65.0 mg, 0.579 mmol) and the reaction heated to 30 °C for 6 hours. Then, NH<sub>4</sub>Cl saturated aqueous solution was added and the solution extracted with EtOAc (3 x 10 mL). The combined organic extracts were dried over Mg<sub>2</sub>SO<sub>4</sub> and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc 7:3) to give **54** (60.0 mg, 0.392 mmol, 73%) as a white solid.

**R<sub>f</sub>** = 0.09 (petroleum ether/EtOAc 3:2); **m.p.** 77 – 79 °C; **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 6.32 (1H, s, NH), 4.18 (1H, d, *J* = 16.7 Hz, H7a), 4.11 (1H, d, *J* = 16.7 Hz, H7b), 3.75 (1H, d, *J* = 11.8 Hz, H6a), 3.53 (1H, d, *J* = 11.8 Hz, H6b), 2.56 (1H, dd, *J* = 16.5, 2.6 Hz, H3a), 2.44 (1H, dd, *J* = 16.5, 2.6 Hz, H3b), 2.11 (1H, t, *J* = 2.6 Hz, H5), 1.34 (3H, s, H1); **<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>) δ 168.4 (C8), 79.0 (C4), 72.2 (C5), 71.9 (C6), 67.7 (C7), 53.9 (C2), 30.0 (C3), 23.5 (C1); **HRMS** (ESI) calcd for [C<sub>8</sub>H<sub>12</sub>NO<sub>2</sub>]<sup>+</sup>: 154.0863, found 154.0863; **IR** ν<sub>max</sub> = 3526 (w, N-H), 3286 (m, C≡C-H), 3212 (w, C-H), 1655 (s, C=O).

## 6.2.42 Synthesis of 8a-methyl-8,8a-dihydro-1H-pyrrolo[2,1-c][1,4]oxazin-4(3H)-one (55)

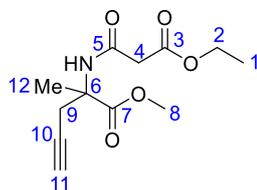


InCl<sub>3</sub> (188 mg, 0.852 mmol) was introduced into 10 mL flask and heated with a heat gun (150 °C) under vacuum for 2 minutes. After being allowed to cool to r.t., THF (4 mL) was added under Argon. The mixture was stirred at r.t. for 10 minutes and then cooled to –78 °C. DIBAL-H (1.0 M in hexane, 0.852 mL, 0.852 mmol) was added drop-wise and the mixture was stirred at –78 °C for 40 minutes. Then, **54** (87.0 mg, 0.570 mmol) in THF (1 mL) was added, followed by BEt<sub>3</sub> (1.0 M in THF, 0.284 mL, 0.284 mmol) and the mixture was stirred at –78 °C for 4.5 hours. A solution of iodine (865 mg, 3.41 mmol) in THF (1 mL) was then added. After 1 hour, the mixture was poured onto NaHCO<sub>3</sub> saturated aqueous solution (5 mL). Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> was added under stirring until complete decoloration and the aqueous layer was extracted with EtOAc (5x). The combined organic extracts were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated *in vacuo*. The crude sample was purified over a short column (silica gel; petroleum ether/EtOAc, 7:3). The crude product (90.0 mg), Cs<sub>2</sub>CO<sub>3</sub> (121 mg, 0.373 mmol) Cul (24.0 mg, 0.124 mmol) and *N,N*-dimethylethyl-1,2-diamine (27.0 uL, 0.249 mmol) in toluene (2 mL) were heated to 80 °C for 3 hours (in a sealed tube). H<sub>2</sub>O (10 mL) was added and the reaction extracted with CH<sub>2</sub>Cl<sub>2</sub> (5 x 10 mL). The combined organic extracts were washed with brine (10 mL), dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated *in vacuo* to give a mixture of desired product and unreacted **55** (ca. 1:1 determined by <sup>1</sup>H NMR). The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 7:3) to give **50** (12.0 mg, 0.0783 mmol, 24%) as a white solid (based on 43% recovered unreacted starting material).

**R<sub>f</sub>** = 0.17 (petroleum ether/EtOAc, 7:3); **m.p.** 77 – 79 °C; **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 6.87 – 6.85 (1H, m, H6), 5.34 – 5.32 (1H, m, H7), 4.28 (1H, d, *J* = 17.1 Hz, H4a), 4.08 (1H, d, *J* = 17.1 Hz, H4b), 3.93 (1H, d, *J* = 11.2 Hz, H3a), 3.58 (1H, d, *J* = 11.2 Hz, H3b), 2.59 (1H, dt, *J* = 15.8, 2.5 Hz, H8a), 2.34 (1H, ddd, *J* = 15.8, 3.1, 1.3 Hz, H8b), 1.40 (3H, s, H1); **<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>) δ 164.0 (C5), 127.3 (C6), 111.3 (C7), 72.4 (C3), 66.1 (C4), 61.5 (C2), 40.5 (C8), 24.0 (C1); **HRMS** (ESI) calcd for [C<sub>8</sub>H<sub>12</sub>N<sub>2</sub>O]<sup>+</sup>: 154.0868, found 154.0870; **IR** *v*<sub>max</sub> = 3102 (w, C-H), 2972 (w, C-H), 1634 (C=O), 1607 (s, C=C), 1090 (s, C-O-C).

## 6.3 Synthesis of NM450 derivatives

### 6.3.1 Synthesis of methyl 2-(3-ethoxy-3-oxopropanamido)-2-methylpent-4-ynoate (**77**)



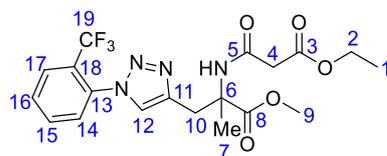
To an ice-cooled solution of **4** (260 mg, 1.84 mmol) in  $\text{CH}_2\text{Cl}_2$  (18 mL) was added triethylamine (0.385 mL, 2.76 mmol) followed by ethyl malonylchloride (0.354 mL, 2.76 mmol) portionwise. The reaction was warmed to r.t. and stirred for 3 hours and then  $\text{NH}_4\text{Cl}$  saturated aqueous solution was added (20 mL). The organic layer was separated and the aqueous further extracted with  $\text{CH}_2\text{Cl}_2$  (2 x 15 mL). The combined organic extracts were dried over  $\text{Na}_2\text{SO}_4$  and concentrated *in vacuo*. The crude sample was purified by flash column chromatography (silica gel; Heptane/EtOAc, gradient 0 – 60%) to give **77** (494 mg, 1.23 mmol, 95%) as a colourless oil.

$R_f = 0.23$  (Petroleum ether/EtOAc, 3:2);  $^1\text{H}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.81 (1H, br s, NH), 4.20 (2H, q,  $J = 7.1$  Hz, H2), 3.75 (3H, s, H8), 3.30 (2H, s, H4), 2.96 (2H, m, H9a & 9b), 1.99 (1H, t,  $J = 2.7$  Hz, H11), 1.60 (3H, s, H12), 1.28 (3H, t,  $J = 7.1$  Hz, H1);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  173.2 (C7), 169.3 (C3), 164.6 (C5), 79.3 (C10), 71.3 (C11), 61.8 (C2), 58.8 (C6), 53.1 (C8), 41.5 (C4), 26.5 (C9), 22.8 (C12), 14.2 (C1); **HRMS** (ESI) calcd for  $[\text{C}_{12}\text{H}_{18}\text{NO}_5]^+$ : 256.1185, found 256.1189; **IR**  $\nu_{\text{max}} = 3272$  (w,  $\text{C}\equiv\text{C-H}$ ), 2984 (w, C-H), 1735 (s, C=O), 1656 (s, C=O), 1124 (s, C-O-C).

### 6.3.2 General procedure 1

To a solution of alkyne **77** (1 eq.) in *t*-BuOH (1.5 mL) and  $\text{H}_2\text{O}$  (1.5 mL) was added copper(II) sulfate pentahydrate (0.1 eq) and sodium ascorbate (0.3 eq.). The corresponding azide (0.9 eq.) was then added and the reaction stirred for 16 hours. Upon completion,  $\text{H}_2\text{O}$  was added and the mixture extracted with EtOAc (3x). The combined organic extracts were dried ( $\text{Na}_2\text{SO}_4$ ), concentrated *in vacuo* and purified by flash column chromatography to give the product.

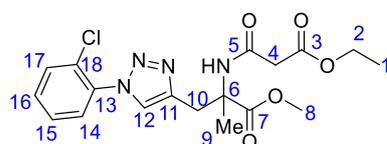
### 6.3.3 Synthesis of methyl 2-(3-ethoxy-3-oxopropanamido)-2-methyl-3-(1-(2-(trifluoromethyl)phenyl)-1*H*-1,2,3-triazol-4-yl)propanoate (**79**)



Following general procedure 1: **77** (65.0 mg, 0.255 mmol), copper(II) sulfate pentahydrate (6.36 mg, 0.0255 mmol), sodium ascorbate (15.1 mg, 0.0762 mmol), 1-azido-2-(trifluoromethyl) benzene (0.5M in TBME, 0.458 mL, 0.230 mmol), *t*-BuOH (1.5 mL) and water (1.5 mL) were used. The crude product was purified by flash column chromatography (silica gel; heptane/EtOAc, 40-100%) to give **79** (78.0 mg, 0.176 mmol, 70%) as a colourless oil.

$R_f$  = 0.32 (heptane/EtOAc, 7:3);  $^1\text{H}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.85 – 7.83 (1H, m, H17), 7.75 – 7.71 (1H, m, H15), 7.68 – 7.66 (3H, m, H12, H16 & NH), 7.56 (1H, d,  $J$  = 8.0 Hz, H14), 4.16 (2H, q,  $J$  = 7.1 Hz, H2), 3.81 (3H, s, H9), 3.68 (1H, d,  $J$  = 14.7 Hz, H10a), 3.52 (1H, d,  $J$  = 14.6 Hz, H10b), 3.27 (2H, s, H4), 1.69 (3H, s, H7), 1.24 (3H, t,  $J$  = 7.2 Hz, H1);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  173.9 (C8), 168.8 (C3), 164.7 (C5), 142.8 (C11), 135.1 (q,  $J$  = 2.4 Hz, C13), 133.2 (C15), 130.4 (C16), 129.1 (C14), 127.4 (q,  $J$  = 3.2 Hz, C17), 125.6 (q,  $J$  = 38 Hz, C18), 125.4 (C12), 125.0 (q,  $J$  = 220 Hz, C19), 61.7 (C2), 60.2 (C6), 53.1 (C9), 42.1 (C4), 32.0 (C10), 23.3 (C7), 14.1 (C1);  $^{19}\text{F}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -59.24 **HRMS** (ESI) calcd for  $[\text{C}_{19}\text{H}_{22}\text{N}_4\text{O}_5\text{F}_3]^+$ : 443.1537, found 443.1523; **IR**  $\nu_{\text{max}}$  = 3315 (w, N-H), 2988 (w, C-H), 1736 (s, C=O), 1677 (s, C=O), 1608 (m, C=C), 1509 (s, C=C), 1316 (s, C-O-C).

### 6.3.4 Synthesis of methyl 3-(1-(2-chlorophenyl)-1*H*-1,2,3-triazol-4-yl)-2-(3-ethoxy-3-oxopropanamido)-2-methylpropanoate (**80**)

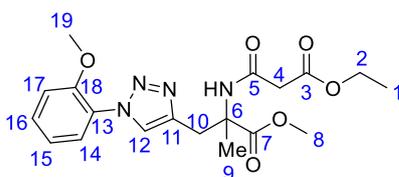


Following general procedure 1: **77** (65.0 mg, 0.255 mmol), copper(II) sulfate pentahydrate (6.36 mg, 0.0255 mmol), sodium ascorbate (15.1 mg, 0.0762 mmol), 1-azido-2-chlorobenzene (0.5M in TBME, 0.458 mL, 0.230 mmol), *t*-BuOH (1.5 mL) and water (1.5 mL) were used. After 16 hours, further copper(II) sulfate pentahydrate (6.36 mg, 0.0255 mmol) was added and the reaction stirred for a further 16 hours. The crude product was purified by flash column chromatography (silica gel; heptane/EtOAc, 20-100%) to give **80** (70.0 mg, 0.172 mmol, 69%) as a colourless oil.

$R_f$  = 0.35 (Petroleum ether/EtOAc, 7:13);  $^1\text{H}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.80 (1H, s, H12), 7.65 (1H, br s, NH), 7.63 – 7.60 (1H, m, H14), 7.58 – 7.54 (1H, m, H17), 7.47 – 7.41 (2H, m, H16, H15),

4.15 (2H, q,  $J = 7.2$  Hz, H2), 3.82 (3H, s, H8), 3.68 (1H, d,  $J = 7.6$  Hz, H10a), 3.51 (1H, d,  $J = 7.6$  Hz, H10b), 3.28 (2H, s, H4), 1.69 (3H, s, H9), 1.24 (3H, t,  $J = 7.1$  Hz, H1);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  173.9 (C7), 168.8 (C3), 164.7 (C5), 142.6 (C11), 135.1 (C18), 130.9 (C17), 130.8 (C16), 128.7 (C13), 128.0 (C15), 127.9 (C14), 125.0 (C12), 61.7 (C2), 60.2 (C6), 53.2 (C8), 42.3 (C4), 32.1 (C10), 23.3 (C9), 14.2 (C1); **HRMS** (ESI) calcd for  $[\text{C}_{18}\text{H}_{22}\text{N}_4\text{O}_5^{35}\text{Cl}]^+$ : 409.1279, found 409.1268; **IR**  $\nu_{\text{max}} = 3303$  (w, N-H), 2986 (w, C-H), 1735 (s, C=O), 1676 (s, C=O), 1670 (s, C=O), 1535 (s, C=C), 1496 (s, C=C), 1456 (s, C=C).

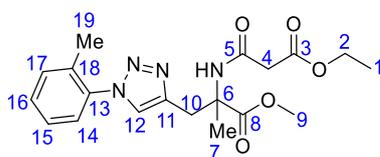
### 6.3.5 Synthesis of methyl 2-(3-ethoxy-3-oxopropanamido)-3-(1-(2-methoxyphenyl)-1*H*-1,2,3-triazol-4-yl)-2-methylpropanoate (**81**)



Following general procedure 1: **77** (65.0 mg, 0.255 mmol), copper(II) sulfate pentahydrate (6.36 mg, 0.0255 mmol), sodium ascorbate (15.1 mg, 0.0762 mmol), 1-azido-2-methoxybenzene (0.5M in TBME, 0.458 mL, 0.230 mmol), *t*-BuOH (1.5 mL) and water (1.5 mL) were used. The crude product was purified by flash column chromatography (silica gel; heptane/EtOAc, 50-100%) to give **81** (66.0 mg, 0.163 mmol, 65%) as a white solid.

$R_f = 0.27$  (Petroleum ether/EtOAc, 7:13);  $^1\text{H}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.92 (1H, s, H12), 7.78 (1H, dd,  $J = 7.9, 1.7$  Hz, H14), 7.62 (1H, br s, NH), 7.43 – 7.38 (1H, m, H16), 7.12 – 7.06 (2H, m, H17, H15), 4.14 (2H, q,  $J = 7.2$  Hz, H2), 3.88 (3H, s, H19), 3.81 (3H, s, H8), 3.61 (1H, d,  $J = 14.6$  Hz, H10a), 3.47 (1H, d,  $J = 14.6$  Hz, H10b), 3.28 (2H, s, H4), 1.69 (3H, s, H9), 1.22 (3H, t,  $J = 7.2$  Hz, H1);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  174.0 (C7), 168.7 (C3), 164.6 (C5), 151.1 (C18), 142.1 (C11), 130.0 (C16), 126.5 (C13), 125.5 (C14), 125.0 (C12), 121.4 (C15), 112.4 (C17), 61.7 (C2), 60.2 (C6), 56.1 (C19), 53.1 (C8), 42.5 (C4), 32.4 (C10), 23.1 (C9), 14.1 (C1); **HRMS** (ESI) calcd for  $[\text{C}_{19}\text{H}_{25}\text{N}_4\text{O}_6]^+$ : 405.1774, found 405.1787; **IR**  $\nu_{\text{max}} = 3251$  (w, N-H) 3056 (w, C-H) 1752 (s, C=O), 1731 (s, C=O), 1670 (s, C=O), 1557 (m, C=C), 1504 (m, C=C), 1453 (m, C=C), 1119 (s, C-O-C).

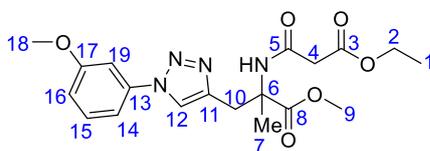
### 6.3.6 Synthesis of methyl 2-(3-ethoxy-3-oxopropanamido)-2-methyl-3-(1-(*o*-tolyl)-1*H*-1,2,3-triazol-4-yl)propanoate (**82**)



Following general procedure 1: **77** (65.0 mg, 0.255 mmol), copper(II) sulfate pentahydrate (6.36 mg, 0.0255 mmol), sodium ascorbate (15.1 mg, 0.0762 mmol), 1-azido-2-methylbenzene (0.5M in TBME, 0.458 mL, 0.230 mmol), *t*-BuOH (1.5 mL) and water (1.5 mL) were used. After 4 days stirring, further copper(II) sulfate pentahydrate (6.36 mg, 0.0255 mmol) was added and stirred for 16 hours. The crude product was purified by flash column chromatography (silica gel; heptane/EtOAc, 20-100%) to give **82** (12.0 mg, 0.0309 mmol, 12%) as a colourless oil.

$R_f$  = 0.36 (Petroleum ether/EtOAc, 7:13);  $^1\text{H}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.64 (1H, br s, NH), 7.57 (1H, s, H12), 7.44 – 7.30 (4H, m, H14, H15, H16 & H17), 4.14 (2H, q,  $J$  = 7.1 Hz, H2), 3.81 (3H, s, H9), 3.66 (1H, d,  $J$  = 14.5 Hz, H10a), 3.48 (1H, d,  $J$  = 14.7 Hz, H10b), 3.28 (2H, s, H4), 2.19 (3H, s, H19), 1.70 (3H, s, H7), 1.23 (3H, t,  $J$  = 7.2 Hz, H1);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  173.9 (C8), 168.8 (C3), 164.6 (C5), 142.5 (C11), 136.7 (C13), 133.7 (C18), 131.6 (C17), 129.9 (C16), 126.9 (C15), 126.1 (C14), 124.5 (C12), 61.7 (C2), 60.3 (C6), 53.1 (C9), 42.4 (C4), 32.3 (C10), 23.2 (C7), 18.0 (C19), 14.2 (C1); **HRMS** (ESI) calcd for  $[\text{C}_{19}\text{H}_{25}\text{N}_4\text{O}_5]^+$ : 389.1825, found 389.1832; **IR**  $\nu_{\text{max}}$  = 3292 (w, N-H), 1738 (s, C=O), 1663 (s, C=O), 1540 (s, C=C), 1188 (m, C-O-C).

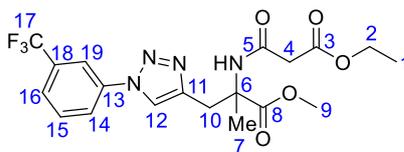
### 6.3.7 Synthesis of methyl 2-(3-methoxy-3-oxopropanamido)-3-(1-(3-methoxyphenyl)-1H-1,2,3-triazol-4-yl)-2-methylpropanoate (**83**)



Following general procedure 1: **77** (65.0 mg, 0.255 mmol), copper(II) sulfate pentahydrate (6.36 mg, 0.0255 mmol), sodium ascorbate (15.1 mg, 0.0762 mmol), 1-azido-3-methoxybenzene (0.5M in TBME, 0.458 mL, 0.230 mmol), *t*-BuOH (1.5 mL) and water (1.5 mL) were used. The crude product was purified by flash column chromatography (silica gel; heptane/EtOAc, 20-100%) to give **83** (67.0 mg, 0.172 mmol, 69%) as a colourless oil.

$R_f$  = 0.37 (Petroleum ether/EtOAc, 7:13);  $^1\text{H}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.91 (1H, s, H12), 7.41 – 7.37 (3H, m, H15, H19 & NH), 7.30 – 7.28 (1H, m, H16), 6.96 – 6.93 (1H, m, H14), 4.17 – 4.12 (2H, m, H2), 3.88 (3H, s, H18), 3.83 (3H, s, H9), 3.71 (1H, d,  $J$  = 14.6 Hz, H10a), 3.48 (1H, d,  $J$  = 14.6 Hz, H10b), 3.28 (2H, s, H4), 1.69 (3H, s, H7), 1.23 (3H, t,  $J$  = 7.2 Hz, H1);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  173.9 (C8), 169.2 (C3), 164.5 (C5), 160.7 (C17), 143.3 (C11), 138.2 (C13), 130.6 (C15), 121.3 (C12), 114.7 (C16), 112.2 (C14), 106.0 (C19), 61.8 (C2), 60.4 (C6), 55.8 (C18), 53.2 (C9), 42.5 (C4), 32.1 (C10), 23.2 (C7), 14.1 (C1); **HRMS** (ESI) calcd for  $[\text{C}_{19}\text{H}_{24}\text{N}_4\text{O}_6\text{Na}]^+$ : 427.1588, found 427.1586; **IR**  $\nu_{\text{max}}$  = 3264 (w, N-H), 3083 (w, C-H), 2952 (w, C-H), 1733 (s, C=O), 1660 (s, C=O), 1648 (s, C=O), 1610 (s, C=C), 1595 (s, C=C), 1498 (m, C=C).

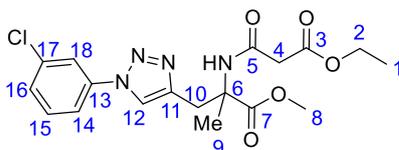
### 6.3.8 Synthesis of methyl 2-(3-ethoxy-3-oxopropanamido)-2-methyl-3-(1-(3-(trifluoromethyl)phenyl)-1*H*-1,2,3-triazol-4-yl)propanoate (**84**)



Following general procedure 1: **77** (65.0 mg, 0.255 mmol), copper(II) sulfate pentahydrate (6.36 mg, 0.0255 mmol), sodium ascorbate (15.1 mg, 0.0762 mmol), 1-azido-3-(trifluoromethyl) benzene (0.5M in TBME, 0.458 mL, 0.230 mmol), *t*-BuOH (1.5 mL) and water (1.5 mL) were used. The crude product was purified by flash column chromatography (silica gel; heptane/EtOAc, 20-100%) to give **84** (93.0 mg, 0.210 mmol, 84%) as a white solid.

$R_f$  = 0.47 (Petroleum ether/EtOAc, 7:13);  $^1\text{H}$  NMR (400MHz,  $\text{CDCl}_3$ )  $\delta$  8.15 (1H, s, H12), 8.11 – 8.10 (1H, m, H19), 8.06 – 8.04 (1H, m, H14), 7.69 – 7.63 (2H, m, H15 & H16), 7.24 (1H, br s, NH), 4.20 – 4.14 (2H, m, H2), 3.84 (3H, s, H9), 3.80 (1H, d,  $J$  = 14.6 Hz, H10a), 3.51 (1H, d,  $J$  = 14.6 Hz, H10b), 3.30 (2H, s, H4), 1.71 (3H, s, H7), 1.25 (3H, t,  $J$  = 7.1 Hz, H1);  $^{13}\text{C}$  (101MHz,  $\text{CDCl}_3$ ) 173.8 (C8), 168.8 (C3), 164.6 (C5), 143.7 (C11), 137.5 (C13), 132.3 (q,  $J$  = 33.2 Hz, C18), 130.4 (C15), 125.0 (q,  $J$  = 3.7 Hz, C16), 123.4 (q,  $J$  = 272.9 Hz, C17), 123.3 (C14), 121.4 (C12), 117.0 (q,  $J$  = 4.0 Hz, C19), 61.7 (C2) 60.4 (C6), 53.2 (C9), 42.8 (C4), 31.7 (C10), 23.2 (C7), 14.0 (C1);  $^{19}\text{F}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -62.9; **HRMS** (ESI) calcd for  $[\text{C}_{19}\text{H}_{22}\text{N}_4\text{O}_5\text{F}_3]^+$ : 443.1542, found 443.1558; **IR**  $\nu_{\text{max}}$  = 3277 (w, N-H), 1736 (s, C=O), 1675 (s, C=O), 1661 (s, C=O), 1535 (m, C=C), 1460 (s, C=C), 1126 (s, C-O-C).

### 6.3.9 Synthesis of methyl 3-(1-(3-chlorophenyl)-1*H*-1,2,3-triazol-4-yl)-2-(3-ethoxy-3-oxopropanamido)-2-methylpropanoate (**85**)

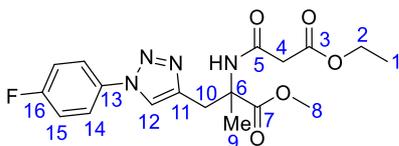


Following general procedure 1: **77** (65.0 mg, 0.255 mmol), copper(II) sulfate pentahydrate (6.36 mg, 0.0255 mmol), sodium ascorbate (15.1 mg, 0.0762 mmol), 1-azido-3-chlorobenzene (0.5M in TBME, 0.458 mL, 0.230 mmol), *t*-BuOH (1.5 mL) and water (1.5 mL) were used. The crude product was purified by flash column chromatography (silica gel; heptane/EtOAc, 20-100%) to give **85** (69.0 mg, 0.168 mmol, 68%) as a white solid.

**LCMS**  $\text{MH}^+$  (409.0);  $R_f$  = 0.34 (60% EtOAc/Heptane);  $^1\text{H}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.01 (1H, s, H12), 7.84 (1H, t,  $J$  = 2.0 Hz, H18), 7.72 – 7.69 (1H, m, H14), 7.45 (1H, t,  $J$  = 8.0 Hz, H15), 7.40 – 7.38 (1H, m, H16), 7.33 (1H, br s, NH), 4.20 – 4.16 (2H, m, H2), 3.83 (3H, s, H8), 3.75

(1H, d,  $J = 14.6$  Hz, H10a), 3.49 (1H, d,  $J = 14.6$  Hz, H10b), 3.30 (2H, s, H4), 1.70 (3H, s, H9), 1.26 (3H, t,  $J = 4.8$  Hz, H1);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  173.8 (C7), 168.7 (C3), 164.5 (C5), 143.5 (C11), 137.9 (C17), 135.5 (C13), 130.7 (C15), 128.5 (C16), 121.2 (C12), 120.5 (C18), 118.2 (C14), 61.7 (C2), 60.3 (C6), 53.1 (C8), 42.7 (C4), 31.8 (C10), 23.1 (C9), 14.0 (C1); **HRMS** (ESI) calcd for  $[\text{C}_{18}\text{H}_{22}\text{N}_4\text{O}_5^{35}\text{Cl}]^+$ : 409.1279, found 409.1277; **IR**  $\nu_{\text{max}}$  = 3255 (w, N-H), 3079 (w, C-H), 2986 (w, C-H), 1744 (s, C=O), 1729 (s, C=O), 1642 (s, C=O), 1597 (s, C=C), 1567 (s, C=C), 1197 (s, C-O-C).

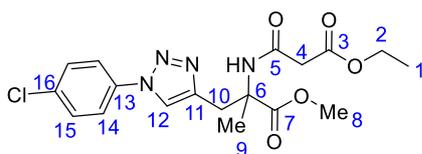
### 6.3.10 Synthesis of methyl 2-(3-ethoxy-3-oxopropanamido)-3-(1-(4-fluorophenyl)-1*H*-1,2,3-triazol-4-yl)-2-methylpropanoate (86)



Following general procedure 1: **77** (65.0 mg, 0.255 mmol), copper(II) sulfate pentahydrate (6.36 mg, 0.0255 mmol), sodium ascorbate (15.1 mg, 0.0762 mmol), 1-azido-4-fluorobenzene (0.5M in TBME, 0.458 mL, 0.230 mmol), *t*-BuOH (1.5 mL) and water (1.5 mL) were used. The crude product was purified by flash column chromatography (silica gel; heptane/EtOAc, 20-100%) to give **86** (81.0 mg, 0.206 mmol, 83%) as a white solid.

**LCMS** MH<sup>+</sup> (393.0);  $R_f = 0.37$  (petroleum ether/EtOAc, 2:3);  $^1\text{H}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.90 (1H, s, H12), 7.77 – 7.72 (2H, m, H14), 7.37 (1H, br s, NH), 7.23 – 7.17 (2H, m, H15), 4.17 – 4.12 (2H, m, H2), 3.83 (3H, s, H8), 3.72 (1H, d,  $J = 14.7$  Hz, H10a), 3.48 (1H, d,  $J = 14.8$  Hz, H10b), 3.28 (2H, s, H4), 1.70 (3H, s, H9), 1.24 (3H, t,  $J = 7.2$  Hz, H1);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  174.0 (C7), 168.8 (C3), 164.6 (C5), 162.3 (d,  $J = 234.4$  Hz, C16), 143.6 (C11), 133.5 (C13), 122.3 (d,  $J = 8.7$  Hz, C14), 121.5 (C12), 116.8 (d,  $J = 23.0$  Hz, C15), 61.8 (C2), 60.4 (C6), 53.2 (C8), 42.8 (C4), 32.0 (C10), 23.3 (C9), 14.2 (C1);  $^{19}\text{F}$  (376 MHz,  $\text{CDCl}_3$ )  $\delta$  -112.7; **HRMS** (ESI) calcd for  $[\text{C}_{18}\text{H}_{22}\text{N}_4\text{O}_5\text{F}]^+$ : 393.1574, found 393.1563; **IR**  $\nu_{\text{max}}$  = 3356 (w, N-H), 2928 (w, C-H), 1736 (s, C=O), 1665 (s, C=O), 1517 (s, C=C).

### 6.3.11 Synthesis of methyl 3-(1-(4-chlorophenyl)-1*H*-1,2,3-triazol-4-yl)-2-(3-ethoxy-3-oxopropanamido)-2-methylpropanoate (87)

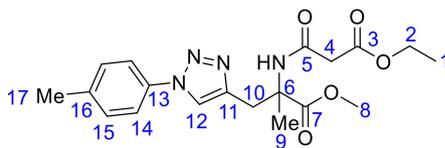


Following general procedure 1: **77** (65.0 mg, 0.255 mmol), copper(II) sulfate pentahydrate (6.36 mg, 0.0255 mmol), sodium ascorbate (15.1 mg, 0.0762 mmol), 1-azido-4-chlorobenzene

(0.5M in TBME, 0.458 mL, 0.230 mmol), *t*-BuOH (1.5 mL) and water (1.5 mL) were used. The crude product was purified by flash column chromatography (silica gel; heptane/EtOAc, 20-100%) to give **87** (76.0 mg, 0.186 mmol, 74%) as a white solid.

**LCMS** MH<sup>+</sup> (409.0); **R<sub>f</sub>** = 0.41 (Heptane/EtOAc, 2:3); **<sup>1</sup>H** (400 MHz, CDCl<sub>3</sub>) δ 7.94 (1H, s, H12), 7.72 (2H, dt, *J* = 9.7, 2.5 Hz, H14), 7.48 (2H, dt, *J* = 9.7, 2.5 Hz, H15), 7.35 (1H, br s, NH), 4.14 (2H, q, *J* = 7.2 Hz, H2), 3.83 (3H, s, H8), 3.74 (1H, d, *J* = 14.6 Hz, H10a), 3.48 (1H, d, *J* = 14.7 Hz, H10b), 3.28 (2H, s, H4), 1.70 (3H, s, H9), 1.24 (3H, t, *J* = 7.1 Hz, H1); **<sup>13</sup>C** (101 MHz, CDCl<sub>3</sub>) δ 173.9 (C7), 168.8 (C3), 164.7 (C5), 143.7 (C11), 135.7 (C16), 134.4 (C13), 130.0 (C15), 121.5 (C14), 121.3 (C12), 61.8 (C2), 60.4 (C6), 53.3 (C8), 42.9 (C4), 32.0 (C10), 23.3 (C9), 14.2 (C1); **HRMS** (ESI) calcd for [C<sub>18</sub>H<sub>22</sub>N<sub>4</sub>O<sub>5</sub><sup>35</sup>Cl]<sup>+</sup>: 409.1279, found 409.1287; **IR**  $\nu_{\max}$  = 3258 (w, N-H), 3078 (w, C-H), 2981 (w, C-H), 1744 (s, C=O), 1729 (s, C=O), 1641 (s, C=O), 1563 (s, C=C), 1502 (s, C=C), 1198 (s, C-O-C).

### 6.3.12 Synthesis of methyl 2-(3-ethoxy-3-oxopropanamido)-2-methyl-3-(1-(*p*-tolyl)-1*H*-1,2,3-triazol-4-yl)propanoate (**88**)



Following general procedure 1: **77** (65.0 mg, 0.255 mmol), copper(II) sulfate pentahydrate (6.36 mg, 0.0255 mmol), sodium ascorbate (15.1 mg, 0.0762 mmol), 1-azido-2-(trifluoromethyl) benzene (0.5M in TBME, 0.458 mL, 0.230 mmol), *t*-BuOH (1.5 mL) and water (1.5 mL) were used. The crude product was purified by flash column chromatography (silica gel; heptane/EtOAc, 20-100%) to give **88** (64.0 mg, 0.165 mmol, 66%) as a white solid.

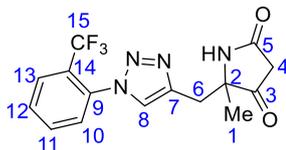
**LCMS** MH<sup>+</sup> (389.0); **R<sub>f</sub>** = 0.32 (Heptane/EtOAc, 3:2); **<sup>1</sup>H** (400 MHz, CDCl<sub>3</sub>) δ 7.86 (1H, s, H12), 7.61 (2H, dt, *J* = 8.9, 2.0 Hz H14), 7.45 (1H, br s, NH), 7.29 (2H, d, *J* = 8.1 Hz, H15), 4.14 (2H, q, *J* = 8.1 Hz, H2), 3.82 (3H, s, H8), 3.67 (1H, d, *J* = 14.6 Hz, H10a), 3.47 (1H, d, *J* = 14.6 Hz, H10b), 3.28 (2H, s, H4), 2.41 (3H, s, H17), 1.69 (3H, s, H9), 1.23 (3H, t, *J* = 7.2 Hz, H1); **<sup>13</sup>C** (101 MHz, CDCl<sub>3</sub>) δ 174.0 (C7), 168.7 (C5), 164.7 (C3), 143.3 (C11), 138.7 (C16), 134.9 (C13), 130.3 (C15), 121.2 (C12), 120.3 (C14), 61.8 (C2), 60.3 (C6), 53.2 (C8), 42.7 (C4), 32.1 (C10), 23.2 (C9), 21.2 (C17), 14.1 (C1); **HRMS** (ESI) calcd for [C<sub>19</sub>H<sub>25</sub>N<sub>4</sub>O<sub>5</sub>]<sup>+</sup>: 389.1825, found 389.1831; **IR**  $\nu_{\max}$  = 3251 (w, N-H), 3078 (w, C-H), 2979 (w, C-H), 1745 (s, C=O), 1732 (s, C=O), 1643 (s, C=O), 1568 (m, C=C), 1520 (m, C=C), 1197 (s, C-O-C).

### 6.3.13 General procedure 2

To a solution of triazole (1 eq.) in THF (0.1M) was added potassium *tert*-butoxide (1.5 eq.) the reaction refluxed for 1-4 hours. Upon completion, 1N HCl (aq.) was added and extracted with

EtOAc (3x). The combined organic extracts were dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated *in vacuo*. The crude intermediate was dissolved in acetonitrile and H<sub>2</sub>O (10:1, 0.1M) and refluxed until no starting material remained (2-4 hours). The solution was concentrated *in vacuo* and purified by flash chromatography to give the product.

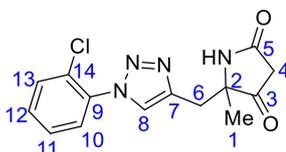
#### 6.3.14 Synthesis of 5-methyl-5-((1-(2-(trifluoromethyl)phenyl)-1H-1,2,3-triazol-4-yl)methyl)pyrrolidine-2,4-dione (**89**)



Following general procedure 2: **79** (70.0 mg, 0.158 mmol), potassium *tert*-butoxide (26.6 mg, 0.237 mmol) and THF (2 mL) were refluxed for 3 hours. After work up, the crude intermediate was dissolved in acetonitrile (2 mL) and H<sub>2</sub>O (0.2 mL) and refluxed for 3.5 hours. The crude product was purified by flash column chromatography (silica gel; heptane/EtOAc, 50 – 100%) to give **89** (34.0 mg, 0.100 mmol, 64%) as a white solid.

**LCMS** MH<sup>+</sup> 339.2; **R<sub>f</sub>** = 0.20 (EtOAc); **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.87 (1H, d, *J* = 7.8 Hz, H13), 7.78 – 7.74 (1H, m, H11), 7.70 (1H, t, *J* = 7.7 Hz H12), 7.66 (1H, s, H8), 7.55 (1H, d, *J* = 7.7 Hz, H10), 6.72 (1H, br s, NH), 3.23 (1H, d, *J* = 15.0 Hz, H6a), 3.14 – 3.07 (2H, m, H4a & H6b), 2.95 (1H, d, *J* = 22.2 Hz, H4b), 1.43 (3H, s, H1); **<sup>13</sup>C NMR** (101 MHz, d<sub>6</sub>-DMSO) δ 211.2 (C3), 169.2 (C5), 141.5 (C7), 134.3 (q, *J* = 1.7 Hz, H9), 133.9 (C11), 131.1 (C12), 130.0 (q, *J* = 191.1 Hz, C15), 129.2 (C10), 127.4 (q, *J* = 4.8 Hz, C13), 126.2 (C8), 124.8 (q, *J* = 31.2 Hz, C14), 66.6 (C2), 40.3 (C4), 33.6 (C6), 23.8 (C1); **<sup>19</sup>F** (376 MHz, d<sub>6</sub>-DMSO) δ 58.1; **HRMS** (ESI) calcd for [C<sub>15</sub>H<sub>14</sub>N<sub>4</sub>O<sub>2</sub>F<sub>3</sub>]<sup>+</sup>: 339.1069, found 339.1069; **IR** *v*<sub>max</sub> = 3241 (m, N-H), 1635 (s, C=O), 1625 (m, C=O), 1503 (s, C=C), 1467 (s, C=C), 1405 (m, C=C).

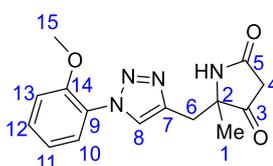
#### 6.3.15 Synthesis of 5-((1-(2-chlorophenyl)-1H-1,2,3-triazol-4-yl)methyl)-5-methylpyrrolidine-2,4-dione (**90**)



Following general procedure 2: **80** (50.0 mg, 0.122 mmol), potassium *tert*-butoxide (27.0 mg, 0.241 mmol) and THF (1.4 mL) were refluxed for 3 hours. After work up, the crude intermediate was dissolved in acetonitrile (3 mL) and H<sub>2</sub>O (0.3 mL) and refluxed for 4 hours. The crude product was purified by flash column chromatography (silica gel; heptane/EtOAc, 20 – 100%) to give **90** (19.0 mg, 0.0624 mmol, 51%) as a white solid.

$R_f = 0.22$  (EtOAc/MeOH, 24:1);  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.81 (1H, s, H8), 7.60 – 7.56 (2H, m, H10 & H13), 7.48 – 7.42 (2H, m, H11 & H12), 7.11 (1H, br s, NH), 3.25 (1H, d,  $J = 15.0$  Hz, H6a), 3.14 (1H, d,  $J = 15.0$  Hz, H6b), 3.06 (1H, d,  $J = 22.2$  Hz, H4a), 2.91 (1H, d,  $J = 22.2$  Hz, H4b), 1.46 (3H, s, H1);  $^{13}\text{C NMR}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  209.5 (C3), 169.4 (C5), 141.6 (C7), 134.8 (C14), 131.1 (C13), 130.9 (C12), 128.7 (C9), 128.2 (C11), 127.8 (C10), 124.9 (C8), 67.6 (C2), 40.2 (C4), 31.1 (C6), 24.0 (C1); **HRMS** calcd for  $[\text{C}_{14}\text{H}_{13}\text{N}_4\text{O}_2^{35}\text{ClNa}]^+$ : 327.0619, found 327.0618; **IR**  $\nu_{\text{max}} = 3250$  (m, N-H), 2931 (w, C-H), 1635 (m, C=O), 1620 (m, C=O), 1493 (s, C=C).

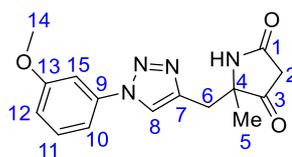
### 6.3.16 Synthesis of 5-((1-(2-methoxyphenyl)-1H-1,2,3-triazol-4-yl)methyl)-5-methylpyrrolidine-2,4-dione (91)



Following general procedure 2: **81** (52.0 mg, 0.129 mmol), potassium *tert*-butoxide (21.0 mg, 0.187 mmol) and THF (1.5 mL) were refluxed for 2 hours. After work up, the crude intermediate was dissolved in acetonitrile (1 mL) and  $\text{H}_2\text{O}$  (0.1 mL) and refluxed for 3 hours. The crude product was purified by flash column chromatography (silica gel; heptane/EtOAc, 20 – 100%) to give **91** (24.0 mg, 0.0799 mmol, 62%) as a white solid.

**LCMS**  $\text{MH}^+$  (300.9);  $R_f = 0.11$  (EtOAc);  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.92 (1H, s, H8), 7.74 (1H, dd,  $J = 7.9, 1.5$  Hz, H10), 7.45 – 7.40 (1H, m, H12), 7.12 – 7.06 (3H, m, H11, H13, NH), 3.88 (3H, s, H15), 3.22 (1H, d,  $J = 15.0$  Hz, H6a), 3.12 (1H, d,  $J = 15.0$  Hz, H6b), 3.04 (1H, d,  $J = 22.2$  Hz, H4a), 2.89 (1H, d,  $J = 22.2$  Hz, H4b), 1.46 (3H, s, H1);  $^{13}\text{C NMR}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  209.8 (C3), 169.5 (C5), 151.2 (C14), 141.1 (C7), 130.4 (C12), 126.2 (C9), 125.5 (C10), 125.0 (C8), 121.4 (C11), 112.4 (C13), 67.9 (C2), 56.1 (C15), 40.3 (C4), 33.9 (C6), 24.0 (C1); **HRMS**  $[\text{C}_{15}\text{H}_{17}\text{N}_4\text{O}_3]^+$ : 301.1301, found 301.1302; **IR**  $\nu_{\text{max}} = 3246$  (w, N-H), 1634 (s, C=O), 1602 (s, C=O), 1508 (s, C=C), 1467 (m, C=C).

### 6.3.17 Synthesis of 5-((1-(3-methoxyphenyl)-1H-1,2,3-triazol-4-yl)methyl)-5-methylpyrrolidine-2,4-dione (92)

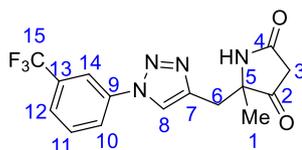


Following general procedure 2: **83** (55.0 mg, 0.136 mmol), potassium *tert*-butoxide (23.0 mg, 0.205 mmol) and THF (1.5 mL) were refluxed for 2 hours. After work up, the crude intermediate

was dissolved in acetonitrile (1 mL) and H<sub>2</sub>O (0.1 mL) and refluxed for 3 hours. The crude product was purified by flash column chromatography (silica gel; heptane/EtOAc, 20 – 100%) to give **92** (32.0 mg, 0.107 mmol, 78%) as a white film.

**LCMS** MH<sup>+</sup> (MH<sup>+</sup> 300.9); **R<sub>f</sub>** = 0.17 (EtOAc); **<sup>1</sup>H** (400 MHz, CDCl<sub>3</sub>) δ 7.83 (1H, s, H8), 7.62 (1H, br s, NH), 7.36 (1H, t, *J* = 8.2 Hz, H11), 7.27 – 7.26 (1H, m, H15), 7.19 (1H, dd, *J* = 7.9, 1.4 Hz, H12), 6.93 (1H, dd, *J* = 8.3, 2.2 Hz, H10), 3.86 (3H, s, H14), 3.20 (1H, d, *J* = 14.9 Hz, H6a), 3.12 (1H, d, *J* = 14.9 Hz, H6b), 3.05 (1H, d, *J* = 22.2 Hz, H2a), 2.90 (1H, d, *J* = 22.2 Hz, H2b), 1.46 (3H, s, H5); **<sup>13</sup>C** (101 MHz, CDCl<sub>3</sub>) δ 209.7 (C3), 170.0 (C1), 160.7 (C13), 142.4 (C7), 137.9 (C9), 130.6 (C11), 121.1 (C8), 114.7 (C12), 112.3 (C10), 106.3 (C15), 67.8 (C4), 55.8 (C14), 40.3 (C2), 33.9 (C6), 23.9 (C5); **HRMS** (ESI) calcd for [C<sub>15</sub>H<sub>17</sub>N<sub>4</sub>O<sub>3</sub>]<sup>+</sup>: 301.1301, found 301.1297; **IR** ν<sub>max</sub> = 3128 (w, N-H), 2976 (w, C-H), 1769 (m, N=N), 1698 (s, C=O), 1608 (s, C=C), 1594 (s, C=C), 1503 (s, C=C).

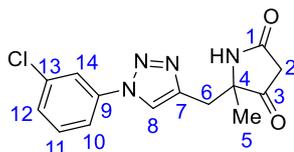
### 6.3.18 Synthesis of 5-methyl-5-((1-(3-(trifluoromethyl)phenyl)-1*H*-1,2,3-triazol-4-yl)methyl)pyrrolidine-2,4-dione (**93**)



Following general procedure 2: **84** (83.0 mg, 0.188 mmol), potassium *tert*-butoxide (32.0 mg, 0.285 mmol) and THF (2 mL) were refluxed for 2.5 hours. After work up, the crude intermediate was dissolved in acetonitrile (1 mL) and H<sub>2</sub>O (0.1 mL) and refluxed for 3 hours. The crude product was purified by flash column chromatography (silica gel; heptane/EtOAc, 20 – 100%) to give **93** (47.0 mg, 0.219 mmol, 74%) as an off-white solid.

**R<sub>f</sub>** = 0.26 (EtOAc); **<sup>1</sup>H** (400 MHz, CDCl<sub>3</sub>) δ 7.97 (1H, s, H8), 7.95 – 7.91 (2H, m, H10 & H14), 7.82 (1H, br s, NH), 7.68 – 7.61 (2H, m, H11 & H12), 3.24 (1H, d, *J* = 14.9 Hz, H6a), 3.16 (1H, d, *J* = 14.9 Hz, H6b), 3.06, (1H, d, *J* = 22.2 Hz, H3a), 2.88 (1H, d, *J* = 22.1 Hz, H3b), 1.47 (3H, s, H1); **<sup>13</sup>C** (101 MHz, CDCl<sub>3</sub>) δ 209.5 (C2), 170.2 (C4), 143.0 (C7), 137.2 (C9), 132.5 (q, *J* = 33.4 Hz, C13), 130.7 (C11), 125.6 (q, *J* = 3.6 Hz, C12), 123.5 (q, *J* = 273.0 Hz, C15), 123.5 (C10), 121.0 (C8), 117.3 (q, *J* = 3.9 Hz, C14), 67.8 (C5), 40.3 (C3), 33.9 (C6), 23.9 (C1); **<sup>19</sup>F** (376 MHz, CDCl<sub>3</sub>) δ -62.9; **HRMS** (ESI) calcd for [C<sub>15</sub>H<sub>14</sub>N<sub>4</sub>O<sub>2</sub>F<sub>3</sub>]<sup>+</sup>: 339.1069, found 339.1064; **IR** ν<sub>max</sub> = 3139 (w, N-H) 3104 (w, C-H), 2941 (w, C-H), 1766 (s, C=O), 1697 (s, C=O), 1498 (m, C=C), 1480 (m, C=C).

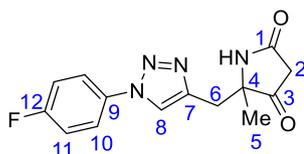
### 6.3.19 Synthesis of 5-((1-(3-chlorophenyl)-1H-1,2,3-triazol-4-yl)methyl)-5-methylpyrrolidine-2,4-dione (94)



Following general procedure 2: **85** (53.0 mg, 0.130 mmol), potassium *tert*-butoxide (21.0 mg, 0.187 mmol) and THF (1.4 mL) were refluxed for 2 hours. After work up, the crude intermediate was dissolved in acetonitrile (1 mL) and H<sub>2</sub>O (0.1 mL) and refluxed for 3 hours. The crude product was purified by flash column chromatography (silica gel; EtOAc) to give **94** (26.0 mg, 0.0853 mmol, 66%) as an off-white solid.

$R_f$  = 0.15 (EtOAc); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.86 (1H, s, H8), 7.70 (1H, t,  $J$  = 1.8 Hz, H14), 7.60 (1H, dt,  $J$  = 7.6, 1.7 Hz, H10), 7.53 (1H, br s, NH), 7.45 – 7.37 (2H, m, H11 & H12), 3.21 (1H, d,  $J$  = 15.0 Hz, H6a), 3.13 (1H, d,  $J$  = 15.0 Hz, H6b), 3.08 (1H, d  $J$  = 22.2 Hz, H2a), 2.91 (1H, d,  $J$  = 22.2 Hz, H2b), 1.46 (3H, s, H5); <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 209.5 (C3), 170.0 (C1), 142.8 (C7), 137.7 (C13), 135.7 (C9), 131.0 (C11), 129.0 (C10), 121.0 (C8), 120.7 (C14), 118.4 (C12), 67.8 (C4), 40.3 (C2), 33.9 (C6), 23.8 (C5); **HRMS** (ESI) calcd for [C<sub>14</sub>H<sub>14</sub>N<sub>4</sub>O<sub>2</sub><sup>35</sup>Cl]<sup>+</sup>: 305.0805, found 305.0814; **IR**  $\nu_{max}$  = 3136 (w, C-H) 3096 (w, C-H), 2924 (w, C-H), 1762 (m, C=O), 1693 (s, C=O), 1594 (s, C=C), 1491 (m, C=C), 1464 (m, C=C).

### 6.3.20 Synthesis of 5-((1-(4-fluorophenyl)-1H-1,2,3-triazol-4-yl)methyl)-5-methylpyrrolidine-2,4-dione (95)

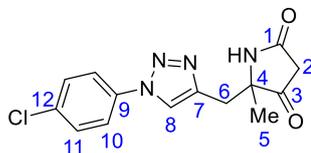


Following general procedure 2: **86** (65.0 mg, 0.165 mmol), potassium *tert*-butoxide (28.0 mg, 0.250 mmol) and THF (1.7 mL) were refluxed for 2 hours. After work up, the crude intermediate was dissolved in acetonitrile (1 mL) and H<sub>2</sub>O (0.1 mL) and refluxed for 3 hours. The crude product was purified by flash column chromatography (silica gel; EtOAc) to give **95** (35.0 mg, 0.121 mmol, 73%) as an off-white solid.

$R_f$  = 0.14 (EtOAc); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.81 (1H, s, H8), 7.67 – 7.64 (2H, m, H10), 7.59 (1H, br s, NH), 7.20 – 7.16 (2H, m, H11), 3.22 (1H, d,  $J$  = 15.0 Hz, H6a), 3.13 (1H, d,  $J$  = 15.0 Hz, H6b), 3.05 (1H, d,  $J$  = 22.2 Hz, H2a), 2.87 (1H, d,  $J$  = 22.2 Hz, H2b), 1.46 (3H, s, H5); <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) 209.6 (C3), 170.0 (C1), 162.6 (d,  $J$  = 249.4 Hz, C12), 142.7 (C7), 133.2 (C9), 122.5 (d,  $J$  = 8.6 Hz, C10), 121.2 (C8), 116.9 (d,  $J$  = 23.0 Hz, C11), 67.8 (C4), 40.4 (C2), 33.9 (C6), 23.9 (C5); <sup>19</sup>F (376 MHz, CDCl<sub>3</sub>) δ -111.8; **HRMS** (ESI) calcd for

[C<sub>14</sub>H<sub>14</sub>N<sub>4</sub>O<sub>2</sub>F]<sup>+</sup>: 289.1101, found 289.1107; IR  $\nu_{\max}$  = 3094 (w, C-H), 2902 (w, C-H), 1762 (m, C=O), 1690 (s, C=O), 1517 (s, C=C).

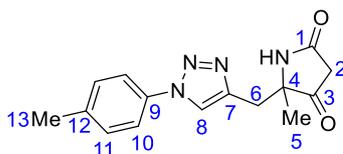
### 6.3.21 Synthesis of 5-((1-(4-chlorophenyl)-1H-1,2,3-triazol-4-yl)methyl)-5-methylpyrrolidine-2,4-dione (96)



Following general procedure 2: **87** (69.0 mg, 0.169 mmol), potassium *tert*-butoxide (28.0 mg, 0.249 mmol) and THF (1.7 mL) were refluxed for 2 hours. After work up, the crude intermediate was dissolved in acetonitrile (1 mL) and H<sub>2</sub>O (0.1 mL) and refluxed for 2 hours. The crude product was purified by flash column chromatography (silica gel; EtOAc) to give **96** (32.0 mg, 0.105 mmol, 62%) as an off-white solid.

R<sub>f</sub> = 0.18 (EtOAc); <sup>1</sup>H (400 MHz, CDCl<sub>3</sub>)  $\delta$  7.83 (1H, s, H8), 7.62 (2H, d, *J* = 8.8 Hz, H10), 7.53 (1H, br s, NH), 7.46 (2H, d, *J* = 8.8 Hz, H11), 3.21 (1H, d, *J* = 15.0 Hz, H6a), 3.13 (1H, d, *J* = 15.0 Hz, H6b), 3.06 (1H, d, *J* = 22.2 Hz, H2a), 2.89 (1H, d, *J* = 22.2 Hz, H2b), 1.46 (3H, s, H5); <sup>13</sup>C (101 MHz, CDCl<sub>3</sub>)  $\delta$  209.6 (C3), 169.9 (C1), 142.8 (C7), 135.3 (C12), 134.8 (C9), 130.1 (C11), 121.6 (C10), 120.9 (C8), 67.8 (C4), 40.3 (C2), 33.9 (C6), 23.9 (C5); HRMS (ESI) calcd for [C<sub>14</sub>H<sub>14</sub>N<sub>4</sub>O<sub>2</sub><sup>35</sup>Cl]<sup>+</sup>: 305.0805, found 305.0811; IR  $\nu_{\max}$  = 3151 (w, C-H), 3080 2921 (w, C-H), 2852 (w, C-H), 1766 (m, C=O), 1690 (s, C=O), 1501 (s, C=C), 1425 (s, C=C).

### 6.3.22 Synthesis of 5-methyl-5-((1-(*p*-tolyl)-1H-1,2,3-triazol-4-yl)methyl)pyrrolidine-2,4-dione (97)

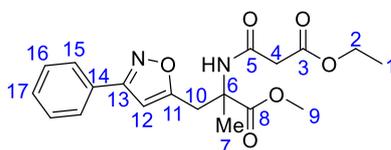


Following general procedure 2: **88** (46.7 mg, 0.125 mmol), potassium *tert*-butoxide (20.0 mg, 0.178 mmol) and THF (1.5 mL) were refluxed for 2 hours. After work up, the crude intermediate was dissolved in acetonitrile (1 mL) and H<sub>2</sub>O (0.1 mL) and refluxed for 3 hours. The crude product was purified by flash column chromatography (silica gel; EtOAc) to give **97** (25.0 mg, 0.0879 mmol, 70%) as an off-white solid.

R<sub>f</sub> = 0.17 (EtOAc); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  7.79 (1H, s, H8), 7.55 (2H, d, *J* = 8.4 Hz, H10), 7.38 (1H, br s, NH), 7.28 (2H, d, *J* = 8.3 Hz, H11), 3.20 (1H, d, *J* = 15.0 Hz, H6a), 3.11 (1H, d, *J* = 15.0 Hz, H6b), 3.06 (1H, d, *J* = 22.2 Hz, H2a), 2.91 (1H, d, *J* = 22.2 Hz, H2b), 2.41 (3H, s, H13), 1.45 (3H, s, H5); <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>)  $\delta$  209.7 (C3), 169.8 (C1), 142.4 (C7), 139.2

(C12), 134.6 (C9), 130.4 (C11), 121.0 (C8), 120.4 (C10), 67.8 (C4), 40.3 (C2), 33.8 (C6), 23.9 (C5), 21.2 (C13); **HRMS** (ESI) calcd for  $[C_{15}H_{17}N_4O_2]^+$ : 285.1352, found 285.1359; **IR**  $\nu_{\max}$  = 3138 (w, C-H), 2975 (w, C-H), 2921 (w, C-H), 1760 (m, C=O), 1693 (s, C=O), 1519 (s, C=C).

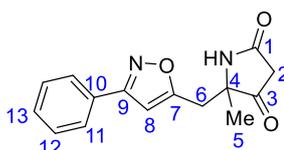
### 6.3.23 Synthesis of methyl 2-(3-ethoxy-3-oxopropanamido)-2-methyl-3-(3-phenylisoxazol-5-yl)propanoate (**98**)



To a solution of **77** (65.0 mg, 0.255 mmol) in *t*-BuOH (1.5 mL) and water (1.5 mL) was added *N*-hydroxybenzimidoyl chloride (36.0 mg, 0.231 mmol), copper(II) sulfate pentahydrate (13.7 mg, 0.0548 mmol), sodium ascorbate (15.3 mg, 0.0772 mmol) followed by potassium hydrogen carbonate (102 mg, 1.02 mmol). The reaction was stirred for 39 hours and then diluted with water and extracted with EtOAc (3 x 10 mL). The combined organic extracts were dried over  $Na_2SO_4$  and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 40 – 60%) to give **98** (46.4 mg, 0.124 mmol, 49%) as a colourless oil.

$R_f$  = 0.18 (petroleum ether/EtOAc, 3:2);  $^1H$  NMR (400 MHz,  $CDCl_3$ )  $\delta$  7.78 (2H, m, H15), 7.69 (1H, br s, NH), 7.46 – 7.43 (3H, m, H16 & H17), 6.38 (1H, s, H12), 4.13 (2H, q,  $J$  = 7.2 Hz, H2), 3.84 (3H, s, H9), 3.81 (1H, d,  $J$  = 15.0 Hz, H10a), 3.58 (1H, d,  $J$  = 15.0 Hz, H10b), 3.30 (2H, s, H4), 1.68 (3H, s, H7), 1.22 (3H, t,  $J$  = 7.2 Hz, H1);  $^{13}C$  NMR (101 MHz,  $CDCl_3$ )  $\delta$  173.4 (C8), 169.1 (C11), 168.9 (C3), 164.8 (C5), 162.4 (C13), 130.1 (C17), 129.1 (C16), 129.0 (C14), 126.9 (C15), 101.9 (C12), 61.9 (C2), 59.4 (C6), 53.4 (C9), 42.1 (C4), 32.8 (C10), 23.6 (C7), 14.1 (C1); **HRMS** (ESI) calcd for  $[C_{19}H_{23}N_2O_6]^+$ : 375.1551, found 375.1548. **IR**  $\nu_{\max}$  = 3303 (w, N-H), 2978 (w, C-H), 1737 (s, C=O), 1670 (m, C=O), 1657 (C=O), 1535 (m, C=C), 1444 (m, C=C).

### 6.3.24 Synthesis of 5-methyl-5-((3-phenylisoxazol-5-yl)methyl)pyrrolidine-2,4-dione (**93**)

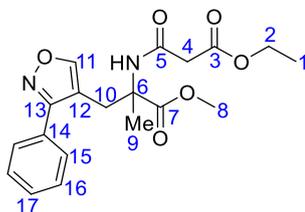


Following general procedure 2: **98** (60.0 mg, 0.160 mmol), potassium *tert*-butoxide (27.0 mg, 0.241 mmol) and THF (2 mL) were refluxed for 2 hours. After work up, the crude intermediate was dissolved in acetonitrile (2 mL) and  $H_2O$  (0.2 mL) and refluxed for 2.5 hours. The crude

product was purified by flash column chromatography (silica gel; heptane/EtOAc, 20 – 100%) to give **99** (26.0 mg, 0.0961 mmol, 38%) as a white solid.

**LCMS** MH<sup>+</sup> (270.1); **R<sub>f</sub>** = 0.16 (Petroleum ether/EtOAc, 7:13); **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.77 – 7.75 (2H, m, H11), 7.45 – 7.44 (3H, m, H12 & H13), 6.89 (1H, br s, NH), 6.42 (1H, s, H8), 3.23 (1H, d, *J* = 15.1 Hz, H6a), 3.16 (1H, d, *J* = 15.1 Hz, H6b), 3.08 (1H, d, *J* = 22.3 Hz, H2a), 2.98 (1H, d, *J* = 22.3 Hz, H2b), 1.45 (1H, s, H5); **<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>) δ 208.2 (C3), 169.7 (C1), 167.5 (C7), 162.7 (C9), 130.4 (C13), 129.1 (C12), 128.6 (C10), 126.9 (C11), 102.3 (C8), 67.0 (C4), 39.8 (C2), 35.4 (C6), 24.0 (C5); **HRMS** (ESI) calcd for [C<sub>15</sub>H<sub>15</sub>N<sub>2</sub>O<sub>3</sub>]<sup>+</sup>: 271.1083, found 271.1088; **IR** *v*<sub>max</sub> = 3088 (w, N-H), 1770 (m, N=N), 1698 (s, C=O), 1690 (s, C=O), 1607 (m, C=C), 1470 (m, C=C), 1443 (m, C=C).

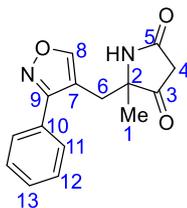
### 6.3.25 Synthesis of methyl 2-(3-ethoxy-3-oxopropanamido)-2-methyl-3-(3-phenylisoxazol-4-yl)propanoate (**100**)



A solution of **77** (65.0 mg, 0.255 mmol) in 1,2-dichloroethane (2.5 mL) was degassed for 15 mins before *N*-hydroxybenzimidoyl chloride (36.0 mg, 0.231 mmol), triethylamine (39.0 μL, 0.281 mmol) and Cp<sup>\*</sup>Ru(COD)Cl (9.70 mg, 0.0255 mmol) were added. The reaction was stirred at r.t for 42 hours. The reaction was concentrated *in vacuo* and purified by flash column chromatography (silica gel; heptane/EtOAc, 10-65% EtOAc) to give **100** (70.0 mg, 0.187 mmol, 73%) as a pale-yellow solid.

**LCMS** MH<sup>+</sup> 375.0; **R<sub>f</sub>** = 0.18 (Petroleum ether/EtOAc, 3:2); **<sup>1</sup>H** (400 MHz, CDCl<sub>3</sub>) δ 8.33 (1H, s, H11), 7.66 (1H, br s, NH), 7.57 – 7.54 (2H, m, H15), 7.49 – 7.46 (3H, m, H16 & H17), 4.20 (2H, q, *J* = 7.2 Hz, H2), 3.65 (1H, d, *J* = 15.2 Hz, H10a), 3.44 (3H, s, H8), 3.28 (1H, d, *J* = 15.1 Hz, H10b), 3.05 (2H, s, H4), 1.57 (3H, s, H9), 1.30 (3H, t, *J* = 7.1 Hz, H1); **<sup>13</sup>C** (101 MHz, CDCl<sub>3</sub>) δ 173.7 (C7), 169.2 (C3), 164.6 (C5), 162.1 (C11), 158.0 (C13), 129.6 (C17), 129.2 (C14), 129.0 (C16), 128.8 (C15), 112.8 (C12), 61.9 (C2), 60.4 (C6), 52.9 (C8), 41.7 (C4), 28.1 (C10), 23.3 (C9), 14.2 (C1); **HRMS** (ESI) calcd for [C<sub>19</sub>H<sub>23</sub>N<sub>2</sub>O<sub>6</sub>]<sup>+</sup>: 375.1556, found 375.1555; **IR** *v*<sub>max</sub> = 3308 (w, N-H), 2985 (w, C-H), 1733 (s, C=O), 1672 (s, C=O), 1655 (s, C=O), 1531 (m, C=C), 1446 (m, C=C), 1118 (s, C-O-C).

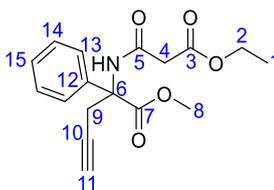
### 6.3.26 Synthesis of 5-methyl-5-((3-phenylisoxazol-4-yl)methyl)pyrrolidine-2,4-dione (101)



Following general procedure 2: **100** (52.0 mg, 0.139 mmol), potassium *tert*-butoxide (16.0 mg, 0.143 mmol) and THF (1.5 mL) were refluxed for 2 hours. After work up, the crude intermediate was dissolved in acetonitrile (1.5 mL) and H<sub>2</sub>O (0.1 mL) and refluxed for 4 hours. The crude product was purified by flash column chromatography (silica gel; heptane/EtOAc, gradient 20-100%) to give **101** (21.0 mg, 0.0777 mmol, 56%) as a white solid.

LCMS MH<sup>+</sup> 271.0; *R*<sub>f</sub> = 0.27 (EtOAc); <sup>1</sup>H (400 MHz, CDCl<sub>3</sub>) δ 8.31 (1H, s, H8), 7.51 (5H, br s, H11, H12 & H13), 6.16 (1H, br s, NH), 3.06 (1H, d, *J* = 15.0 Hz, H6a), 2.91 (1H, d, *J* = 22.2 Hz, H4a), 2.84 (1H, d, *J* = 15.0 Hz, H6b), 2.60 (1H, d, *J* = 22.2 Hz, H4b), 1.30 (3H, s, H1), ; <sup>13</sup>C (101 MHz, CDCl<sub>3</sub>) δ; 209.0 (C3), 169.4 (C5), 162.5 (C8), 157.6 (C9), 130.1 (C13), 129.3 (C10), 129.3 (C12), 128.6 (C11), 112.0 (C7), 68.4 (C2), 40.3 (C4), 30.5 (C6), 24.1 (C1), **HRMS** (ESI) calcd for [C<sub>15</sub>H<sub>15</sub>N<sub>2</sub>O<sub>3</sub>]<sup>+</sup>: 271.1083, found 271.1092; **IR** *v*<sub>max</sub> = 3259 (w, N-H), 2932 (w, C-H), 1645 (s, C=O), 1551 (s, C=O), 1444 (s, C=C).

### 6.3.27 Synthesis of methyl 2-(3-ethoxy-3-oxopropanamido)-2-phenylpent-4-ynoate (102)

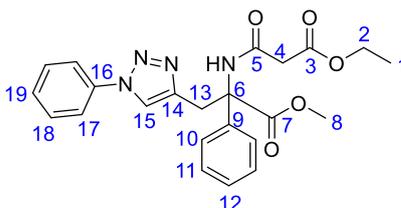


To an ice-cooled solution of **11** (100 mg, 0.490 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (5 mL) was added triethylamine (0.103 mL, 0.740 mmol), followed by ethyl 3-chloro-3-oxopropanoate (0.094 mL, 0.740 mmol) dropwise. The reaction was stirred for 3 hours 45 minutes and then diluted with NH<sub>4</sub>Cl saturated aqueous solution. The organic layer was separated, and the aqueous further extracted with CH<sub>2</sub>Cl<sub>2</sub> (2 x 10 mL). The organic extracts were combined, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated *in vacuo*. The sample was purified by flash column chromatography (silica gel; heptane/EtOAc, gradient 10 – 60%) to give **102** (146 mg, 0.460 mmol, 94 %) as a yellow oil.

*R*<sub>f</sub> = 0.36 (Petroleum ether/EtOAc, 3:2); <sup>1</sup>H (400 MHz, CDCl<sub>3</sub>) δ 8.54 (1H, br s, NH), 7.47 – 7.44 (2H, m, H13), 7.40 – 7.36 (2H, m, H14), 7.34 – 7.30 (1H, m, H15), 4.25 (2H, q, *J* = 7.2

Hz, H2), 3.72 (3H, s, H8), 3.66 (1H, dd,  $J = 16.7, 2.6$  Hz, H9a), 3.50 (1H, dd,  $J = 16.7, 2.6$  Hz, H9b), 3.43 – 3.33 (2H, m, H4), 1.94 (1H, t,  $J = 2.6$  Hz, H11), 1.32 (3H, t,  $J = 7.1$  Hz, H1);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  171.6 (C7), 169.3 (C3), 164.2 (C5), 137.6 (C12), 128.9 (C14), 128.5 (C15), 125.9 (C13), 79.5 (C10), 71.4 (C11), 64.8 (C6), 61.9 (C2), 53.6 (C8), 41.8 (C4), 24.7 (C9), 14.2 (C1); **HRMS** (ESI) calcd for  $[\text{C}_{17}\text{H}_{20}\text{NO}_5]^+$ : 318.1341, found 318.1345; **IR**  $\nu_{\text{max}} = 3287$  (w, N-H), 2983 (w, C-H), 1735 (s, C=O), 1677 (s, C=O), 1519 (s, C=C), 1498 (s, C=C), 1436 (s, C=C).

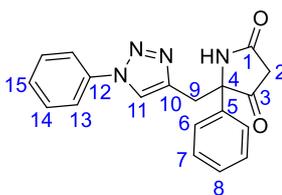
### 6.3.28 Synthesis of methyl 2-(3-ethoxy-3-oxopropanamido)-2-phenyl-3-(1-phenyl-1*H*-1,2,3-triazol-4-yl)propanoate (103)



Following general procedure 1: **102** (72.0 mg, 0.227 mmol), copper(II) sulfate pentahydrate (5.66 mg, 0.0227 mmol), sodium ascorbate (13.5 mg, 0.0681 mmol), azidobenzene (0.5M in TBME, 0.408 mL, 0.207 mmol), *t*-BuOH (1.5 mL) and water (1.5 mL) were used. The crude product was purified by flash column chromatography (silica gel; heptane/EtOAc, 20-100%) to give **103** (52.0 mg, 0.119 mmol, 52 %) as a white solid.

**LCMS**  $\text{MH}^+$  437.0;  $R_f = 0.61$  (Petroleum ether/EtOAc, 7:13);  $^1\text{H}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.07 (1H, br s, NH), 7.84 (1H, s, H15), 7.75 – 7.72 (2H, m, H17), 7.53 – 7.48 (4H, m, H10 & H11), 7.44 – 7.37 (3H, m, H18 & H19), 7.34 – 7.30 (1H, m, H12), 4.30 (1H, d,  $J = 14.3$  Hz, H13a), 4.17 – 4.10 (3H, m, H2 & H13b), 3.80 (3H, s, H8), 3.36 – 3.27 (2H, m, H4), 1.23 (3H, t,  $J = 7.2$  Hz, H1);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  172.4 (C7), 168.7 (C3), 164.1 (C5), 143.3 (C14), 138.7 (C16), 137.2 (C9), 129.8 (C18), 128.9 (C11), 128.6 (C19), 128.3 (C12), 126.0 (C10), 121.4 (C15), 120.3 (C17), 65.7 (C6), 61.9 (C2), 53.8 (C8), 42.9 (C4), 30.0 (C13), 14.1 (C1); **HRMS** (ESI) calcd for  $[\text{C}_{23}\text{H}_{25}\text{N}_4\text{O}_5]^+$ : 437.1825, found 437.1834; **IR**  $\nu_{\text{max}} = 3246$  (w, N-H), 2961 (w, C-H), 1739 (s, C=O), 1651 (s, C=O), 1598 (m, C=C), 1521 (m, C=C), 1504 (m, C=C).

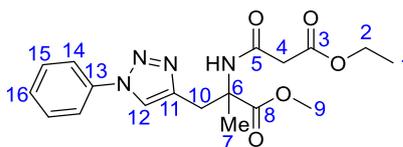
### 6.3.29 Synthesis of 5-phenyl-5-((1-phenyl-1*H*-1,2,3-triazol-4-yl)methyl)pyrrolidine-2,4-dione (104)



Following general procedure 2: **103** (43.0 mg, 0.0985 mmol), potassium *tert*-butoxide (17.0 mg, 0.151 mmol) and THF (1 mL) were refluxed for 2.5 hours. After work up, the crude intermediate was dissolved in acetonitrile (1 mL) and H<sub>2</sub>O (0.1 mL) and refluxed for 3 hours. The crude product was purified by flash column chromatography (silica gel; heptane/EtOAc, 20 – 100%) to give **104** (23.0 mg, 0.0692 mmol, 71%) as white film.

**R<sub>f</sub>** = 0.49 (EtOAc); **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ 7.76 (1H, br s, NH), 7.57 – 7.51 (4H, m, H6 & H13), 7.50 (1H, s, H11), 7.47 – 7.44 (2H, m, H14), 7.41 – 7.35 (3H, m, H7 & H15), 7.33 – 7.29 (1H, m, H8), 3.66 – 3.58 (2H, m, H9a & 9b), 3.11 (1H, d, *J* = 21.8 Hz, H2a), 2.98 (1H, d, *J* = 21.8 Hz, H2b); **<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>) δ 205.8 (C3), 169.8 (C1), 142.5 (C10), 137.4 (C5), 136.8 (C12), 129.9 (C14), 129.1 (C7), 129.0 (C15), 128.7 (C8), 125.5 (C6), 120.8 (C11), 120.5 (C13), 73.1 (C4), 40.0 (C2), 34.2 (C9); **HRMS** (ESI) calcd for [C<sub>19</sub>H<sub>17</sub>N<sub>4</sub>O<sub>2</sub>]<sup>+</sup>: 333.1346, found 333.1340; **IR**  $\nu_{\text{max}}$  = 3256 (w, N-H), 1773 (m, C=O), 1698 (m, N=N), 1667 (s, C=O), 1500 (m, C=C), 1412 (m, C=C).

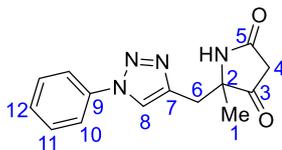
### 6.3.30 Synthesis of methyl 2-(3-ethoxy-3-oxopropanamido)-2-methyl-3-(1-phenyl-1*H*-1,2,3-triazol-4-yl)propanoate (**78**)



Following general procedure 1: **77** (400 mg, 1.57 mmol), copper(II) sulfate pentahydrate (39.1 mg, 0.156 mmol), sodium ascorbate (93.0 mg, 0.469 mmol), azidobenzene (0.5M in TBME, 2.82 mL, 1.41 mmol), *t*-BuOH (8 mL) and water (8 mL) were used. The crude product was purified by flash column chromatography (silica gel; heptane/EtOAc, 30-100%) to give **78** (450 mg, 1.20 mmol, 77%) as a white solid.

**LCMS** MH<sup>+</sup> 375.0; **R<sub>f</sub>** = 0.39 (Petroleum ether/EtOAc, 7:13); **<sup>1</sup>H** (400 MHz, CDCl<sub>3</sub>) δ 7.91 (1H, s, H12), 7.76 – 7.74 (2H, m, H14), 7.52 – 7.49 (2H, m, H15), 7.44 – 7.39 (2H, m, H16 & NH), 4.14 (2H, q, *J* = 7.2 Hz, H2), 3.82 (3H, s, H9), 3.70 (1H, d, *J* = 14.6 Hz, H10a), 3.48 (1H, d, *J* = 14.6 Hz, H10b), 3.28 (2H, s, H4), 1.69 (3H, s, H7), 1.22 (3H, t, *J* = 7.2 Hz, H1); **<sup>13</sup>C** (101 MHz, CDCl<sub>3</sub>) δ 174.0 (C8), 168.8 (C3), 164.7 (C5), 143.4 (C11), 137.2 (C13), 129.8 (C15), 128.7 (C16), 121.2 (C12), 120.4 (C14), 61.8 (C2), 60.3 (C6), 53.2 (C9), 42.7 (C4), 32.1 (C10), 23.2 (C7), 14.1 (C1); **HRMS** (ESI) calcd for [C<sub>18</sub>H<sub>23</sub>N<sub>4</sub>O<sub>5</sub>]<sup>+</sup>: 375.1668, found 375.1670; **IR**  $\nu_{\text{max}}$  = 3256 (w, N-H), 3157 (w, C-H), 3083 (w, C-H), 1732 (s, C=O), 1645 (s, C=O), 1564 (s, C=C), 1505 (s, C=C), 1426 (m, C=C).

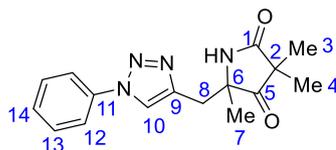
### 6.3.31 Synthesis of 5-methyl-5-((1-phenyl-1*H*-1,2,3-triazol-4-yl)methyl)pyrrolidine-2,4-dione (NM450)



Following general procedure 2: **78** (450 mg, 1.20 mmol) potassium *tert*-butoxide (202 mg, 1.80 mmol) and THF (12 mL) were refluxed for 3 hours. After work up, the crude intermediate was dissolved in acetonitrile (10 mL) and water (1 mL) and refluxed for 3 hours. The crude product was purified by flash column chromatography (silica gel; heptane/EtOAc, 20-100%) to give **NM450** (240 mg, 0.888 mmol, 74%) as a white solid.

**LCMS** MH<sup>+</sup> (271); **R<sub>f</sub>** = 0.13 (EtOAc); **<sup>1</sup>H** (400 MHz, CDCl<sub>3</sub>) δ 7.83 (1H, s, H8), 7.68 (2H, d, *J* = 7.8 Hz, H10), 7.52 – 7.48 (2H, m, H11), 7.45 – 7.41 (1H, m, H12), 7.34 (1H, br s, NH), 3.21 (1H, d, *J* = 15.0 Hz, H6a), 3.13 (1H, d, *J* = 15.0 Hz, H6b), 3.07 (1H, d, *J* = 22.2 Hz, H4a), 2.92 (1H, d, *J* = 22.2 Hz, H4b), 1.46 (3H, s, H1); **<sup>13</sup>C** (101 MHz, CDCl<sub>3</sub>) δ; 209.6 (C3), 169.7 (C5), 142.5 (C7), 136.9 (C9), 129.9 (C12), 129.0 (C11), 121.0 (C8), 120.5 (C10), 67.7 (C2), 40.3 (C4), 33.8 (C6), 23.9 (C1); **HRMS** (ESI) calcd for [C<sub>14</sub>H<sub>14</sub>N<sub>4</sub>O<sub>2</sub>]<sup>+</sup>: 271.1190, found 271.1191; **IR** ν<sub>max</sub> = 3137 (w, C-H), 2899, (w, C-H), 1759 (m, C=O), 1690 (m, C=O), 1600 (m, C=C), 1505 (m, C=C).

### 6.3.32 Synthesis of 3,3,5-trimethyl-5-((1-phenyl-1*H*-1,2,3-triazol-4-yl)methyl)pyrrolidine-2,4-dione (105)

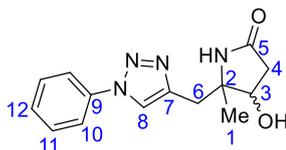


To a solution of **NM450** (25.0 mg, 0.0924 mmol) in DMF (1 mL) was added potassium carbonate (26.0 mg, 0.188 mmol) followed by iodomethane (0.0120 mL, 0.184 mmol). After stirring at r.t. for 18 hours, iodomethane (0.0120 mL, 0.184 mmol) was added and the reaction heated to 60 °C for 7 hours. The reaction was quenched with NH<sub>4</sub>Cl saturated aqueous solution and extracted with EtOAc (3 x 10 mL). The combined organic extracts were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel; heptane/EtOAc, 20 – 100%) to give **105** (10.0 mg, 0.0335 mmol, 36%) as a yellow solid.

**LCMS** MH<sup>+</sup> (299.0); **R<sub>f</sub>** = 0.33 (EtOAc); **<sup>1</sup>H** (400 MHz CDCl<sub>3</sub>) δ 7.80 (1H, s, H10), 7.70 – 7.68 (2H, m, H12), 7.54 – 7.50 (2H, m, H13), 7.46 – 7.42 (1H, m, H14), 6.96 (1H, br s NH), 3.19 – 3.10 (2H m, H8a & 8b), 1.46 (3H, s, H7), 1.30 (3H, s, H3), 1.06 (3H, s, H4); **<sup>13</sup>C** (101 MHz

CDCl<sub>3</sub>)  $\delta$  216.6 (C5), 176.7 (C1), 142.6 (C9), 136.9 (C11), 130.0 (C13), 129.1 (C14), 121.2 (C10), 120.6 (C12), 65.7 (C6), 46.3 (C2), 34.2 (C8), 25.0 (C7), 22.8 (C3), 20.4 (C4); **HRMS** (ESI) calcd for [C<sub>16</sub>H<sub>19</sub>N<sub>4</sub>O<sub>2</sub>]<sup>+</sup>: 299.1508, found 299.1512; **IR**  $\nu_{\max}$  = 3149 (w, N-H), 2971 (w, C-H), 2932 (w, C-H), 1693 (s, C=O), 1655 (s, C=O), 1596 (s, C=C), 1502 (s, C=C).

### 6.3.33 Synthesis of 4-hydroxy-5-methyl-5-((1-phenyl-1*H*-1,2,3-triazol-4-yl)methyl)pyrrolidin-2-one (106)

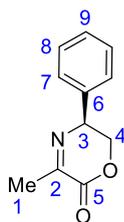


To an ice-cooled solution of **NM450** (40.0 mg, 0.148 mmol) in MeOH (1.5 mL) was added sodium borohydride (108 mg, 0.285 mmol) and the reaction stirred for 3 days at r.t. Then, NH<sub>4</sub>Cl saturated aqueous solution was added and the reaction extracted with EtOAc (3 x 10 mL). The combined organic extracts were dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated *in vacuo* to give **106** (*dr ca.* 77:23 by <sup>1</sup>H NMR, 40.0 mg, 0.147 mmol, 99%) as a white solid.

**LCMS** MH<sup>+</sup> (273.0); **R<sub>f</sub>** = 0.45, 0.39 (10% MeOH in CH<sub>2</sub>Cl<sub>2</sub>); **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>)  $\delta$  7.87 (1H, s, H8), 7.74 – 7.72 (2H, m, H10), 7.56 – 7.52 (2H, m, H11), 7.46 (1H, t, *J* = 7.4 Hz, H12), 6.69 (1H, br s, NH), 5.11 (1H, m, OH), 4.59 (1H, t, *J* = 8.1 Hz, H3'), 4.30 (1H, br s, OH'), 4.19 – 4.16 (1H, m, H3), 3.25 (1H, d, *J* = 14.4 Hz, H6a), 3.09 (3H, m, H6b, H6a' & H6b'), 2.76 (1H, dd, *J* = 17.6, 7.0 Hz, H4a), 2.66 (1H, dd, *J* = 16.9, 7.9 Hz, H4a'), 2.51 (1H, dd, *J* = 16.8, 8.5 Hz, H4b'), 2.28 (1H, dd, *J* = 17.6, 3.2 Hz, H4b), 1.22 (3H, s, H1); **<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>)  $\delta$  175.5 (C5), 144.2 (C7), 136.9 (C9), 130.0 (C11), 129.2 (C12), 121.3 (C8), 120.6 (C10), 74.0 (C3), 63.7 (C2), 39.3 (C4), 32.3, (C6), 25.7 (C1); **HRMS** (ESI) calcd for [C<sub>14</sub>H<sub>17</sub>N<sub>4</sub>O<sub>2</sub>]<sup>+</sup>: 273.1352, found 273.1344; **IR**  $\nu_{\max}$  = 3161 (br, OH), 2975 (w, C-H), 2917 (w, C-H), 1667 (s, C=O), 1597 (m, C=C), 1502 (m, C=C).

## 6.4 Synthesis of NM466 derivatives

### 6.4.1 Synthesis of (S)-3-methyl-5-phenyl-5,6-dihydro-2*H*-1,4-oxazin-2-one (113)

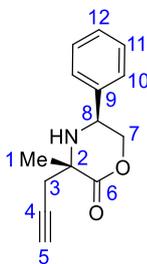


To a solution of methyl pyruvate (1.20 mL, 13.2 mmol) and molecular sieves (6.00 g) in trifluoroethanol (36 mL) was added (*S*)-phenyl glycinol (1.64 g, 12.0 mmol). The reaction was

sealed in vial and irradiated to 120 °C for 50 minutes. The reaction was filtered through celite and purified by flash column chromatography (silica gel, petroleum ether/EtOAc, 4:1) to give **113** (1.04 g, 5.47 mmol, 46%) as a white solid.

$R_f$  = 0.25 (petroleum ether/EtOAc, 4:1);  $^1\text{H}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.43 – 7.39 (2H, m, H8) 7.37 – 7.33 (3H, m, H7 & H9), 4.87 – 4.82 (1H, m, H3), 4.56 (1H, dd,  $J$  = 11.6, 4.5 Hz, H4a), 4.26 (1H, m, H4b), 2.41 (3H, d,  $J$  = 2.4 Hz, H1);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  160.4 (C2), 155.6 (C5), 136.9 (C6), 129.0 (C8), 128.4 (C9), 127.2 (C7), 71.6 (C4), 59.8 (C3), 21.9 (C1); **HRMS** (ESI) calcd for  $[\text{C}_{11}\text{H}_{12}\text{NO}_2]^+$ : 190.0868, found 190.0864; **IR**  $\nu_{\text{max}}$  = 3034 (w, C-H), 3007 (w, C-H), 1721 (s, C=O), 1641 (s, C=N), 1455 (s, C=C);  $[\alpha]_{\text{D}}^{25}$  = +214.0 (c = 0.1 in  $\text{CHCl}_3$ ).

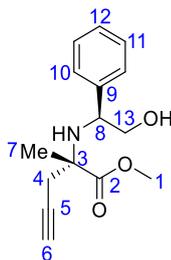
#### 6.4.2 Synthesis of (3S,5S)-3-methyl-5-phenyl-3-(prop-2-yn-1-yl)morpholin-2-one (115)



To a solution of **113** (960 mg, 5.08 mmol) in DMF (25 mL) cooled to -20 °C was added Zn (495 mg, 7.62 mmol) followed by propargyl bromide (810  $\mu\text{L}$ , 7.62 mmol). The reaction was slowly warmed to 0 °C over 3 hours 40 mins. Upon completion, the reaction was cooled to -15 °C and quenched with  $\text{NH}_4\text{Cl}$  saturated aqueous solution. The aqueous was extracted with EtOAc (3 x 20 mL) and the combined organic extracts washed with brine, dried over  $\text{Na}_2\text{SO}_4$  and concentrated *in vacuo*. Crude  $^1\text{H}$  NMR analysis indicated an 87:13 mixture of alkyne and allene isomers. The crude mixture was purified by flash column chromatography (silica gel; petroleum ether/EtOAc 22:3) to give **115** (606 mg, 2.64 mmol, 52%).

$R_f$  = 0.25 (petroleum ether/EtOAc 22:3);  $^1\text{H}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.46 – 7.44 (2H, m, H10), 7.41 – 7.33 (3H, m, H11 & H12), 4.36 – 4.26 (3H, m, H7 & H8), 3.10 (1H, dd,  $J$  = 16.9, 2.4 Hz, H3a), 2.64 (1H, dd,  $J$  = 16.9, 2.6 Hz, H3b), 2.34 (1H, br s, NH), 2.09 (1H, t,  $J$  = 2.6 Hz, H5), 1.57 (3H, s, H1);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  172.2 (C6), 137.7 (C9), 129.0 (C12), 128.9 (C11), 127.4 (C10), 78.5 (C4), 75.4 (C7), 73.0 (C5), 59.0 (C2), 53.2 (C8), 28.7 (C3), 26.8 (C1); **HRMS** (ESI) calcd for  $[\text{C}_{14}\text{H}_{16}\text{NO}_2]^+$ : 230.1181, found 230.1171; **IR**  $\nu_{\text{max}}$  = 3239 (s,  $\text{C}\equiv\text{C-H}$ ), 2916 (w, C-H), 2163 (w,  $\text{C}\equiv\text{C}$ ), 1730 (s, C=O), 1455 (s, C=C), 1430 (s, C=C), 1146 (s, C-O-C);  $[\alpha]_{\text{D}}^{25}$  = -26.0 (c = 0.1 in  $\text{CHCl}_3$ ).

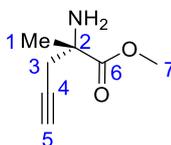
#### 6.4.2 Synthesis of methyl (S)-2-(((S)-2-hydroxy-1-phenylethyl)amino)-2-methylpent-4-ynoate (**116**)



To a solution of **115** (430 mg, 1.88 mmol) in MeOH (18.8 mL) was added  $\text{SOCl}_2$  (272  $\mu\text{L}$ , 3.75 mmol) and the reaction stirred for 2 hours at r.t. Then, the reaction was concentrated *in vacuo* and  $\text{NaHCO}_3$  saturated aqueous solution and EtOAc were added and stirred for 15 minutes. The organic layer was separated and the aqueous extracted twice more. The combined organic extracts were dried over  $\text{Na}_2\text{SO}_4$  and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc 7:3) to give **116** (285 mg, 1.09 mmol, 58%) as a colourless oil.

$R_f = 0.17$  (petroleum ether/EtOAc 7:3);  $^1\text{H}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.34 – 7.30 (2H, m, H10), 7.27 – 7.23 (3H, m, H11 & H12), 3.84 (1H, dd,  $J = 9.2, 4.7$  Hz, H13a), 3.57 (1H, dd,  $J = 10.8, 4.4$  Hz, H13b), 3.48 – 3.41 (1H, m, H8), 3.38 (3H, s, H1), 3.02 (1H, br s, OH), 2.54 (1H, dd,  $J = 16.6, 2.6$  Hz, H4a), 2.47 (1H, dd,  $J = 16.6, 2.6$  Hz, H4b), 2.30 (1H, br s, NH), 2.00 (1H, t,  $J = 2.6$  Hz, H6), 1.41 (3H, s, H7);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  175.0 (C2), 141.1 (C9), 128.6 (C10), 127.7 (C12), 127.2 (C11), 79.9 (C5), 71.4 (C6), 67.1 (C13), 60.9 (C3), 59.8 (C8), 52.1 (C1), 30.1 (C4), 22.6 (C7); **HRMS** (ESI) calcd for  $[\text{C}_{15}\text{H}_{20}\text{NO}_3]^+$ : 262.1443, found 262.1432; **IR**  $\nu_{\text{max}} = 3284$  (w,  $\text{C}\equiv\text{C-H}$ ), 2948 (w, C-H), 2342 (w,  $\text{C}\equiv\text{C}$ ), 1728 (s, C=O), 1454 (m, C=C), 1113 (s, C-O-C);  $[\alpha]_{\text{D}}^{25} = +41$  (c = 0.1 in  $\text{CHCl}_3$ ).

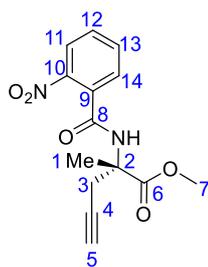
#### 6.4.3 Synthesis of methyl (S)-2-amino-2-methylpent-4-ynoate ((S)-**4**)



To a solution of **116** (250 mg, 0.957 mmol) in MeOH (3 mL) and  $\text{CH}_2\text{Cl}_2$  (6 mL) cooled to 0 °C was added  $\text{Pb}(\text{OAc})_4$  (593 mg, 1.33 mmol) and the reaction stirred for 1 hour warming to r.t. Then, 2N HCl (aq., 10 mL) was added and stirred for 10 minutes. The resulting suspension was filtered through a silica plug, eluting with MeOH and the solvents removed *in vacuo*. The aqueous residue was extracted with EtOAc (3 x 10 mL), basified with  $\text{Na}_2\text{CO}_3$  to pH 10-11 and extracted with EtOAc (3 x 10 mL). The combined organic extracts were dried over  $\text{Na}_2\text{SO}_4$  and concentrated *in vacuo* to give (S)-**4** (110 mg, 0.779 mmol, 81%) as a colourless oil.

$R_f = 0.15$  (petroleum ether/EtOAc 5:3);  $^1\text{H}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  3.72 (3H, s, H7), 2.64 (1H, dd,  $J = 2.6, 16.5$  Hz, H3a), 2.44 (1H dd,  $J = 2.6, 16.5$  Hz, H3b), 2.04 (1H, t,  $J = 2.6$  Hz, H5), 1.87 (2H, br s, NH), 1.36 (3H, s, H1);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  176.6 (C6), 79.8 (C4), 71.5 (C5), 57.5 (C2), 52.6 (C7), 31.0 (C3), 26.0 (C1); **HRMS** (ESI) calcd for  $[\text{C}_7\text{H}_{12}\text{NO}_2]^+$ : 142.0868, found 142.0862; **IR**  $\nu_{\text{max}} = 3295$  (m, N-H), 2953 (w, C-H), 2358 (w,  $\text{C}\equiv\text{C}$ ), 1732 (s,  $\text{C}=\text{O}$ ), 1116 (s, C-O-C)  $[\alpha]_{\text{D}}^{25} = -13.0$  ( $c = 0.1$  in  $\text{CHCl}_3$ ).

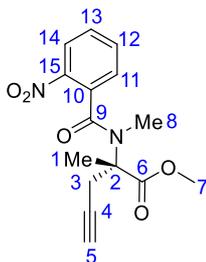
#### 6.4.4 Synthesis of methyl (S)-2-methyl-2-(2-nitrobenzamido)pent-4-ynoate ((S)-111)



To a solution of (S)-4 (35.0 mg, 0.248 mmol) in  $\text{CH}_2\text{Cl}_2$  (2.5 mL) cooled to 0 °C was added triethylamine (41.0  $\mu\text{L}$ , 0.298 mmol) followed by 2-nitrobenzoyl chloride (39.0  $\mu\text{L}$ , 0.298 mmol) and the reaction stirred for 2 hours warming to r.t. Then, the reaction was quenched with  $\text{NH}_4\text{Cl}$  saturated aqueous solution and extracted with EtOAc (3 x 10 mL). The combined organic extracts were dried over  $\text{Na}_2\text{SO}_4$  and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc 1:1) to give **111** (48.0 mg, 0.166 mmol, 67%) as a colourless oil.

$R_f = 0.22$  (petroleum ether/EtOAc 11:9);  $^1\text{H}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.09 – 8.06 (1H, m, H11), 7.70 – 7.66 (1H, m, H13), 7.61 – 7.57 (2H, m, H12 & H14), 6.74 (1H, br s, NH), 3.85 (3H, s, H7), 3.34 (1H, dd,  $J = 16.9, 2.6$  Hz, H3a), 2.98 (1H, dd,  $J = 16.9, 2.6$  Hz, H3b), 2.06 (1H, t,  $J = 2.6$  Hz, H5), 1.78 (3H, s, H1);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  173.3 (C6), 165.7 (C8), 146.5 (C10), 133.8 (C13), 132.8 (C9), 130.7 (C14), 129.0 (C12), 124.7 (C11), 79.5 (C4), 71.4 (C5), 60.1 (C2), 53.4 (C7), 26.6 (C3), 22.4 (C1); **HRMS** (ESI) calcd for  $[\text{C}_{14}\text{H}_{15}\text{N}_2\text{O}_5]^+$ : 291.0981; found: 291.0993; **IR**  $\nu_{\text{max}} = 3280$  (w,  $\text{C}\equiv\text{C-H}$ ), 3187 (w, C-H), 3005 (w, C-H), 2159 (w,  $\text{C}\equiv\text{C}$ ) 1751 (s,  $\text{C}=\text{O}$ ), 1640 (s,  $\text{C}=\text{O}$ ), 1563 (s,  $\text{C}=\text{C}$ ), 1524 (s,  $\text{C}=\text{C}$ );  $[\alpha]_{\text{D}}^{25} = -35.0$  ( $c = 0.1$  in  $\text{CHCl}_3$ ).

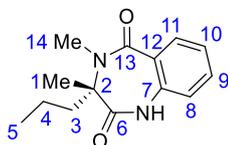
#### 6.4.5 Synthesis of methyl (S)-2-methyl-2-(N-methyl-2-nitrobenzamido)pent-4-ynoate ((S)-112)



To a solution of **111** (37.0 mg, 0.127 mmol) in DMF (1.3 mL) at 0 °C was added NaH (60% dispersion in mineral oil, 7.60 mg, 0.190 mmol) and the reaction stirred for 15 mins before MeI (16.0  $\mu$ L, 0.254 mmol) was added. After 1 hr stirring at r.t. the reaction was quenched with  $\text{NH}_4\text{Cl}$  saturated aqueous solution and extracted with EtOAc (3 x 10 mL). The combined organic extracts were washed with brine, dried over  $\text{Na}_2\text{SO}_4$  and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc 1:1) to give **112** (32.0 mg, 0.105 mmol, 83%) as a white solid.

$R_f$  = 0.27 (petroleum ether/EtOAc 1:1);  $^1\text{H}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.19 (1H, dd,  $J$  = 8.3, 0.7 Hz, H14), 7.72 (1H, td,  $J$  = 7.5, 1.6 Hz, H12), 7.59 – 7.55 (1H, m, H13), 7.43 (1H, d,  $J$  = 7.4 Hz, H11), 3.78 (3H, s, H7), 3.50 (1H, dd,  $J$  = 17.3, 2.2 Hz, H3a), 3.02 (3H, s, H8), 2.86 (1H, br s, H3b), 2.07 (1H, t  $J$  = 2.6 Hz, H5), 1.73 (3H, s, H1);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  173.3 (C6), 168.1 (C9), 145.0 (C15), 134.6 (C12), 133.3 (C10), 129.8 (C13), 128.0, (C11), 124.7 (C14), 80.3 (C4), 70.9 (C5), 61.9 (C2), 52.6 (C7), 34.4 (C8), 25.0 (C3), 21.4 (C1); **HRMS** (ESI) calcd for  $[\text{C}_{15}\text{H}_{17}\text{N}_2\text{O}_5]^+$ : 305.1137, found 305.1130; **IR**  $\nu_{\text{max}}$  = 3283 (w, N-H), 2923 (w, C-H), 1740 (s, C=O), 1640 (s, C=O), 1530 (s, C=C);  $[\alpha]_D^{25}$  = -121.0 (c = 0.1 in  $\text{CHCl}_3$ ).

#### 6.4.6 Synthesis of (S)-3,4-dimethyl-3-propyl-3,4-dihydro-1H-benzo[e][1,4]diazepine-2,5-dione ((S)-NM466)



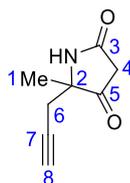
To a degassed solution of (S)-**112** (23.0 mg, 0.0755 mmol) in MeOH (1 mL) was added Pd/C (8.00 mg, 800  $\mu$ mol) and the reaction stirred under an atmosphere of  $\text{H}_2$  for 2 hours 15 minutes. The reaction was filtered through celite, eluting with MeOH and the solvents removed *in vacuo*. The crude intermediate was dissolved in DMF (0.6 mL) and NaH (60% dispersion in mineral oil, 6.00 mg, 0.151 mmol) was added and the reaction stirred for 1.5 hours. Then, the reaction was quenched with  $\text{NH}_4\text{Cl}$  saturated aqueous solution and extracted with EtOAc (3 x 10 mL). The combined organic extracts were washed with brine, dried over  $\text{Na}_2\text{SO}_4$  and concentrated

*in vacuo*. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 3:2) to give (*S*)-**NM466** (13.6 mg, 0.0552 mmol, 73%) as a white solid.

$R_f$  = 0.14 (petroleum ether/EtOAc, 3:2);  $^1\text{H}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  8.25 (1H, br s, NH), 7.97 (1H, dd,  $J$  = 1.4, 7.9 Hz, H11), 7.44 (1H, td,  $J$  = 1.5, 11.4 Hz, H9), 7.23 – 7.19 (1H, m, H10), 6.90 (1H, d,  $J$  = 8.2 Hz, H8), 3.23 (3H, s, H14), 1.68 (3H, s, H1), 1.63 – 1.57 (1H, H3a), 1.49 – 1.41 (1H, m, H3b), 1.26 – 1.19 (2H, m, H4), 0.65 (3H, t,  $J$  = 7.3 Hz, H5);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  173.3 (C6), 168.3 (C13), 134.9 (C7), 132.5 (C9), 131.4 (C11), 127.8 (C12), 124.8 (C10), 119.1 (C8), 62.9 (C2), 38.4 (C3), 32.6 (C14), 23.1 (C1), 17.4 (C4), 14.1 (C5); **HRMS** (ESI) calcd for  $[\text{C}_{14}\text{H}_{19}\text{N}_2\text{O}_2]^+$ : 247.1441, found 247.1430; **IR**  $\nu_{\text{max}}$  = 3151 (w, C-H), 2958 (w, C-H), 2873 (w, C-H), 1676 (s, C=O), 1603 (s, C=O), 1487 (s, C=C), 1432 (m, C=C);  $[\alpha]_{\text{D}}^{25}$  = +269.0 ( $c$  = 0.1 in  $\text{CHCl}_3$ ).

## 6.5 Synthesis of alkynes for high-throughput analogue synthesis

### 6.5.1 Synthesis of 5-methyl-5-(prop-2-yn-1-yl)pyrrolidine-2,4-dione (**60**)

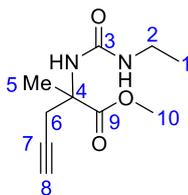


Following general procedure 2: **77** (250 mg, 0.980 mmol), potassium tert-butoxide (165 mg, 1.47 mmol) and THF (10 mL) were refluxed for 2 hours. After work up, the crude intermediate was dissolved in acetonitrile (10 mL) and  $\text{H}_2\text{O}$  (1 mL) and refluxed for 4 hours. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 1:2) to give **60** (110 mg, 0.724 mmol, 74%) as a yellow solid.

$R_f$  = 0.22 (petroleum ether/EtOAc; 1:2);  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  7.26 (1H, br s, NH), 3.08 (2H, d,  $J$  = 2.3 Hz, H4), 2.58 (1H, dd,  $J$  = Hz, H6a), 2.50 (1H, dd,  $J$  = HZ, H6a), 2.10 (1H, t,  $J$  = 2.6 Hz, H8), 1.42 (3H, s, H1);  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ )  $\delta$  208.3 (C5), 170.5 (C3) 78.0 (C7), 72.5 (C8), 67.1 (C2), 40.6 (C4), 28.9 (C6), 23.1 (C1); **HRMS** (ESI) calcd for  $[\text{C}_8\text{H}_{10}\text{NO}_2]^+$ : 152.0706, found 152.0704. **IR**  $\nu_{\text{max}}$  = 3243 (w,  $\text{C}\equiv\text{C-H}$ ), 3155 (w, C-H), 2160 (w,  $\text{C}\equiv\text{C}$ ), 1769 (s, C=O), 1682 (s, C=O).

These data are in accordance with the literature.<sup>232</sup>

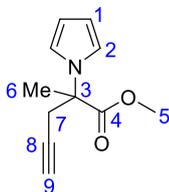
### 6.5.2 Synthesis of methyl 2-(3-ethylureido)-2-methylpent-4-ynoate (138)



To a solution of amine **4** (250 mg, 1.77 mmol) in  $\text{CH}_2\text{Cl}_2$  (18 mL) at 0 °C was added ethylisocyanate (280  $\mu\text{L}$ , 3.54 mmol) and the reaction stirred for 18 hours. Then, the solvent was removed *in vacuo* and the product purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 2:3) to give **138** (345 mg, 1.63 mmol, 92%) as a white solid.

$R_f$  = 0.30 (Petroleum ether/EtOAc, 2:3);  $^1\text{H}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  5.39 (1H, br s, NH), 4.87 (1H, br t,  $J$  = 5.3 Hz, NH), 3.74 (3H, s, H10), 3.20 – 3.13 (2H, m H2), 3.00 (1H, dd,  $J$  = 16.9, 2.7 Hz, H6a), 2.90 (1H, dd,  $J$  = 16.9, 2.7 Hz, H6b), 1.99 (1H, t,  $J$  = 2.7 Hz, H8), 1.54 (3H, s, H5), 1.10 (3H, t,  $J$  = 7.2 Hz, H1);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  174.6 (C9), 157.1 (C3), 80.3 (C7), 71.0 (C8), 58.2 (C4), 52.8 (C10), 35.3 (C2), 27.4 (C6), 23.7 (C5), 15.4 (C1); **HRMS** (ESI) calcd for  $[\text{C}_{10}\text{H}_{17}\text{N}_2\text{O}_3]^+$ : 213.1239, found 213.1240; **IR**  $\nu_{\text{max}}$  = 3385 (m, N-H), 3324 (m, N-H), 3267 (m,  $\text{C}\equiv\text{C-H}$ ), 2974 (w, C-H), 1719 (s,  $\text{C}=\text{O}$ ), 1634 (s,  $\text{C}=\text{O}$ ), 1557 (s,  $\text{C}=\text{O}$ ), 1130 (s, C-O-C).

### 6.5.3 Synthesis of methyl 2-methyl-2-(1H-pyrrol-1-yl)pent-4-ynoate (139)



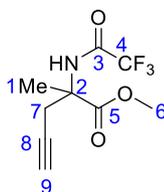
To a solution of amine **4** (250 mg, 1.77 mmol) in 1,2-dichloroethane (9.8 mL),  $\text{H}_2\text{O}$  (5.9 mL) and acetic acid (1.95 mL) was added NaOAc (174 mg, 2.13 mmol). The reaction was sealed in a vial and heated to 90 °C for 5 minutes, then 2,5-dimethoxy tetrahydrofuran (252  $\mu\text{L}$ , 1.95 mmol) was added and the reaction heated at the same temperature for 19 hours. The reaction was cooled to room temperature, diluted with EtOAc and washed with brine. The organic layer was dried over  $\text{Na}_2\text{SO}_4$  and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 9:1) to yield **139** (232 mg, 1.21 mmol, 68%) as a white solid.

$R_f$  = 0.56 (petroleum ether/EtOAc, 4:1);  $^1\text{H}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  6.82 (2H, t,  $J$  = 2.2 Hz, H2), 6.21 (2H, t,  $J$  = 2.2 Hz, H1), 3.74 (3H, s, H5), 3.07 (1H, dd,  $J$  = 16.8, 2.6 Hz, H7a), 2.99 (1H, dd,  $J$  = 16.8, 2.6 Hz, H7b), 2.04 (1H, t,  $J$  = 2.6 Hz, H9), 1.93 (3H, s, H6);  $^{13}\text{C}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta$  172.0 (C4), 118.8 (C2), 108.8 (C1), 78.5 (C8), 72.3 (C9), 63.1 (C3), 53.2 (C5), 30.1 (C7),

22.9 (C6); **HRMS** (ESI) calcd for  $[C_{11}H_{14}NO_2]^+$  192.1019, found 192.1012; **IR**  $\nu_{max}$  = 3254 (s, C≡C-H), 2952 (w, C-H), 2159 (w, C≡C), 1788 (s, C=O), 1108 (s, C-O-C).

These data are in accordance with the literature.<sup>232</sup>

### 6.5.2 Synthesis of methyl 2-methyl-2-(2,2,2-trifluoroacetamido)pent-4-ynoate (**140**)



To a solution of **4** (210 mg, 1.49 mmol) in  $CH_2Cl_2$  (15 mL) at 0 °C was added trifluoroacetic anhydride (414  $\mu$ L, 2.97 mmol) and triethylamine (414  $\mu$ L, 2.97 mmol) and the reaction stirred for 16 hours. The reaction was diluted with  $NH_4Cl$  saturated aqueous solution (20 mL) and extracted with  $CH_2Cl_2$  (3 x 20 mL). The combined organic extracts were dried over  $Na_2SO_4$  and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 4:1) to give **140** (314 mg, 1.32 mmol, 89%) as a colourless oil.

$R_f$  = (0.44, petroleum ether/EtOAc, 4:1);  $^1H$  (400 MHz,  $CDCl_3$ )  $\delta$  7.30 (1H, br s, NH), 3.84 (3H, s, H6), 3.23 (1H, dd,  $J$  = 17.0, 2.6 Hz, H7a), 2.85 (1H, dd,  $J$  = 17.0, 2.6 Hz, H7b), 2.04 (1H, t,  $J$  = 2.6 Hz, H9), 1.69 (3H, s, H1);  $^{13}C$  (101 MHz,  $CDCl_3$ )  $\delta$  172.3 (C5), 156.3 (q,  $J$  = 37.4 Hz, C3), 115.5 (q,  $J$  = 288.4 Hz, C4), 77.8 (C8), 72.1 (C9), 60.2 (C2), 53.8 (C6), 26.2 (C7), 22.2 (C1);  $^{19}F$  (376 MHz,  $CDCl_3$ )  $\delta$  -76.2; **HRMS** (ESI) calcd for  $[C_9H_{11}NO_3F_3]^+$ : 238.0686, found 238.0676; **IR**  $\nu_{max}$  = 3296 (w, C≡C-H), 2956 (w, C-H), 1750 (s, C=O), 1717 (s, C=O), 1150 (s, C-O-C).

### 6.6 X-Ray crystallography and library screening

CFI<sub>25</sub> protein was expressed and purified at the Structural Genomics Consortium, Oxford, by Dr Michael Fairhead following the in-house protocol "PREPX" (Parallel rapid expression and purification of proteins for crystallography). The final protein buffer consisting of 10 mM HEPES pH 7.5, 0.5 M NaCl, 5% glycerol, 0.5 mM TCEP. Crystals were grown in SWISSCI 3 Lens crystallization sitting-drop plates at 20 °C by mixing 50–100 nl of 20 mg/ml protein solution in a 1:1 ratio with 50–100 nl reservoir solution consisting of 0.1 M sodium acetate pH 4.2–4.8, 0.1 M zinc acetate, 5–16%(w/v) PEG 3000 and placing the drops over 20 ml reservoir solution. First crystals appeared after 2 days but the high-resolution diffracting ones appeared in 5–6 days.

Activin A was cloned, expressed and purified as previously described.<sup>276</sup> Activin A crystals were grown from 200 nL protein (3.3 mg/mL dissolved in 20 % MeCN), 50 nL seed solution and 200 nL well solution. The well solution consisted of 1.4 M ammonium sulfate, 100 mM HEPES pH 7.4, 8 % DMSO, and the seed solution consisted of 200 nL of 1.55 M ammonium sulfate, 4 % PEG300, 100 mM HEPES pH 7.4, 8 % DMSO solution and 200 nL of mature activin A (3.6 mg/mL) dissolved in 20 % MeCN. The resulting crystals were soaked with a 40 mM compound solution.

All compounds were dissolved in DMSO at a concentration of 500mM. In cases where compounds could not be dissolved at this concentration, the solution was further diluted to 250mM. Crystal soaking was performed by acoustic transfer using a Labcyte Echo 550. CFI<sub>25</sub> crystals diffracted to 1.5–1.8 Å resolution in space group P3221, with typical unit-cell parameters  $a = 59.6$ ,  $c = 213.6$  Å,  $\alpha = 90.00$ ,  $\gamma = 120.00$  and with two CFI<sub>25</sub> molecules in the asymmetric unit. X-ray diffraction data were collected on beamline I04-1 at Diamond Light Source and were processed using the Diamond autoprocessing pipeline, which utilizes xia2,<sup>332</sup> DIALS,<sup>333</sup> XDS,<sup>334</sup> POINTLESS,<sup>335</sup> and CCP4.<sup>336</sup> Electron-density maps were generated using XChemExplorer *via* DIMPLE.<sup>337</sup> Ligand restraints were generated with AceDRG<sup>338</sup> and ligand binding was detected with PanDDA,<sup>255</sup> with ligands built into PanDDA event maps. Iterative refinement and manual model correction was performed using REFMAC<sup>339</sup> and Coot,<sup>340</sup> respectively



## 7 References

- 1 N. Westwood and A. Nelson, in *Chemical and Biological Synthesis: Enabling Approached for Understanding biology*, eds. N. Westwood and A. Nelson, Royal Society of Chemistry, 2018, pp. 1–7.
- 2 C. J. O' Connor, L. Laraia and D. R. Spring, *Chem. Soc. Rev.*, 2011, **40**, 4332–4345.
- 3 J. A. Zorn and J. A. Wells, *Nat. Chem. Biol.*, 2010, **6**, 179–188.
- 4 D. R. Spring, *Chem. Soc. Rev.*, 2005, **34**, 472.
- 5 W. R. J. D. Galloway, A. Isidro-Llobet and D. R. Spring, *Nat. Commun.*, 2010, **1**, 1–13.
- 6 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, *Nat. Chem. Biol.*, 2015, **11**, 536–541.
- 7 P. J. Carter and G. A. Lazar, *Nat. Rev. Drug Discov.*, 2017, **17**, 197–223.
- 8 B. Munos, *Nat. Rev. Drug Discov.*, 2009, **8**, 959–968.
- 9 N. Mateu, S. L. Kidd, T. J. Osberger, H. L. Stewart, S. Bartlett, W. R. J. D. Galloway, A. J. P. North, H. F. Sore and D. R. Spring, in *Chemical and Biological Synthesis: Enabling Approached for Understanding Biology*, eds. N. Westwood and A. Nelson, Royal Society of Chemistry, 2018, pp. 8–44.
- 10 J. W. Scannell, A. Blanckley, H. Boldon and B. Warrington, *Nat. Rev. Drug Discov.*, 2012, **11**, 191–200.
- 11 R. H. Griesenauer and M. S. Kinch, *Drug Discov. Today*, 2017, **22**, 1593–1597.
- 12 C. Cookson, Pharma industry's return on R&D investment falls sharply, <https://www.ft.com/content/b020be56-e00a-11e7-a8a4-0a1e63a52f9c>, (accessed 29 October 2018).
- 13 A. Mullard, *Nat. Rev. Drug Discov.* , 2018, **17**, 81–85.

- 14 J. A. DiMasi, H. G. Grabowski and R. W. Hansen, *J. Health Econ.*, 2016, **47**, 20–33.
- 15 C. H. Wong, K. W. Siah and A. W. Lo, *Biostatistics*, 2018, doi.org/10.1093/biostatistics/kxx069.
- 16 K. Smietana, M. Siatkowski and M. Møller, *Nat. Rev. Drug Discov.*, 2016, **15**, 379–380.
- 17 R. K. Harrison, *Nat. Rev. Drug Discov.*, 2016, **15**, 817–818.
- 18 I. Khanna, *Drug Discov. Today*, 2012, **12**, 1088–1102.
- 19 I. Kola and J. Landis, *Nat. Rev. Drug Discov.*, 2004, **3**, 711–716.
- 20 S. L. Schreiber, J. D. Kotz, M. Li, J. Aubé, C. P. Austin, J. C. Reed, H. Rosen, E. L. White, L. A. Sklar, C. W. Lindsley, B. R. Alexander, J. A. Bittker, P. A. Clemons, A. De Souza, M. A. Foley, M. Palmer, A. F. Shamji, M. J. Wawer, O. McManus, M. Wu, B. Zou, H. Yu, J. E. Golden, F. J. Schoenen, A. Simeonov, A. Jadhav, M. R. Jackson, A. B. Pinkerton, T. D. Y. Chung, P. R. Griffin, B. F. Cravatt, P. S. Hodder, W. R. Roush, E. Roberts, D. H. Chung, C. B. Jonsson, J. W. Noah, W. E. Severson, S. Ananthan, B. Edwards, T. I. Oprea, P. J. Conn, C. R. Hopkins, M. R. Wood, S. R. Stauffer and K. A. Emmitte, *Cell*, DOI:10.1016/j.cell.2015.05.023.
- 21 M. Lindsay, *Nat. Rev. Drug Discov.*, 2003, **2**, 831–8.
- 22 M. R. Nelson, T. Johnson, L. Warren, A. R. Hughes, S. L. Chissoe, C.-F. Xu and D. M. Waterworth, *Nat Rev Genet*, 2016, **17**, 197–206.
- 23 D. Altshuler, M. J. Daly and E. S. Lander, *Science*, 2008, **322**, 881–888.
- 24 M. E. Bunnage, E. L. P. Chekler and L. H. Jones, *Nat. Chem. Biol.*, 2013, **9**, 195–199.
- 25 J. Blagg and P. Workman, *Cancer Cell*, 2017, **32**, 9–25.
- 26 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**, 529–533.
- 27 P. Filippakopoulos and S. Knapp, *Nat. Rev. Drug Discov.*, 2014, **13**, 337–356.
- 28 O. Mirguet, R. Gosmini, J. Toum, C. A. Clément, M. Barnathan, J.-M. Brusq, J. E. Mordaunt, R. M. Grimes, M. Crowe, O. Pineau, M. Ajakane, A. Daugan, P. Jeffrey, L. Cutler, A. C. Haynes, N. N. Smithers, C.-W. Chung, P. Bamborough, I. J. Uings, A. Lewis, J. Witherington, N. Parr, R. K. Prinjha and E. Nicodème, *J. Med. Chem.*, 2013,

- 56**, 7501–7515.
- 29 R. Santos, O. Ursu, A. Gaulton, A. Patrícia Bento, R. S. Donadi, C. G. Bologna, A. Karlsson, B. Al-Lazikani, A. Hersey, T. I. Oprea and J. P. Overington, *Nat. Rev. Drug Discov.*, 2017, **16**, 19–34.
- 30 P. Agarwal, P. Sanseau and L. R. Cardon, *Nat. Rev. Drug Discov.*, 2013, **12**, 575–576.
- 31 M. Congreve, C. W. Murray and T. L. Blundell, *Drug Discov. Today*, 2005, **10**, 895–907.
- 32 W. P. Walters and M. Namchuk, *Nat. Rev. Drug Discov.*, 2003, **2**, 259–266.
- 33 K. M. G. O. Connell, W. R. J. D. Galloway, B. M. Ibbeson, A. Isidro-Ilobet, C. J. O. Connor and D. R. Spring, in *Solid-Phase Organic Synthesis: Concepts, Strategies, and Applications*, ed. Y. L. P. H. Toy, Wiley, 2012, pp. 131–150.
- 34 D. S. Tan, *Nat. Chem. Biol.*, 2005, **1**, 74–84.
- 35 J. Clardy and C. Walsh, *Nature*, 2004, **432**, 829–837.
- 36 T. Rodrigues, D. Reker, P. Schneider and G. Schneider, *Nat. Chem.*, 2016, 531–541.
- 37 C. V. Galliford and K. A. Scheidt, *Angew. Chemie Int. Ed.*, 2007, **46**, 8748–8758.
- 38 A. L. Harvey, *Drug Discov. Today*, 2008, **13**, 894–901.
- 39 D. J. Newman and G. M. Cragg, *J. Nat. Prod.*, 2016, **79**, 629–661.
- 40 R. J. Spandl, A. Bender and D. R. Spring, *Org. Biomol. Chem.*, 2008, **6**, 1149–1158.
- 41 A. L. Harvey, R. Edrada-Ebel and R. J. Quinn, *Nat. Rev. Drug Discov.*, 2015, **14**, 111–129.
- 42 G. M. Cragg and D. J. Newman, *Biochim. Biophys. Acta - Gen. Subj.*, 2014, **1830**, 3670–3695.
- 43 I. Paterson, M. Xuan and S. M. Dalby, *Angew. Chemie Int. Ed.*, 2014, **53**, 7286–7289.
- 44 I. Paterson and E. Anderson, *Science*, 2005, **310**, 451.
- 45 D. R. Spring, *Org. Biomol. Chem.*, 2003, **1**, 3867–3870.
- 46 G. R. Crabtree, S. L. Schreiber, A. Borchardt, S. D. Liberles, S. R. Biggar, A. J. You, R. J. Jackman, G. M. Whitesides, A. Furka, F. Sebestyén, M. Asgedom, G. Dibó, K. S. Lam, S. E. Salmon, E. M. Hersh, V. J. Hruby, W. M. Kazmierski, R. J. Knapp, H. B. Wood, B. Ganem, D. A. McGowan and G. A. Berchtold, *Stereoselective Synthesis of*

*over Two Million Compounds Having Structural Features Both Reminiscent of Natural Products and Compatible with Miniaturized Cell-Based Assays*, 1998, vol. 120.

- 47 S. R. Langdon, N. Brown and J. Blagg, *J. Chem. Inf. Model.*, 2011, **51**, 2174–2185.
- 48 M. Dean, G. Brown and J. Boströ, *J. Med. Chem.*, 2015, **59**, 4443–4458.
- 49 R. S. Bohacek, C. McMartin and W. C. Guida, *Med. Res. Rev.*, 1996, **16**, 3–50.
- 50 A. H. Lipkus, Q. Yuan, K. A. Lucas, S. A. Funk, W. F. Bartelt, R. J. Schenck and A. J. Trippe, *J. Org. Chem.*, 2008, **73**, 4443–4451.
- 51 M. Dow, M. Fisher, T. James, F. Marchetti and A. Nelson, *Org. Biomol. Chem.*, 2012, **10**, 17–28.
- 52 P. D. Leeson and B. Springthorpe, *Nat. Rev. Drug Discov.*, 2007, **6**, 881–90.
- 53 C. Lipinski and A. Hopkins, *Nature*, 2004, **432**, 855–861.
- 54 C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, *Adv. Drug Deliv. Rev.*, 2001, **46**, 3–26.
- 55 T. Kogej, N. Blomberg, P. J. Greasley, S. Mundt, M. J. Vainio, J. Schamberger, G. Schmidt and J. Hüser, *Drug Discov. Today*, 2013, **18**, 1014–1024.
- 56 D. J. Payne, M. N. Gwynn, D. J. Holmes and D. L. Pompliano, *Nat. Rev. Drug Discov.*, 2007, **6**, 29–40.
- 57 W. R. J. D. Galloway and D. R. Spring, *Divers. Oriented Synth.*, 2013, **1**, 21–28.
- 58 T. J. Ritchie and S. J. F. Macdonald, *Drug Discov. Today*, 2009, **14**, 1011–1020.
- 59 P. A. Clemons, N. E. Bodycombe, H. A. Carrinski, J. A. Wilson, A. F. Shamji, B. K. Wagner, A. N. Koehler and S. L. Schreiber, *PNAS*, 2010, **107**, 18787–18792.
- 60 F. Lovering, J. Bikker and C. Humblet, *J. Med. Chem.*, 2009, **52**, 6752–6756.
- 61 P. J. Hajduk, W. R. J. D. Galloway and D. R. Spring, *Nature*, 2011, **470**, 42–43.
- 62 S. L. Schreiber, *Nature*, 2009, **457**, 153–154.
- 63 R. Doveston, S. Marsden and A. Nelson, *Drug Discov. Today*, 2014, **19**, 813–819.
- 64 A. Nadin, C. Hattotuwigama and I. Churcher, *Angew. Chemie Int. Ed.*, 2012, **51**, 1114–1122.
- 65 T. James, P. Maclellan, G. M. Burslem, I. Simpson, J. A. Grant, S. Warriner, V. Sridharan and A. Nelson, *Org. Biomol. Chem.*, 2014, **12**, 2584–2591.
- 66 M. Lüthy, M. C. Wheldon, C. Haji-Cheteh, M. Atobe, P. S. Bond, P. O'brien, R. E.

- Hubbard and I. J. S. Fairlamb, *Bioorg. Med. Chem.*, 2015, **23**, 2680–2694.
- 67 H. van Hattum and H. Waldmann, *J. Am. Chem. Soc.*, 2014, **136**, 11853–9.
- 68 S. Basu, B. Ellinger, S. Rizzo, C. Deraeve, M. Schürmann, H. Preut, H.D. Arndt and H. Waldmann, *PNAS*, 2011, **108**, 6805–6810.
- 69 S. Wetzel, R. S. Bon, K. Kumar and H. Waldmann, *Angew. Chemie - Int. Ed.*, 2011, **50**, 10800–10826.
- 70 A. Karawajczyk, F. Giordanetto, J. Benningshof, D. Hamza, T. Kalliokoski, K. Pouwer, R. Morgentin, A. Nelson, G. Müller, A. Piechot and D. Tzalis, *Drug Discov. Today*, 2015, **20**, 1310–1316.
- 71 D. Lee, J. K. Sello and S. L. Schreiber, *Org. Lett.*, 2000, **2**, 709–712.
- 72 S. L. Schreiber, *Science*, 2000, **287**, 1964–1969.
- 73 W. R. J. D. Galloway, J. E. Stokes and D. Spring, in *Small Molecule Medicinal Chemistry: Strategies and Technologies*, eds. W. Czechtizky and P. Hamley, John Wiley & Sons, Inc., 2016, pp. 79–101.
- 74 W. R. J. D. Galloway, A. Bender, M. Welch and D. R. Spring, *Chem. Commun.*, 2009, 2446–2462.
- 75 W. R. Galloway and D. R. Spring, *Expert Opin. Drug Discov.*, 2009, **4**, 467–72.
- 76 M. D. Burke and S. L. Schreiber, *Angew. Chemie Int. Ed.*, 2004, **43**, 46–58.
- 77 R. J. Spandl, M. Díaz-Gavilán, K. M. G. O'Connell, G. L. Thomas and D. R. Spring, *Chem. Rec.*, 2008, **8**, 129–142.
- 78 M. D. Burke, E. M. Berger and S. L. Schreiber, *Science*, 2003, **302**, 613–618.
- 79 D. Morton, S. Leach, C. Cordier, S. Warriner and A. Nelson, *Angew. Chemie Int. Ed.*, 2009, **48**, 104–109.
- 80 T. E. Nielsen and S. L. Schreiber, *Angew. Chemie Int. Ed.*, 2008, **47**, 48–56.
- 81 H. S. G. Beckmann, F. Nie, C. E. Hagerman, H. Johansson, Y. S. Tan, D. Wilcke and D. R. Spring, *Nat. Chem.*, 2013, **5**, 861–867.
- 82 W. R. J. D. Galloway, M. Díaz-Gavilán, A. Isidro-Llobet and D. R. Spring, *Angew. Chemie Int. Ed.*, 2009, **48**, 1194–1196.
- 83 J. E. Biggs-Houck, A. Younai and J. T. Shaw, *Curr. Opin. Chem. Biol.*, 2010, **14**, 371–382.

- 84 R. J. Spandl, H. Rudyk and D. R. Spring, *Chem. Commun.*, 2008, 3001–3003.
- 85 G. Muncipinto, in *Diversity-Oriented Synthesis: Basics and Applications in Organic Synthesis, Drug Discovery, and Chemical Biology*, ed. A. Trabocchi, John Wiley & Sons Inc, 2013, vol. 8, pp. 129–142.
- 86 W. Tan, X.-T. Zhu, S. Zhang, G.-J. Xing, R.-Y. Zhu and F. Shi, *RSC Adv.*, 2013, **3**, 10875–10886.
- 87 E. Jacoby, A. Schuffenhauer, M. Popov, K. Azzaoui, B. Havill, U. Schopfer, C. Engeloch, J. Stanek, P. Acklin, P. Rigollier, F. Stoll, G. Koch, P. Meier, D. Orain, R. Giger, J. Hinrichs, K. Malagu, J. Zimmermann and H.-J. Roth, *Curr. Top. Med. Chem.*, 2005, **5**, 397–411.
- 88 B. M. Ibbeson, L. Laraia, E. Alza, C. J. O' Connor, Y. S. Tan, H. M. L. Davies, G. McKenzie, A. R. Venkitaraman and D. R. Spring, *Nat. Commun.*, 2014, **5**, 3155–3163.
- 89 E. Lenci, G. Menchi, A. Guarna and A. Trabocchi, *J. Org. Chem.*, 2015, **80**, 2182–2191.
- 90 R. W. Heidebrecht, C. Mulrooney, C. P. Austin, R. H. Barker, J. A. Beaudoin, K. Chih-Chien Cheng, E. Comer, S. Dandapani, J. Dick, J. R. Duvall, E. H. Ekland, D. A. Fidock, M. E. Fitzgerald, M. Foley, R. Guha, P. Hinkson, M. Kramer, A. K. Lukens, D. Masi, L. A. Marcaurette, X.-Z. Su, C. J. Thomas, M. Weïwer, R. C. Wiegand, D. Wirth, M. Xia, J. Yuan, J. Zhao, M. Palmer, B. Munoz and S. Schreiber, *ACS Med. Chem. Lett.*, 2011, **3**, 112–117.
- 91 N. Kato, E. Comer, T. Sakata-Kato, A. Sharma, M. Sharma, M. Maetani, J. Bastien, N. M. Brancucci, J. A. Bittker, V. Corey, D. Clarke, E. R. Derbyshire, G. L. Dornan, S. Duffy, S. Eckley, M. A. Itoe, K. M. J. Koolen, T. A. Lewis, P. S. Lui, A. K. Lukens, E. Lund, S. March, E. Meibalan, B. C. Meier, J. A. McPhail, B. Mitasev, E. L. Moss, M. Sayes, Y. Van Gessel, M. J. Wawer, T. Yoshinaga, A.-M. Zeeman, V. M. Avery, S. N. Bhatia, J. E. Burke, F. Catteruccia, J. C. Clardy, P. A. Clemons, K. J. Dechering, J. R. Duvall, M. A. Foley, F. Gusovsky, C. H. M. Kocken, M. Marti, M. L. Morningstar, B. Munoz, D. E. Neafsey, A. Sharma, E. A. Winzeler, D. F. Wirth, C. A. Scherer and S. L. Schreiber, *Nature*, 2016, **538**, 344–349.
- 92 S.-Y. Kuo, A. B. Castoreno, L. N. Aldrich, K. G. Lassen, G. Goel, V. Dančik, P. Kuballa, I. Latorre, K. L. Conway, S. Sarkar, D. Maetzel, R. Jaenisch, P. A. Clemons, S. L. Schreiber, A. F. Shamji, R. J. Xavier, M. A. Fischbach and E. Latz, *PNAS*, 2015, **112**, e4281–e4287.
- 93 L. N. Aldrich, S. Y. Kuo, A. B. Castoreno, G. Goel, P. Kuballa, M. G. Rees, B. A.

- Seashore-Ludlow, J. H. Cheah, I. J. Latorre, S. L. Schreiber, A. F. Shamji and R. J. Xavier, *J. Am. Chem. Soc.*, 2015, **137**, 5563-5568.
- 94 D. H. C. Chou, J. R. Duvall, B. Gerard, H. Liu, B. A. Pandya, B. C. Suh, E. M. Forbeck, P. Faloon, B. K. Wagner and L. A. Marcaurelle, *ACS Med. Chem. Lett.*, 2011, **2**, 698–702.
- 95 S. Di Micco, R. Vitale, M. Pellicchia, M. F. Rega, R. Riva, A. Basso and G. Bifulco, *J. Med. Chem.*, 2009, **52**, 7856–7867.
- 96 J. Kim, J. Jung, J. Koo, W. Cho, W. S. Lee, C. Kim, W. Park and S. B. Park, *Nat. Commun.*, 2016, **7**, 1–10.
- 97 A. N. Koehler, A. F. Shamji and S. L. Schreiber, *J. Am. Chem. Soc.*, 2003, **125**, 8420–8421.
- 98 E. Wyatt, W. Galloway, G. Thomas, M. Welch, O. Loiseleur, A. Plowright and D. Spring, *Chem. Commun.*, 2008, **40**, 4962–4964.
- 99 G. L. Thomas, R. J. Spandl, F. G. Glansdorp, M. Welch, A. Bender, J. Cockfield, J. A. Lindsay, C. Bryant, D. F. J. Brown, O. Loiseleur, H. Rudyk, M. Ladlow and D. R. Spring, *Angew. Chemie Int. Ed.*, **47**, 2808 - 2812.
- 100 B. Z. Stanton, L. F. Peng, N. Maloof, K. Nakai, X. Wang, J. L. Duffner, K. M. Taveras, J. M. Hyman, S. W. Lee, A. N. Koehler, J. K. Chen, J. L. Fox, A. Mandinova and S. L. Schreiber, *Nat. Chem. Biol.*, 2009, **5**, 154–6.
- 101 S. B. Shuker, P. J. Hajduk, R. P. Meadows and S. W. Fesik, *Science*, 1996, **274**, 1531–1534.
- 102 J. Schoepfer, W. Jahnke, G. Berellini, S. Buonamici, S. Cotesta, S. W. Cowan-Jacob, S. Dodd, P. Drueckes, D. Fabbro, T. Gabriel, J.-M. Groell, R. M. Grotzfeld, A. Q. Hassan, || Chrystè Le Henry, V. Iyer, D. Jones, F. Lombardo, A. Loo, P. W. Manley, X. Pellé, Pellé, G. Rummel, B. Salem, M. Warmuth, A. A. Wylie, T. Zoller, A. L. Marzinzik and P. Furet, *J. Med. Chem.*, 2018, **61**, 8120–8135.
- 103 C. De Fusco, P. Brear, J. Iegre, K. H. Georgiou, H. F. Sore, M. Hyvönen and D. R. Spring, *Bioorg. Med. Chem.*, 2017, **25**, 3471–3482.
- 104 P. O. Nikiforov, S. Surade, M. Blaszczyk, V. Delorme, P. Brodin, A. R. Baulard, T. L. Blundell and C. Abell, *Org. Biomol. Chem.*, 2016, **14**, 2318–2326.
- 105 L. Ruddigkeit, R. Van Deursen, L. C. Blum and J.-L. Reymond, *J. Chem. Inf. Model.*, 2012, **52**, 2864–2875.

- 106 W. P. Janzen, *Chem. Biol.*, 2014, **21**, 1162–1170.
- 107 A. Schuffenhauer, S. Ruedisser, A. L. Marzinzik, W. Jahnke, M. Blommers, P. Selzer and E. Jacoby, *Curr. Top. Med. Chem.*, 2005, **5**, 751–762.
- 108 M. M. Hann, A. R. Leach and G. Harper, *J. Chem. Inf. Comput. Sci.*, 2001, **41**, 856–864.
- 109 A. R. Leach and M. M. Hann, *Curr. Opin. Chem. Biol.*, 2011, **15**, 489–496.
- 110 D. A. Erlanson, S. W. Fesik, R. E. Hubbard, W. Jahnke and H. Jhoti, *Nat. Rev. Drug Discov.*, 2016, **15**, 605–619.
- 111 C. W. Murray and D. C. Rees, *Nat. Chem.*, 2009, **1**, 187–192.
- 112 A. L. Hopkins and C. R. Groom, *Drug Discov. Today*, 2004, **9**, 430–431.
- 113 S. D. Bembenek, B. A. Tounge and C. H. Reynolds, *Drug Discov. Today*, 2009, **14**, 278–283.
- 114 D. C. Rees, M. Congreve, C. W. Murray and R. Carr, *Nat. Rev. Drug Discov.*, 2004, **3**, 660–672.
- 115 G. M. Keserű, D. A. Erlanson, G. G. Ferenczy, M. M. Hann, C. W. Murray and S. D. Pickett, *J. Med. Chem.*, 2016, **59**, 8189–8206.
- 116 D. A. Erlanson, R. S. Mcdowell and T. O'brien, *J. Med. Chem.*, 2004, **47**, 3463–3482.
- 117 C. A. Shepherd, A. L. Hopkins and I. Navratilova, *Prog. Biophys. Mol. Biol.*, 2014, **116**, 113–123.
- 118 J. Wielens, S. Headey, D. I. Rhodes, R. J. Mulder, O. Dolezal, J. J. Deadman, J. Newman, D. K. Chalmers, M. W. Parker, T. S. Peat and M. J. Scanlon, *J. Biomol. Screen.*, 2013, **18**, 147–159.
- 119 J. Schiebel, N. Radeva, H. Köster, A. Metz, T. Krotzky, M. Kuhnert, W. E. Diederich, A. Heine, L. Neumann, C. Atmanene, D. Roecklin, V. Vivat-Hannah, J.-P. Renaud, R. Meinecke, N. Schlinck, A. Sitte, F. Popp, M. Zeeb and G. Klebe, *ChemMedChem*, 2015, **10**, 1511–1521.
- 120 B. J. Davis and D. A. Erlanson, *Bioorg. Med. Chem. Lett.*, 2013, **23**, 2844–2852.
- 121 C. W. Murray and D. C. Rees, *Angew. Chemie Int. Ed.*, 2016, **55**, 488–492.
- 122 C. Abell and C. Dagostin, in *Fragment-Based Drug Discovery*, eds. S. Howard and C. Abell, Royal Society of Chemistry, 2015, pp. 1–18.
- 123 P. J. Hajduk and J. Greer, *Nat. Rev. Drug Discov.*, 2007, **6**, 211–219.

- 124 D. A. Erlanson, B. J. Davis and W. Jahnke, *Cell Chem. Biol.*, 2018, **26**, 9 - 15.
- 125 G. Szabó, G. I. Túrós, S. Kolok, M. Vastag, Z. Sánta, M. Dékány, G. I. Lévy, I. Greiner, M. Natsumi, W. Tatsuya and G. M. Keserű, *J. Med. Chem.*, 2018, **62**, 234 - 246.
- 126 M. Congreve, R. Carr, C. Murray and H. Jhoti, *Drug Discov. Today*, 2003, **8**, 876–877.
- 127 D. C. Fry, C. Wartchow, B. Graves, C. Janson, C. Lukacs, U. Kammlott, C. Belunis, S. Palme, C. Klein and B. Vu, *ACS Med. Chem. Lett.*, 2013, **4**, 660–665.
- 128 J. Iegre, P. Brear, C. De Fusco, M. Yoshida, S. L. Mitchell, M. Rossmann, L. Carro, H. F. Sore, M. Hyvönen, M. Hyvönen and D. R. Spring, *Chem. Sci.*, 2018, **9**, 3041–3049.
- 129 A. Mousnier, A. S. Bell, D. P. Swieboda, J. Morales-Sanfrutos, I. Pérez-Dorado, J. A. Brannigan, J. Newman, M. Ritzefeld, J. A. Hutton, A. Guedán, A. S. Asfor, S. W. Robinson, I. Hopkins-Navratilova, A. J. Wilkinson, S. L. Johnston, R. J. Leatherbarrow, T. J. Tuthill, R. Solari and E. W. Tate, *Nat. Chem.*, 2018, **10**, 599–606.
- 130 G. Bollag, J. Tsai, J. Zhang, C. Zhang, P. Ibrahim, K. Nolop and P. Hirth, *Nat. Rev. Drug Discov.*, 2012, **11**, 873–886.
- 131 A. J. Souers, J. D. Levenson, E. R. Boghaert, S. L. Ackler, N. D. Catron, J. Chen, B. D. Dayton, H. Ding, S. H. Enschede, W. J. Fairbrother, D. C. S. Huang, S. G. Hymowitz, S. Jin, S. L. Khaw, P. J. Kovar, L. T. Lam, J. Lee, H. L. Maecker, K. C. Marsh, K. D. Mason, M. J. Mitten, P. M. Nimmer, A. Oleksijew, C. H. Park, C. M. Park, D. C. Phillips, A. W. Roberts, D. Sampath, J. F. Seymour, M. L. Smith, G. M. Sullivan, S. K. Tahir, C. Tse, M. D. Wendt, Y. Xiao, J. C. Xue, H. Zhang, R. A. Humerickhouse, S. H. Rosenberg and S. W. Elmore, *Nat. Med.*, **19**, 202 - 208 .
- 132 Dan Erlanson, Fragments in the clinic: 2018 edition,  
<http://practicalfragments.blogspot.com/2018/10/fragments-in-clinic-2018-edition.html> ,  
accessed 7 November 2018.
- 133 C. W. Murray, O. Callaghan, G. Chessari, A. Cleasby, M. Congreve, M. Frederickson, M. J. Hartshorn, R. Mcmenamin, S. Patel and N. Wallis, *J. Med. Chem.*, 2007, 1116–1123.
- 134 H. Jhoti, A. Cleasby, M. Verdonk and G. Williams, *Curr. Opin. Chem. Biol.*, 2007, **11**, 485–493.
- 135 C. Tse, A. R. Shoemaker, J. Adickes, M. G. Anderson, J. Chen, S. Jin, E. F. Johnson, K. C. Marsh, M. J. Mitten, P. Nimmer, L. Roberts, S. K. Tahir, Y. Xiao, X. Yang, H. Zhang, S. Fesik, S. H. Rosenberg and S. W. Elmore, *Cancer Res.*, 2008, **68**, 3421–

3428.

- 136 G. Chessari and A. J. Woodhead, *Drug Discov. Today*, 2009, **14**, 668–675.
- 137 P. D. Leeson and S. A. St-Gallay, *Nat. Rev. Drug Discov.*, 2011, **10**, 749–765.
- 138 P. Bamborough, H. Diallo, J. D. Goodacre, L. Gordon, A. Lewis, J. T. Seal, D. M. Wilson, M. D. Woodrow and C.-W. Chung, *J. Med. Chem.*, 2012, **55**, 587–596.
- 139 M. Moustakim and P. E. Brennan, in *Chemical and Biological Synthesis: Enabling Approaches for Understanding Biology*, 2018, pp. 114–137.
- 140 N. Igoe, E. D. Bayle, O. Fedorov, C. Tallant, P. Savitsky, C. Rogers, D. R. Owen, G. Deb, T. C. P. Somervaille, D. M. Andrews, N. Jones, A. Cheasty, H. Ryder, P. E. Brennan, S. Mü, S. Knapp and P. V Fish, *J. Med. Chem.*, 2017, **60**, 668–680.
- 141 A. M. Taylor, A. Cõ, M. C. Hewitt, R. Pastor, Y. Leblanc, C. G. Nasveschuk, F. A. Romero, T. D. Crawford, N. Cantone, H. Jayaram, J. Setser, J. Murray, M. H. Beresini, G. De, L. Boenig, Z. Chen, A. R. Conery, R. T. Cummings, L. A. Dakin, E. M. Flynn, O. W. Huang, S. Kaufman, P. J. Keller, J. R. Kiefer, T. Lai, Y. Li, J. Liao, W. Liu, H. Lu, E. Pardo, V. Tsui, J. Wang, Y. Wang, Z. Xu, F. Yan, D. Yu, L. Zawadzke, X. Zhu, X. Zhu, R. J. Sims, A. G. Cochran, S. Bellon, J. E. Audia, S. Magnuson and B. K. Albrecht, *ACS Med. Chem. Lett.*, 2016, **7**, 531–536.
- 142 Z.-F. Tao, L. Hasvold, L. Wang, X. Wang, A. M. Petros, C. H. Park, E. R. Boghaert, N. D. Catron, J. Chen, P. M. Colman, P. E. Czabotar, K. Deshayes, W. J. Fairbrother, J. A. Flygare, S. G. Hymowitz, S. Jin, R. A. Judge, M. F. T. Koehler, P. J. Kovar, G. Lessene, M. J. Mitten, C. O. Ndubaku, P. Nimmer, H. E. Purkey, A. Oleksijew, D. C. Phillips, B. E. Sleeb, B. J. Smith, M. L. Smith, S. K. Tahir, K. G. Watson, Y. Xiao, J. Xue, H. Zhang, K. Zobel, S. H. Rosenberg, C. Tse, J. D. Levenson, S. W. Elmore and A. J. Souers, *ACS Med. Chem. Lett.*, 2014, **5**, 1088–1093.
- 143 M. Lainchbury, T. P. Matthews, T. Mchardy, K. J. Boxall, M. I. Walton, P. D. Eve, A. Hayes, M. R. Valenti, A. K. De, H. Brandon, G. Box, G. W. Aherne, J. C. Reader, # Florence, I. Raynaud, S. A. Eccles, M. D. Garrett and I. Collins, *J. Med. Chem.*, 2012, **55**, 10229–10240.
- 144 A. D. Morley, A. Pugliese, K. Birchall, J. Bower, P. Brennan, N. Brown, T. Chapman, M. Drysdale, I. H. Gilbert, S. Hoelder, A. Jordan, S. V. Ley, A. Merritt, D. Miller, M. E. Swarbrick and P. G. Wyatt, *Drug Discov. Today*, 2013, **18**, 1221–1227.
- 145 R. J. Hall, P. N. Mortenson and C. W. Murray, *Prog. Biophys. Mol. Biol.*, 2014, **116**, 82–91.

- 146 D. J. Foley, P. G. E. E. Craven, P. M. Collins, R. G. Doveston, A. Aimon, R. Talon, I. Churcher, F. von Delft, S. P. Marsden and A. Nelson, *Chem. - A Eur. J.*, 2017, **23**, 15227–15232.
- 147 M. J. Harner, B. A. Chauder, J. Phan and S. W. Fesik, *J. Med. Chem.*, 2014, **57**, 9687–9692.
- 148 M. Lüthy, M. C. Wheldon, C. Haji-Cheteh, M. Atobe, P. S. Bond, P. O'Brien, R. E. Hubbard and I. J. S. S. Fairlamb, *Bioorganic Med. Chem.*, 2015, **23**, 2680–2694.
- 149 O. A. Davis, R. A. Croft and J. A. Bull, *Chem. Comm.*, 2015, **51**, 15446–15449.
- 150 D. Antermite, P. A. Dominic and J. A. Bull, **20**, 3948-3952.
- 151 D. G. Twigg, N. Kondo, S. L. Mitchell, W. R. J. D. Galloway, H. F. Sore, A. Madin and D. R. Spring, *Angew. Chemie Int. Ed.*, 2016, **55**, 12479–12483.
- 152 N. Palmer, T. M. Peakman, D. Norton and D. C. Rees, *Org. Biomol. Chem.*, 2016, **14**, 1599–1610.
- 153 Y. Wang, J. Y. Wach, P. Sheehan, C. Zhong, C. Zhan, R. Harris, S. C. Almo, J. Bishop, S. J. Haggarty, A. Ramek, K. N. Berry, C. O'Herin, A. N. Koehler, A. W. Hung and D. W. Young, *ACS Med. Chem. Lett.*, 2016, **7**, 852–856.
- 154 D. C. Blakemore, L. Castro, I. Churcher, D. C. Rees, A. W. Thomas, D. M. Wilson and A. Wood, *Nat. Chem.*, 2018, **10**, 383–394.
- 155 N. Mateu, D. R. Spring, H. F. Sore, T. J. Osberger and S. L. Kidd, *Front. Chem.*, 2018, **6**, 460.
- 156 J. Mayol-Llinàs, W. Farnaby and A. Nelson, *Chem. Commun.*, 2017, **53**, 12345–12348.
- 157 S. Haftchenary, S. D. Nelson, L. Furst, S. Dandapani, S. J. Ferrara, Ž. V. Bošković, S. Figueroa Lazú, A. M. Guerrero, J. C. Serrano, D. K. Crews, C. Brackeen, J. Mowat, T. Brumby, M. Bauser, S. L. Schreiber and A. J. Phillips, *ACS Comb. Sci.*, 2016, **18**, 569–574.
- 158 H. Hassan, S. P. Marsden and A. Nelson, *Bioorganic Med. Chem.*, 2018, **26**, 3030–3033.
- 159 Y. Liu, S. J. Han, W. B. Liu and B. M. Stoltz, *Acc. Chem. Res.*, 2015, **48**, 740–751.
- 160 K. W. Quasdorf and L. E. Overman, *Nature*, 2014, **516**, 181–191.
- 161 C. J. Douglas and L. E. Overman, *PNAS*, 2004, **101**, 5363–5367.

- 162 E. Vitaku, E. A. Ilardi and J. T. Njardarson, Top 200 Pharmaceuticals in 2012 by US Retail Sales,  
[http://njardarson.lab.arizona.edu/sites/njardarson.lab.arizona.edu/files/Top200Pharmaceutical Products by US Retail Sales in 2012\\_0.pdf](http://njardarson.lab.arizona.edu/sites/njardarson.lab.arizona.edu/files/Top200PharmaceuticalProductsbyUSRetailSalesin2012_0.pdf). (accessed 9 October 2018).
- 163 A. Kimishima, H. Umihara, A. Mizoguchi, S. Yokoshima and T. Fukuyama, *Org. Lett.*, 2014, **16**, 6244–6247.
- 164 D. C. Behenna, Y. Liu, T. Yurino, J. Kim, D. E. White, S. C. Virgil and B. M. Stoltz, *Nat. Chem.*, 2011, **4**, 130–133.
- 165 I. Marek, Y. Minko, M. Pasco, T. Mejuch, N. Gilboa, H. Chechik and J. P. Das, *J. Am. Chem. Soc.*, 2014, **136**, 2682–2694.
- 166 F. Gao, K. P. McGrath, Y. Lee and A. H. Hoveyda, *J. Am. Chem. Soc.*, 2010, **132**, 14315–14320.
- 167 J. T. Mohr, D. C. Behenna, A. M. Harned and B. M. Stoltz, *Angew. Chemie Int. Ed.*, 2005, **44**, 6924–6927.
- 168 A. M. R. Smith and K. K. (Mimi) Hii, *Chem. Rev.*, 2011, **111**, 1637–1656.
- 169 J. Clayden, M. Donnard, J. Lefranc and D. J. Tetlow, *Chem. Commun.*, 2011, **47**, 4624–4639.
- 170 F. Zhou, F. M. Liao, J. S. Yu and J. Zhou, *Synth.*, 2014, **46**, 2983–3003.
- 171 H. Kobayashi, T. Misawa, K. Matsuno and Y. Demizu, *J. Org. Chem.*, 2017, **82**, 10722–10726.
- 172 C. Cativiela and M. D. Díaz-de-Villegas, *Tetrahedron: Asymmetry*, 2007, **18**, 569–623.
- 173 S. Kotha, D. Goyal and A. S. Chavan, *J. Org. Chem.*, 2013, 12288–12313.
- 174 S. H. Kang, S. Y. Kang, H.-S. Lee and A. J. Buglass, *Chem. Rev.*, 2005, **105**, 4537–4558.
- 175 Y. Ohfuné and T. Shinada, *European J. Org. Chem.*, 2005, **2005**, 5127–5143.
- 176 WHO Model List of Essential Medicines,  
[http://www.who.int/medicines/publications/essentialmedicines/20th\\_EML2017.pdf?ua=1](http://www.who.int/medicines/publications/essentialmedicines/20th_EML2017.pdf?ua=1), (accessed 13 November 2018).
- 177 P. C. May, R. A. Dean, S. L. Lowe, F. Martenyi, S. M. Sheehan, L. N. Boggs, S. A. Monk, B. M. Mathes, D. J. Mergott, B. M. Watson, S. L. Stout, D. E. Timm, E. S. Labell, C. R. Gonzales, M. Nakano, S. S. Jhee, M. Yen, L. Ereshefsky, T. D.

- Lindstrom, D. O. Calligaro, P. J. Cocke, D. G. Hall, S. Friedrich, M. Citron and J. E. Audia, *J. Neurosci.*, 2011, **31**, 16507–16516.
- 178 P. C. May, B. A. Willis, S. L. Lowe, R. A. Dean, S. A. Monk, P. J. Cocke, J. E. Audia, L. N. Boggs, A. R. Borders, R. A. Brier, D. O. Calligaro, T. A. Day, L. Ereshefsky, J. A. Erickson, H. Gevorkyan, C. R. Gonzales, D. E. James, S. S. Jhee, S. F. Komjathy, L. Li, T. D. Lindstrom, B. M. Mathes, F. Martenyi, S. M. Sheehan, S. L. Stout, D. E. Timm, G. M. Vaught, B. M. Watson, L. L. Winneroski, Z. Yang and D. J. Mergott, *J. Neurosci.*, 2015, **35**, 1199–1210.
- 179 O. Delgado, A. Monteagudo, M. Van Gool, A. a Trabanco and S. Fustero, *Org. Biomol. Chem.*, 2012, **10**, 6758–66.
- 180 G. Saxty, S. J. Woodhead, V. Berdini, T. G. Davies, M. L. Verdonk, P. G. Wyatt, R. G. Boyle, D. Barford, R. Downham, M. D. Garrett and R. A. Carr, *J. Med. Chem.*, 2007, **50**, 2293–2296.
- 181 M. Addie, P. Ballard, D. Buttar, C. Crafter, G. Currie, B. R. Davies, J. Debreczeni, H. Dry, P. Dudley, R. Greenwood, P. D. Johnson, J. G. Kettle, C. Lane, G. Lamont, A. Leach, R. W. A. Luke, J. Morris, D. Ogilvie, K. Page, M. Pass, S. Pearson and L. Ruston, *J. Med. Chem.*, 2013, **56**, 2059–2073.
- 182 T. McHardy, J. J. Caldwell, K.-M. Cheung, L. J. Hunter, K. Taylor, M. Rowlands, R. Ruddle, A. Henley, A. De, H. Brandon, M. Valenti, T. G. Davies, L. Fazal, L. Seavers, F. I. Raynaud, S. A. Eccles, G. W. Aherne, M. D. Garrett and I. Collins, *J. Med. Chem.*, 2010, **53**, 2239–2249.
- 183 A. Donald, T. McHardy, M. G. Rowlands, L.-J. K. Hunter, T. G. Davies, V. Berdini, R. G. Boyle, G. W. Aherne, M. D. Garrett and I. Collins, *J. Med. Chem.*, 2007, **50**, 2289–2292.
- 184 T. A. Yap, M. I. Walton, K. M. Grimshaw, R. H. Te Poele, P. D. Eve, M. R. Valenti, A. K. De Haven Brandon, V. Martins, A. Zetterlund, S. P. Heaton, K. Heinzmann, P. S. Jones, R. E. Feltell, M. Reule, S. J. Woodhead, T. G. Davies, J. F. Lyons, F. I. Raynaud, S. A. Eccles, P. Workman, N. T. Thompson and M. D. Garrett, *Clin. Cancer Res.*, 2012, **18**, 3912–3923.
- 185 M. C. McLeod and J. Aubé, *Tetrahedron*, 2016, **72**, 3766–3774.
- 186 A. W. Hung, A. Ramek, Y. Wang, T. Kaya, J. A. Wilson, P. A. Clemons and D. W. Young, *PNAS*, 2011, **108**, 6799–6804.
- 187 M. Sivaprakasam, F. Couty, O. David, J. Marrot, R. Sridhar, B. Srinivas and K. Rama Rao, *European J. Org. Chem.*, 2007, **2007**, 5734–5739.

- 188 D. Foley, R. Doveston, I. Churcher, A. Nelson and S. P. Marsden, *Chem. Commun.*, 2015, **51**, 11174–11177.
- 189 B. Q. Zhang, Y. Luo, Y. H. He and Z. Guan, *Tetrahedron*, 2014, **70**, 1961–1966.
- 190 J. Zhang and J. Ji, *ACS Catal.*, 2011, **1**, 1360–1363.
- 191 Y. W. Zhu, J. L. Qian, W. Bin Yi and C. Cai, *Tetrahedron Lett.*, 2013, **54**, 638–641.
- 192 S. Fustero, J. Miró, M. Sánchez-Roselló and C. Del Pozo, *Chem. - A Eur. J.*, 2014, **20**, 14126–14131.
- 193 S. Fustero, P. Bello, J. Miró, M. Sánchez-Roselló, M. a Maestro, J. González and C. del Pozo, *Chem. Commun.*, 2013, **49**, 1336–8.
- 194 G. J. Jiang, Q. H. Zheng, M. Dou, L. G. Zhuo, W. Meng and Z. X. Yu, *J. Org. Chem.*, 2013, **78**, 11783–11793.
- 195 L. W. Bieber, M. F. Da Silva, R. C. Da Costa and L. O. S. Silva, *Tetrahedron Lett.*, 1998, **39**, 3655–3658.
- 196 A. Jögi and U. Mäeorg, *Molecules*, 2001, **6**, 964–968.
- 197 H. Fuwa, *Mar. Drugs*, 2016, **14**, 1–40.
- 198 T. Biftu, R. Sinha-Roy, P. Chen, X. Qian, D. Feng, J. T. Kuethe, G. Scapin, Y. D. Gao, Y. Yan, D. Krueger, A. Bak, G. Eiermann, J. He, J. Cox, J. Hicks, K. Lyons, H. He, G. Salituro, S. Tong, S. Patel, G. Doss, A. Petrov, J. Wu, S. S. Xu, C. Sewall, X. Zhang, B. Zhang, N. A. Thornberry and A. E. Weber, *J. Med. Chem.*, 2014, **57**, 3205–3212.
- 199 F. Xu, M. J. Zacuto, Y. Kohmura, J. Rosen, A. Gibb, M. Alam, J. Scott and D. Tschaen, *Org. Lett.*, 2014, **16**, 5422–5425.
- 200 M. Yamashita, T. Tomozawa, M. Kakuta, A. Tokumitsu, H. Nasu and S. Kubo, *Antimicrob. Agents Chemother.*, 2009, **53**, 186–192.
- 201 M. J. Zacuto, D. Tomita, Z. Pirzada and F. Xu, *Org. Lett.*, 2010, **12**, 684–687.
- 202 R. D. Taylor, M. MacCoss and A. D. G. Lawson, *J. Med. Chem.*, 2014, **57**, 5845–5859.
- 203 M. T. Crimmins and A. L. Choy, *J. Am. Chem. Soc.*, 1999, **121**, 5653–5660.
- 204 T. A. Reekie, M. E. Kavanagh, M. Longworth and M. Kassiou, *Synthesis*, 2013, **45**, 3211–3227.
- 205 A. Nortcliffe and C. J. Moody, *Bioorganic Med. Chem.*, 2015, **23**, 2730–2735.
- 206 S. Basu, B. Ellinger, S. Rizzo, C. Deraeve, M. Schürmann, H. Preut, H.-D. Arndt and

- H. Waldmann, *PNAS*, 2011, **108**, 6805–10.
- 207 M. Mori, N. Sakakibara and A. Kinoshita, *J. Org. Chem.*, 1998, **63**, 6082–6083.
- 208 K. Tonogaki and M. Mori, *Tetrahedron Lett.*, 2002, **43**, 2235–2238.
- 209 P. Xing, Z. G. Huang, Y. Jin and B. Jiang, *Tetrahedron Lett.*, 2013, **54**, 699–702.
- 210 D. V. Vorobyeva, A. K. Mailyan, A. S. Peregudov, N. M. Karimova, T. P. Vasilyeva, I. S. Bushmarinov, C. Bruneau, P. H. Dixneuf and S. N. Osipov, *Tetrahedron*, 2011, **67**, 3524–3532.
- 211 N. Agenet, V. Gandon, K. P. C. Vollhardt, M. Malacria and C. Aubert, *J. Am. Chem. Soc.*, 2007, **129**, 8860–8871.
- 212 D. D. Young and A. Deiters, *Angew. Chemie Int. Ed.*, 2007, **46**, 5187–5190.
- 213 G. Domínguez and J. Pérez-Castells, *Chem. - A Eur. J.*, 2016, **22**, 6720–6739.
- 214 A. L. McIver and A. Deiters, *Org. Lett.*, 2010, **12**, 1288–1291.
- 215 D. D. Young, J. A. Teske and A. Deiters, *Synthesis*, 2009, 3785–3790.
- 216 S. Alvarez, S. Medina, G. Dom, J. Pe, F. De Farmacia, D. Química, U. San, P. Ceu, U. Montepíncipe and B. Monte, *J. Org. Chem.*, 2013, **79**, 9995–10001.
- 217 N. Nicolaus, S. Strauss, J. M. Neudörfl, A. Prokop and H. G. Schmalz, *Org. Lett.*, 2009, **11**, 341–344.
- 218 N. Nicolaus and H. G. Schmalz, *Synlett*, 2010, 2071–2074.
- 219 L. V. R. Boñaga, H.-C. Zhang and B. E. Maryanoff, *Chem Comm*, 2004, 2394–2395.
- 220 P. A. S. Lowden, in *Aziridines and Epoxides in Organic Synthesis*, eds A. K. Yudin, Wiley, 2006, 399 - 442.
- 221 F. Reck, F. Zhou, M. Girardot, G. Kern, C. J. Eyermann, N. J. Hales, R. R. Ramsay and M. B. Gravestock, *J. Med. Chem.*, 2005, **48**, 499–506.
- 222 D. E. Gutstein, R. Krishna, D. Johns, H. K. Surks, H. M. Dansky, S. Shah, Y. B. Mitchel, J. Arena and J. A. Wagner, *Clin. Pharmacol. Ther.*, 2012, **91**, 109–22.
- 223 F. M. D. Ismail, D. O. Levitsky and V. M. Dembitsky, *Eur. J. Med. Chem.*, 2009, **44**, 3373–3387.
- 224 C. A. Flentge, J. T. Randolph, P. P. Huang, L. L. Klein, K. C. Marsh, J. E. Harlan and D. J. Kempf, *Bioorganic Med. Chem. Lett.*, 2009, **19**, 5444–5448.
- 225 M. Ho, J. K. K. Chung and N. Tang, *Tetrahedron*, 1993, **34**, 6513–6516.

- 226 P. Wessig and J. Schwarz, *Synlett*, 1997, 893–894.
- 227 N. Kuntala, J. R. Telu, V. Banothu, S. B. Nallapati, J. S. Anireddy and S. Pal, *Med. Chem. Commun.*, 2015, **6**, 1612–1619.
- 228 A. Grauer and B. König, *European J. Org. Chem.*, 2009, 5099–5111.
- 229 S. Nicolai, C. Piemontesi and J. Waser, *Angew. Chemie Int. Ed.*, 2011, **50**, 4680–4683.
- 230 L. Jiang, G. E. Job, A. Klapars and S. L. Buchwald, *Org. Lett.*, 2003, **5**, 3667–3669.
- 231 A. Tam, U. Arnold, M. B. Soellner and R. T. Raines, *J. Am. Chem. Soc.*, 2007, **129**, 12670–12671.
- 232 N. Mateu, S. L. Kidd, L. Kalash, H. F. Sore, A. Madin, A. Bender and D. R. Spring, *Chem. – A Eur. J.*, 2018, **24**, 13681–13687.
- 233 W. H. B. Sauer and M. K. Schwarz, *J. Chem. Inf. Comput. Sci.*, 2003, **43**, 987–1003.
- 234 M. Congreve, R. Carr, C. Murray and H. Jhoti, *Drug Discov. Today*, 2003, **8**, 876–877.
- 235 H. Jhoti, G. Williams, D. C. Rees and C. W. Murray, *Nat. Rev. Drug Discov.*, 2013, **12**, 644–644.
- 236 D. F. Veber, S. R. Johnson, H.-Y. Cheng, B. R. Smith, K. W. Ward and K. D. Kopple, *J. Med. Chem.*, 2002, **45**, 2615–2623.
- 237 D. E. Clark, *J. Pharm. Sci.*, 1999, **88**, 815–821.
- 238 H. Köster, T. Craan, S. Brass, C. Herhaus, M. Zentgraf, L. Neumann, A. Heine and G. Klebe, *J. Med. Chem.*, 2011, **54**, 7784–7796.
- 239 D. A. Erlanson, Pushing the Rule of 3,  
<http://practicalfragments.blogspot.com/2011/11/pushing-rule-of-3.html>. (accessed 3 November 2018).
- 240 M. M. Hann, *Med. Chem. Comm.*, 2011, **2**, 349 - 355.
- 241 M. J. Hartshorn, C. W. Murray, A. Cleasby, M. Frederickson, I. J. Tickle and H. Jhoti, *J. Med. Chem.*, 2005, **48**, 403–413.
- 242 H. L. Silvestre, T. L. Blundell, C. Abell and A. Ciulli, *PNAS*, 2013, **110**, 12984–12989.
- 243 C. W. Murray, M. G. Carr, O. Callaghan, G. Chessari, M. Congreve, S. Cowan, J. E. Coyle, R. Downham, E. Figueroa, M. Frederickson, B. Graham, R. McMenemy, M. A. O'Brien, S. Patel, T. R. Phillips, G. Williams, A. J. Woodhead and A. J. Woolford, *J. Med. Chem.*, 2010, **53**, 5942–5955.

- 244 J. Schiebel, N. Radeva, S. G. Krimmer, X. Wang, M. Stieler, F. R. Ehrmann, K. Fu, A. Metz, F. U. Huschmann, M. S. Weiss, U. Mueller, A. Heine and G. Klebe, *ACS Chem. Biol.*, 2016, **11**, 1693 - 1701.
- 245 T. G. Davies and I. J. Tickle, in *Fragment-Based Drug Discovery and X-Ray Crystallography*, ed. M. Davies, Thomas G. Hyvönen, Springer, Berlin, Heidelberg, 2011, pp. 33–59.
- 246 X. Yin, A. Scalia, L. Leroy, C. M. Cuttitta, G. M. Polizzo, D. L. Ericson, C. G. Roessler, O. Campos, M. Y. Ma, R. Agarwal, R. Jackimowicz, M. Allaire, A. M. Orville, R. M. Sweet and A. S. Soares, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2014, **70**, 1177–1189.
- 247 D. Patel, J. D. Bauman and E. Arnold, *Prog. Biophys. Mol. Biol.*, 2014, **116**, 92–100.
- 248 O. B. Cox, T. Krojer, P. Collins, O. Monteiro, R. Talon, A. Bradley, O. Fedorov, J. Amin, B. D. Marsden, J. Spencer, F. Von Delft and P. E. Brennan, *Chem. Sci.*, 2016, **7**, 2322–2330.
- 249 C. Yeow Koh, L. Kallur Siddaramaiah, R. M. Ranade, J. Nguyen, T. Jian, Z. Zhang, J. Robert Gillespie, F. S. Buckner, C. L. M. J. Verlinde, E. Fan and W. G. J. Hol, *Acta Crystallogr. D Biol. Crystallogr.*, 2015, **4**, 1684–1698.
- 250 J. Tsing Ng, C. Dekker, M. Kroemer, M. Osborne and F. Von Delft, *Acta Crystallogr D Biol Crystallogr*, 2014, 2702–2718.
- 251 P. M. Collins, J. Tsing Ng, R. Talon, K. Nekrosiute, T. Krojer, A. Douangamath, J. Brandao-Neto, N. Wright, N. M. Pearce and F. Von Delft, *Acta Crystallogr D Struct Biol*, 2017, **73**, 246–255.
- 252 N. Wright and F. von Delft, *Unpubl. Work*.
- 253 F. von Delft, Details of methods - X-ray data collection: beamline I04-1, <https://www.diamond.ac.uk/Instruments/Mx/Fragment-Screening/Methods-for-Fragment-Screening.html>, (accessed 24 October 2018).
- 254 T. Krojer, R. Talon, N. Pearce, P. Collins, A. Douangamath, J. Brandao-Neto, A. Dias, B. Marsden and F. von Delft, *Acta Crystallogr D Struct Biol*, 2017, **73**, 267–278.
- 255 N. M. Pearce, T. Krojer, A. R. Bradley, P. Collins, R. P. Nowak, R. Talon, B. D. Marsden, S. Kelm, J. Shi, C. M. Deane and F. Von Delft, *Nat. Commun.*, **8**, 1 - 8.
- 256 D. Bowkett, R. Talon, C. Tallant, C. Schofield, F. von Delft, S. Knapp, G. Bruton and P. E. Brennan, *ChemMedChem*, 2018, **13**, 1051–1057.

- 257 P. J. McIntyre, P. M. Collins, L. Vrzal, K. Birchall, L. H. Arnold, C. Mpamhanga, P. J. Coombs, S. G. Burgess, M. W. Richards, A. Winter, V. Veverka, F. Von Delft, A. Merritt and R. Bayliss, *ACS Chem. Biol.*, 2017, **12**, 2906–2914.
- 258 F. Delbart, M. Brams, F. Gruss, S. Noppen, S. Peigneur, S. Boland, P. Chaltin, J. Brandao-Neto, F. von Delft, W. G. Touw, R. P. Joosten, S. Liekens, J. Tytgat and C. Ulens, *J. Biol. Chem.*, 2018, **293**, 2534–2545.
- 259 A. S. Mildvan, Z. Xia, H. F. Azurmendi, V. Saraswat, P. M. Legler, M. A. Massiah, S. B. Gabelli, M. A. Bianchet, L.-W. Kang and L. M. Amzel, *Arch. Biochem. Biophys.*, 2005, **433**, 129–143.
- 260 M. Coseno, G. Martin, C. Berger, G. Gilmartin, W. Keller and S. Doubl  , *Nucleic Acids Res.*, 2008, **36**, 3474–3483.
- 261 S. Kim, J. Yamamoto, Y. Chen, M. Aida, T. Wada, H. Handa and Y. Yamaguchi, *Genes to Cells*, 2010, **15**, 1003–1013.
- 262 Q. Yang, M. Coseno, G. M. Gilmartin and S. Doubl  , *Structure*, 2011, **19**, 368–377.
- 263 B. Tian and J. L. Manley, *Nat. Rev. Mol. Cell Biol.*, 2016, **18**, 18–30.
- 264 J. I. Fletcher, J. D. Swarbrick, D. Maksel, K. R. Gayler and P. R. Gooley, *Structure*, 2002, **10**, 205–213.
- 265 Q. Yang, G. M. Gilmartin and S. Doubl  , *PNAS*, 2010, **107**, 10062–7.
- 266 A. R. Gruber, G. Martin, W. Keller and M. Zavolan, *RNA Biol.*, 2012, **9**, 1405–1412.
- 267 C. P. Masamha, Z. Xia, J. Yang, T. R. Albrecht, M. Li, A. Bin Shyu, W. Li and E. J. Wagner, *Nature*, 2014, **510**, 412–416.
- 268 V. A. Gennarino, C. E. Alcott, C. A. Chen, A. Chaudhury, M. A. Gillentine, J. A. Rosenfeld, S. Parikh, J. W. Wheless, E. R. Roeder, D. D. G. Horovitz, E. K. Roney, J. L. Smith, S. W. Cheung, W. Li, J. R. Neilson, C. P. Schaaf and H. Y. Zoghbi, *Elife*, 2015, **4**, 1–13.
- 269 C. R. S  ndergaard, A. E. Garrett, T. Carstensen, G. Pollastri and J. E. Nielsen, *J. Med. Chem.*, 2009, **52**, 5673–5684.
- 270 A. R. Bogdan and N. W. Sach, *Adv. Synth. Catal.*, 2009, **351**, 849–854.
- 271 S. Mohammed, R. A. Vishwakarma and S. B. Bharate, *RSC Adv.*, 2015, **5**, 3470–3473.
- 272 Fahmi Himo, Timothy Lovell, Robert Hilgraf, Vsevolod V. Rostovtsev, Louis Noodleman, K. Barry Sharpless and Valery V. Fokin, *J. Am. Chem. Soc.*, 2005, **127**,

- 210–216.
- 273 S. Grecian and V. V. Fokin, *Angew. Chemie Int. Ed.*, 2008, **47**, 8285–8287.
- 274 D. V. Vorobyeva, N. M. Karimova, I. L. Odinets, G.-V. Röschenthaler and S. N. Osipov, *Org. Biomol. Chem.*, 2011, **9**, 7335.
- 275 J. D. Hicks, A. M. Hyde, A. Martinez Cuezva and S. L. Buchwald, *J. Am. Chem. Soc.*, 2009, **131**, 16720–16734.
- 276 X. Wang, G. Fischer and M. Hyvönen, *Nat. Commun.*, 2016, **7**, 1–11.
- 277 L. Attisano, J. L. Wrana, E. Montalvo and J. Massagué, *Activation of Signalling by the Activin Receptor Complex*, 1996, **16**, 1066 - 1073
- 278 J. L. Wrana, L. Attisano, R. Wieser, F. Ventura and J. Massagué, *Nature*, 1994, **370**, 341–347.
- 279 S. Dennler, S. Itoh, D. Vivien, P. Ten Dijke, S. Phane Huet and J.-M. Gauthier, *Direct binding of Smad3 and Smad4 to critical TGF $\beta$ -inducible elements in the promoter of human plasminogen activator inhibitor-type 1 gene*, 1998, **17**, 3091 - 3100.
- 280 M. Antsiferova and S. Werner, *J. Cell Sci.*, 2012, **125**, 3929–3937.
- 281 S. A. Pangas, C. J. Jorgez, M. Tran, J. Agno, X. Li, C. W. Brown, T. R. Kumar and M. M. Matzuk, *Mol. Endocrinol.*, 2007, **21**, 2458–2471.
- 282 C. E. Murry and G. Keller, *Cell*, 2008, **132**, 661–680.
- 283 M. Antsiferova, M. Huber, M. Meyer, A. Piwko-Czuchra, T. Ramadan, A. S. MacLeod, W. L. Havran, R. Dummer, D. Hohl and S. Werner, *Nat. Commun.*, 2011, **2**, 1–10.
- 284 S. Werner and C. Alzheimer, *Cytokine Growth Factor Rev.*, 2006, **17**, 157–171.
- 285 D. M. Hyman, D. W. Rasco, J. R. Infante, J. F. Liu, E. Welkowsky, D. L. Thai and C. M. Haqq, *J. Clin. Oncol.*, 2016, **34**, 5536–5536.
- 286 J. J. Tao, J. F. Liu, D. W. Rasco, W. Navarro, C. M. Haqq and D. M. Hyman, *Cancer Res.*, 2018, **78**, CT011.
- 287 J. Greenwald, M. E. Vega, G. P. Allendorph, W. H. Fischer, W. Vale and S. Choe, *Mol. Cell*, 2004, **15**, 485–489.
- 288 N. Baurin, F. Aboul-Ela, X. Barril, B. Davis, M. Drysdale, B. Dymock, H. Finch, C. Fromont, C. Richardson, H. Simmonite and R. E. Hubbard, *J. Chem. Inf. Comput. Sci.*, 2004, **44**, 2157–2166.
- 289 B. Lamoree and R. E. Hubbard, *Essays Biochem.*, 2017, **61**, 453–461.

- 290 S. D. Roughley and A. M. Jordan, *J. Med. Chem.*, 2011, **54**, 3451–3479.
- 291 M. Hartenfeller, M. Eberle, P. Meier, C. Nieto-Oberhuber, K.-H. Altmann, G. Schneider, E. Jacoby and S. Renner, *J. Chem. Inf. Model.*, 2011, **51**, 3093–3098.
- 292 Enamine, Press Release: Enamine supplies DSI poised fragment and analogue libraries to Diamond Light Source XChem facility and SGC Oxford screening efforts Diamond-SGC-iNEXT Poised Library supports rapid fragment elaboration for chemical probe and drug development, <http://www.enamine.net>, (accessed 27 October 2018).
- 293 M. R. Friedfeld, M. Shevlin, J. M. Hoyt, S. W. Krska, M. T. Tudge and P. J. Chirik, *Science*, **342**, 1076–1080.
- 294 D. A. DiRocco, K. Dykstra, S. Krska, P. Vachal, D. V. Conway and M. Tudge, *Angew. Chemie Int. Ed.*, 2014, **53**, 4802–4806.
- 295 D. W. Robbins and J. F. Hartwig, *Science*, 2011, **333**, 1423–1427.
- 296 M. Shevlin, *ACS Med. Chem. Lett.*, 2017, **8**, 601–607.
- 297 S. M. Preshlock, B. Ghaffari, P. E. Maligres, S. W. Krska, R. E. Maleczka and M. R. Smith, *J. Am. Chem. Soc.*, 2013, **135**, 7572–7582.
- 298 A. Buitrago Santanilla, E. L. Regalado, T. Pereira, M. Shevlin, K. Bateman, L.-C. Campeau, J. Schneeweis, S. Berritt, Z.-C. Shi, P. Nantermet, Y. Liu, R. Helmy, C. J. Welch, P. Vachal, I. W. Davies, T. Cernak and S. D. Dreher, *Science*, 2015, **347**, 49–53.
- 299 A. Bellomo, N. Celebi-Olcum, X. Bu, N. Rivera, R. T. Ruck, C. J. Welch, K. N. Houk and S. D. Dreher, *Angew. Chemie Int. Ed.*, 2012, **51**, 6912–6915.
- 300 T. Cernak, N. J. Gesmundo, K. Dykstra, Y. Yu, Z. Wu, Z. C. Shi, P. Vachal, D. Sperbeck, S. He, B. A. Murphy, L. Sonatore, S. Williams, M. Madeira, A. Verras, M. Reiter, C. H. Lee, J. Cuff, E. C. Sherer, J. Kuethe, S. Goble, N. Perrotto, S. Pinto, D. M. Shen, R. Nargund, J. Balkovec, R. J. DeVita and S. D. Dreher, *J. Med. Chem.*, 2017, **60**, 3594–3605.
- 301 S. Chow, S. Liver and A. Nelson, *Nature Chem. Rev.*, 2018, **2**, 174 - 183.
- 302 N. J. Gesmundo, B. Sauvagnat, P. J. Curran, M. P. Richards, C. L. Andrews, P. J. Dandliker and T. Cernak, *Nature*, 2018, **557**, 228–232.
- 303 H. N. Weller, D. S. Nirschl, J. L. Paulson, S. L. Hoffman and W. H. Bullock, *ACS Comb. Sci.*, 2012, **14**, 520–526.
- 304 G. Karageorgis, S. Warriner and A. Nelson, *Nat. Chem.*, 2014, **6**, 872–6.

- 305 G. Karageorgis, M. Dow, A. Aimon, S. Warriner and A. Nelson, *Angew. Chemie Int. Ed.*, 2015, **54**, 13538–13544.
- 306 J. B. Murray, S. D. Roughley, N. Matassova and P. A. Brough, *J. Med. Chem.*, 2014, **57**, 2845–2850.
- 307 L. Liu, *J. Med. Chem.*, 2014, **57**, 2843–2844.
- 308 Opentrons, <https://opentrons.com/>, (accessed 24 October 2018).
- 309 A. Aimon, T. Krojer and F. Von Delft, *Unpubl. results*.
- 310 A. Bradley, Opentrons repo for parallel chemistry reactions, <https://github.com/xchem/opentrons>, (accessed 8 December 2018).
- 311 H. C. Kolb and K. B. Sharpless, *Drug Discov. Today*, 2003, **8**, 1128–1137.
- 312 V. V. Rostovtsev, L. G. Green, V. V. Fokin and K. B. Sharpless, *Angew. Chemie Int. Ed.*, 2002, **41**, 2596–2599.
- 313 Y. Wu, L. B. Olsen, Y. H. Lau, C. Hatt Jensen, M. Rossmann, Y. R. Baker, H. F. Sore, S. Collins and D. R. Spring, *ChemBioChem*, 2016, **17**, 689–692.
- 314 Y. H. Lau, Y. Wu, M. Rossmann, B. Xiong Tan, P. de Andrade, Y. S. Tan, C. Verma, G. J. McKenzie, A. R. Venkitaraman, M. Hyvönen and D. R. Spring, *Angew. Chemie Int. Ed.*, 2015, **54**, 15410–15413.
- 315 Y. H. Lau, P. de Andrade, S.-T. Quah, M. Rossmann, L. Laraia, N. Sköld, T. J. Sum, P. J. E. Rowling, T. L. Joseph, C. Verma, M. Hyvönen, L. S. Itzhaki, A. R. Venkitaraman, C. J. Brown, D. P. Lane and D. R. Spring, *Chem. Sci.*, 2014, **5**, 1804–1809.
- 316 Y. Chevolut, C. Bouillon, S. Vidal, F. Morvan, A. Meyer, J.-P. Cloarec, A. Jochum, J.-P. Praly, J.-J. Vasseur and E. Souteyrand, *Angew. Chemie Int. Ed.*, 2007, **46**, 2398–2402.
- 317 P. M. E. Gramlich, S. Warncke, J. Gierlich and T. Carell, *Angew. Chemie Int. Ed.*, 2008, **47**, 3442–3444.
- 318 A. J. Link and D. A. Tirrell, *J. Am. Chem. Soc.*, 2003, **125**, 1164–1165.
- 319 A. C. W. Pike, E. F. Garman, T. Krojer, F. von Delft and E. P. Carpenter, *Acta Crystallogr D, Struct Biol*, 2016, **72**, 303–318.
- 320 C. Girard, E. Önen, M. Aufort, S. Beauvière, E. Samson and J. Herscovici, *Org. Lett.*, 2006, **8**, 1689–1692.

- 321 B. H. Lipshutz and B. R. Taft, *Angew. Chemie Int. Ed.*, 2006, **45**, 8235–8238.
- 322 B. S. Gourlay, P. P. Molesworth, J. H. Ryan and J. A. Smith, *Tetrahedron Lett.*, 2006, **47**, 799–801.
- 323 N. Elming and N. Clauson-Kaas, *Acta Chem. Scand.*, 1952, **6**, 867–874.
- 324 S. V Ley, I. R. Baxendale, R. N. Bream, P. S. Jackson, A. G. Leach, D. A. Longbottom, M. Nesi, J. S. Scott, R. I. Storer and S. J. Taylor, *J. Chem. Soc., Perkin Trans. 1*, 2000, 3815–4195.
- 325 N. Nikbin, M. Ladlow and S. V Ley, *Org. Process Res. Dev.*, 2007, **11**, 458–462.
- 326 A. Hinchcliffe, C. Hughes, D. A. Pears and M. R. Pitts, *Org. Process Res. Dev.*, 2007, **11**, 477–481.
- 327 M. Baumann, I. R. Baxendale, S. V Ley, C. D. Smith and G. K. Tranmer, *Org. Lett.*, 2000, **8**, 5231–5234.
- 328 A. Chighine, G. Sechi and M. Bradley, *Drug Discov. Today*, 2007, **12**, 459–464.
- 329 Y. N. Belokon, R. Gareth and M. North, 2000, **41**, 7245–7248.
- 330 H. Sogawa, M. Shiotsuki and F. Sanda, *J. Polym. Sci. Part A Polym. Chem.*, 2012, **50**, 2008–2018.
- 331 L. Campbell-Verduyn, P. H. Elsinga, L. Mirfeizi, R. a Dierckx and B. L. Feringa, *Org. Biomol. Chem.*, 2008, **6**, 3461–3463.
- 332 G. Winter, *J. Appl. Cryst.*, 2010, **43**, 186–190.
- 333 D. G. Waterman, G. Winter, R. J. Gildea, J. M. Parkhurst, A. S. Brewster, N. K. Sauter and G. Evans, *Acta Crystallogr. Sect. D Struct. Biol.*, 2016, **72**, 558–575.
- 334 W. Kabsch, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2010, **66**, 125–132.
- 335 P. Evans, *Acta Crystallogr. Sect. D Struct. Biol.*, 2006, **62**, 72–82.
- 336 M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin and K. S. Wilson, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2011, **67**, 235–242.
- 337 M. Wojdyr, R. Keegan, G. Winter, A. Ashton, A. Lebedev and E. Krissinel, *Acta Crystallogr. Sect. A Found. Adv.*, 2014, **70**, C1447.
- 338 F. Long, R. A. Nicholls, P. Emsley, S. Gražulis, A. Merkys, A. Vaitkus and G. N. Murshudov, *Acta Crystallogr. Sect. D Struct. Biol.*, 2017, **73**, 112–122.

- 339 G. N. Murshudov, P. Skubák, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long and A. A. Vagin, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2011, **67**, 355–367.
- 340 P. Emsley, B. Lohkamp, W. G. Scott and K. Cowtan, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2010, **66**, 486–501.



## 8 Appendix

### 8.1 The full NSQC library

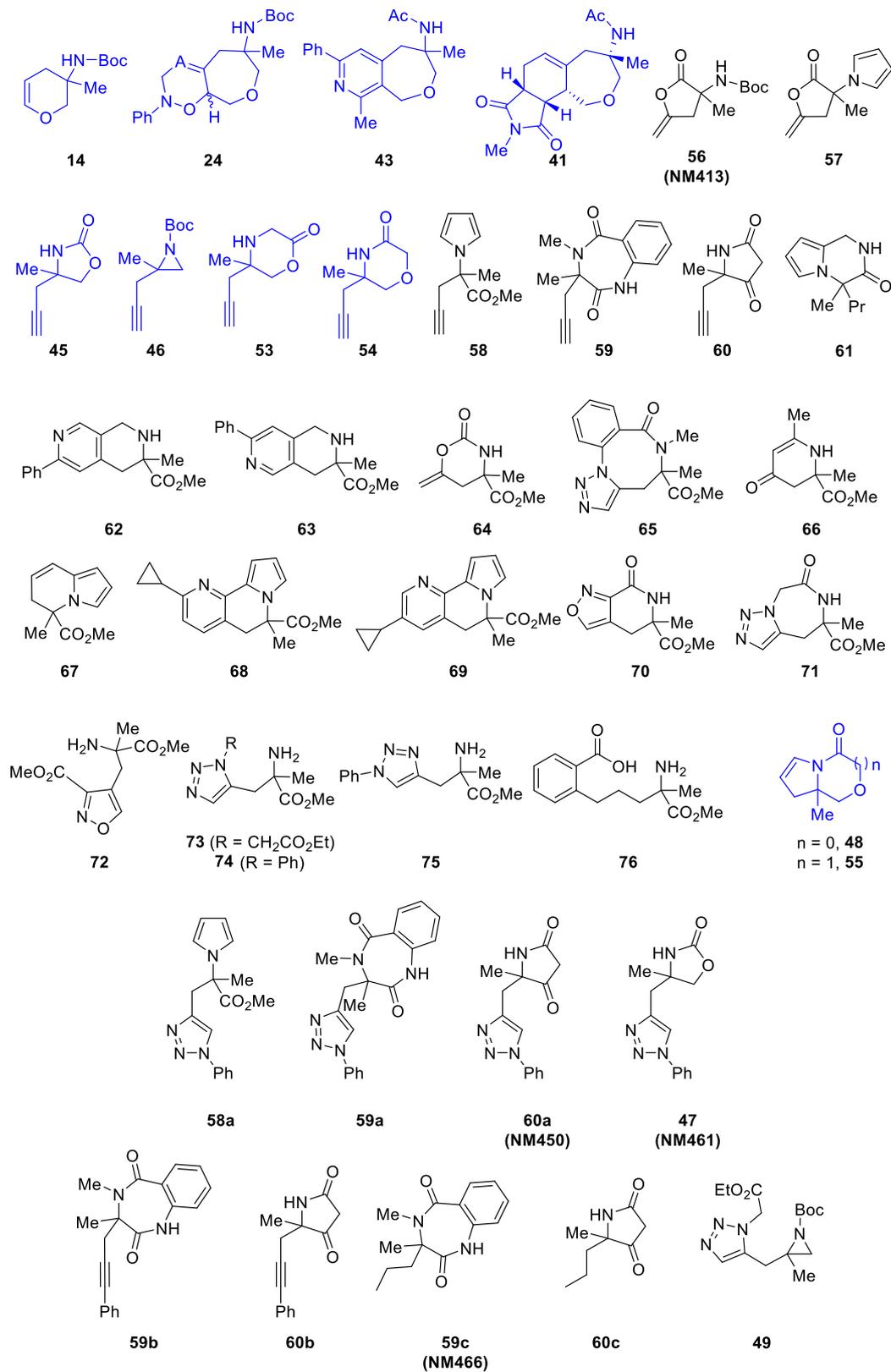


Figure 8.1. The full NSQC DOS library.

## 8.2 Computational analysis

### A) Principal Moment of Inertia (PMI)

#### General details

Principal Moment of Inertia (PMI) was performed using Molecular Operating Environment (MOE) software package version 2012.10 from the Chemical Computing Group. Merck molecular force field 94X (MMFF94x), an all-atom force field parameterised for small organic molecules with the Generalised Born solvation model, was used to minimise the energy potential of the library members. A LowModeMD search was employed for the conformation generation. Detailed settings for conformational search are listed below.

Rejection Limit	100
RMS Gradient	0.005
Iteration Limit	10000
MM Iteration Limit	500
RMSD Limit	0.15
Energy window	3
Conformation Limit	100

Only the conformer with the lowest energy was retained for principal moment of inertia (PMI) calculations. Normalized PMI ratios (I1/I3 and I2/I3) of these conformers were obtained from MOE and then plotted on a triangular graph, with the canonical coordinates (0,1), (0.5,0.5) and (1,1) representing a perfect rod, disc and sphere respectively (Figure 3).

### Compound collections analysed and PMI plots

#### Collection 1: DOS Library

*Table 8.1. Chemical structures of DOS Library in SMILES format.*

Compound	SMILES
14	<chem>CC1(NC(OC(C)(C)C)=O)COC=CC1</chem>

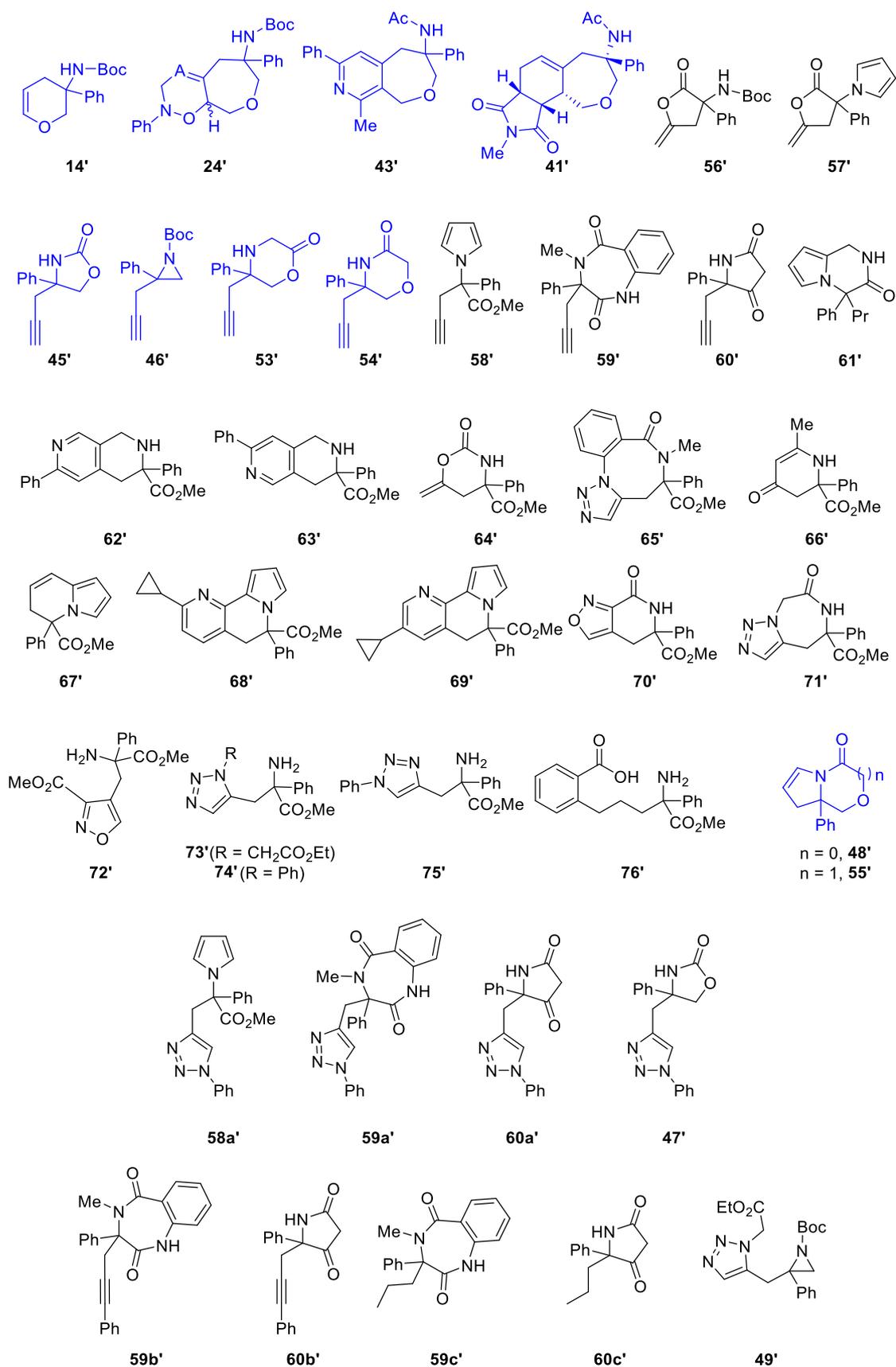
24	<chem>CC(C1)(NC(OC(C)(C)C)=O)COCC2C1=CCN(C3=CC=CC=C3)O2</chem>
41	<chem>O=C(N1C)[C@]([C@@]2([H])C1=O)([H])CC=C3[C@@H]2COC[C@](NC(C)=O)(C)C3</chem>
43	<chem>CC1(NC(C)=O)COCC2=C(C)N=C(C3=CC=CC=C3)C=C2C1</chem>
45	<chem>C#CCC(CO1)(C)NC1=O</chem>
46	<chem>C#CCC1(C)N(C(OC(C)(C)C)=O)C1</chem>
47	<chem>O=C1OCC(CC2=CN(C3=CC=CC=C3)N=N2)(C)N1</chem>
48	<chem>O=C1OCC2(C)N1C=CC2</chem>
49	<chem>O=C(OCC)CN1C(CC2(C)N(C(OC(C)(C)C)=O)C2)=CN=N1</chem>
53	<chem>C#CCC(CO1)(C)NCC1=O</chem>
54	<chem>O=C1COCC(CC#C)(C)N1</chem>
55	<chem>O=C1COCC2(C)N1C=CC2</chem>
56	<chem>O=C1C(NC(OC(C)(C)C)=O)(C)CC(O1)=C</chem>
57	<chem>O=C1C(N2C=CC=C2)(C)CC(O1)=C</chem>
58	<chem>C#CCC(C)(C(OC)=O)N1C=CC=C1</chem>
59	<chem>O=C(N(C)C1(CC#C)C)C2=C(C=CC=C2)NC1=O</chem>
60	<chem>O=C(NC1(CC#C)C)CC1=O</chem>
61	<chem>O=C(NC1)C(C)(CCC)N2C1=CC=C2</chem>
62	<chem>O=C(C1(C)CC2=C(C=NC(C3=CC=CC=C3)=C2)CN1)OC</chem>
63	<chem>O=C(C1(C)CC2=C(C=C(C3=CC=CC=C3)N=C2)CN1)OC</chem>
64	<chem>C=C(CC(C(OC)=O)(C)N1)OC1=O</chem>
65	<chem>O=C1N(C)C(C(OC)=O)(C)CC2=CN=NN2C3=C1C=CC=C3</chem>
66	<chem>O=C1C=C(C)NC(C)(C(OC)=O)C1</chem>
67	<chem>CC1(C(OC)=O)CC=CC2=CC=CN12</chem>
68	<chem>O=C(C1(C)CC2=CC=C(C3CC3)N=C2C4=CC=CN14)OC</chem>
69	<chem>O=C(C1(C)CC2=CC(C3CC3)=CN=C2C4=CC=CN14)OC</chem>
70	<chem>O=C1NC(C)(C(OC)=O)CC2=CON=C21</chem>
71	<chem>O=C(CN1C(C2)=CN=N1)NC2(C(OC)=O)C</chem>
72	<chem>NC(C)(C(OC)=O)CC1=CON=C1C(OC)=O</chem>
73	<chem>NC(CC1=CN=NN1CC(OCC)=O)(C)C(OC)=O</chem>
74	<chem>NC(CC1=CN=NN1C2=CC=CC=C2)(C)C(OC)=O</chem>
75	<chem>NC(CC1=CN(C2=CC=CC=C2)N=N1)(C(OC)=O)C</chem>
76	<chem>NC(CCCC1=CC=CC=C1C(O)=O)(C)C(OC)=O</chem>
58a	<chem>CC(CC1=CN(C2=CC=CC=C2)N=N1)(C(OC)=O)N3C=CC=C3</chem>

<b>59a</b>	<chem>O=C(N(C)C1(CC2=CN(C3=CC=CC=C3)N=N2)C)C4=C(C=CC=C4)NC1=O</chem>
<b>59b</b>	<chem>O=C(N(C)C1(CC#CC2=CC=CC=C2)C)C3=C(C=CC=C3)NC1=O</chem>
<b>59c</b>	<chem>O=C(N(C)C1(CCC)C)C2=C(C=CC=C2)NC1=O</chem>
<b>60a</b>	<chem>O=C(NC1(CC2=CN(C3=CC=CC=C3)N=N2)C)CC1=O</chem>
<b>60b</b>	<chem>O=C(NC1(CC#CC2=CC=CC=C2)C)CC1=O</chem>
<b>60c</b>	<chem>O=C(NC1(C)CCC)CC1=O</chem>

**Table 8.2.** Normalised PMI ratio (npr) values of conformers of the DOS Library with the lowest energy.

Compound	npr1	npr2	Compound	npr1	npr2
14	0.2626	0.9435	64	0.6608	0.9315
24	0.3226	0.8555	65	0.4033	0.7284
41	0.2417	0.9216	66	0.348	0.7927
43	0.2727	0.8894	67	0.4224	0.8973
45	0.2325	0.9648	68	0.3633	0.8393
46	0.4231	0.7895	69	0.3696	0.7741
47	0.0846	0.9619	70	0.2735	0.8451
48	0.4928	0.7482	71	0.3122	0.813
49	0.3812	0.8038	72	0.2989	0.8118
53	0.2465	0.9311	73	0.3747	0.7195
54	0.4579	0.8303	74	0.3506	0.7691
55	0.6285	0.7595	75	0.1877	0.9268
56	0.2034	0.9381	76	0.4495	0.9293
57	0.3332	0.9538	58a	0.2723	0.9423
58	0.646	0.695	59a	0.3108	0.8538
59	0.5321	0.8555	59b	0.3231	0.8831
60	0.4319	0.7592	59c	0.6047	0.9251
61	0.313	0.8822	60a	0.221	0.9554
62	0.2007	0.9149	60b	0.1535	0.9378
63	0.1761	0.9628	60c	0.538	0.8653

**Collection 2: DOS Library Ph.** Virtual collection of 40 small molecules based on DOS Library featuring a Ph substituent at the quaternary stereocenter.



**Figure 8.2.** The Ph NSQC library.

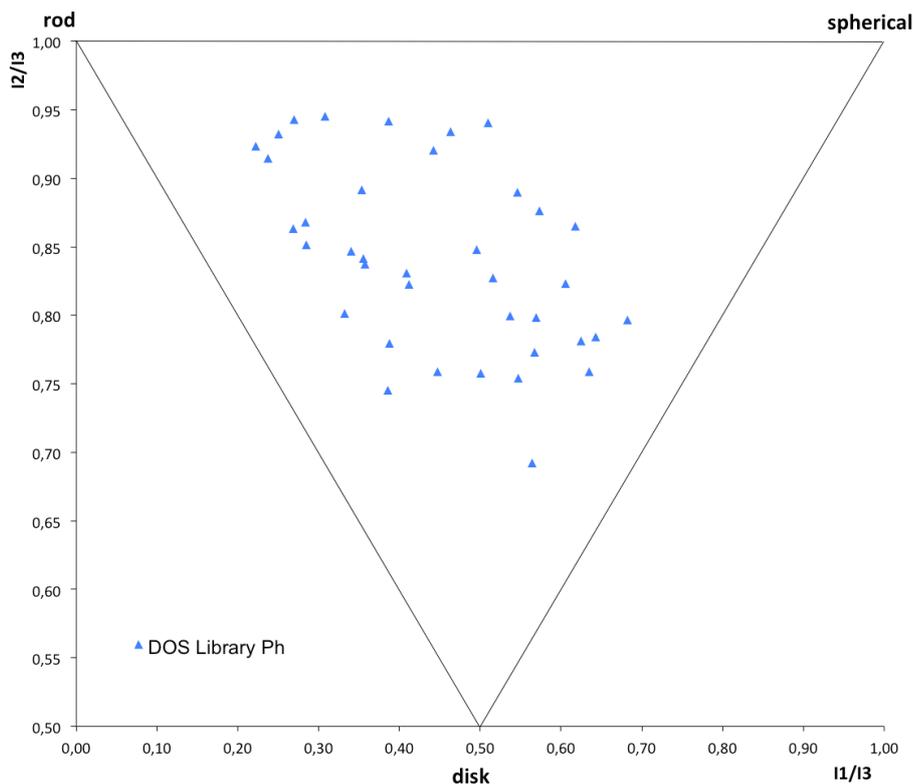
**Table 8.3.** Chemical structures of DOS Library Ph in SMILES format.

Compound	SMILES
14'	<chem>CC(OC(NC1(C2=CC=CC=C2)COC=CC1)=O)(C)C</chem>
24'	<chem>CC(OC(NC(C1)(C2=CC=CC=C2)COCC3C1=CCN(C4=CC=CC=C4)O3)=O)(C)C</chem>
41'	<chem>O=C(N1C)[C@]([C@@]2([H])C1=O)([H])CC=C3[C@@H]2COC[C@](N(C(C)=O)(C4=CC=CC=C4)C3</chem>
43'	<chem>CC1=C(COCC(NC(C)=O)(C2=CC=CC=C2)C3)C3=CC(C4=CC=CC=C4)=N1</chem>
45'	<chem>C#CCC(CO1)(C2=CC=CC=C2)NC1=O</chem>
46'	<chem>C#CCC1(C2=CC=CC=C2)N(C(OC(C)(C)C)=O)C1</chem>
47'	<chem>O=C1OCC(CC2=CN(C3=CC=CC=C3)N=N2)(C4=CC=CC=C4)N1</chem>
48'	<chem>O=C1OCC2(C3=CC=CC=C3)N1C=CC2</chem>
49'	<chem>O=C(OCC)CN1C(CC2(C3=CC=CC=C3)N(C(OC(C)(C)C)=O)C2)=CN=N1</chem>
53'	<chem>C#CCC(CO1)(C2=CC=CC=C2)NCC1=O</chem>
54'	<chem>O=C1COCC(CC#C)(C2=CC=CC=C2)N1</chem>
55'	<chem>O=C1COCC2(C3=CC=CC=C3)N1C=CC2</chem>
56'	<chem>O=C1C(NC(OC(C)(C)C)=O)(C2=CC=CC=C2)CC(O1)=C</chem>
57'	<chem>O=C1C(N2C=CC=C2)(C3=CC=CC=C3)CC(O1)=C</chem>
58'	<chem>C#CCC(C1=CC=CC=C1)(C(OC)=O)N2C=CC=C2</chem>
59'	<chem>O=C(N(C)C1(CC#C)C2=CC=CC=C2)C3=C(C=CC=C3)NC1=O</chem>
60'	<chem>O=C(NC1(CC#C)C2=CC=CC=C2)CC1=O</chem>
61'	<chem>O=C(NC1)C(C2=CC=CC=C2)(CCC)N3C1=CC=C3</chem>
62'	<chem>O=C(C1(C2=CC=CC=C2)CC3=C(C=NC(C4=CC=CC=C4)=C3)CN1)OC</chem>
63'	<chem>O=C(C1(C2=CC=CC=C2)CC3=C(C=C(C4=CC=CC=C4)N=C3)CN1)OC</chem>
64'	<chem>C=C(CC(C(OC)=O)(C1=CC=CC=C1)N2)OC2=O</chem>
65'	<chem>O=C1N(C)C(C(OC)=O)(C2=CC=CC=C2)CC3=CN=NN3C4=C1C=CC=C4</chem>
66'	<chem>O=C1C=C(C)NC(C2=CC=CC=C2)(C(OC)=O)C1</chem>
67'	<chem>O=C(C1(C2=CC=CC=C2)CC=CC3=CC=CN13)OC</chem>
68'	<chem>O=C(C1(C2=CC=CC=C2)CC3=CC=C(C4CC4)N=C3C5=CC=CN15)OC</chem>

<b>69'</b>	<chem>O=C(C1(C2=CC=CC=C2)CC3=CC(C4CC4)=CN=C3C5=CC=CN15)O</chem> <chem>C</chem>
<b>70'</b>	<chem>O=C1NC(C2=CC=CC=C2)(C(OC)=O)CC3=CON=C31</chem>
<b>71'</b>	<chem>O=C(CN1C(C2)=CN=N1)NC2(C(OC)=O)C3=CC=CC=C3</chem>
<b>72'</b>	<chem>NC(C1=CC=CC=C1)(C(OC)=O)CC2=CON=C2C(OC)=O</chem>
<b>73'</b>	<chem>NC(CC1=CN=NN1CC(OCC)=O)(C2=CC=CC=C2)C(OC)=O</chem>
<b>74'</b>	<chem>NC(CC1=CN=NN1C2=CC=CC=C2)(C3=CC=CC=C3)C(OC)=O</chem>
<b>75'</b>	<chem>NC(CC1=CN(C2=CC=CC=C2)N=N1)(C(OC)=O)C3=CC=CC=C3</chem>
<b>76'</b>	<chem>NC(CCCC1=CC=CC=C1C(O)=O)(C2=CC=CC=C2)C(OC)=O</chem>
<b>58a'</b>	<chem>O=C(C(CC1=CN(C2=CC=CC=C2)N=N1)(C3=CC=CC=C3)N4C=CC=C4)OC</chem>
<b>59a'</b>	<chem>O=C(N(C)C1(CC2=CN(C3=CC=CC=C3)N=N2)C4=CC=CC=C4)C5=C(C=CC=C5)NC1=O</chem>
<b>59b'</b>	<chem>O=C(N(C)C1(CC#CC2=CC=CC=C2)C3=CC=CC=C3)C4=C(C=CC=C4)NC1=O</chem>
<b>59c'</b>	<chem>O=C(N(C)C1(CCC)C2=CC=CC=C2)C3=C(C=CC=C3)NC1=O</chem>
<b>60a'</b>	<chem>O=C(NC1(CC2=CN(C3=CC=CC=C3)N=N2)C4=CC=CC=C4)CC1=O</chem>
<b>60b'</b>	<chem>O=C(NC1(CC#CC2=CC=CC=C2)C3=CC=CC=C3)CC1=O</chem>
<b>60c'</b>	<chem>O=C(NC1(C2=CC=CC=C2)CCC)CC1=O</chem>

**Table 8.4. Normalised PMI ratio (npr) values of conformers of the DOS Library Ph with the lowest energy.**

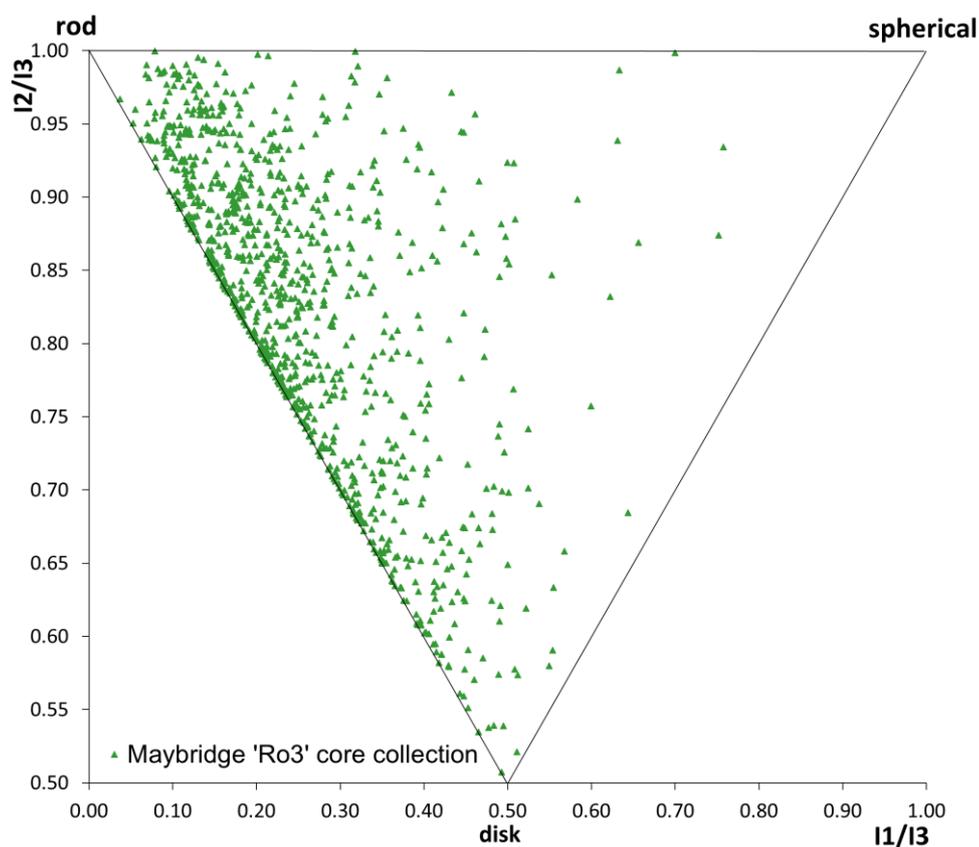
<b>Compound</b>	<b>npr1</b>	<b>npr2</b>	<b>Compound</b>	<b>npr1</b>	<b>npr2</b>
<b>14'</b>	0.5367	0.7998	<b>65</b>	0.6252	0.7811
<b>24'</b>	0.222	0.9232	<b>66</b>	0.6827	0.7967
<b>41'</b>	0.3576	0.8375	<b>67</b>	0.643	0.784
<b>43'</b>	0.2845	0.8513	<b>68</b>	0.3867	0.9416
<b>45'</b>	0.4475	0.7587	<b>69</b>	0.4424	0.9206
<b>46'</b>	0.3872	0.7793	<b>70</b>	0.5677	0.773
<b>47'</b>	0.2682	0.8633	<b>71</b>	0.606	0.8234
<b>48'</b>	0.3555	0.8412	<b>72</b>	0.4118	0.4118
<b>53'</b>	0.5471	0.754	<b>73</b>	0.4118	0.8226
<b>54'</b>	0.3858	0.7454	<b>74</b>	0.496	0.8478
<b>55'</b>	0.5096	0.9404	<b>75</b>	0.3317	0.8011
<b>56'</b>	0.4091	0.8306	<b>76</b>	0.3531	0.8914
<b>57'</b>	0.5158	0.8271	<b>23d'</b>	0.5009	0.7579
<b>58'</b>	0.6347	0.7591	<b>58a</b>	0.2371	0.9147
<b>59'</b>	0.6175	0.8648	<b>59a</b>	0.2698	0.943
<b>60'</b>	0.5456	0.8896	<b>59b</b>	0.3082	0.9453
<b>61'</b>	0.5641	0.6922	<b>59c</b>	0.5731	0.8761
<b>62'</b>	0.2509	0.2509	<b>60a</b>	0.2834	0.8679
<b>63'</b>	0.2509	0.932	<b>60b</b>	0.3405	0.8466
<b>64'</b>	0.4634	0.9343	<b>60c</b>	0.5695	0.7983



**Figure 8.3.** PMI plot of DOS Library Ph.

### **Collection 3: Maybridge 'Ro3' core fragment collection**

This library is based on the core 1000-member collection within the Maybridge Fragment library. Details of the library (including SMILES and SDF) are available from '<http://www.maybridge.com/>' under the 'Ro3 Fragment library section. More details can be found at: '[http://www.maybridge.com/images/pdfs/MB\\_Ro3\\_fragment\\_flyer\\_2011\\_EUR\\_v7.pdf](http://www.maybridge.com/images/pdfs/MB_Ro3_fragment_flyer_2011_EUR_v7.pdf)'



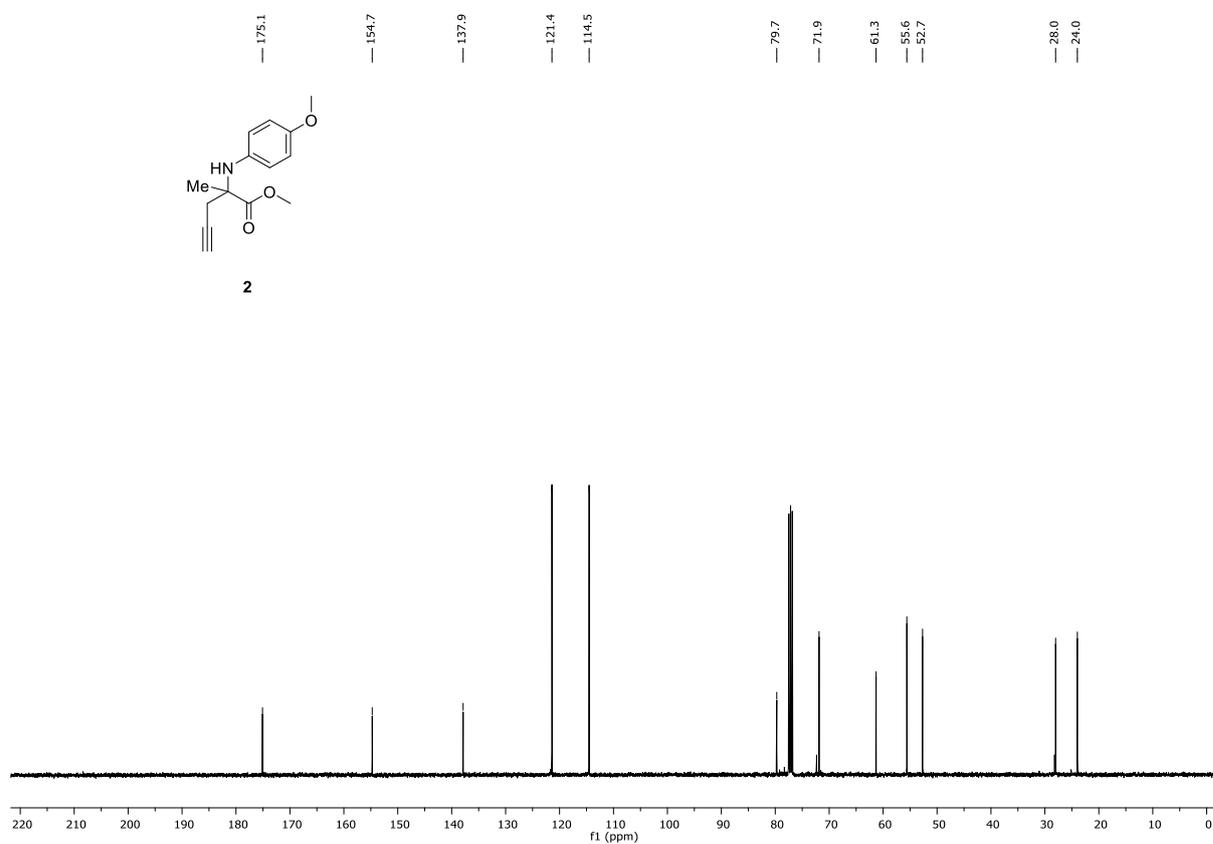
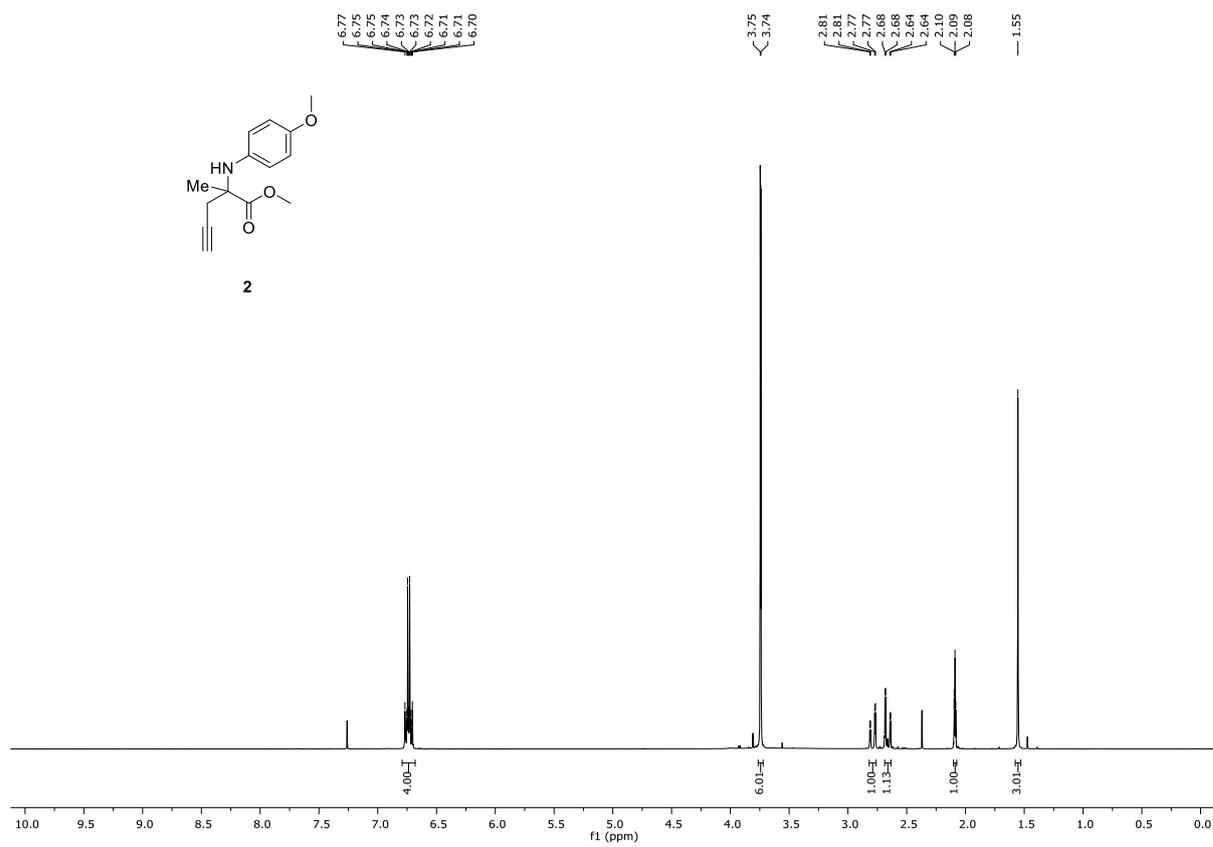
**Figure 8.4.** PMI plot of Maybridge 'Ro3' core collection.

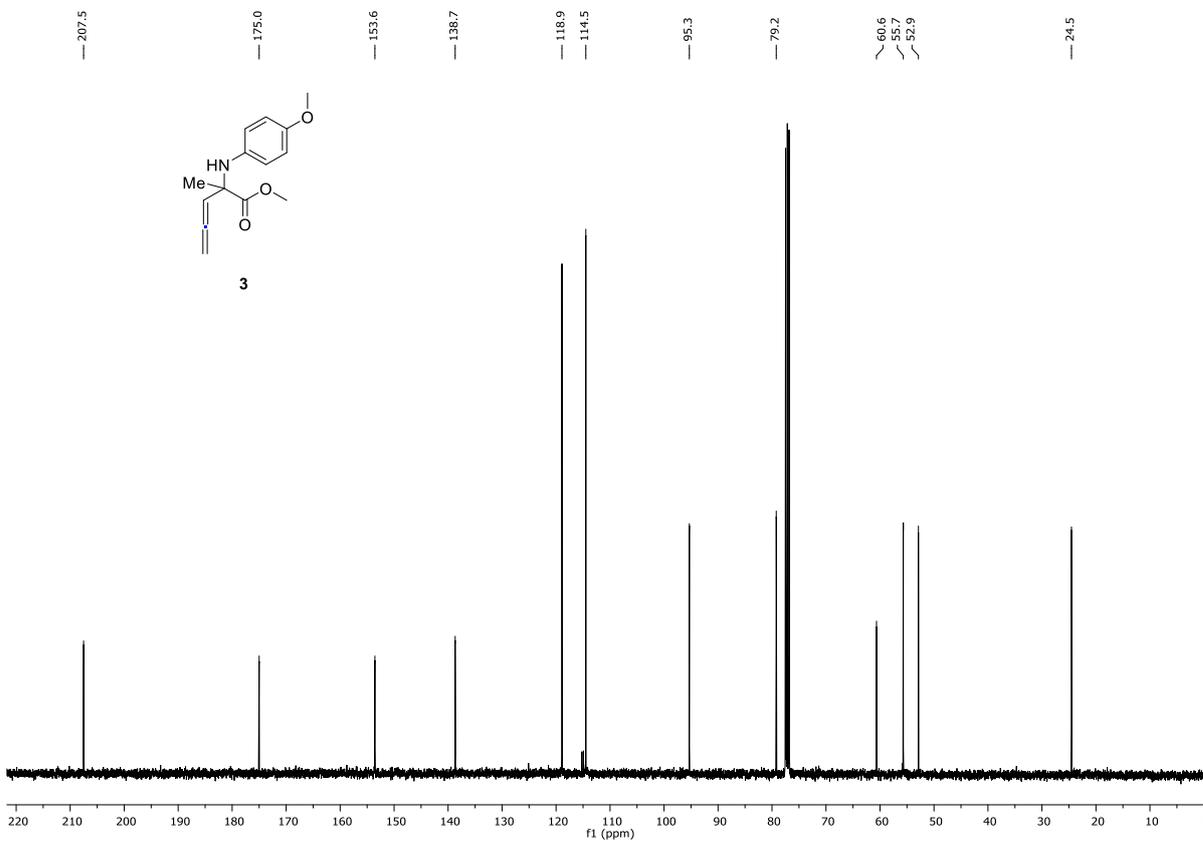
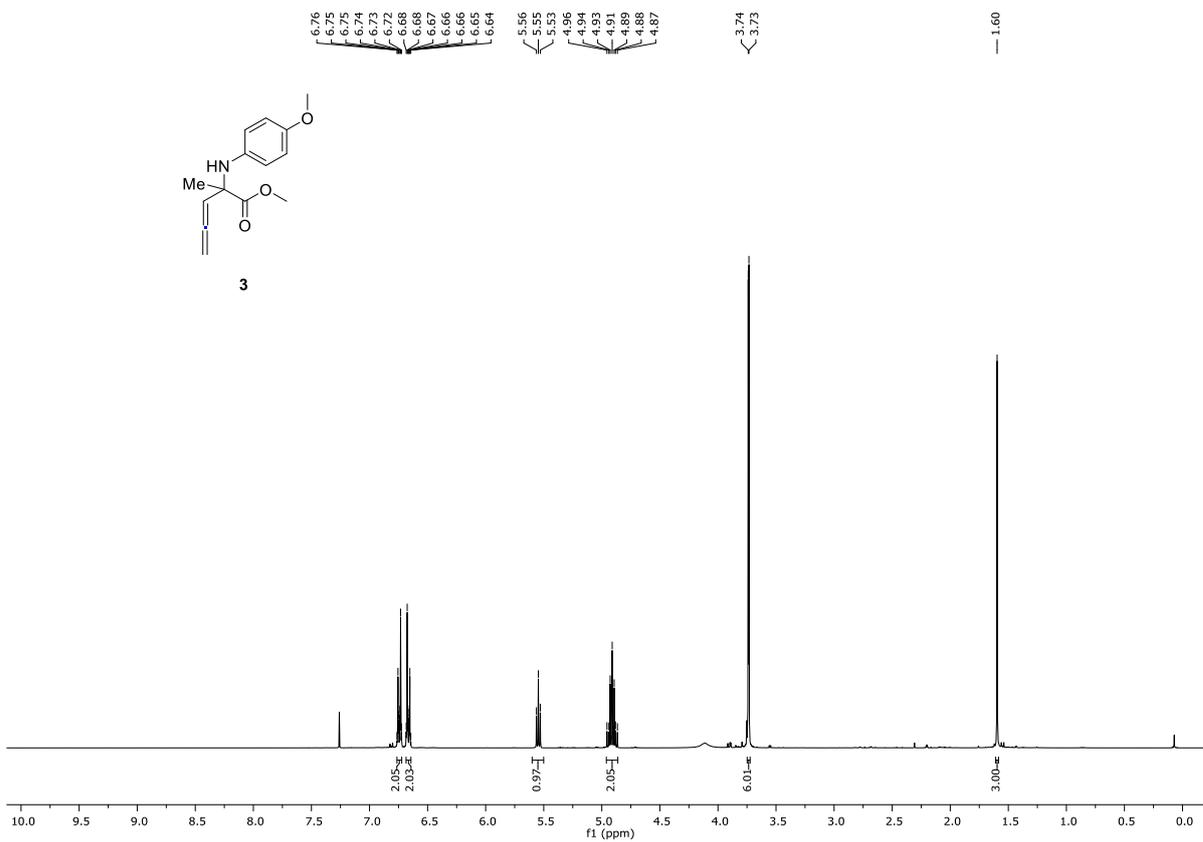
## B) Computational evaluation of physicochemical properties

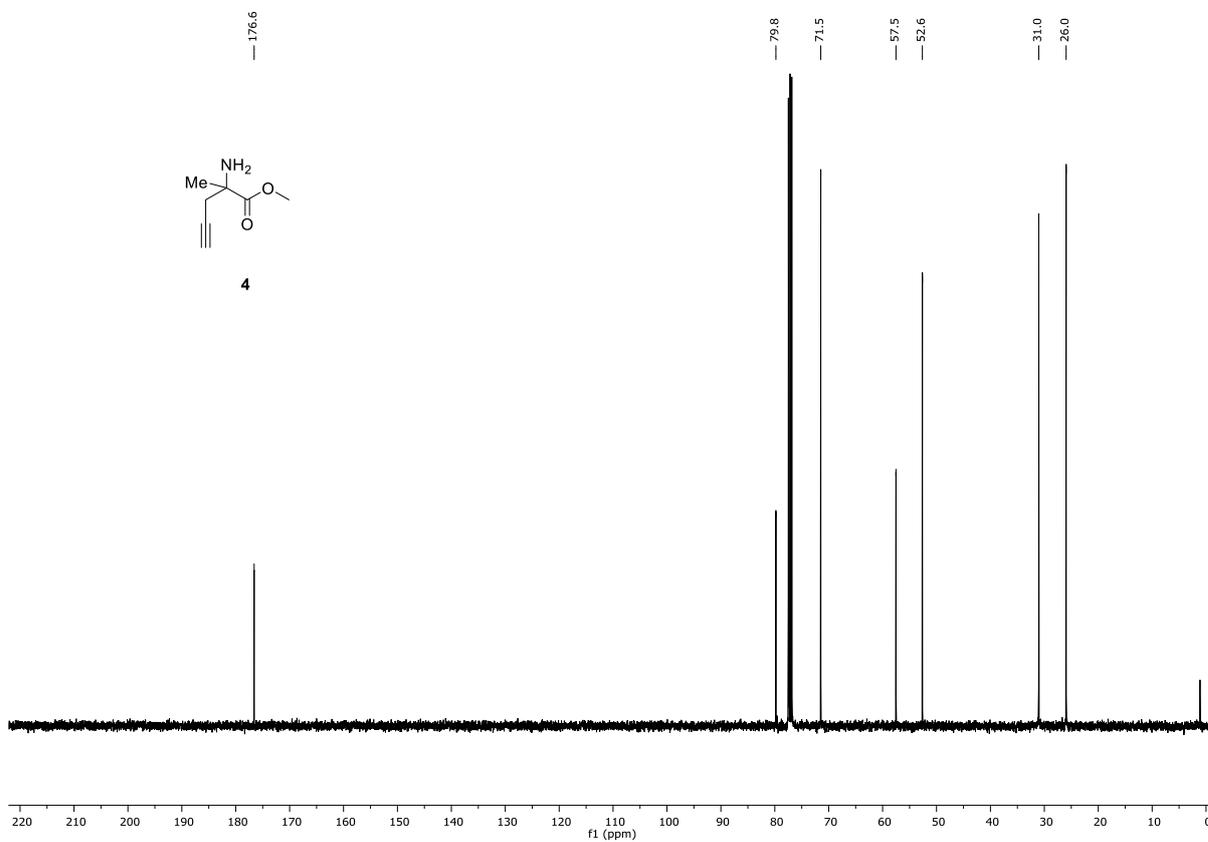
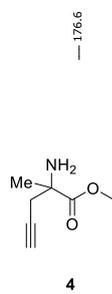
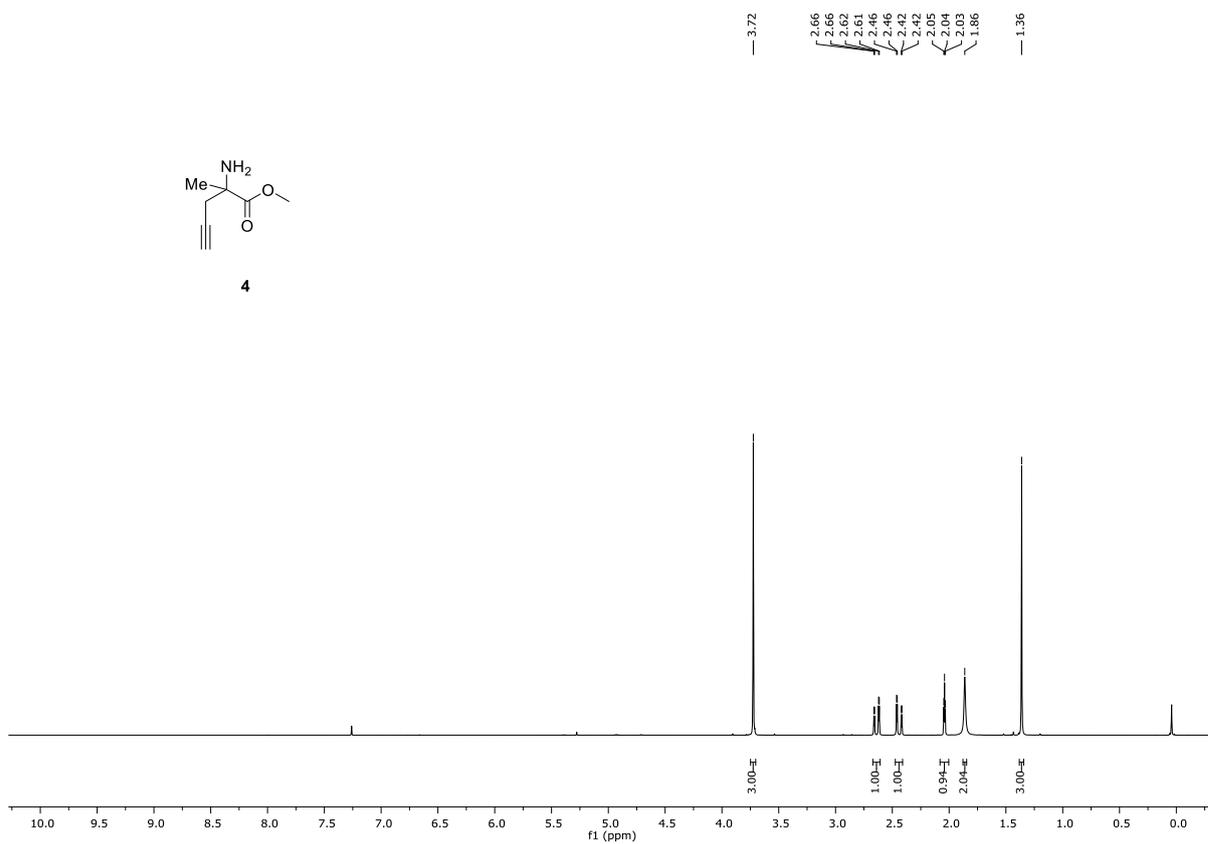
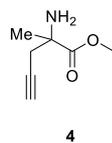
### General details

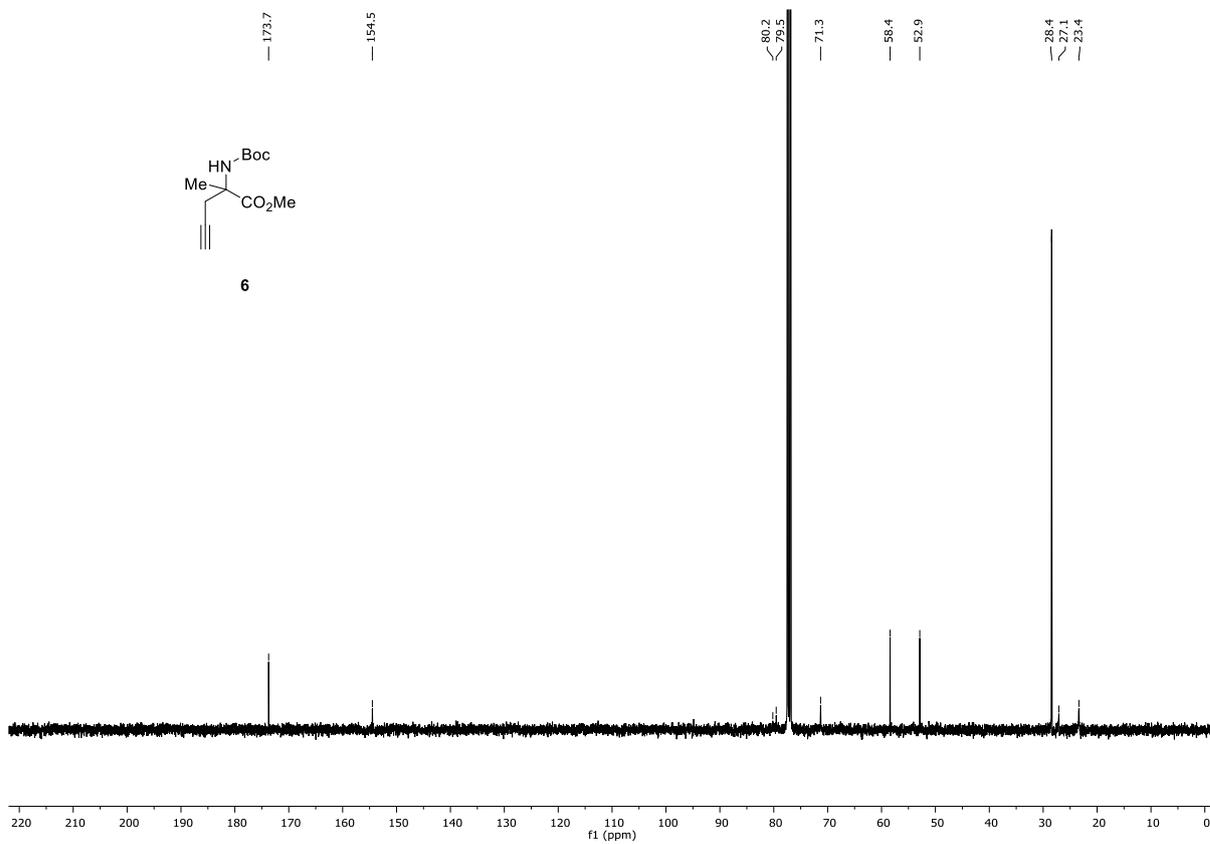
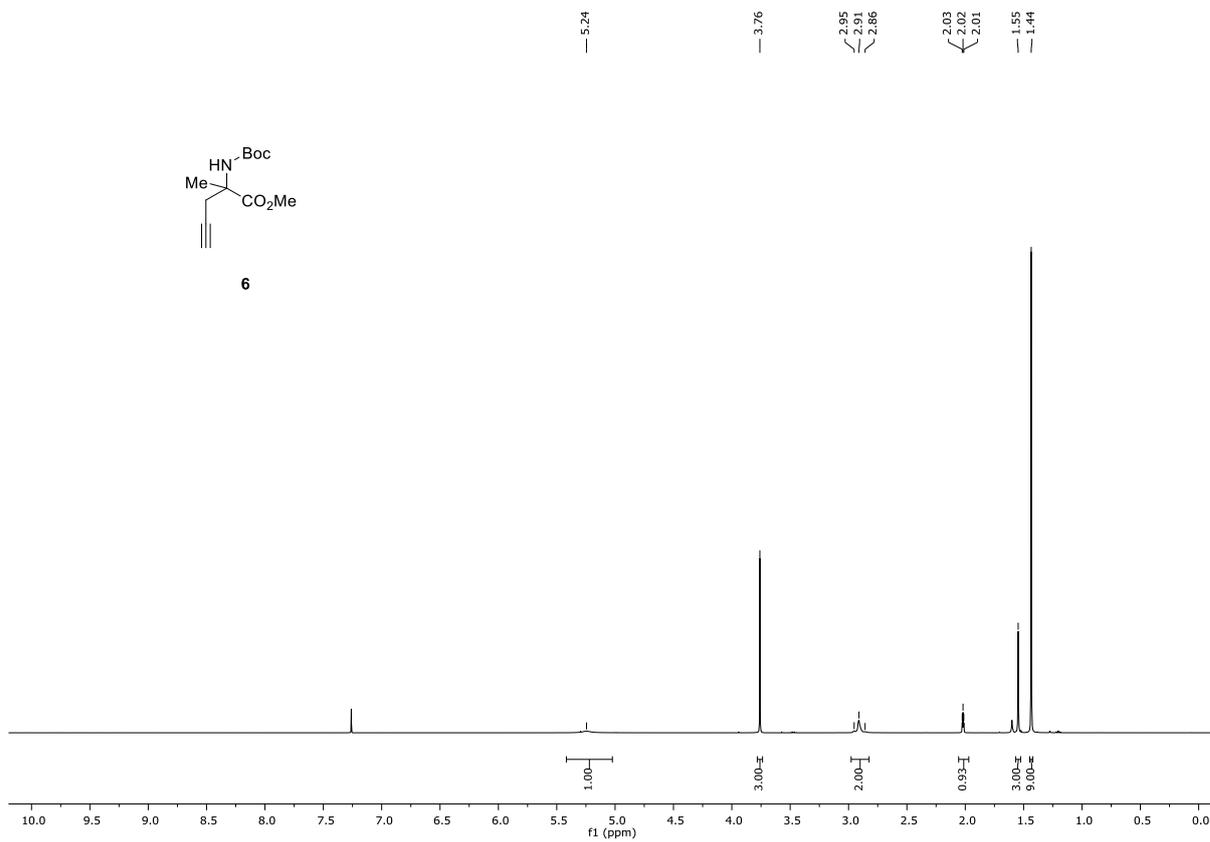
Computational analysis was carried out using the same MOE software package and settings described above in section A. The DOS library compounds were analysed for the following properties: SlogP, molecular weight (MW), number of hydrogen-bond acceptors (HBA), number of hydrogen-bond donors (HBD), number of chiral centres and fraction aromatic (the number of aromatic atoms expressed as a fraction of the total number of heavy atoms).

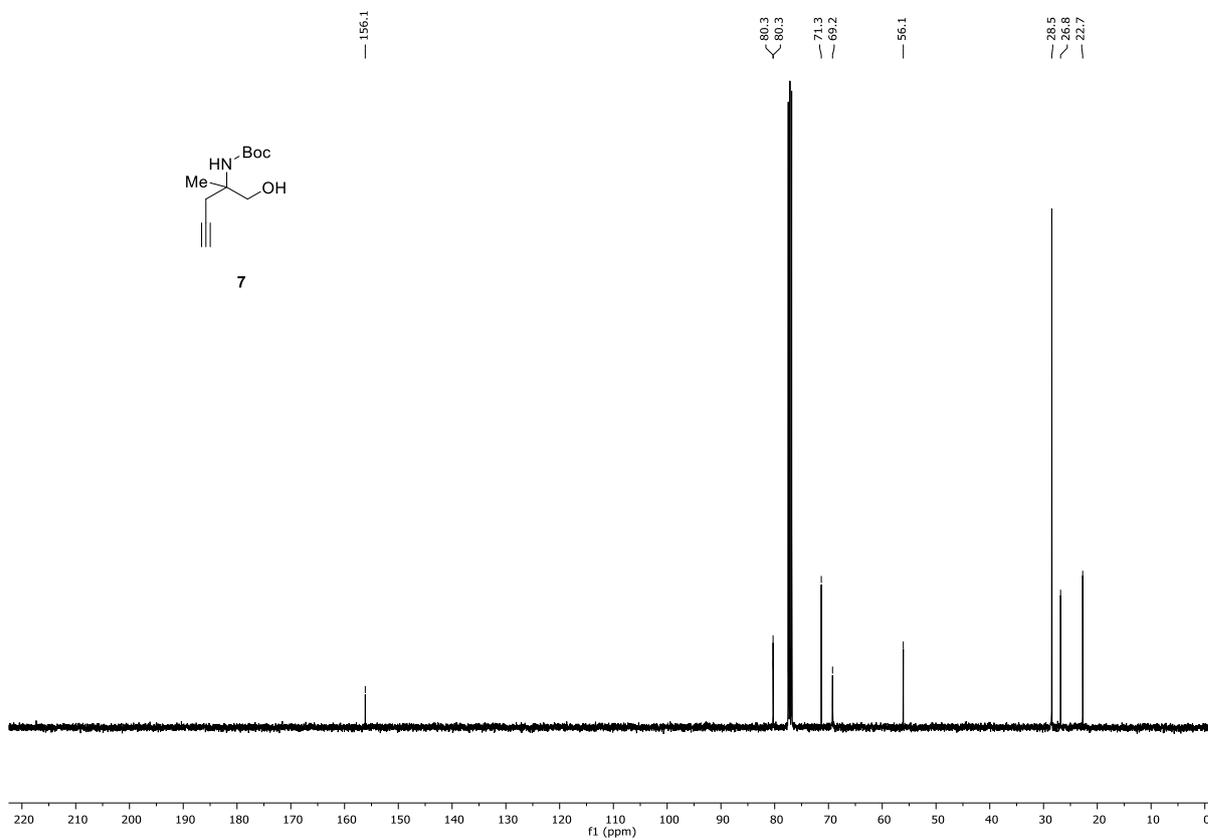
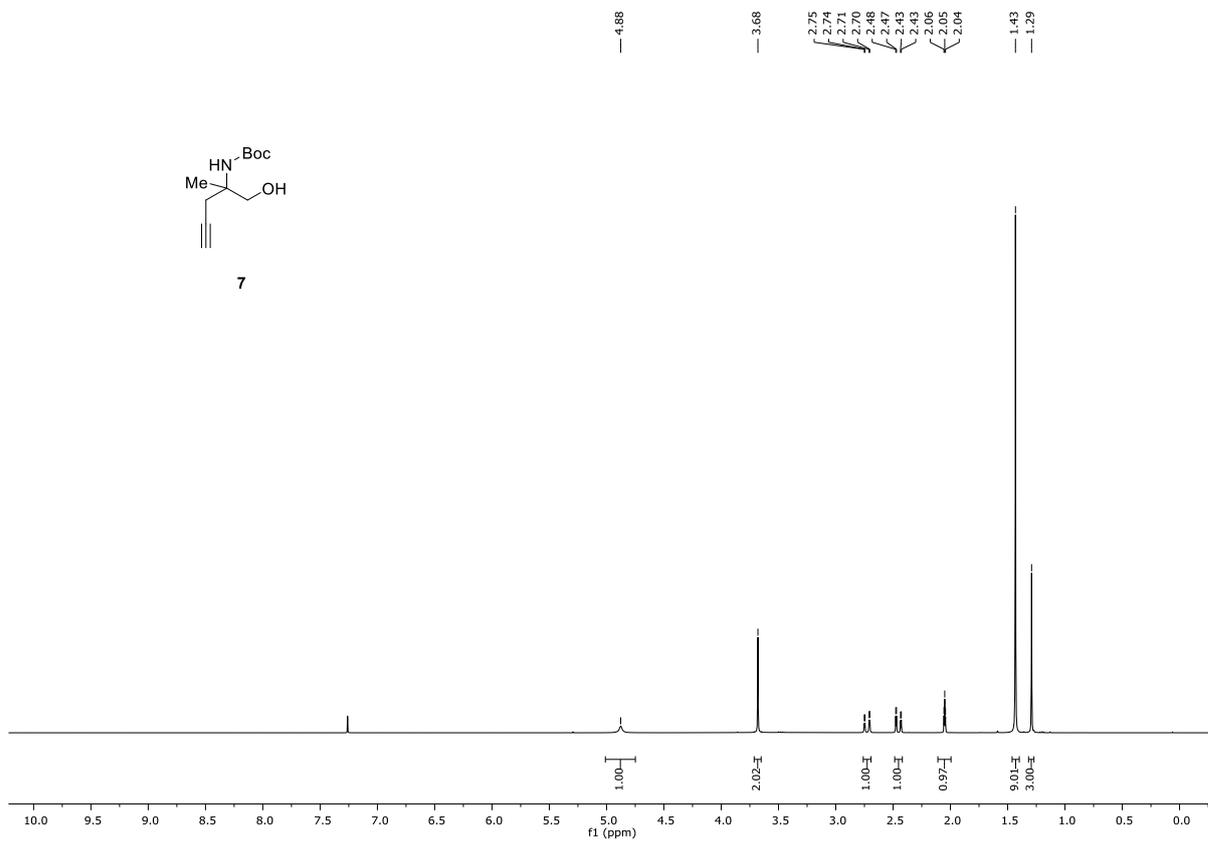
## 8.3 Selected spectra





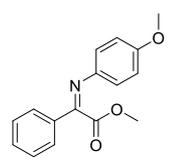




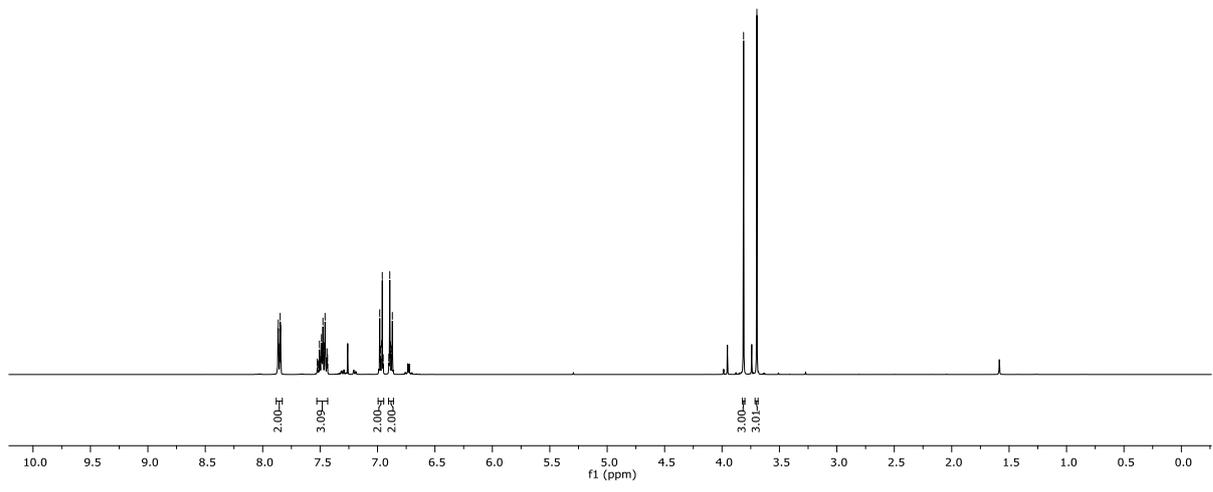


7.87  
7.86  
7.86  
7.85  
7.85  
7.84  
7.82  
7.82  
7.51  
7.50  
7.49  
7.49  
7.48  
7.48  
7.47  
7.46  
7.46  
7.44  
7.44  
6.98  
6.98  
6.96  
6.96  
6.90  
6.89  
6.89  
6.88  
6.87

3.81  
3.70

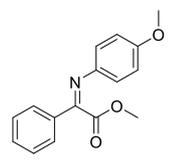


8

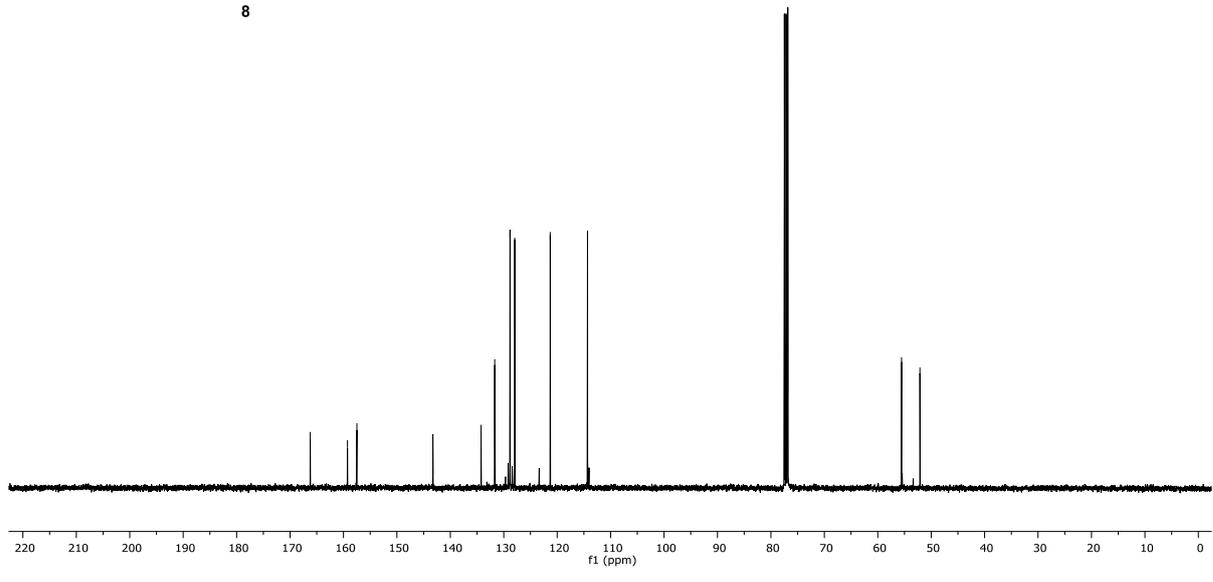


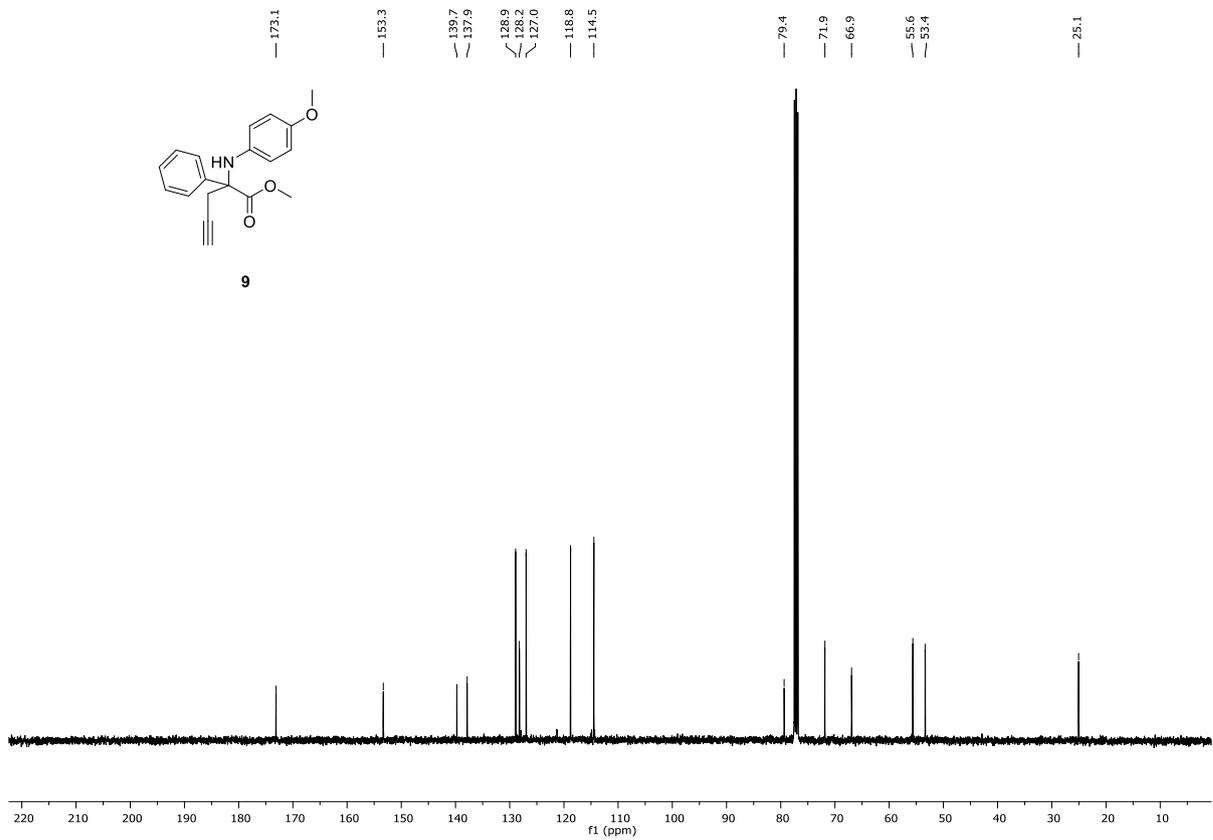
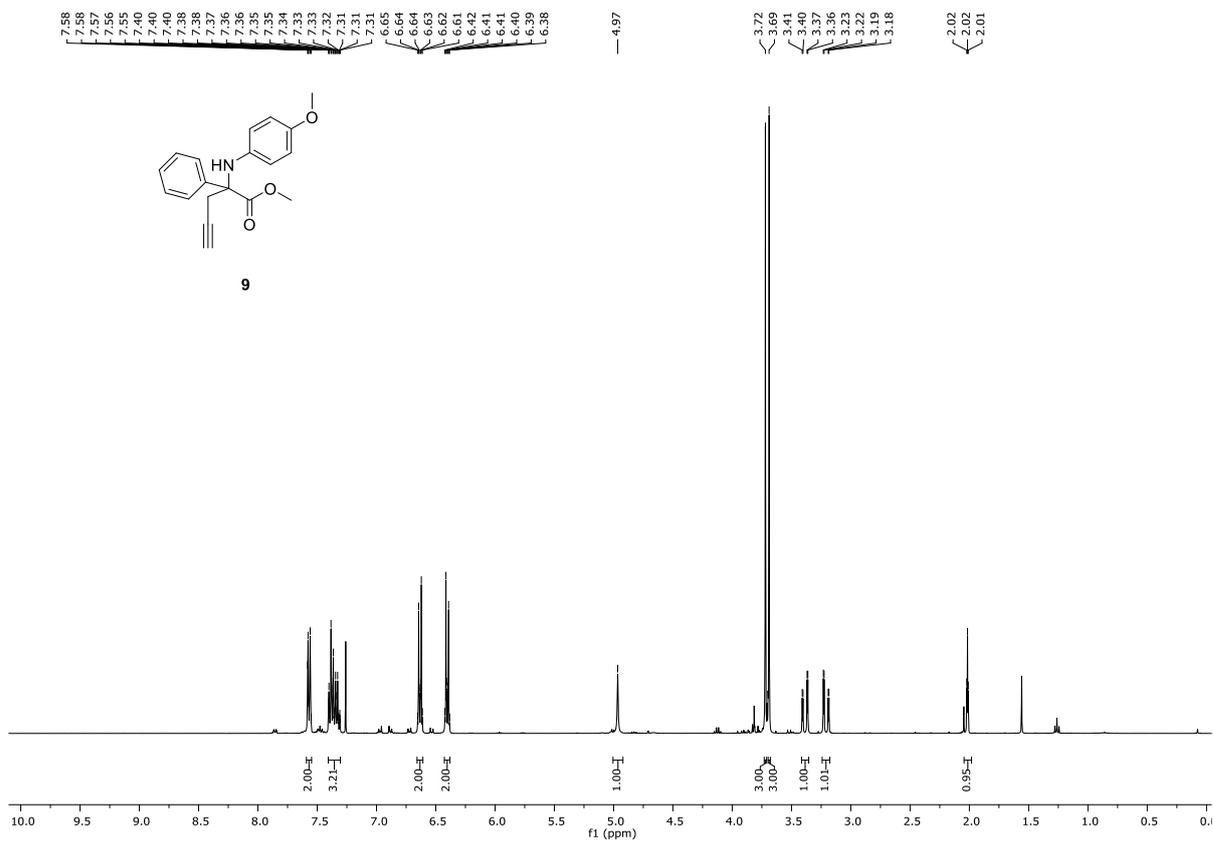
166.2  
159.3  
157.5  
134.3  
131.7  
128.8  
127.9  
121.3  
114.3

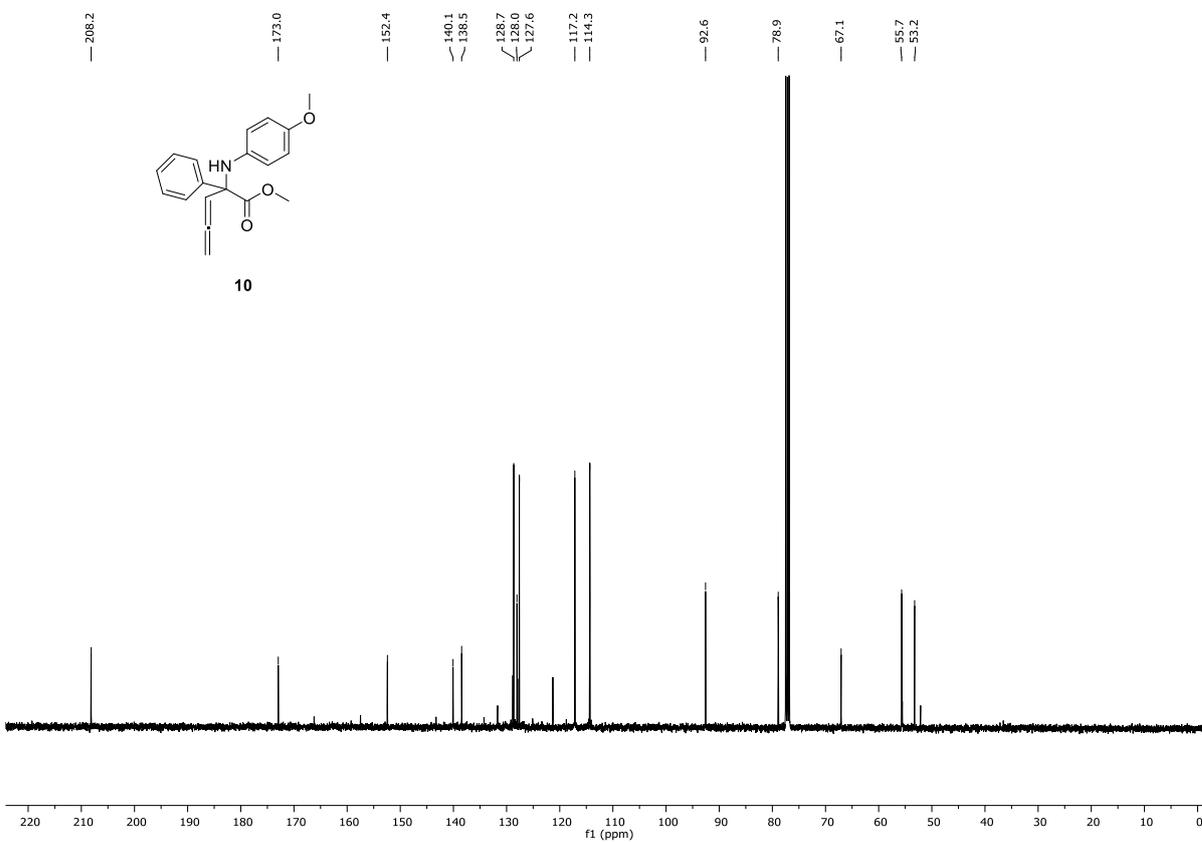
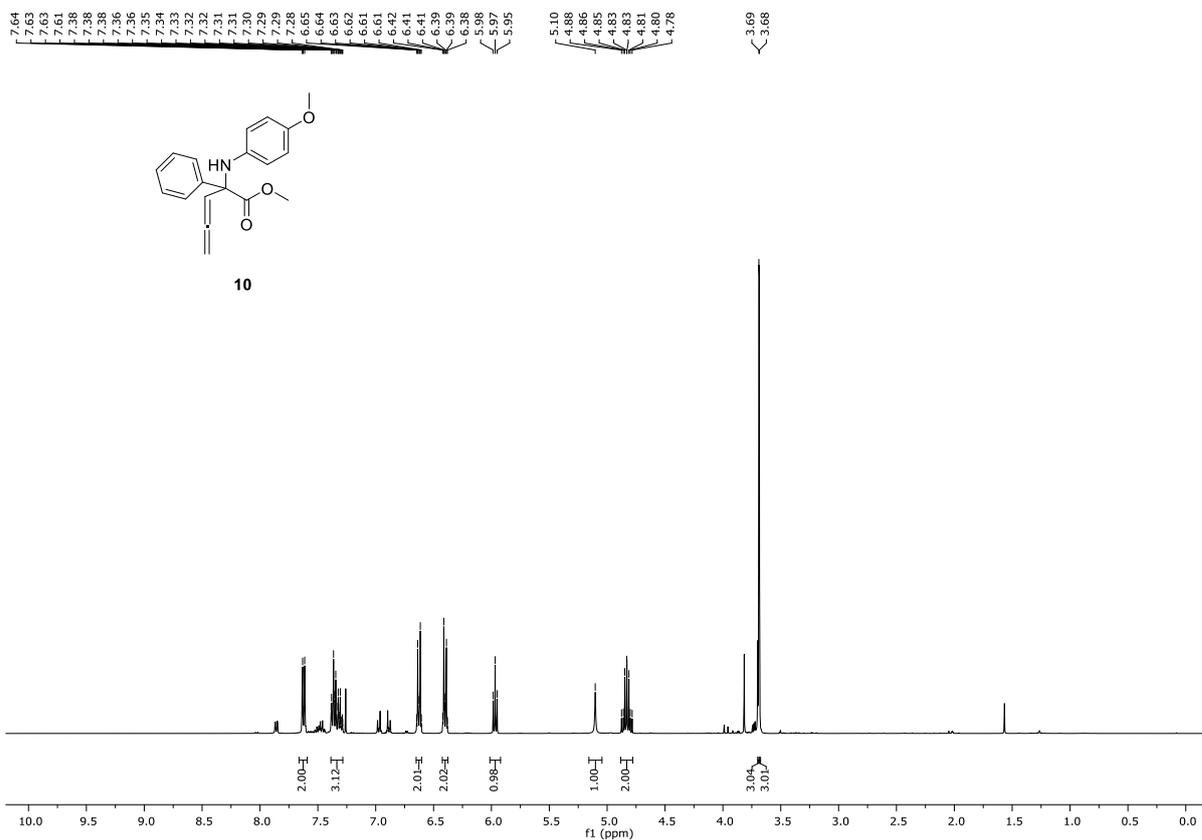
55.6  
52.1

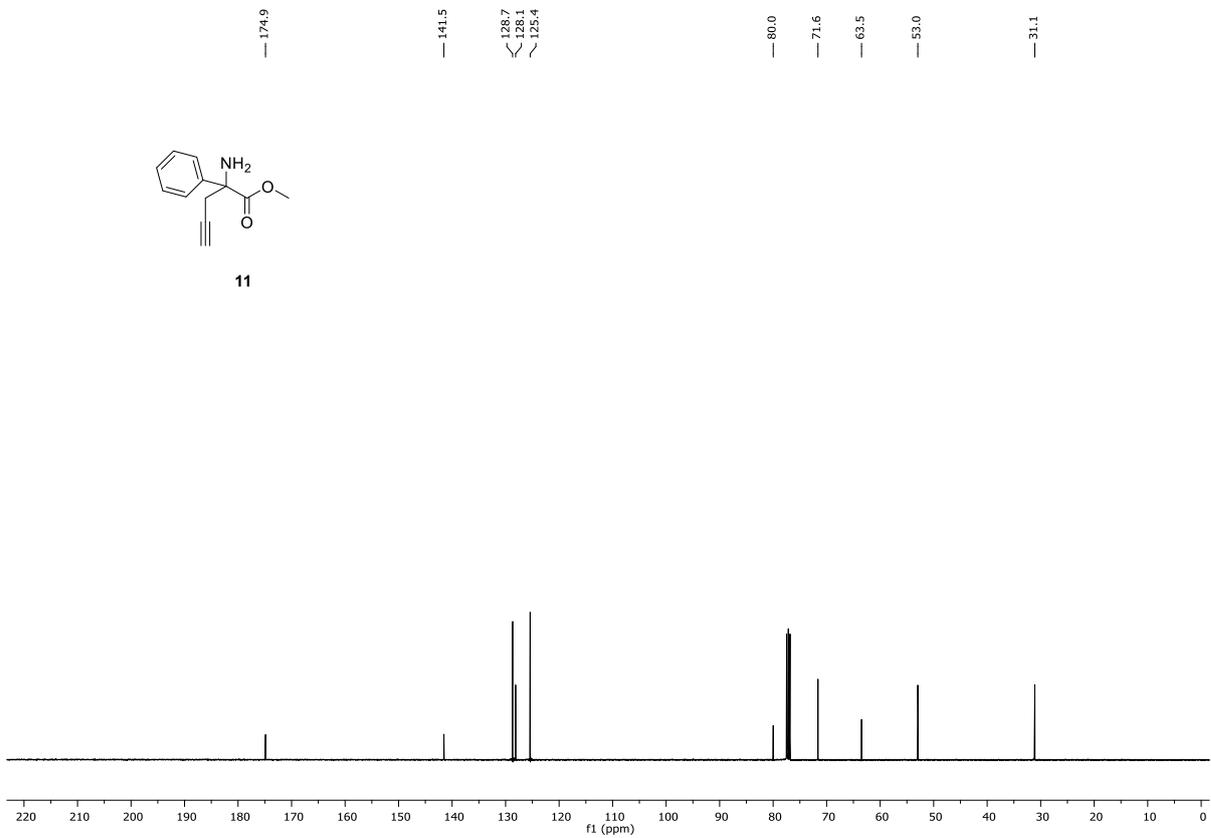
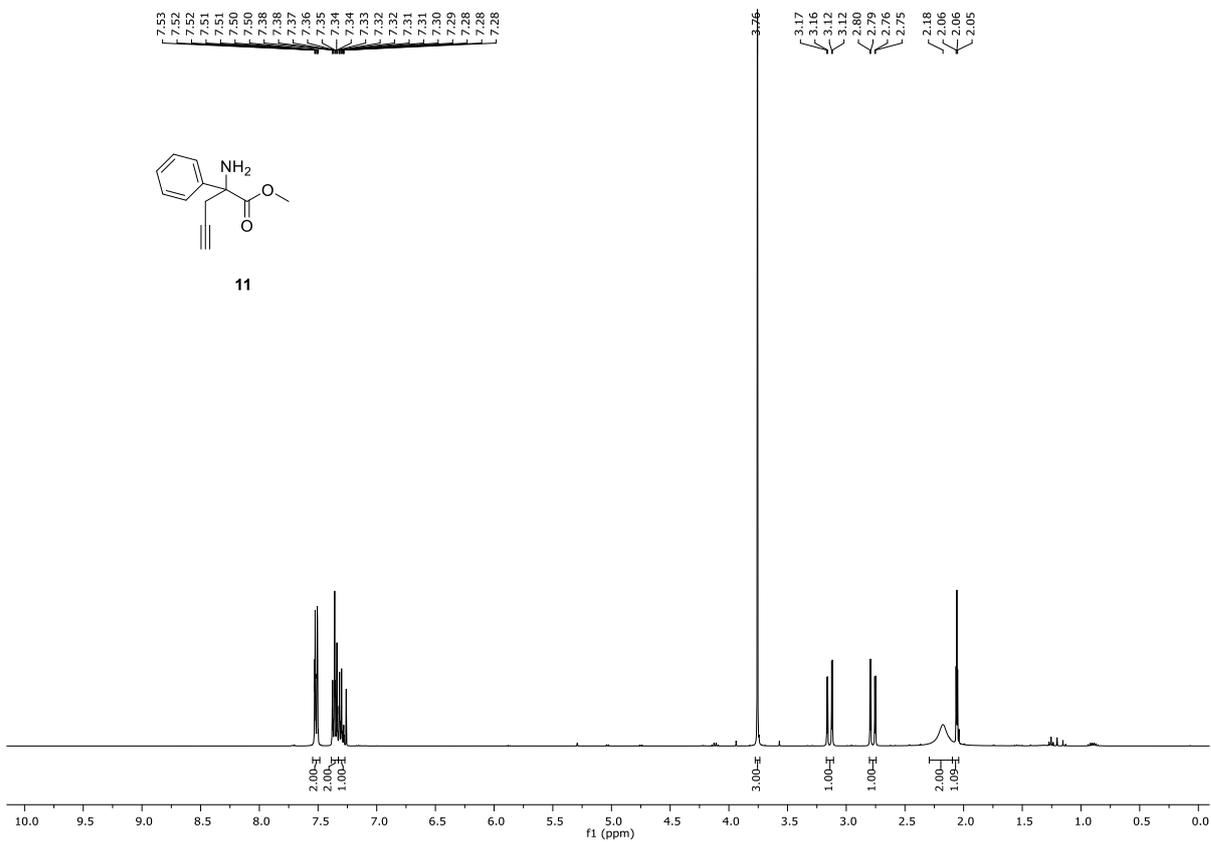


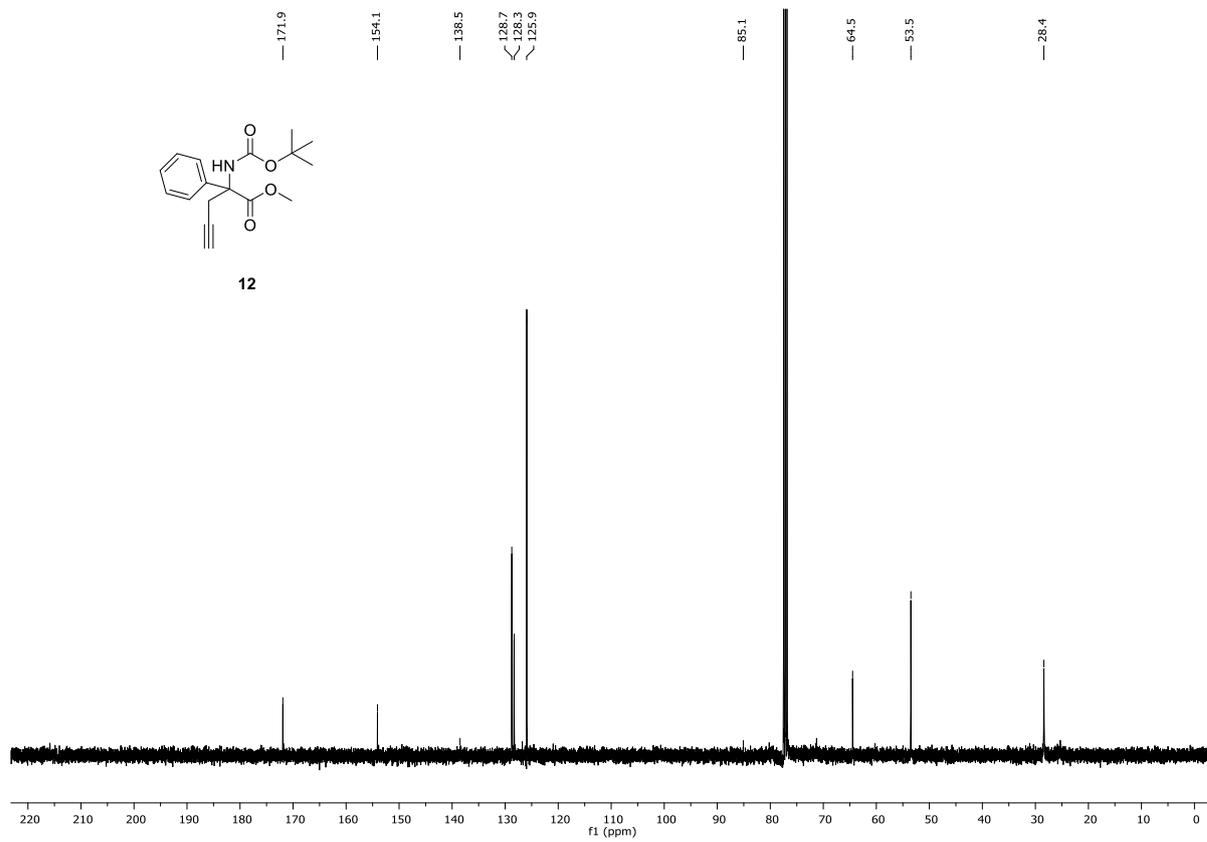
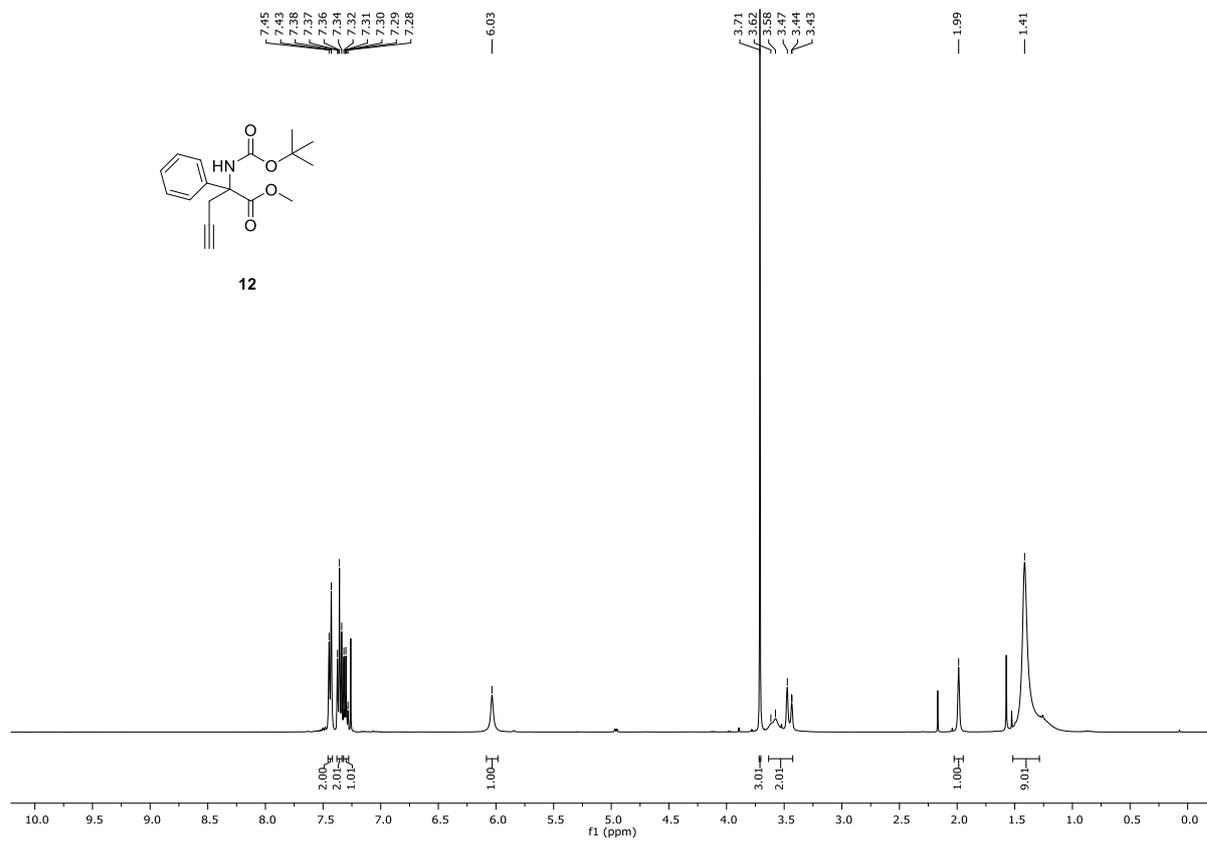
8

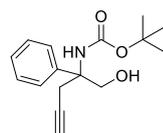












13

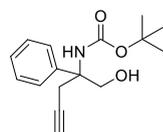
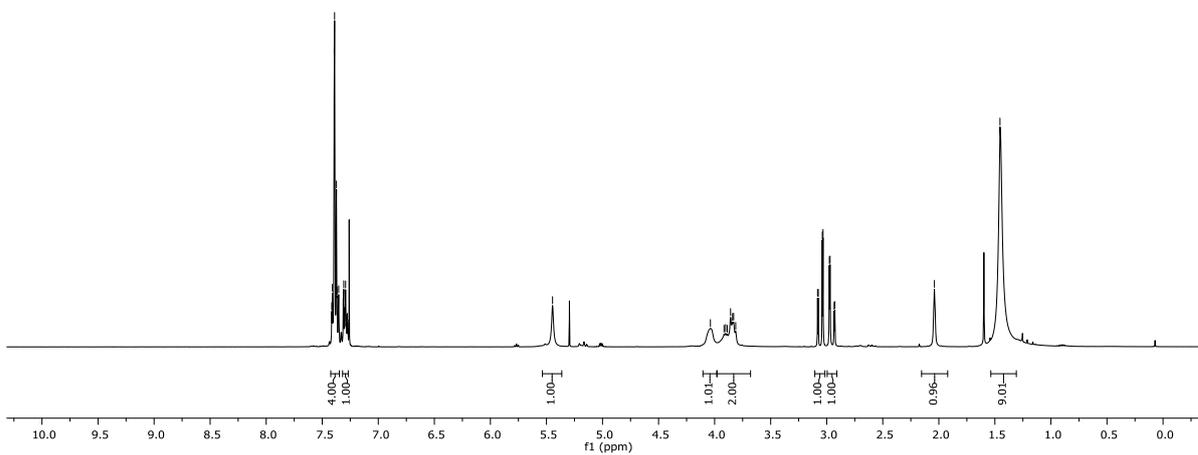
7.42  
7.41  
7.40  
7.39  
7.38  
7.37  
7.36  
7.35  
7.31  
7.31  
7.30  
7.30  
7.29  
7.28  
7.28  
7.27

5.45

4.04  
3.92  
3.90  
3.89  
3.86  
3.84  
3.83  
3.81  
3.08  
3.06  
3.04  
3.03  
2.98  
2.97  
2.93  
2.93

2.04

1.45



13

156.1

141.2

128.7

127.7

126.0

80.5

79.9

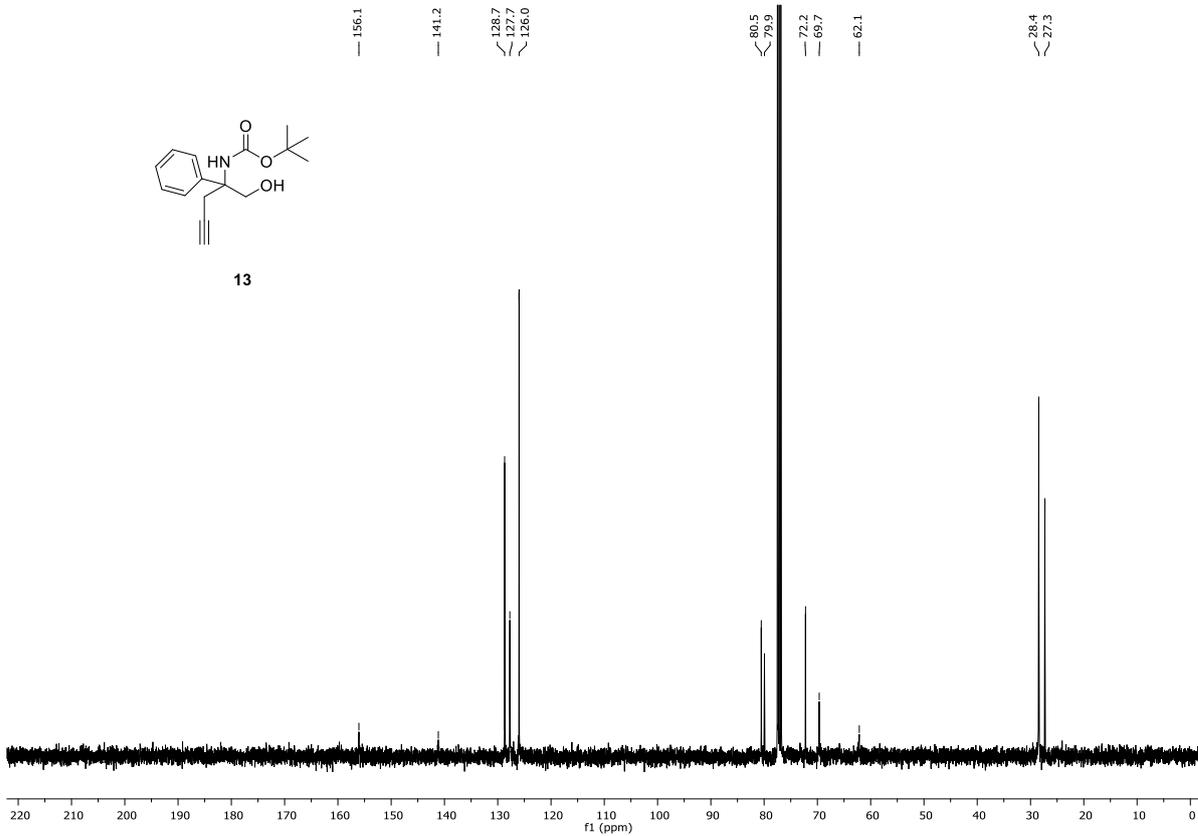
72.2

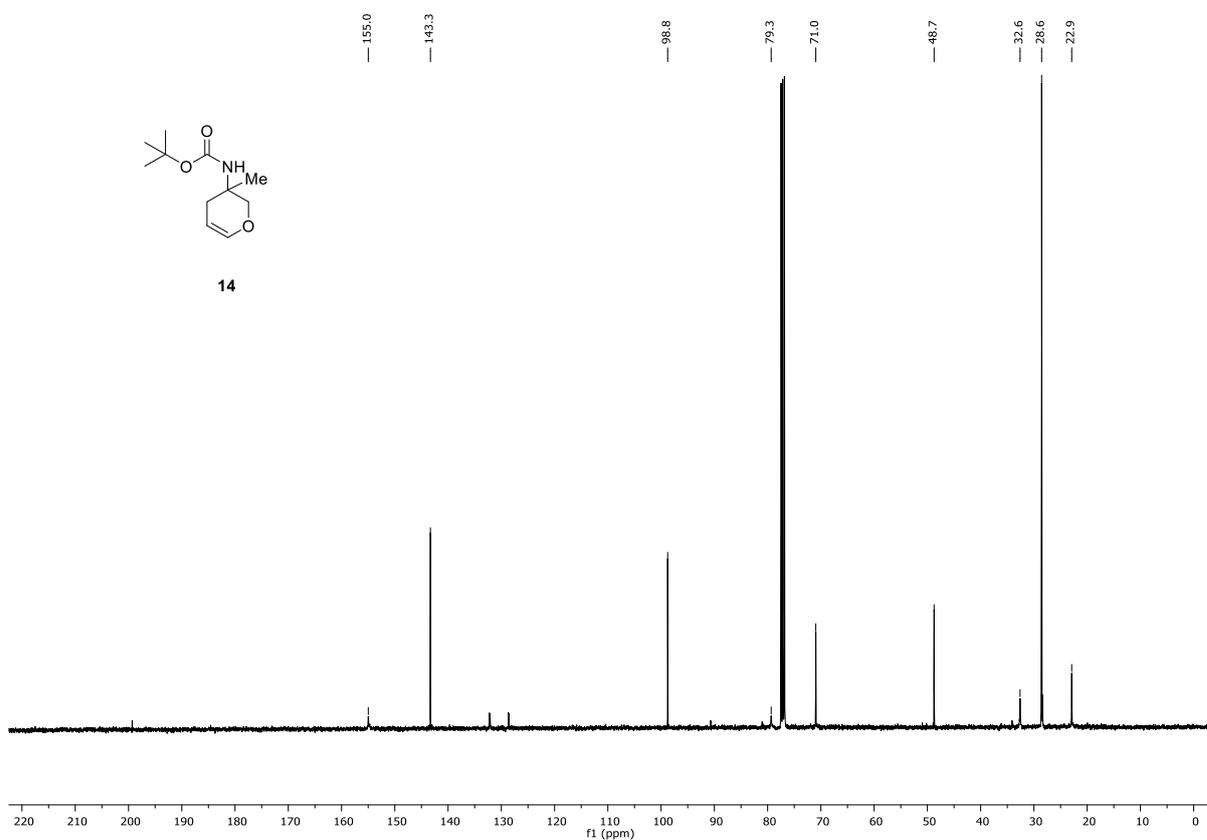
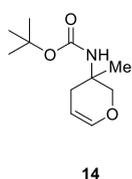
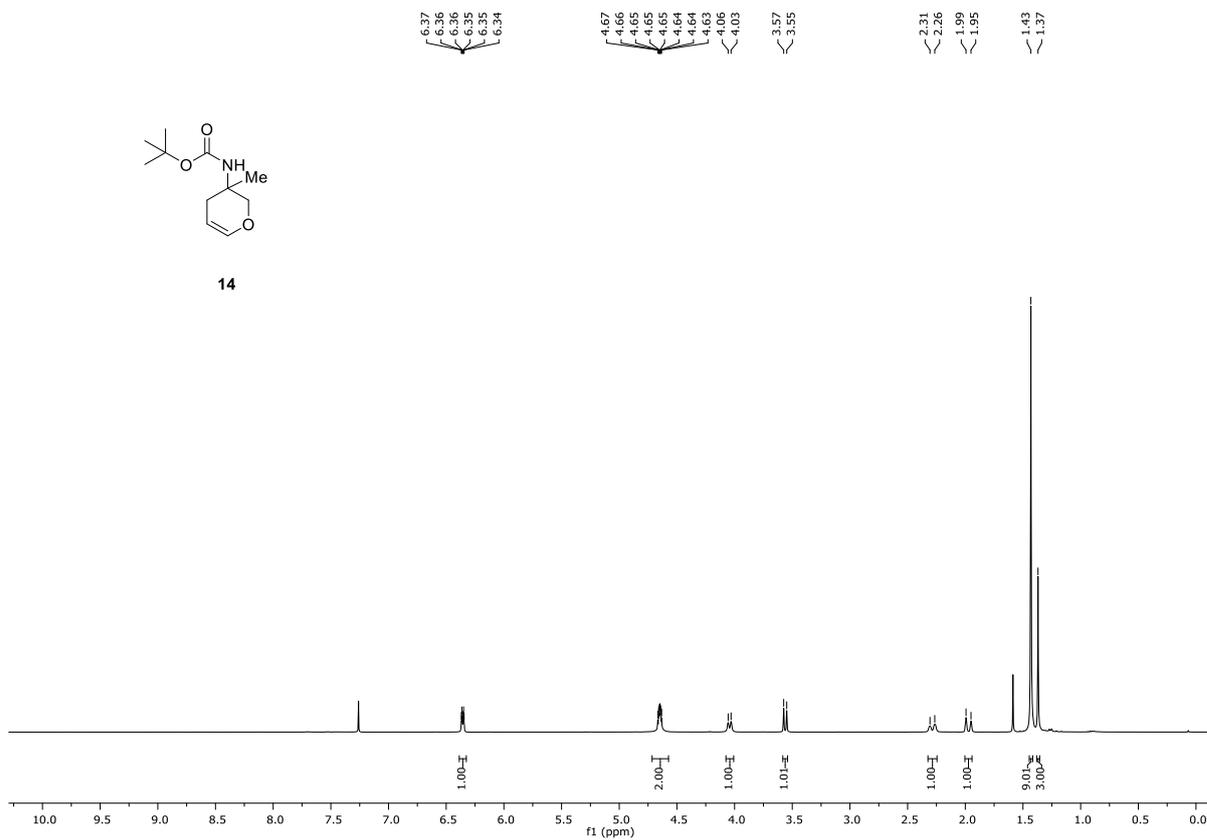
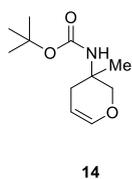
69.7

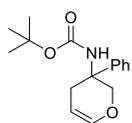
62.1

28.4

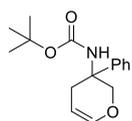
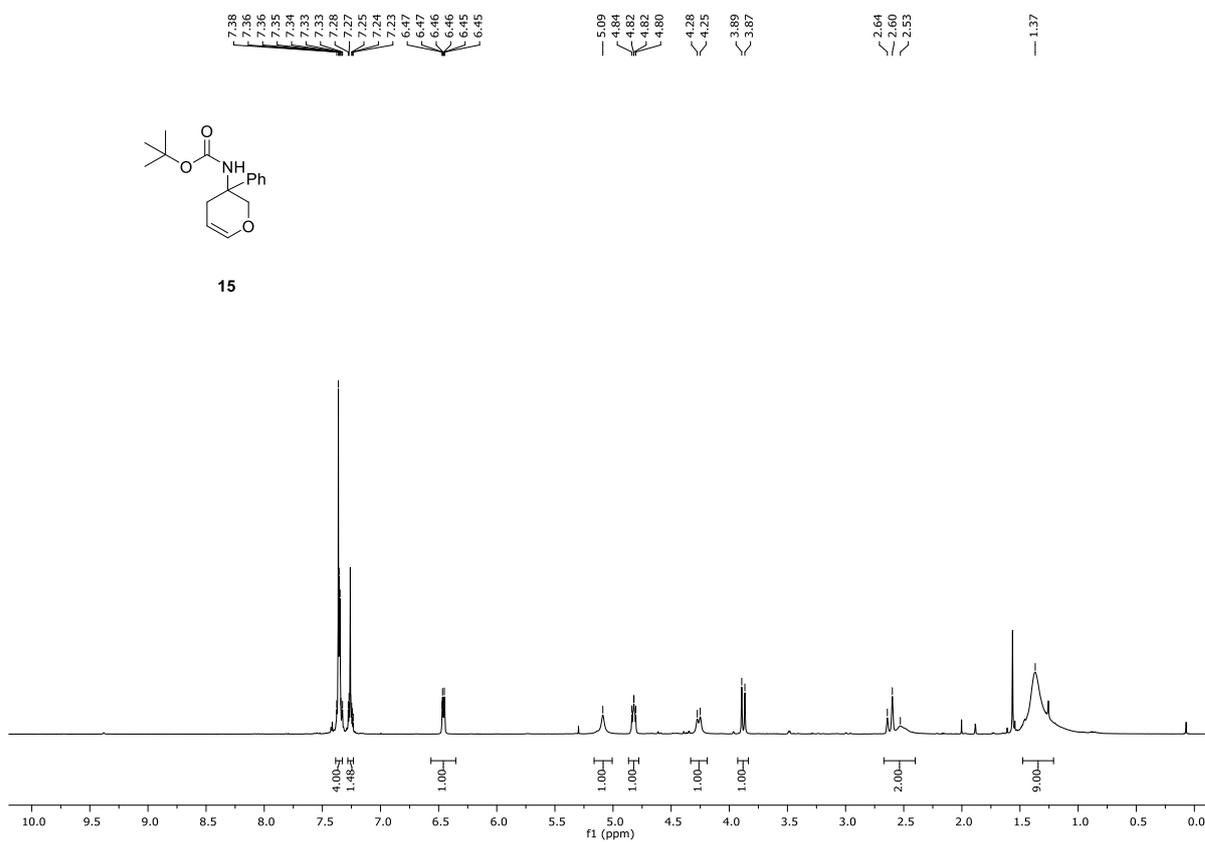
27.3



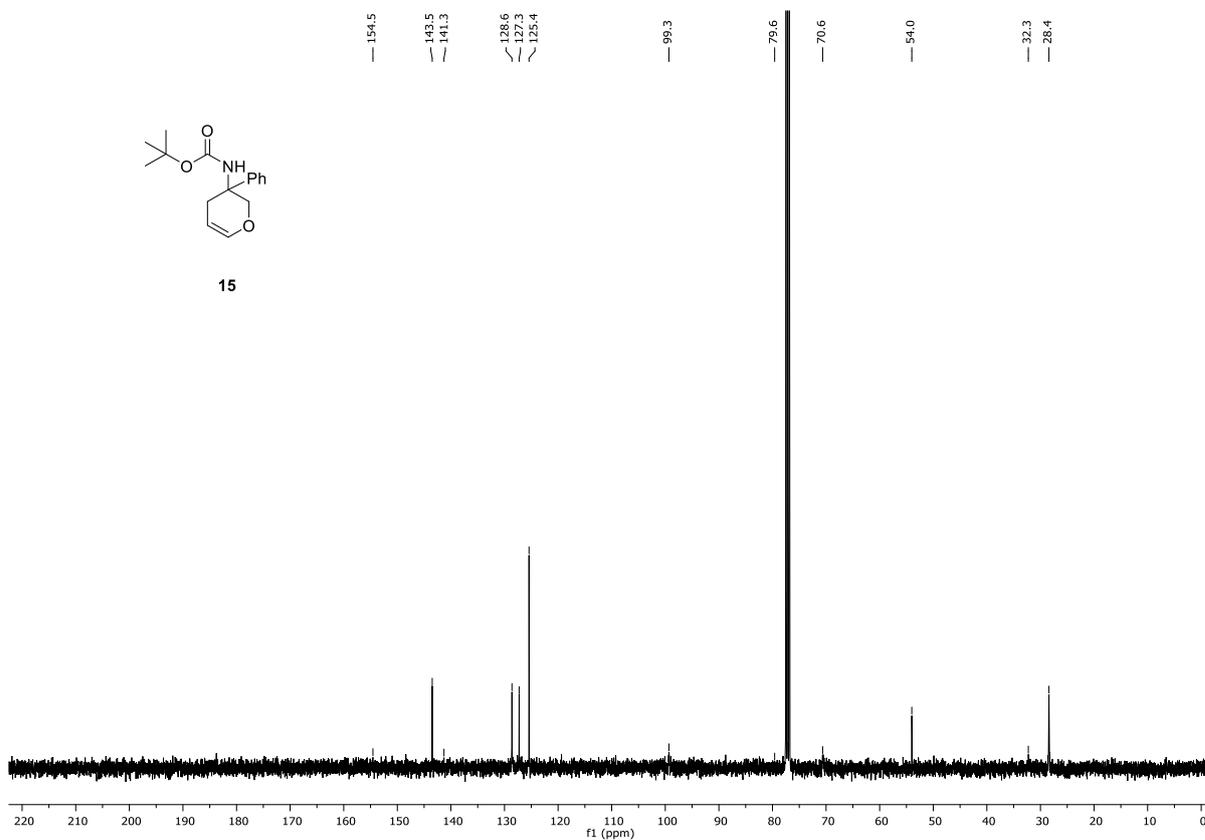


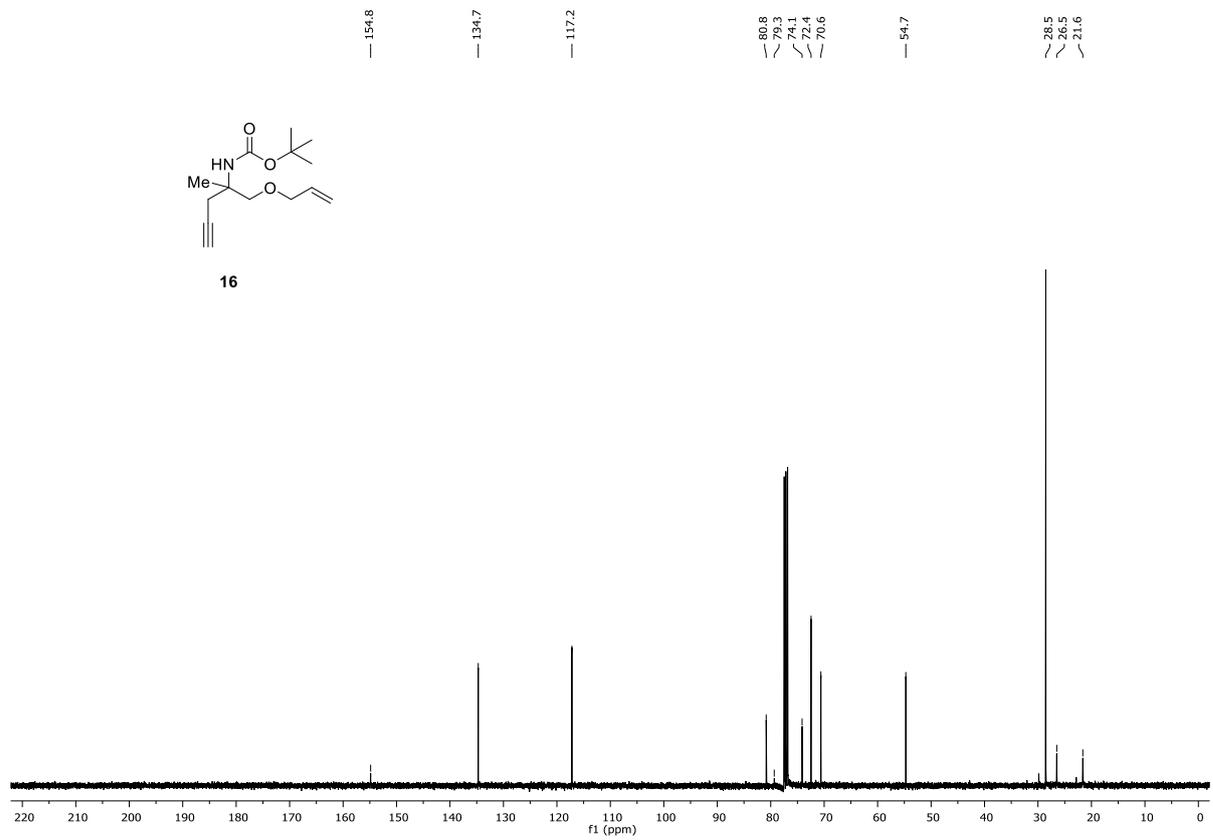
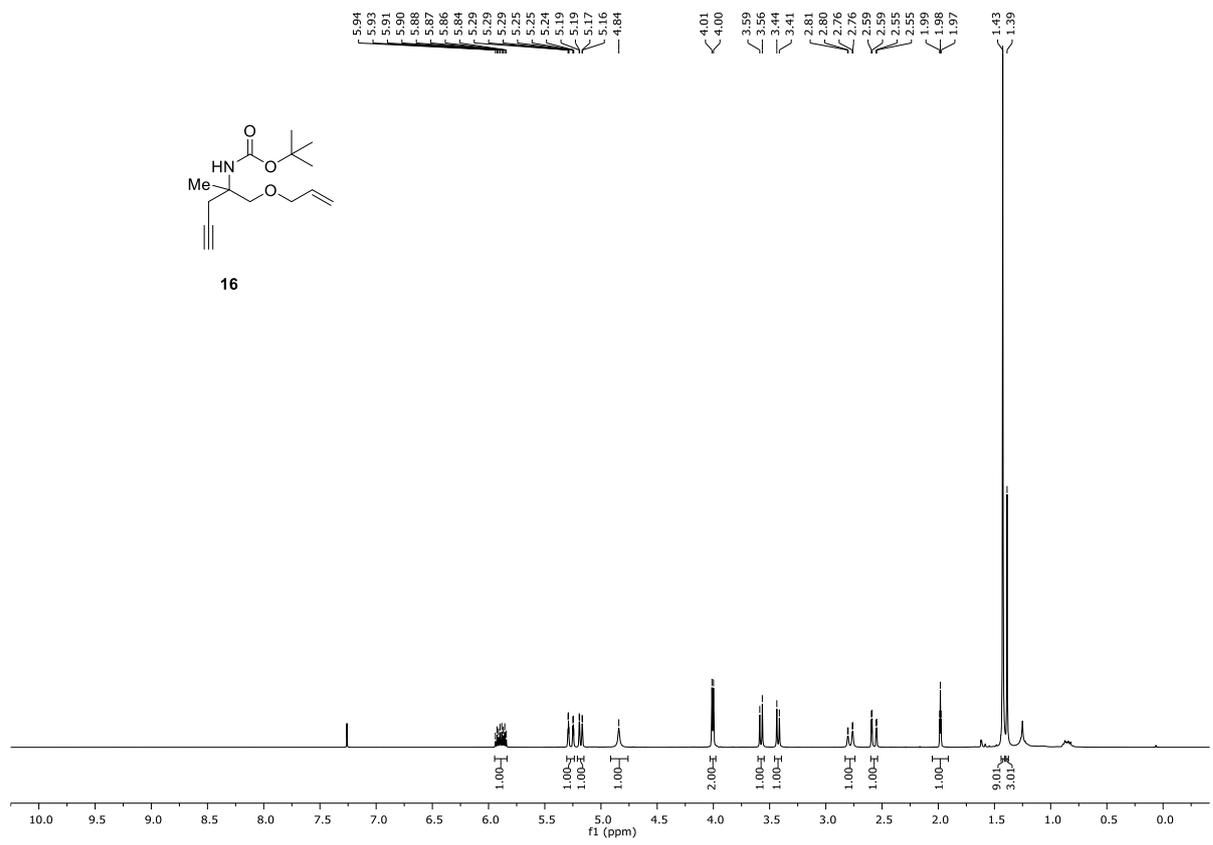


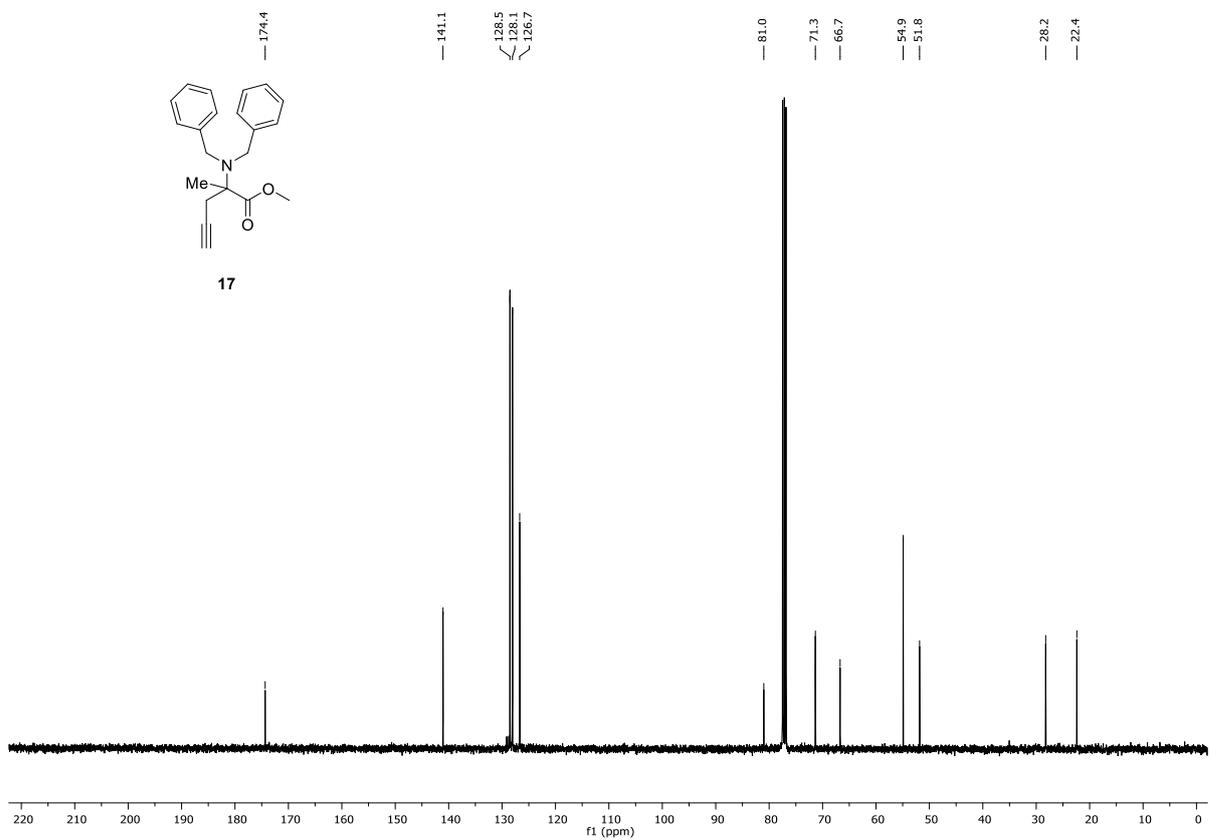
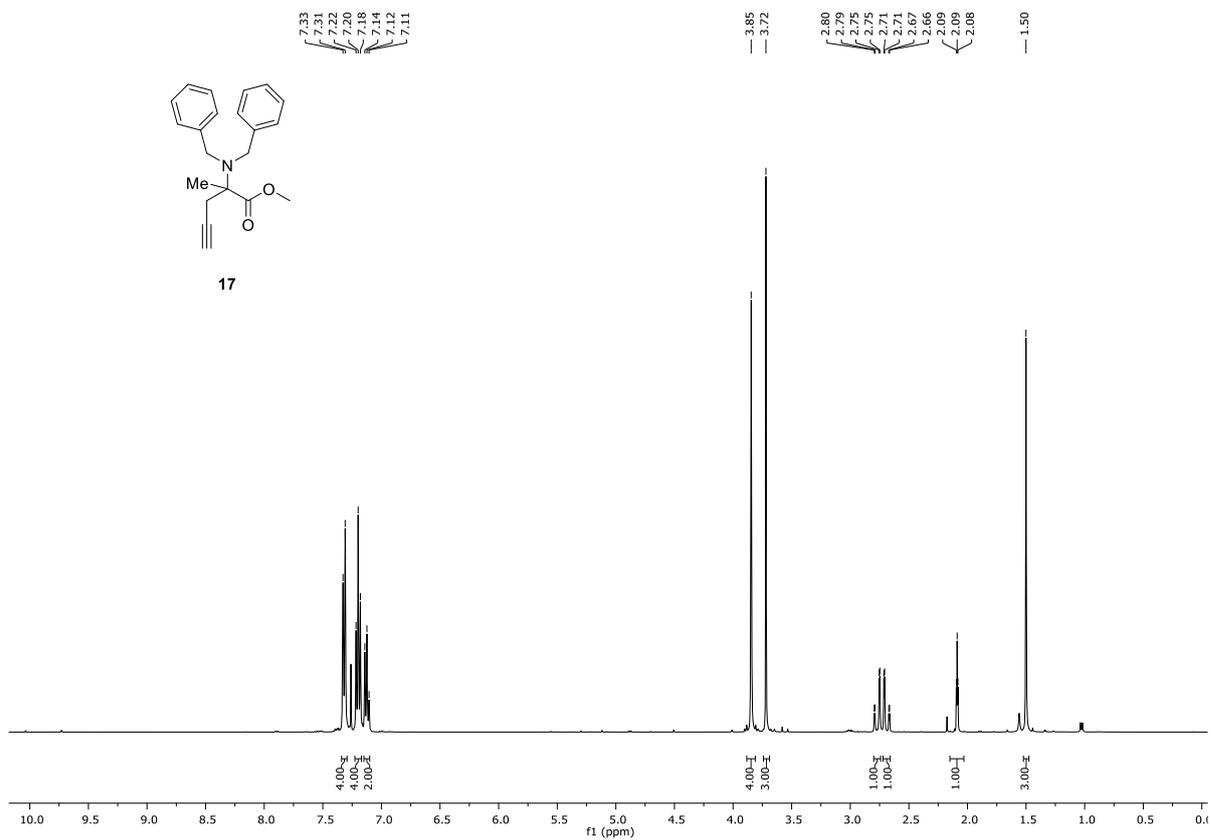
15

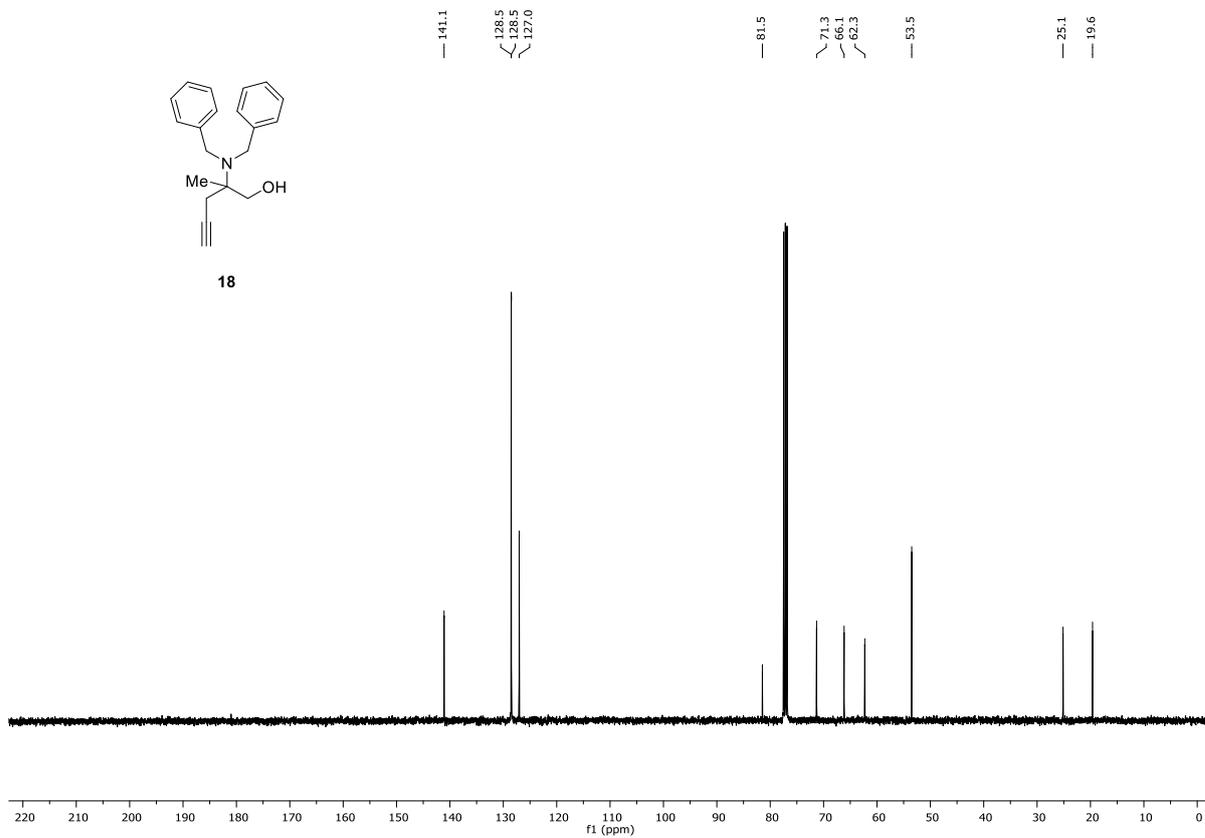
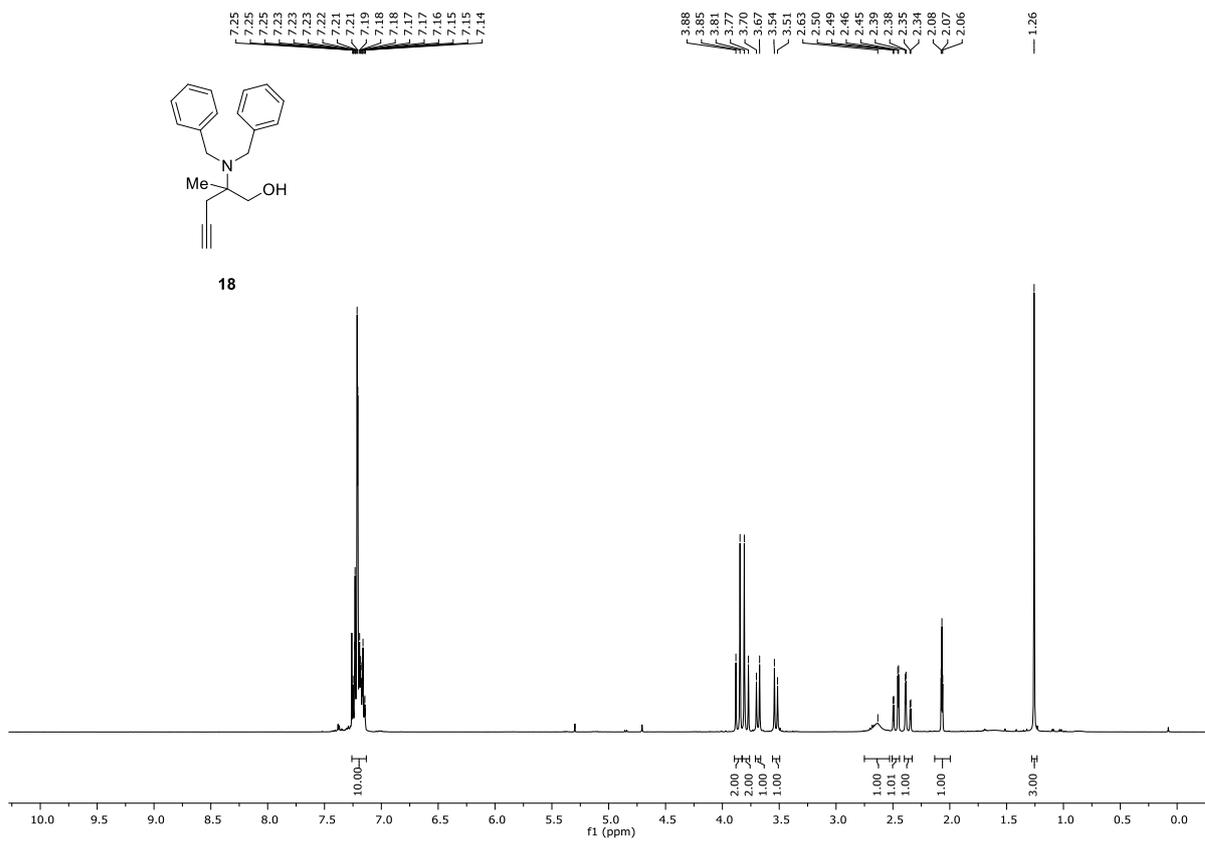


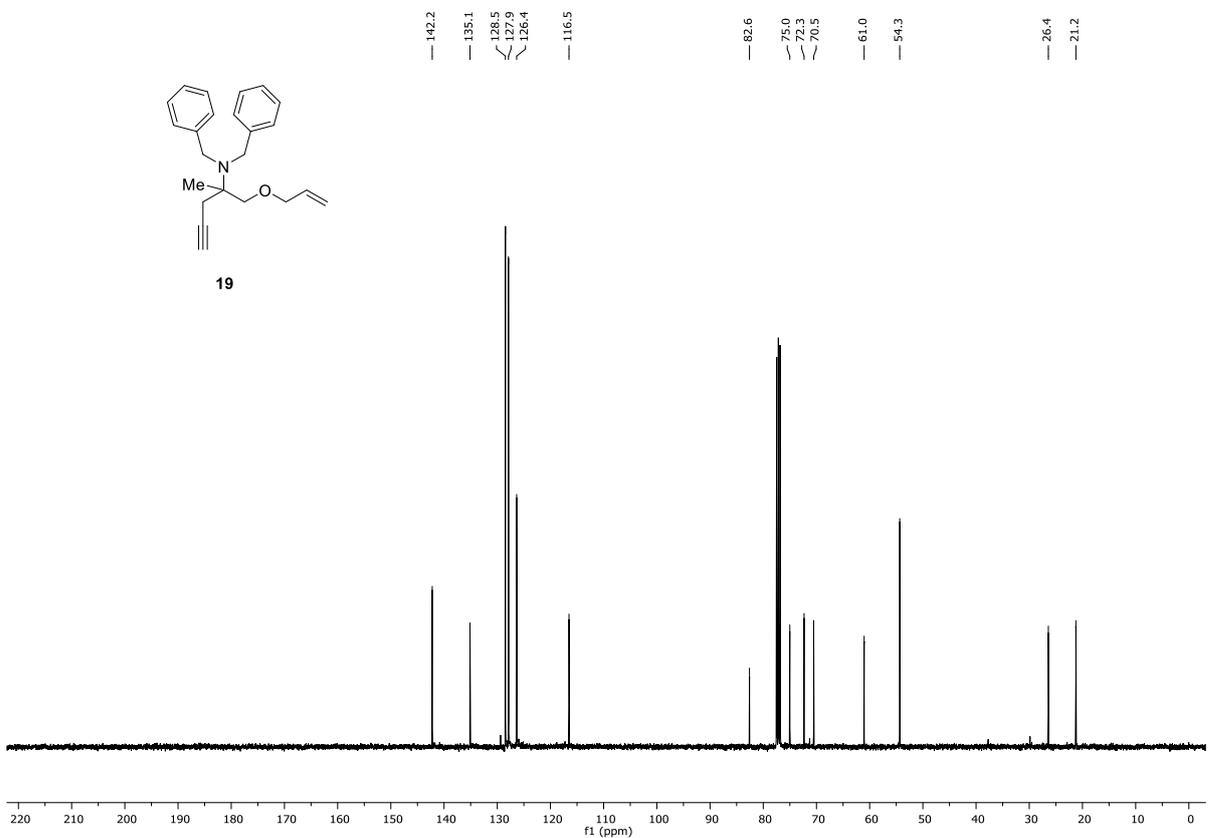
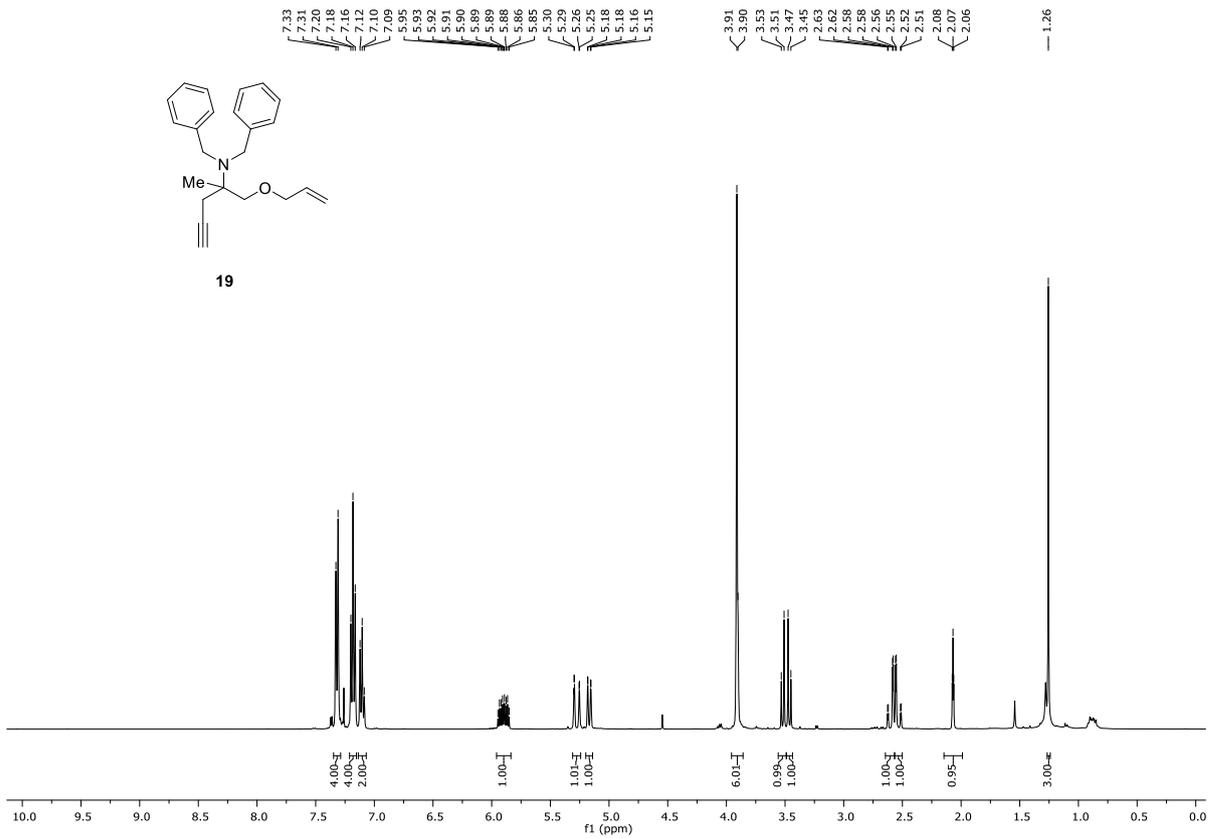
15

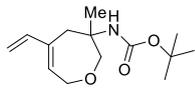




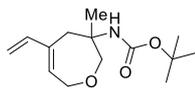
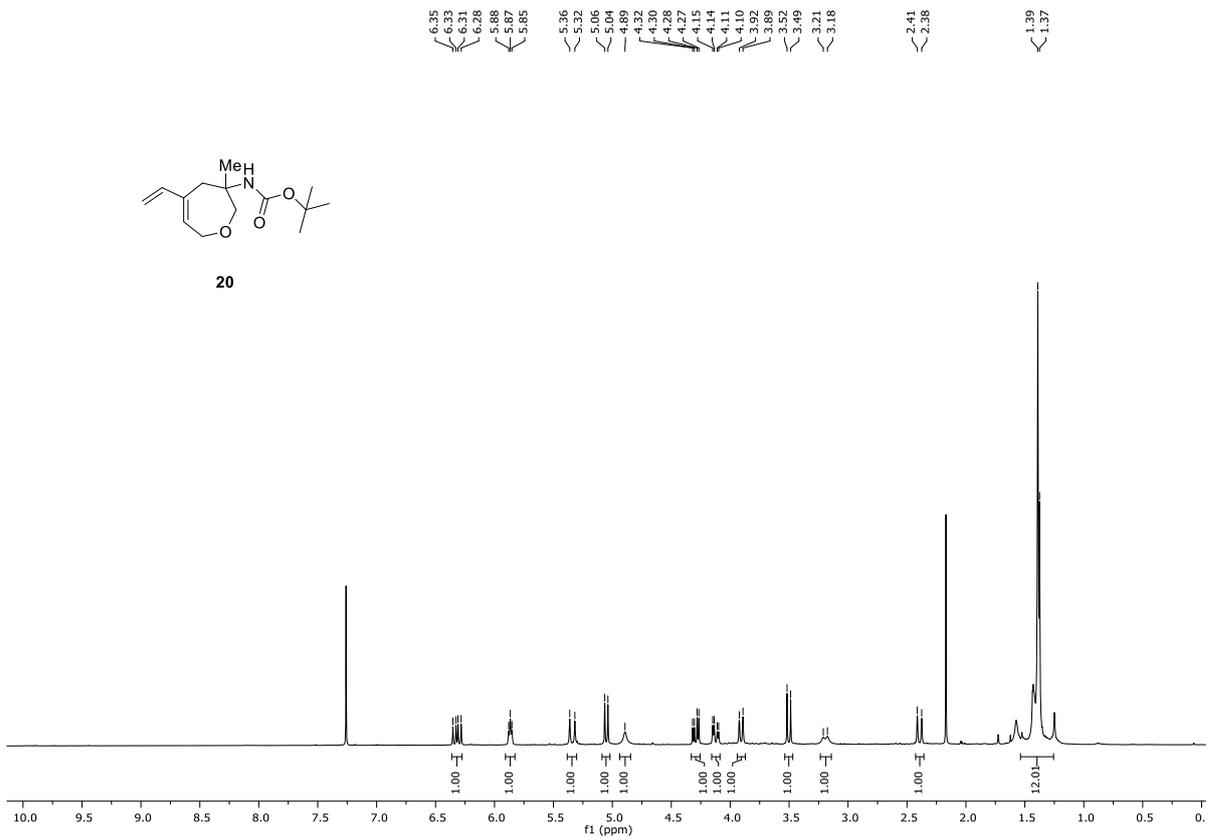




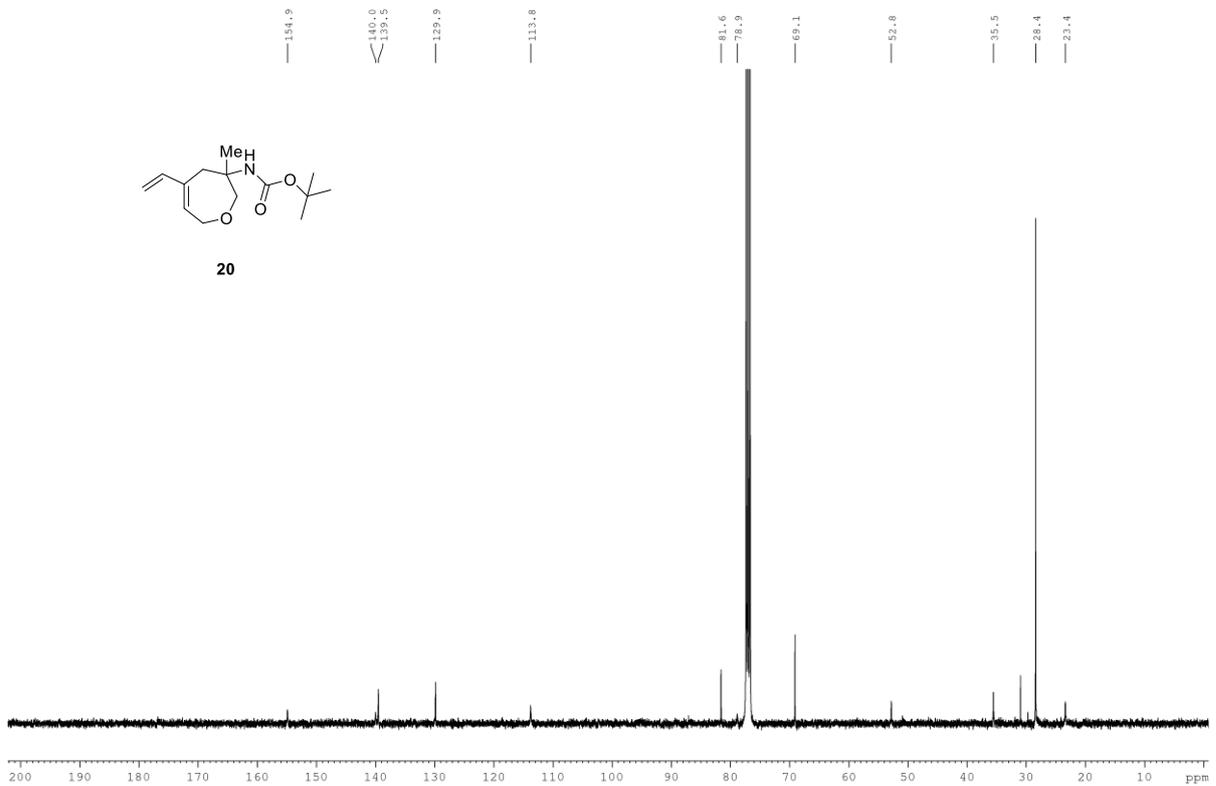


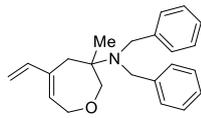


20

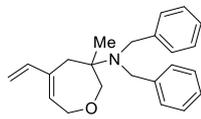
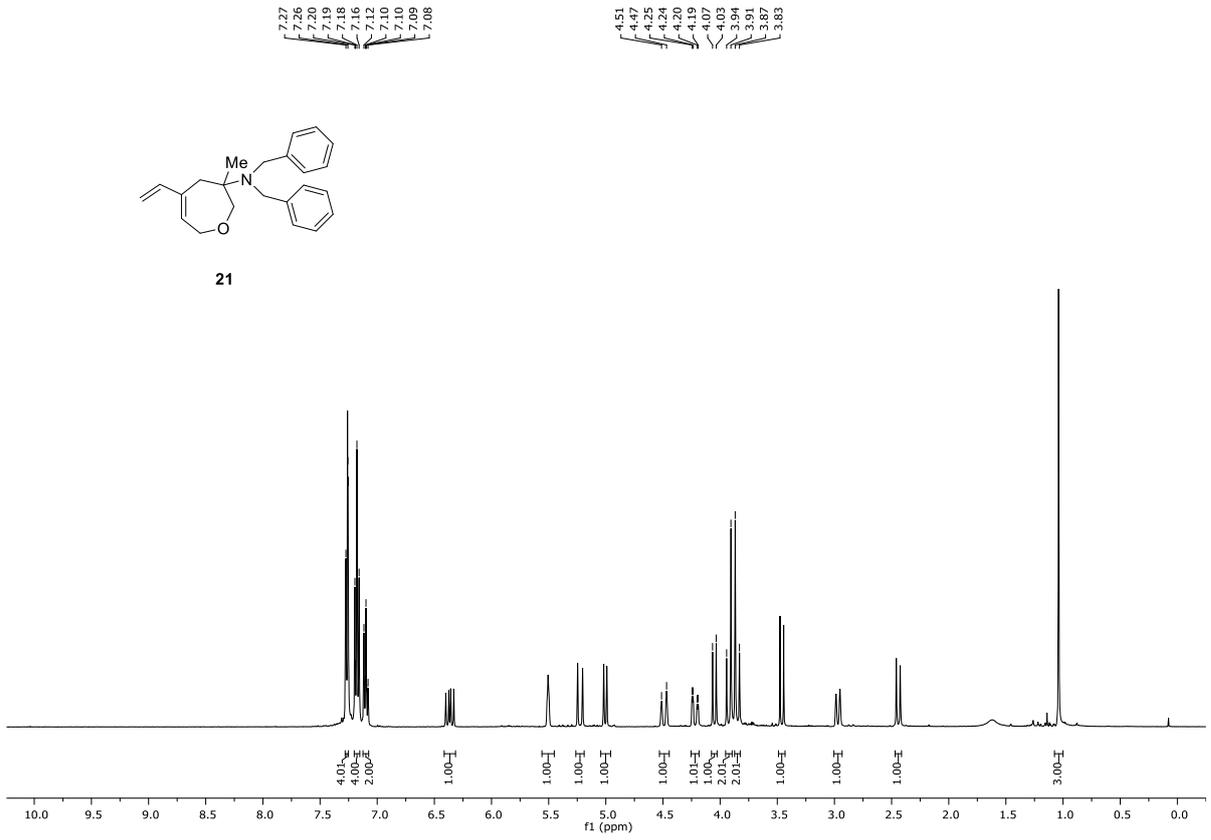


20

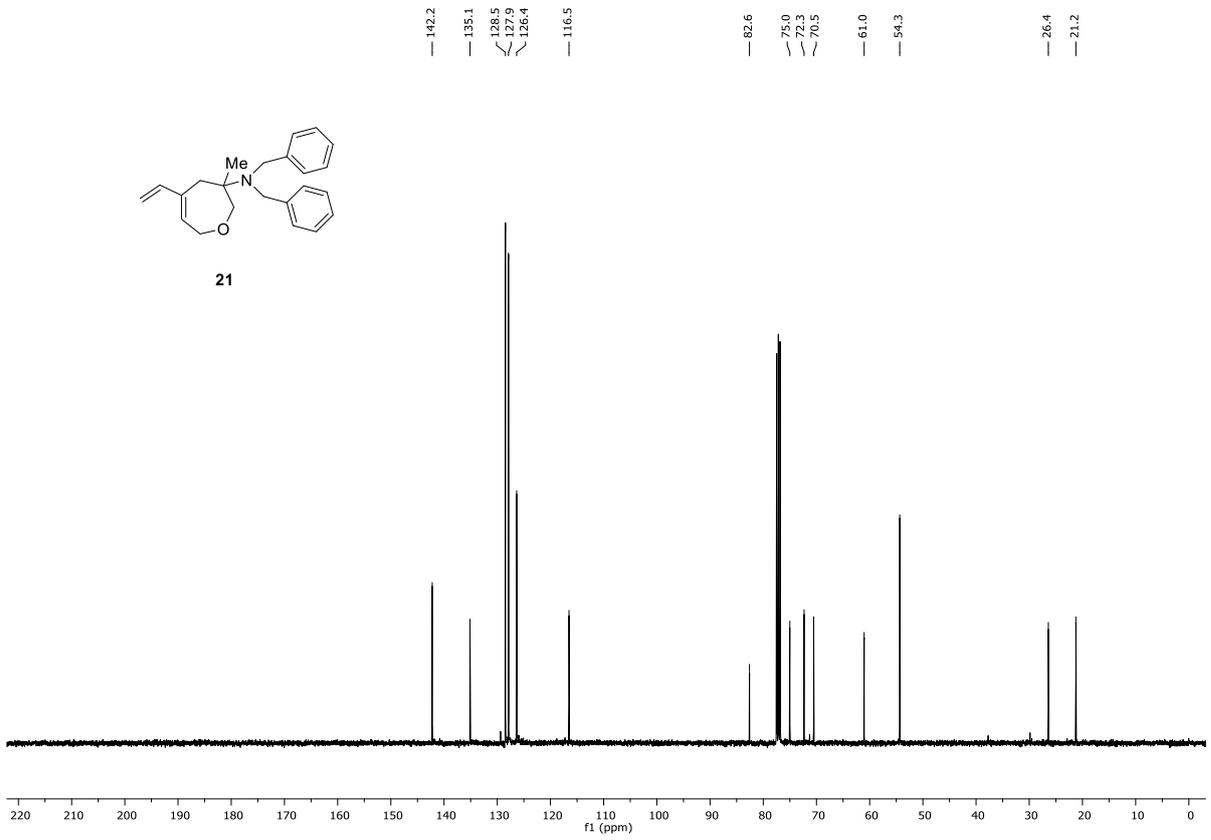


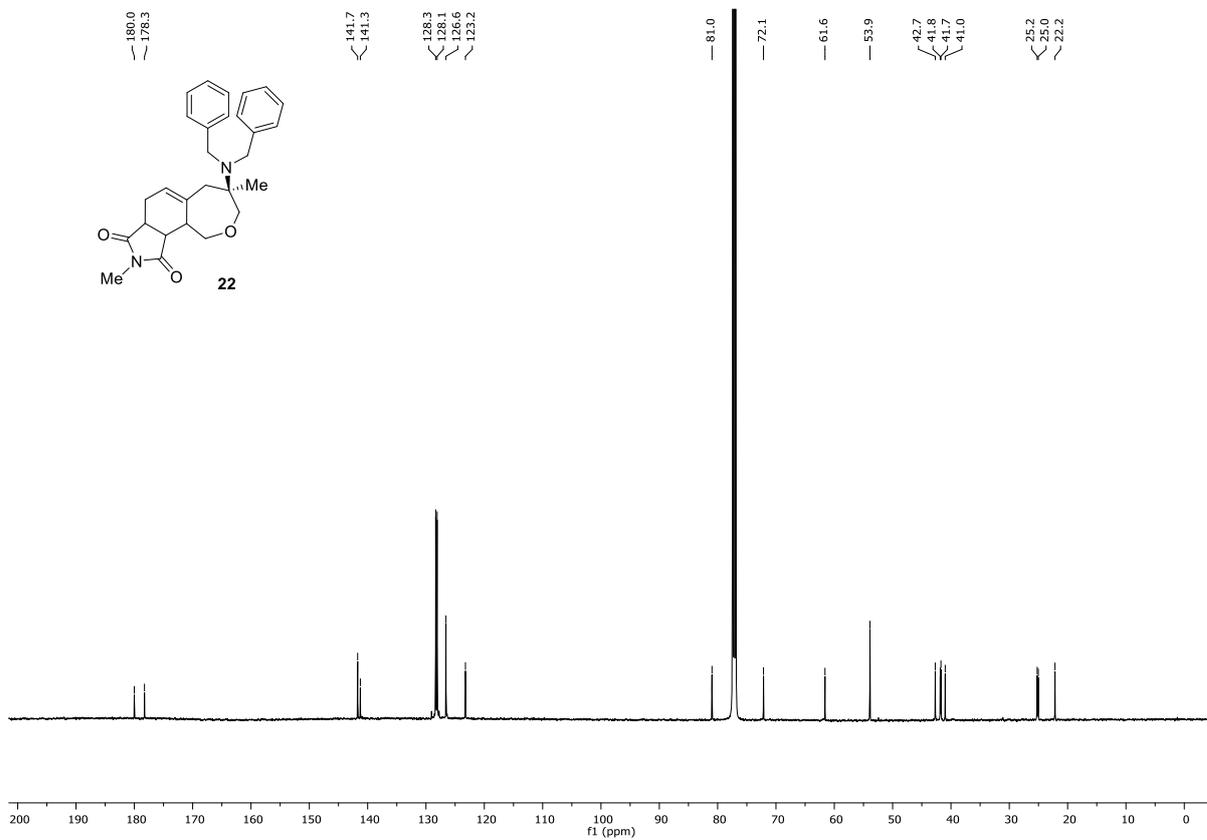
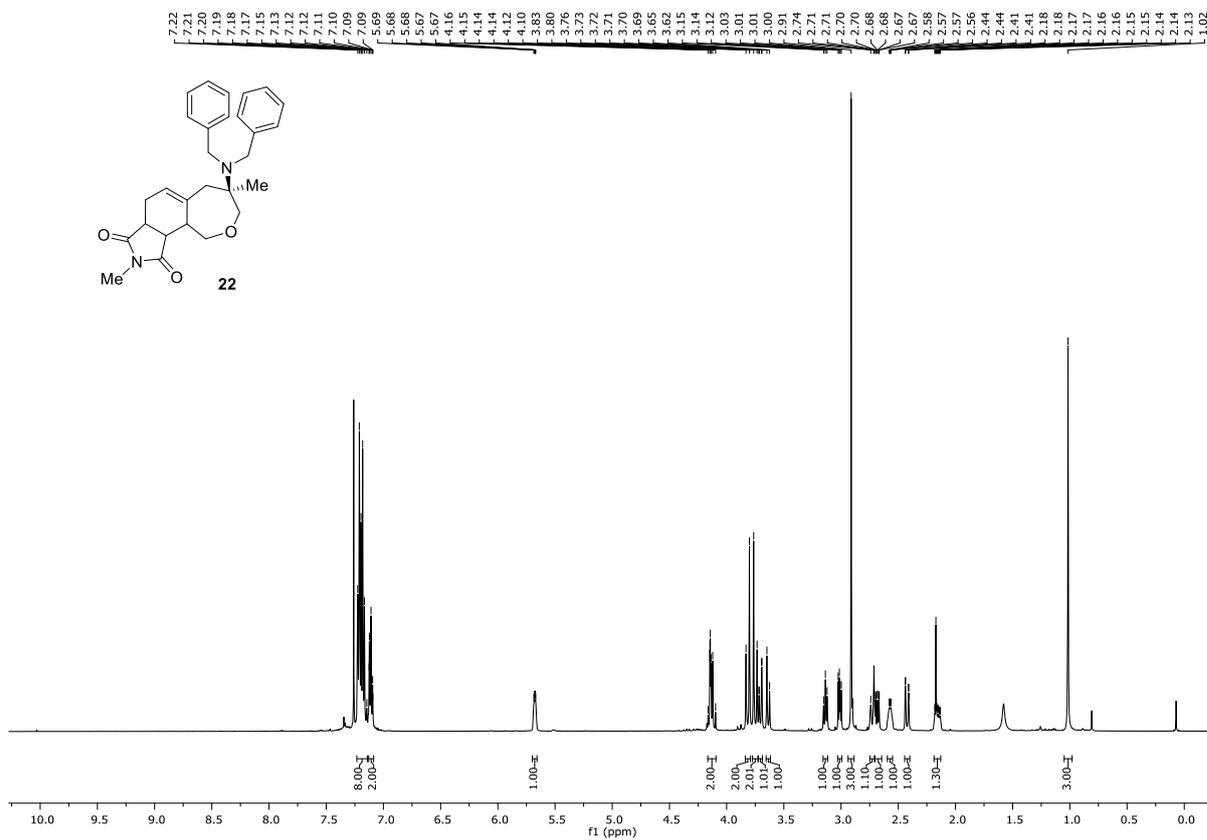


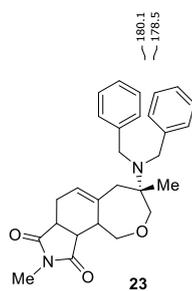
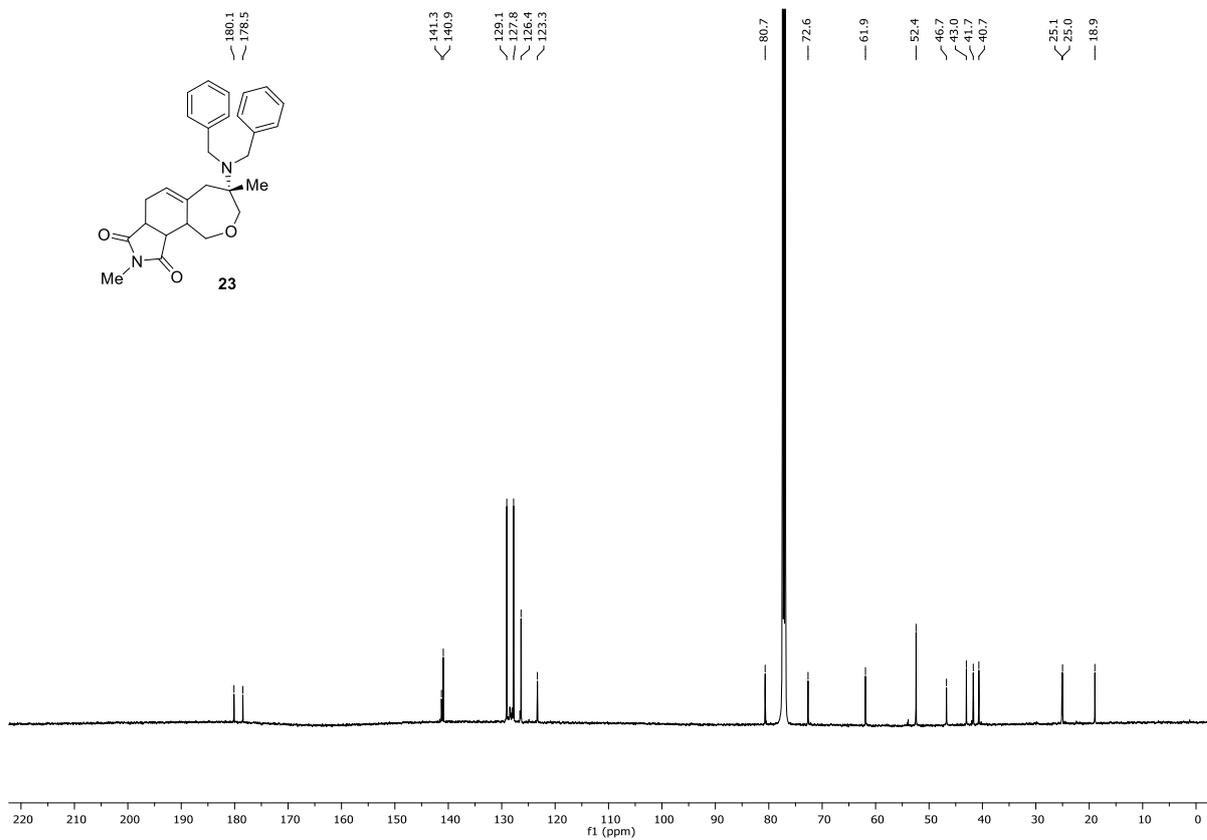
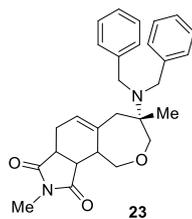
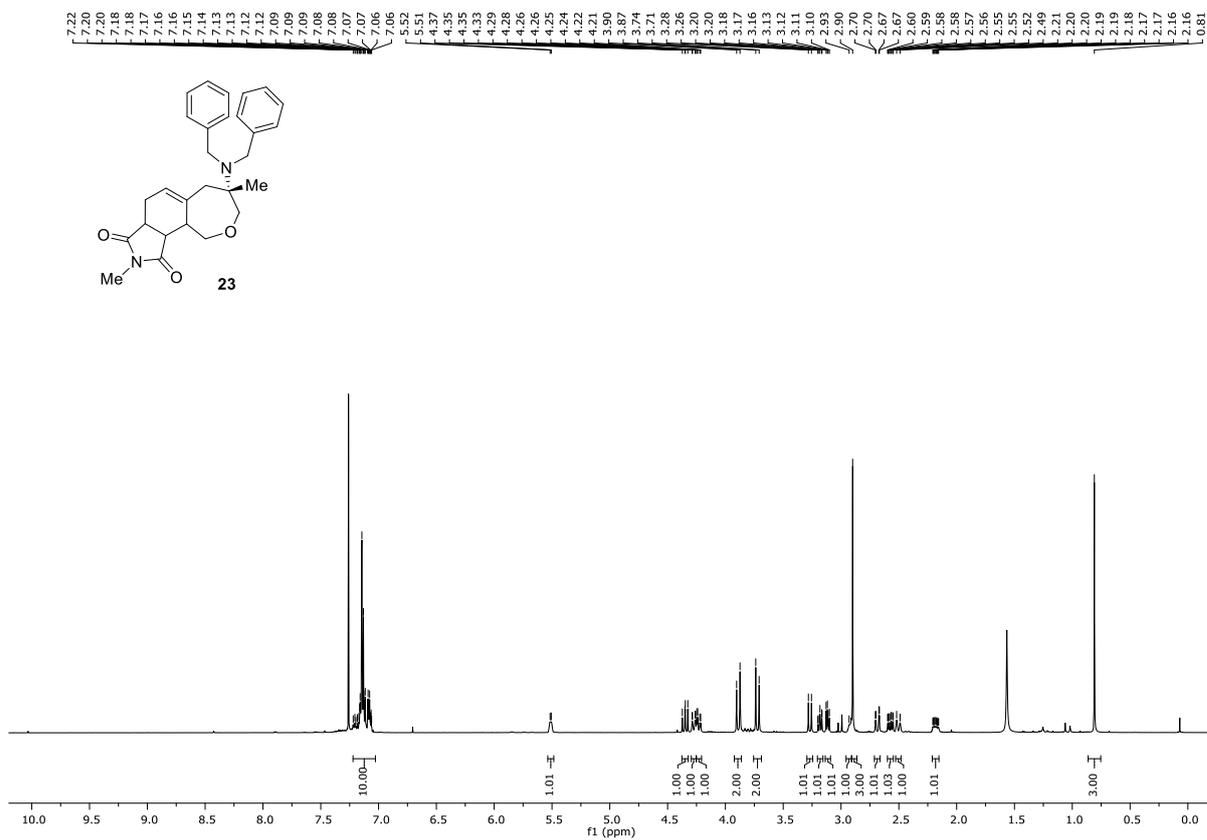
21

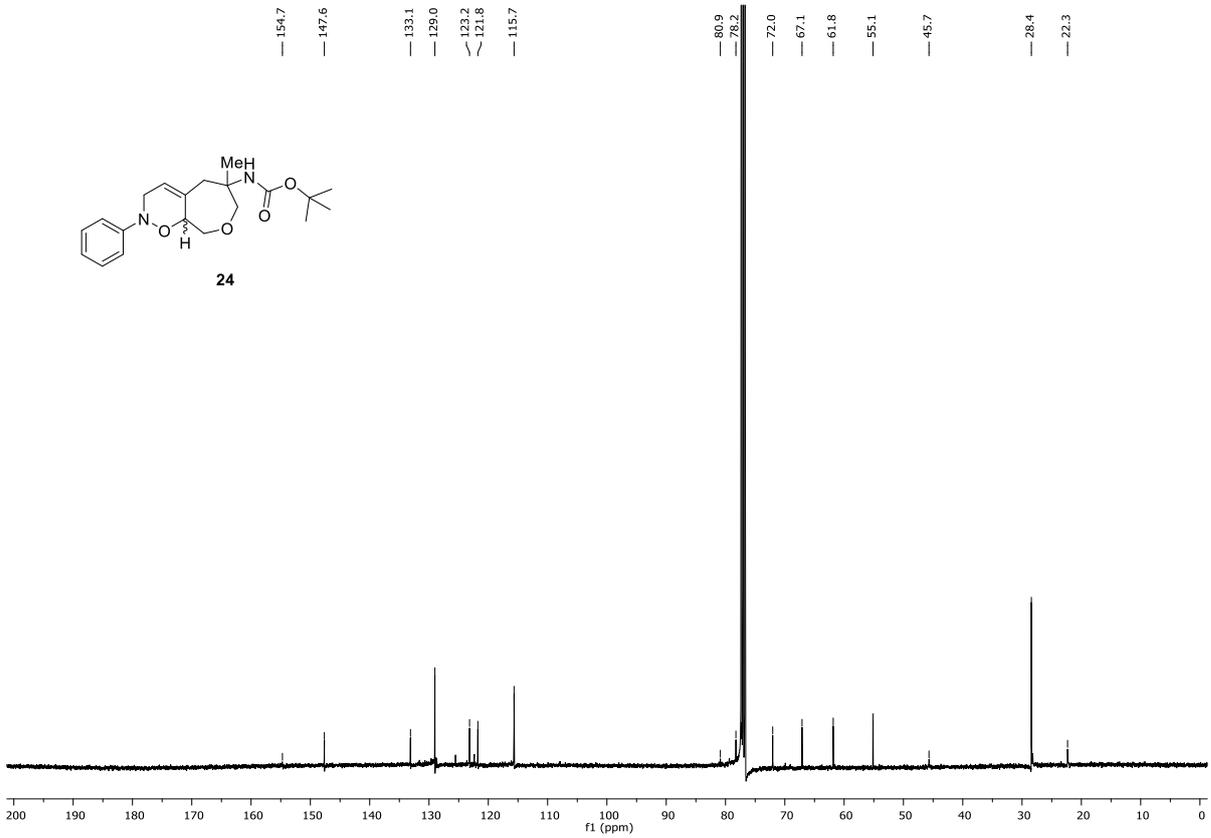
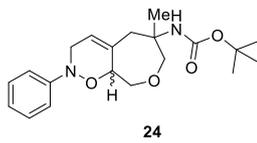
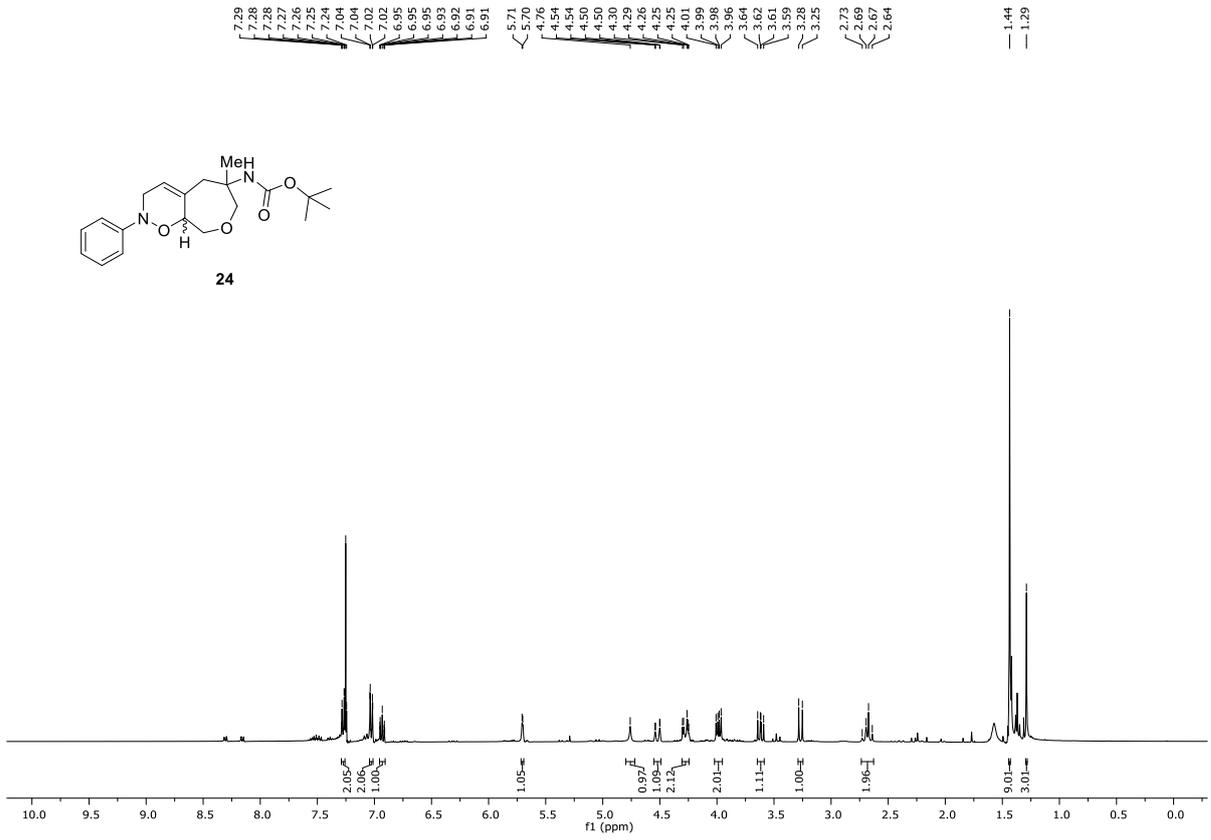
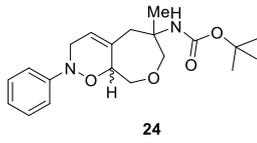


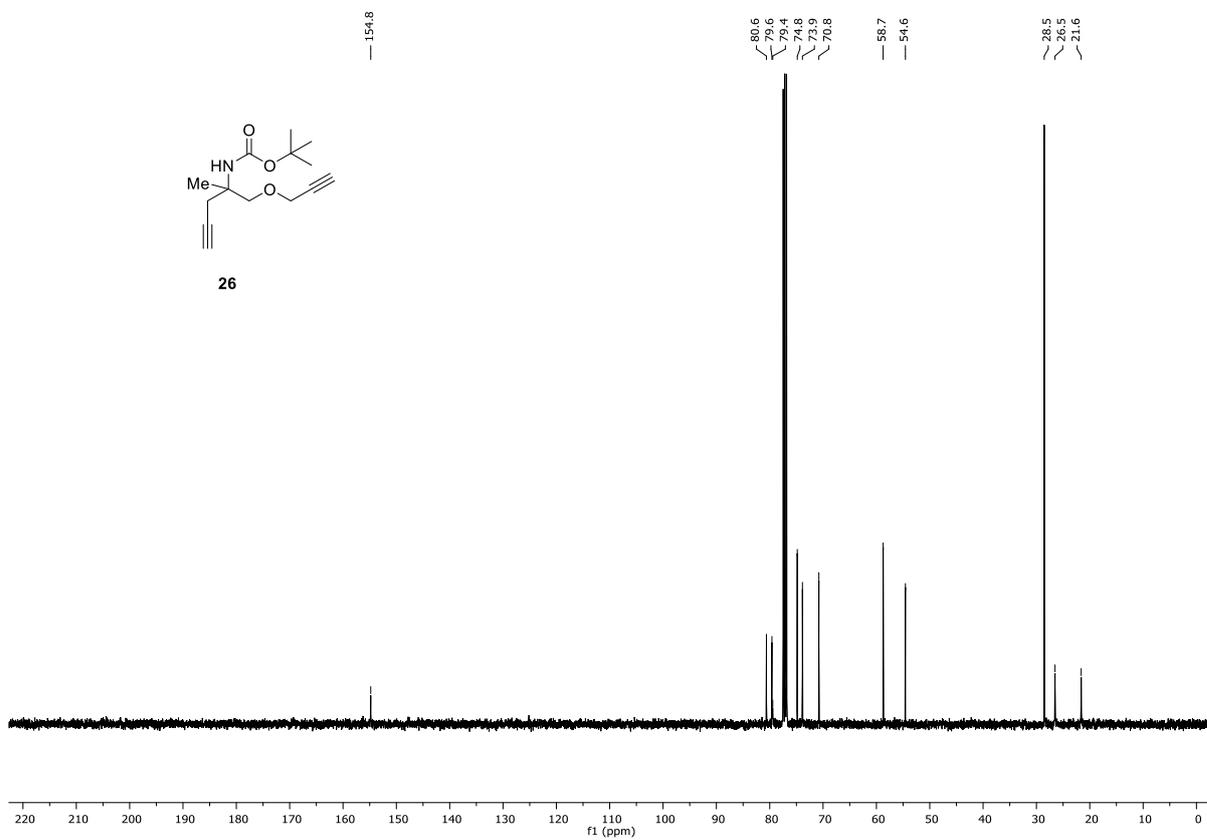
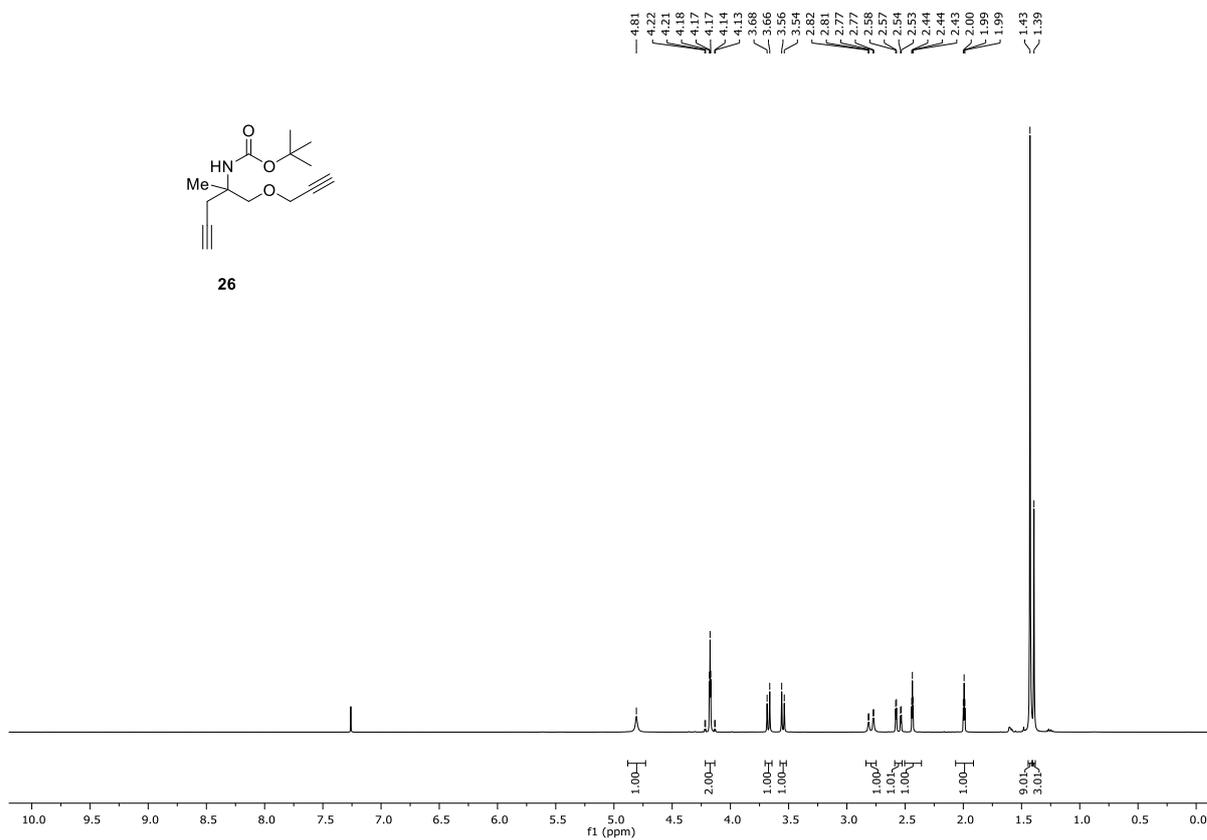
21

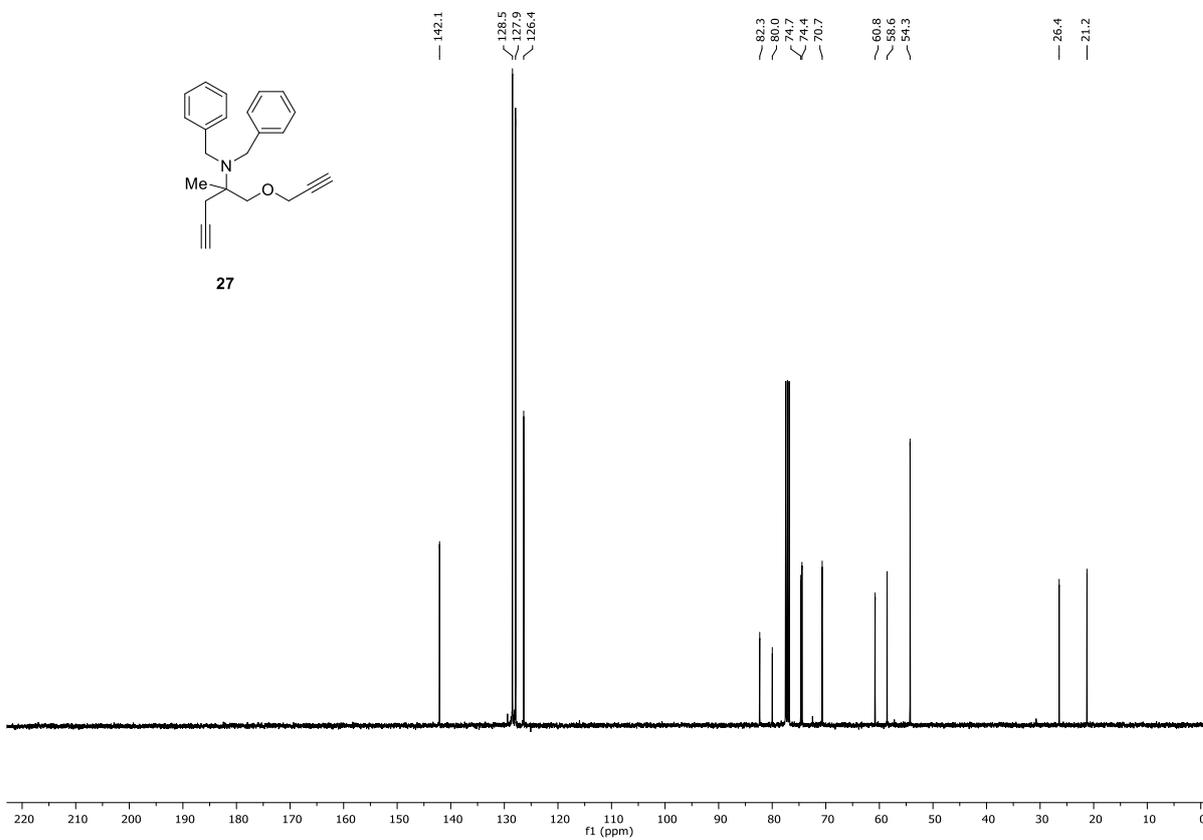
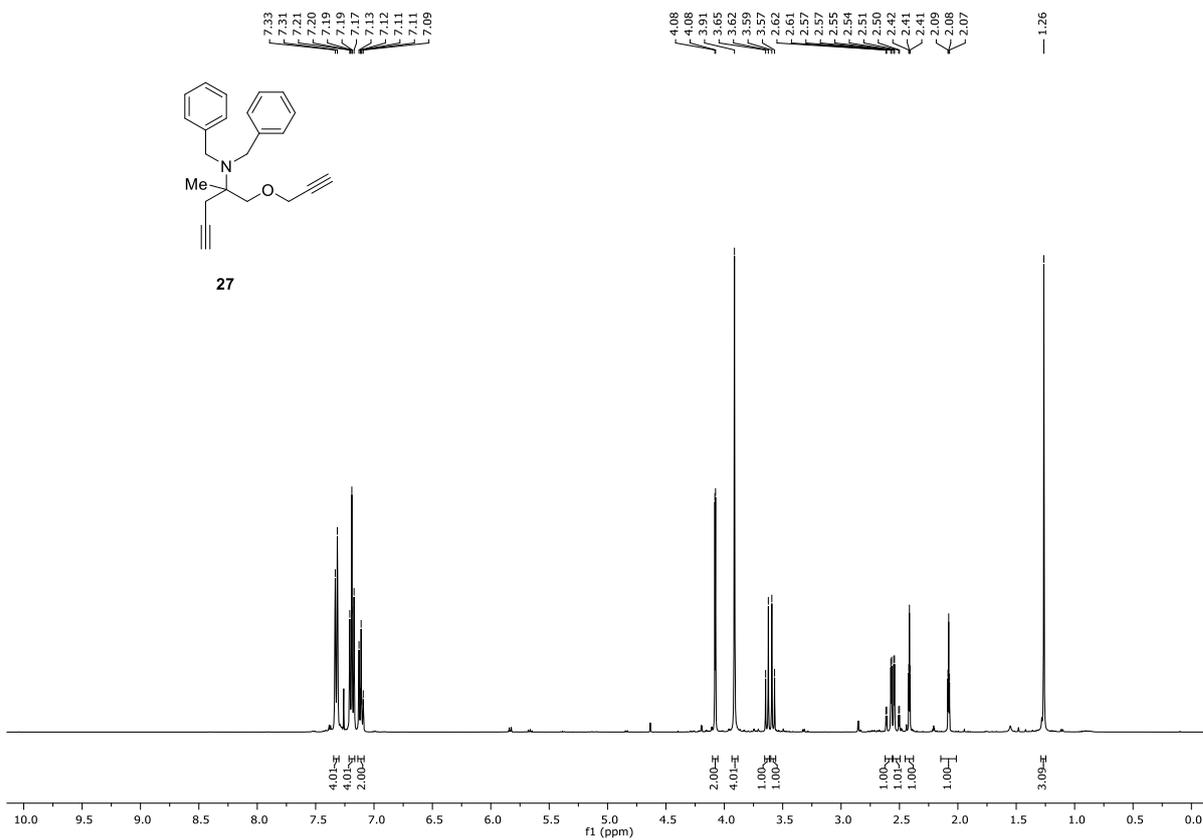


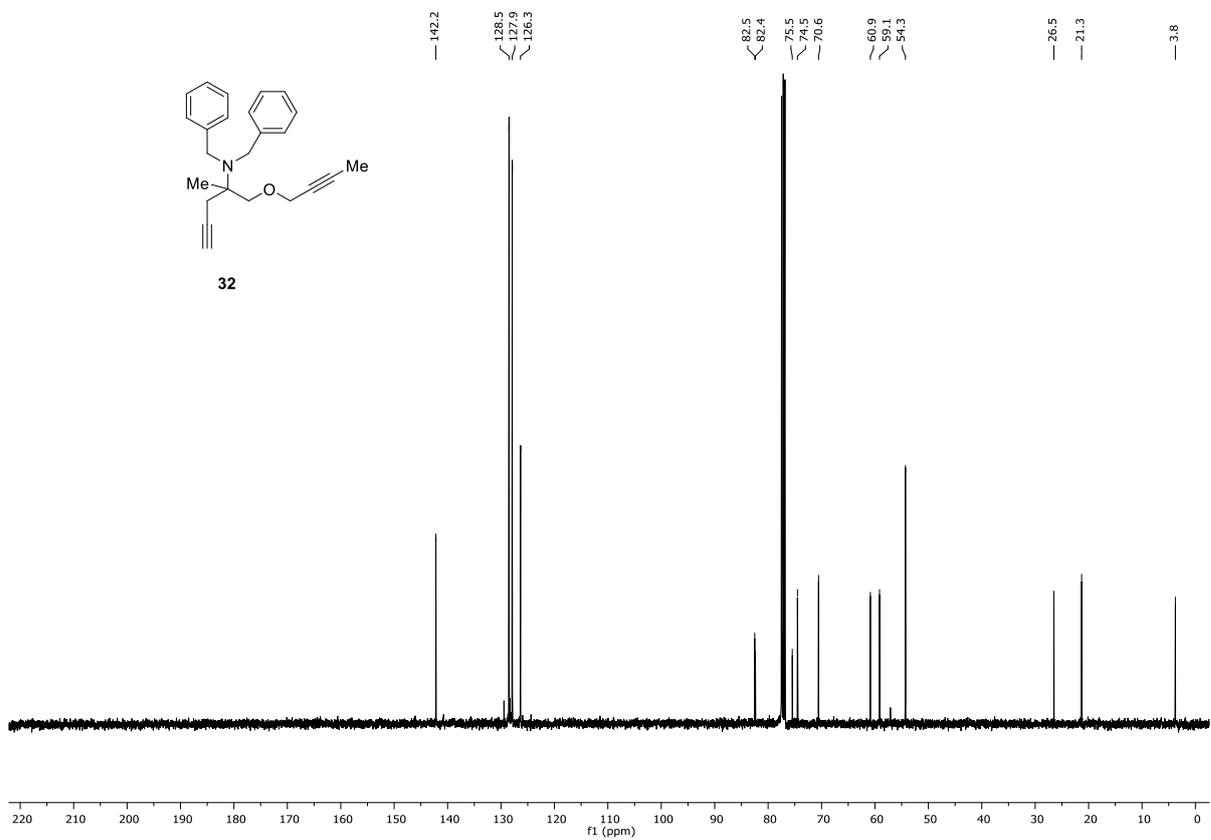
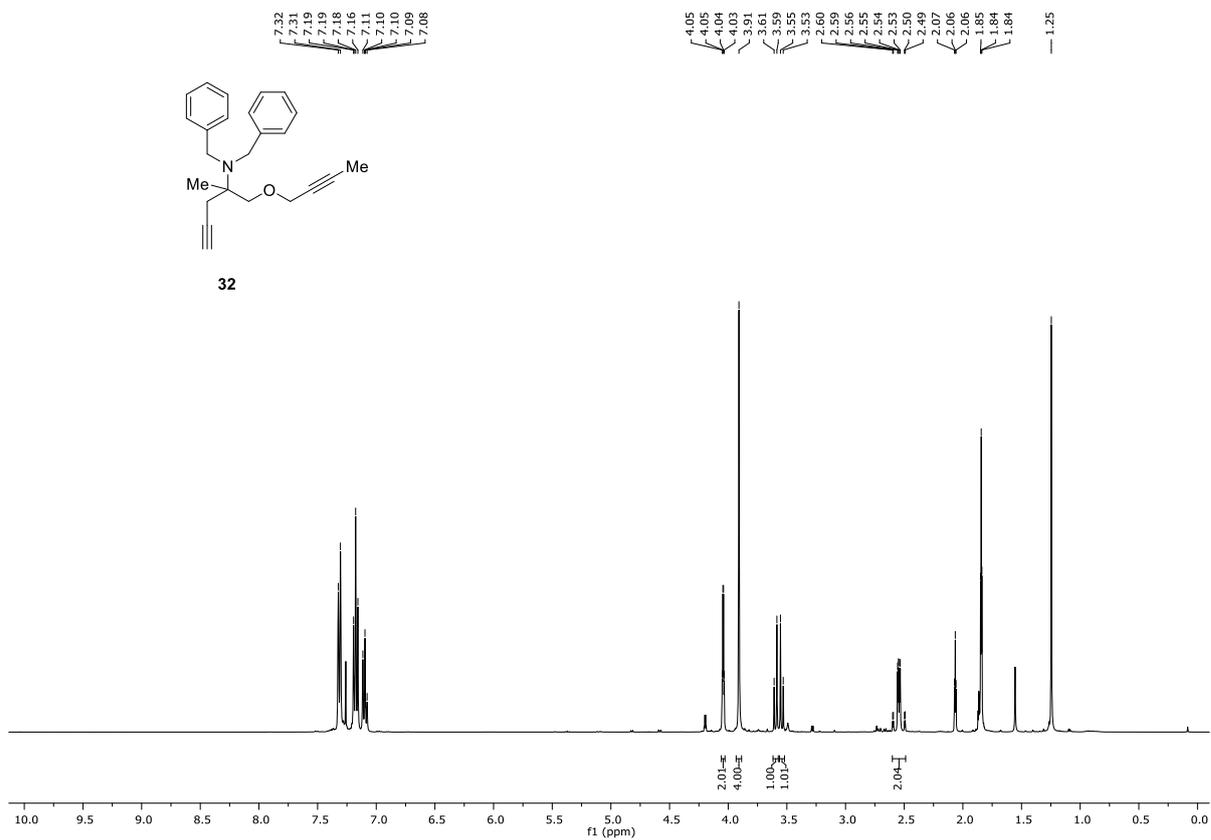


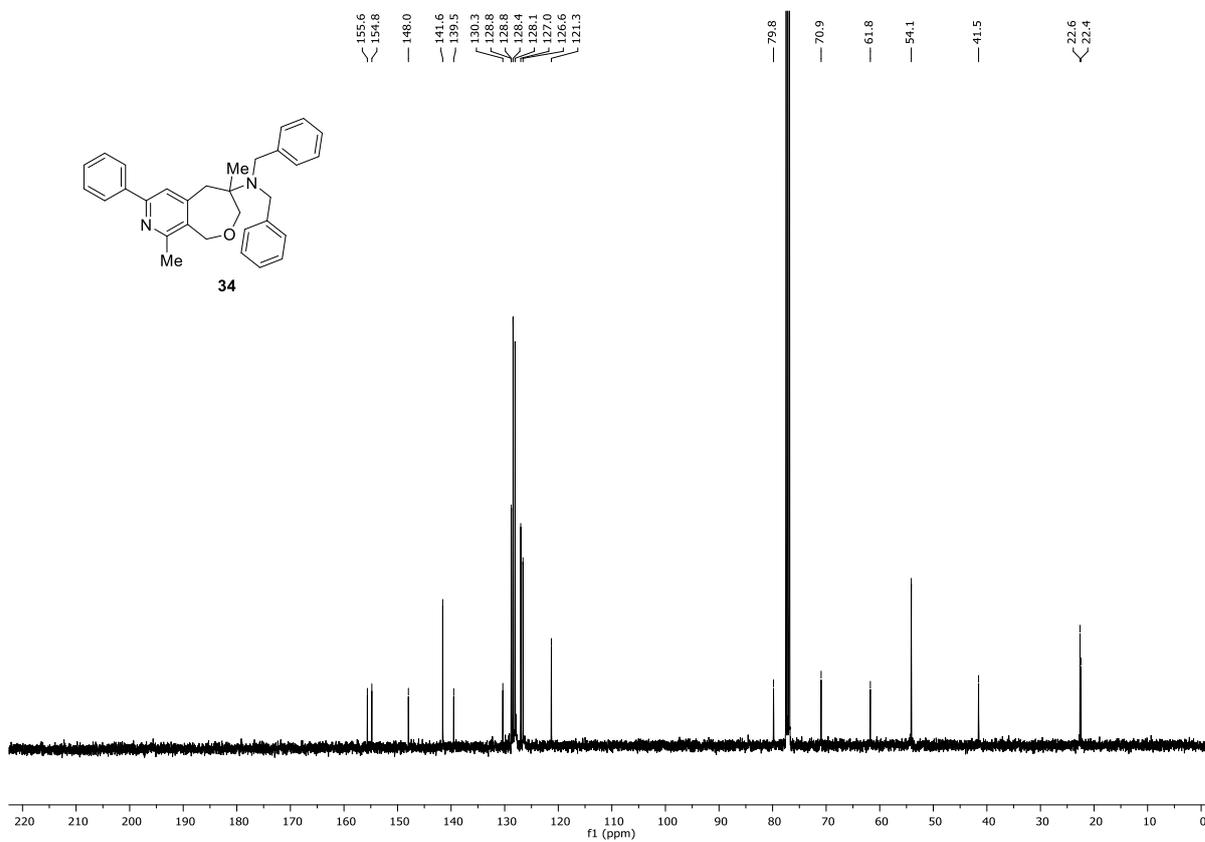
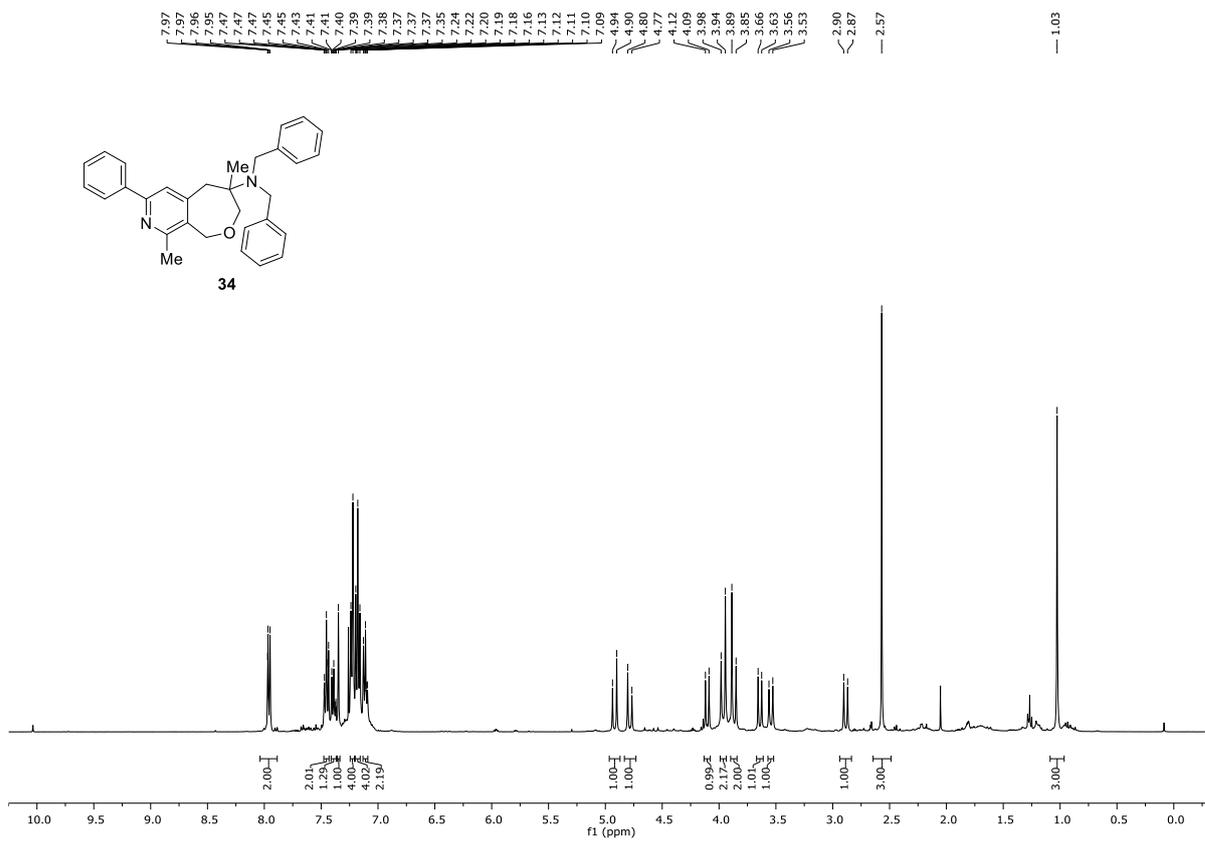


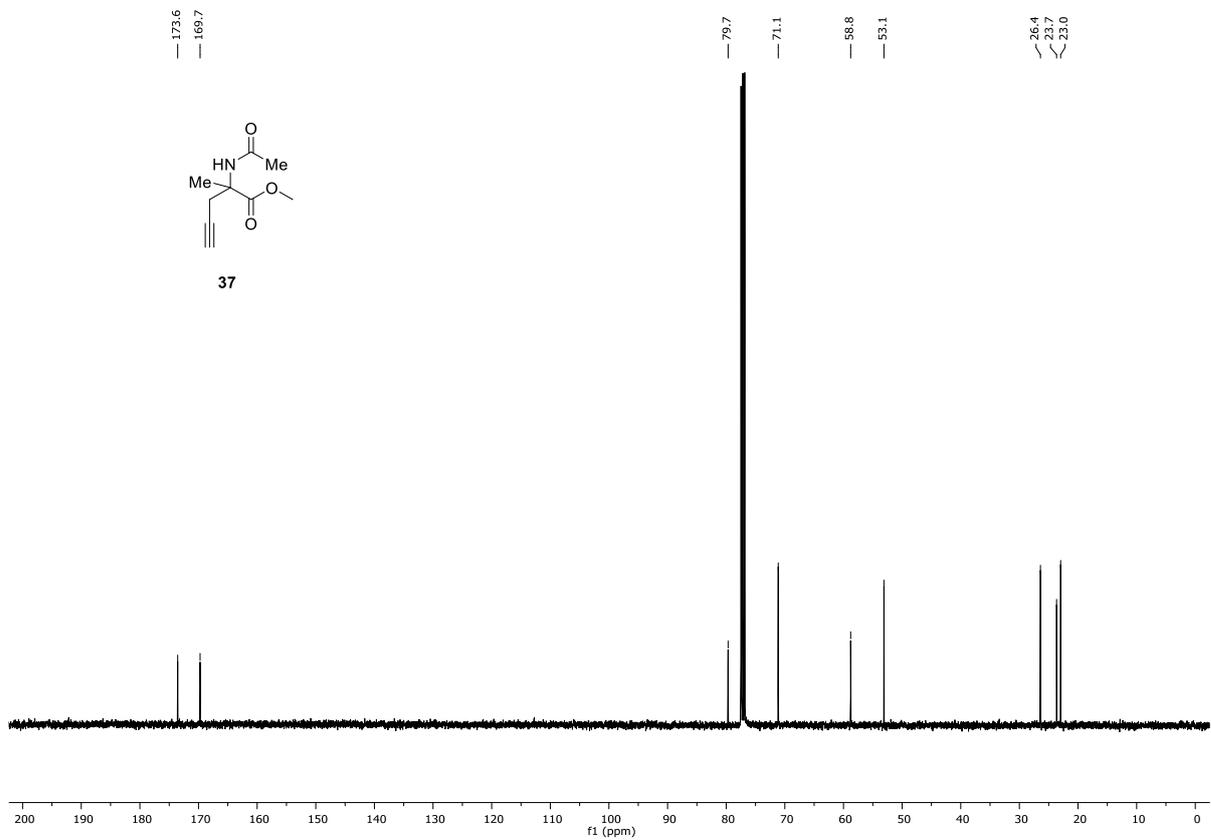
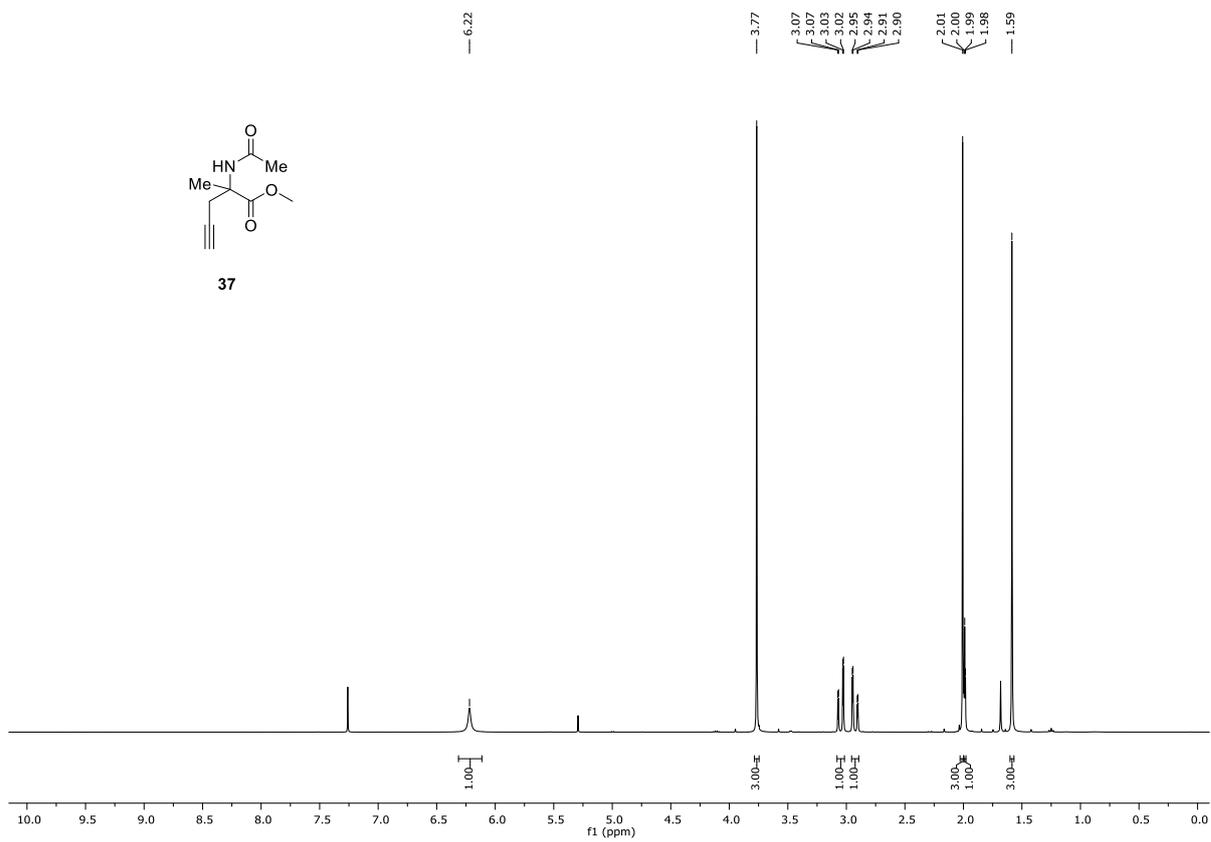


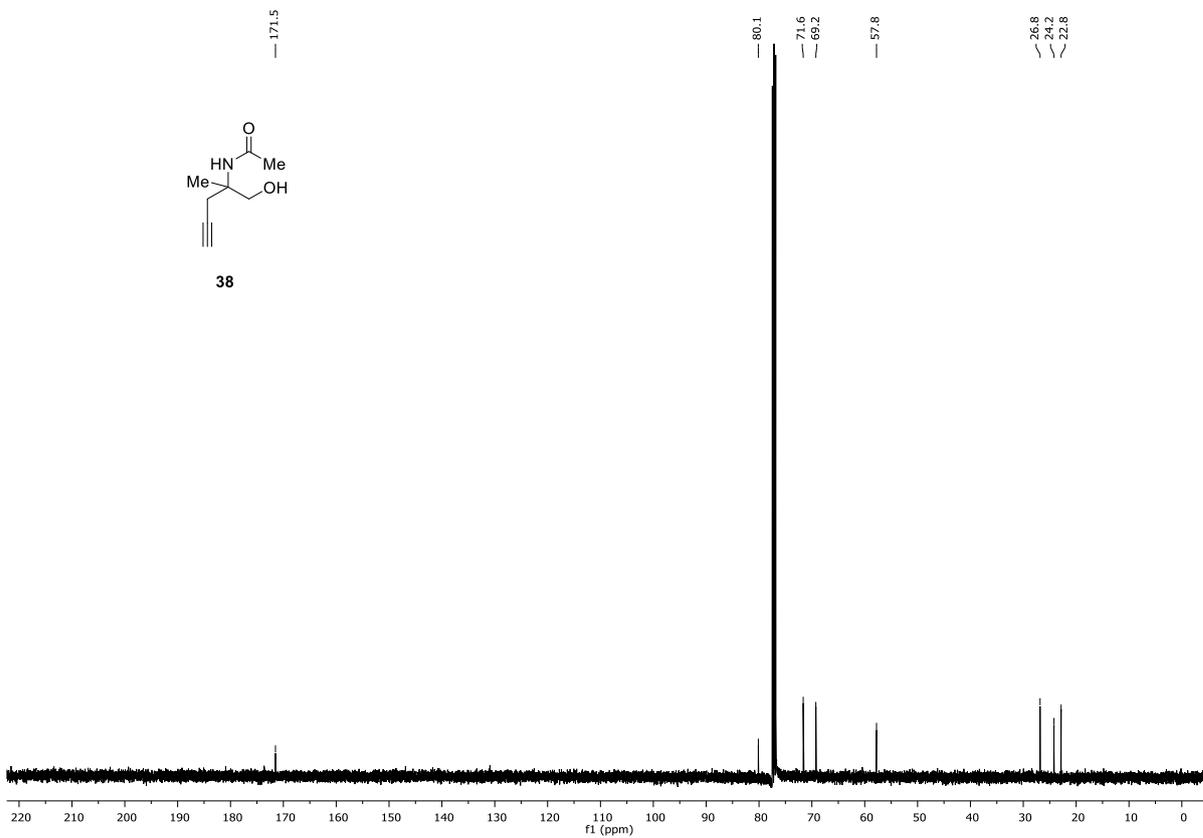
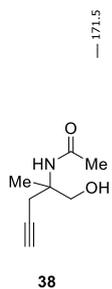
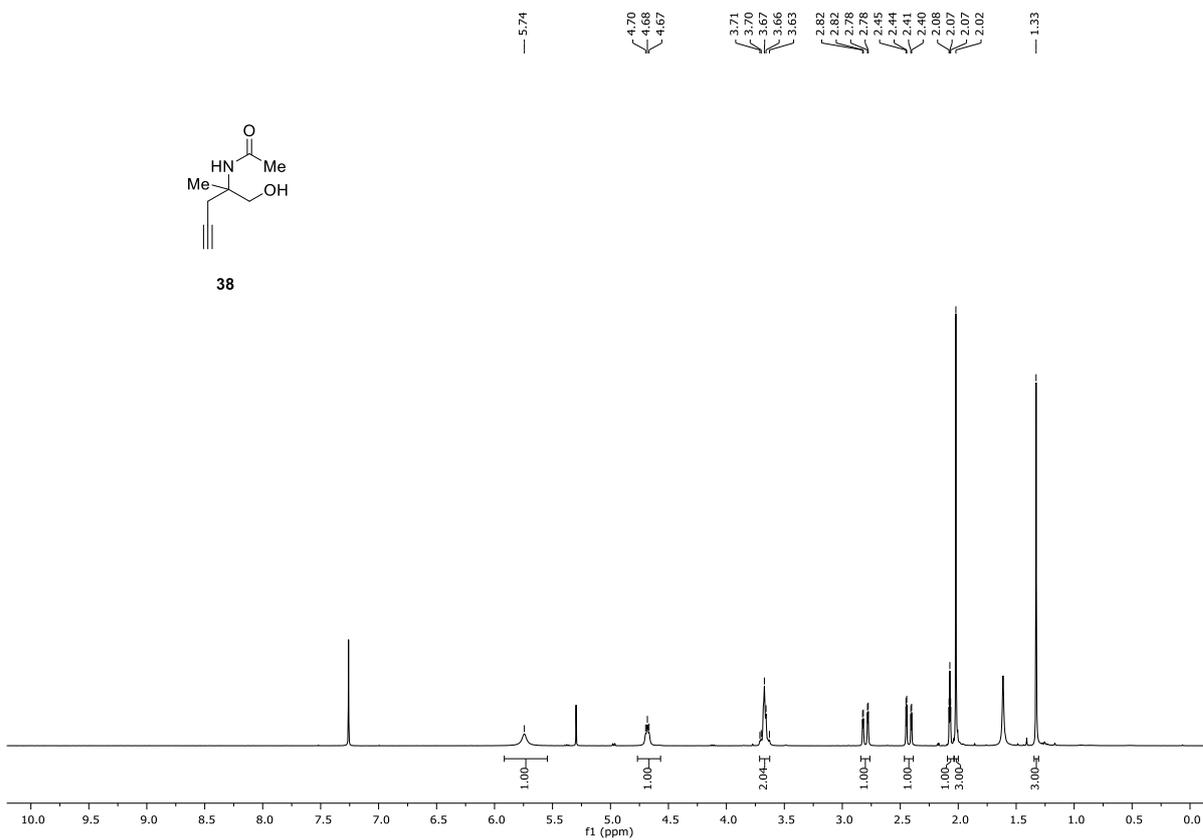
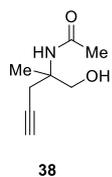


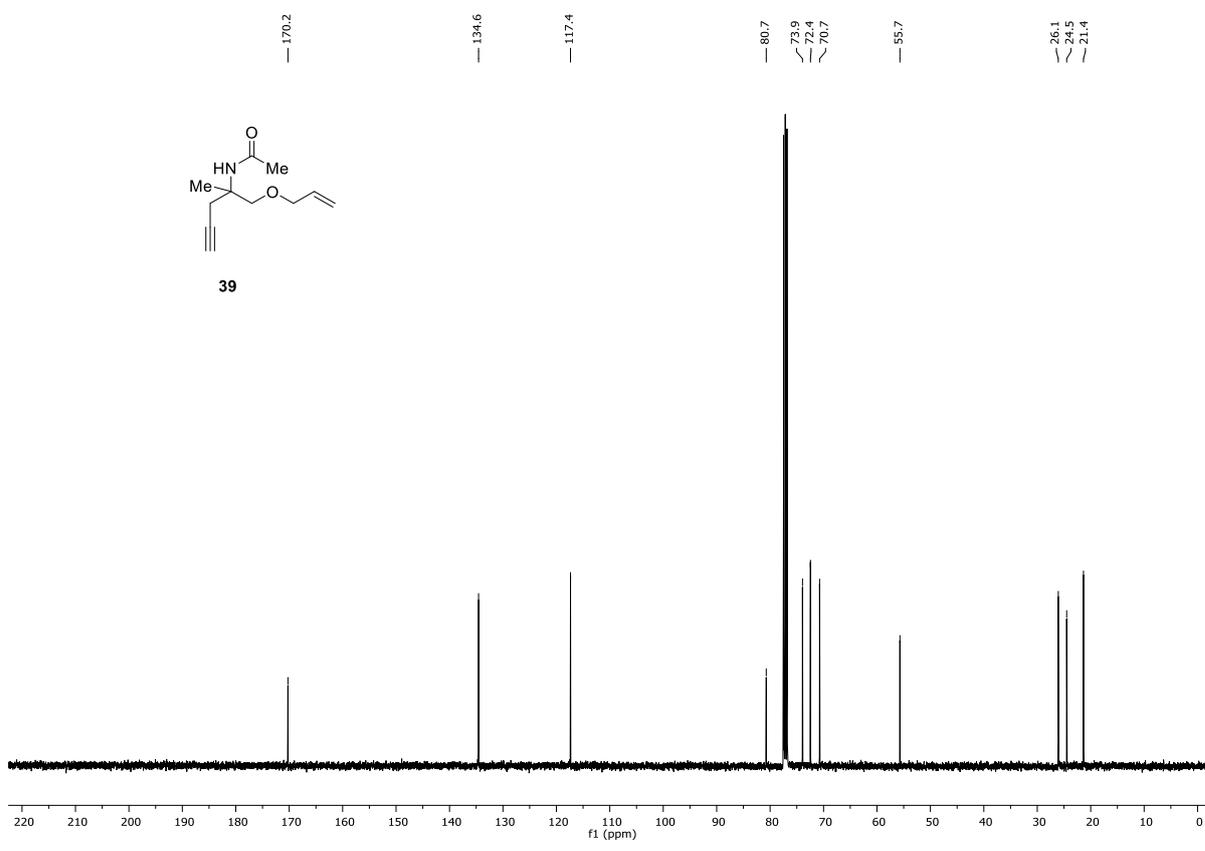
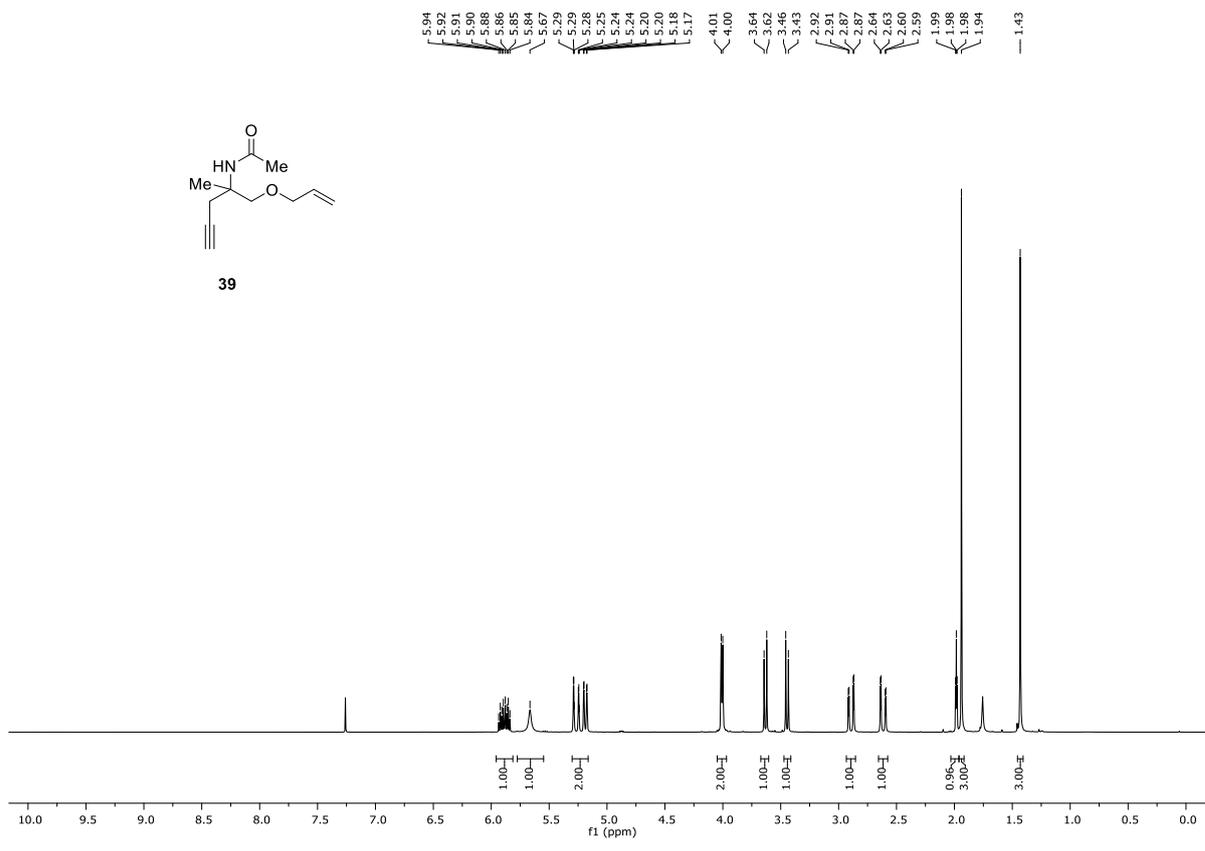




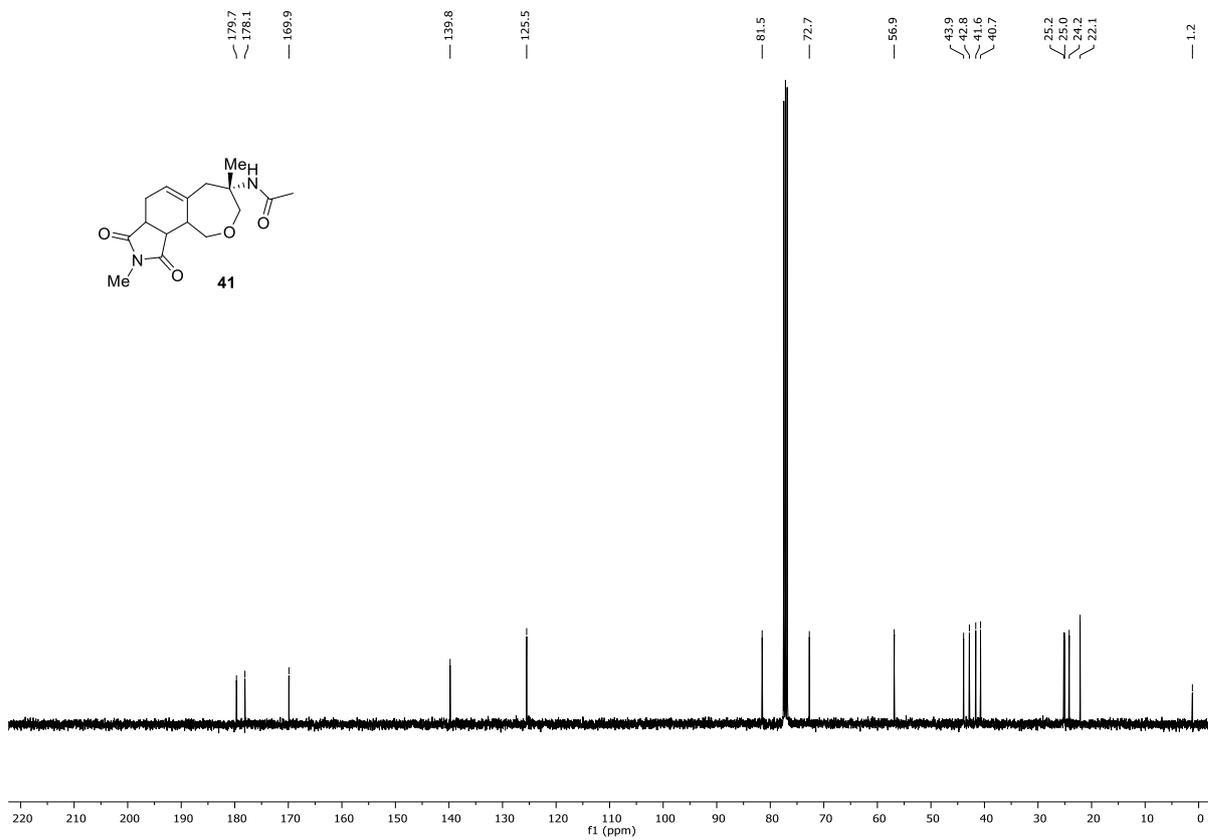
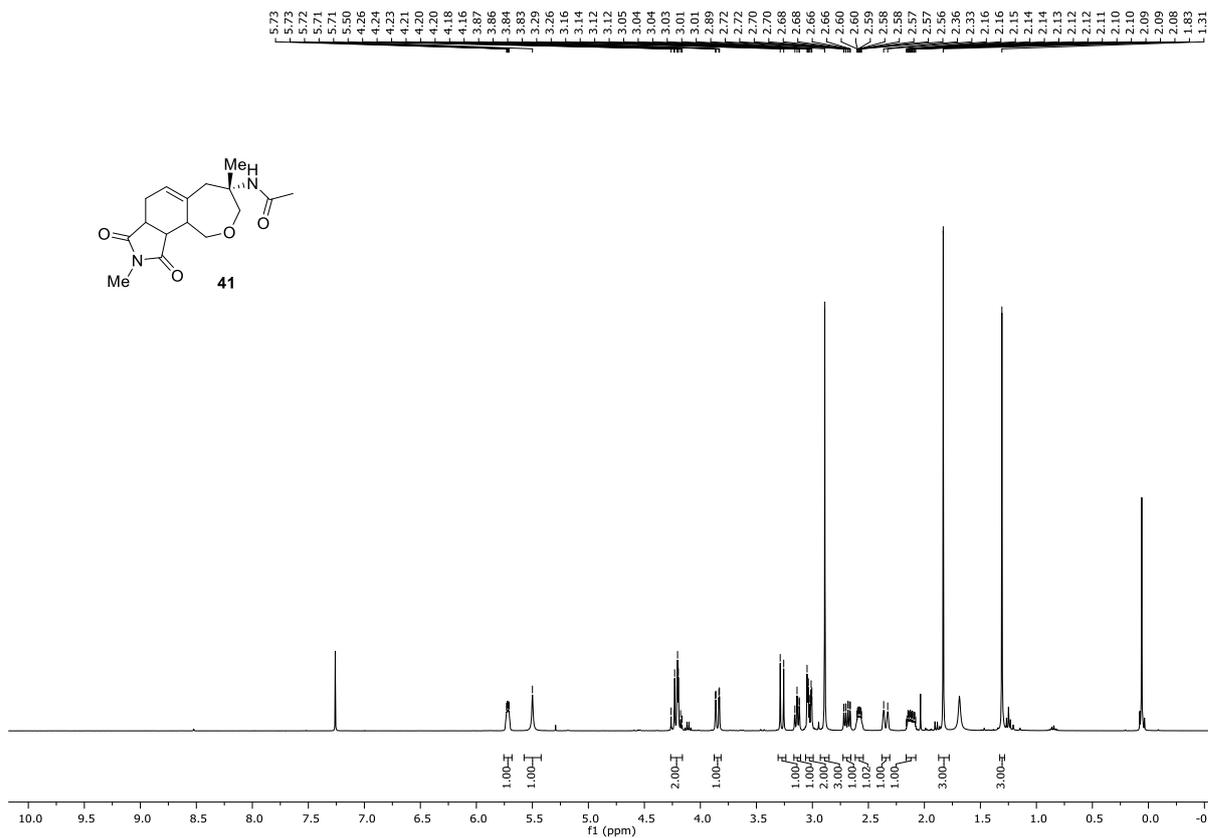


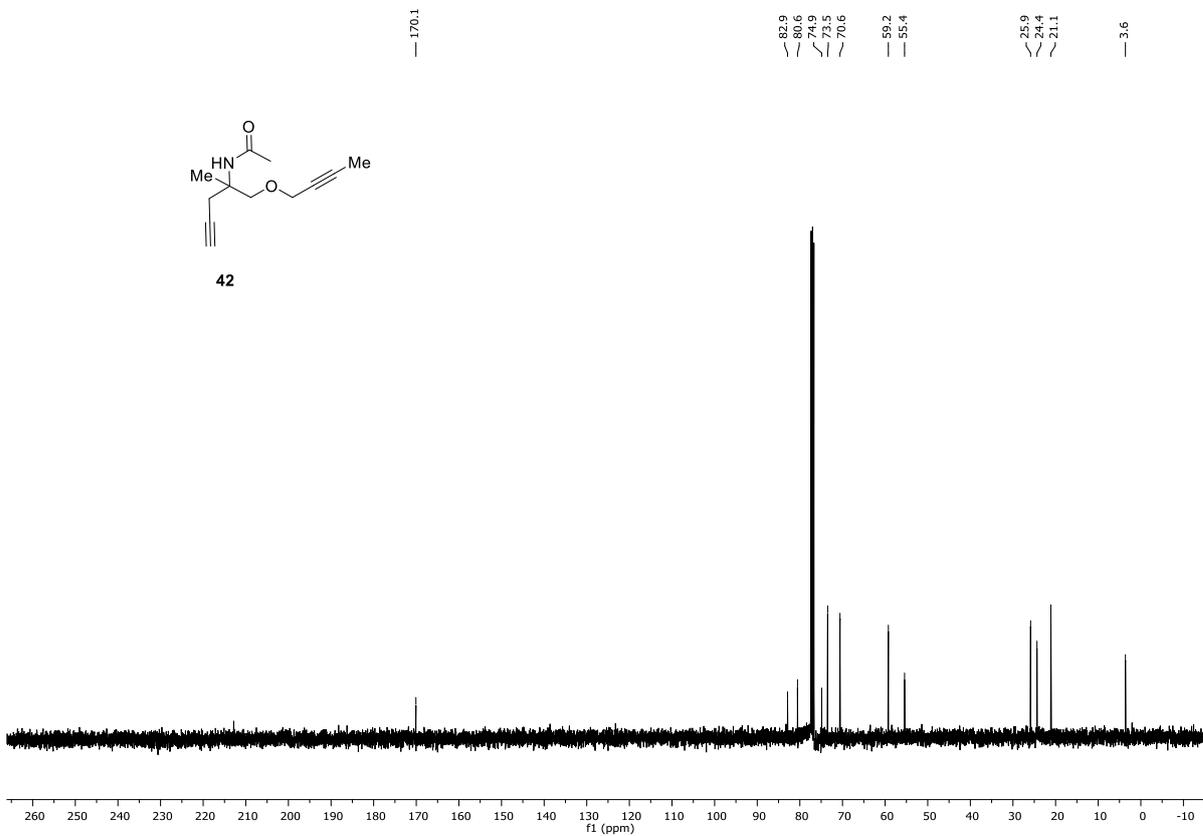
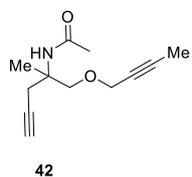
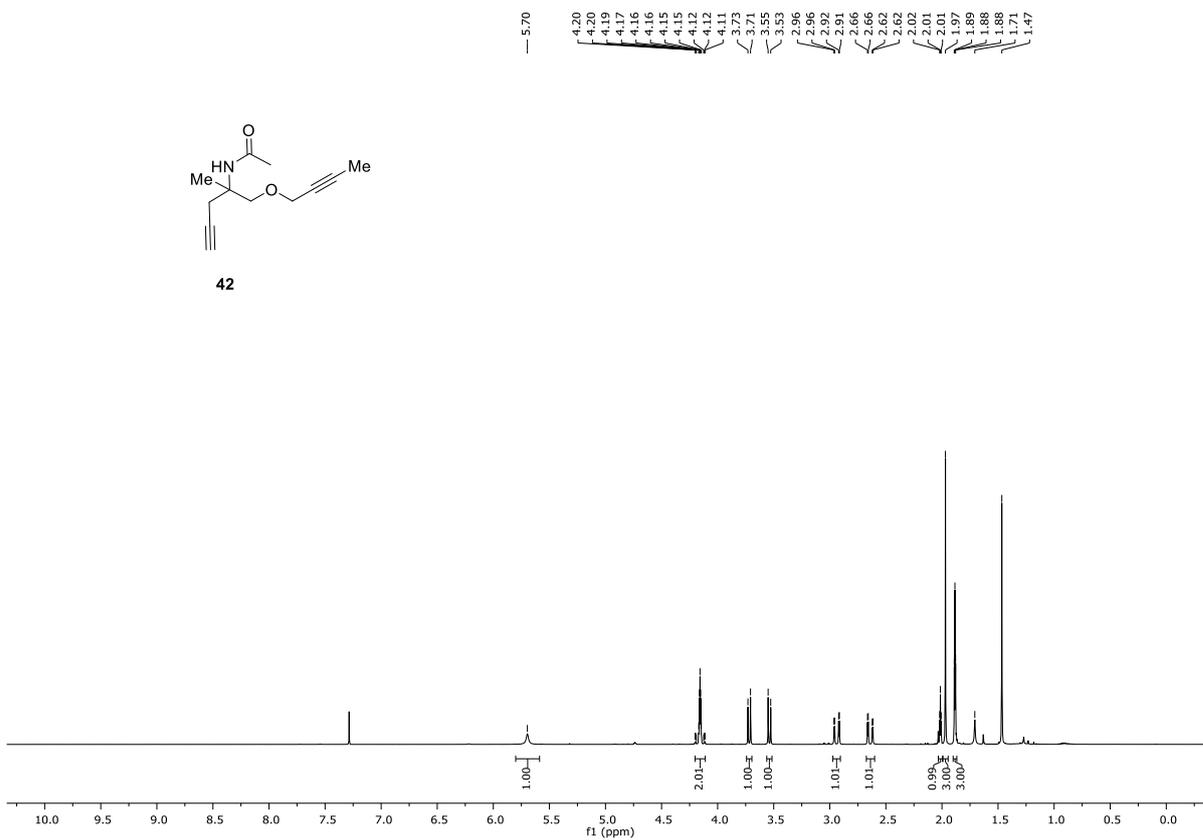
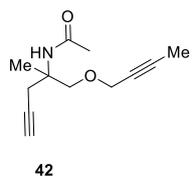


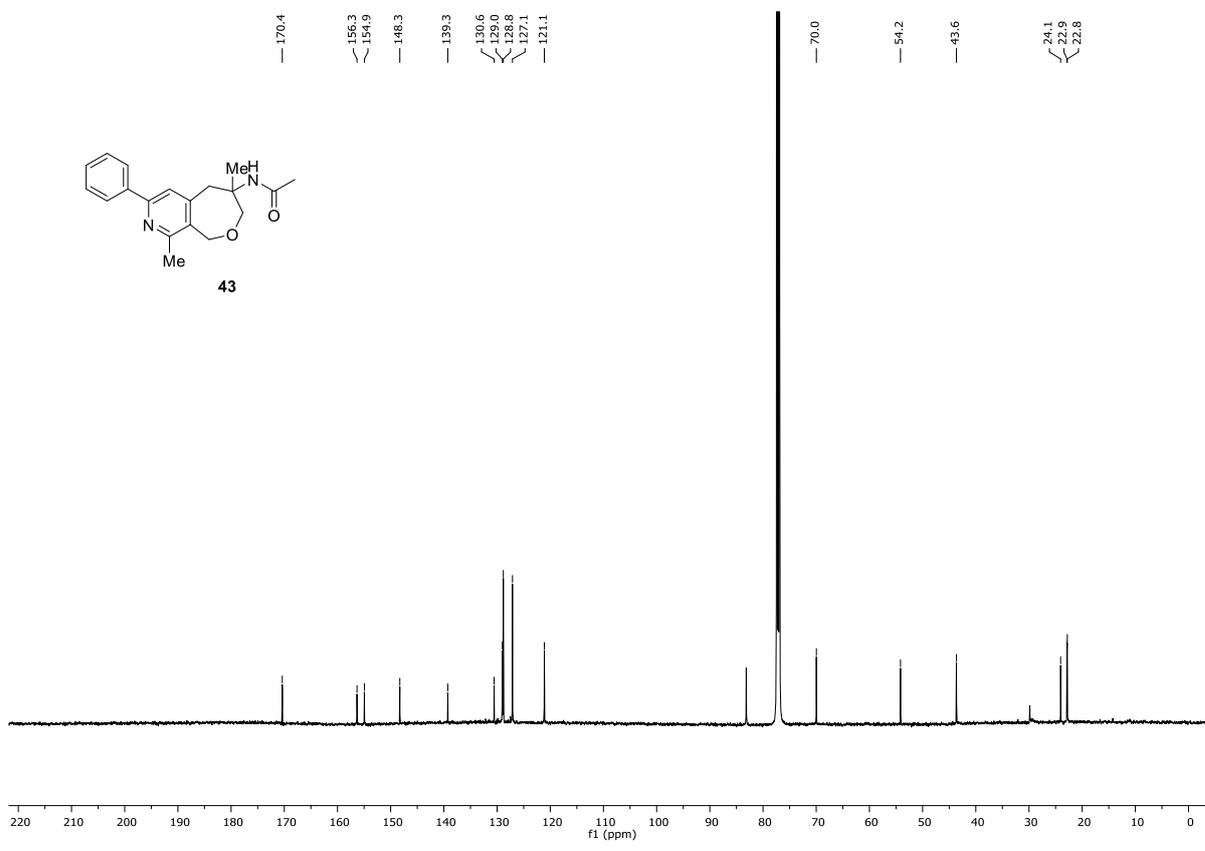
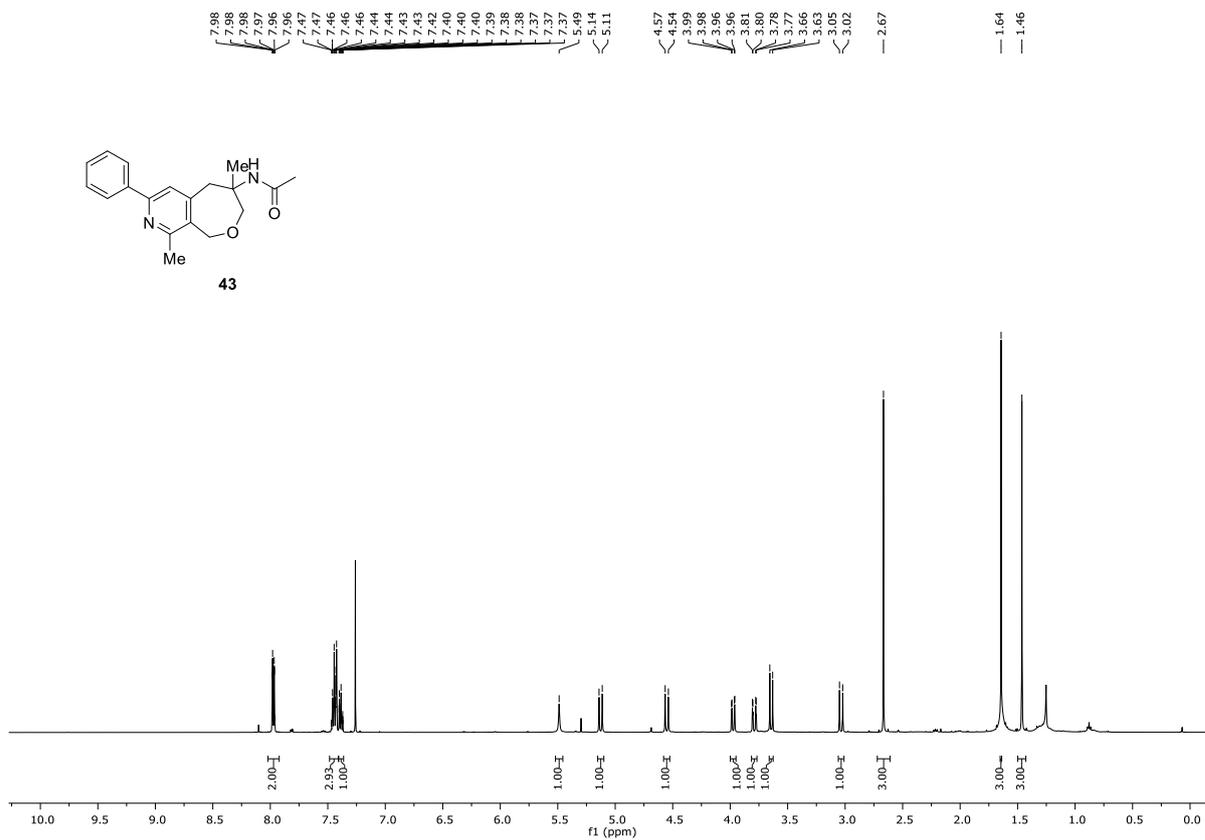


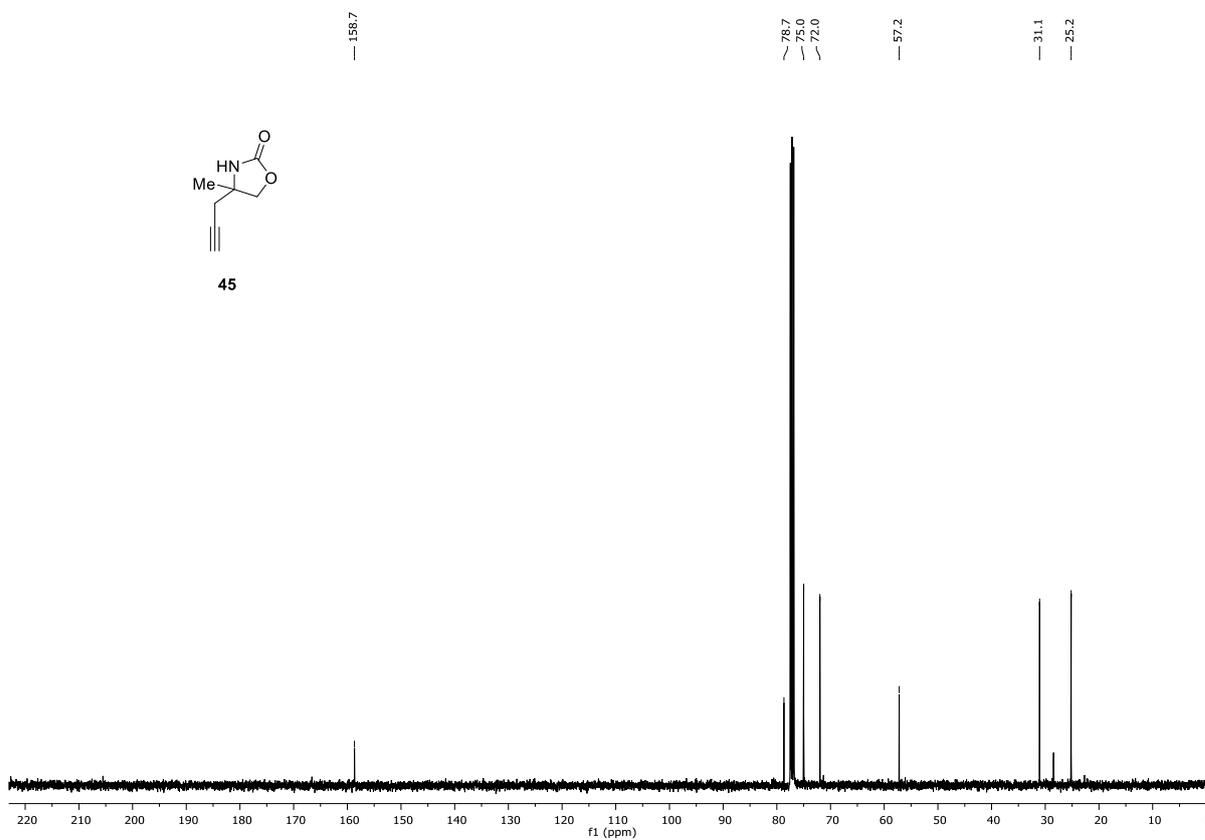
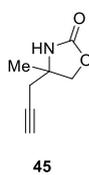
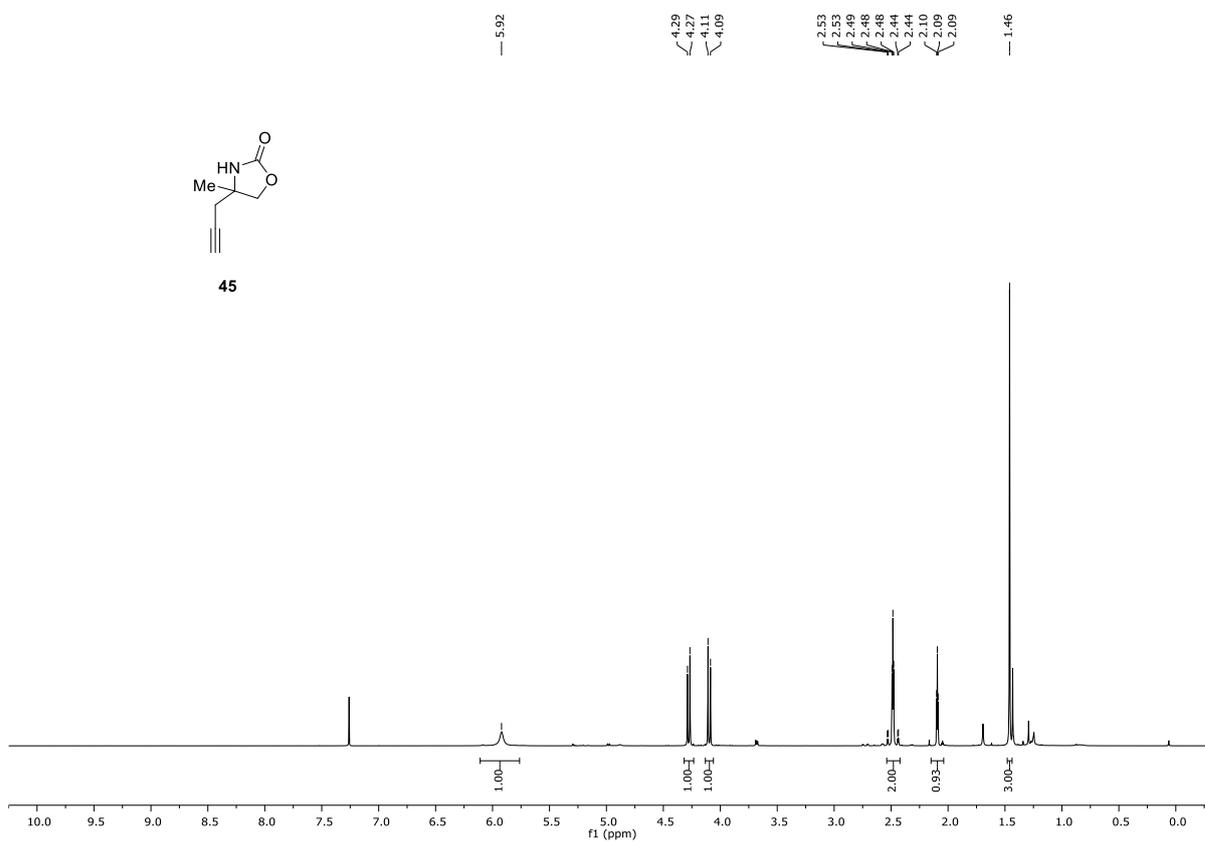
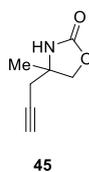


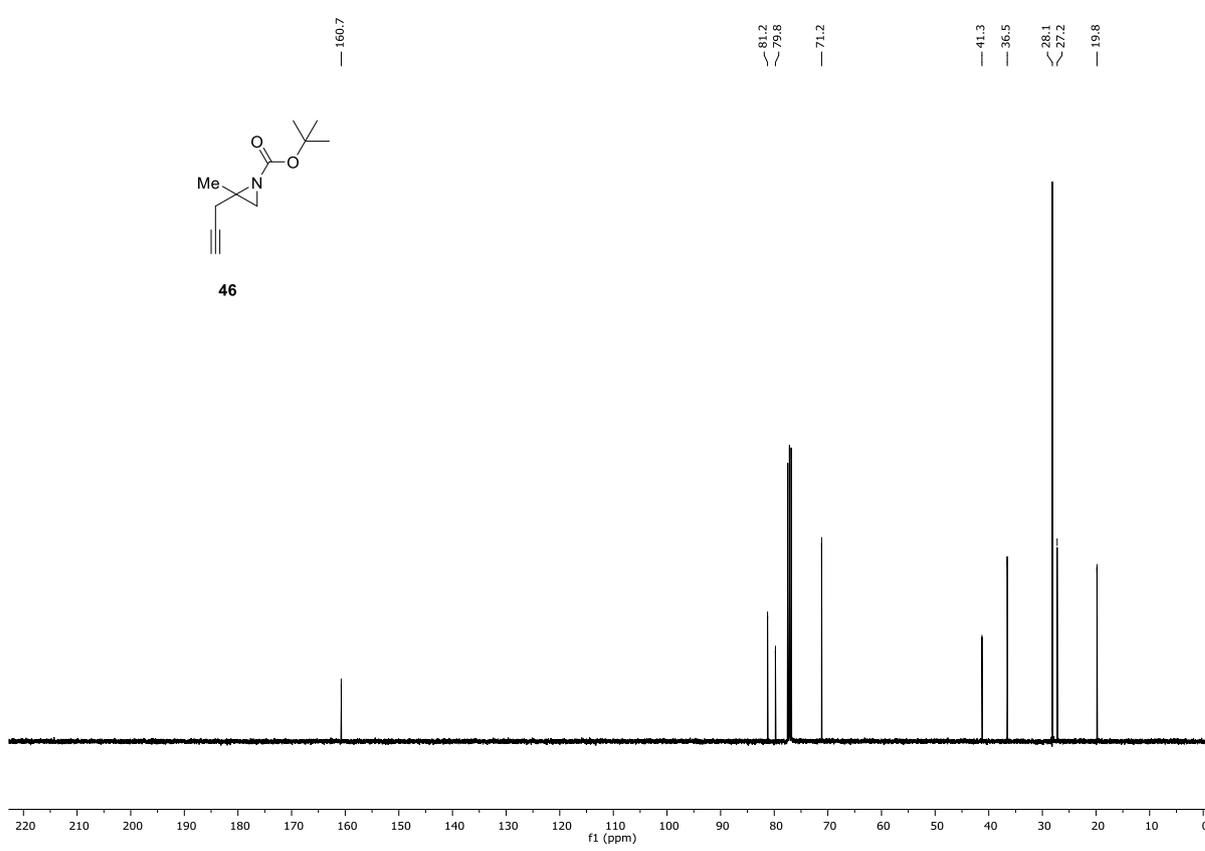
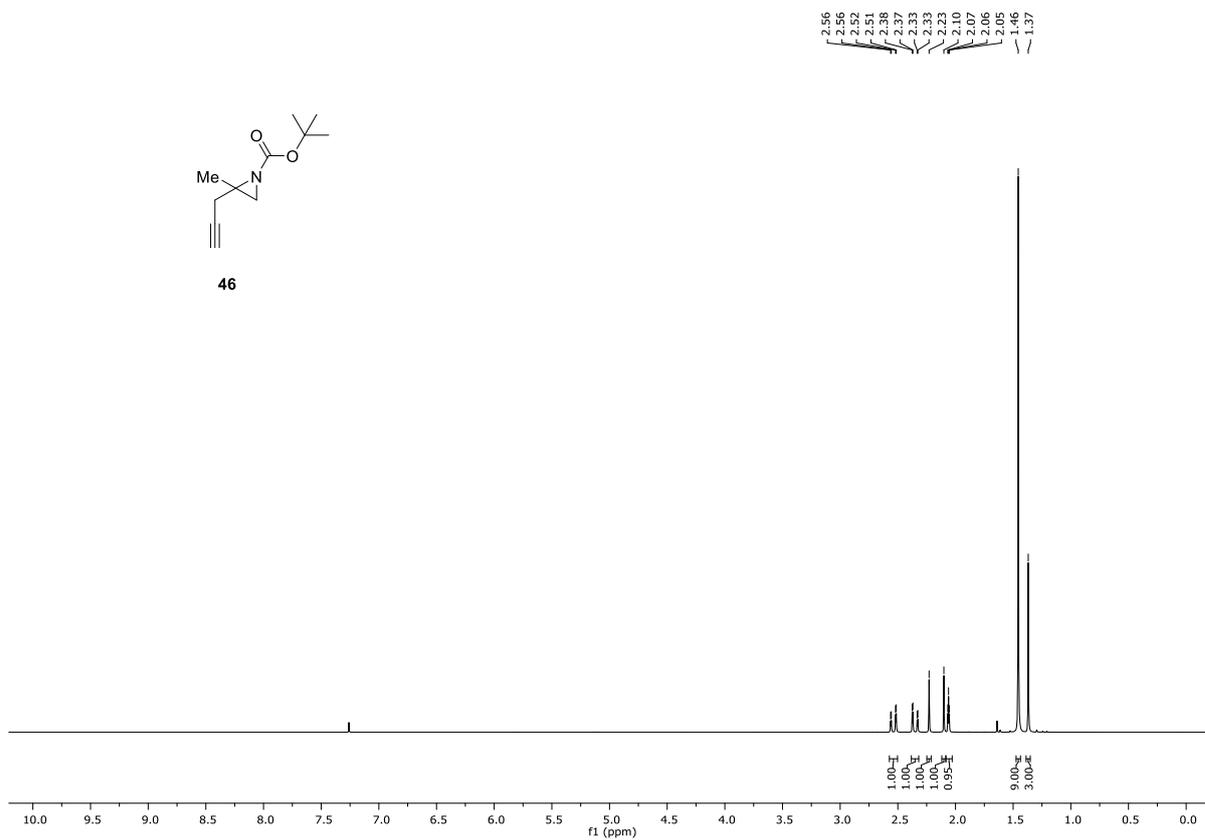


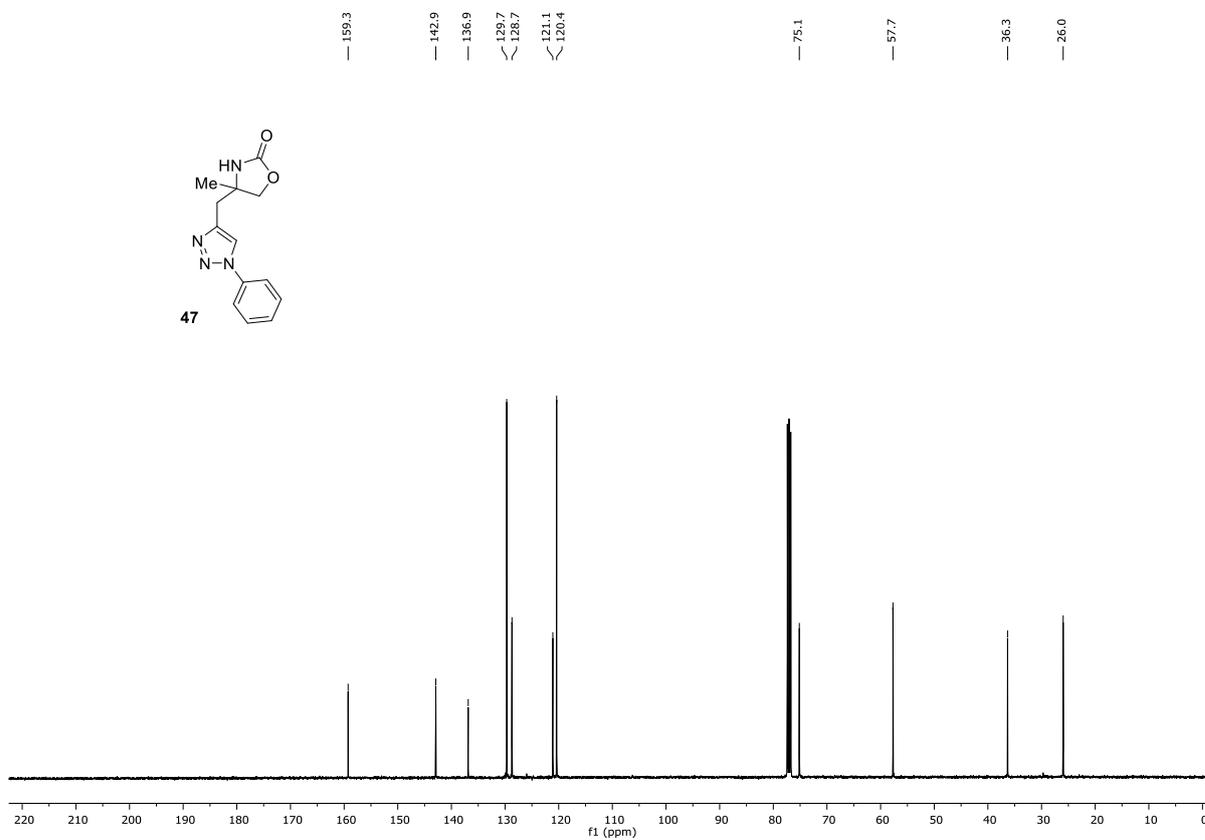
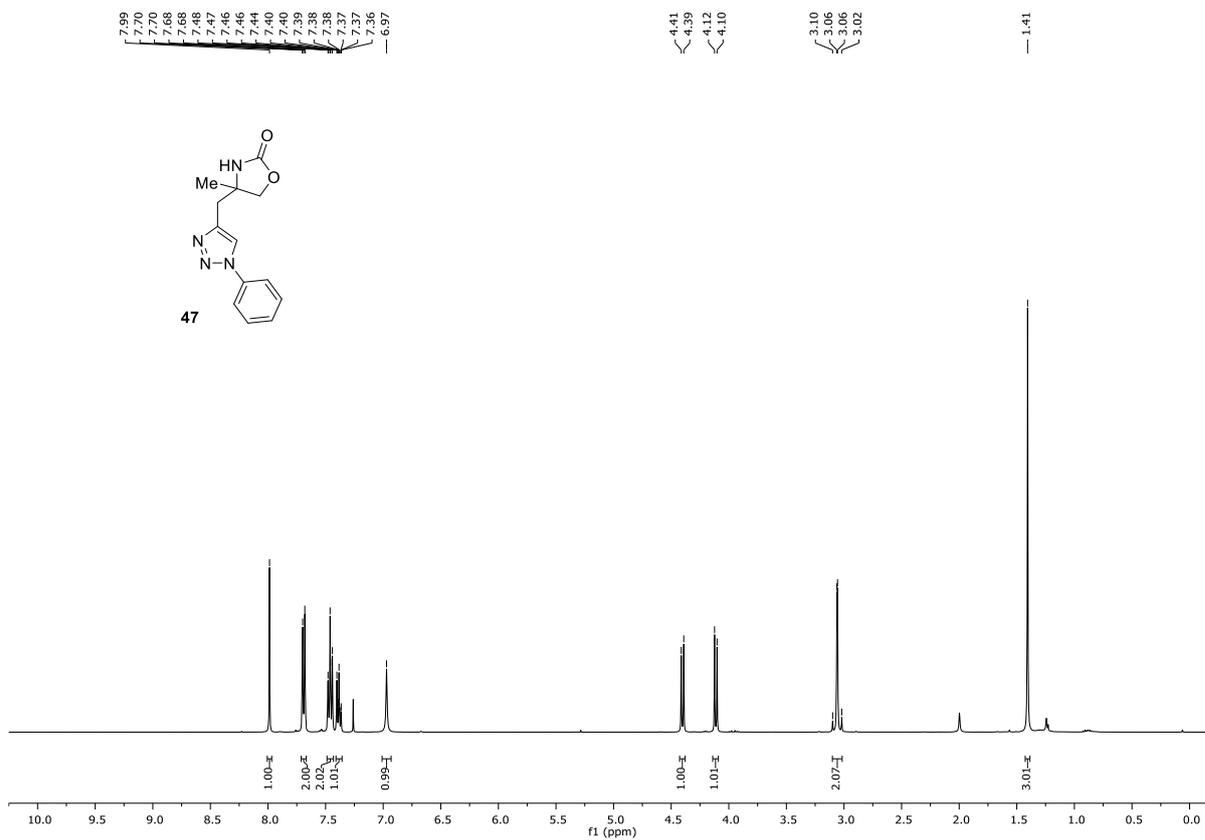


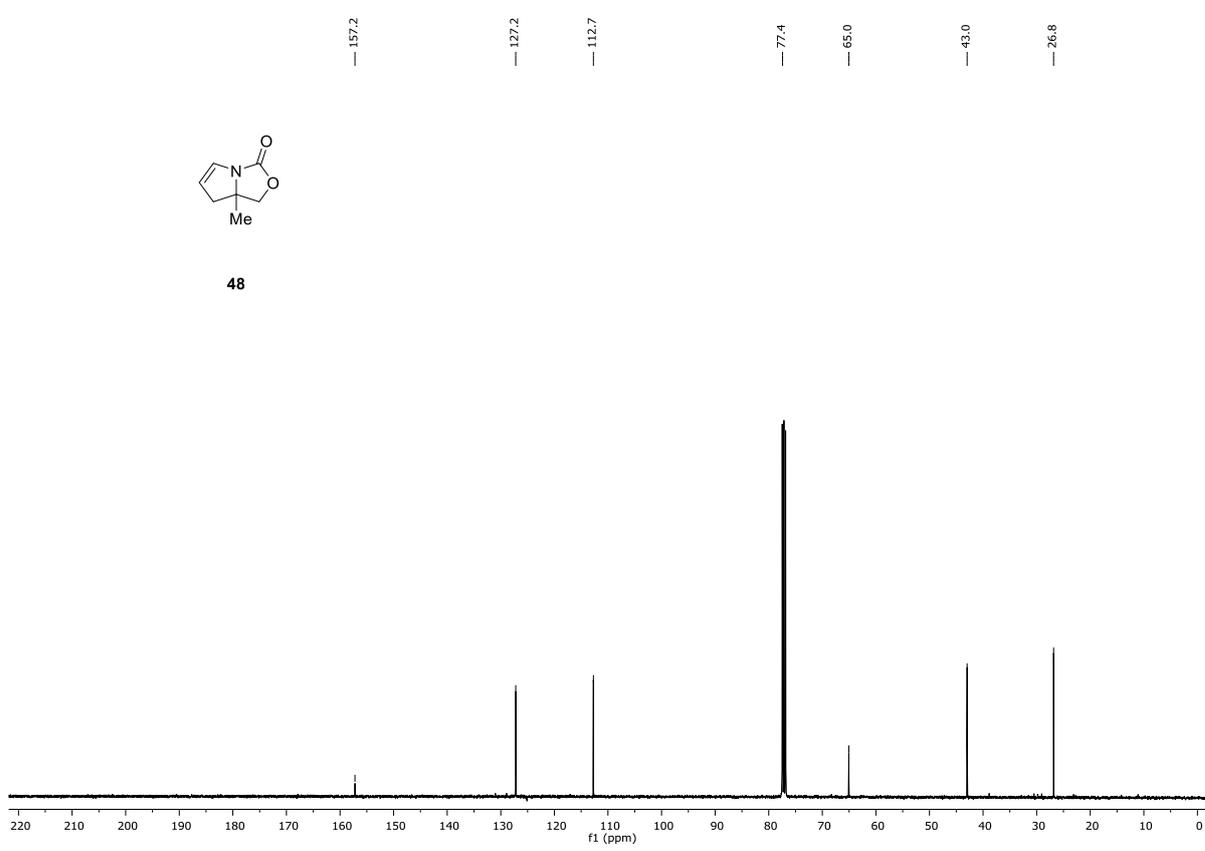
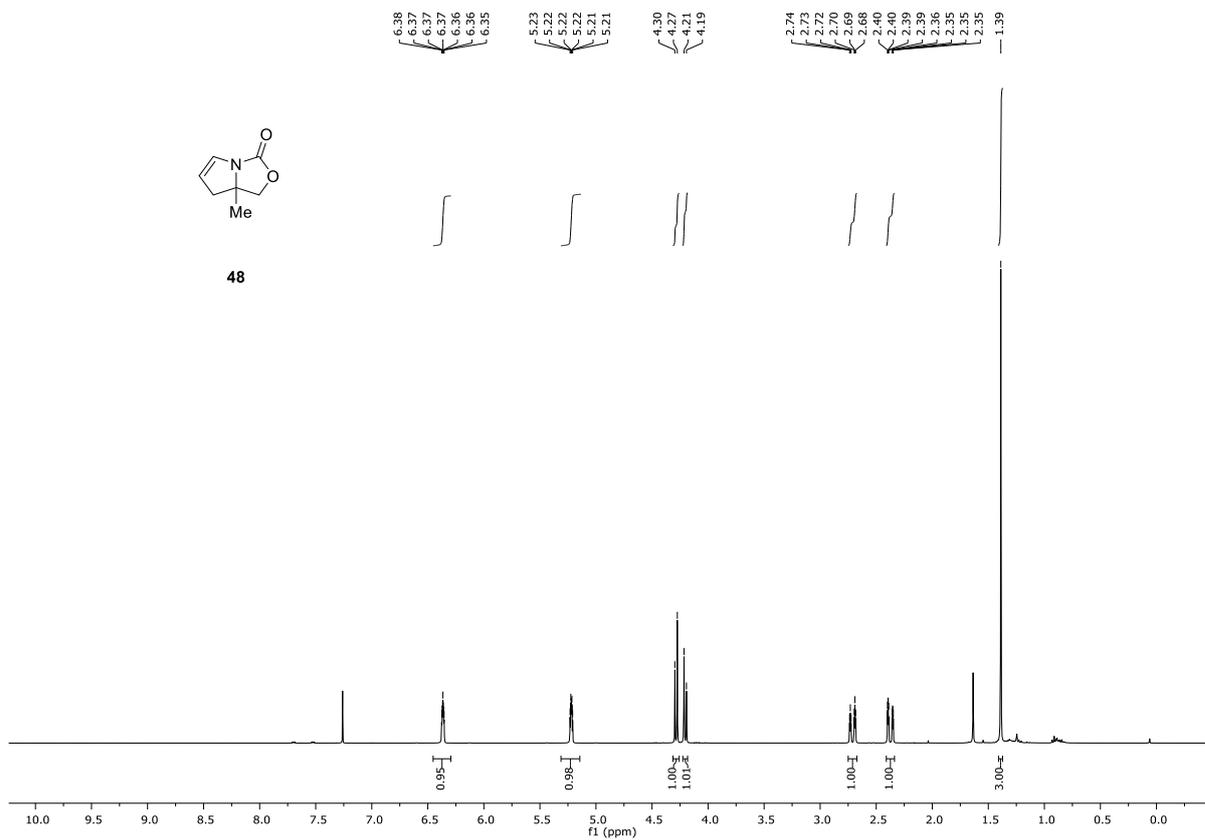


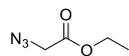




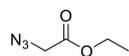
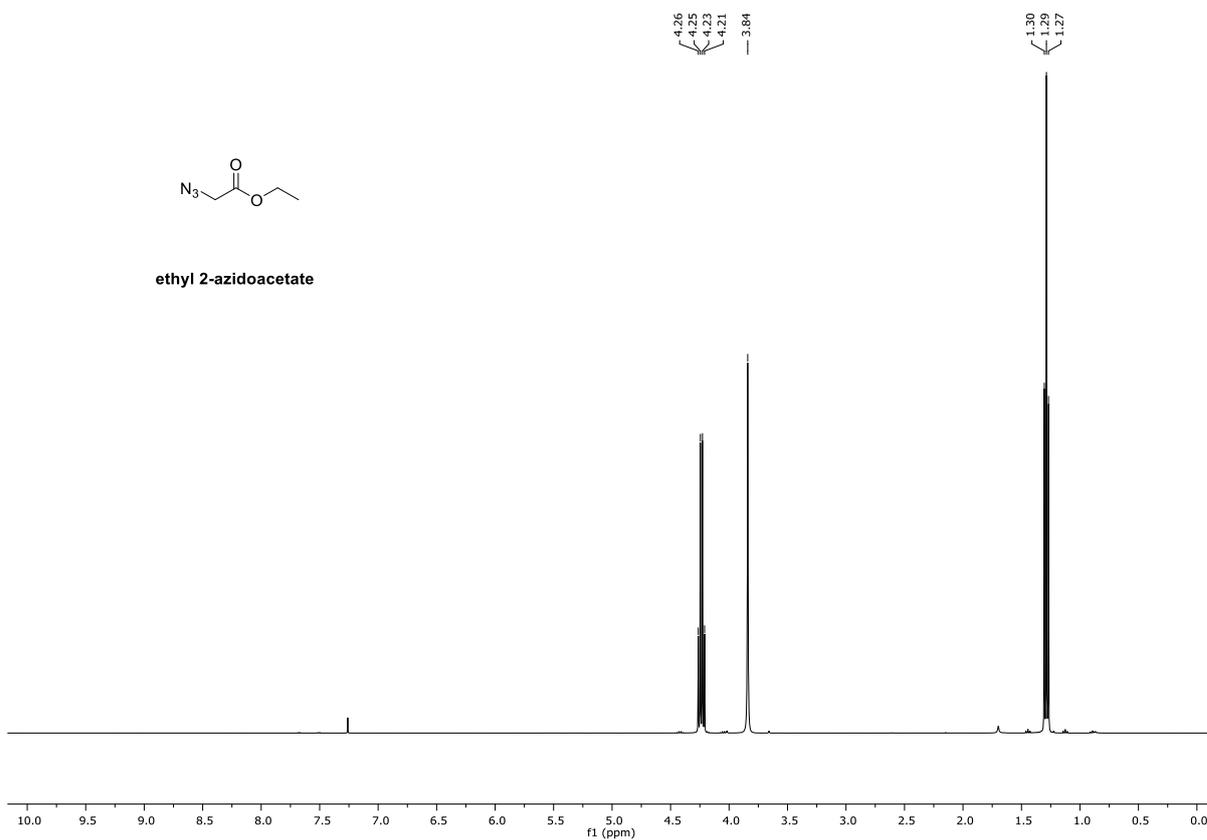




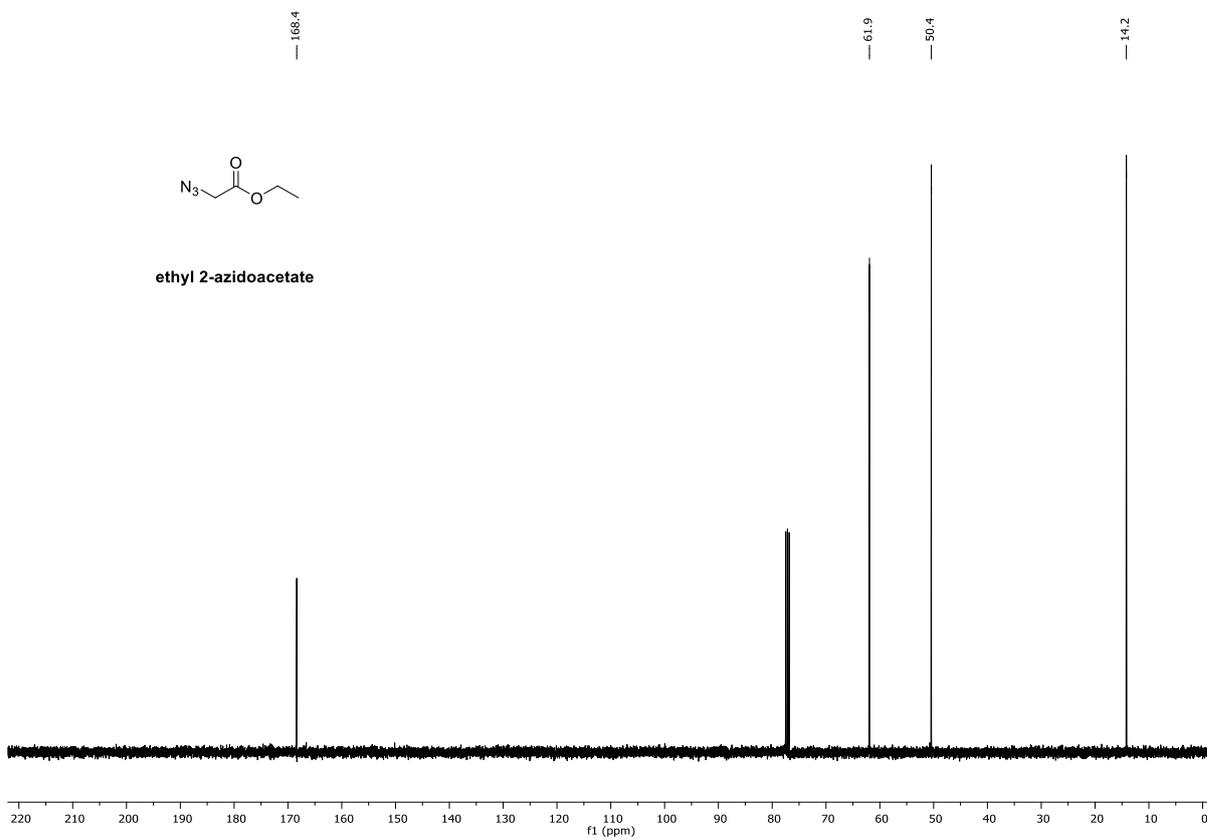


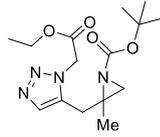


ethyl 2-azidoacetate

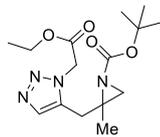
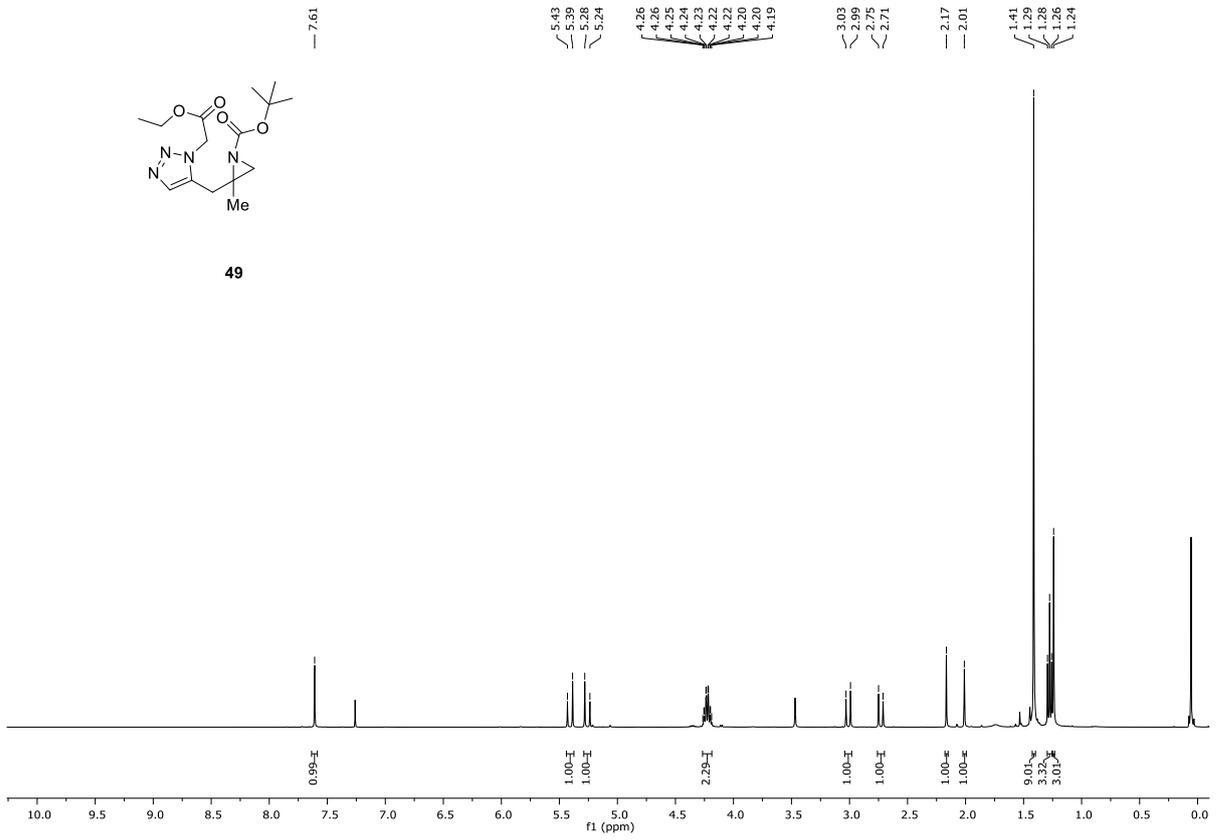


ethyl 2-azidoacetate

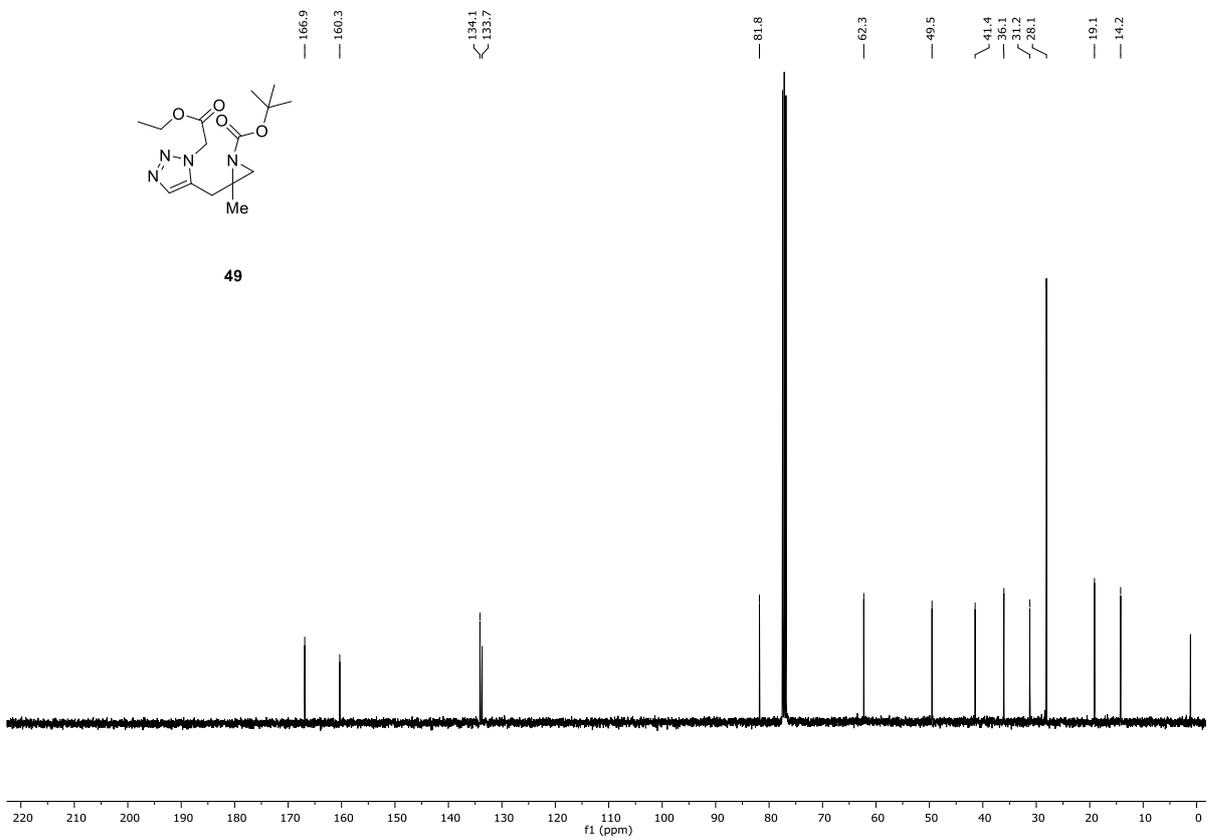


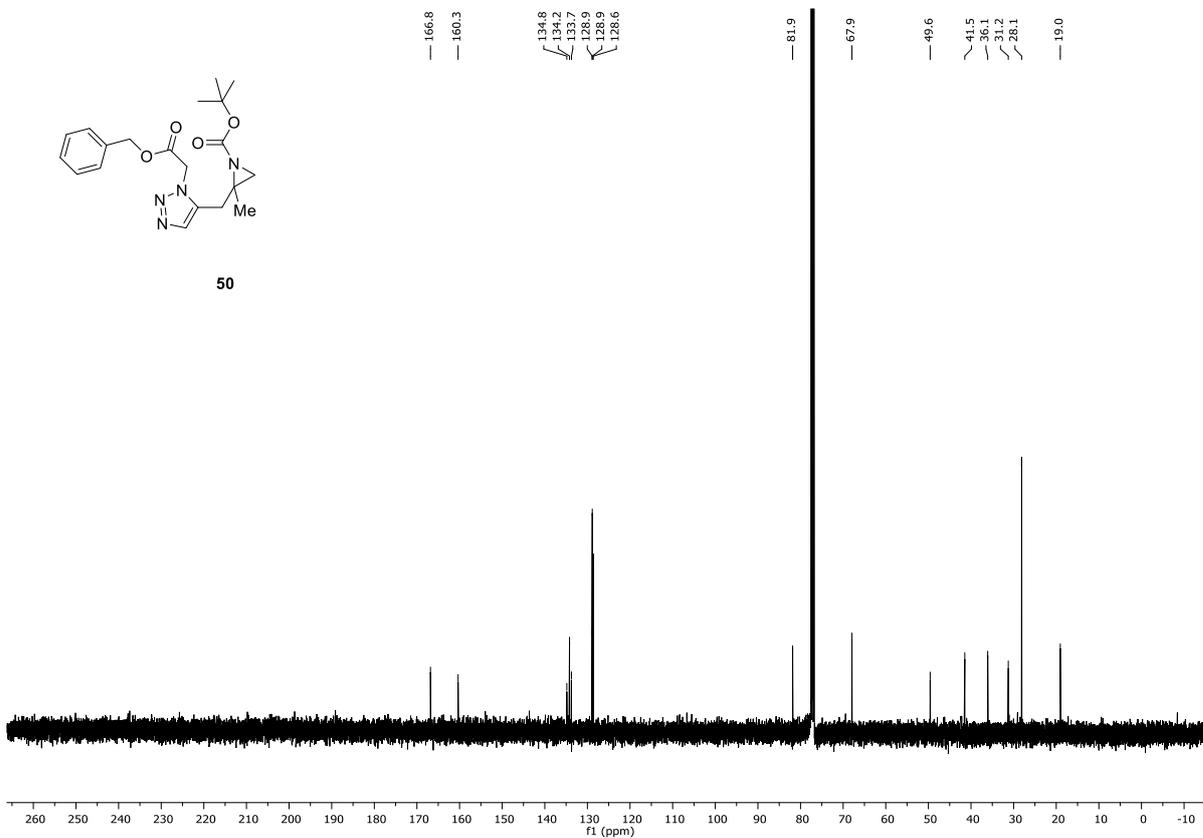
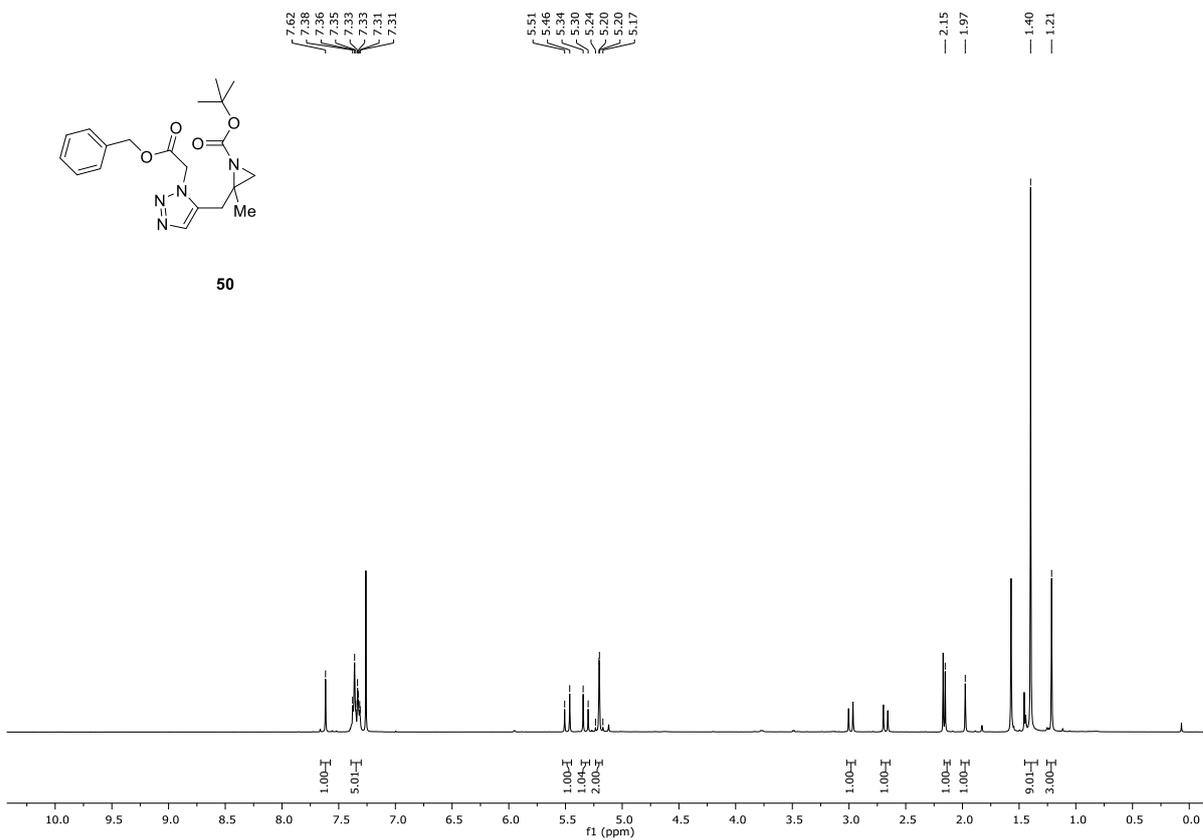


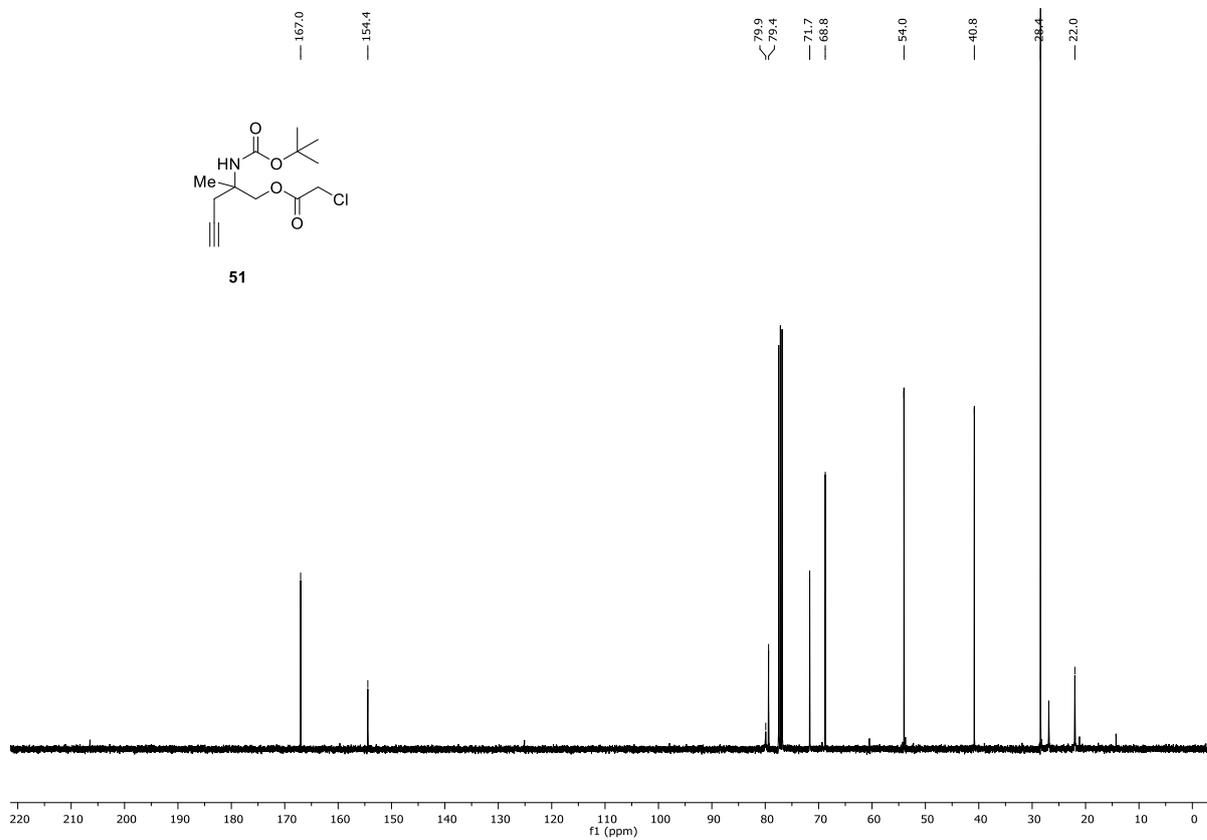
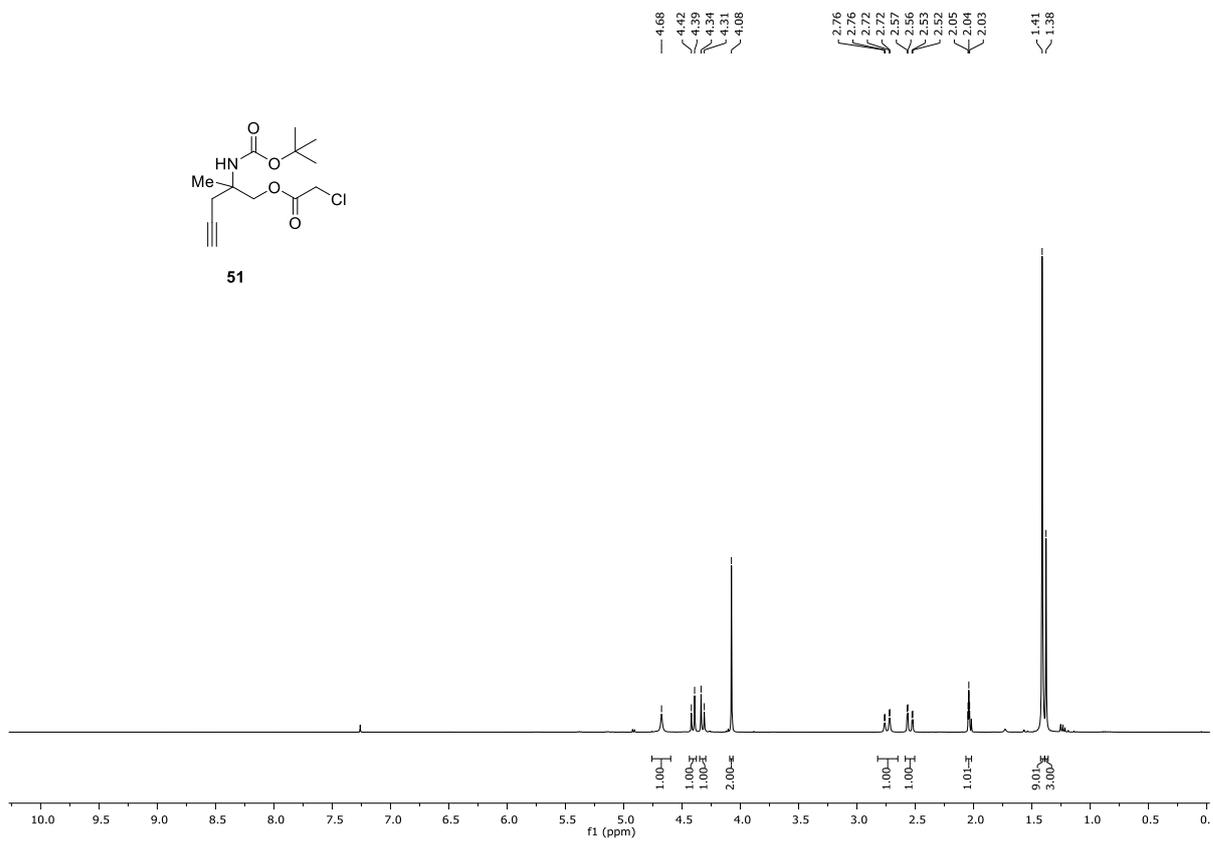
49

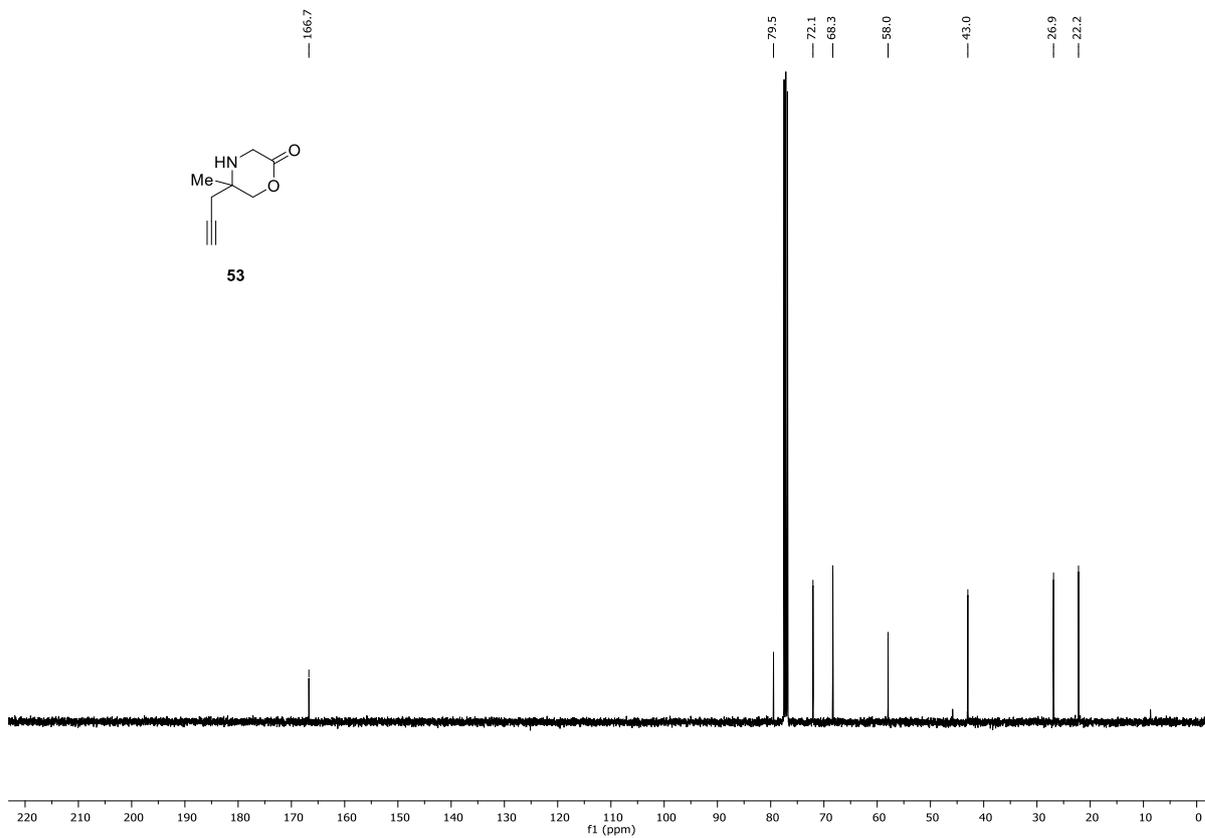
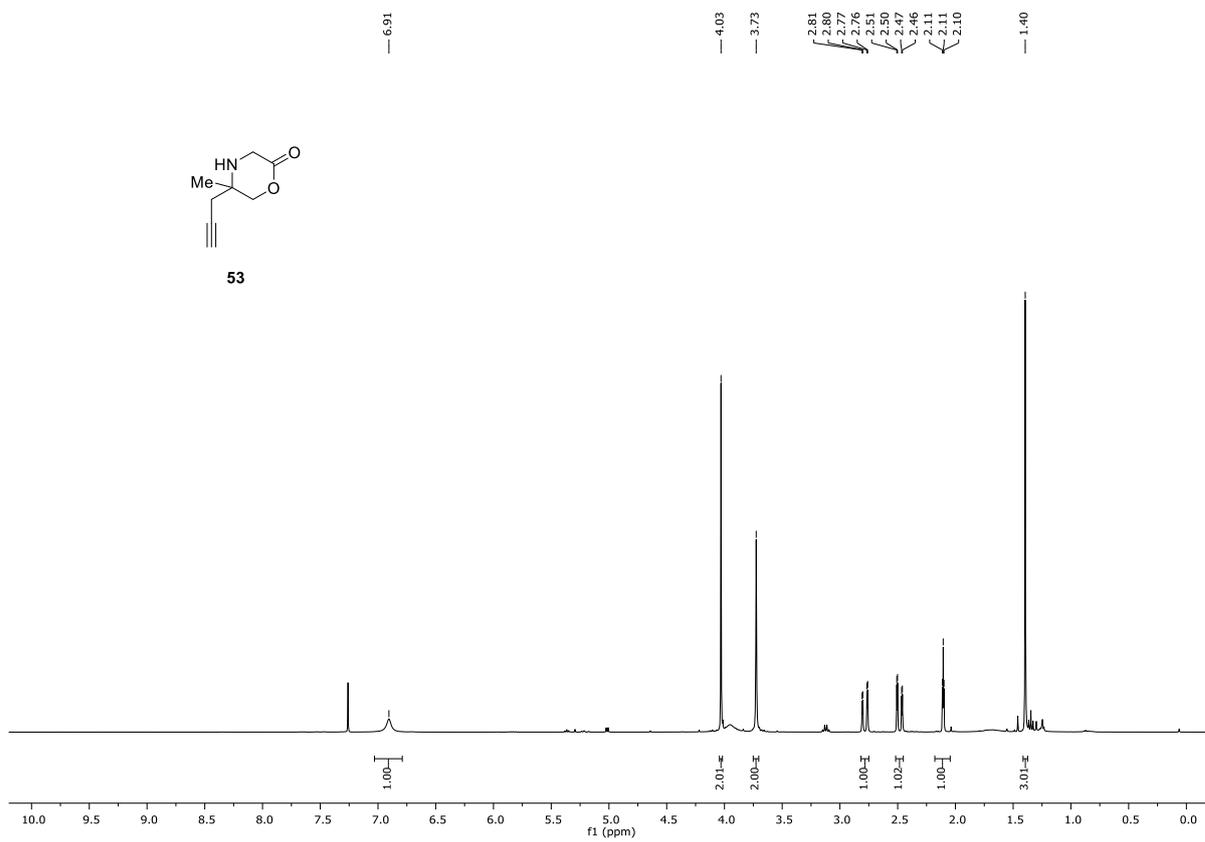


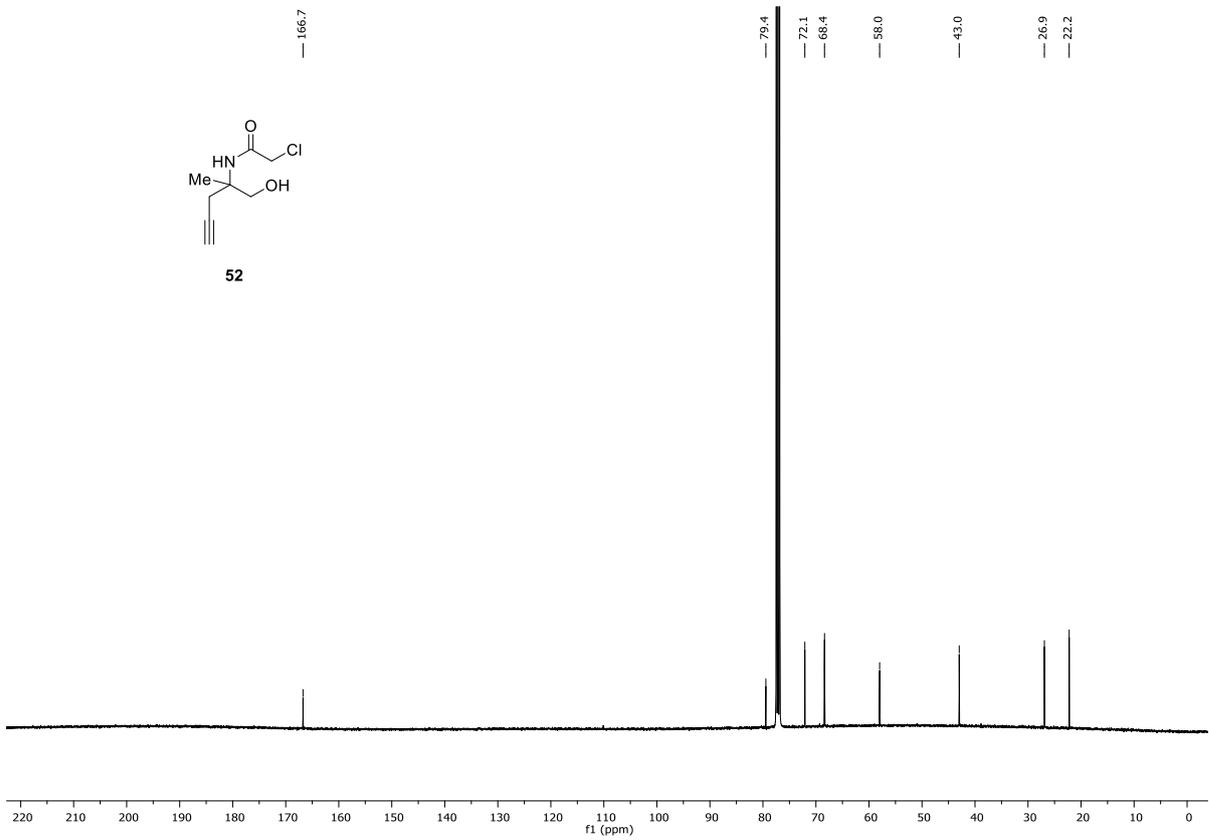
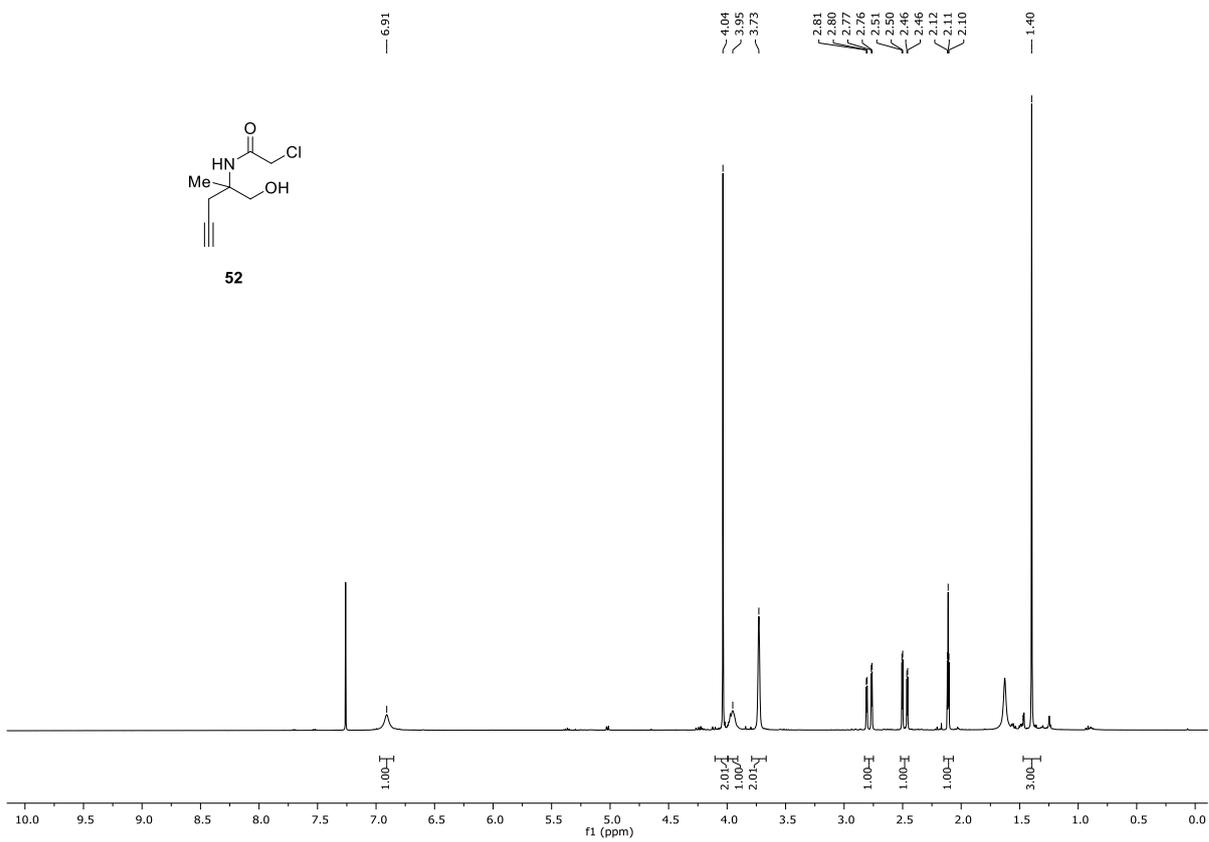
49

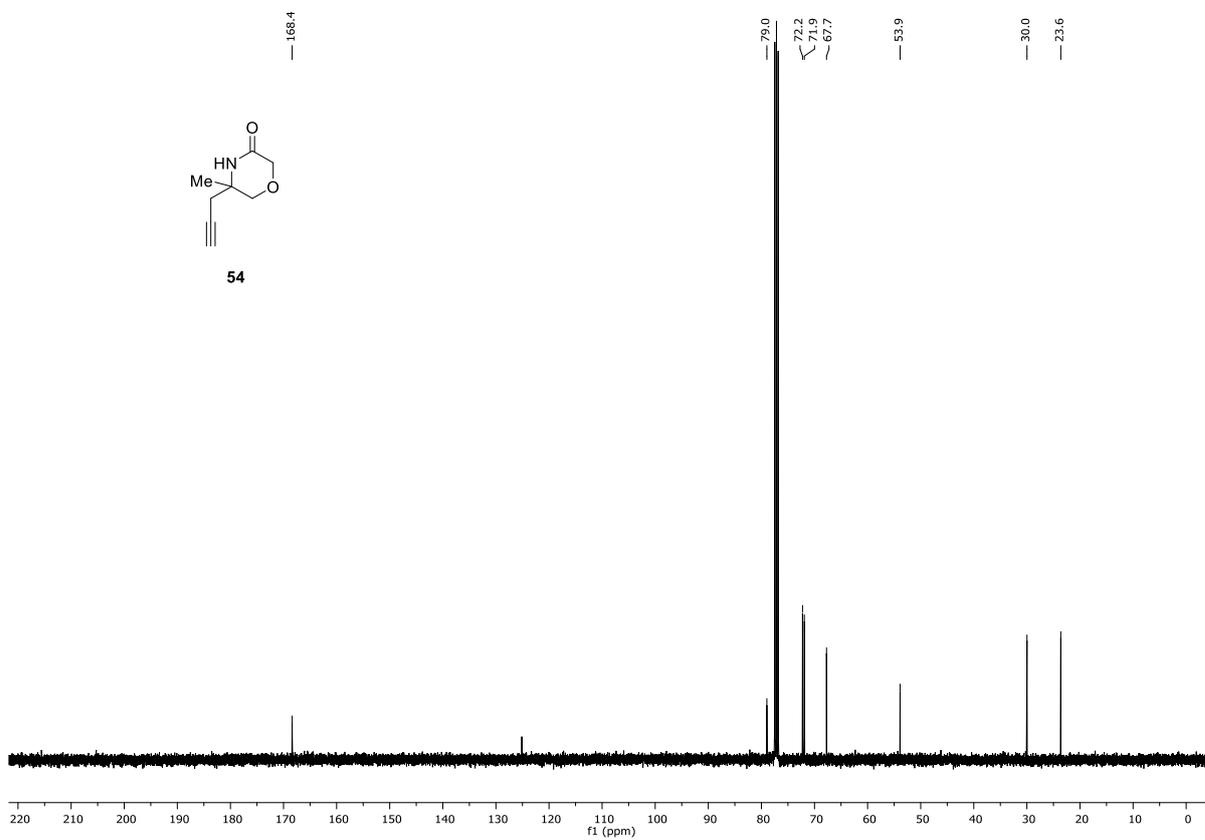
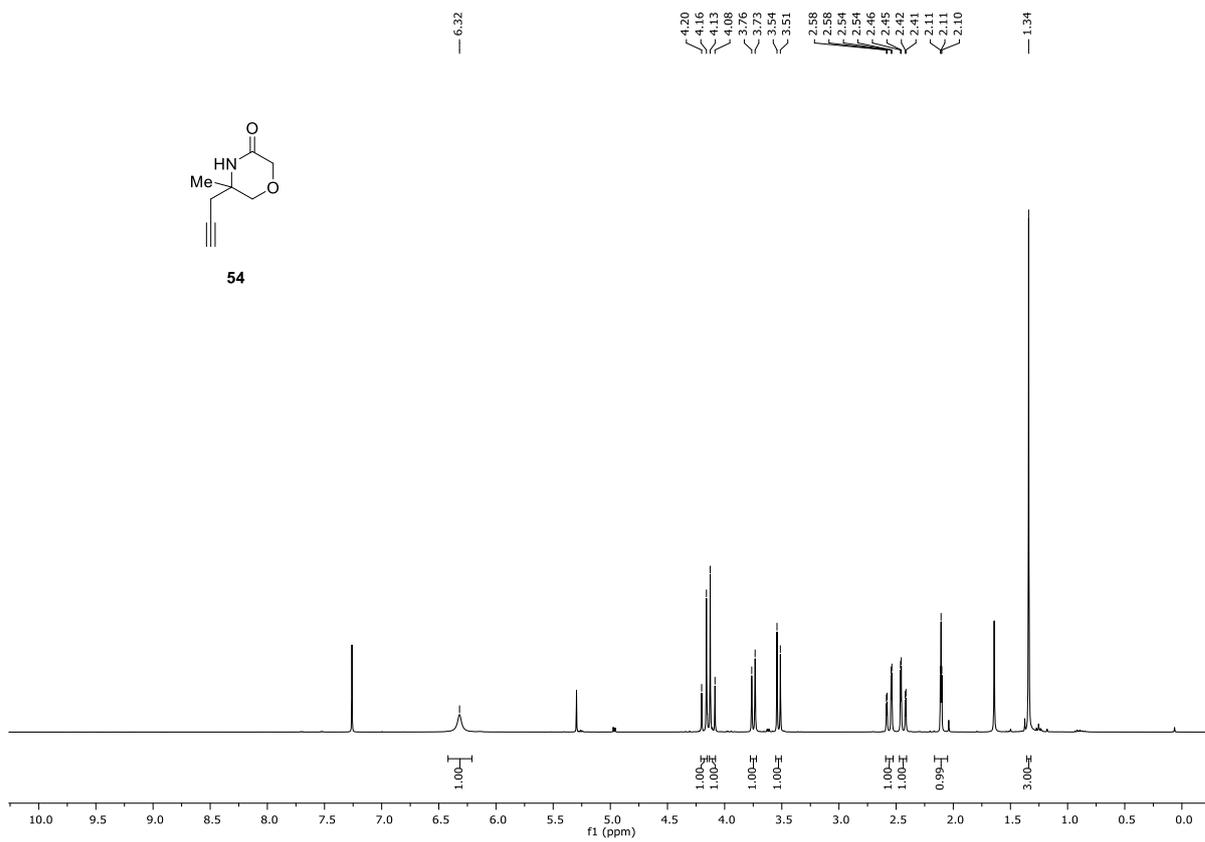


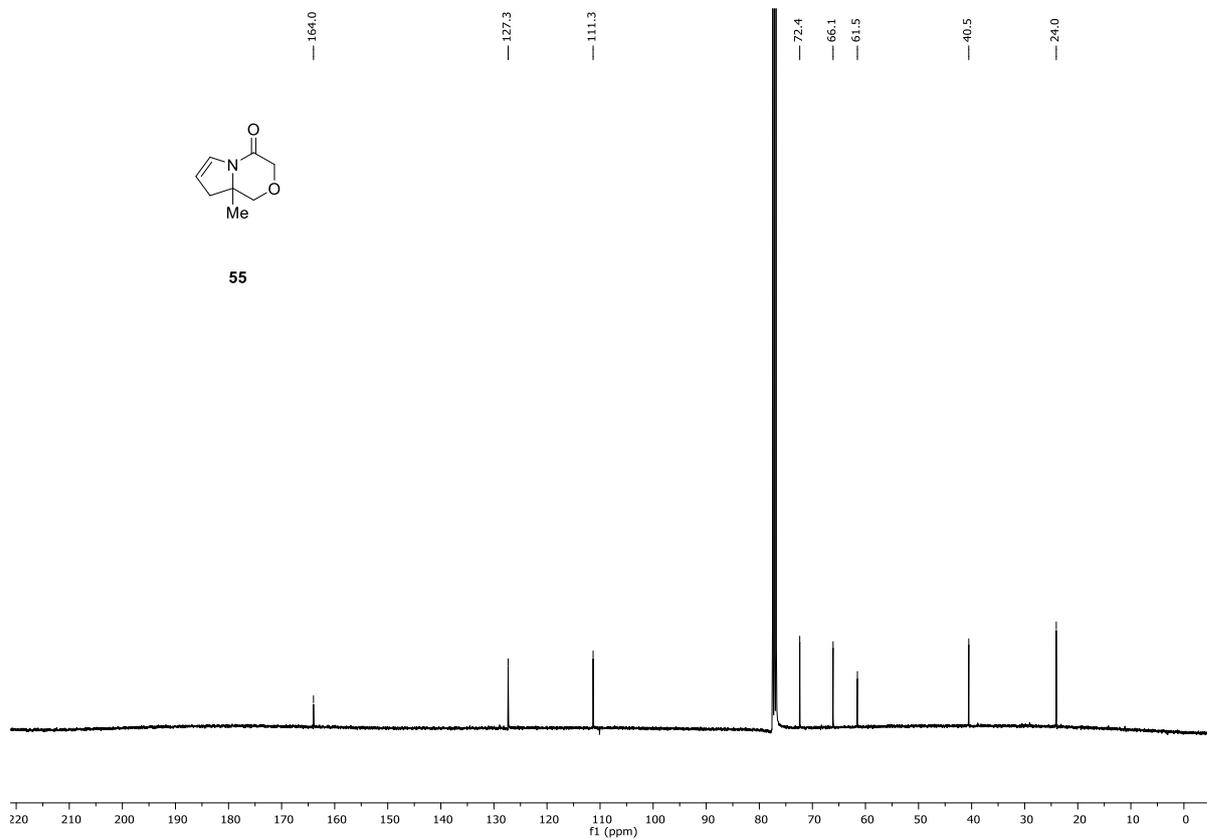
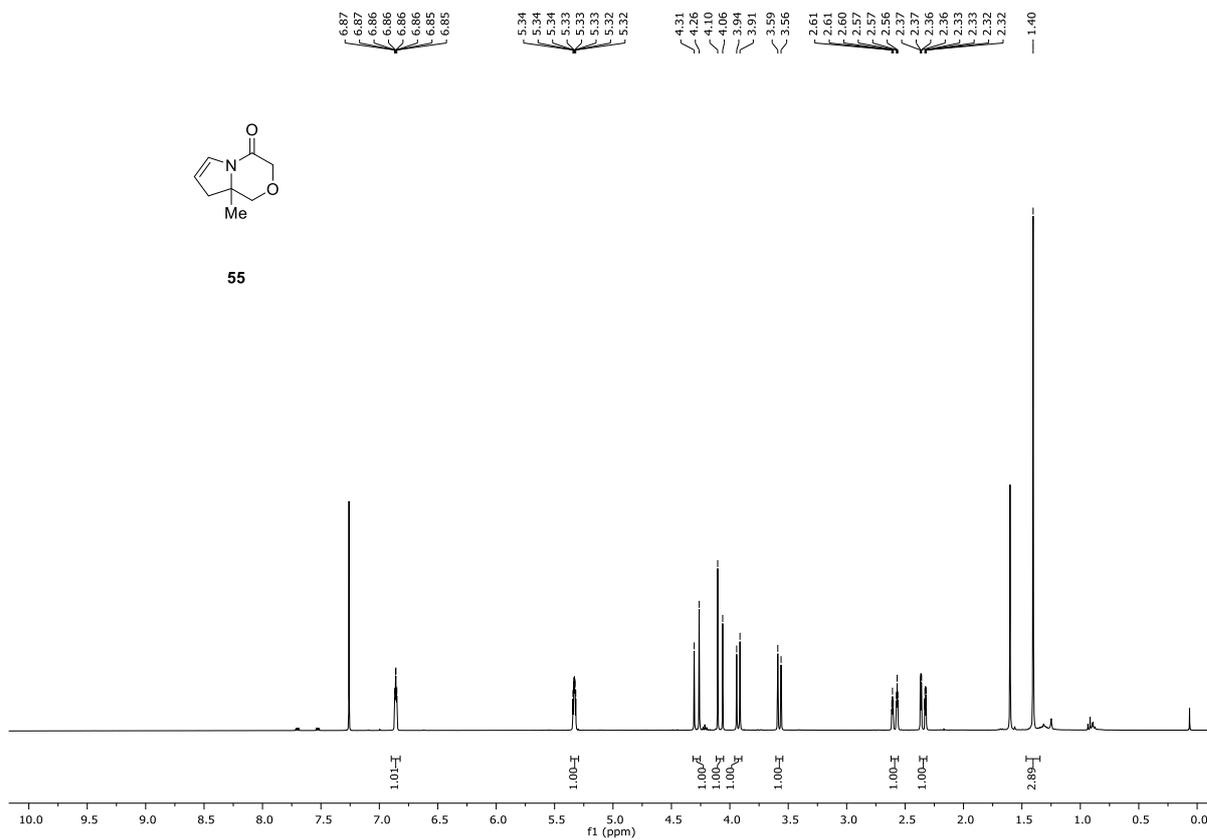


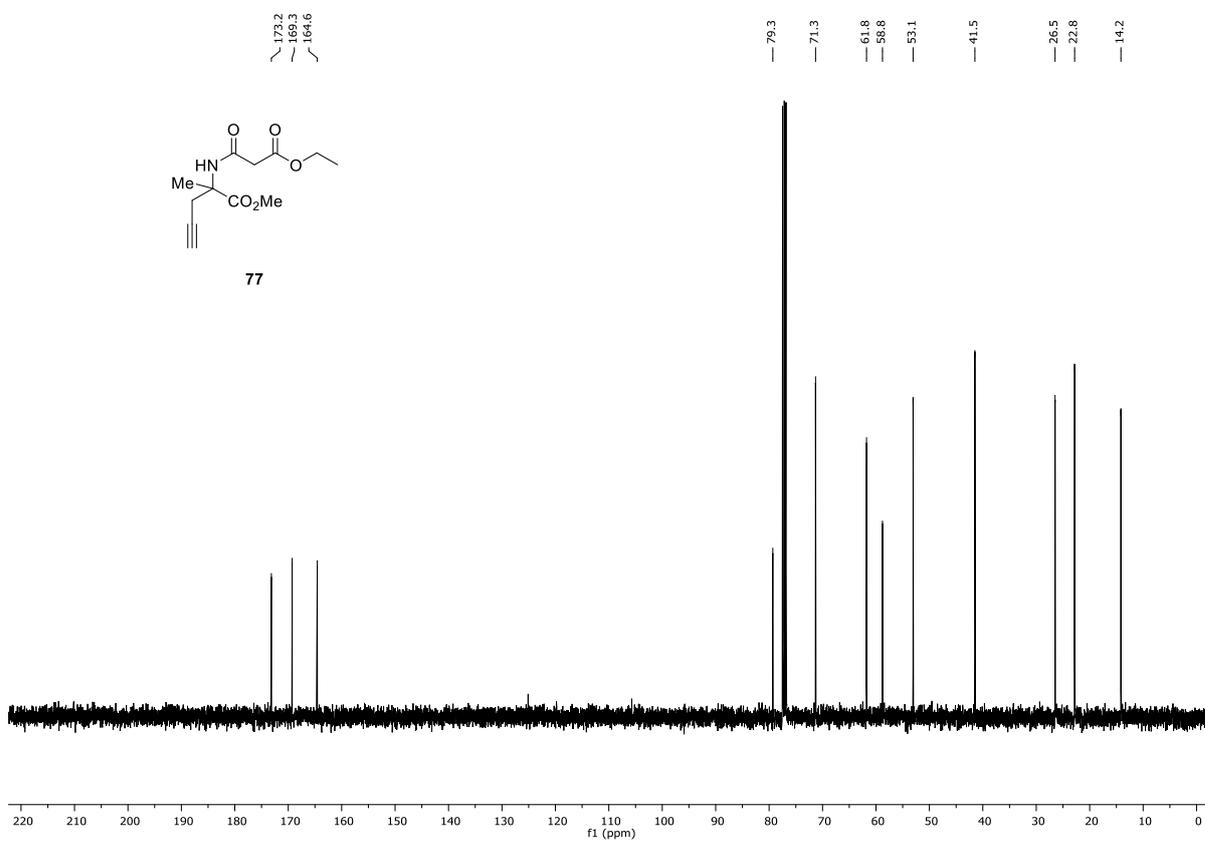
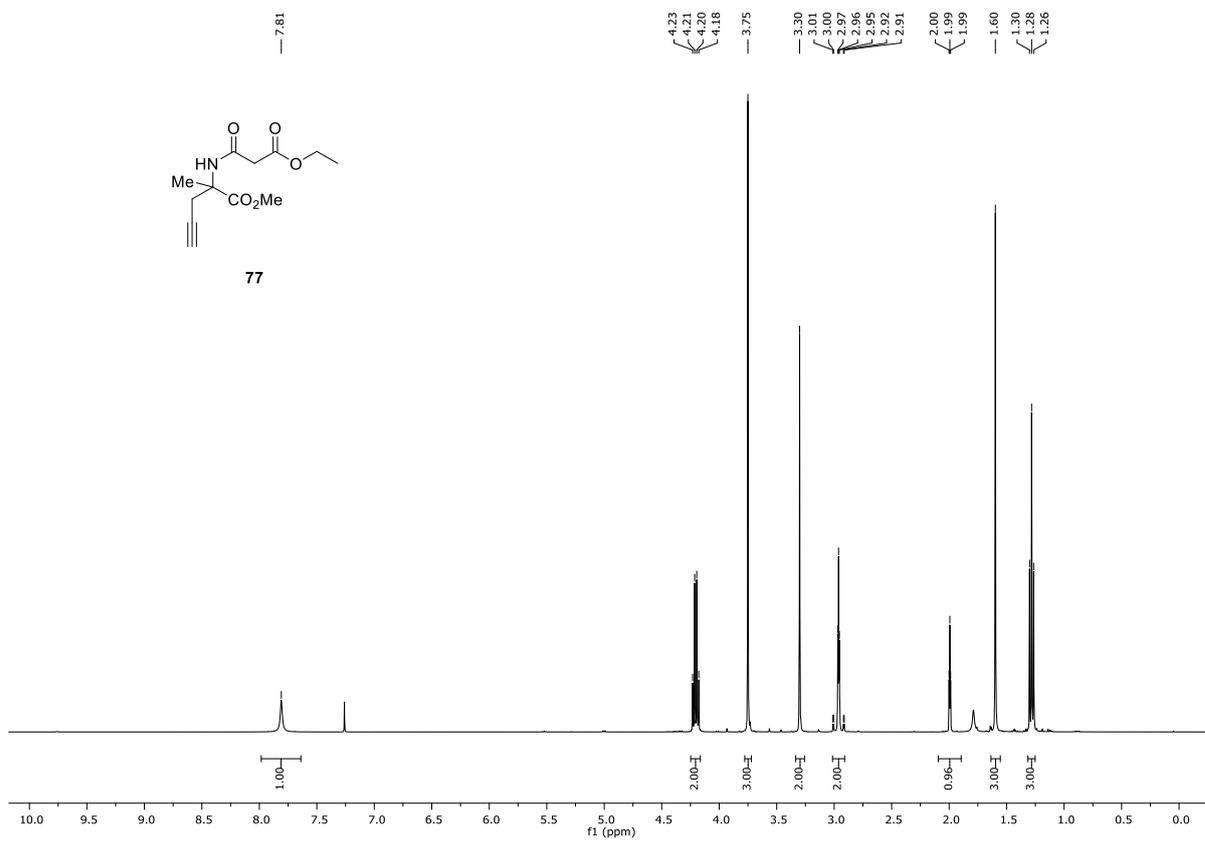


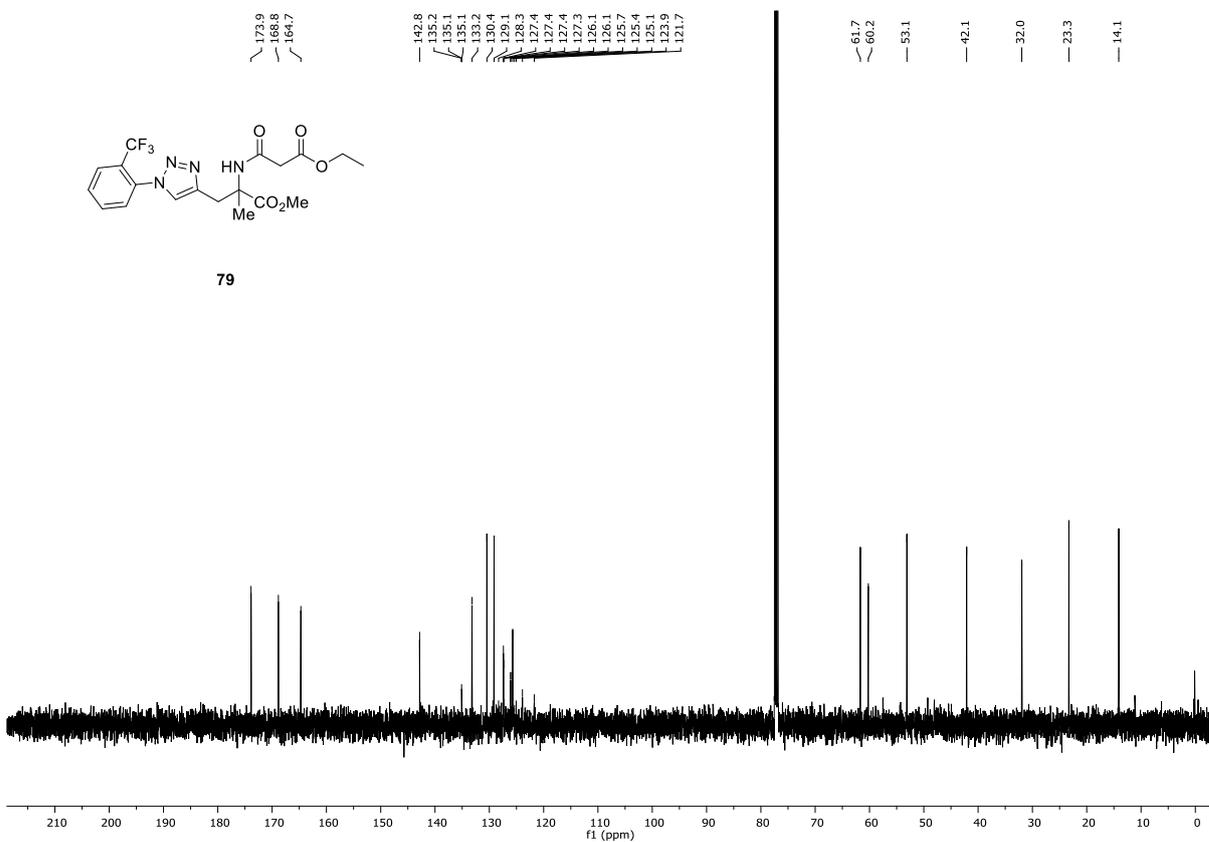
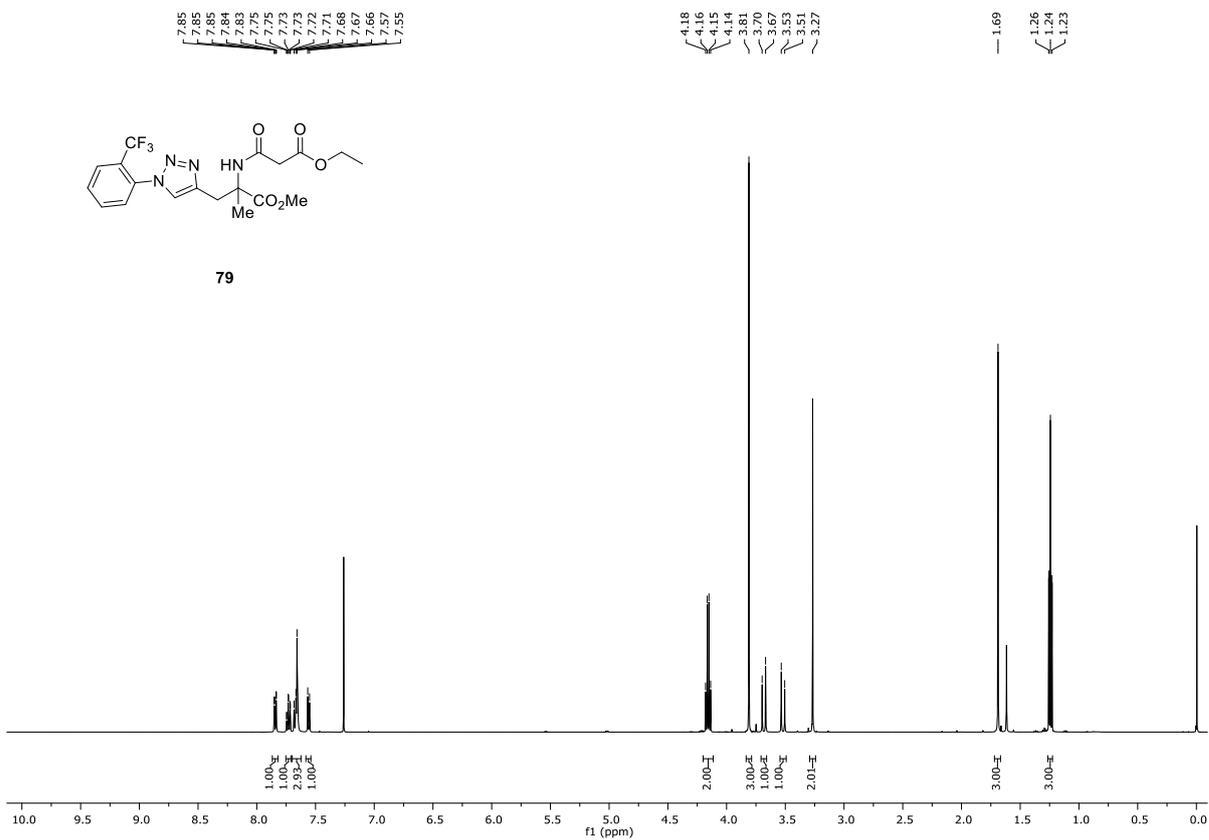


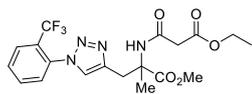




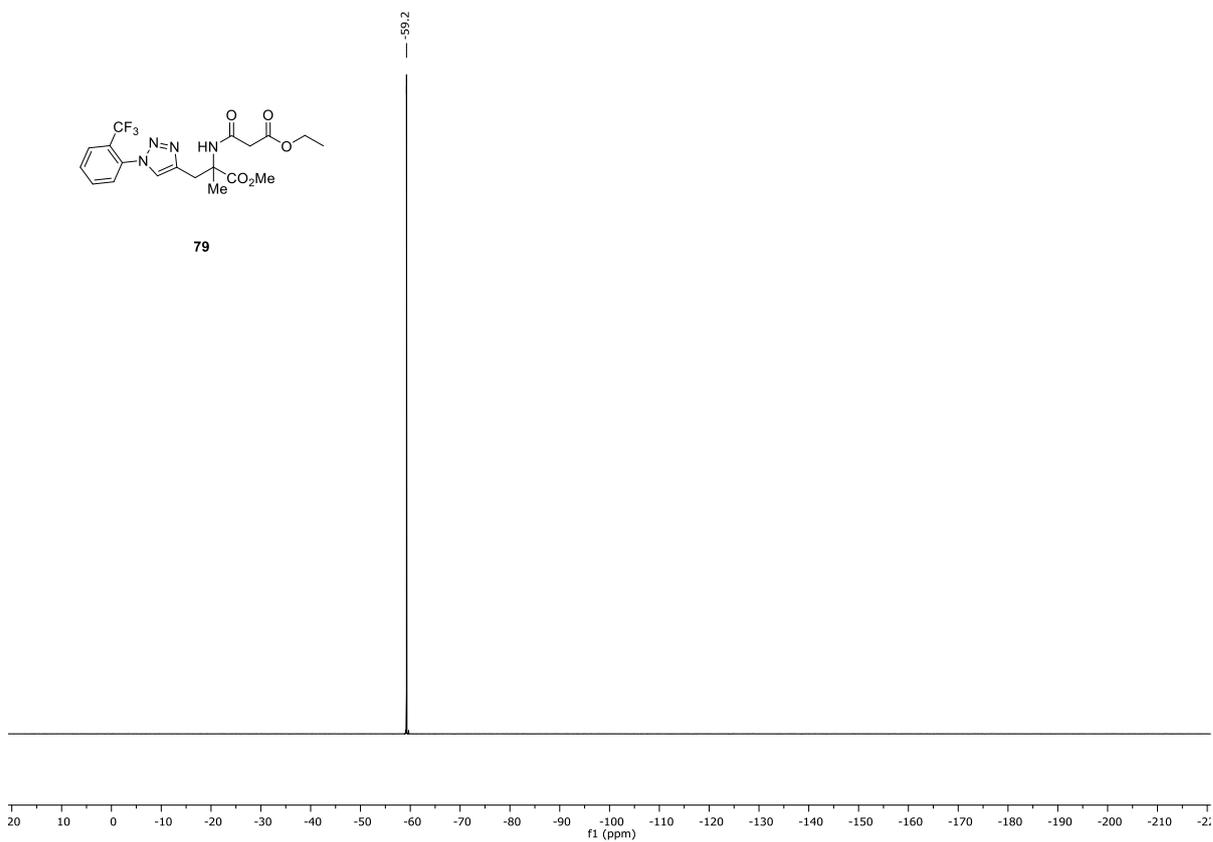


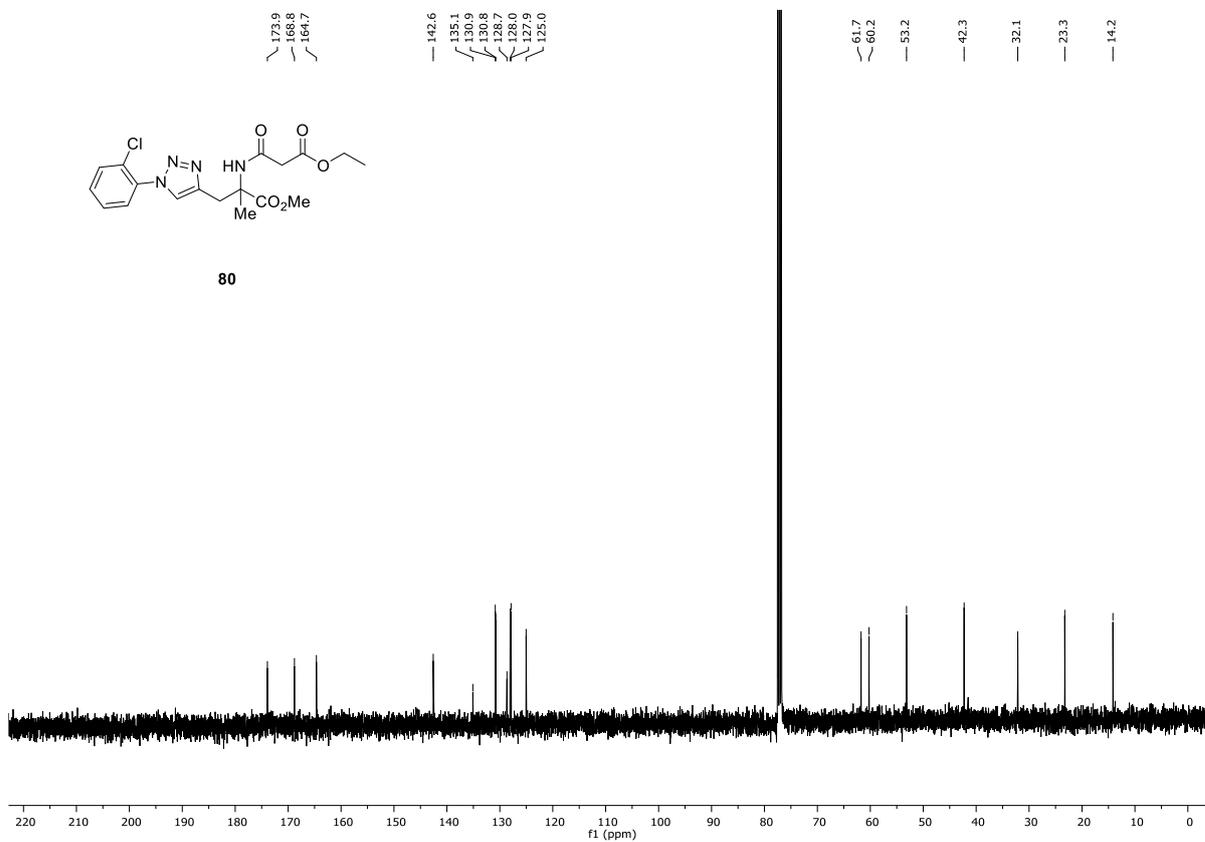
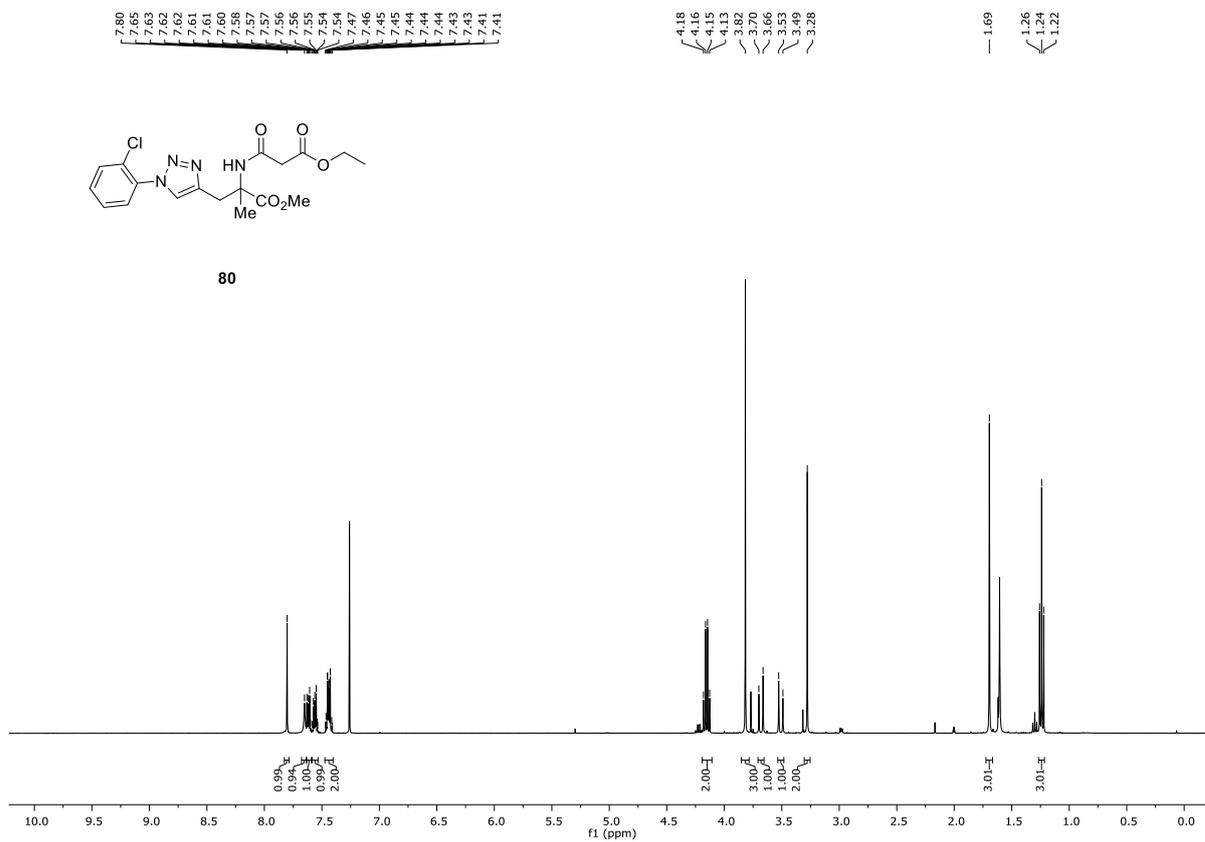


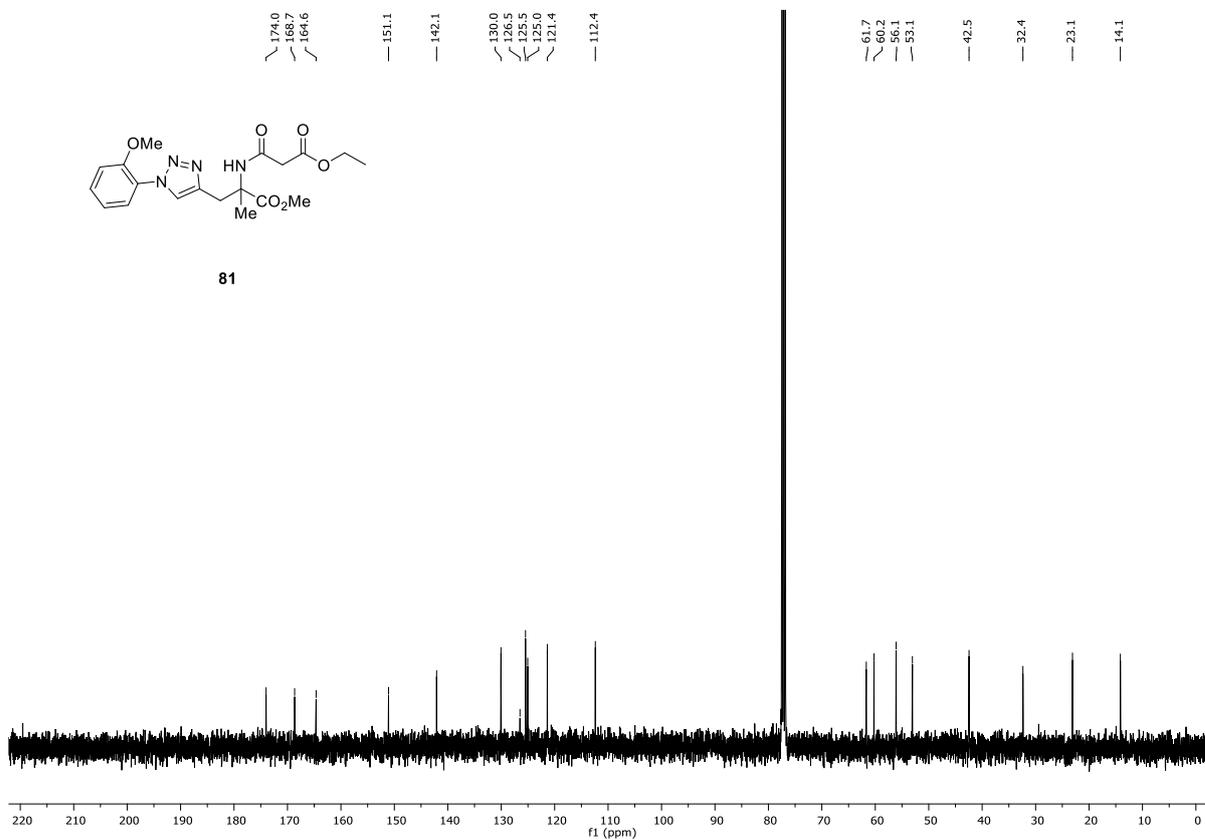
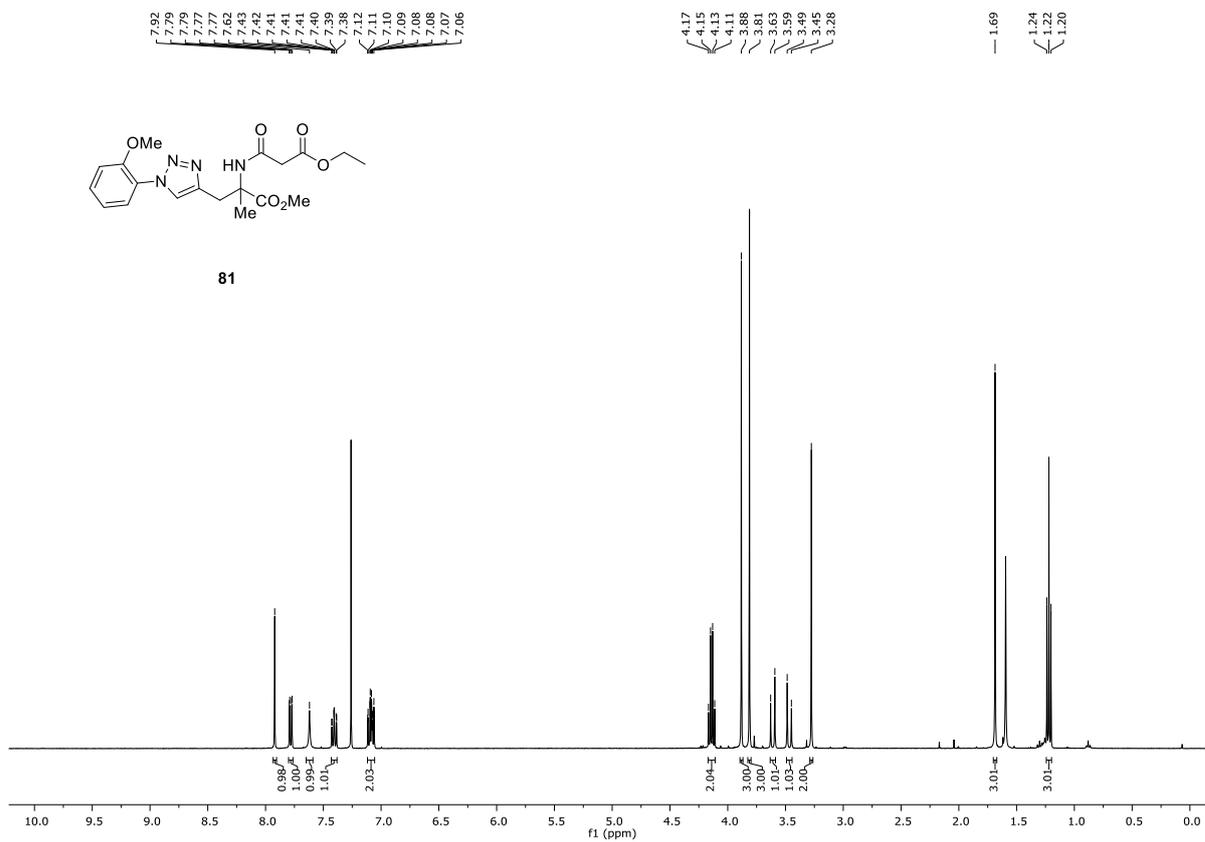


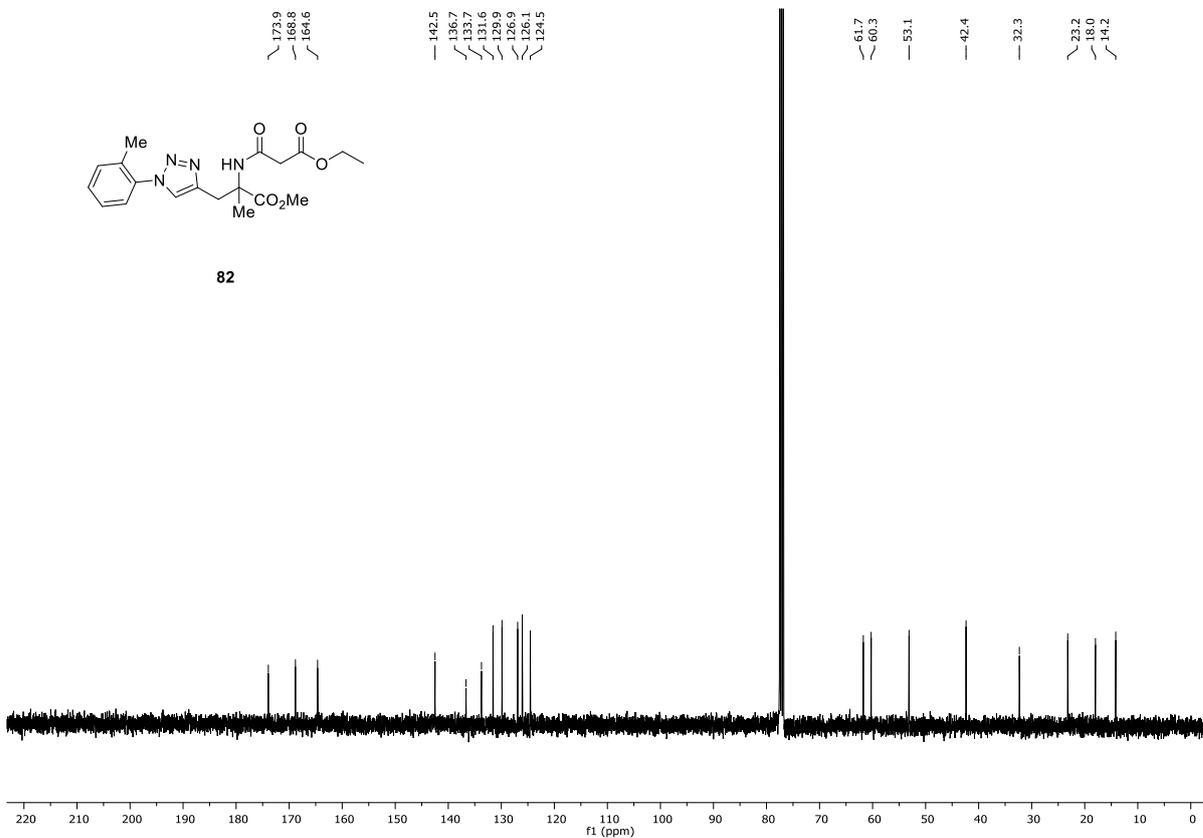
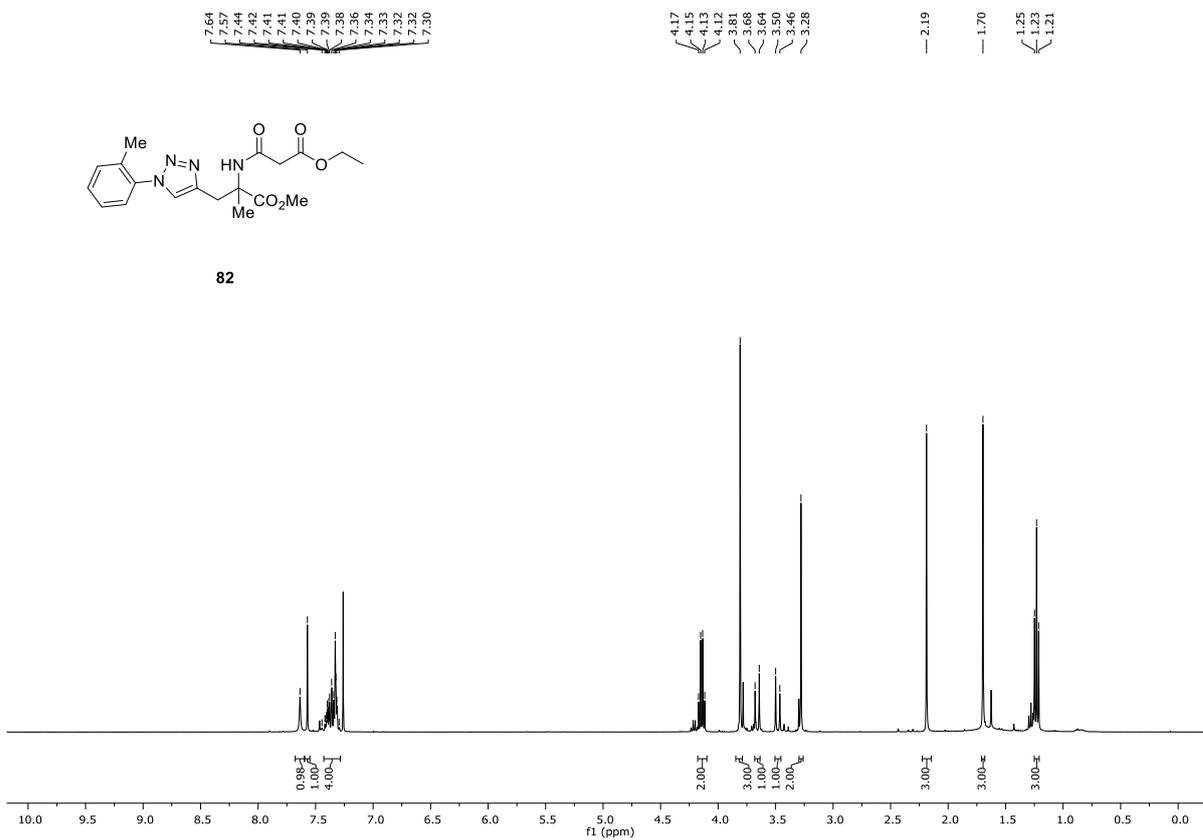


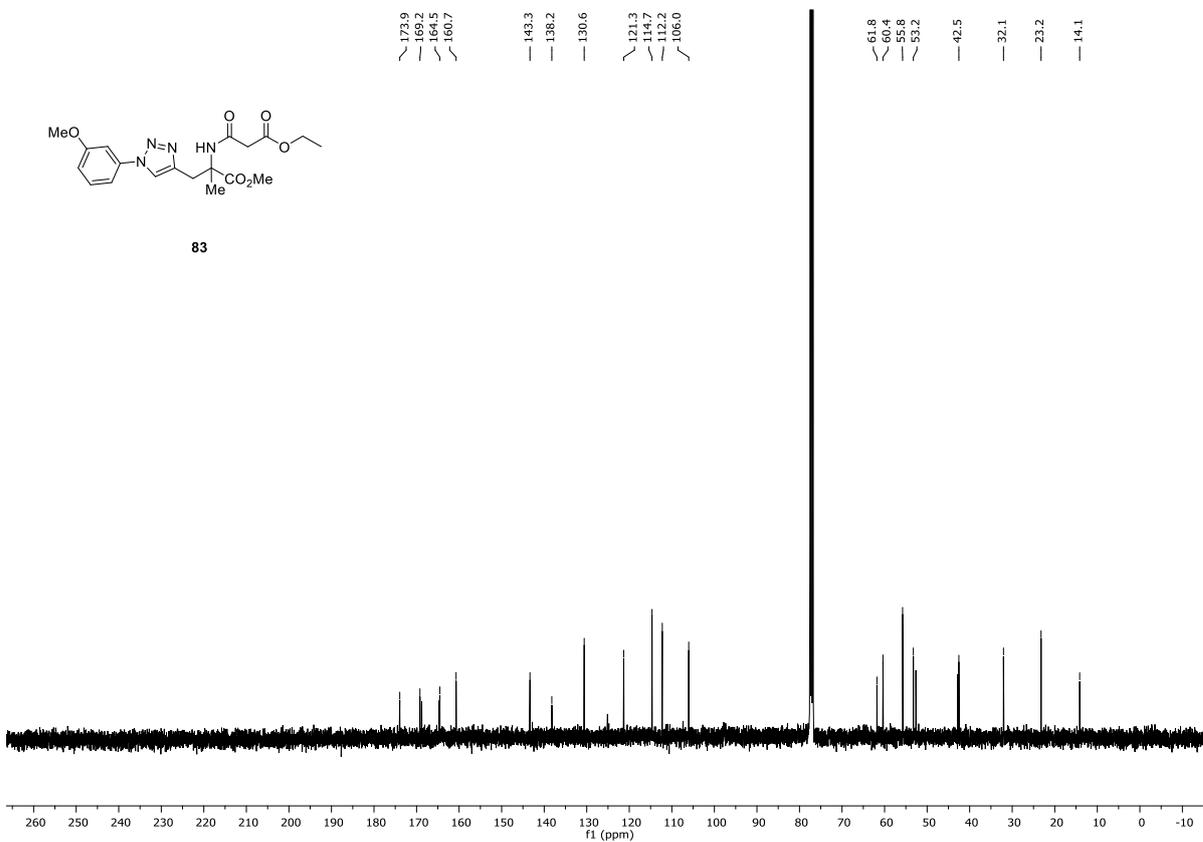
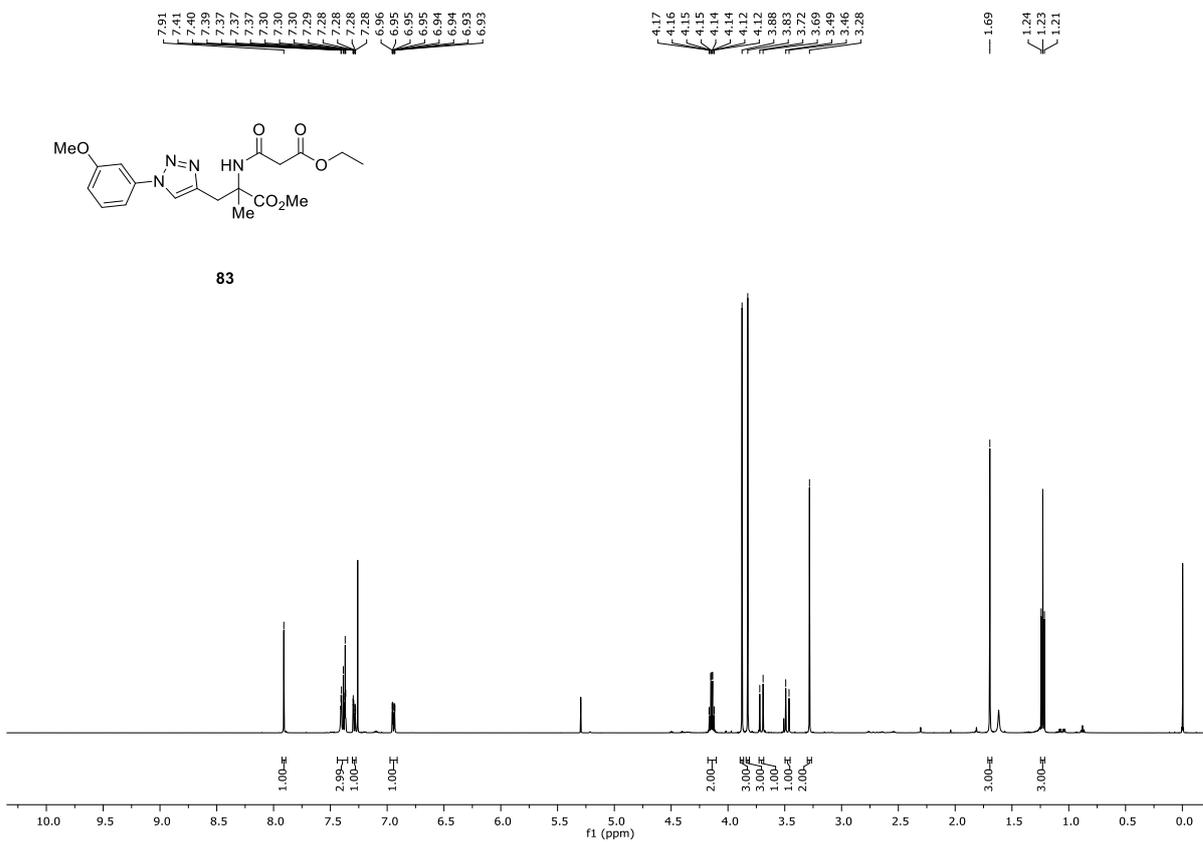
79

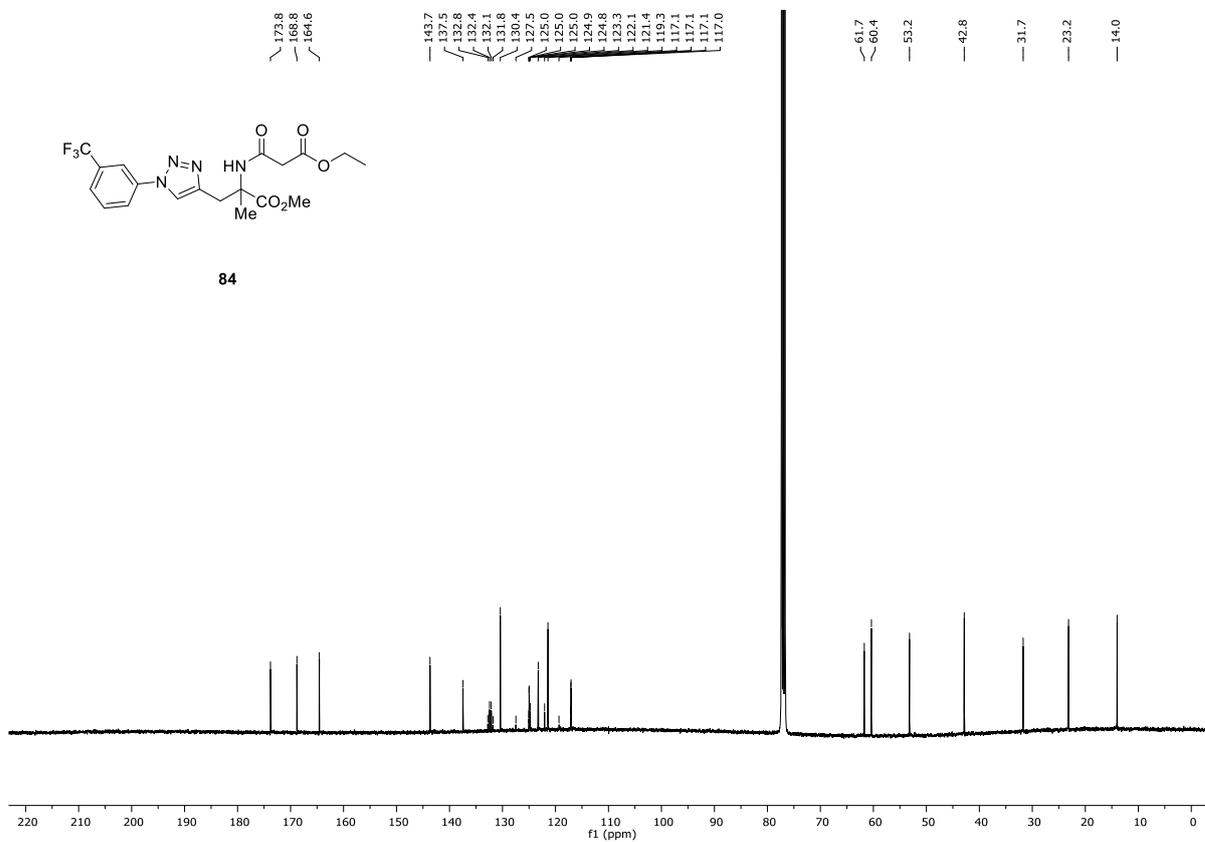
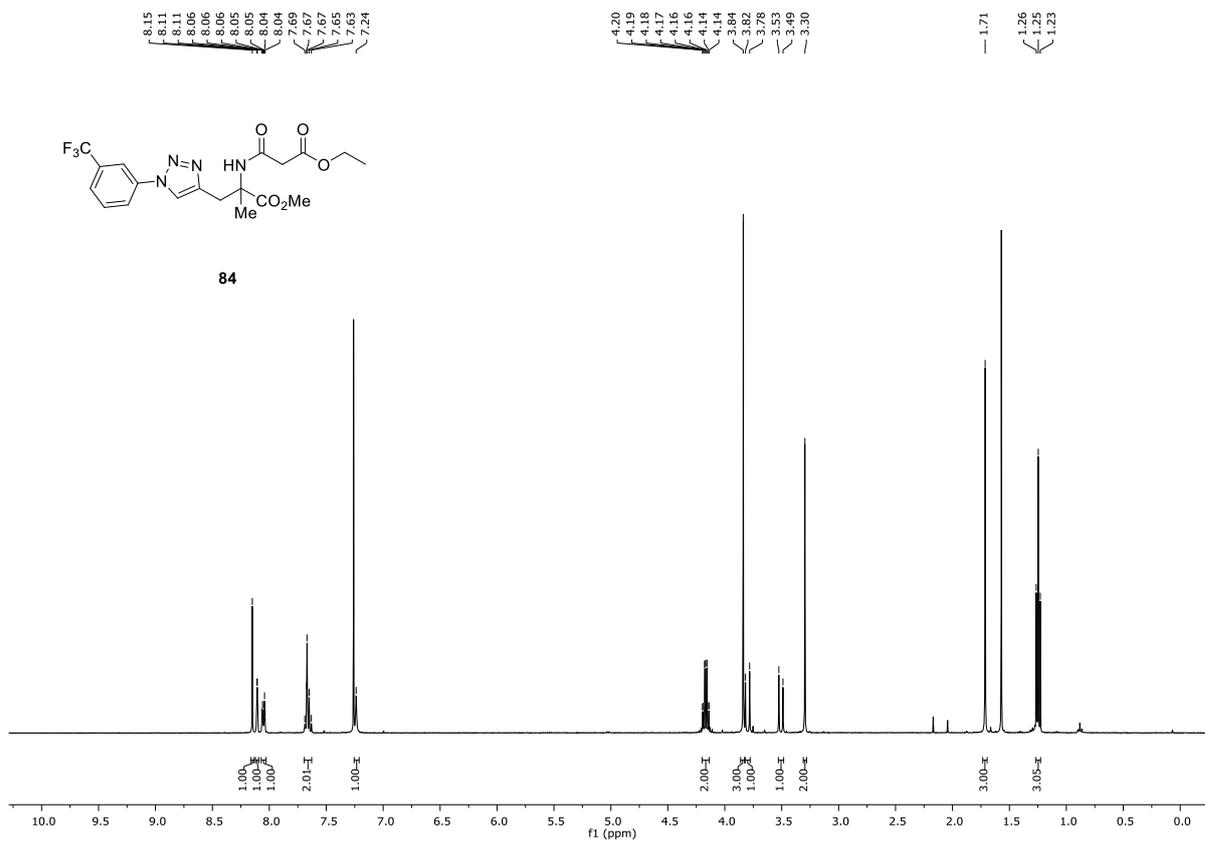


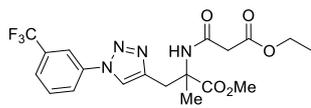




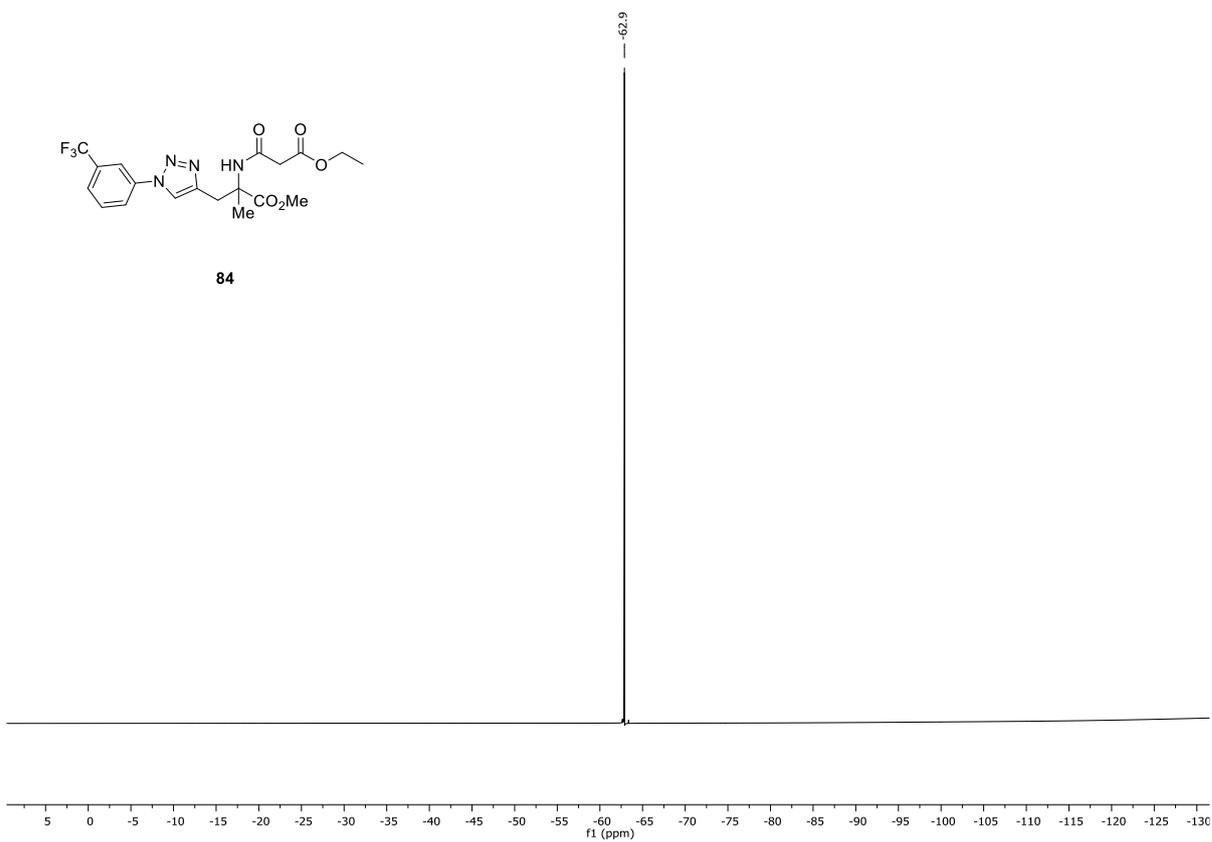


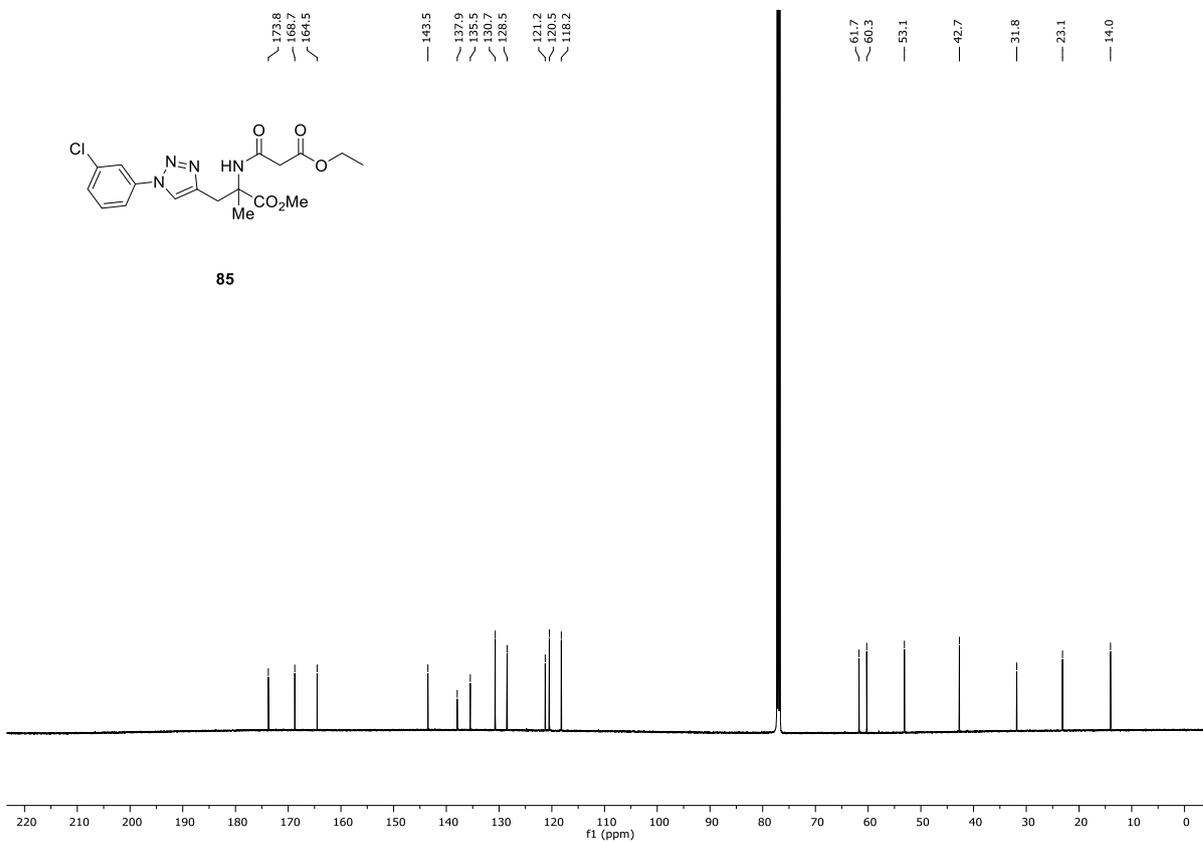
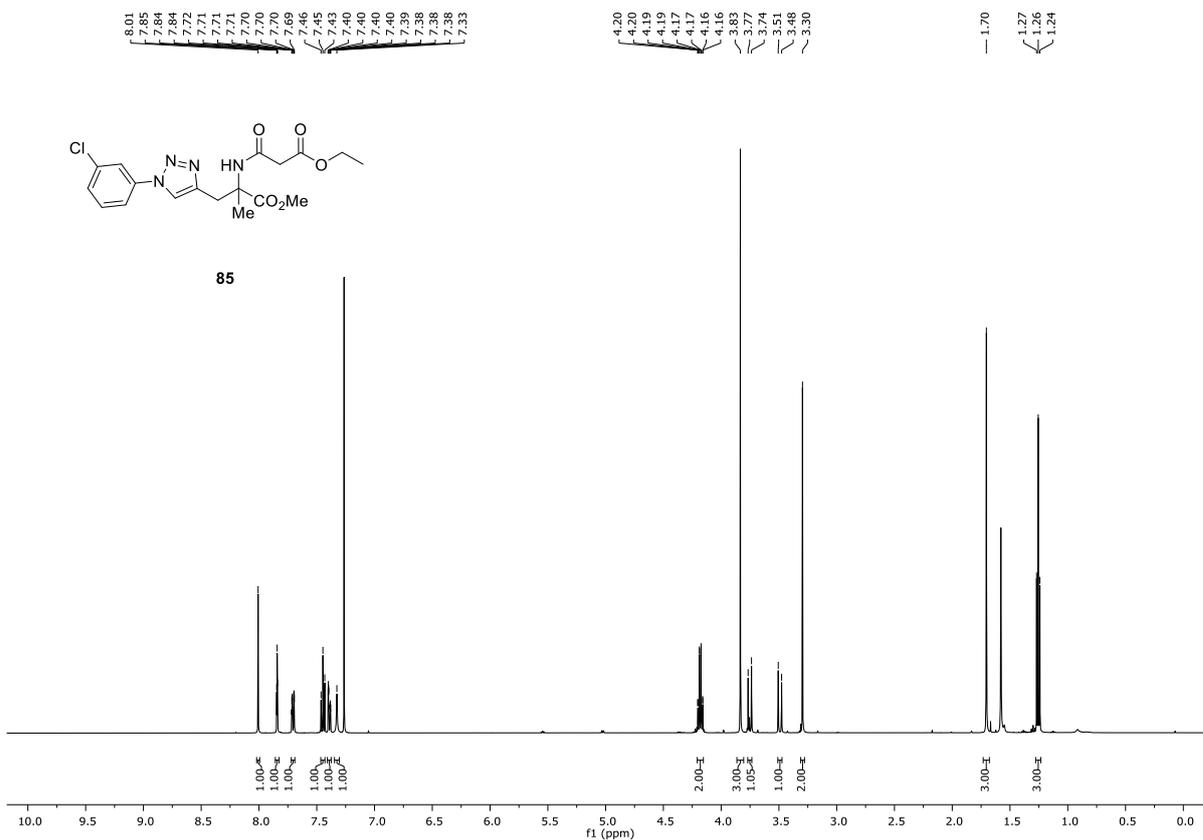


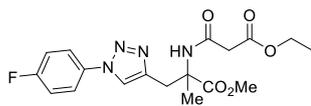




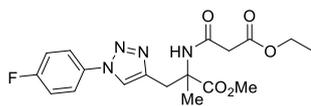
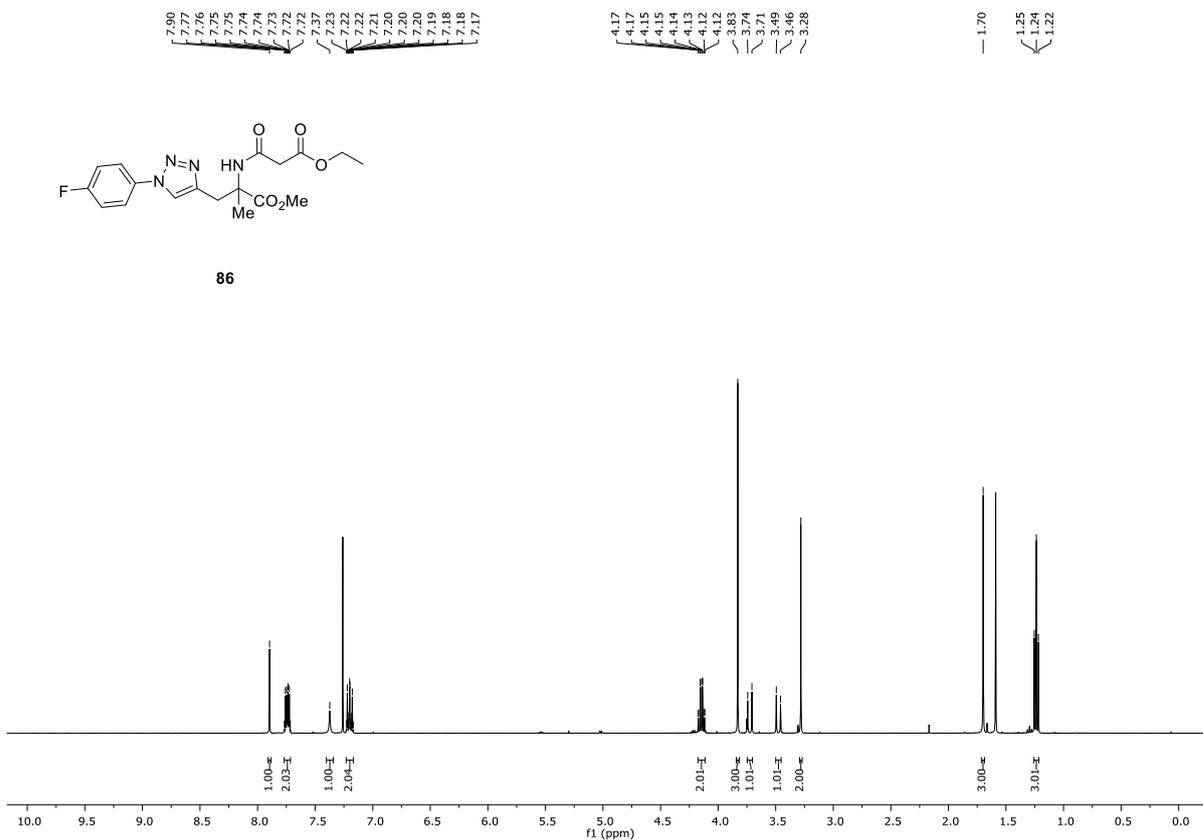
84



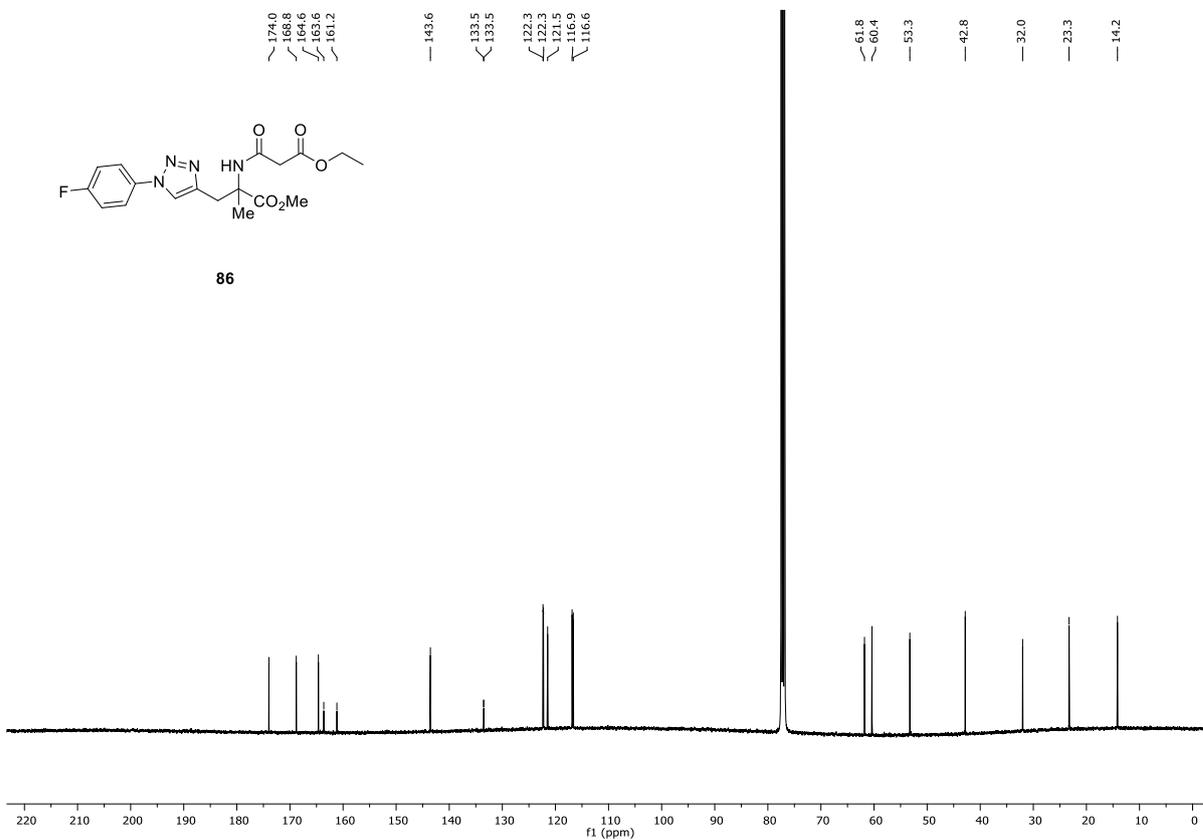


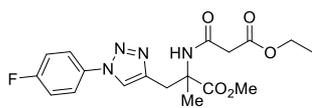


86



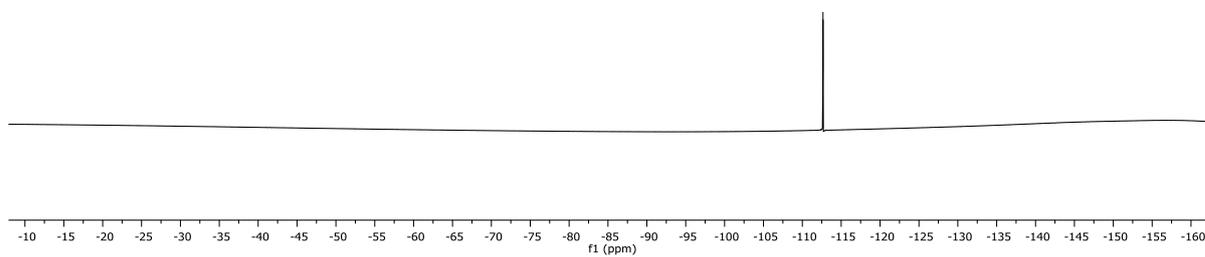
86

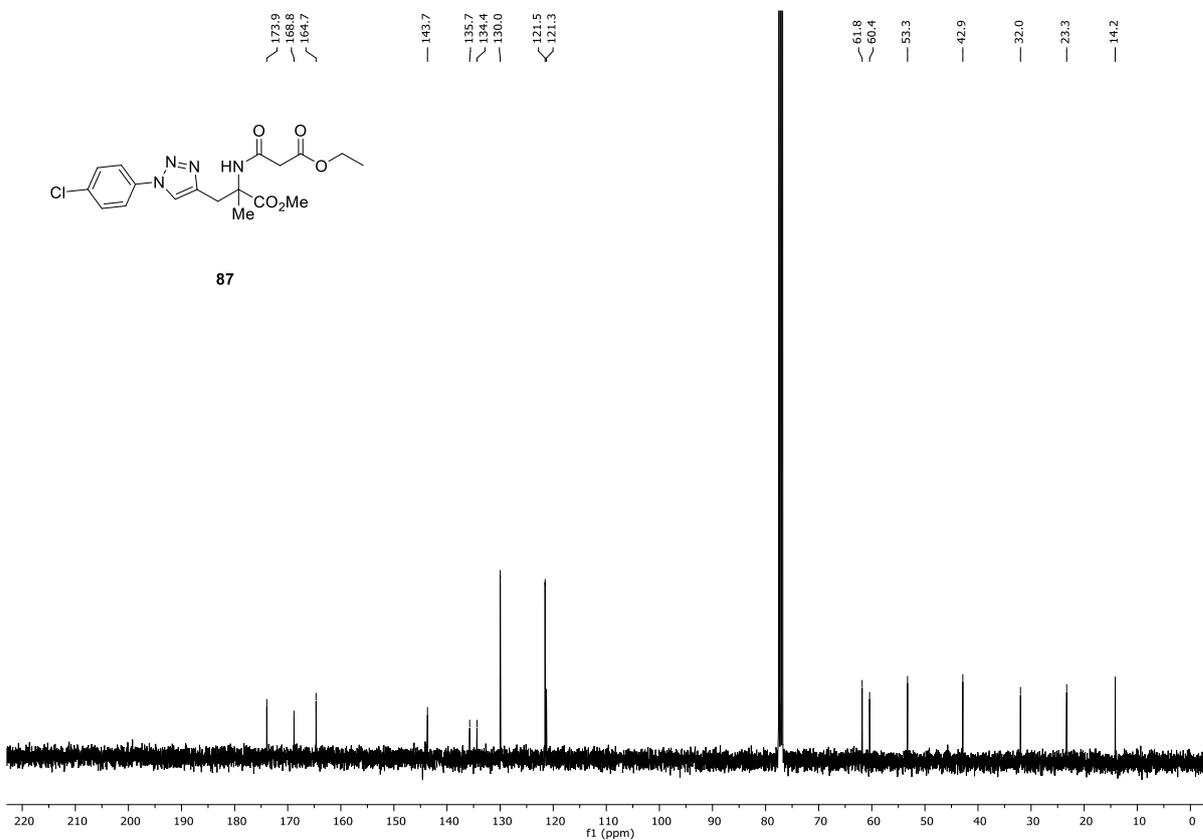
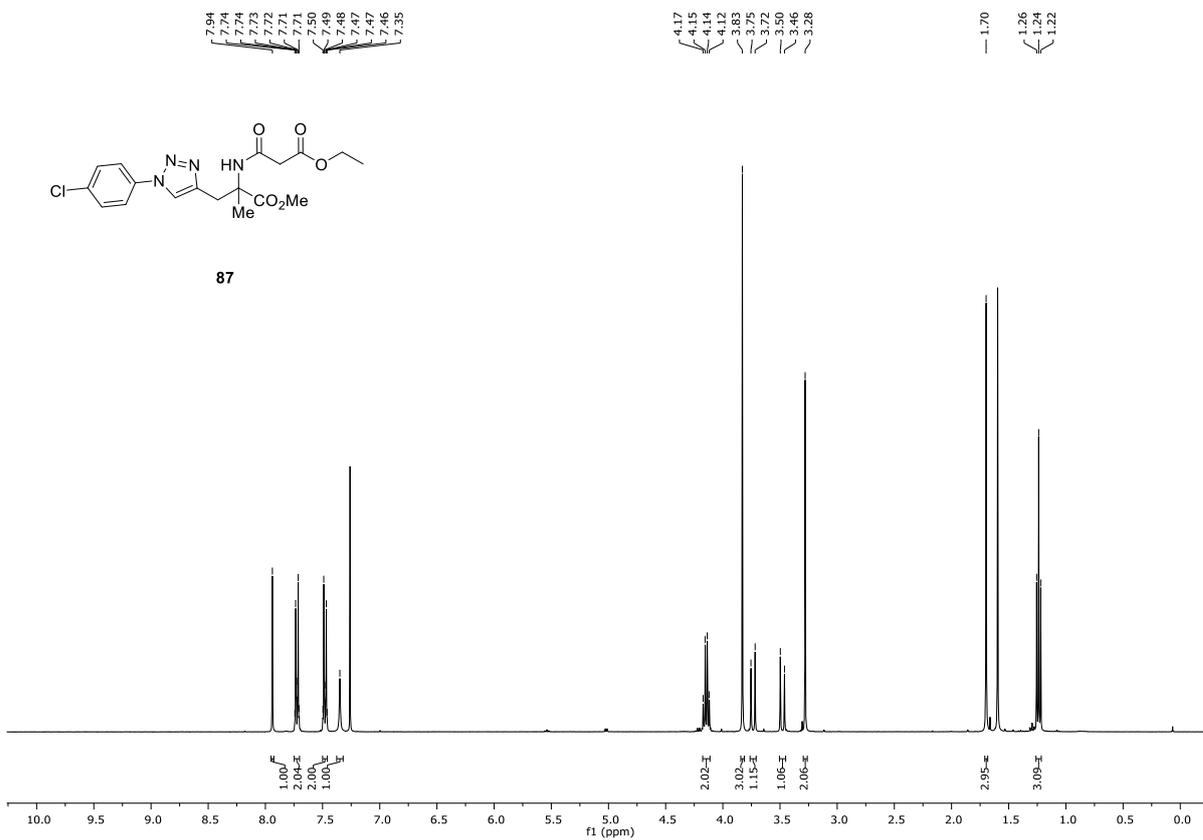


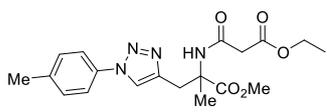


86

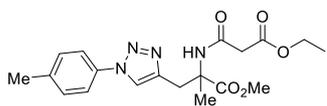
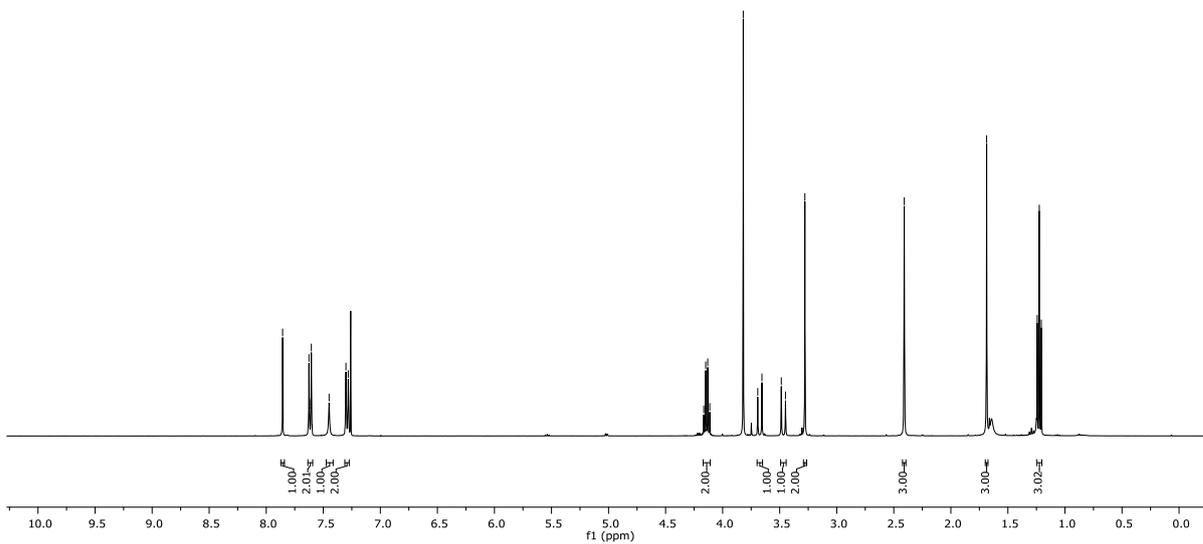
—112.7



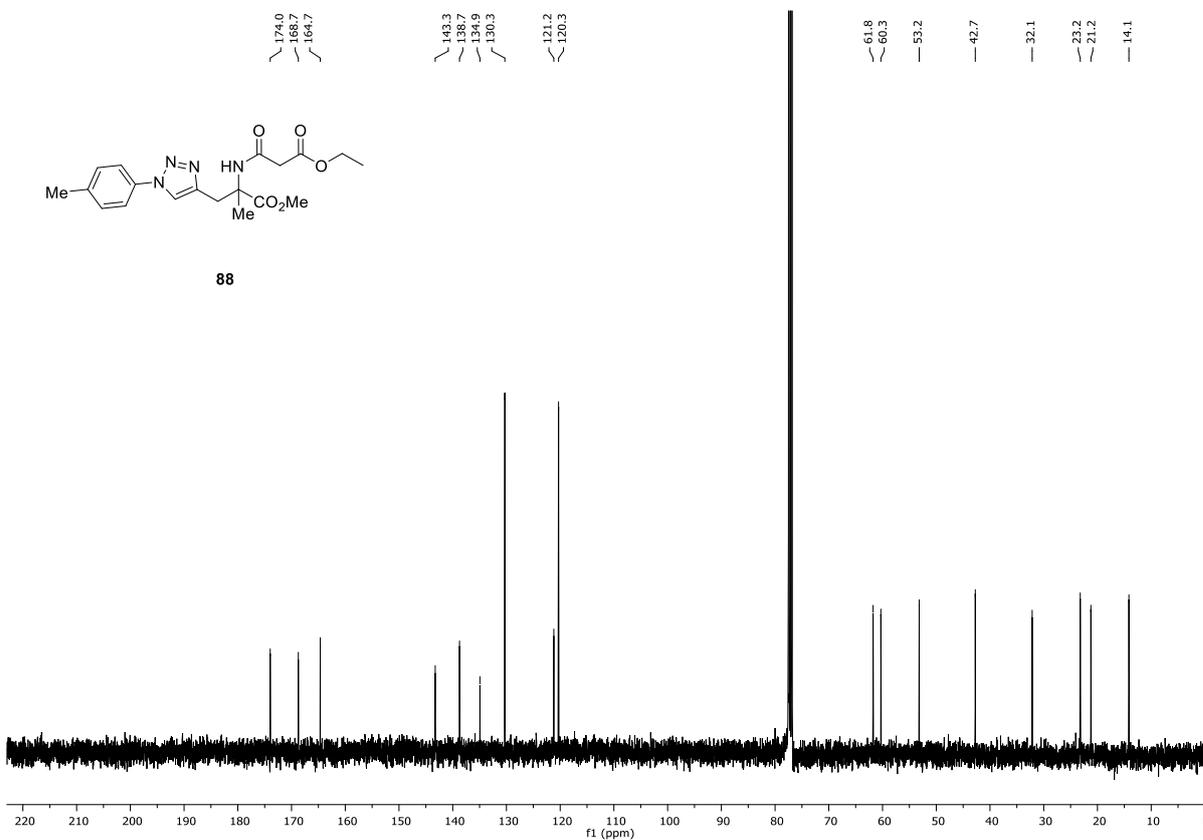


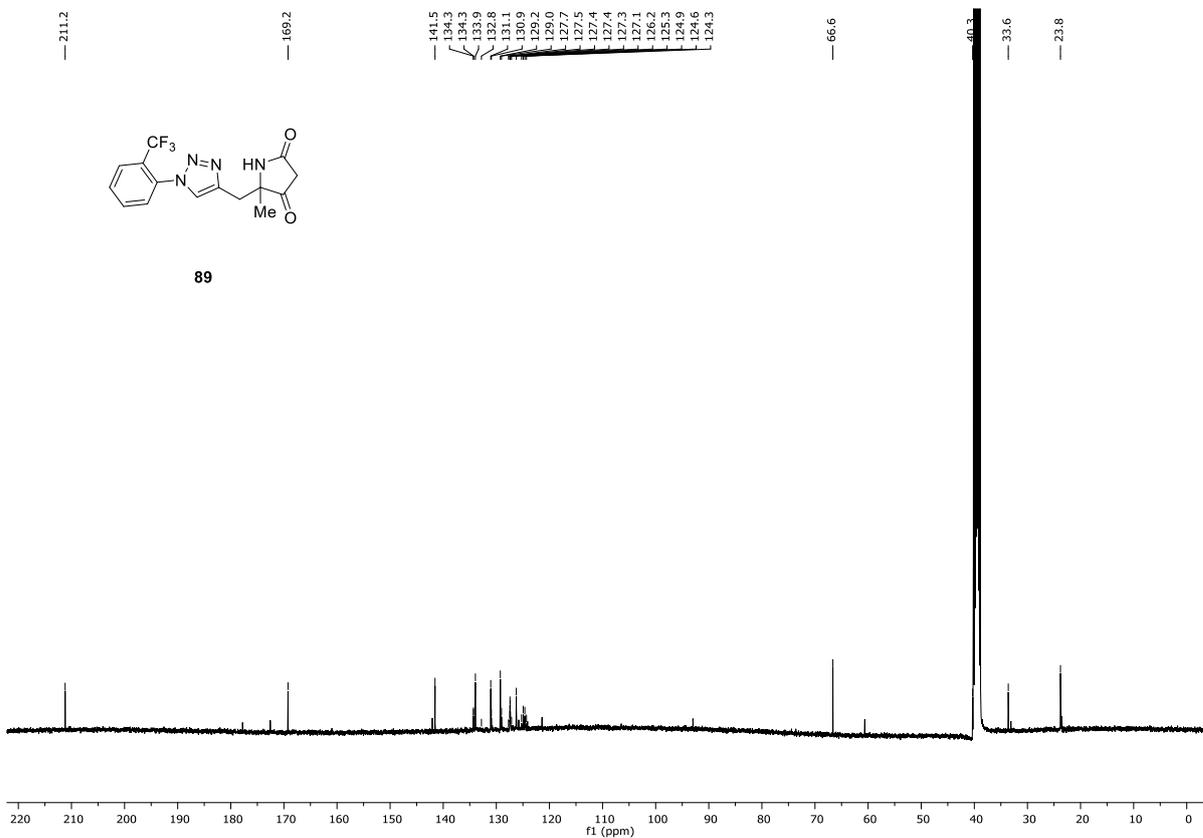
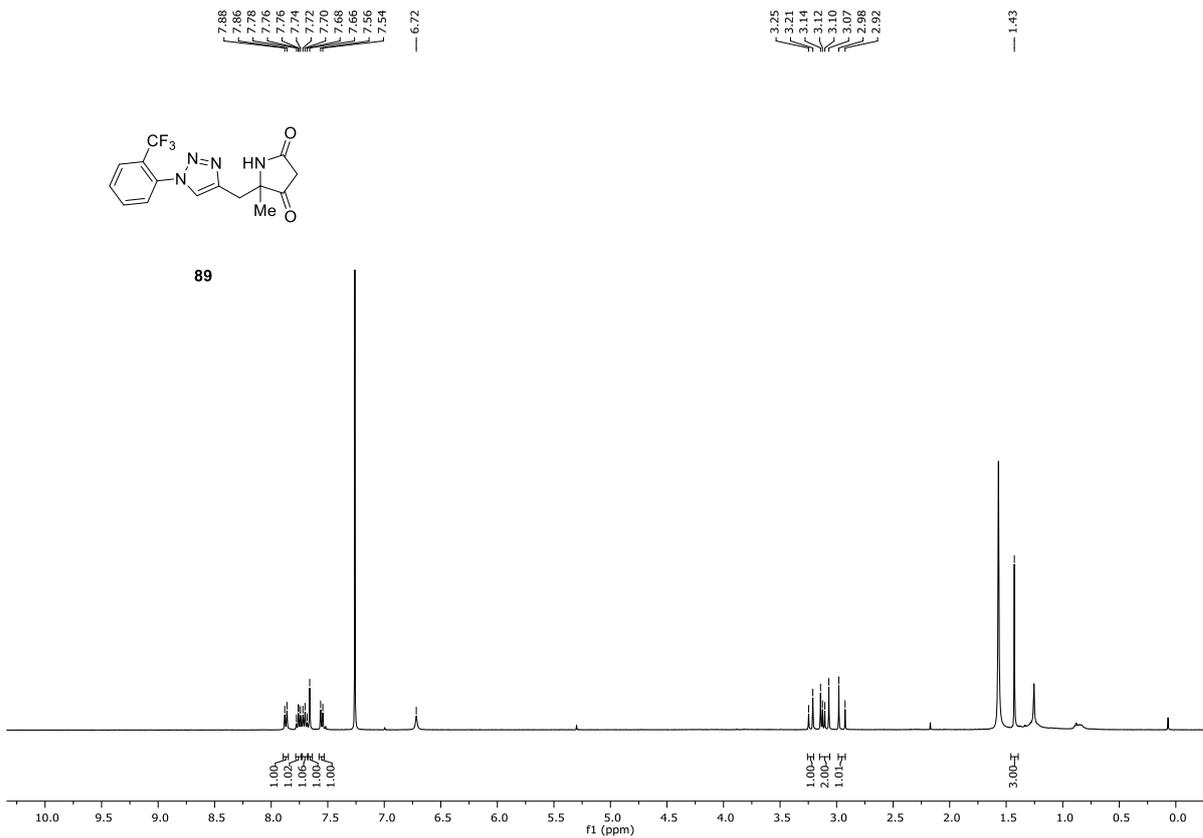


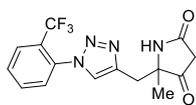
88



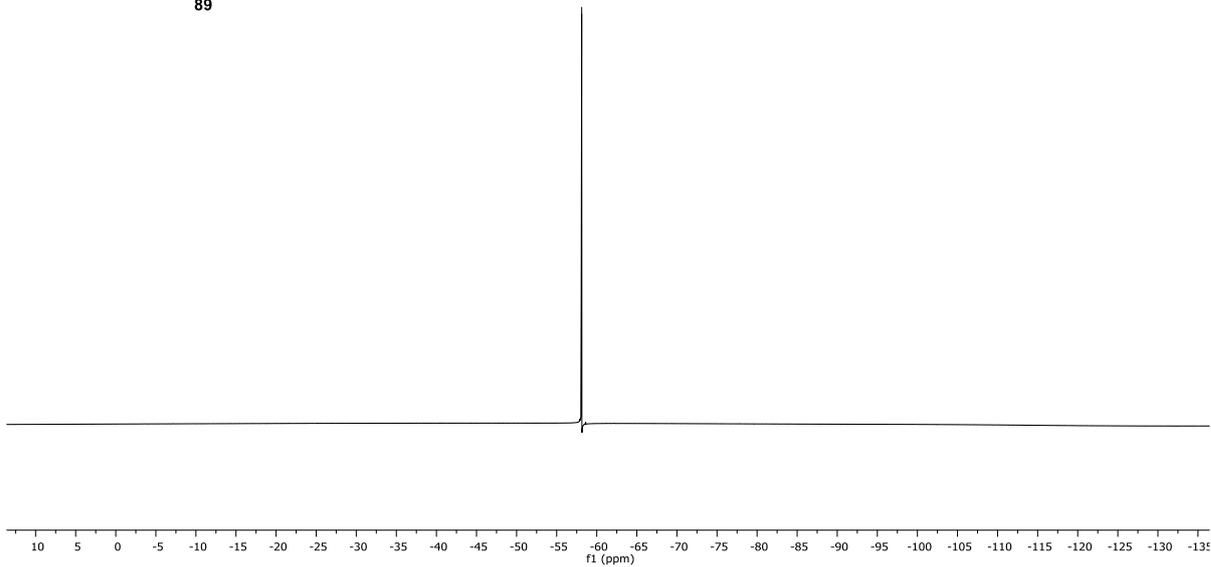
88

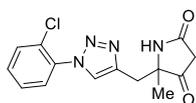




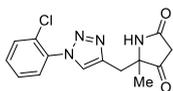
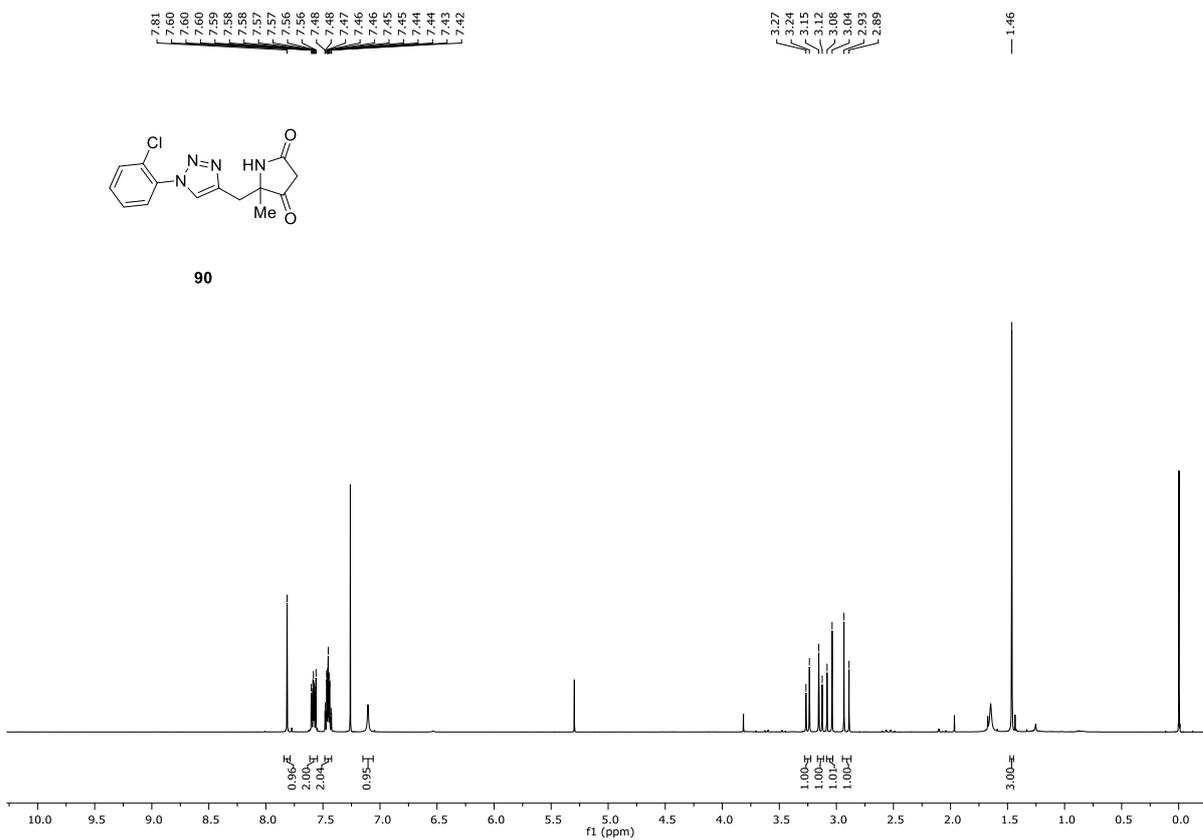


89

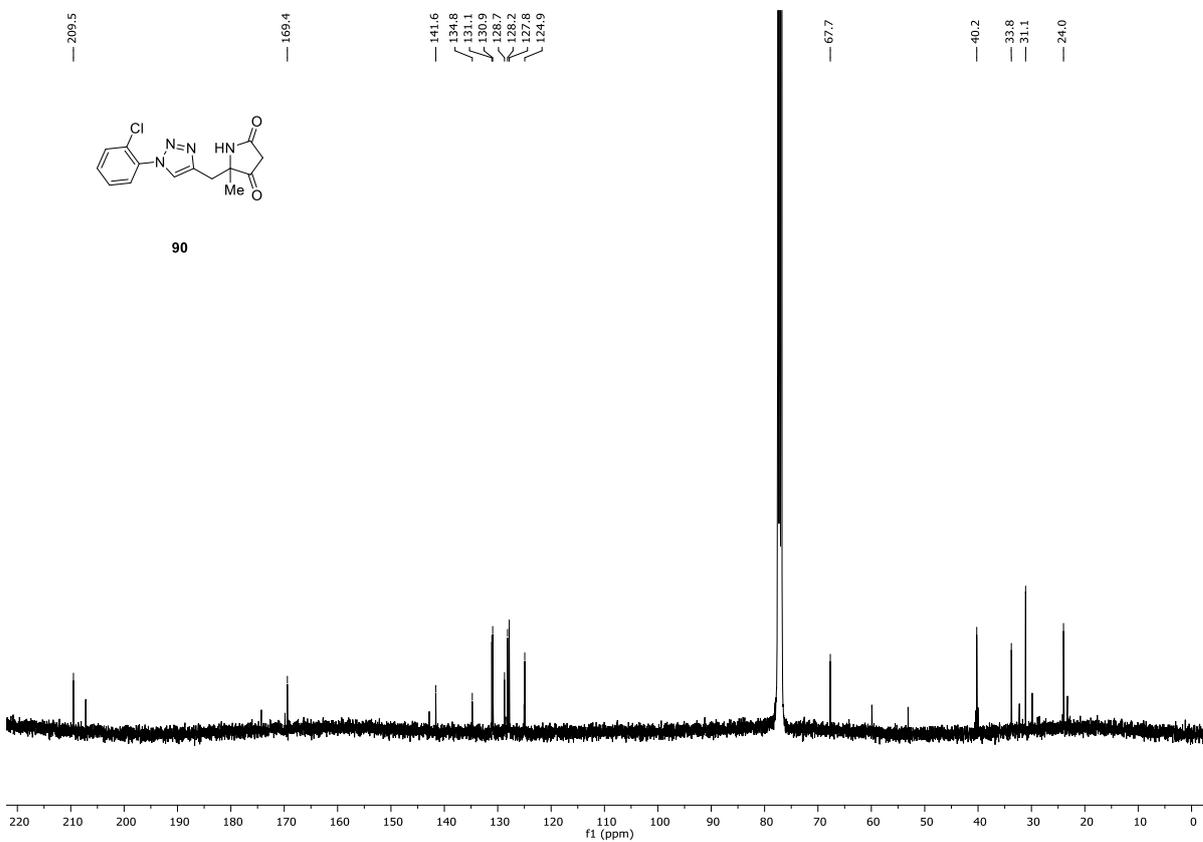


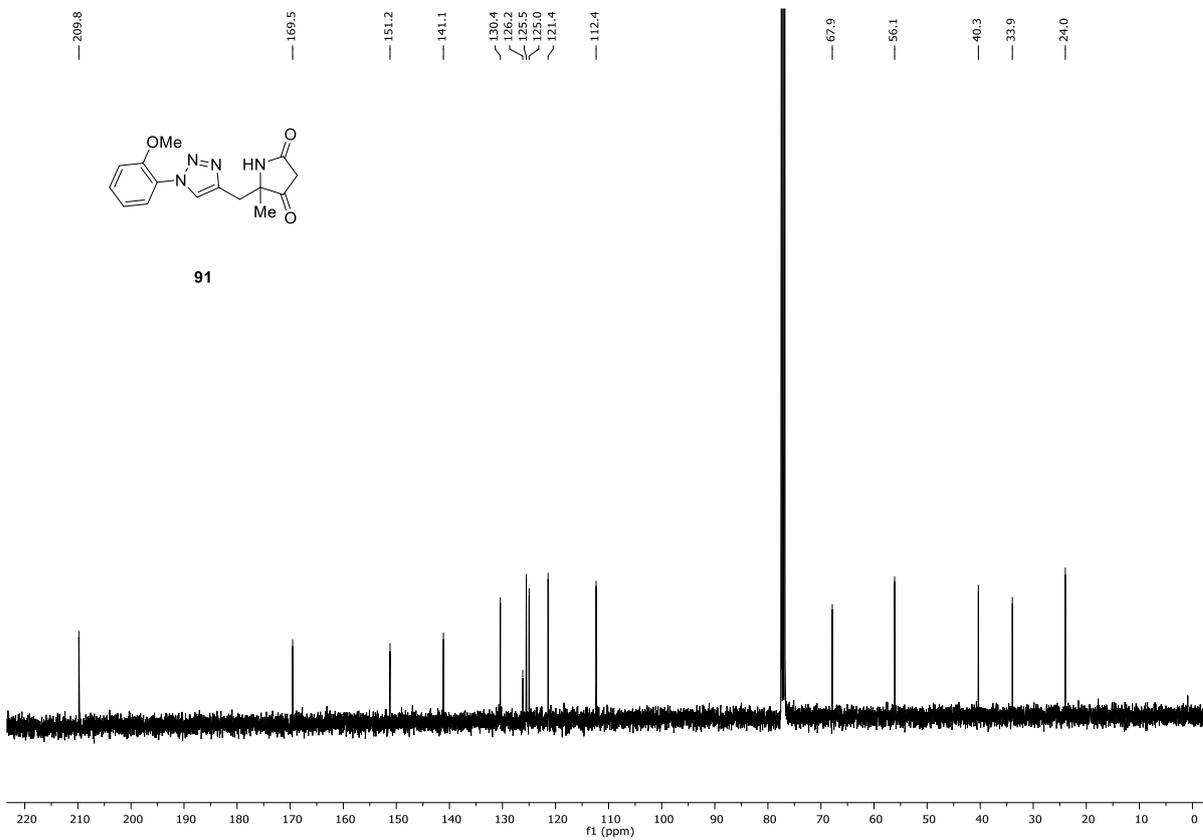
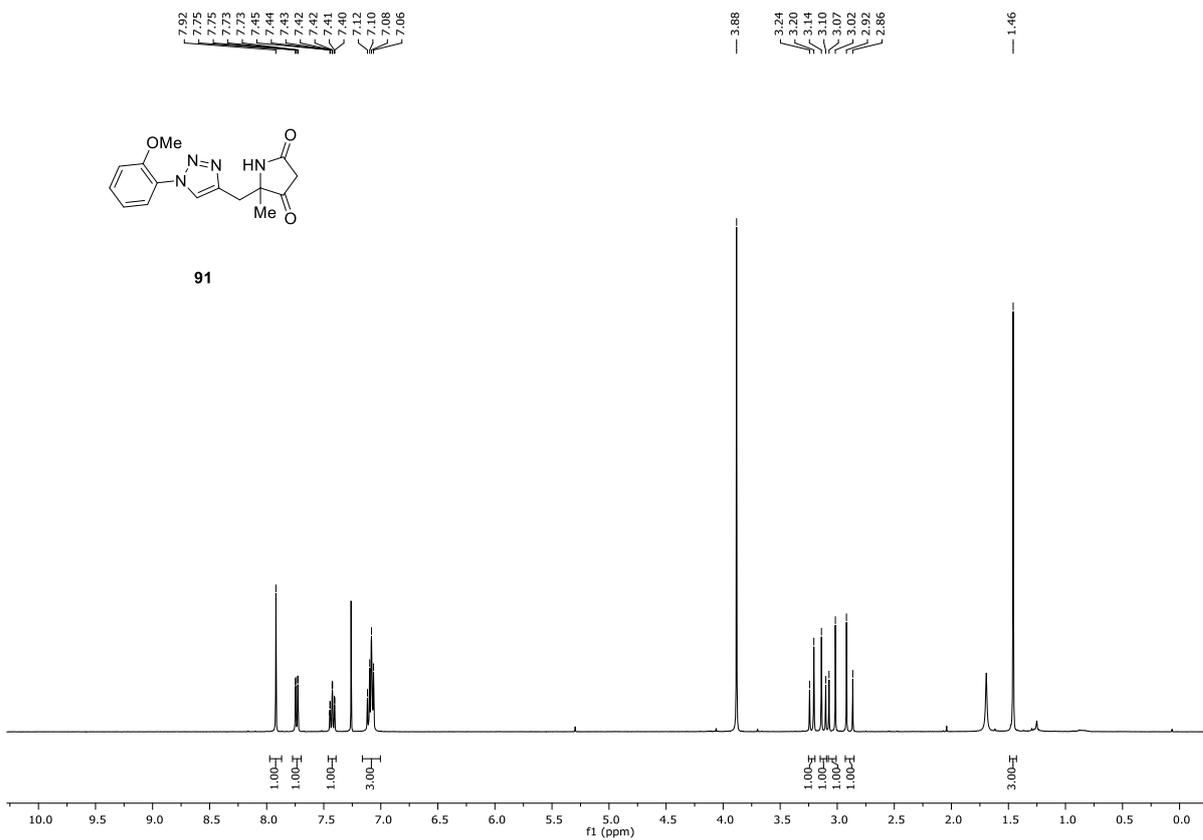


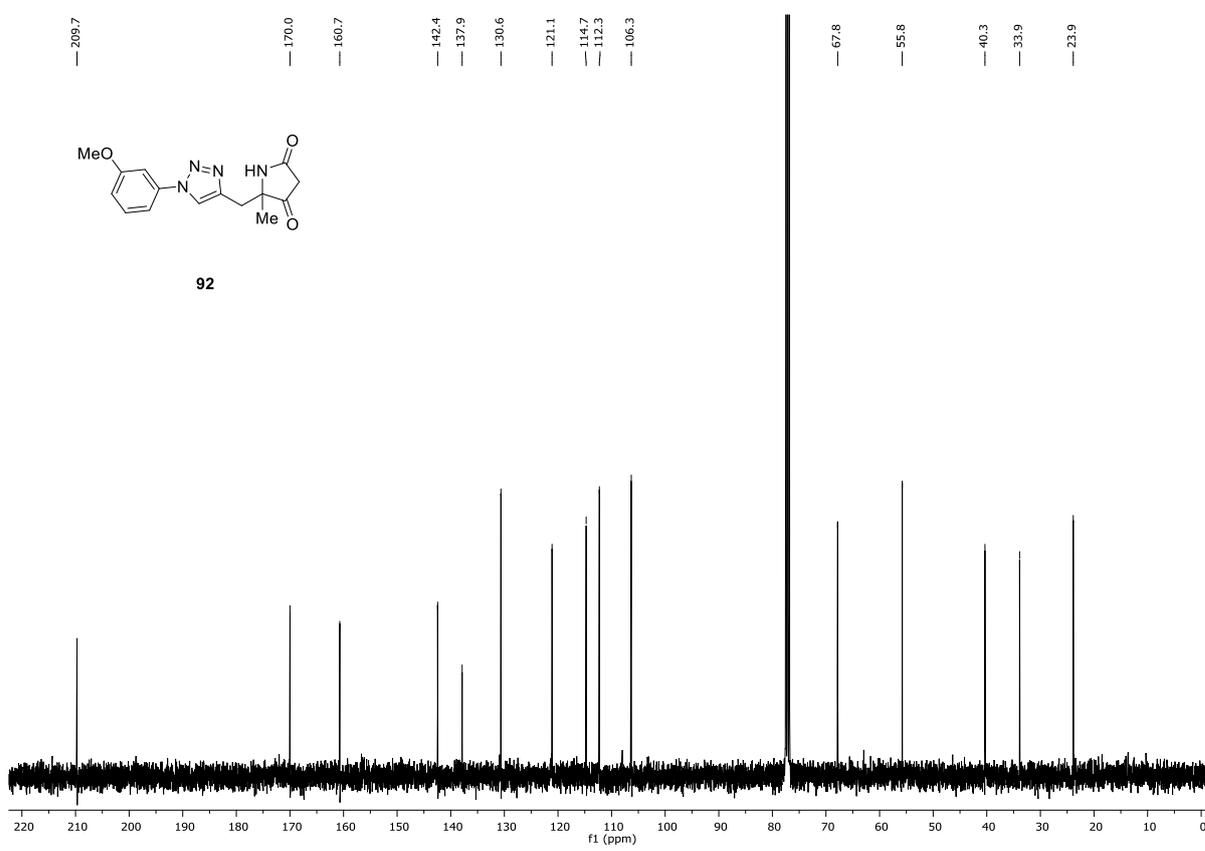
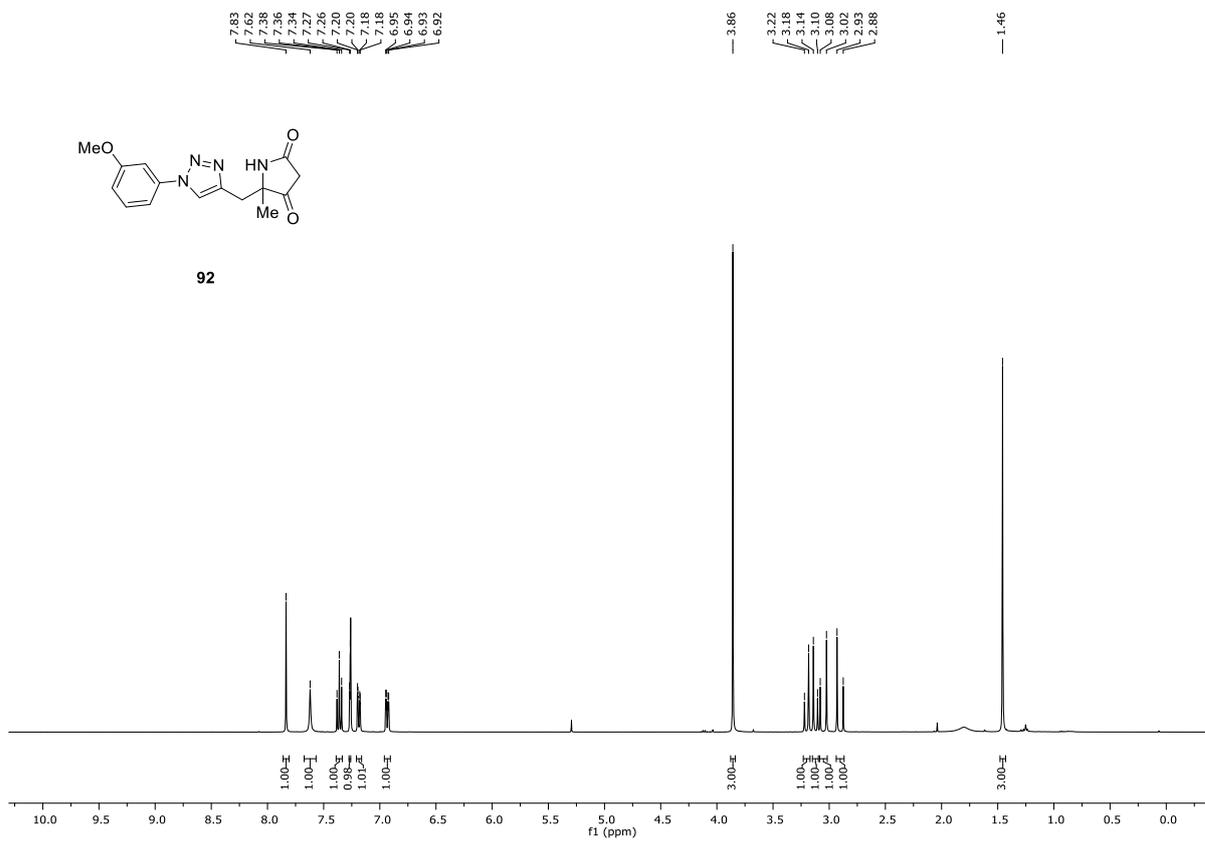
90

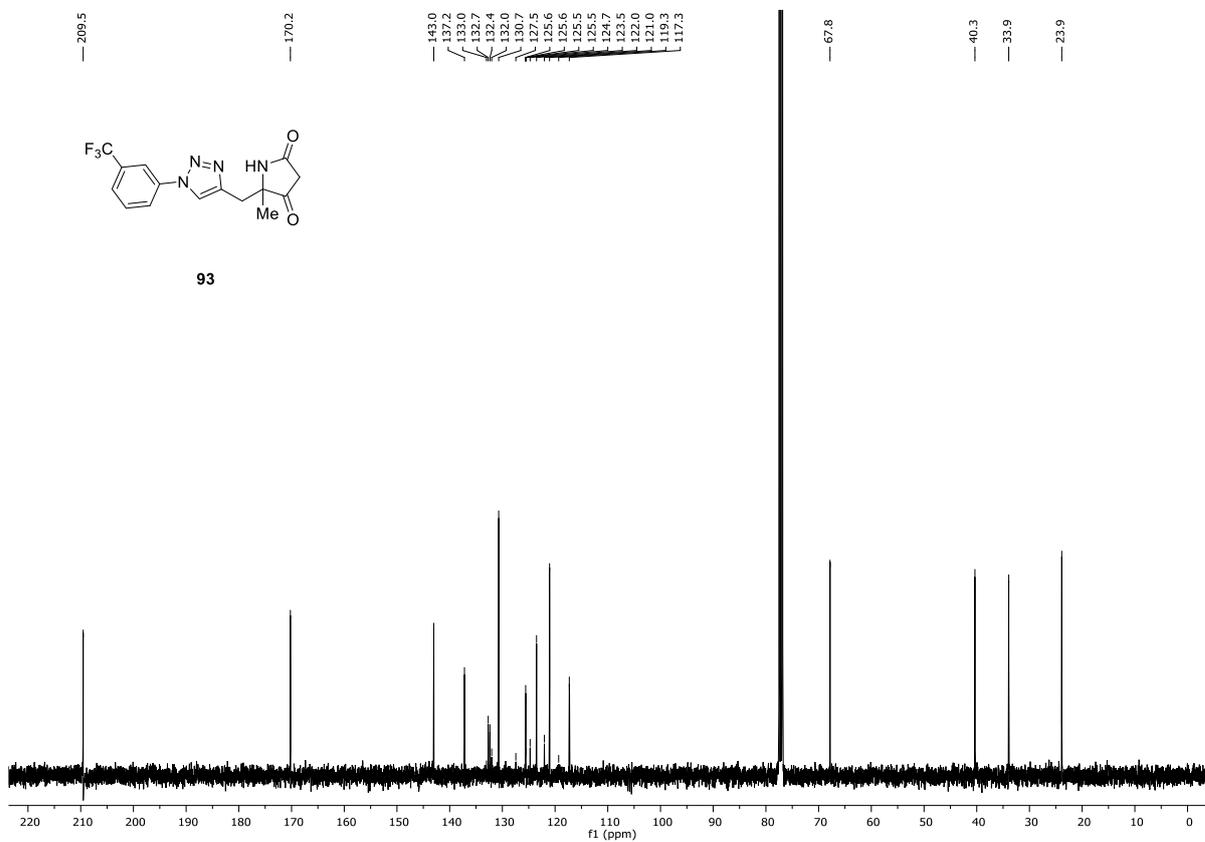
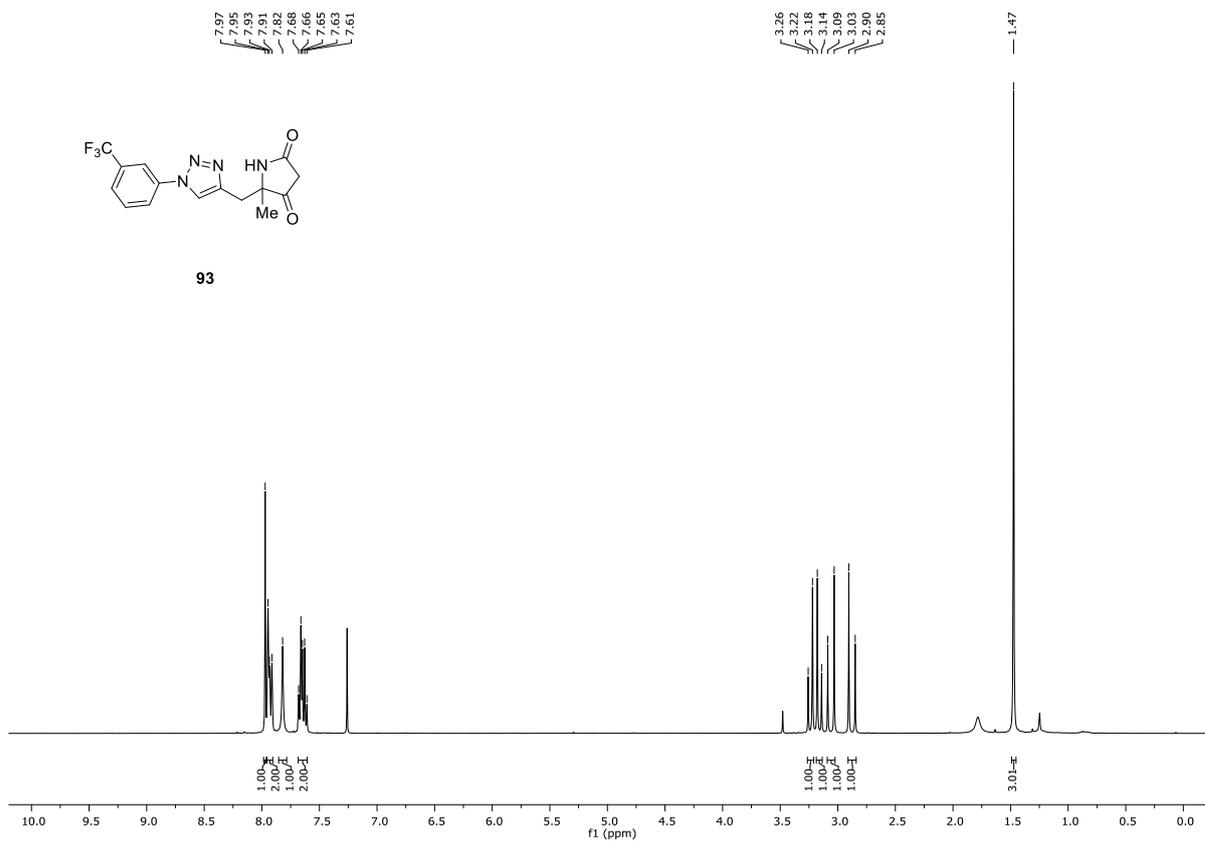


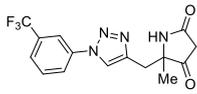
90



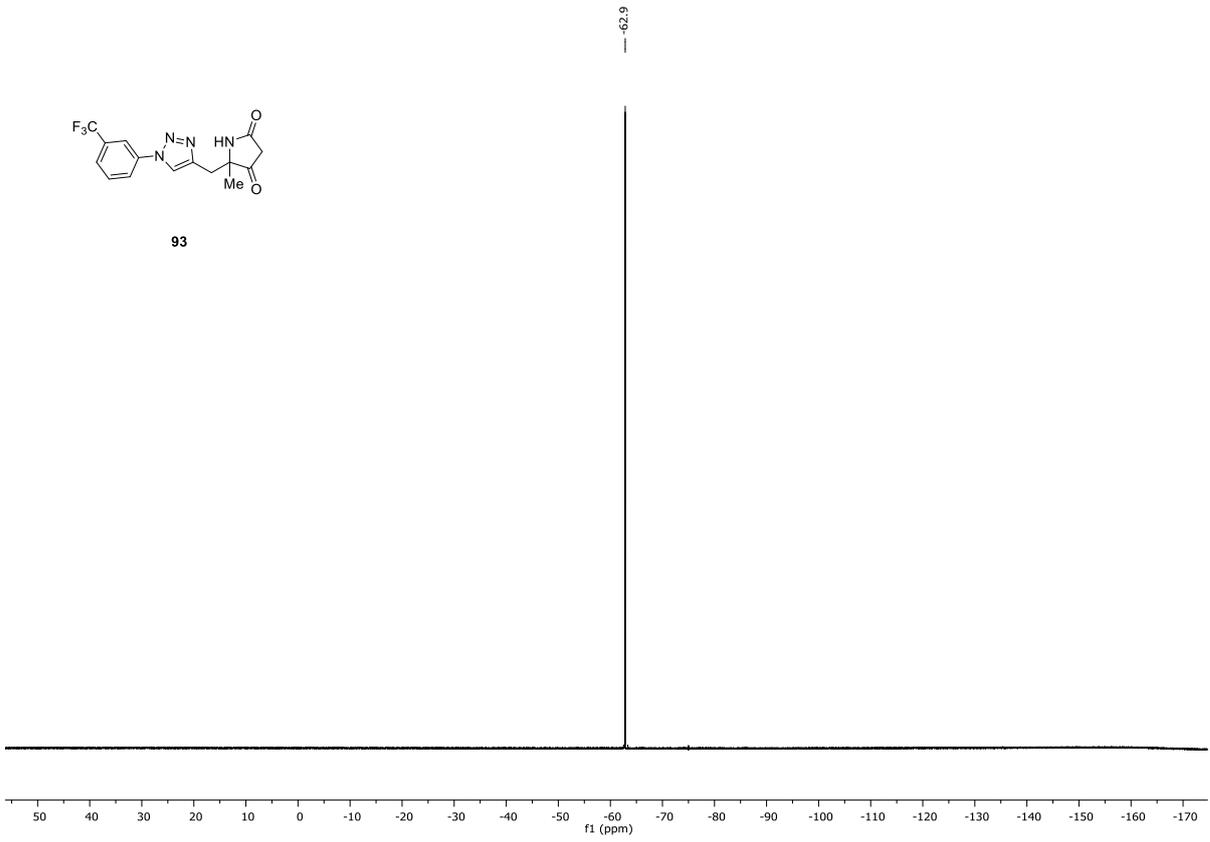


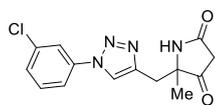




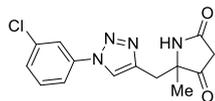
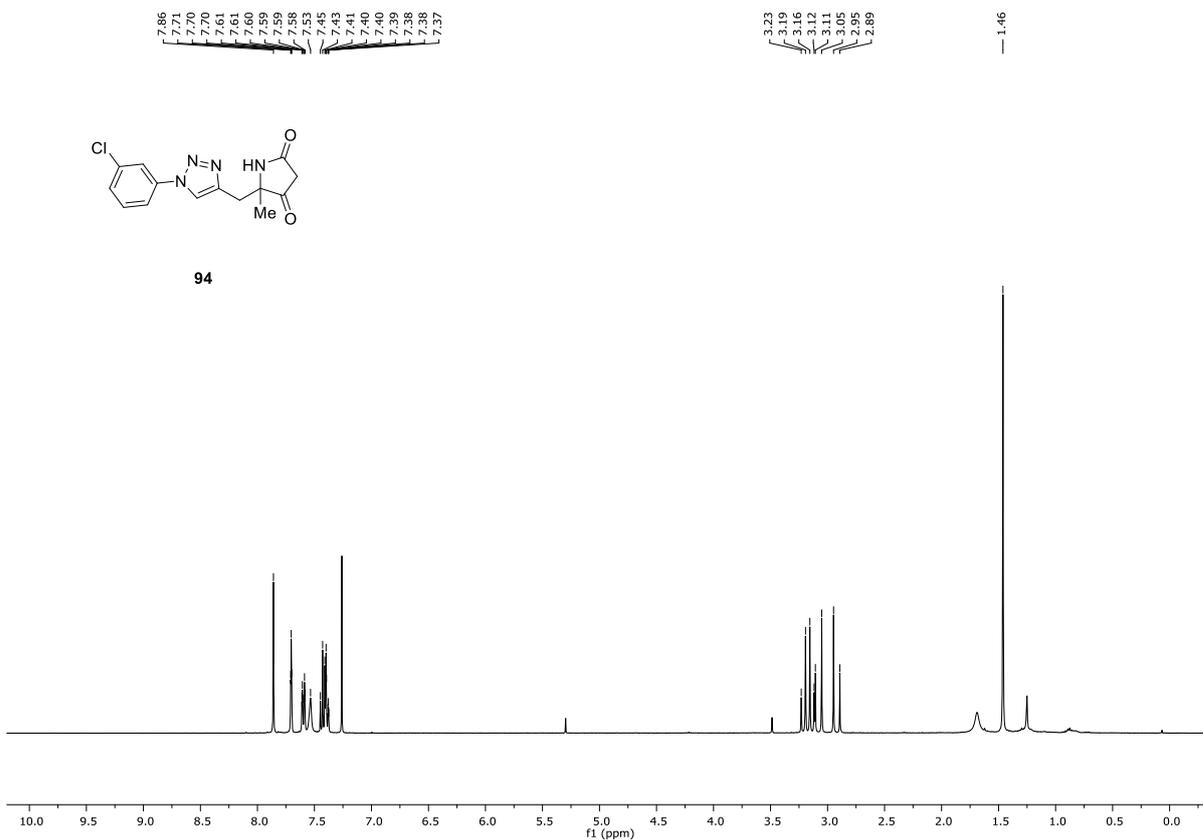


93

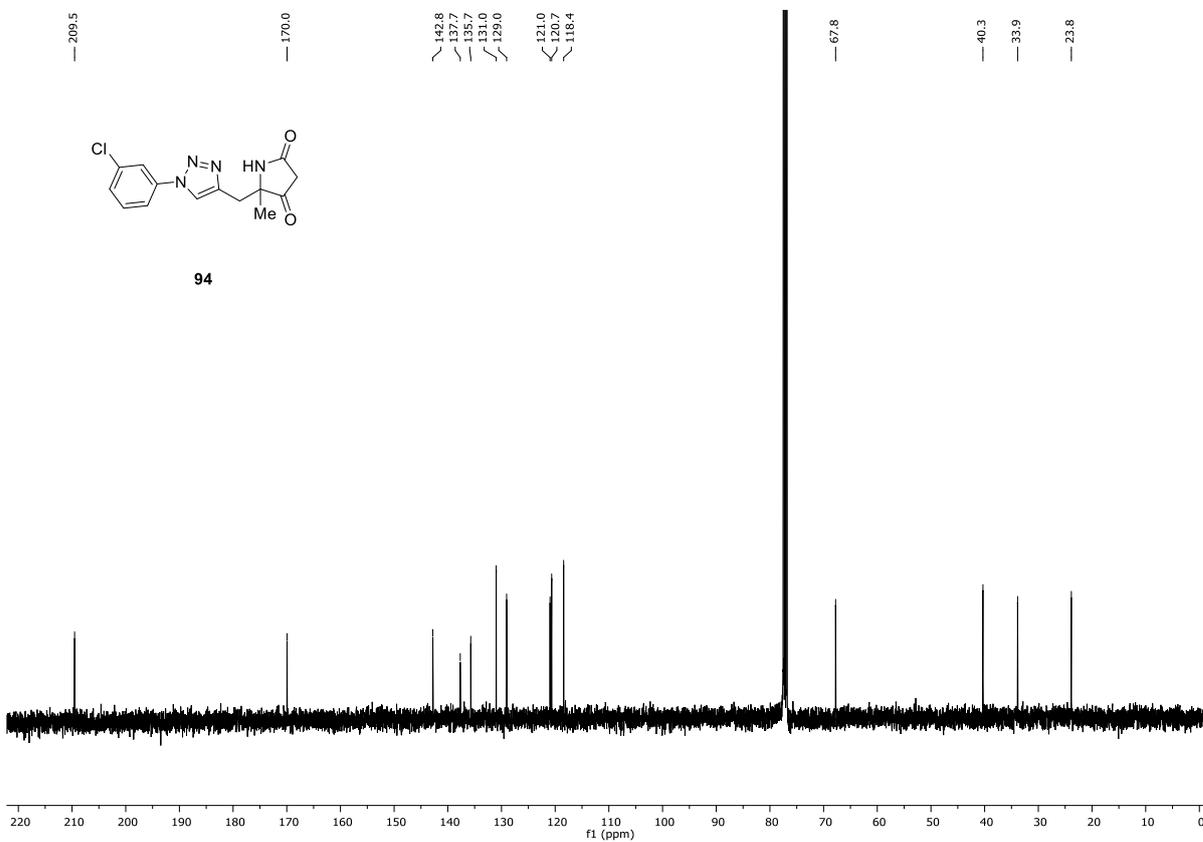


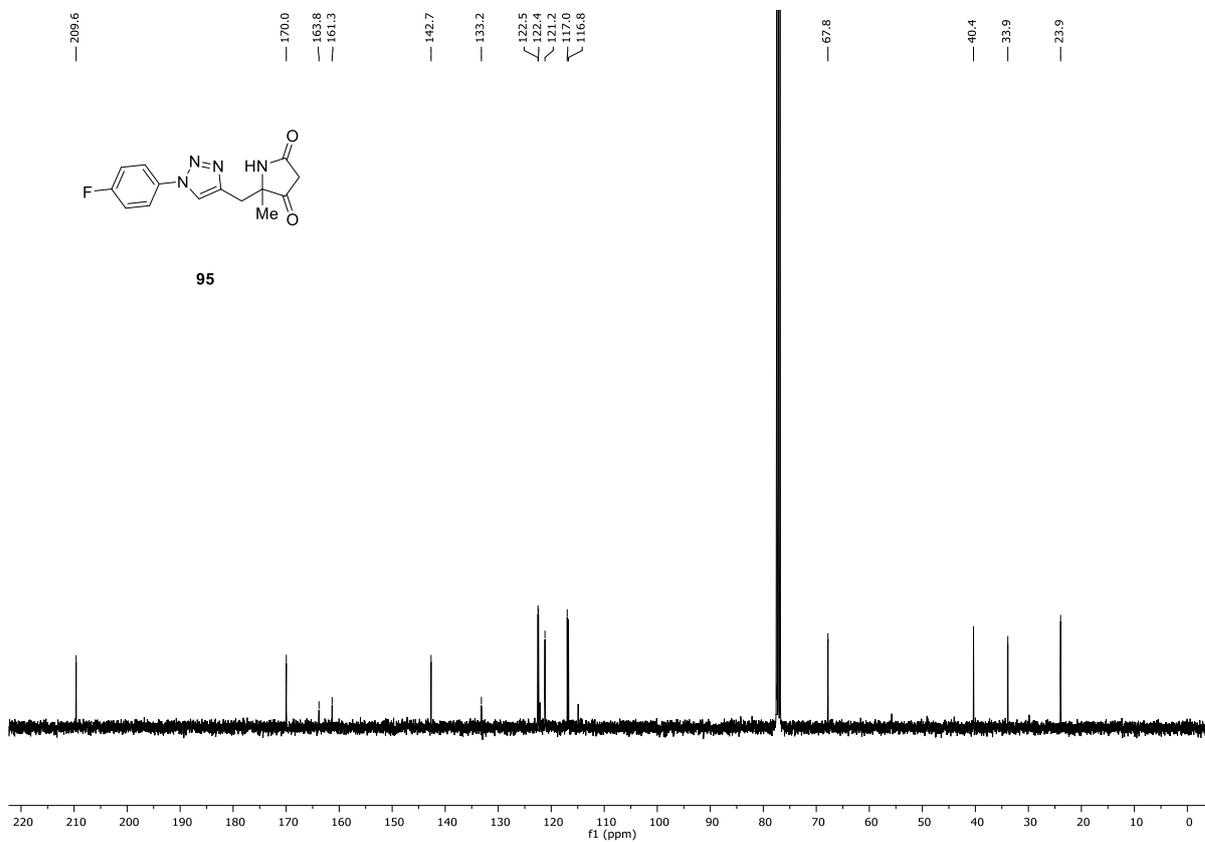
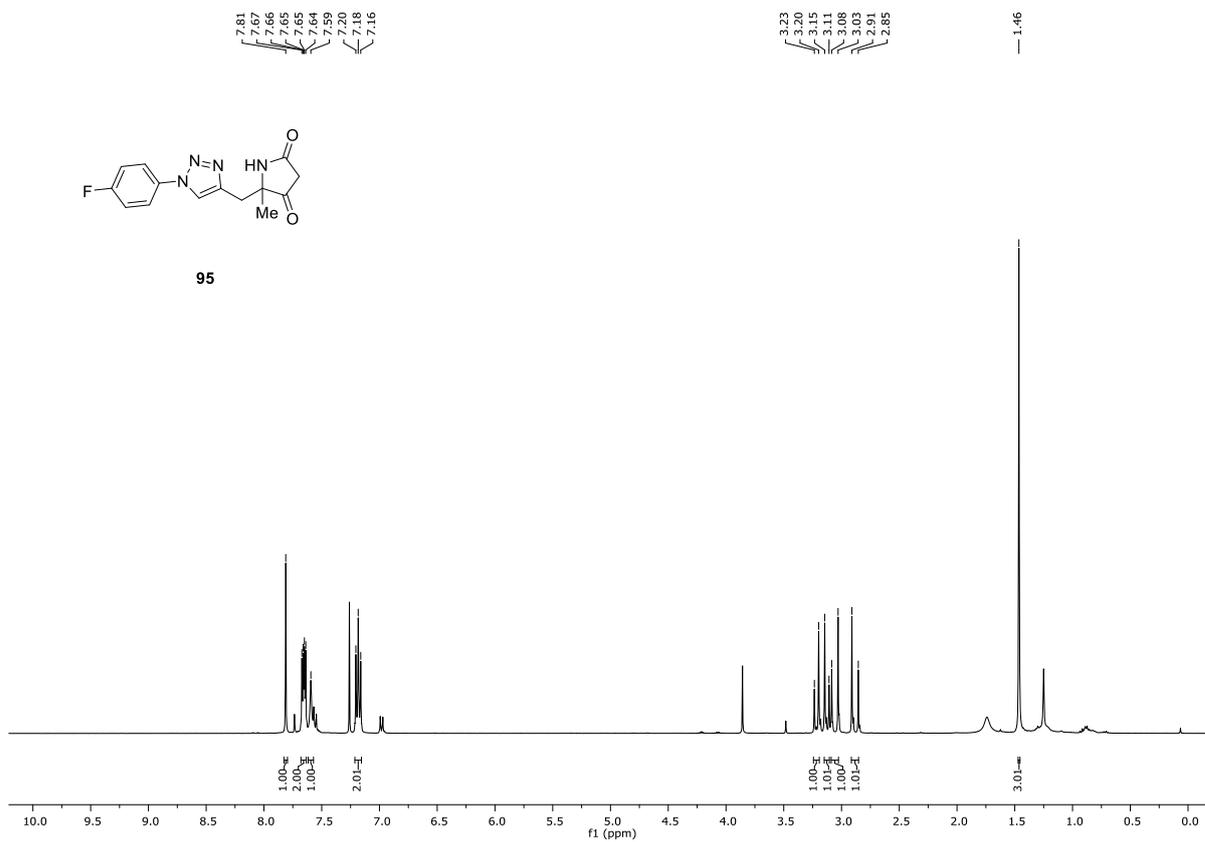


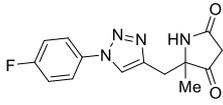
94



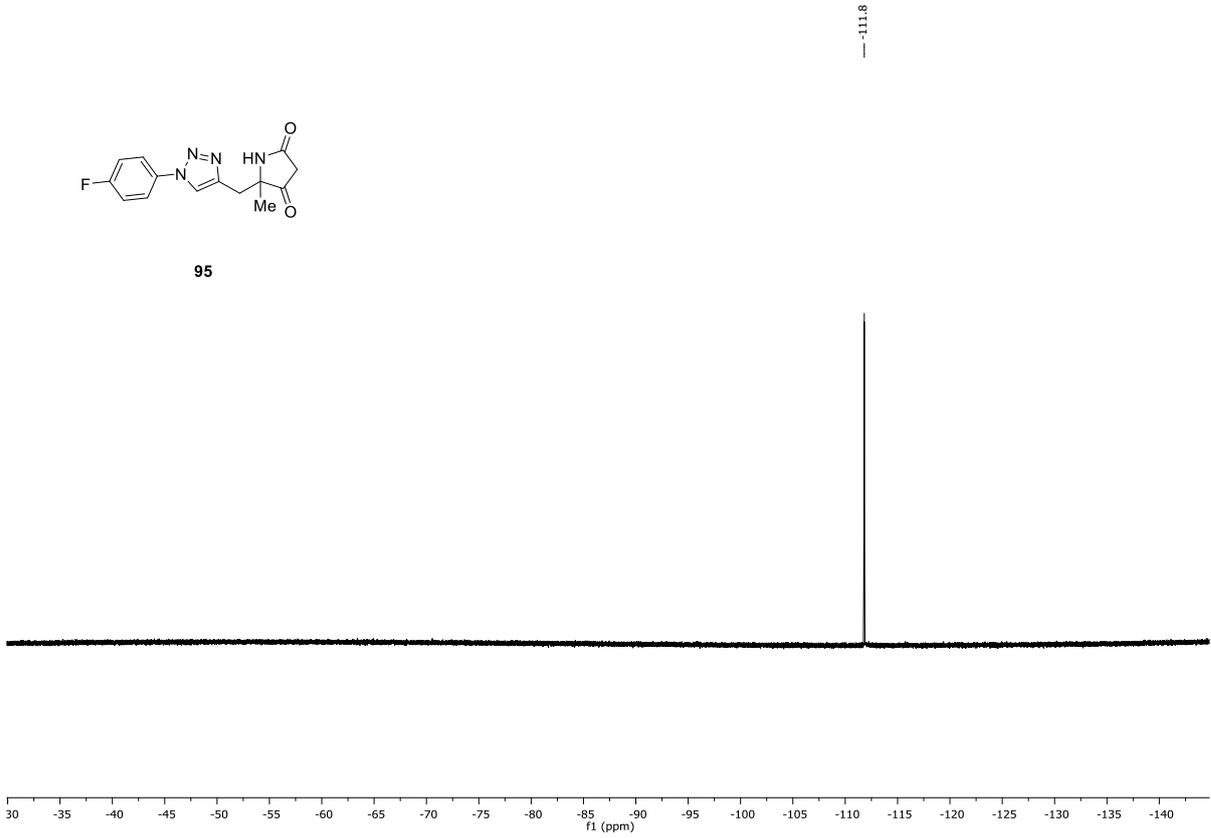
94

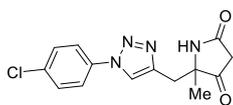




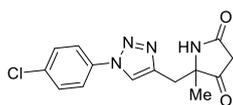
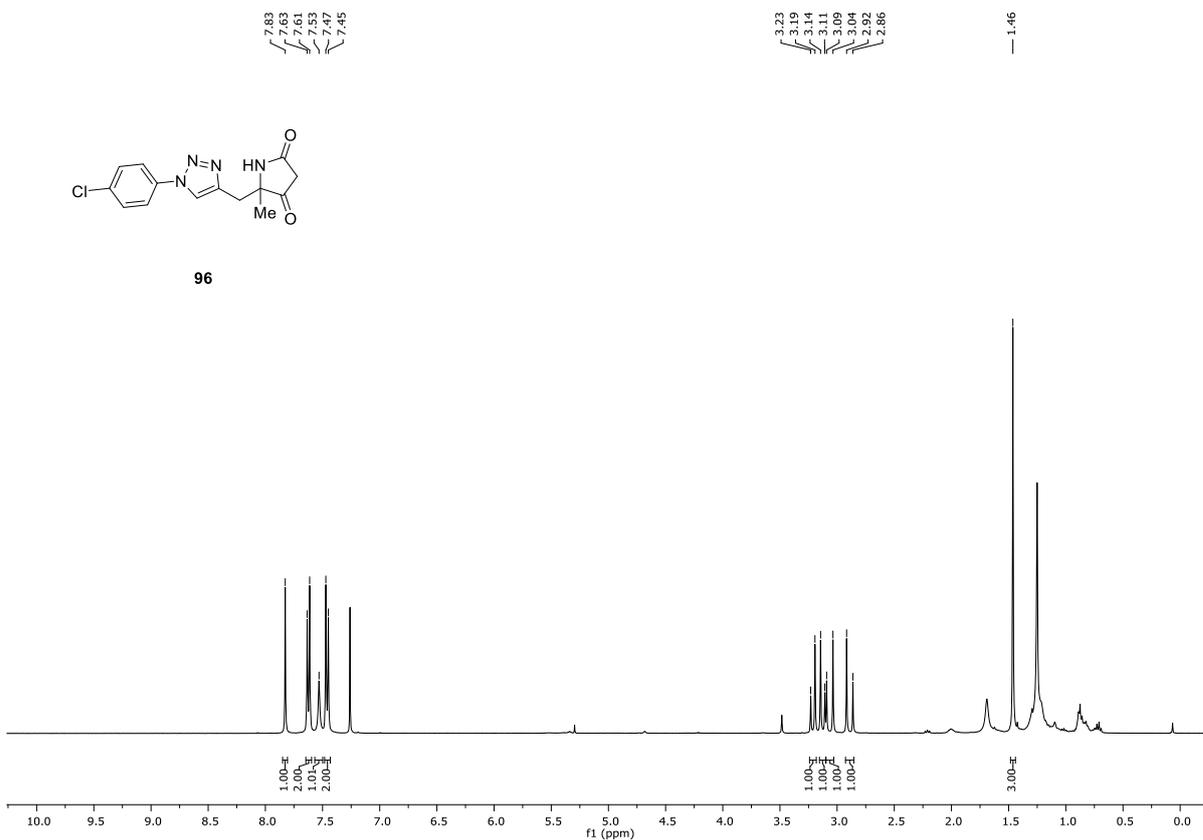


95

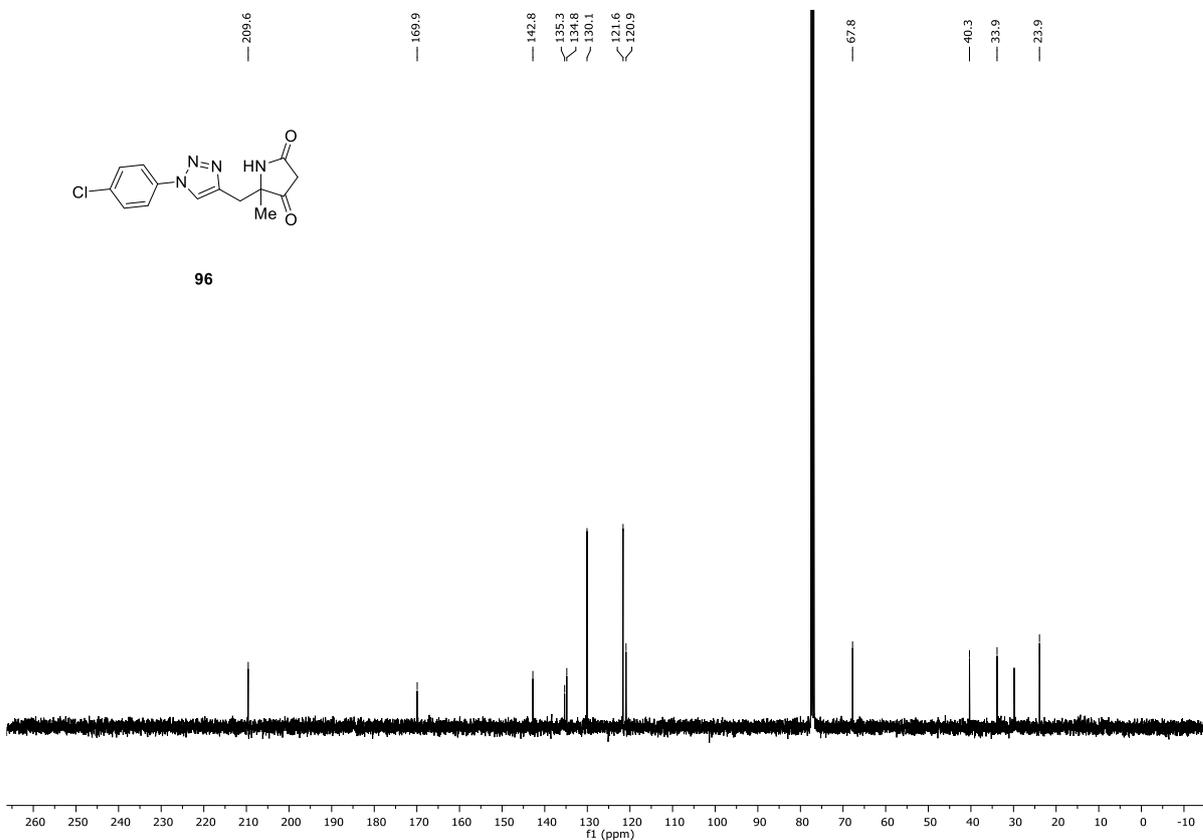


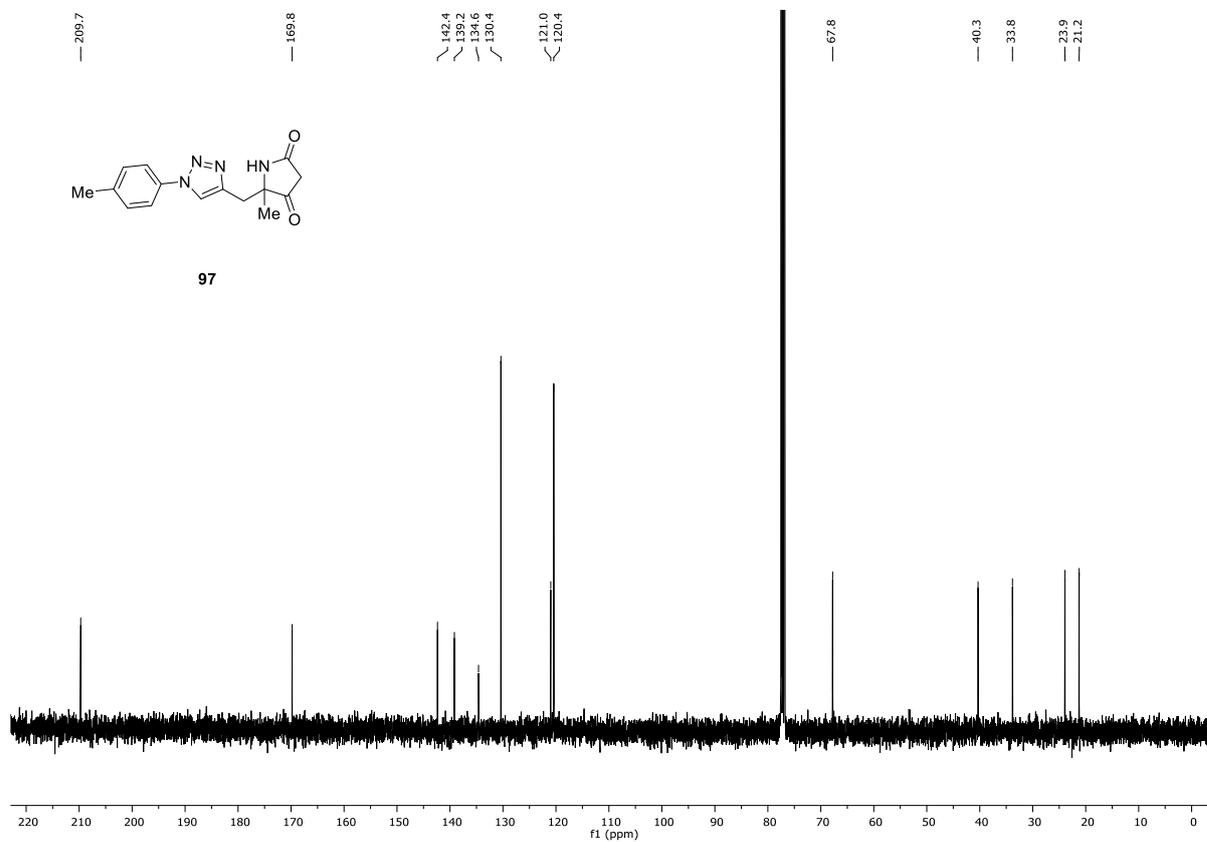
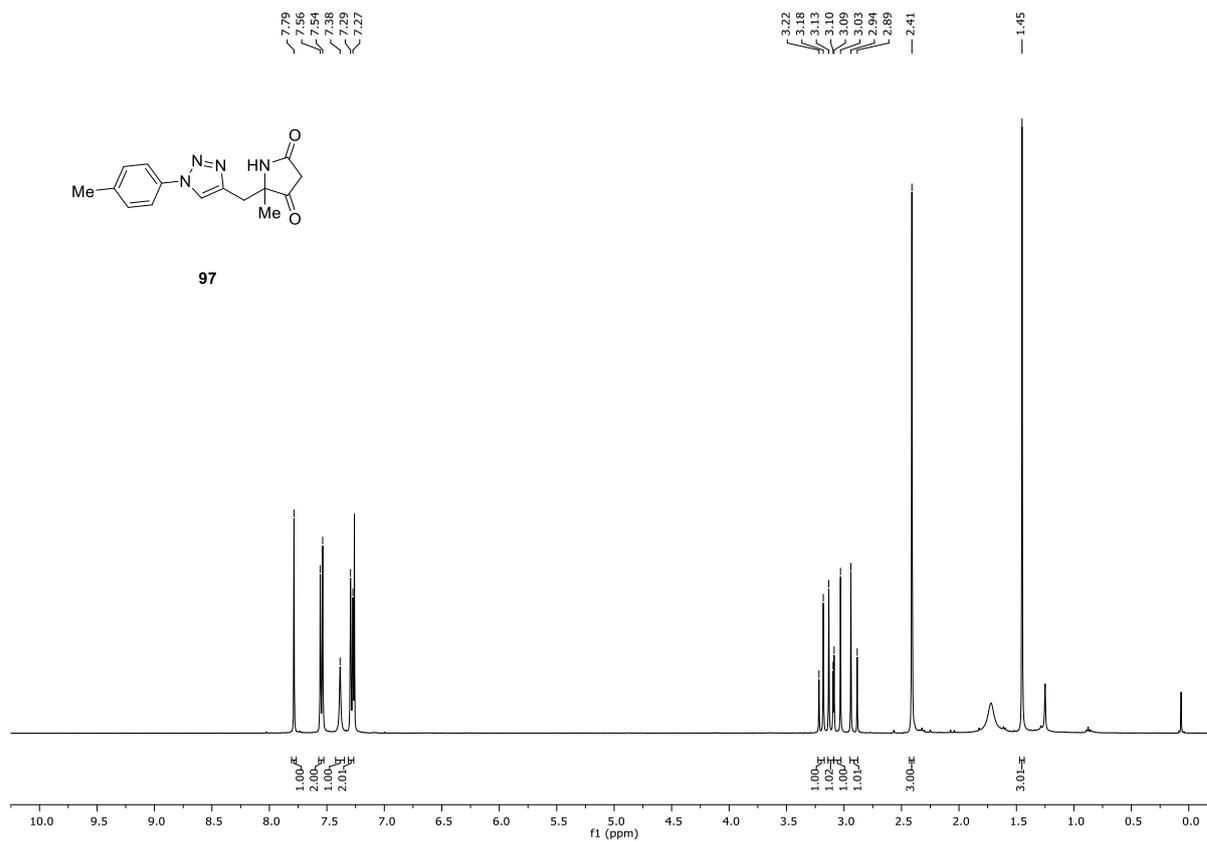


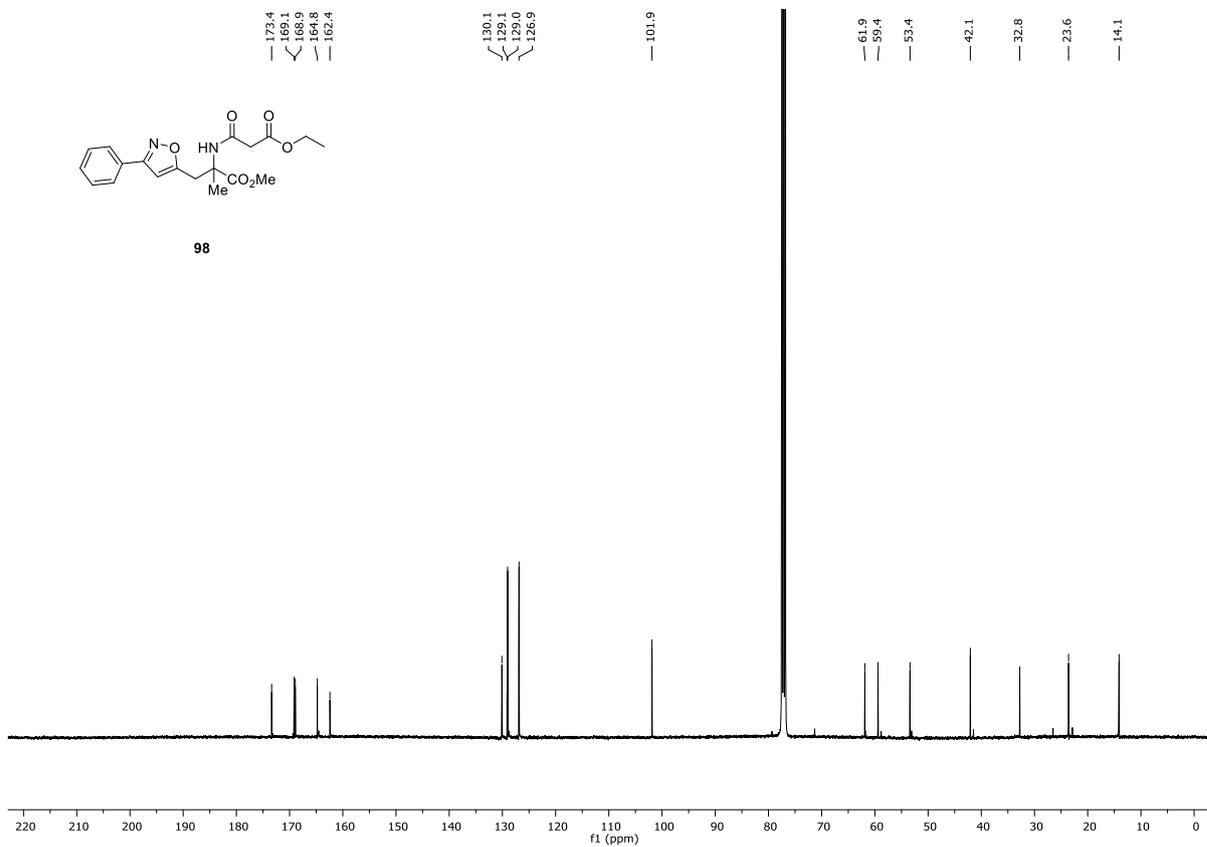
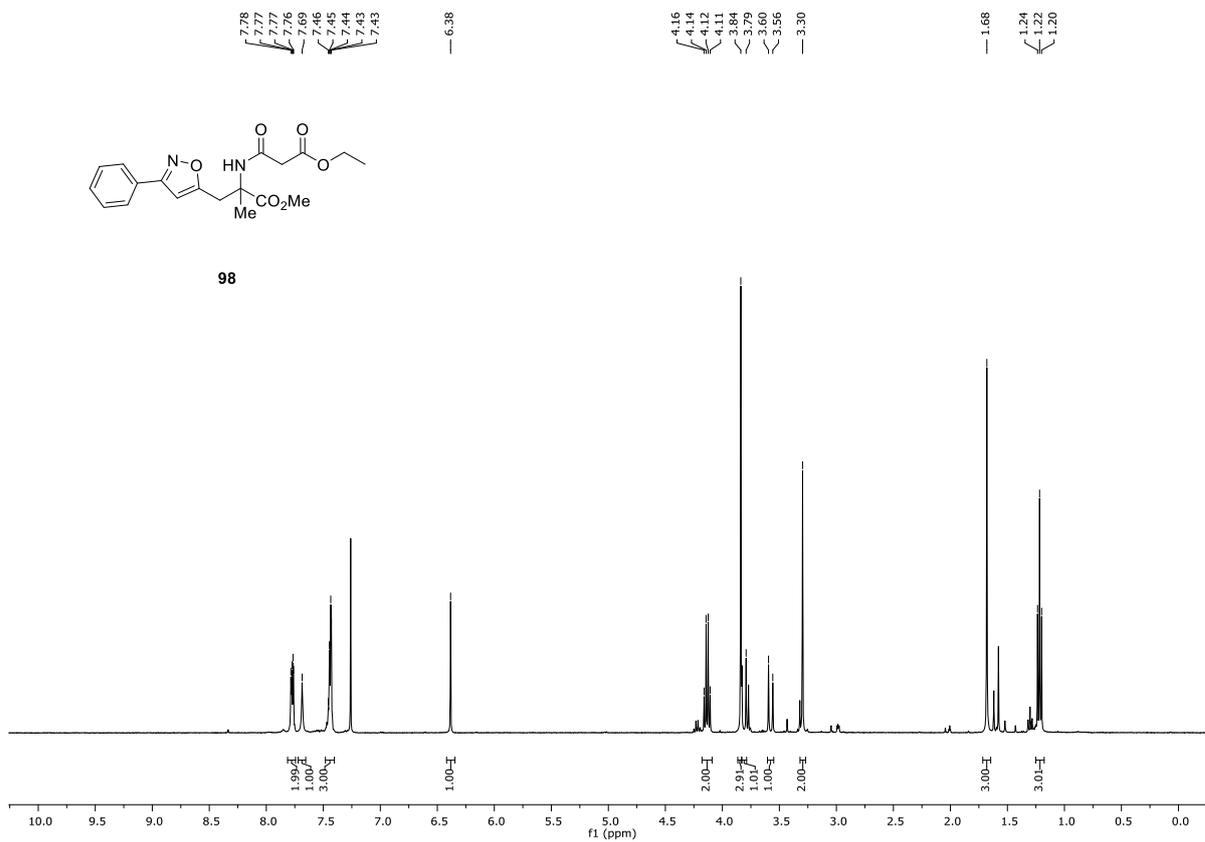
96

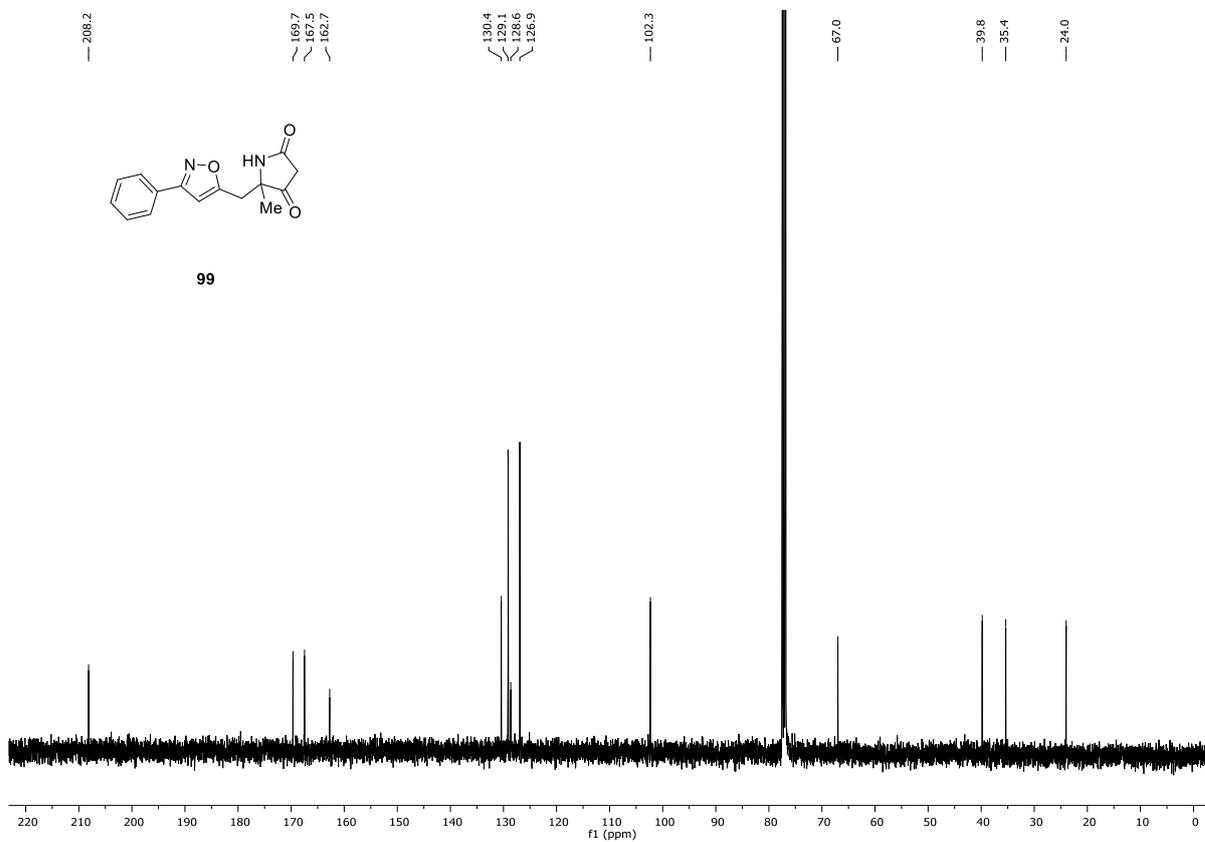
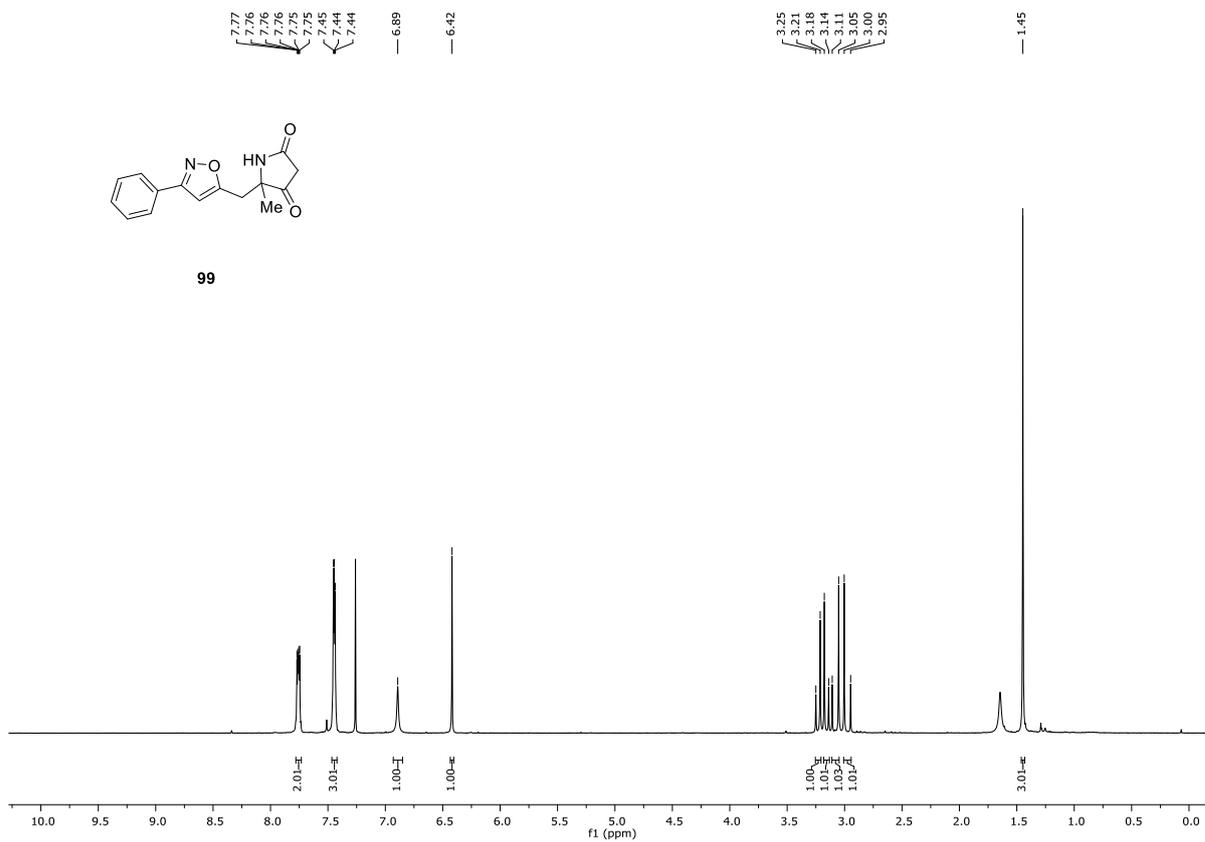


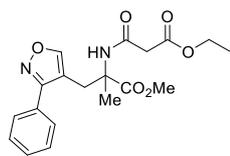
96



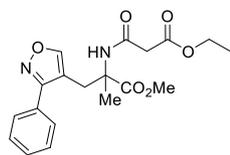
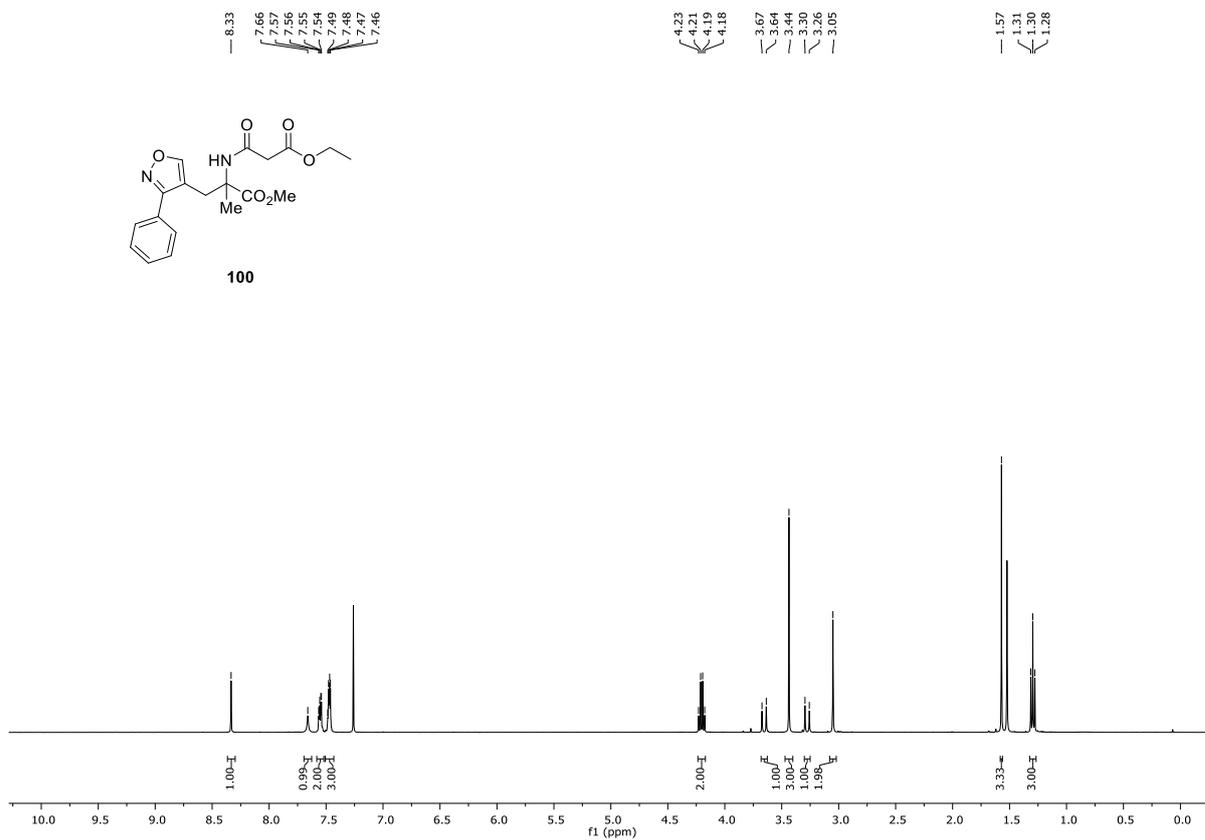




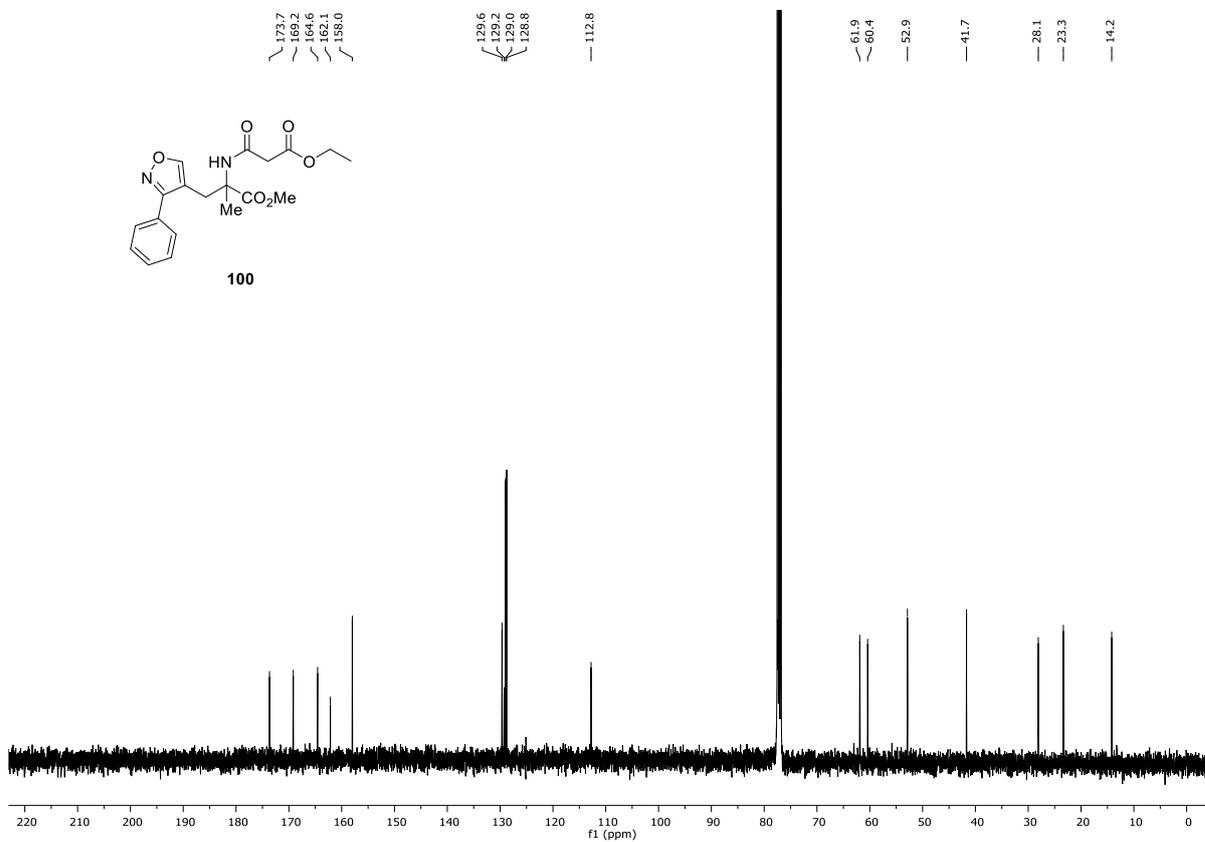


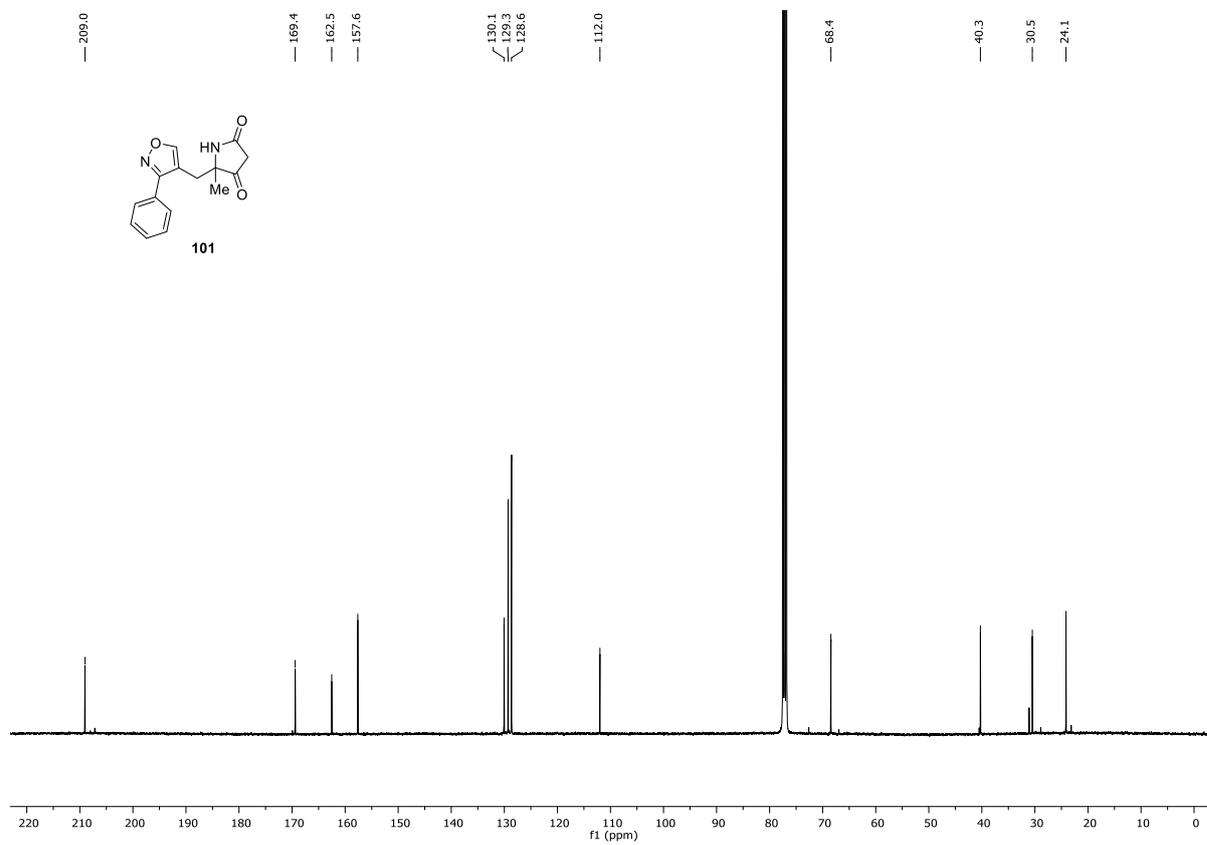
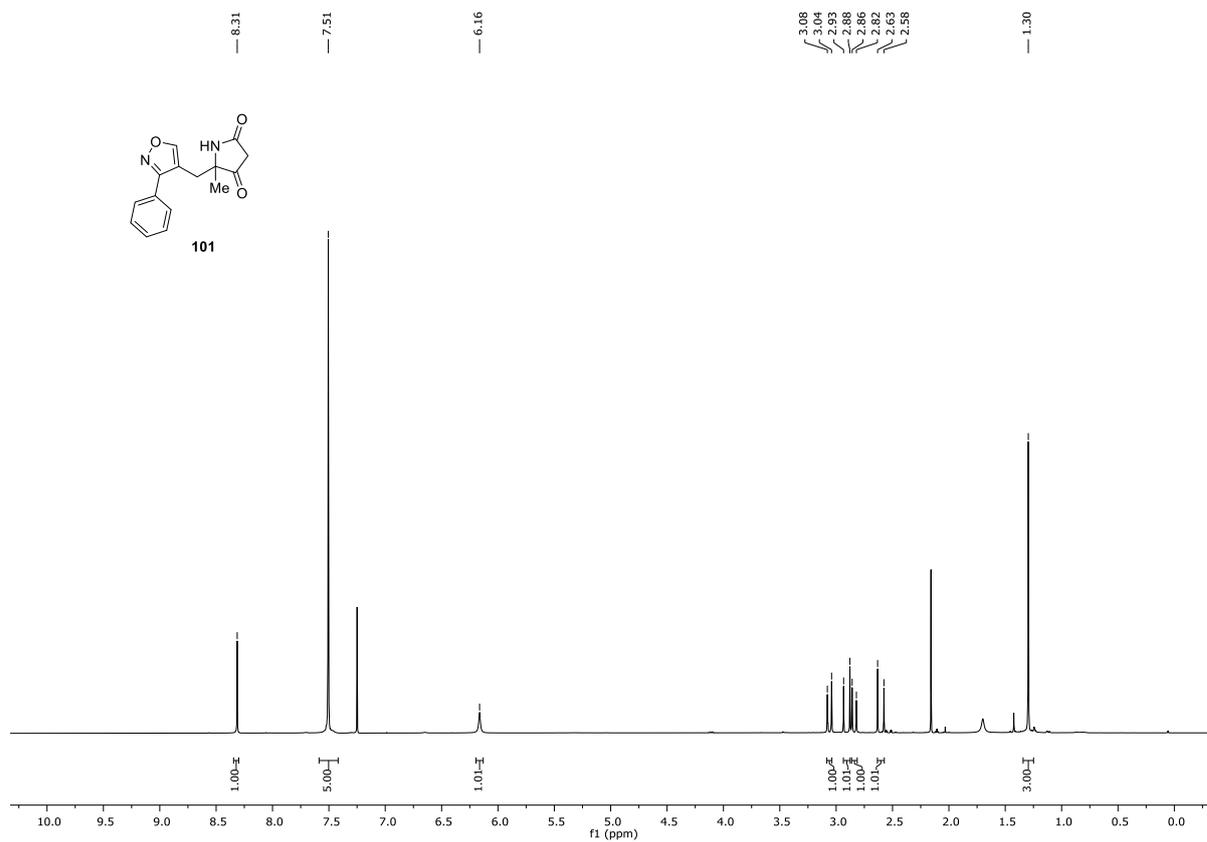


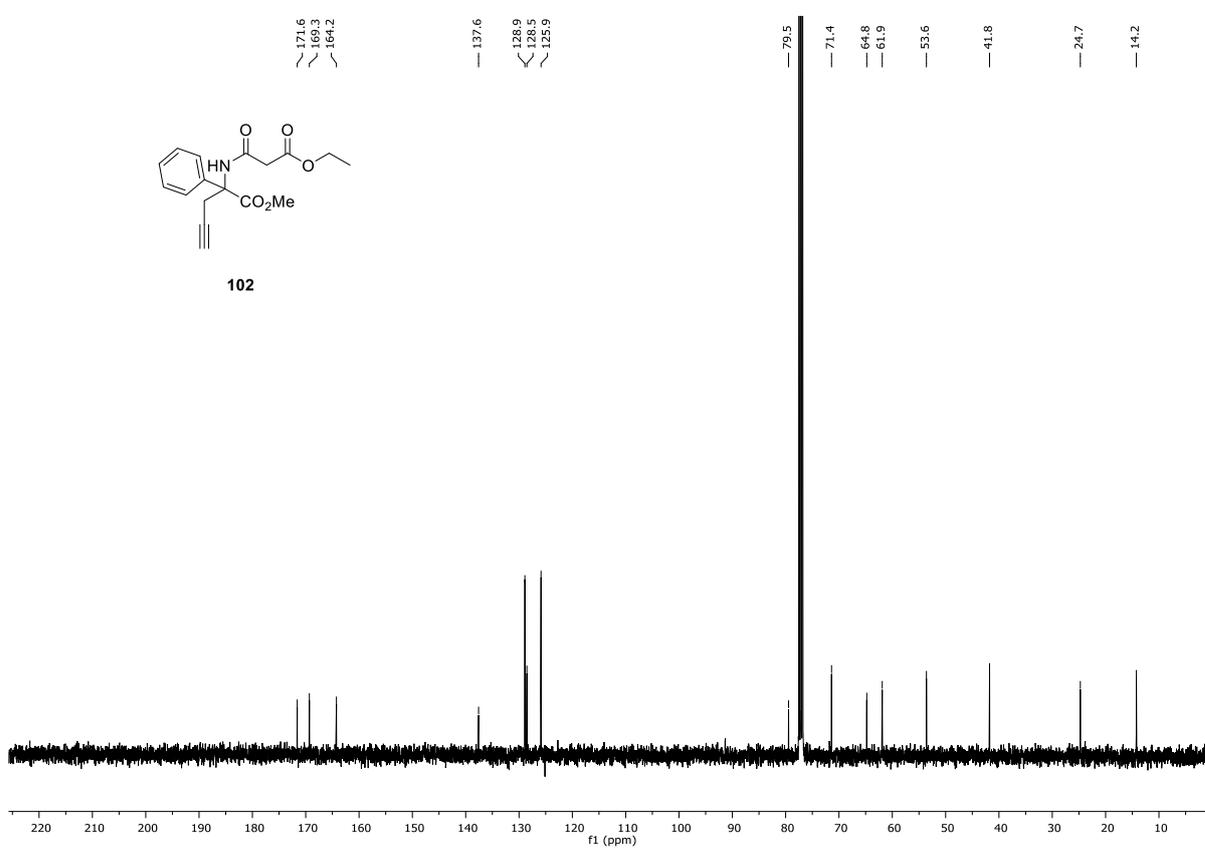
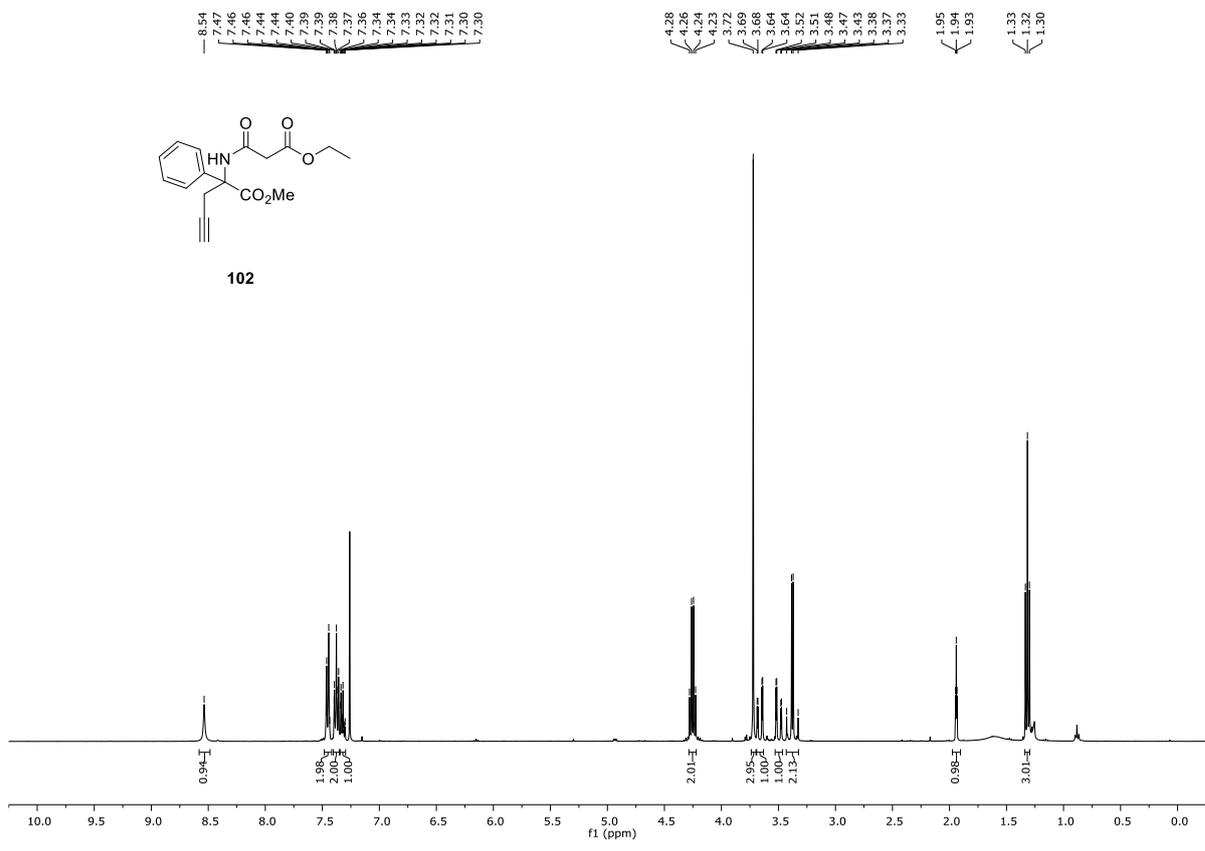
100

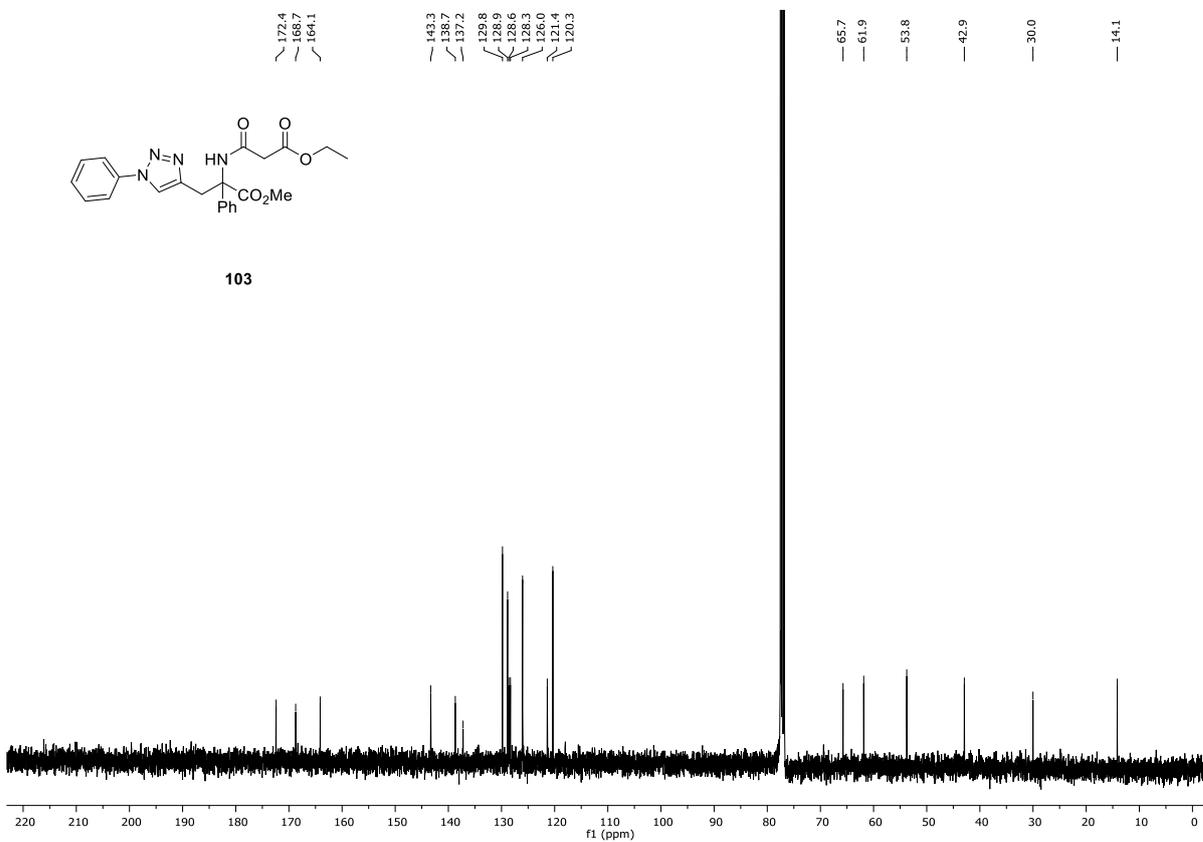
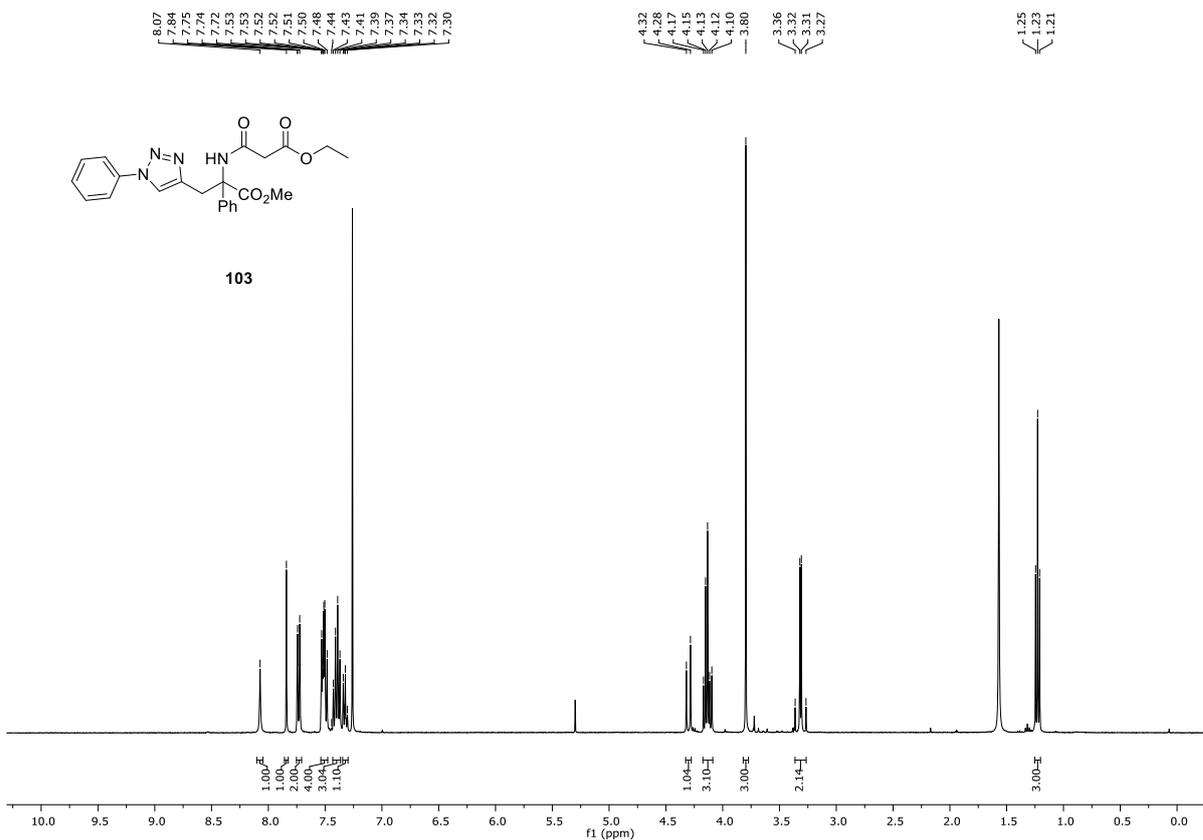


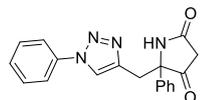
100



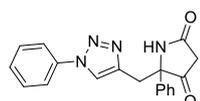
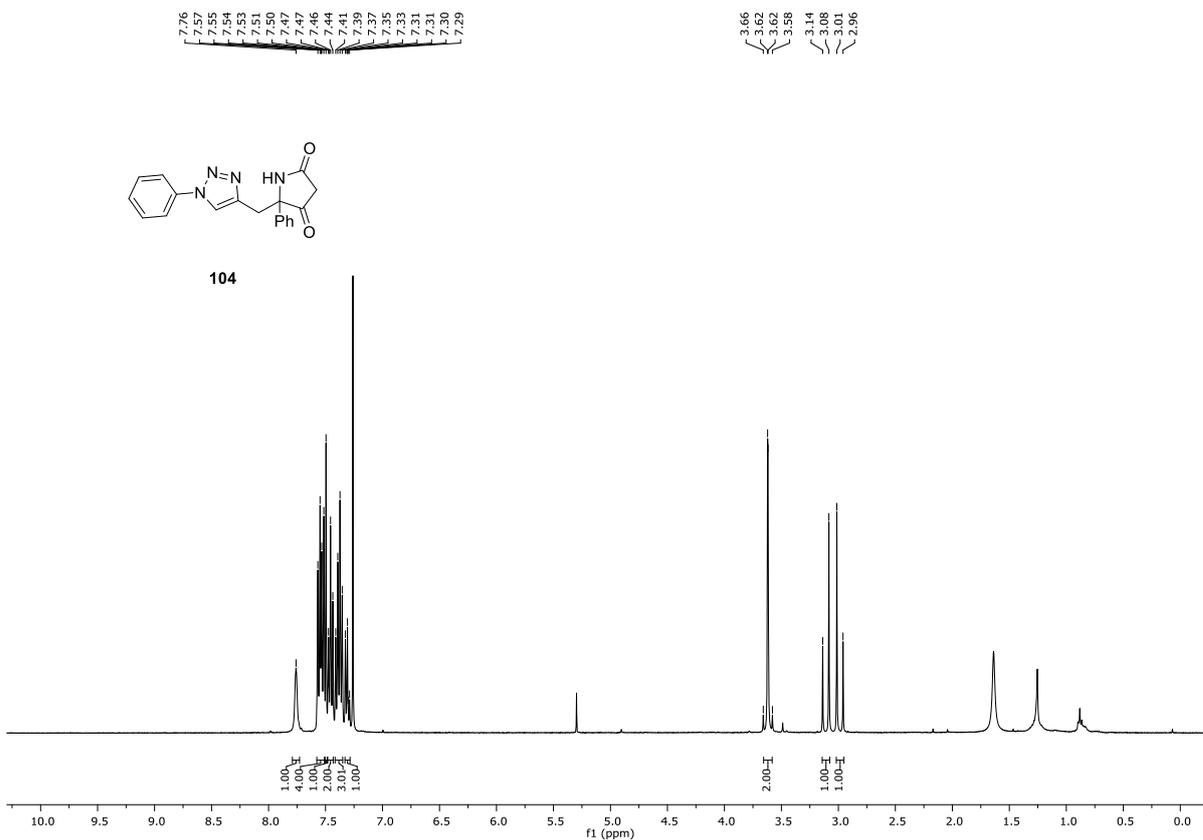




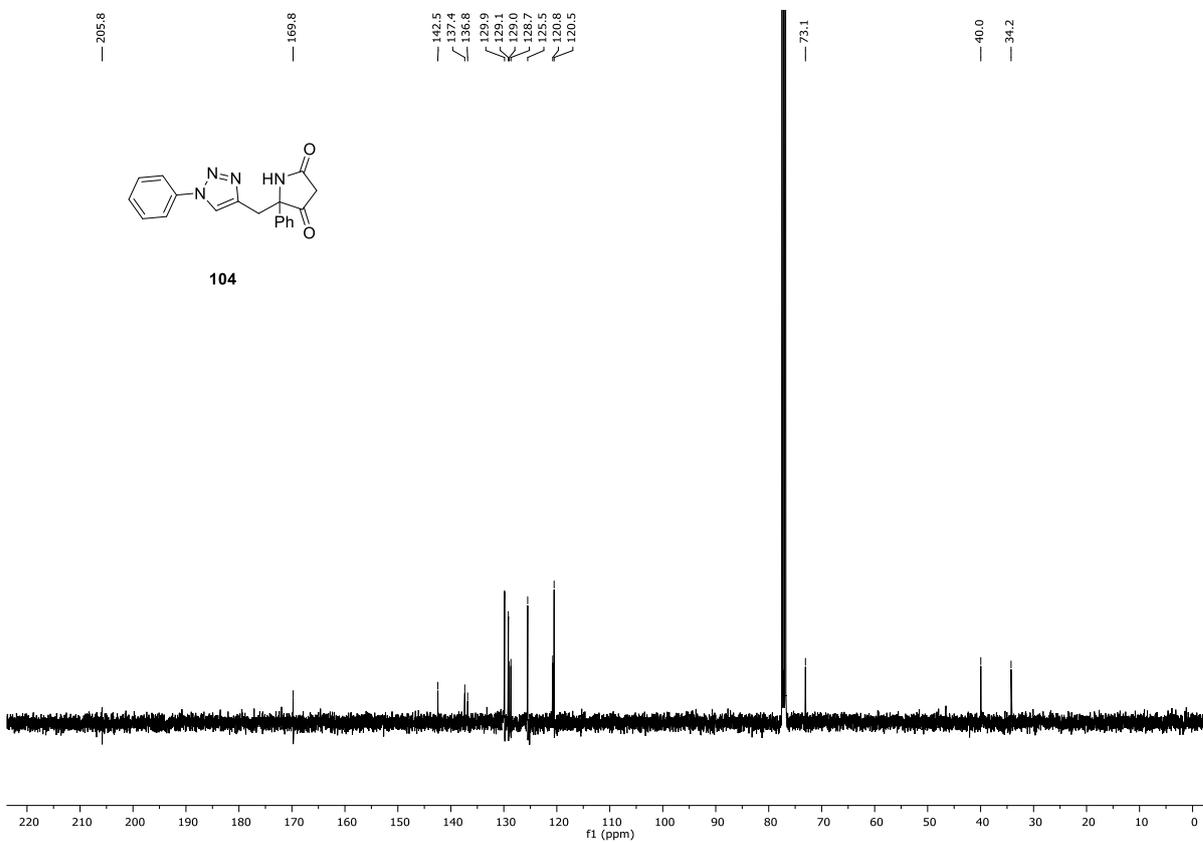


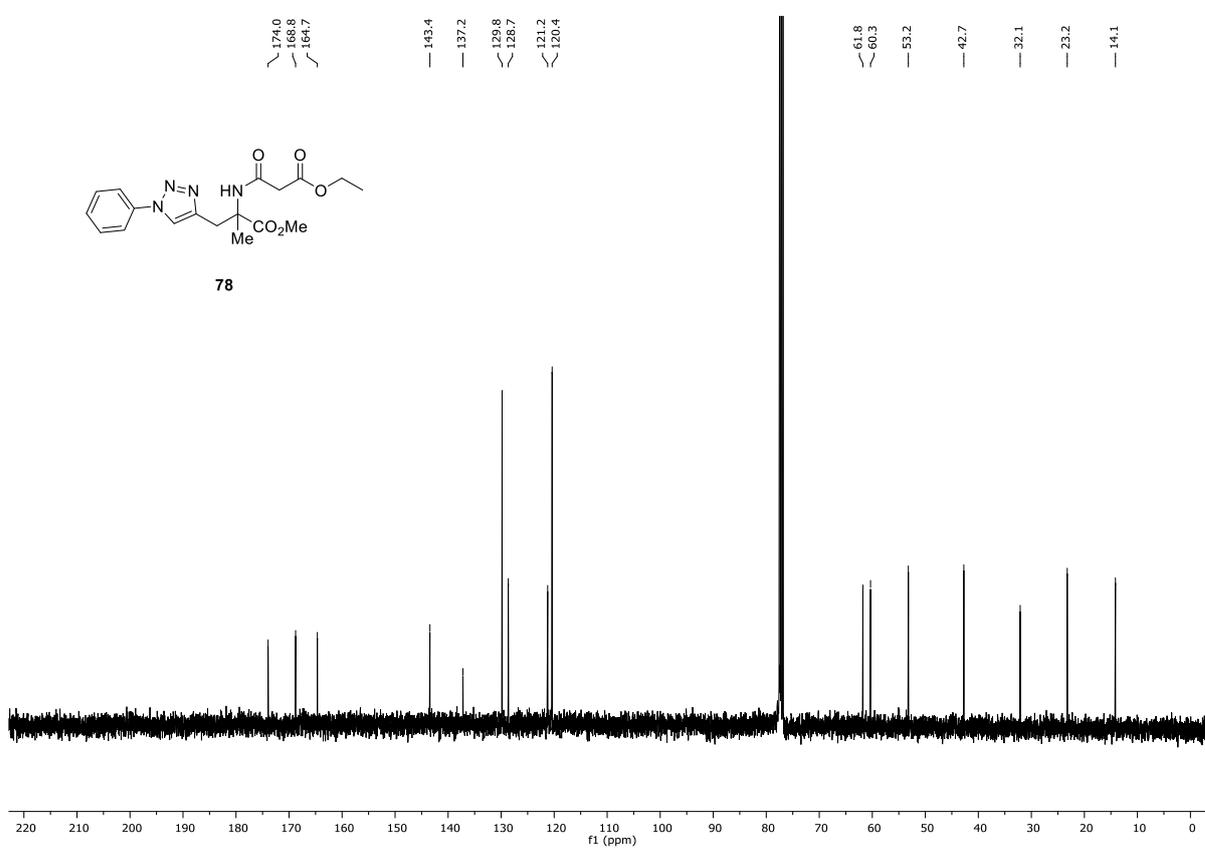
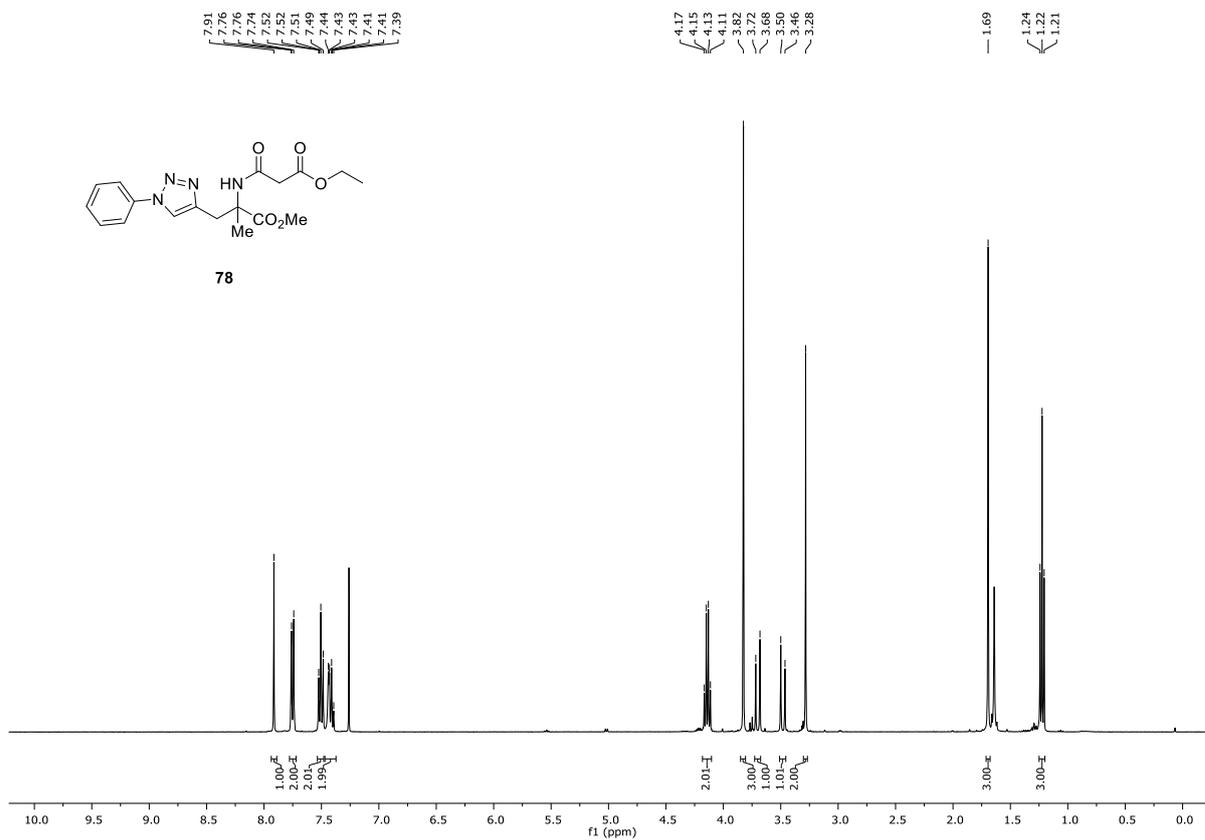


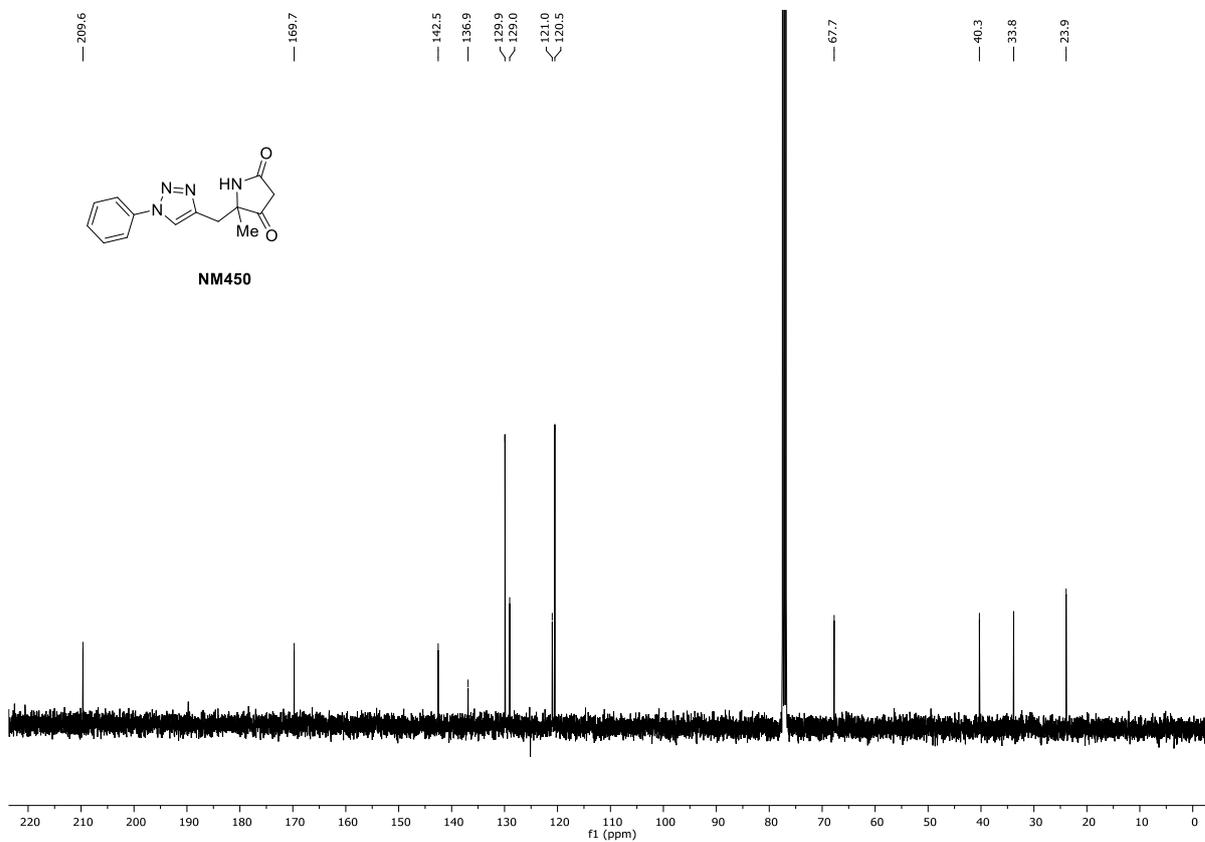
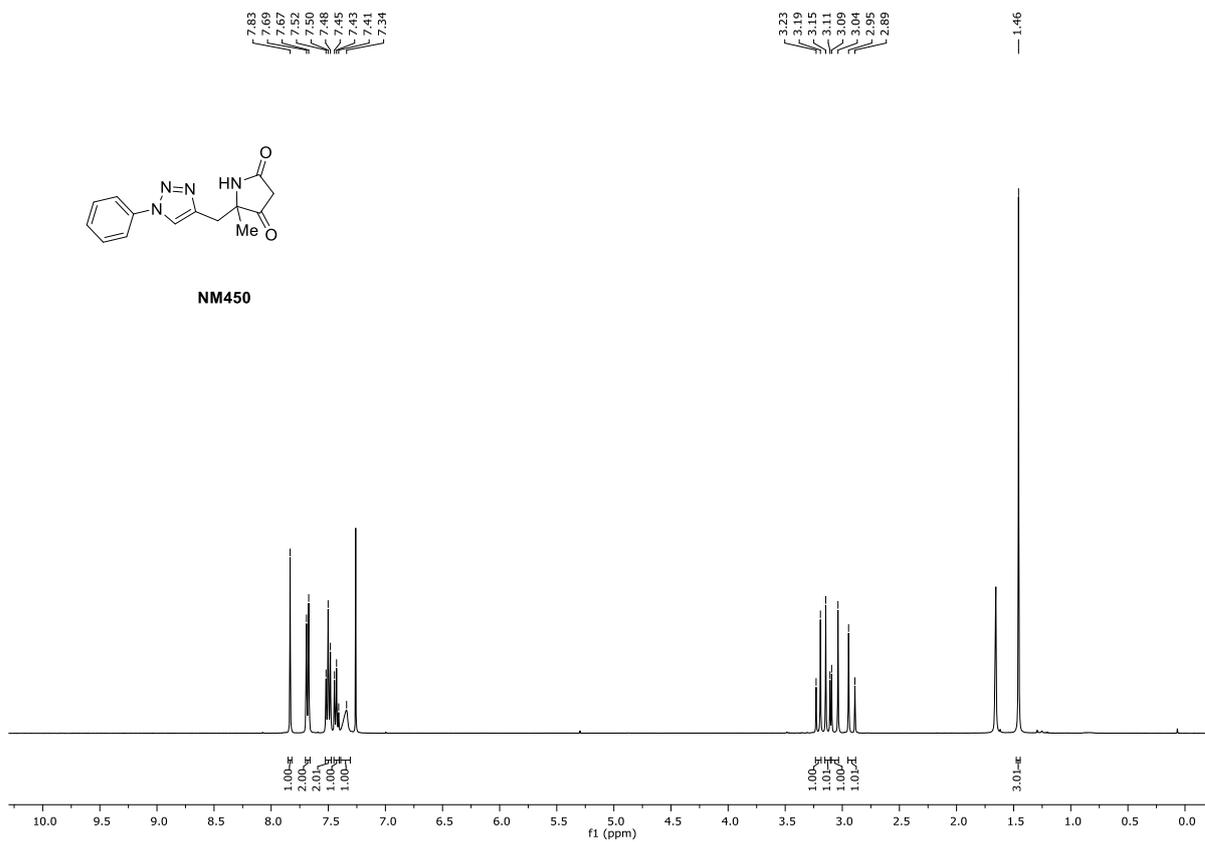
104

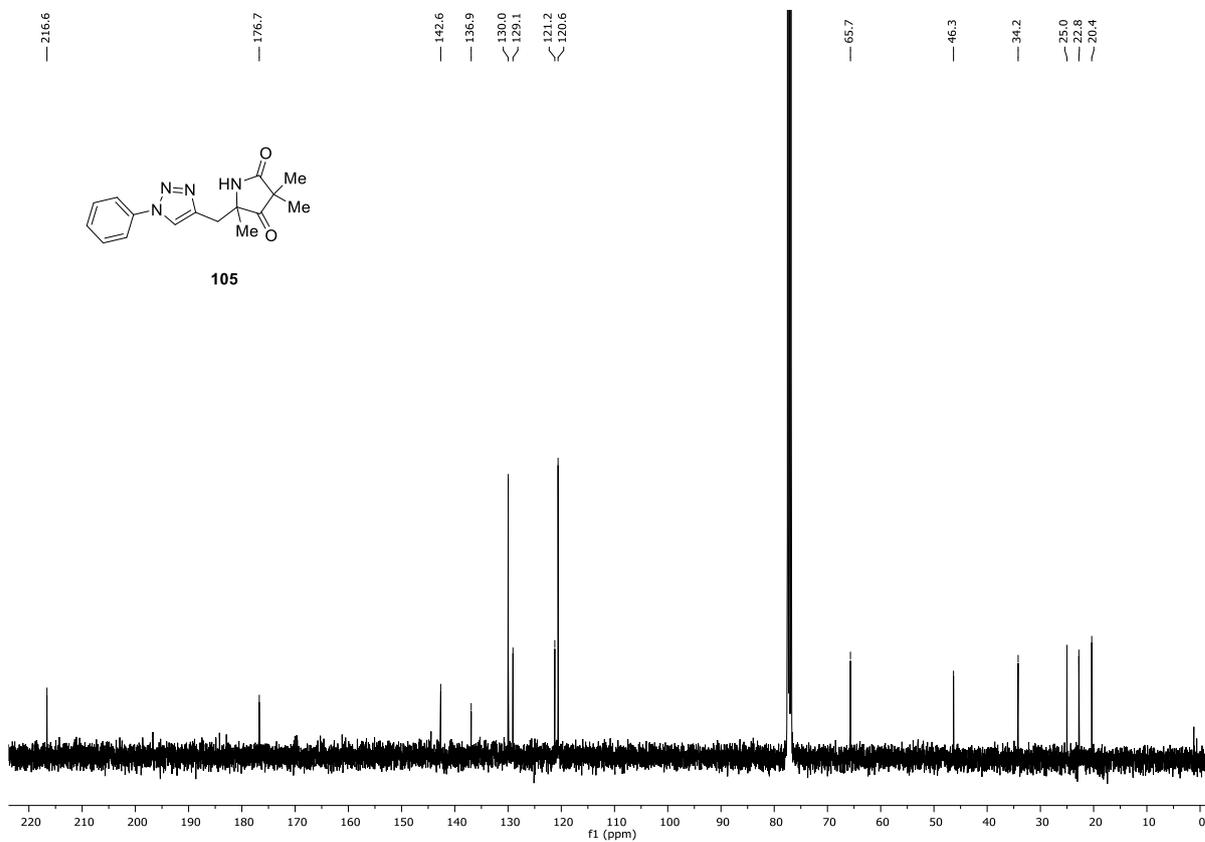
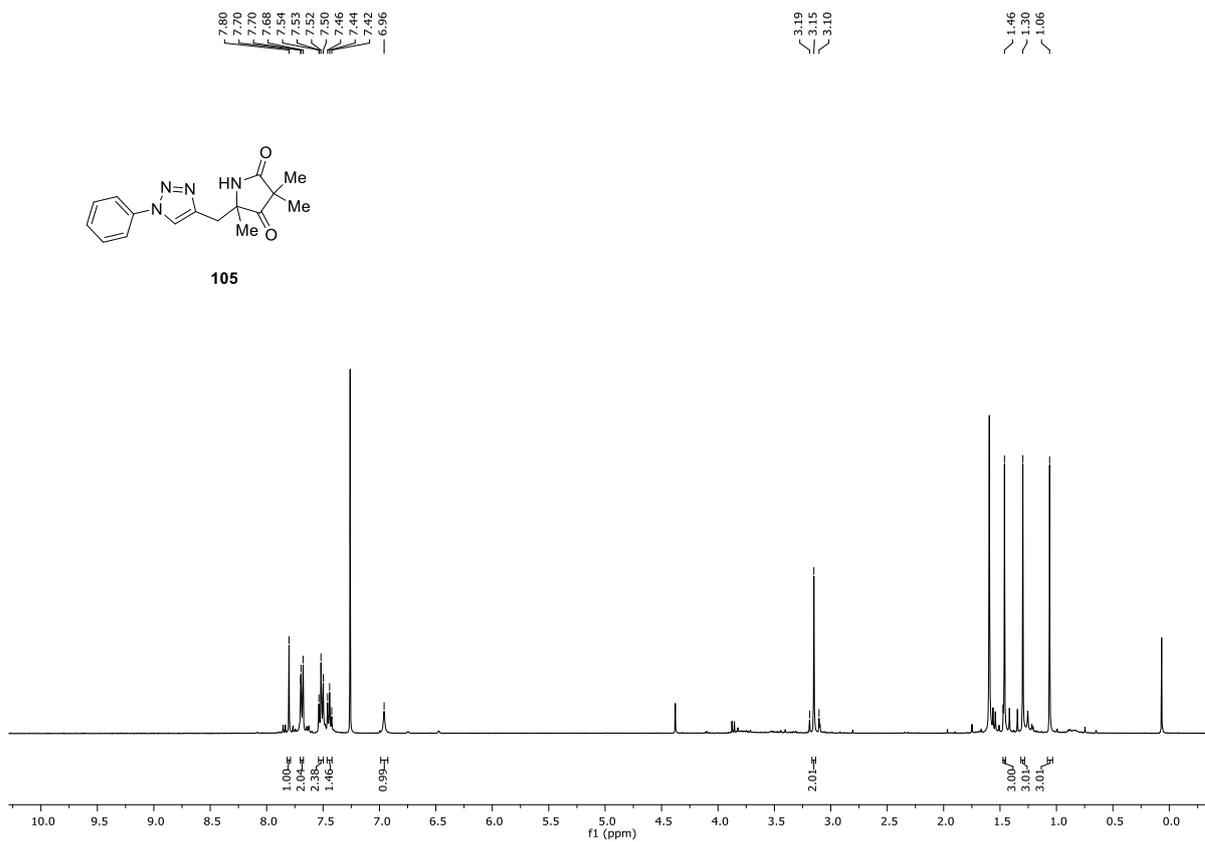


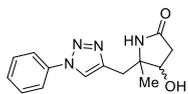
104







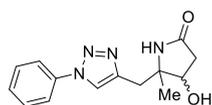
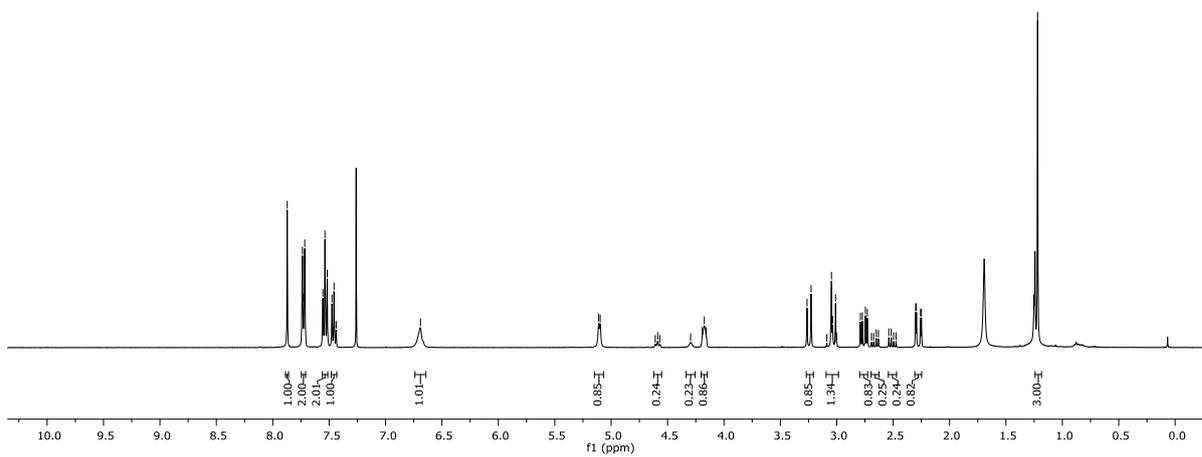




106

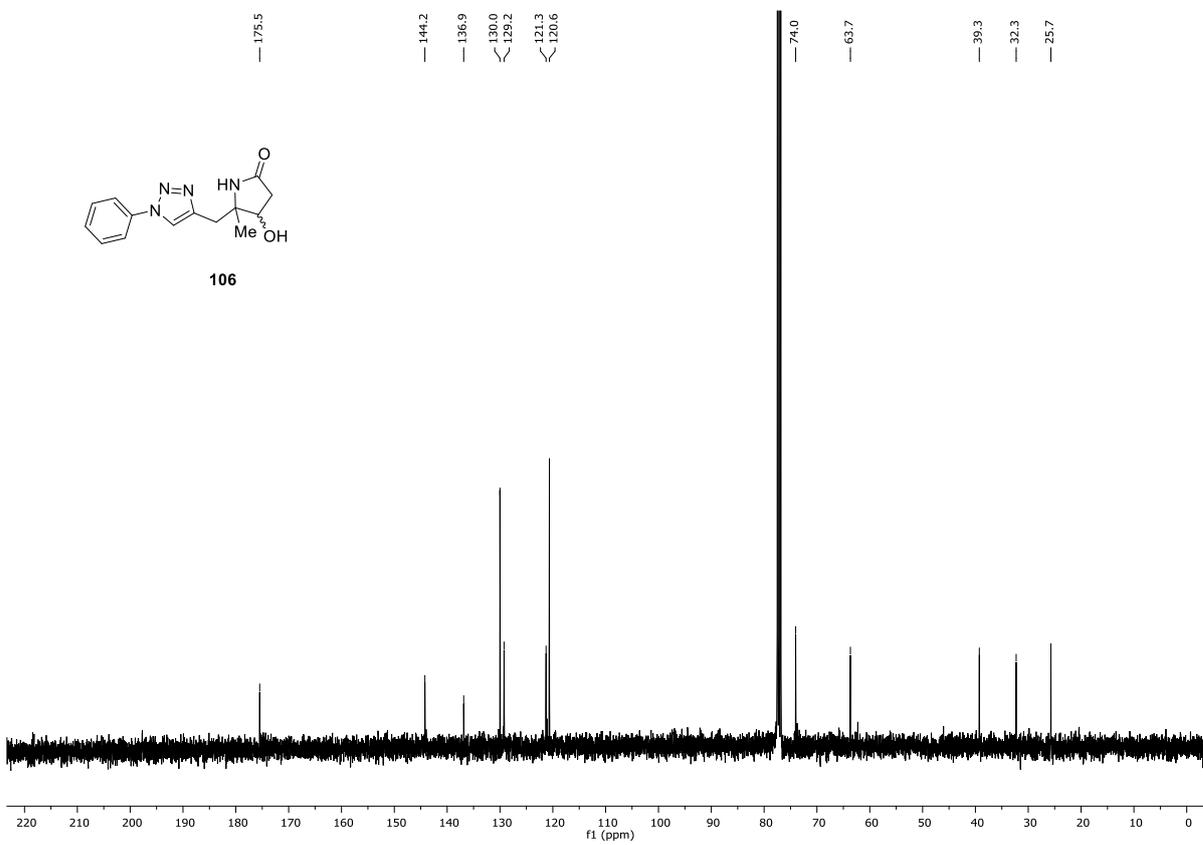
7.87  
7.74  
7.74  
7.72  
7.56  
7.55  
7.54  
7.52  
7.47  
7.46  
7.44  
— 6.69

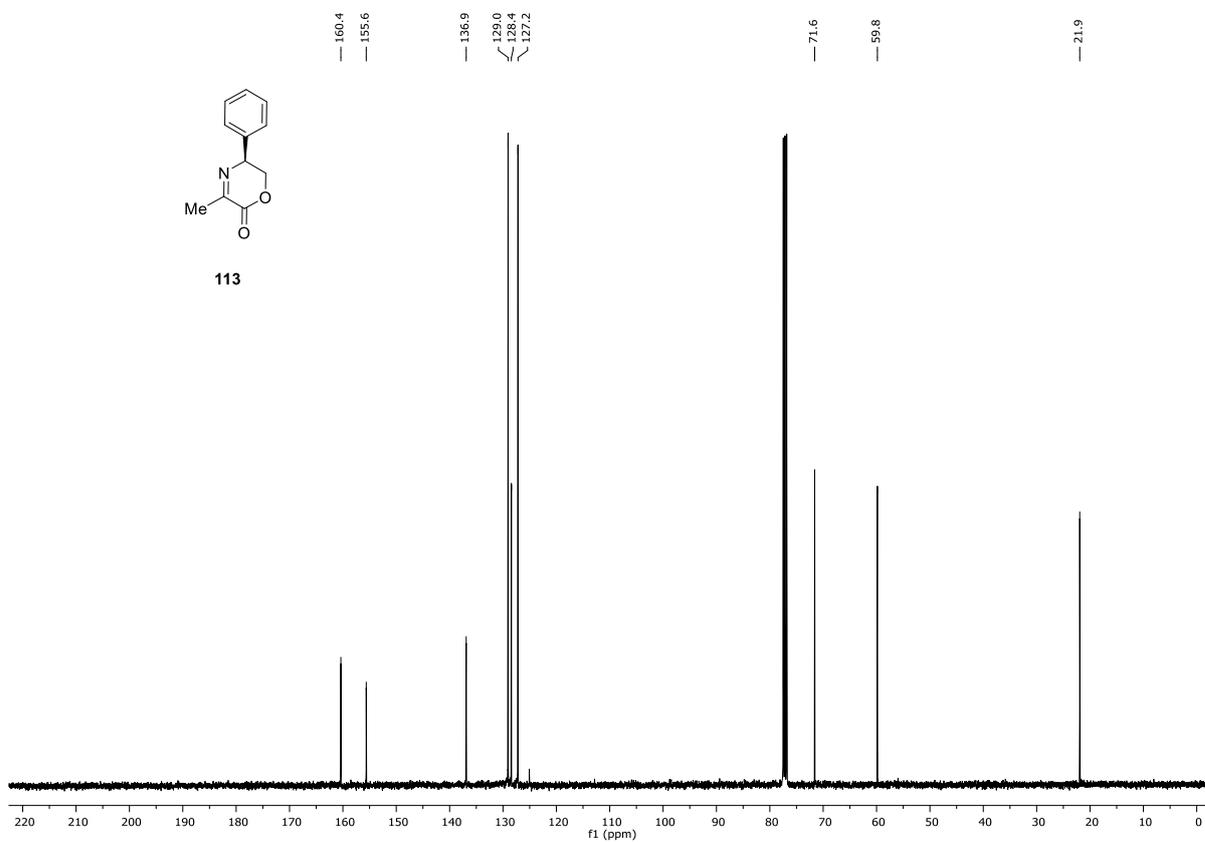
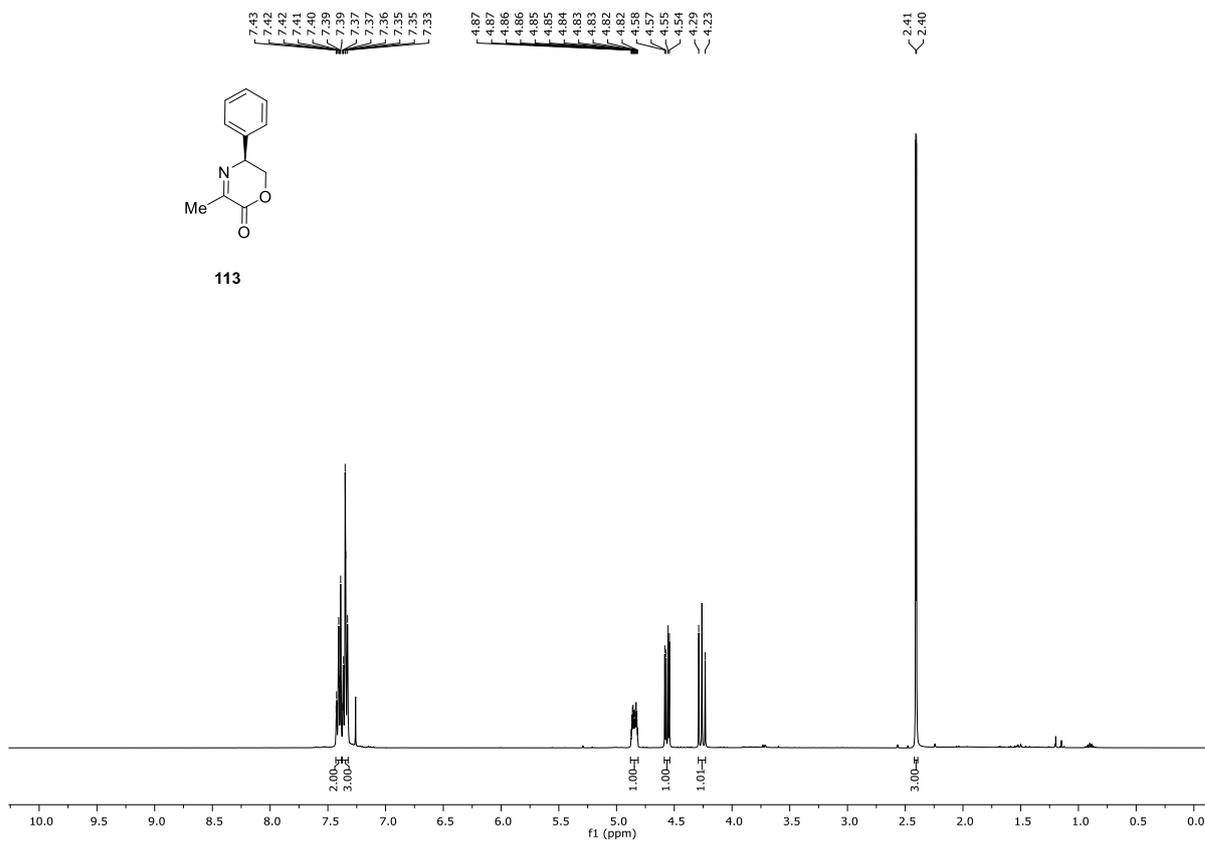
5.11  
4.61  
4.59  
4.57  
4.30  
4.19  
4.17  
4.16  
3.26  
3.26  
3.09  
3.05  
3.04  
3.01  
3.00  
2.79  
2.77  
2.75  
2.73  
2.69  
2.67  
2.65  
2.63  
2.54  
2.52  
2.50  
2.47  
2.30  
2.29  
2.26  
2.25  
1.22

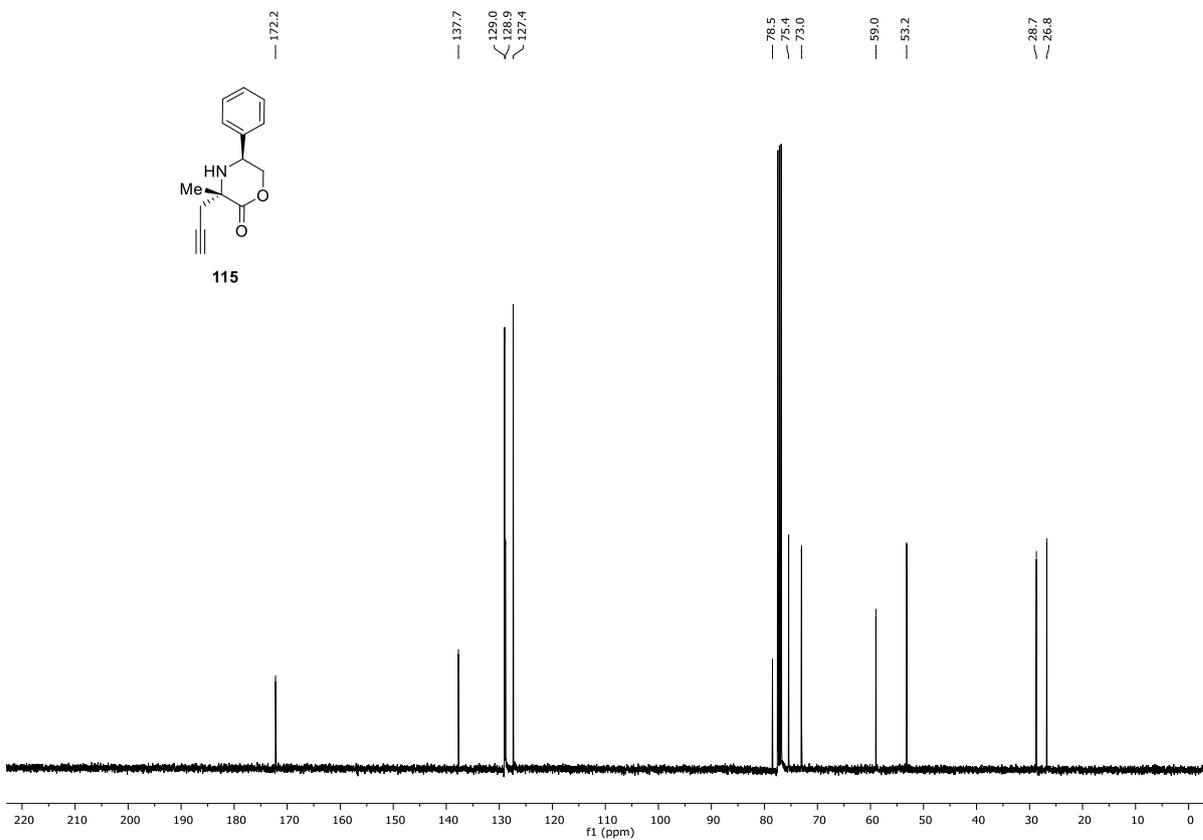
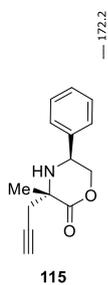
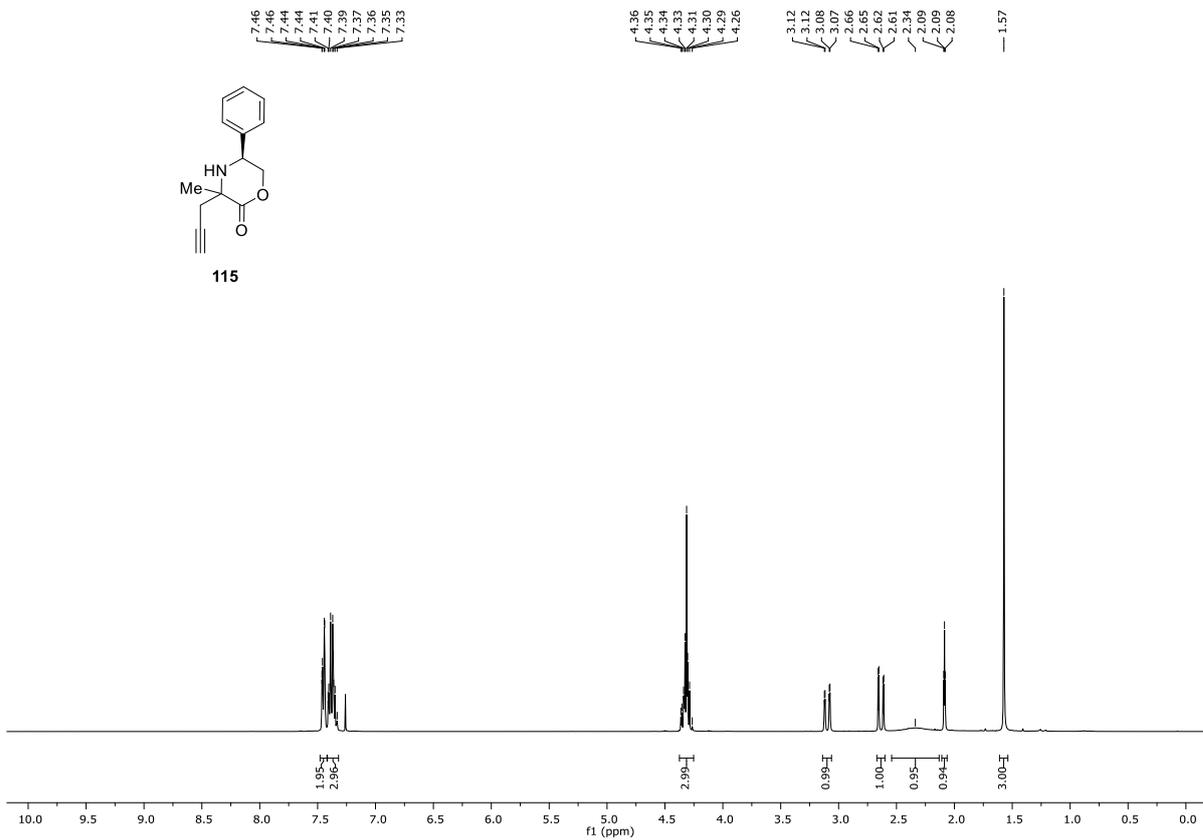
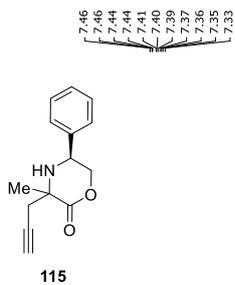


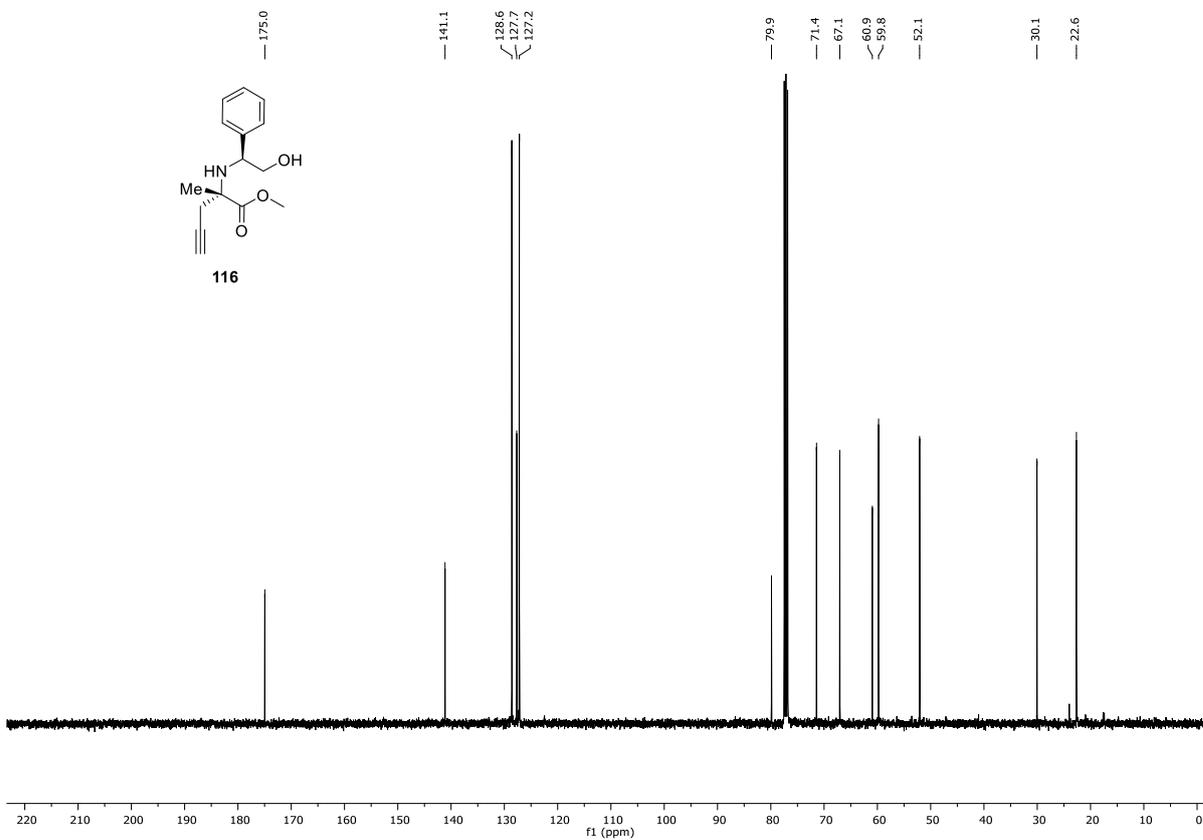
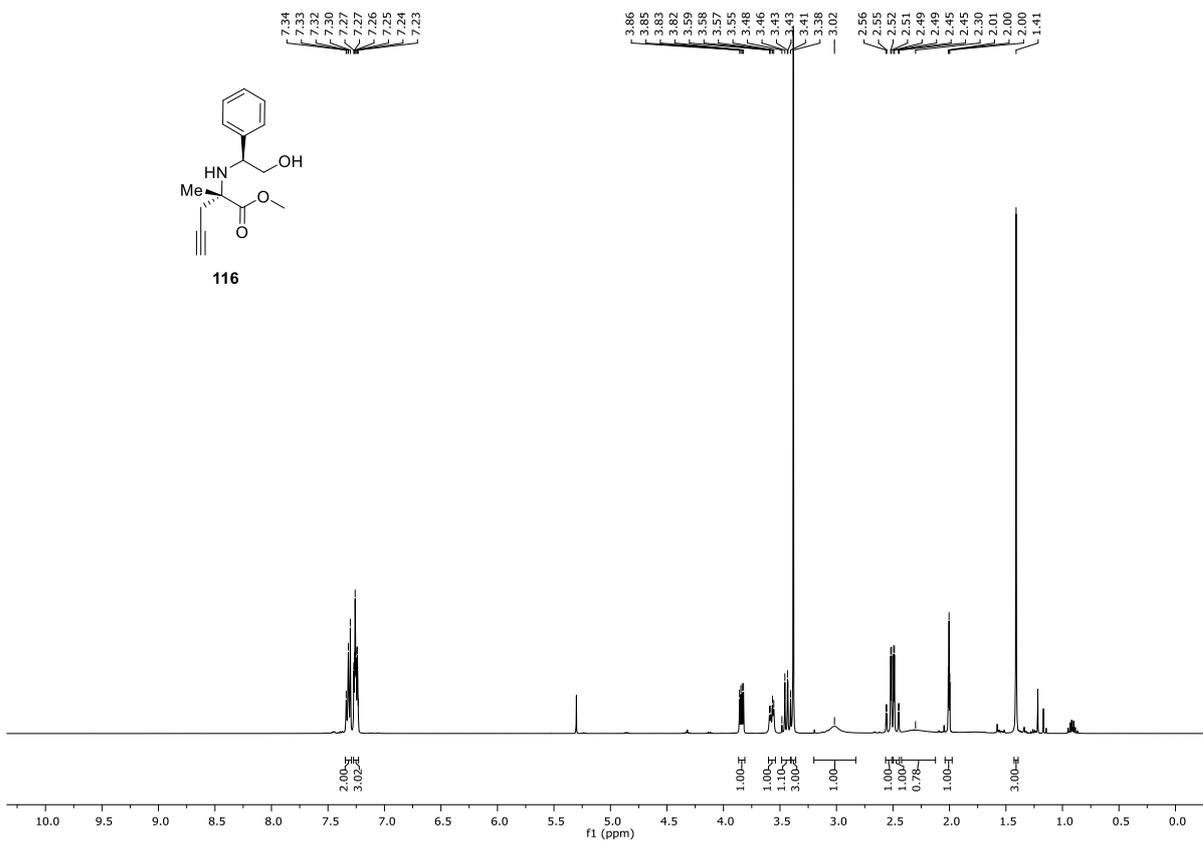
106

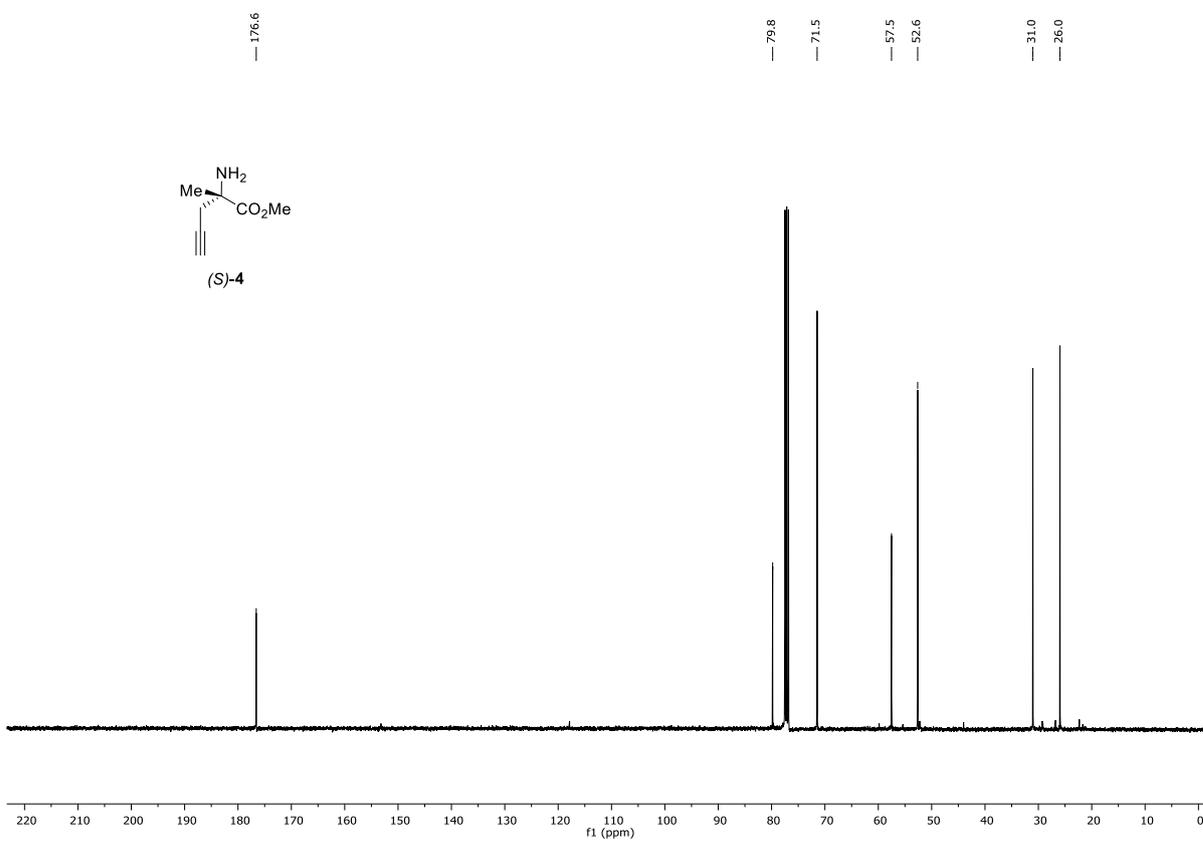
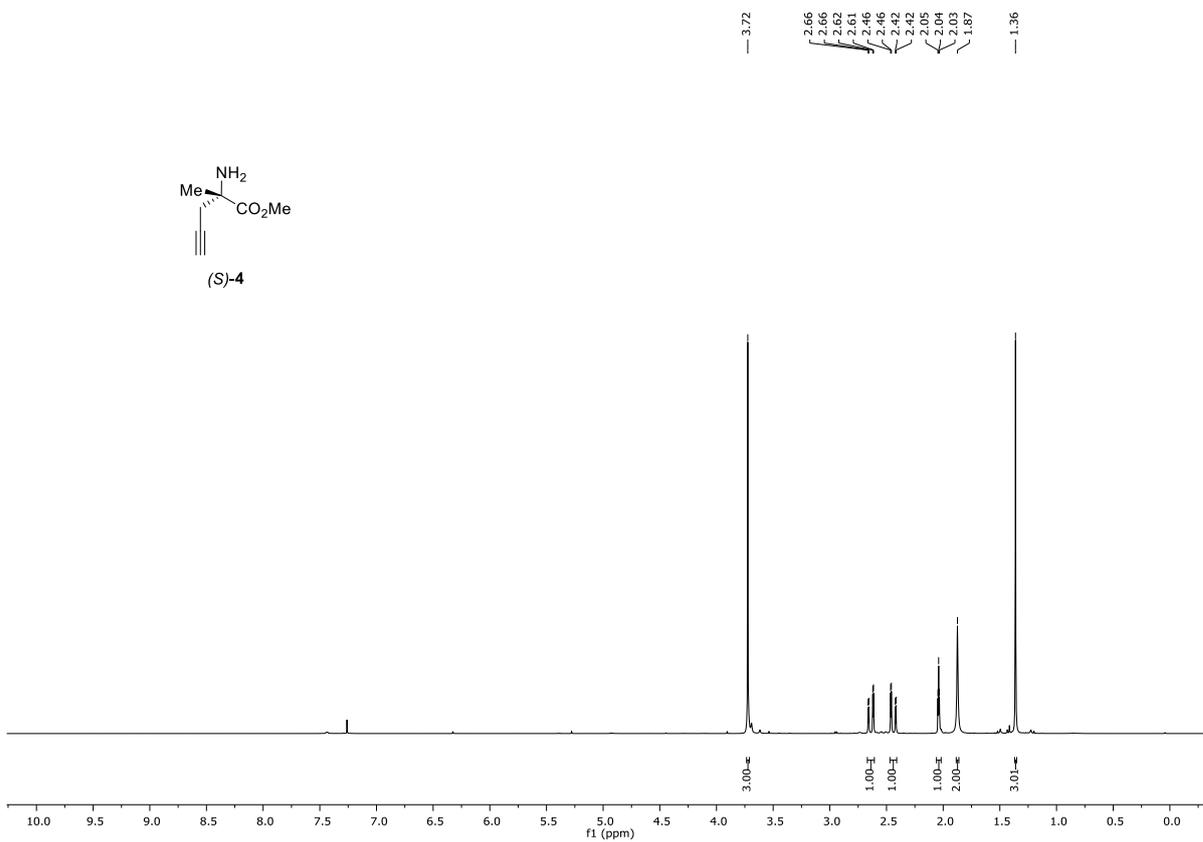
175.5  
144.2  
136.9  
130.0  
129.2  
121.3  
120.6  
74.0  
63.7  
39.3  
32.3  
25.7

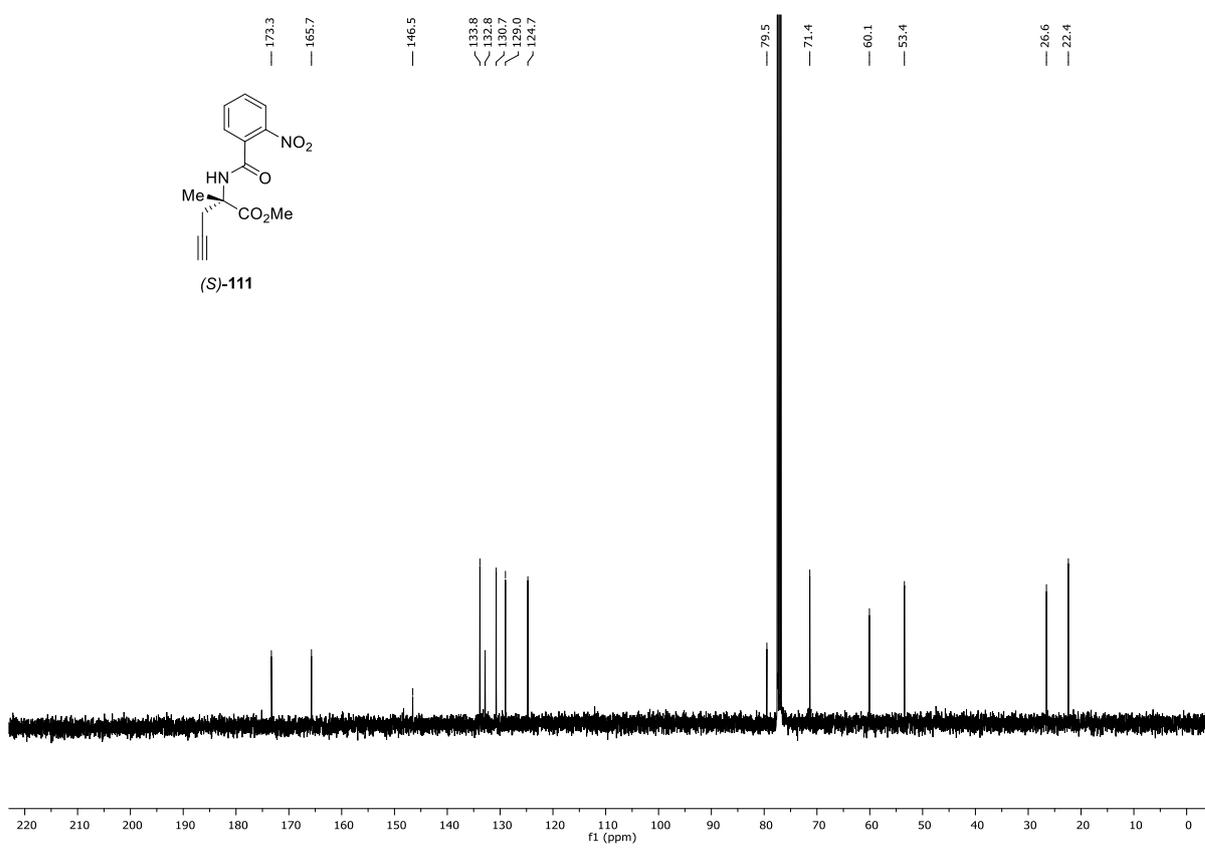
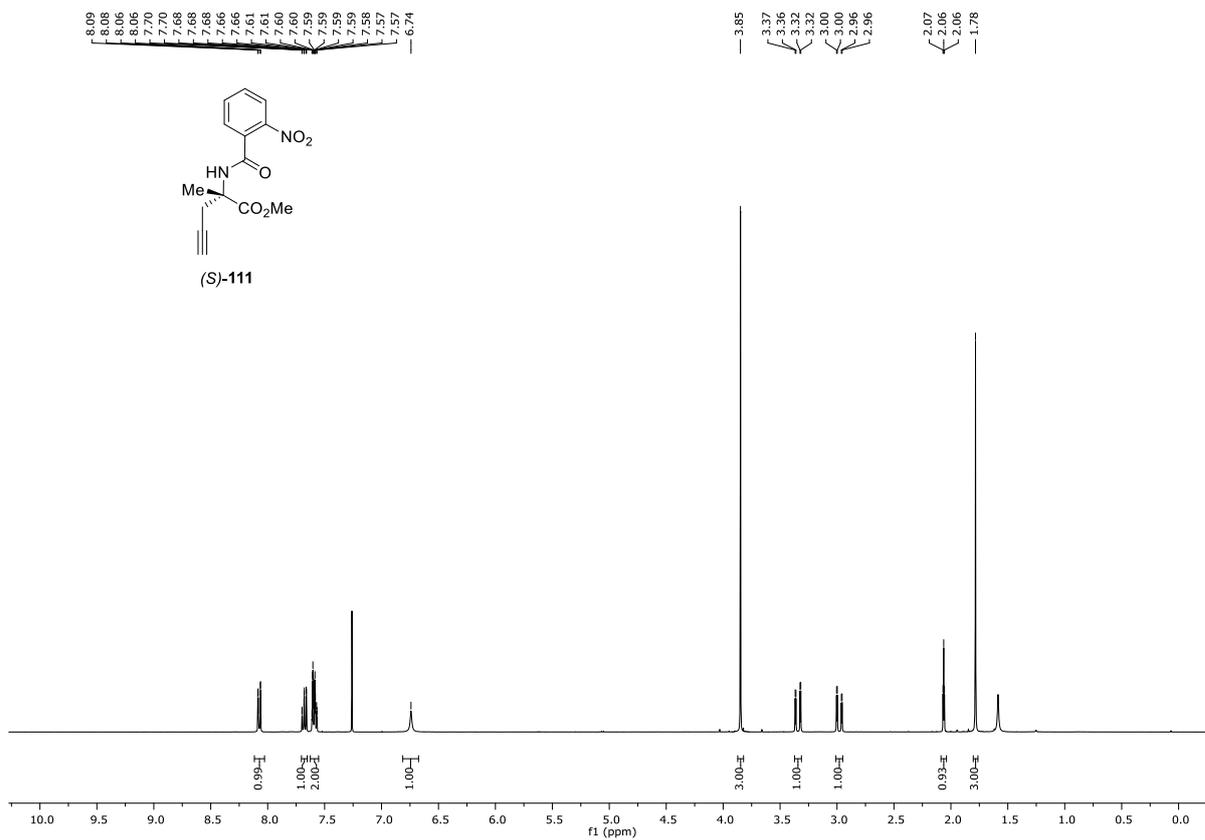


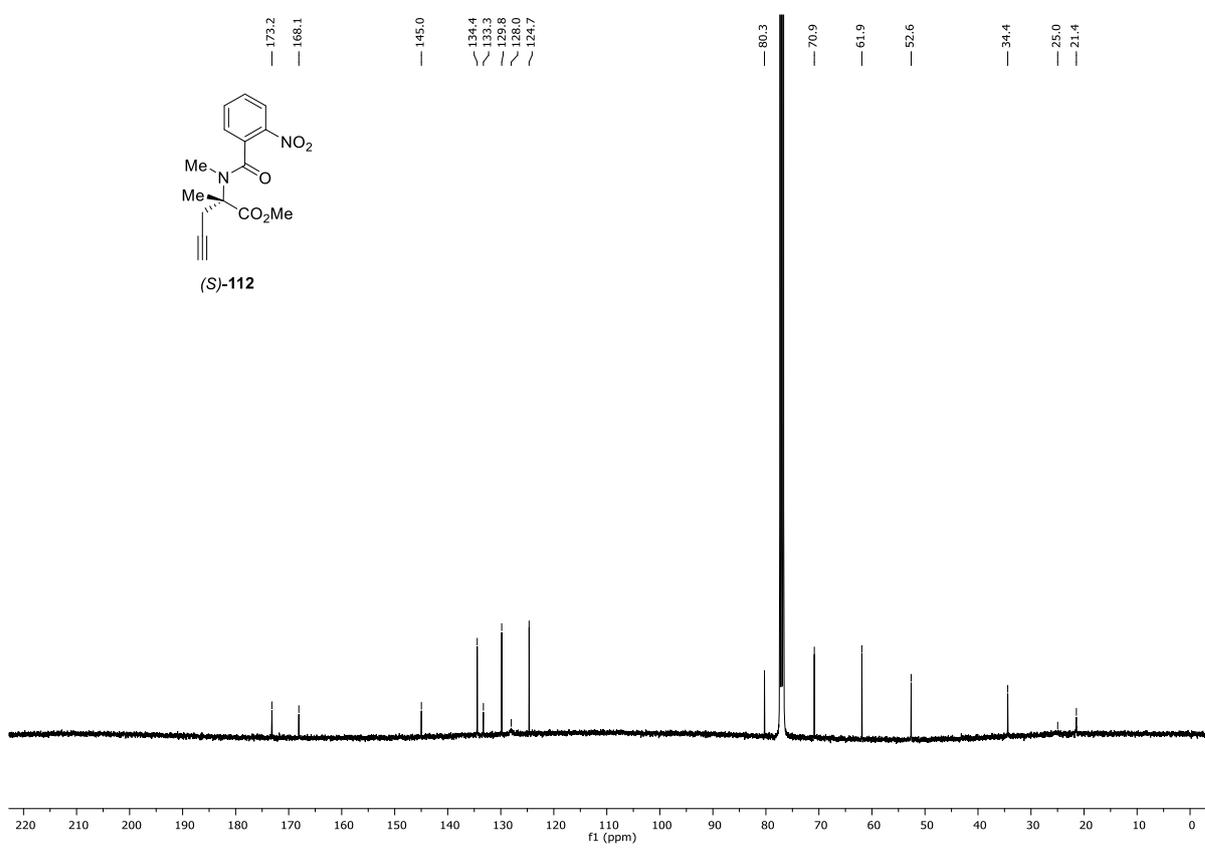
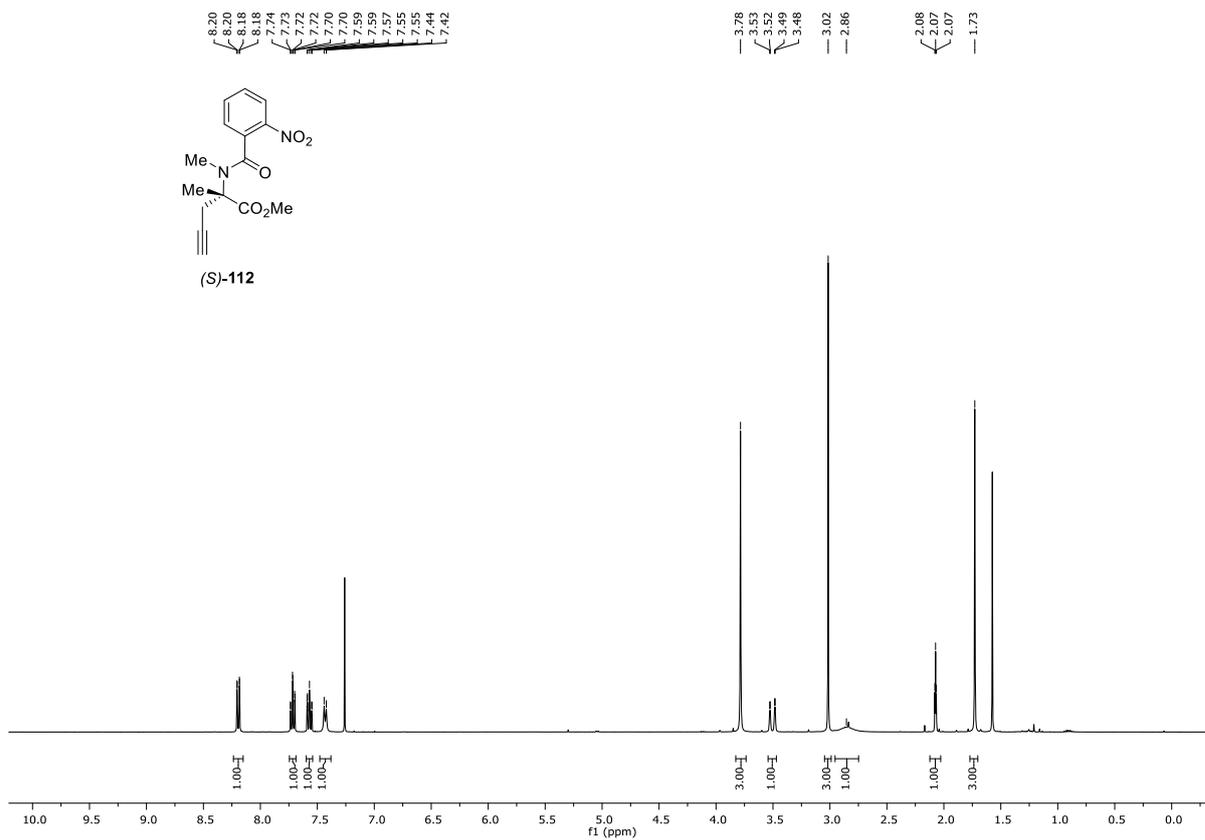


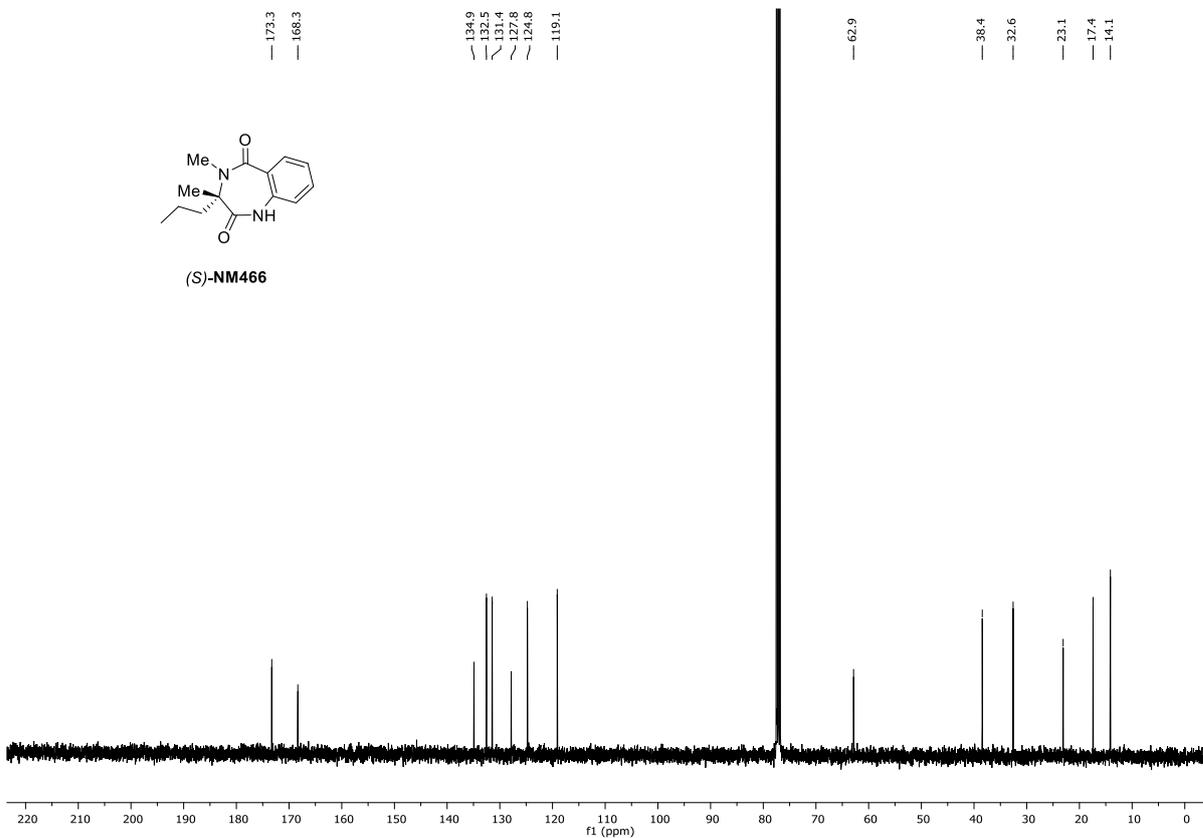
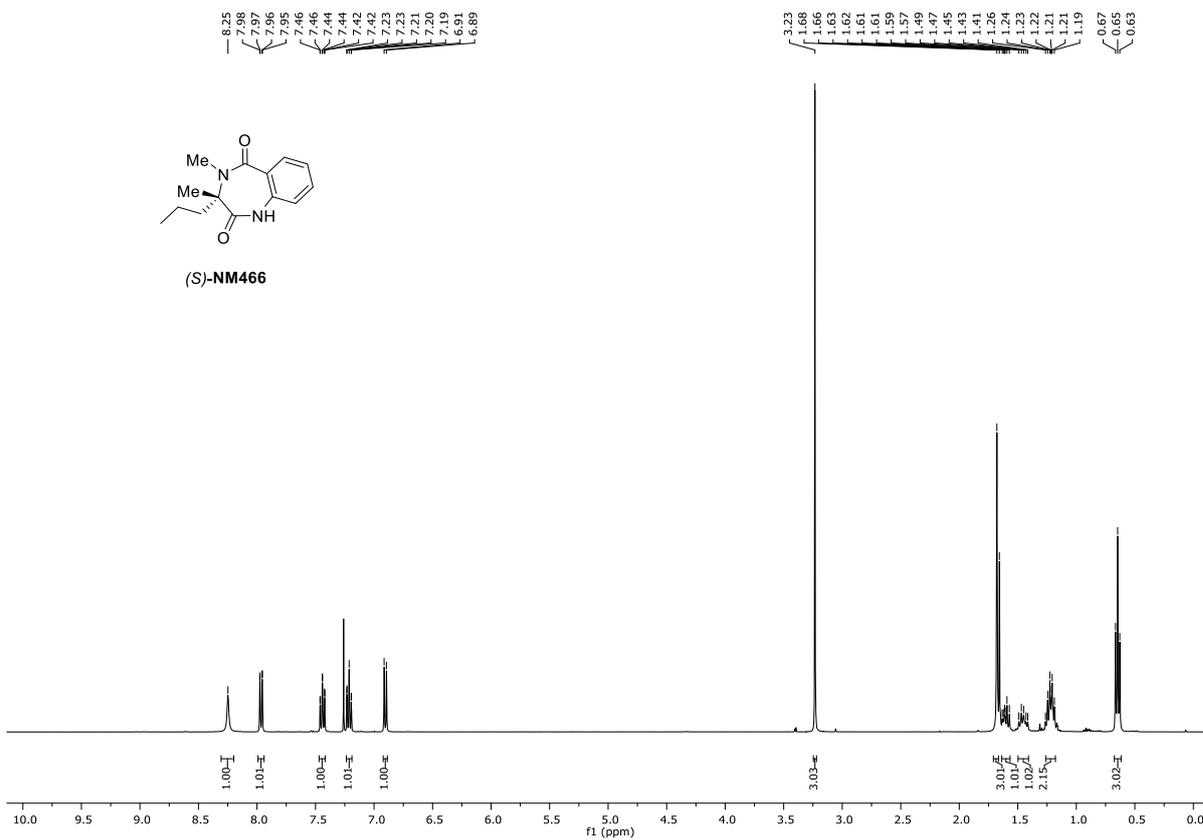


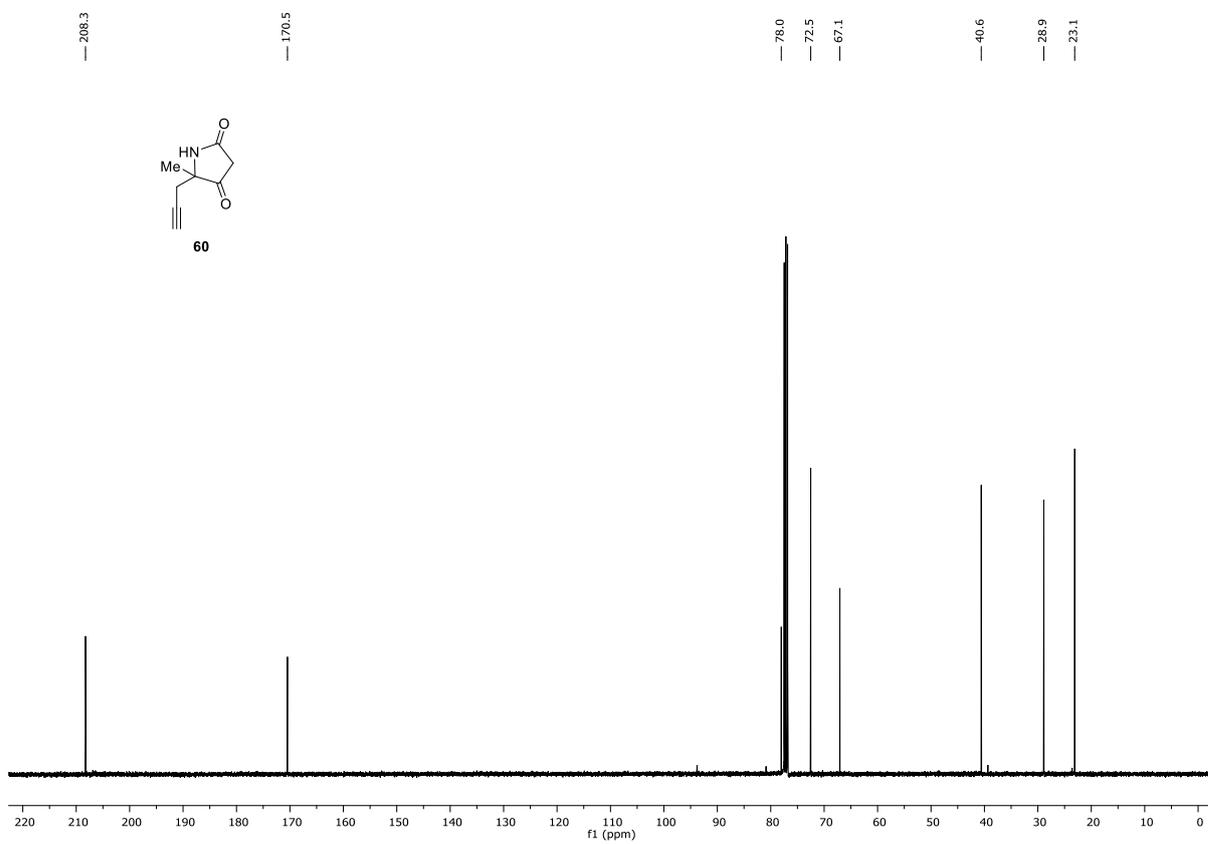
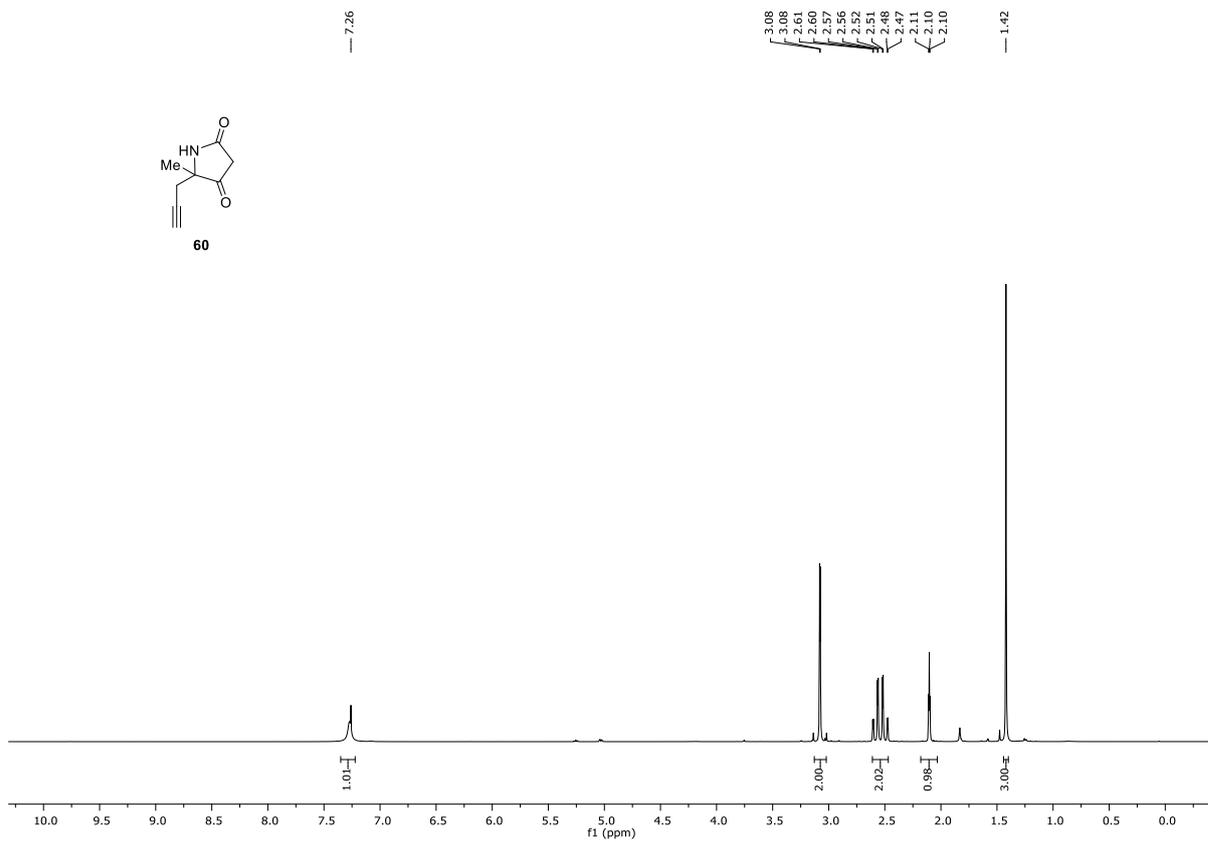


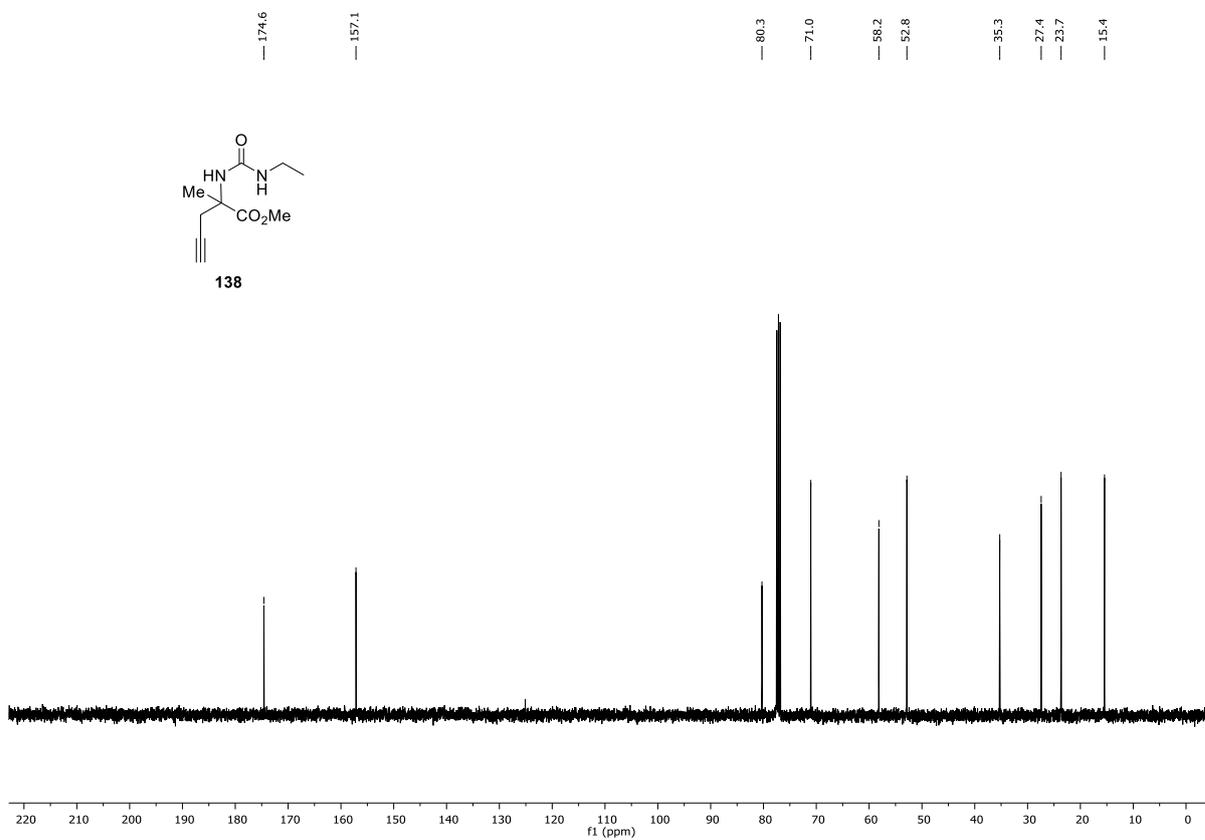
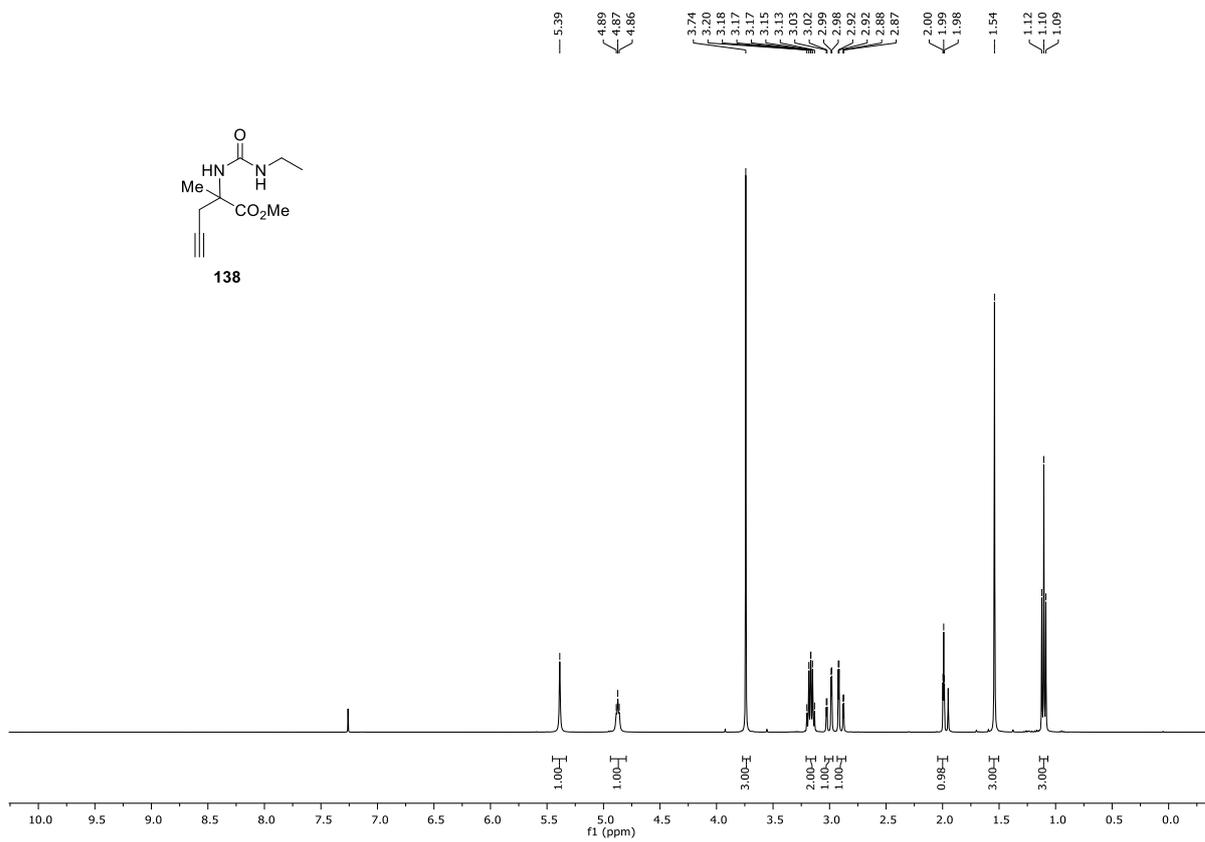


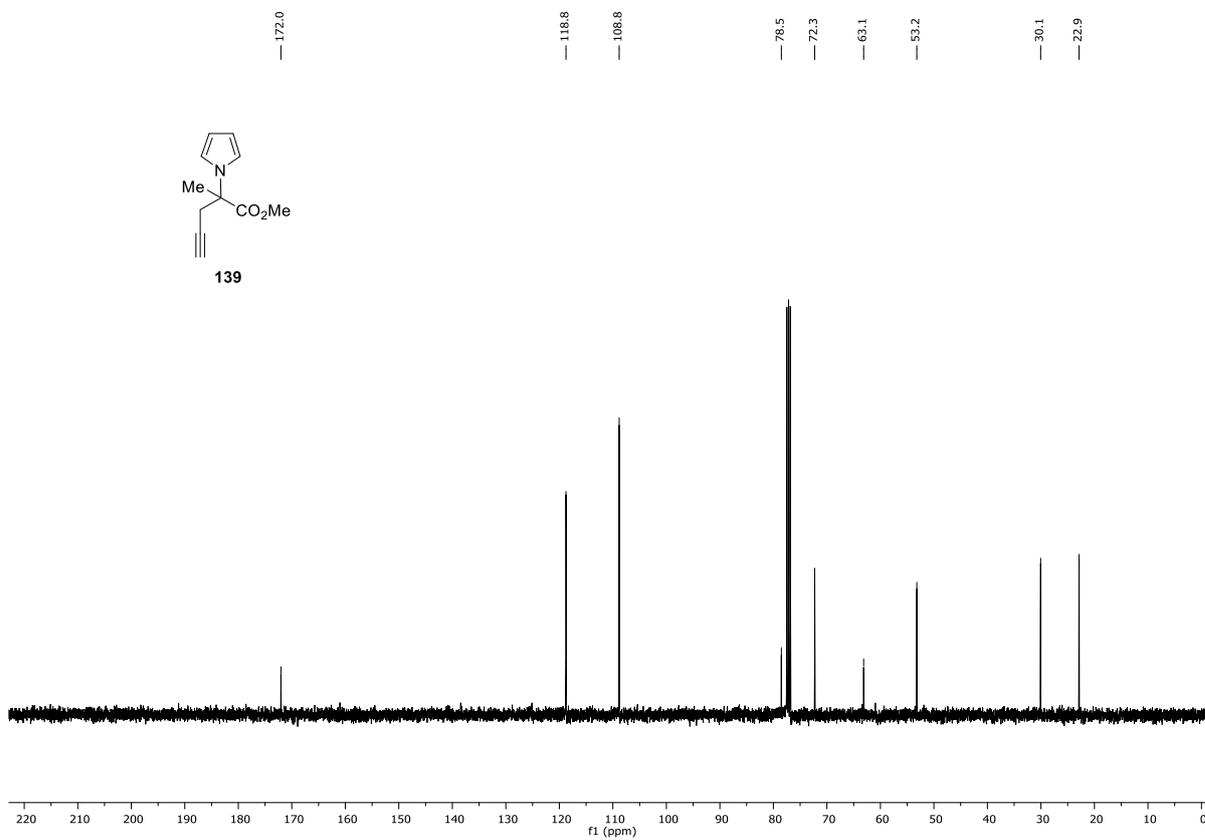
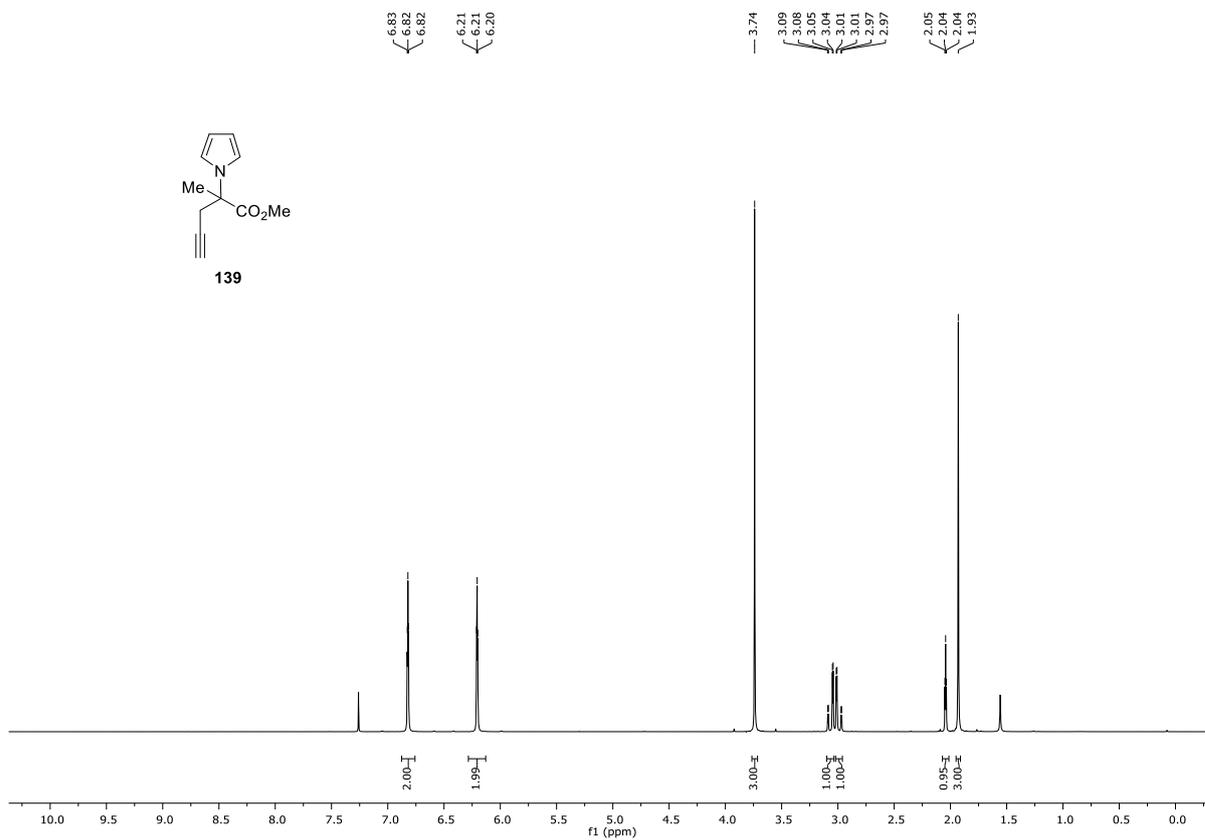


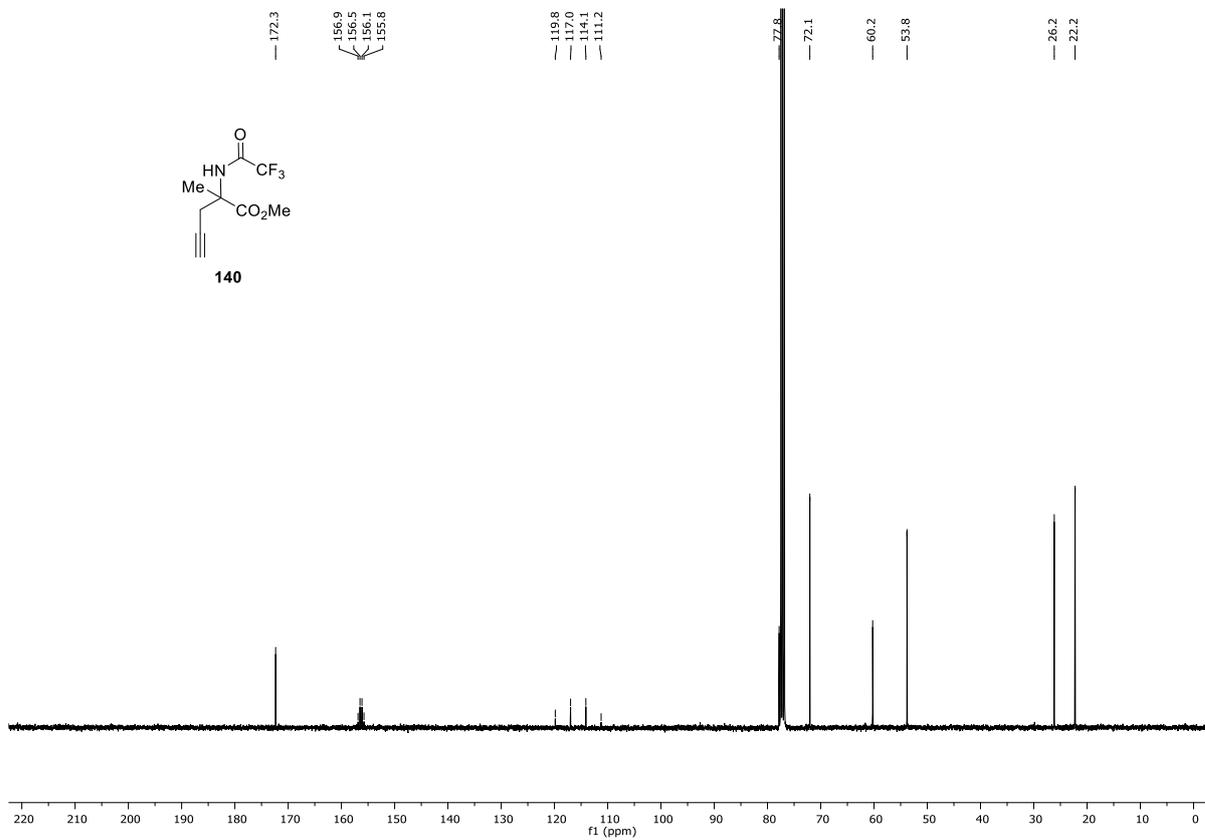
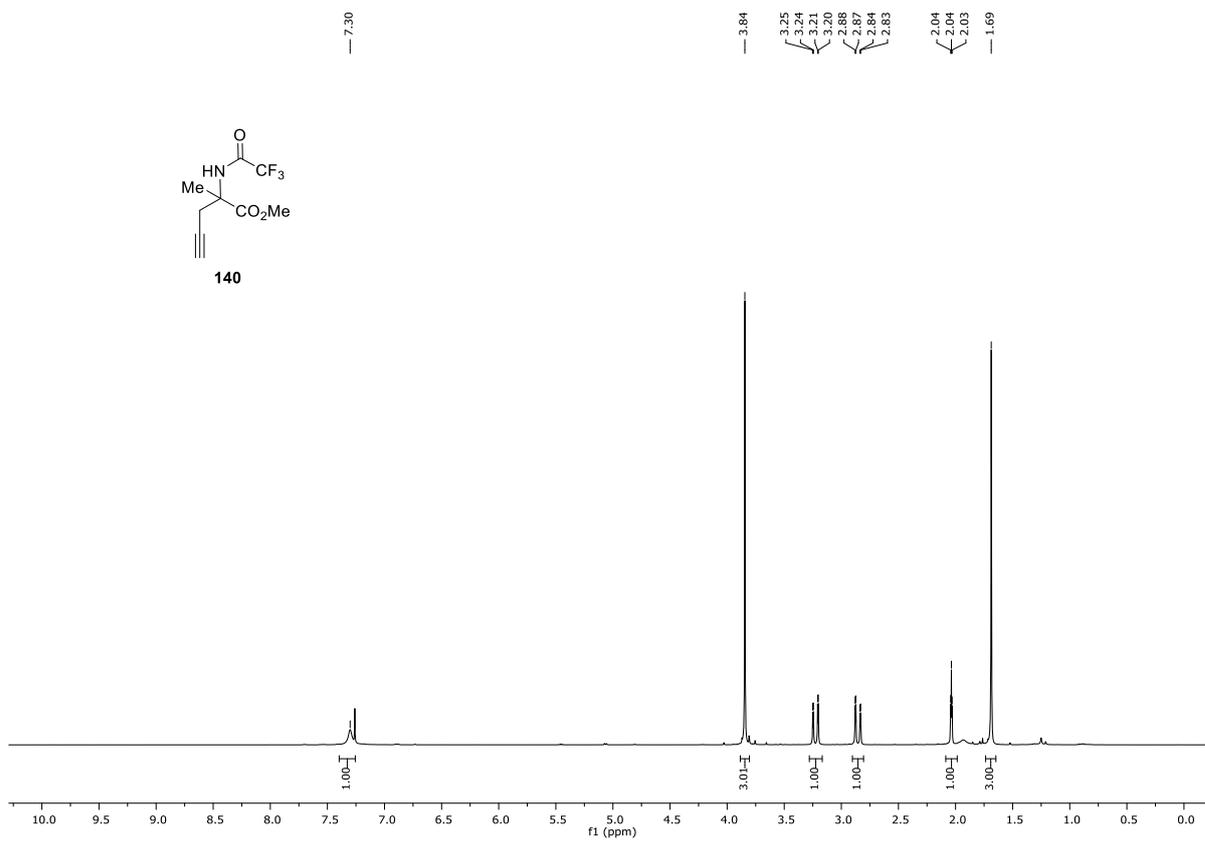


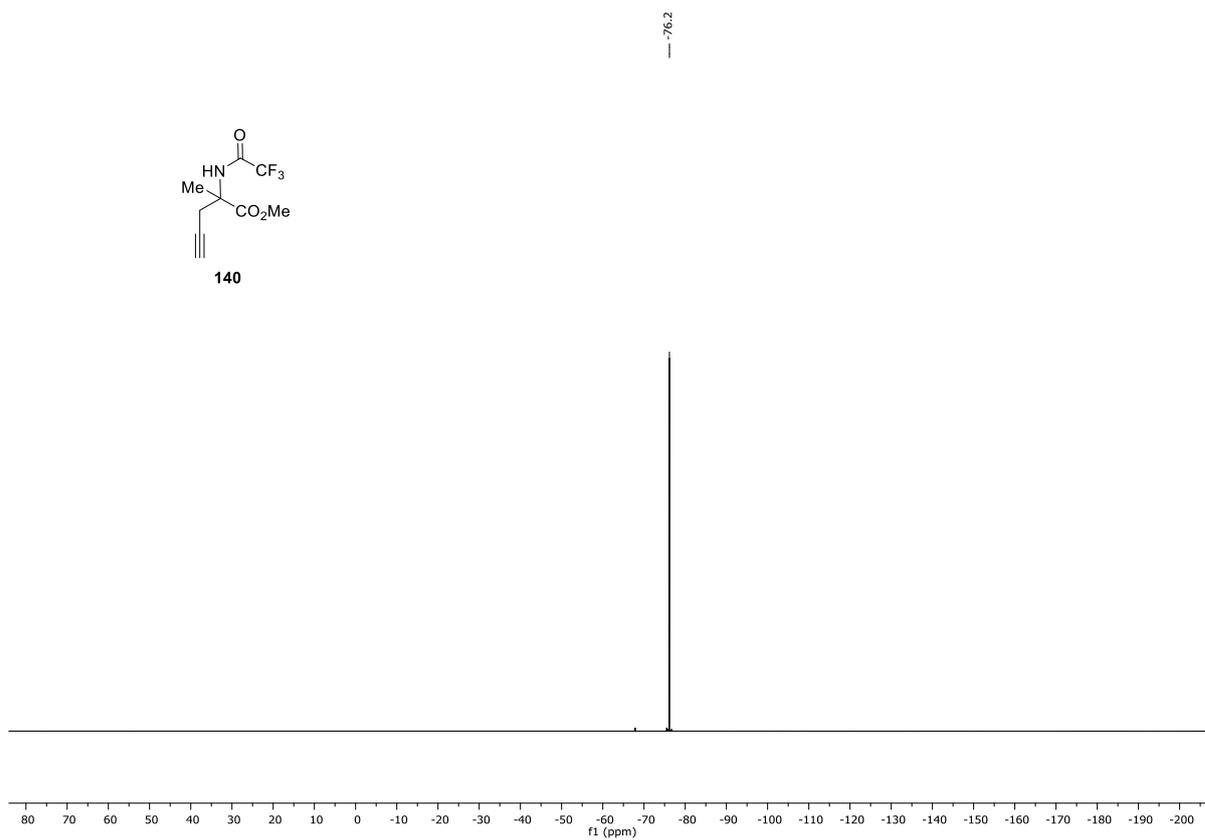












## 8.4 Publications



# Recent Applications of Diversity-Oriented Synthesis Toward Novel, 3-Dimensional Fragment Collections

Sarah L. Kidd, Thomas J. Osberger, Natalia Mateu, Hannah F. Sore and David R. Spring\*

Department of Chemistry, University of Cambridge, Cambridge, United Kingdom

Fragment-based drug discovery (FBDD) is a well-established approach for the discovery of novel medicines, illustrated by the approval of two FBDD-derived drugs. This methodology is based on the utilization of small “fragment” molecules (<300 Da) as starting points for drug discovery and optimization. Organic synthesis has been identified as a significant obstacle in FBDD, however, in particular owing to the lack of novel 3-dimensional (3D) fragment collections that feature useful synthetic vectors for modification of hit compounds. Diversity-oriented synthesis (DOS) is a synthetic strategy that aims to efficiently produce compound collections with high levels of structural diversity and three-dimensionality and is therefore well-suited for the construction of novel fragment collections. This Mini-Review highlights recent studies at the intersection of DOS and FBDD aiming to produce novel libraries of diverse, polycyclic, fragment-like compounds, and their application in fragment-based screening projects.

**Keywords:** fragment-based drug discovery, diversity-oriented synthesis, medicinal chemistry, organic synthesis, compound collections

## OPEN ACCESS

### Edited by:

Seung Bum Park,  
Seoul National University, South Korea

### Reviewed by:

Wei Zhang,  
University of Massachusetts Boston,  
United States  
Yu Zhou,  
Shanghai Institute of Materia Medica  
(CAS), China

### \*Correspondence:

David R. Spring  
spring@ch.cam.ac.uk

### Specialty section:

This article was submitted to  
Organic Chemistry,  
a section of the journal  
Frontiers in Chemistry

**Received:** 30 June 2018

**Accepted:** 14 September 2018

**Published:** 16 October 2018

### Citation:

Kidd SL, Osberger TJ, Mateu N,  
Sore HF and Spring DR (2018) Recent  
Applications of Diversity-Oriented  
Synthesis Toward Novel,  
3-Dimensional Fragment Collections.  
*Front. Chem.* 6:460.  
doi: 10.3389/fchem.2018.00460

## INTRODUCTION

Within the biomedical community there remains a pressing need for new molecules to seed early stage drug discovery programs. Diversity-oriented synthesis (DOS) emerged in the early 2000s in response to this challenge, a strategy which involves the efficient and deliberate construction of multiple scaffolds in a divergent manner (Lee et al., 2000; Schreiber, 2000; Spring, 2003; Burke and Schreiber, 2004). Nowadays, applications of this methodology span much of the spectrum of chemical space with examples describing the synthesis of fragment (Hung et al., 2011), small molecule (Wyatt et al., 2008; Lenci et al., 2015; Caputo et al., 2017), peptide (Kotha et al., 2013; Contreras-Cruz et al., 2017; Zhang et al., 2017) and macrocyclic (Isidro-Llobet et al., 2011; Kopp et al., 2012; Beckmann et al., 2013; Dow et al., 2017) collections all abundant within the literature. Furthermore, as the field of DOS has evolved, research themes have focused on addressing key calls from within the drug discovery community, namely the deficiencies within compound screening libraries (Lipkus et al., 2008; Dow et al., 2012), the identification of new bioactive molecules against challenging biological targets (Stanton et al., 2009; Kato et al., 2016; Kim et al., 2016) and populating underexplored areas of chemical space with novel structural entities (Thomas et al., 2008; Morton et al., 2009; Pizzirani et al., 2010). Until recently, however, the majority of DOS successes have been achieved in high-throughput screening (HTS) contexts (Chou et al., 2011; Laraia et al., 2014; Aldrich et al., 2015; Kuo et al., 2015).

Recent applications of DOS, however, exemplify how this methodology can be utilized to address significant challenges currently faced within the field of fragment-based drug discovery (FBDD). FBDD is now a widely adopted technique across both industry and academia, with two marketed drugs having emerged from this methodology [vemurafenib (Bollag et al., 2012), venetoclax (Souers et al., 2013)] and dozens of clinical candidates (Erlanson et al., 2016). This process involves the screening of small “fragment” molecule libraries (<300 Da) to identify efficient, but none the less weakly binding molecules, which are in turn subsequently elaborated to generate potent lead compounds (Erlanson and Jahnke, 2016). “Rule of three” guidelines are commonly employed within FBDD and library construction, relating to a molecular weight <300 Da, the number of hydrogen and acceptors/donors  $\leq 3$  and a cLogP  $\leq 3$  (Congreve et al., 2003). Importantly, due to the additional physicochemical constraints imposed on these screening libraries compared to traditional HTS approaches, it is broadly accepted that this method allows far more efficient sampling of chemical space, since there are far fewer possible fragment-sized molecules (Murray and Rees, 2009; Hall et al., 2014).

Despite significant advances in the foundational technologies of FBDD which have aided its implementation, reports associated with the synthetic intractability of hit fragments support the view that organic synthesis may be a rate-limiting step in the FBDD cycle, and in fact across drug discovery as a whole (Murray and Rees, 2016; Blakemore et al., 2018). In a similar vein to traditional drug discovery, deficiencies in commercially available fragment screening collections have been noted, in particular relating to the overrepresentation of  $sp^2$ -rich flat molecules (Hajduk et al., 2011; Hung et al., 2011) that feature limited numbers of synthetic handles for fragment elaboration. This latter feature is especially important within FBDD since this process relies on the merging, growth or linkage of small fragment molecules to develop initial weak hits (typically  $\mu M$  or  $mM$  range) hits into potent lead compounds. Without these vital functional handles, this process is significantly more time consuming, requiring the development of new synthetic routes to modify relatively simple fragment scaffolds. Furthermore, the incompatibility of many existing synthetic methodologies with amines, heterocycles, and unprotected polar functionalities limits their utilization. Consequently, there is a need for new strategies and technologies that enable non-traditional disconnections, late-stage functionalization as well as the incorporation of 3D elements into drug-like scaffolds.

Thus, appeals from within scientific community have been made for the development of novel and flexible synthetic methodologies that enable access to new fragments and their derivatives, including those with increased 3-dimensionality and heterocyclic architecture (Keseru et al., 2016; Murray and Rees, 2016). Despite the debates within the literature on the requirements of 3D character within fragment libraries, population of these underrepresented areas can be considered to complement existing flatter libraries, whilst providing access to alternative growth vectors, and therefore remains an important avenue of research (Morley et al., 2013; Fuller et al., 2016). From the perspective of library construction, the 3D character of the

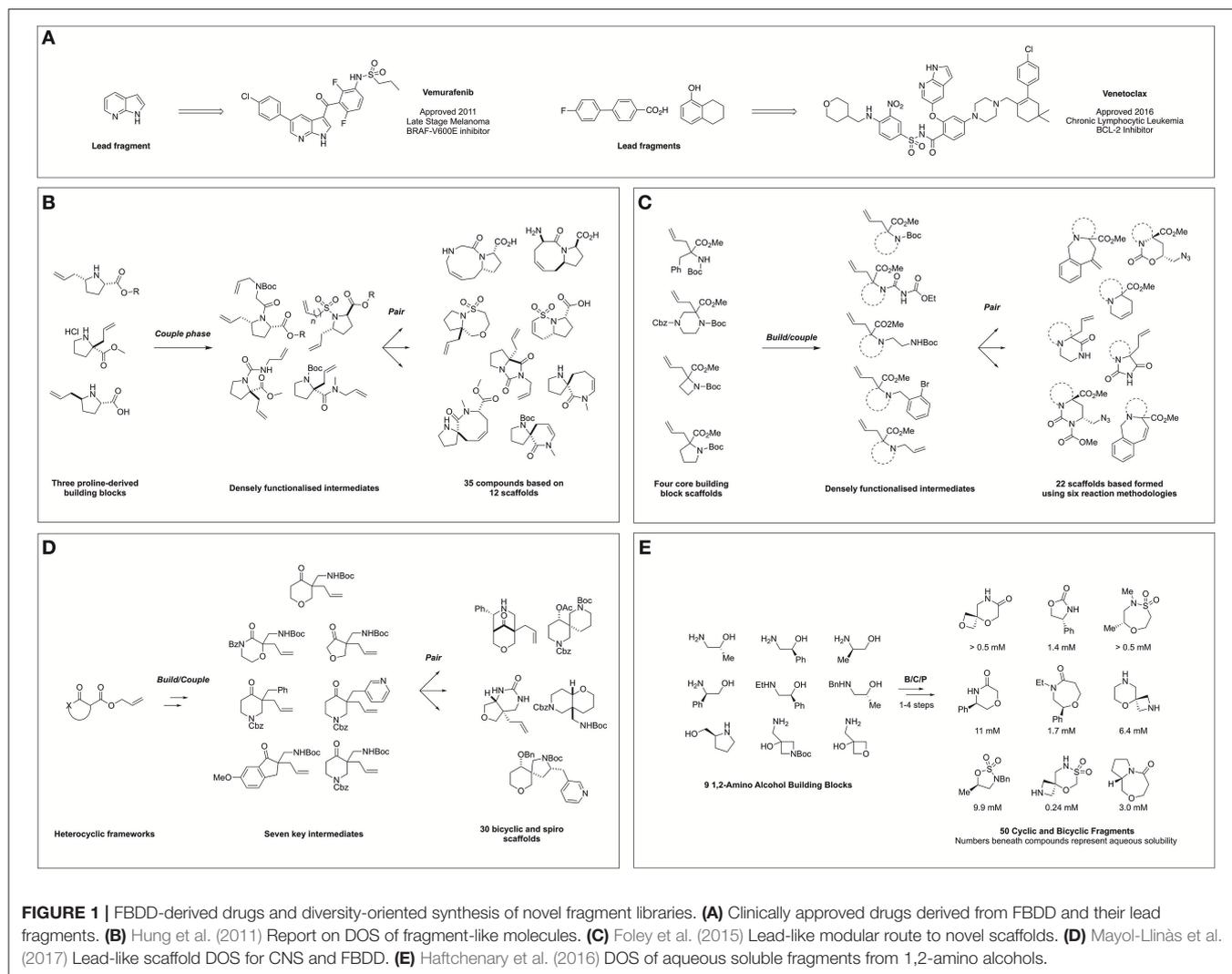
resulting libraries is commonly judged by the number of chiral centers and the fraction of  $sp^3$  carbons ( $F_{sp^3}$ ) within a molecule (Lovering et al., 2009), in addition to visual representations of the molecular shape space distribution using principal moment of inertia (PMI) analysis (Sauer and Schwarz, 2003; Kopp et al., 2012).

With a growing demand for novel heterocycles and 3D-shaped molecules for use within FBDD campaigns, many studies centering on the synthesis of 3D fragments around single heterocycles have been reported, for example using C-H activation methodologies (Davis et al., 2015; Palmer et al., 2016; Antermite et al., 2018). This mini-review aims to highlight the suitability of DOS approaches for addressing these challenges through the production of multiple scaffolds with a broader coverage of chemical space. One important feature of this strategy is the utilization of highly efficient and modular synthetic routes, commonly in the form of a build/couple/pair (B/C/P) algorithm (Nielsen and Schreiber, 2008). This involves (1) *the build phase*—construction of common starting materials, (2) *the couple phase*—intermolecular coupling of the building blocks with readily synthesized or commercial materials to form reactive intermediates and (3) *the pair phase*—intramolecular reaction or cyclisation of these precursors to afford distinct scaffolds. Thus, the flexibility of these strategies often results in methodologies that can provide efficient access to analogs of a desired scaffold. Herein, we discuss recent applications of the DOS strategy to the construction of novel and diverse 3D fragment collections and their applications in FBDD.

## THE APPLICATION OF DOS TO ACCESS NOVEL FRAGMENTS WITH MULTIPLE GROWTH VECTORS

The first publication conceptually merging DOS and FBDD appeared in 2011 in which Hung et al. (2011) described the application of DOS for the generation of a 3D fragment collection utilizing allyl proline-based precursors as the basis for library design. The researchers exploited three proline-derived building blocks in a B/C/P sequence to facilitate the formation of a series of fused and spiro bicyclic compounds (**Figure 1B**). This was achieved through installation of a second olefin *via* *N*-substitution using a variety of linker types, furnishing distinct linear precursors. Subsequently subjecting these intermediates to various intramolecular cyclizations such as ring closing metathesis (RCM) and oxo-Michael reactions, yielded 20 compounds based on 12 frameworks. Furthermore, due to the modular nature of this approach a complete matrix of stereoisomers of the 5–6, 5–7, 5–8, and 5–9 bicyclic frameworks could be constructed. Finally, the scaffolds were derivatized in a post-pair phase manner through functional group interconversion or olefin reduction to increase the diversity and the saturation, affording a total of 35 fragments.

Importantly, polar functional handles were installed throughout the library, enabling potential fragment growth from different vectors during hit-to-lead efforts. The applicability of the resultant library to FBDD was demonstrated *via*



chemoinformatic analysis, which highlighted rule of three compliance whilst principle moment of inertia (PMI) plots suggested a broad coverage of 3D molecular shape space.

Amino acid-derived reagents represent valuable building blocks for use within DOS methodologies owing to their polar and chiral nature, and their exploitation within these techniques has become more prevalent within the field. Work by Foley et al. (2015) described the application of four  $\alpha,\alpha$ -amino acid derived building blocks to generate a library of diverse bicyclic and tricyclic fragments (Foley et al., 2015). Through variation in the building block structure and the nature of the pair-phase cyclisation the researchers constructed 22 different heterocyclic scaffolds in a synthetically efficient manner (Figure 1C). Firstly, five different nitrogen substituents were installed on the four amino acid building blocks: a *tert*-butyl carbamate, an acyl urea, a 1,2-diamine, a *o*-bromobenzylamine or a second allyl olefin. In turn, pair phase reactions were then explored through reactivity of these functionalities with either the preinstalled ester or allyl moieties. This included iodine-mediated cyclisation

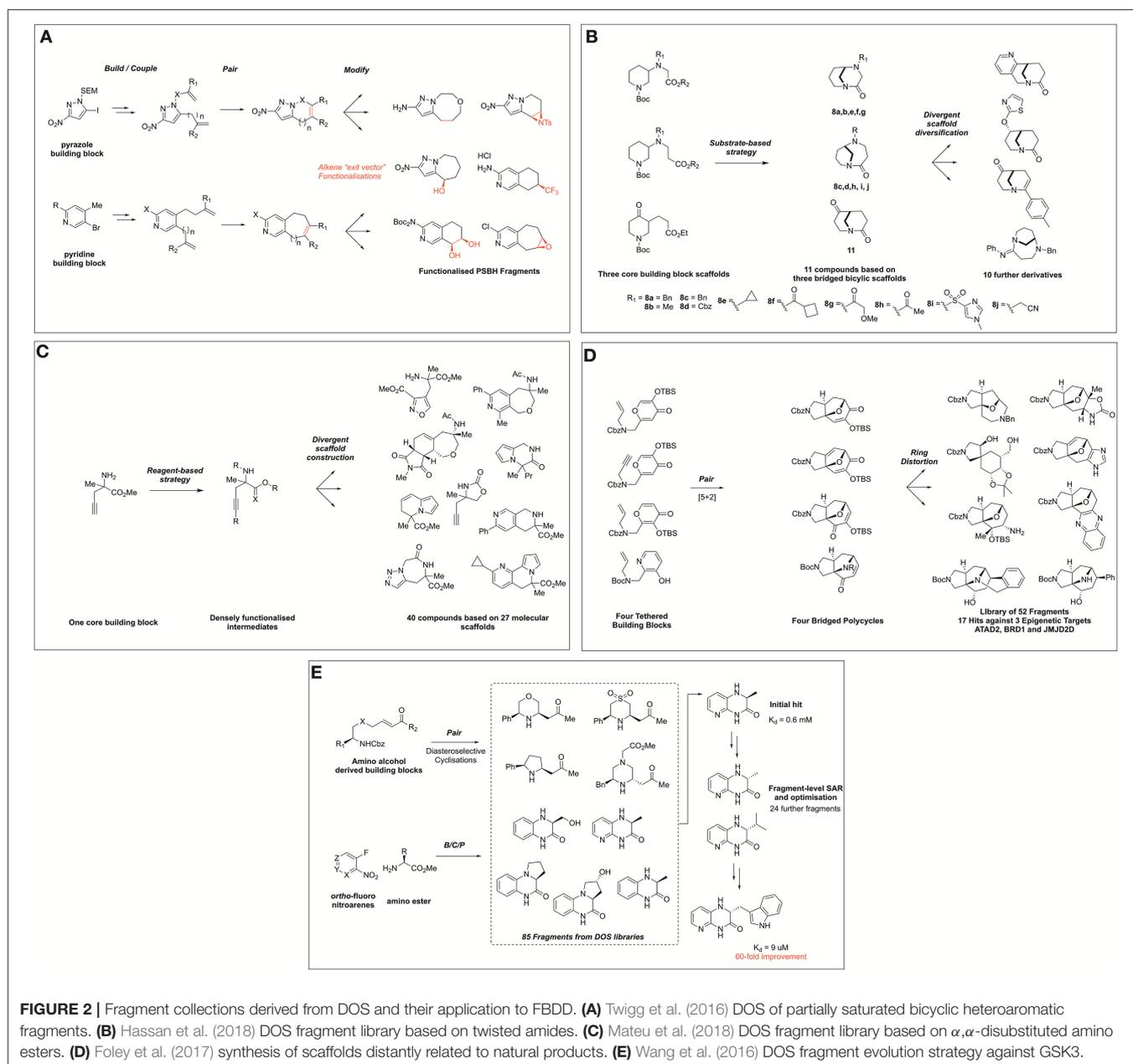
followed by azide addition and reactivity of the electrophilic ester moiety with *N*-based nucleophiles. Finally, ring closure *via* either Pd-mediated Heck reaction or Ru-mediated metathesis afforded further tri- and bicyclic fragments. The final collection of 22 scaffolds featured biologically relevant moieties such as ureas, hydantoin, and lactams, in addition to multiple functional synthetic handles. Subsequent virtual enumeration led to a library of 1,110 compounds that were predicted to possess lead-like properties and with considerable 3D character, (average  $Fsp^3 = 0.57$ ) and several examples meeting the criteria for FBDD.

In a similar vein, Mayol-Llinàs and co-workers also explored the use of cyclic  $\alpha$ -allyl quaternary ketones in a divergent and modular synthetic process to generate a library of 30 structurally distinct scaffolds featuring spiro, fused and bridged architectures (Figure 1D) (Mayol-Llinàs et al., 2017). Instead of amino acid-based precursors, Tsuji-type decarboxylative allylation was utilized to generate seven quaternary allylated building blocks. One example was selected for pilot studies, during which a variety of transformations were applied in a

reagent-based approach to yield 12 different scaffolds through exploitation of four key reactive moieties within the intermediate. This included an intramolecular Mannich reaction, a sequence of hydroboration-oxidation followed by either reduction or sulfonylation and then cyclisation, base-mediated cyclisation and Pd-catalyzed aminoarylation. Then, the remaining six precursors were subjected to the most promising conditions, yielding an additional 18 scaffolds. Virtual library enumeration was also conducted using six synthetic transformations, including reductive amination, urea formation and sulfonylation using 98 medicinal chemistry relevant capping groups. Multiparameter optimization analysis (Wager et al., 2010) was used to assess the amenability of this work to a CNS-based drug discovery

context. In addition, it was noted the resulting library possessed lead-like properties (Doveston et al., 2014) and that many of the compounds and derivatives would be applicable to a FBDD setting.

A recent report from Haftchenary at the Broad Institute detailed the synthesis of a fragment collection based on chiral 1,2-amino alcohols (Figure 1E) (Haftchenary et al., 2016). Beginning from a library of nine readily available amino alcohols, a range of 5-, 6-, and 7-membered scaffolds were synthesized in 1–4 steps using established synthetic procedures. The resulting fragment collection included medically important heterocycles such as oxazolidinones, morpholinones, and sulfamidate and sultam-based rings, along with fused and spiro-bicyclic compounds.



**FIGURE 2 |** Fragment collections derived from DOS and their application to FBDD. **(A)** Twigg et al. (2016) DOS of partially saturated bicyclic heteroaromatic fragments. **(B)** Hassan et al. (2018) DOS fragment library based on twisted amides. **(C)** Mateu et al. (2018) DOS fragment library based on  $\alpha,\alpha$ -disubstituted amino esters. **(D)** Foley et al. (2017) synthesis of scaffolds distantly related to natural products. **(E)** Wang et al. (2016) DOS fragment evolution strategy against GSK3.

Importantly for a screening context, the aqueous solubility of each of the 50 final fragments was measured, with values ranging from 0.085 to >15 mM, within the range for many fragment screening techniques.

Further to the goal of populating fragment space with  $sp^3$ -enriched compound collections that possess favorable fragment-like properties and synthetic exit vectors, Spring and coworkers disclosed a DOS-related approach to the synthesis of partially saturated bicyclic heteroaromatic (PSBH) molecules (**Figure 2A**) (Twigg et al., 2016). The synthetic route centered on the functionalization of pyrazole and pyridine-based building blocks featuring possessing amino-, or nitro groups, which were incorporated as potential solubilizing moieties; alternatively, a chloro substituent was incorporated as a hydrophobic element. The build and couple stages comprised of Suzuki cross coupling and alkylations to install various alkene functionalities, which were paired using RCM to afford bicyclic scaffolds, each featuring a positionally defined endocyclic alkene vector for further functionalization.

The endocyclic alkene was then modified in a post-pairing event to exemplify its utility as a synthetic growth vector in these fragments. A variety of alkene transformations, including dihalogenation, epoxidation, aziridination, cyclopropanation, halohydrin formation, and hydroboration-oxidation, were performed to yield a range of further functionalized scaffolds. The resulting library of compounds was then subjected to analysis of its physicochemical properties, which compared favorably to commercial screening libraries in relation to key properties such as number of chiral centers (0.88 vs. 0.27 or 0.18) and fraction aromatic (0.43 vs. 0.42 and 0.52), while maintaining rule of three compliance.

Hassan et al. recently disclosed an interesting example via the exploitation of twisted bicyclic amide compounds for the generation of a 3D fragment screening library (Hassan et al., 2018). In this work five 3-( $\omega$ -carboxylate)-substituted piperidine starting materials were manipulated to produce a 22-member polycyclic library (**Figure 2B**). Using substrate-based DOS methodology, five analogous starting materials based on three common structures were constructed and *via*  $Bu_2SnO$ -mediated cyclisation these were transformed to afford bicyclo[4.3.1]decane and bicyclo[3.3.1]nonane scaffolds in moderate yield. The generality of this methodology was exemplified through the synthesis of six further compounds through modification of the *N*-substituent the bicyclic ring systems.

In turn, these three key scaffolds were ultimately then divergently modified through manipulation of either the ketone or amide functionalities to generate a further nine compounds. The ketone moiety within the bicyclo[3.3.1]nonane was first modified by the use of either gold- or palladium-mediated reactions to afford tetra- or tricyclic heteroaromatic fused motifs. Alternatively, this moiety could also be reduced and a variety of heteroaromatics or alkyl moiety installed *via* either  $S_NAr$  or alkylation conditions in a diastereoselective fashion. Finally, the twisted amide within these scaffolds could also be manipulated to form either a chloroamine intermediate, followed by Suzuki-coupling to install an aryl substituent or simply by amidine formation. The resultant library was shown

to possess fragment lead-like properties with a high  $Fsp^3$  (0.63) and generally 17 or fewer heavy atoms and a  $clogP < 2.5$ . Furthermore, PMI analysis of the shape distribution suggested the library possessed significant 3D character to complement existing fragment collections for screening purposes.

The most recent and final example of synthetic efforts within this field by Mateu et al. (2018) report the use of  $\alpha,\alpha$ -disubstituted amino esters for the DOS of fragments incorporating a *N*-substituted quaternary carbon, an important and underrepresented motif within screening collections (**Figure 2C**). Using a single building block, 40 structurally diverse molecules based on 27 molecular frameworks were constructed in a synthetically efficient manner using an average of only three synthetic steps to access the entire library. This involved exploiting the three reactive handles within the building block in different combinations and utilizing a broad range of chemistries such as [2+2+2] cyclotrimerizations, Au-, Ru-, and Cu-mediated cyclizations and regioselective click chemistry to afford mono-, bi-, and tri-fused heterocycles featuring this important motif. Importantly, the authors also demonstrated the versatility of this synthetic methodology through the synthesis of an alternative quaternary *R*-substituent and the asymmetric synthesis of one library member.

Subsequent computational assessment of the resulting library *via* PMI analysis revealed a broad distribution of molecular shape space, in addition to favorable comparisons to a commercially available fragment collection in terms of 3-dimensional shape space coverage. Additionally, the mean values of the physicochemical properties of the library demonstrated the compatibility of the library for fragment screening, falling within the Rule of three guidelines, whilst exhibiting more favorable properties when again compared to existing commercial libraries. The authors note promising hits identified by X-ray fragment screening at the XChem screening facility, against proteins from three distinct families (a hydrolase, a TGF  $\beta$  growth factor, and a peptidase).

## DEMONSTRATION OF DOS METHODOLOGIES FOR THE IDENTIFICATION OF NOVEL BINDERS FOR CHALLENGING BIOLOGICAL TARGETS

In addition to populating new areas of fragment chemical space, DOS-derived fragment libraries can play a significant role in the identification of novel binders to seed future FBDD programs. Recent work by Foley et al. (2017) demonstrated the application of DOS-derived fragment libraries in the identification of novel hits against three epigenetic proteins from two distinct mechanistic classes (ATAD2, BRD1, and JMJD2D), *via* X-ray crystallographic screening methods. The researchers took inspiration from natural product frameworks, utilizing intramolecular [5+2] cyclizations to forge bridged structures incorporating natural product-related heteroaromatic frameworks (**Figure 2D**). Ring distortion reactions on these four initial structures using either expansion, cleavage, annulation, or substitution methodologies, were performed to divergently

modify the precursors, ultimately affording a library of 52 fragments based on 23 different scaffolds with bridged architectures and a high  $sp^3$  content. Interestingly, when this library was screened against the three epigenetic targets *via* high-throughput X-ray crystallography methods, 17 hits were identified against the three proteins, including those binding in novel regions of the proteins to those described previously. Moreover, comparisons could be drawn between the natural product-like fragment library and that obtained from commercial sources, whereby a significantly higher hit rate against ATAD2 was observed with the 3D fragments synthesized where seven hits were identified from a 52-member library vs. the commercially available fragment library where nine hits were identified from a 700-member library. Although the authors did not report any biophysical data for the fragments, the identification of novel X-ray hits from these efforts demonstrate the promise of a merged DOS-FBDD approach.

Finally, Young and co-workers recently demonstrated the successful use of the DOS strategy to optimize fragments against the serine/threonine kinase GSK3 $\beta$  (Wang et al., 2016), which is overexpressed in cancer and Alzheimer's disease (Luo, 2009; Hernandez et al., 2012). To initiate the investigation, a set of 86 fragments was compiled from DOS libraries constructed *via* three distinct B/C/P pathways (Figure 2E). The first DOS fragment library utilized allylproline building blocks and has been previously discussed in this review (Figure 1A). The second DOS library coupled enones with amino alcohol and related building blocks. The final scaffolds were accessed *via* catalytic, diastereoselective aza-Michael additions to afford stereochemically diverse disubstituted heterocycles. The third DOS library incorporated into this study was generated from *ortho*-nitrofluoro arenes and  $\alpha$ -amino ester building blocks. Intermolecular coupling products were obtained *via*  $S_NAr$ , and pairing products were accessed by reduction of the nitro group followed by spontaneous cyclisation onto the ester functionality. This modular approach yielded a small collection of enantiomerically enriched bicyclic piperazinone compounds.

Using this fragment collection, screening against GSK3 $\beta$  was performed using differential scanning fluorimetry (DSF) to detect fragment binding. Initial results identified a benzopiperizinone-library member to exhibit good thermal stabilization and subsequent assays showed 46% inhibition of GSK3 $\beta$  at 1 mM concentration. A library of derivatives based on this initial hit were then synthesized using the modular and rapid DOS chemistry initially developed. Thus, the single enantiomer variants and other derivatives could easily be constructed to generate structure-activity relationships (SAR). Preliminary fragment-level SAR indicated the (*R*)-enantiomer of the chiral center to be more potent, and further studies identified the substituent at this site as an important potential growth vector. Fragment growth by incorporating large aryl groups into the scaffold *via* the same B/C/P pathway yielded the lead compound, with a large indolyl unit connected to the core heterocycle. This fragment exhibited a  $K_d = 9 \mu M$ , a 60-fold improvement over

the initial fragment hit. Ultimately an X-ray crystal structure of the lead compound with GSK3 $\beta$  was obtained, revealing it binds in the ATP pocket of this kinase.

This study demonstrated the successful implementation of a DOS-based FBDD workflow to evolve fragments against an important kinase target. Key to the success of this project was the utility of the DOS concept as a tool to generate skeletally and stereochemically diverse initial libraries, and later as an efficient, modular route to analogs for SAR and fragment growth.

## FUTURE PERSPECTIVES

The studies discussed herein have demonstrated the utility of DOS as an effective approach for populating new areas of fragment space, in areas largely complementary to existing fragment collections. In each case, the resulting libraries featured high structural and shape diversity, increased 3D character and exemplified synthetic vectors for fragment growth. The latter two examples discussed detail applications of these libraries for the identification of novel fragment binders and inhibitors against challenging protein targets, ultimately demonstrating the utility of DOS within drug discovery efforts.

It is worth noting the increasing application of computational virtual library enumeration, an element of which has featured in several of the publications discussed. It is envisioned that these methodologies will only increase in their utility when coupled to *in silico*-based screening techniques to guide library design and prioritization of synthesis. Moreover, a focus on applications of newly developed methodologies to DOS, for example C-H activation, and site selective late-stage modifications of complex scaffolds would enable population of underexplored areas of chemical space and further derivatization of the resulting scaffolds. Finally, an outstanding requirement within this field is the establishment of new translational collaborations between academic and industrial groups to enable the routine screening of the novel libraries.

## AUTHOR CONTRIBUTIONS

SLK and TJO conceived and wrote the manuscript. All other authors (NM, HFS and DRS) provided comments and discussion on the manuscript to aid its preparation.

## FUNDING

Our research is supported by the EPSRC, BBSRC, MRC, Wellcome Trust, and ERC (FP7/2007-2013; 279337/DOS). SK thanks AstraZeneca for funding.

## ACKNOWLEDGMENTS

The authors would like to thank Dr D. Twigg from Astex for useful discussions on the topic.

## REFERENCES

- Aldrich, L. N., Kuo, S. Y., Castoreno, A. B., Goel, G., Kuballa, P., Rees, M. G., et al. (2015). Discovery of a small-molecule probe for V-ATPase function. *J. Am. Chem. Soc.* 137, 5563–5568. doi: 10.1021/jacs.5b02150
- Antermite, D., Dominic, P. A., and Bull, J. A. (2018). Regio- and stereoselective palladium-catalyzed C(sp<sup>3</sup>)-H arylation of pyrrolidines and piperidines with C(3) directing groups. *Organ. Lett.* 20, 3948–3952. doi: 10.1021/acs.orglett.8b01521
- Beckmann, H. S., Nie, F., Hagerman, C. E., Johansson, H., Tan, Y. S., Wilcke, D., et al. (2013). A strategy for the diversity-oriented synthesis of macrocyclic scaffolds using multidimensional coupling. *Nat. Chem.* 5, 861–867. doi: 10.1038/nchem.1729
- Blakemore, D. C., Castro, L., Churcher, I., Rees, D. C., Thomas, A. W., Wilson, D. M., et al. (2018). Organic synthesis provides opportunities to transform drug discovery. *Nat. Chem.* 10, 383–394. doi: 10.1038/s41557-018-0021-z
- Bollag, G., Tsai, J., Zhang, J., Zhang, C., Ibrahim, P., Nolop, K., et al. (2012). Vemurafenib: the first drug approved for BRAF-mutant cancer. *Nat. Rev. Drug Discov.* 11, 873–886. doi: 10.1038/nrd3847
- Burke, M. D., and Schreiber, S. L. (2004). A planning strategy for diversity-oriented synthesis. *Angew. Chem. Int. Ed.* 43, 46–58. doi: 10.1002/anie.200300626
- Caputo, S., Banfi, L., Basso, A., Galatini, A., Moni, L., Riva, R., et al. (2017). Diversity-oriented synthesis of various enantiopure heterocycles by coupling organocatalysis with multicomponent reactions. *Eur. J. Organ. Chem.* 2017, 6619–6628. doi: 10.1002/ejoc.201701328
- Chou, D. H. C., Duvall, J. R., Gerard, B., Liu, H., Pandya, B. A., Suh, B. C., et al. (2011). Synthesis of a novel suppressor of  $\beta$ -cell apoptosis via diversity-oriented synthesis. *ACS Med. Chem. Lett.* 2, 698–702. doi: 10.1021/ml200120m
- Congreve, M., Carr, R., Murray, C., and Jhoti, H. (2003). A 'rule of three' for fragment-based lead discovery? *Drug Discov. Today* 8, 876–877. doi: 10.1016/S1359-6446(03)02831-9
- Contreras-Cruz, D. A., Sánchez-Carmona, M. A., Vengoechea-Gómez, F. A., Peña-Ortiz, D., and Miranda, L. D. (2017). Diversity-oriented synthesis of cyclopropyl peptides from Ugi-derived dehydroalanines. *Tetrahedron* 73, 6146–6156. doi: 10.1016/j.tet.2017.09.005
- Davis, O. A., Croft, R. A., and Bull, J. A. (2015). Synthesis of diversely functionalised 2,2-disubstituted oxetanes: fragment motifs in new chemical space. *Chem. Commun.* 51, 15446–15449. doi: 10.1039/C5CC05740J
- Doveston, R., Marsden, S., and Nelson, A. (2014). Towards the realisation of lead-oriented synthesis. *Drug Discov. Today* 19, 813–819. doi: 10.1016/j.drudis.2013.11.006
- Dow, M., Fisher, M., James, T., Marchetti, F., and Nelson, A. (2012). Towards the systematic exploration of chemical space. *Organ. Biomol. Chem.* 10, 17–28. doi: 10.1039/C1OB06098H
- Dow, M., Marchetti, F., Abrahams, K. A., Vaz, L., Besra, G. S., Warriner, S., et al. (2017). Modular synthesis of diverse natural product-like macrocycles: discovery of hits with antimycobacterial activity. *Chem. A Eur. J.* 23, 7207–7211. doi: 10.1002/chem.201701150
- Erlanson, D. A., Fesik, S. W., Hubbard, R. E., Jahnke, W., and Jhoti, H., (2016). Twenty years on: the impact of fragments on drug discovery. *Nat. Rev. Drug Discov.* 15, 605–619. doi: 10.1038/nrd.2016.109
- Erlanson, D. A., and Jahnke, W. (2016). *Fragment-Based Drug Discovery Lessons and Outlook*. Weinheim: Wiley-VCH Verlag GmbH and Co.
- Foley, D., Doveston, R., Churcher, I., Nelson, A., and Marsden, S. P. (2015). A systematic approach to diverse, lead-like scaffolds from  $\alpha,\alpha$ -disubstituted amino acids. *Chem. Commun.* 51, 11174–11177. doi: 10.1039/C5CC03002A
- Foley, D. J., Craven, P. G. E., Collins, P. M., Doveston, R. G., Aimon, A., Talon, R., et al. (2017). Synthesis and demonstration of the biological relevance of sp<sup>3</sup>rich Scaffolds distantly related to natural product frameworks. *Chem. A Eur. J.* 23, 15227–15232. doi: 10.1002/chem.201704169
- Fuller, N., Spadola, L., Cowen, S., Patel, J., Schonherr, H., Cao, Q., et al. (2016). An improved model for fragment-based lead generation at AstraZeneca. *Drug Discov. Today* 21, 1272–1283. doi: 10.1016/j.drudis.2016.04.023
- Haftchenary, S., Nelson, S. D., Furst, L., Dandapani, S., Ferrara, S. J., Bošković, Ž. V., et al. (2016). Efficient routes to a diverse array of amino alcohol-derived chiral fragments. *Comb. Sci.* 18, 569–574. doi: 10.1021/acscombsci.6b00050
- Hajduk, P. J., Galloway, W. R., and Spring, D. R. (2011). Drug discovery: a question of library design. *Nature* 470, 42–43. doi: 10.1038/470042a
- Hall, R. J., Mortenson, P. N., and Murray, C. W. (2014). Efficient exploration of chemical space by fragment-based screening. *Prog. Biophys. Mol. Biol.* 116, 82–91. doi: 10.1016/j.pbiomolbio.2014.09.007
- Hassan, H., Marsden, S. P., and Nelson, A. (2018). Design and synthesis of a fragment set based on twisted bicyclic lactams. *Bioorgan. Med. Chem.* 26, 3030–3033. doi: 10.1016/j.bmc.2018.02.027
- Hernandez, F., Lucas, J. J., and Avila, J. (2012). GSK3 and tau: two convergence points in Alzheimer's disease. *J. Alzheimer's Dis.* 33, S141–S144. doi: 10.3233/JAD-2012-129025
- Hung, A. W., Ramek, A., Wang, Y., Kaya, T., Wilson, J. A., Clemons, P. A., et al. (2011). Route to three-dimensional fragments using diversity-oriented synthesis. *Proc. Natl. Acad. Sci. U.S.A.* 108, 6799–6804. doi: 10.1073/pnas.1015271108
- Isidro-Llobet, A., Murillo, T., Bello, P., Cilibrizzi, A., Hodgkinson, J. T., Galloway, W. R. J. D., et al. (2011). Diversity-oriented synthesis of macrocyclic peptidomimetics. *Proc. Natl. Acad. Sci. U.S.A.* 108, 6793–6798. doi: 10.1073/pnas.1015267108
- Kato, N., Comer, E., Sakata-Kato, T., Sharma, A., Sharma, M., Maetani, M., et al. (2016). Diversity-oriented synthesis yields novel multistage antimalarial inhibitors. *Nature* 538, 344–349. doi: 10.1038/nature19804
- Keseru, G. M., Erlanson, D. A., Ferenczy, G. G., Hann, M. M., Murray, C. W., and Pickett, S. D. (2016). Design principles for fragment libraries: maximizing the value of learnings from pharma Fragment-Based Drug Discovery (FBDD) programs for use in academia. *J. Med. Chem.* 59, 8189–8206. doi: 10.1021/acs.jmedchem.6b00197
- Kim, J., Jung, J., Koo, J., Cho, W., Lee, W. S., Kim, C., et al. (2016). Diversity-oriented synthetic strategy for developing a chemical modulator of protein-protein interaction. *Nat. Commun.* 7:13196. doi: 10.1038/ncomms13196
- Kopp, F., Stratton, C. F., Akella, L. B., and Tan, D. S. (2012). A diversity-oriented synthesis approach to macrocycles via oxidative ring expansion. *Nat. Chem. Biol.* 8, 358–365. doi: 10.1038/nchembio.911
- Kotha, S., Goyal, D., and Chavan, A. S. (2013). Diversity-oriented approaches to unusual  $\alpha$ -amino acids and peptides: step economy, atom economy, redox economy, and beyond. *J. Organ. Chem.* 78, 12288–12313. doi: 10.1021/jo4020722
- Kuo, S.-Y., Castoreno, A. B., Aldrich, L. N., Lassen, K. G., Goel, G., Dančik, V., et al. (2015). Small-molecule enhancers of autophagy modulate cellular disease phenotypes suggested by human genetics. *Proc. Natl. Acad. Sci. U.S.A.* 112, 4281–e4287. doi: 10.1073/pnas.1512289112
- Laraia, L., Stokes, J., Emery, A., McKenzie, G. J., Venkitaraman, A. R., and Spring, D. R. (2014). High content screening of diverse compound libraries identifies potent modulators of tubulin dynamics. *ACS Med. Chem. Lett.* 5, 598–603. doi: 10.1021/ml5000564
- Lee, D., Sello, J. K., and Schreiber, S. L. (2000). Pairwise use of complexity-generating reactions in diversity-oriented organic synthesis. *Organ. Lett.* 2, 709–712. doi: 10.1021/ol005574n
- Lenci, E., Menchi, G., Guarna, A., and Trabocchi, A. (2015). Skeletal diversity from carbohydrates: use of mannose for the diversity-oriented synthesis of polyhydroxylated compounds. *J. Organ. Chem.* 80, 2182–2191. doi: 10.1021/jo502701c
- Lipkus, A. H., Yuan, Q., Lucas, K. A., Funk, S. A., Bartelt, W. F., Schenck, R. J., et al. (2008). Structural diversity of organic chemistry. A Scaffold Analysis of the CAS Registry. *Organ. Chem.* 73, 4443–4451. doi: 10.1021/jo801276
- Lovering, F., Bikker, J., and Humblet, C. (2009). Escape from flatland: increasing saturation as an approach to improving clinical success. *J. Med. Chem.* 52, 6752–6756. doi: 10.1021/jm901241e
- Luo, J. (2009). Glycogen synthase kinase 3beta (GSK3beta) in tumorigenesis and cancer chemotherapy. *Cancer Lett.* 273, 194–200. doi: 10.1016/j.canlet.2008.05.045
- Mateu, N., Kidd, S. L., Kalash, L., Sore, H. F., Madin, A., Bender, A., et al. (2018). Synthesis of structurally diverse N-substituted quaternary-carbon-containing small molecules from  $\alpha,\alpha$ -disubstituted propargyl amino esters. *Chem. Eur. J.* 24, 13681–13687. doi: 10.1002/chem.201803143
- Mayol-Llinàs, J., Farnaby, W., and Nelson, A. (2017). Modular synthesis of thirty lead-like scaffolds suitable for CNS drug discovery. *Chem. Commun.* 53, 12345–12348. doi: 10.1039/C7CC06078E

- Morley, A. D., Pugliese, A., Birchall, K., Bower, J., Brennan, P., Brown, N., et al. (2013). Fragment-based hit identification: thinking in 3D. *Drug Discov. Today* 18, 1221–1227. doi: 10.1016/j.drudis.2013.07.011
- Morton, D., Leach, S., Cordier, C., Warriner, S., and Nelson, A. (2009). Synthesis of natural-product-like molecules with over eighty distinct scaffolds. *Angew. Chem. Int. Ed.* 48, 104–109. doi: 10.1002/anie.200804486
- Murray, C. W., and Rees, D. C. (2009). The rise of fragment-based drug discovery. *Nat. Chem.* 1, 187–192. doi: 10.1038/nchem.217
- Murray, C. W., and Rees, D. C. (2016). Opportunity knocks: organic chemistry for Fragment-Based Drug Discovery (FBDD). *Angew. Chem. Int. Ed.* 55, 488–492. doi: 10.1002/anie.201506783
- Nielsen, T. E., and Schreiber, S. L. (2008). Towards the optimal screening collection: a synthesis strategy. *Angew. Chem. Int. Ed.* 47, 48–56. doi: 10.1002/anie.200703073
- Palmer, N., Peakman, T. M., Norton, D., and Rees, D. C. (2016). Design and synthesis of dihydroisoquinolones for fragment-based drug discovery (FBDD). *Organ. Biomol. Chem.* 14, 1599–1610. doi: 10.1039/C5OB02461G
- Pizzirani, D., Kaya, T., Clemons, P. A., and Schreiber, S. L. (2010). Stereochemical and skeletal diversity arising from amino propargylic alcohols. *Organ. Lett.* 12, 2822–2825. doi: 10.1021/ol100914b
- Sauer, W. H. B., and Schwarz, M. K. (2003). Molecular shape diversity of combinatorial libraries: a prerequisite for broad bioactivity. *J. Chem. Inf. Comput. Sci.* 43, 987–1003. doi: 10.1021/ci025599w
- Schreiber, S. L. (2000). Target-oriented and diversity-oriented organic synthesis in drug discovery. *Science* 287, 1964–1969. doi: 10.1126/science.287.5460.1964
- Souers, A. J., Levenson, J. D., Boghaert, E. R., Ackler, S. L., Catron, N. D., Chen, J., et al. (2013). ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets. *Nat. Med.* 19, 202–208. doi: 10.1038/nm.3048
- Spring, D. R. (2003). Diversity-oriented synthesis; a challenge for synthetic chemists. *Organ. Biomol. Chem.* 1, 3867–3870. doi: 10.1039/b310752n
- Stanton, B. Z., Peng, L. F., Maloof, N., Nakai, K., Wang, X., Duffner, J. L., et al. (2009). A small molecule that binds Hedgehog and blocks its signaling in human cells. *Nat. Chem. Biol.* 5, 154–156. doi: 10.1038/nchembio.142
- Thomas, G. L., Spandl, R. J., Glansdorp, F. G., Welch, M., Bender, A., Cockfield, J., et al. (2008). Anti-MRSA agent discovery using diversity-oriented synthesis. *Angew. Chem. Int. Ed.* 47, 2808–2812. doi: 10.1002/anie.200705415
- Twigg, D. G., Kondo, N., Mitchell, S. L., Galloway, W. R. J. D., Sore, H. F., Madin, A., et al. (2016). Partially saturated bicyclic heteroaromatics as an sp<sup>3</sup>-enriched fragment collection. *Angew. Chem. Int. Ed.* 55, 12479–12483. doi: 10.1002/anie.201606496
- Wager, T. T., Hou, X., Verhoest, P. R., and Villalobos, A. (2010). Moving beyond rules: the development of a central nervous system multiparameter optimization (CNS MPO) approach to enable alignment of druglike properties. *ACS Chem. Neurosci.* 1, 435–449. doi: 10.1021/cn100008c
- Wang, Y., Wach, J. Y., Sheehan, P., Zhong, C., Zhan, C., Harris, R., et al. (2016). Diversity-oriented synthesis as a strategy for fragment evolution against GSK3 $\beta$ . *ACS Med. Chem. Lett.* 7, 852–856. doi: 10.1021/acsmchemlett.6b00230
- Wyatt, E., Galloway, W. R. J. D., Thomas, G., Welch, M., Loiseleur, O., Plowright, A., et al. (2008). Identification of an anti-MRSA dihydrofolate reductase inhibitor from a diversity-oriented synthesis. *Chem. Commun.* 40, 4962–4964. doi: 10.1039/b812901k
- Zhang, J., Mulumba, M., Ong, H., and Lubell, W. D. (2017). Diversity-oriented synthesis of cyclic azapeptides by A3-macrocyclization provides high-affinity CD36-modulating peptidomimetics. *Angew. Chem. Int. Ed.* 56, 6284–6288. doi: 10.1002/anie.201611685

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Kidd, Osberger, Mateu, Sore and Spring. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

## Synthetic Methods

Synthesis of Structurally Diverse N-Substituted Quaternary-Carbon-Containing Small Molecules from  $\alpha,\alpha$ -Disubstituted Propargyl Amino EstersNatalia Mateu<sup>+,\* [a]</sup> Sarah L. Kidd<sup>+, [a]</sup> Leen Kalash,<sup>[a]</sup> Hannah F. Sore,<sup>[a]</sup> Andrew Madin,<sup>[b]</sup> Andreas Bender,<sup>[a]</sup> and David R. Spring<sup>\*, [a]</sup>

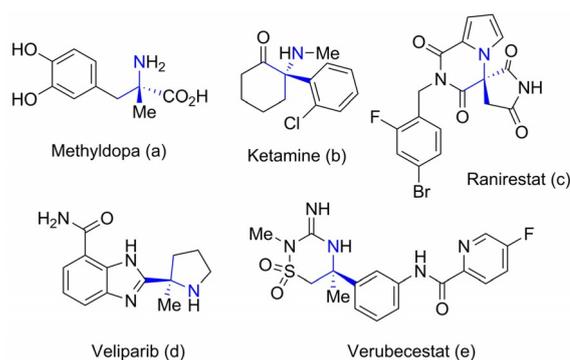
Dedicated to Sir Jack E. Baldwin on the occasion of his 80th birthday

**Abstract:** N-containing quaternary stereocenters represent important motifs in medicinal chemistry. However, due to their inherently sterically hindered nature, they remain under-represented in small molecule screening collections. As such, the development of synthetic routes to generate small molecules that incorporate this particular feature are highly desirable. Herein, we describe the diversity-oriented synthesis (DOS) of a diverse collection of structurally distinct small

molecules featuring this three-dimensional (3D) motif. The subsequent derivatisation and the stereoselective synthesis exemplified the versatility of this strategy for drug discovery and library enrichment. Chemoinformatic analysis revealed the enhanced  $sp^3$  character of the target library and demonstrated that it represents an attractive collection of biologically diverse small molecules with high scaffold diversity.

## Introduction

N-containing quaternary stereocenters are important motifs in medicinal chemistry and are present in significant essential medicines including the antihypertensive methyldopa (Aldomet<sup>®</sup>) and the anaesthetic ketamine (Ketalar<sup>®</sup>) (Figure 1 a,b).<sup>[1]</sup> The presence of this particular stereocenter in the three small molecules currently under clinical evaluation, Ranirestat,<sup>[2]</sup> Veliparib<sup>[3]</sup> and Verubecestat<sup>[4]</sup> (Figure 1 c–e), further highlight its relevance in drug discovery contexts. The biological activity of such compounds has been generally shown to be intrinsically related to their absolute configuration.<sup>[5–10]</sup> For instance, investigations have concluded that the presence of a N-containing quaternary (*R*)-stereocenter within the core structure is necessary for interaction of the substituents with the pockets of the



**Figure 1.** Examples of bioactive compounds containing a N-substituted quaternary carbon.

$\beta$ -secretase (BACE1) binding site.<sup>[9,10]</sup> Moreover, the presence of such restricted elements within a small molecule can provide conformational restrictions owing to their sterically hindered nature, thereby increasing the molecular complexity of a given molecule, which has been shown to be desirable chemical feature for bioactive molecules.<sup>[11,12]</sup>

Studies have shown that increased complexity (e.g.  $sp^3$ -rich cycles or quaternary carbons) can correlate to an increase in the selectivity, potency and metabolic stability of drug candidates, along with the successful progression from clinical stages to drug approval.<sup>[13,14]</sup> However, due to their congested nature, the incorporation of N-containing quaternary stereocenters into small molecules poses a synthetic challenge for organic chemists.<sup>[15–19]</sup> As a result, small molecules bearing this particular feature are still under-represented in probe and drug discovery screening collections.

[a] Dr. N. Mateu,<sup>+</sup> S. L. Kidd,<sup>+</sup> L. Kalash, Dr. H. F. Sore, Dr. A. Bender, Prof. D. R. Spring  
Department of Chemistry, University of Cambridge  
Lensfield Rd, Cambridge, CB2 1EW (UK)  
E-mail: nm462@cam.ac.uk  
spring@ch.cam.ac.uk

[b] Dr. A. Madin  
AstraZeneca (UK) Ltd., 310 Cambridge Science Park  
Milton Rd, Cambridge, CB4 0FZ (UK)

[\*] These authors contributed equally to this work.

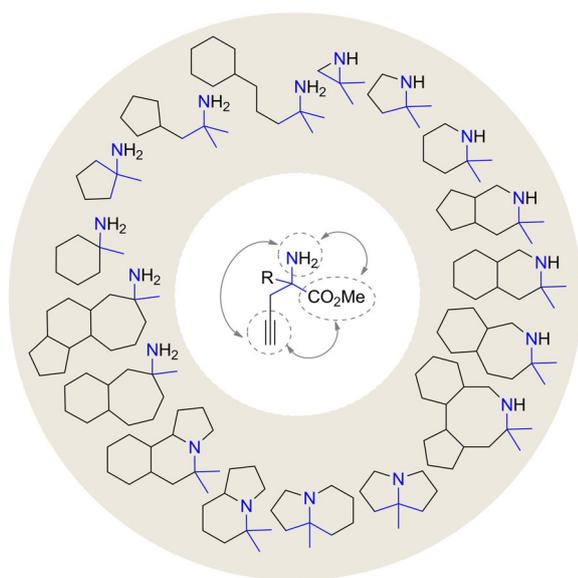
Supporting information and the ORCID identification numbers for the authors of this article can be found under:  
<https://doi.org/10.1002/chem.201803143>.

© 2018 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

From a synthetic perspective, despite all the advances made in the field of asymmetric catalysis to generate  $\alpha$ -branched amino-containing compounds,<sup>[16–19]</sup> the majority of the reported reactions are based on simple scaffolds with restricted structural diversity and there has been a limited exploration into the construction of such motifs into more complex molecules. It has, however, been demonstrated that the use of  $\alpha,\alpha$ -disubstituted amino esters as subunits to prepare complex molecules bearing this sterically hindered motif can be particularly effective. Thus, some examples in the literature have validated this strategy by the target-oriented synthesis (TOS) of biologically active small molecules.<sup>[20]</sup> More recently, allyl-containing cyclic quaternary amino esters have also been used for the diversity-oriented synthesis (DOS) of small molecule collections.<sup>[21,22]</sup> However, structurally diverse,  $sp^3$ -rich and complex libraries are still under-represented in probe and drug discovery screening collections.

In addition to these deficiencies, within the literature, significant synthetic challenges have been identified in relation to the development of chemistries which tolerate polar functionalities such as amines, N-heterocycles and unprotected polar groups.<sup>[23]</sup> Moreover, within a fragment-based drug discovery context, there have been recent calls for the design and synthesis of novel compounds that feature multiple 3D growth vectors whilst incorporating polar functionality for molecular recognition.<sup>[24]</sup> In response to this, notable examples of efficient syntheses of diverse  $sp^3$ -rich libraries have been recently published,<sup>[25–32]</sup> however there is still an outstanding need for diverse libraries featuring this key N-containing quaternary stereocenter motif.

With these points in mind, we envisioned that  $\alpha,\alpha$ -disubstituted propargyl amino esters could act as a versatile and pluri-potent platform for the DOS of structurally diverse screening collection featuring a stereogenic N-containing quaternary carbon (Figure 2). In this manner, the reactivity of its three



**Figure 2.** Synthetic versatility of  $\alpha,\alpha$ -disubstituted amino esters for the DOS of N-substituted quaternary carbon containing small molecules.

functional vectors, namely the terminal alkyne, the amine as well and the ester moieties could be exploited in both inter- and intramolecular reactions.<sup>[33]</sup>

## Results and Discussion

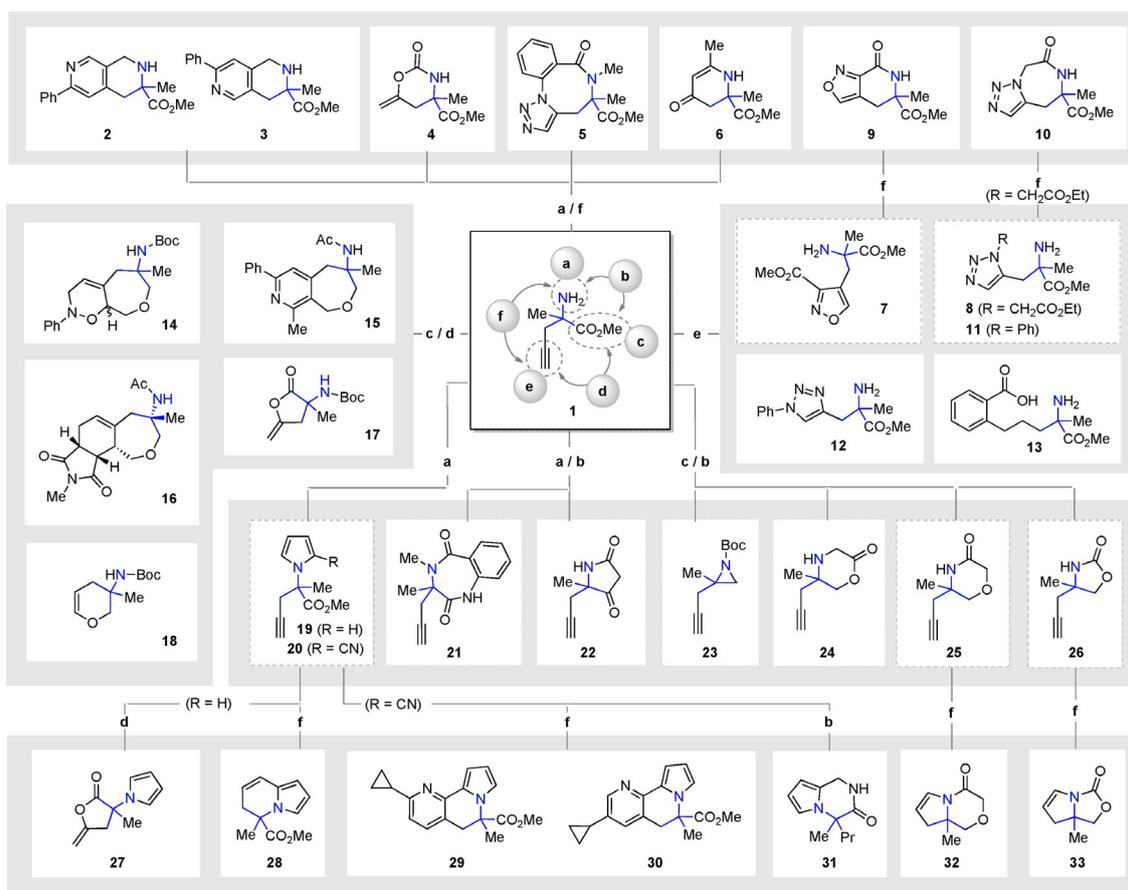
### Synthesis of the DOS library

To validate our hypothesis, we initially selected the racemic  $\alpha$ -methyl propargylglycine **1** ( $R=Me$ ) as a model substrate to perform the reagent-based DOS outlined in this work (Scheme 1). Our investigations began via alkylation of the amino group in **1** through the introduction of a *tert*-butyl carbamate, propargyl, 2-azido benzoyl or an acyl group to generate the first set of highly functionalised intermediates (**S2–6**, for all intermediates see Scheme S1). Subsequent pairing of the functionalities present in these intermediates with the terminal alkyne could be achieved through different metal-catalysed cyclisations (Scheme 1, steps a/f). Accordingly, Co-induced [2+2+2] cyclootrimerisation afforded **2** and **3**, whilst hydroamination using gold catalysis yielded **4** and **6**. Finally, intramolecular Ru-mediated click chemistry gave rise to compound **5**.

In a similar fashion, Ru-mediated 1,5-click chemistry could be used to generate acyclic compounds **7** and **8**, through intermolecular reaction of a protected form of **1** with the commercially available chloro-oxime or the corresponding azido ester. In turn, lactamisation of the ester moieties within **7** and **8** with the quaternary amine (Scheme 1, steps e/f) afforded the rigidified scaffolds **9** and **10**. Further exploitation of the terminal alkyne handle within **1** via Ru- or Cu-catalysis afforded the acyclic 1,5- and 1,4-phenyl substituted triazoles **11** and **12**, respectively. Finally, Sonogashira cross-coupling of the terminal alkyne followed by hydrogenation afforded **13** (Scheme 1, step e).

Next, an exploration into intramolecular cyclisations between the ester functionality within **1** and the terminal alkyne was undertaken (Scheme 1, steps c/d). Reduction of the carboxylic moiety into the corresponding alcohol provided the opportunity to introduce additional unsaturated moieties through the alkylation of the newly generated hydroxyl group to afford the allylic- and propargylic intermediates. Following this, metal-catalysed intramolecular cyclisations such as Co-induced [2+2+2] cyclootrimerisations, when starting from the dialkyne intermediate or a sequence of ring-closing enyne metathesis followed by Diels–Alder cycloaddition, when starting from the allylic intermediate, afforded novel bi- and tricyclic oxepine-containing scaffolds **14**, **15** and **16**. Additionally, ester hydrolysis of **1** to the carboxylic acid followed by the Cu-promoted intramolecular cyclisation afforded **17**, whilst Ru-mediated cyclisation of the amino alcohol intermediate with the terminal alkyne generated **18**.

Smaller propargyl-containing fragments featuring different polar functionalities could also be synthesised via pairing the amino and ester moieties within **1**. Pyrrole **19** was prepared via Paal-Knorr reaction and a nitrile-based secondary branch point was introduced to afford **20** (Scheme 1, step a, see Sup-



**Scheme 1.** Synthesis of structurally diverse N-substituted quaternary carbon containing small molecules from  $\alpha,\alpha$ -disubstituted amino ester **1**: a) amine derivatization; b) pairing of the amino and ester moieties; c) ester derivatization; d) pairing of the alkyne and ester moieties; e) alkyne derivatizations; f) pairing of the amino and alkyne moieties. For reaction conditions see Supporting Information.

porting Information). Benzodiazepine **21** was prepared starting from the azido benzoyl intermediate followed by azide reduction to facilitate the lactamisation reaction with the ester functionality (Scheme 1, steps a/b). Again, acylation of **1** proved useful in generating acyclic intermediates which could in turn be cyclized with the ester functionality. For example, Dieckmann condensation of the malonyl derivative followed by decarboxylation, yielded the substituted pyrrolidone **22**. Moreover, reduction of the ester to a hydroxyl also provided the opportunity to cyclise between this moiety and the amino group (Scheme 1, steps c/b). In this manner, using a single *N*-Boc protected amino alcohol, aziridine **23** and oxazolidinone **26** could be both synthesized. Finally, two further propargyl-containing scaffolds were formed by *N*-acylation followed by hydroxyl-mediated cyclisation or vice versa, *O*-acylation followed by nitrogen-mediated cyclisation, to yield morpholinones **24** and **25**, respectively.

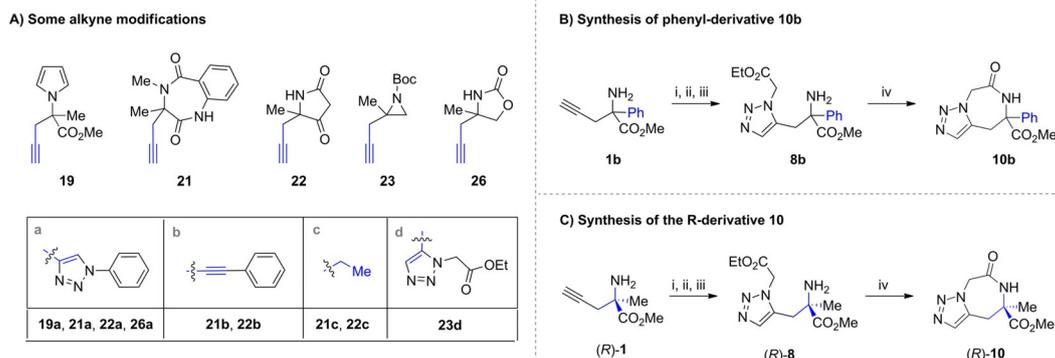
The reactivity of both **19** and **20** could be further exploited to generate five different scaffolds in a synthetically efficient manner. Unexpectedly, when extending the reaction time for the formation of **19**, an alternative intramolecular cyclisation was observed between the terminal alkyne and the in situ hydrolysed methyl ester, giving rise to **27**. Moreover, gold catalysis promoted the intramolecular cyclisation between the terminal alkyne and a pyrrole within **19** to afford dihydroindolizine

**28**. We next sought to explore the reactivity of cyano functionality within **20**. Accordingly, a novel one-step approach to synthesise tricyclic pyrrole-containing scaffolds was developed via Co-catalysed [2+2+2] cyclotrimerization methodology. Thus, starting from **20**, CpCo(CO)<sub>2</sub> and cyclopropylacetylene could be utilised to furnish fused heterocycles **29** and **30**. Further scaffold exploration using this approach is currently being investigated in our research group. Additionally, Pd-catalysed reduction of the nitrile moiety within **20** allowed the intramolecular cyclisation affording the propyl-substituted pyrrolipiperazine **31**.

Finally, the terminal alkyne within **25** and **26** could be further modified to generate two conformationally restricted small molecules **32** and **33** via vinyl iodide formation followed by Buchwald coupling.<sup>[34]</sup>

To demonstrate the synthetic potential of the propargyl-containing warheads for library enumeration, the alkyne handle within **19**, **21–23** and **26** was further modified (Figure 3A). Thus, Cu-catalysed click reaction gave the 1,4-substituted triazoles **19 a**, **21 a**, **22 a**, and **26 a**; Sonogashira cross-coupling afforded **21 b** and **22 b**; hydrogenation of the alkyne produced saturated derivatives **21 c** and **22 c** and, finally, Ru-mediated click reaction afforded the 1,5-triazole **23 d**.

Using the strategy outlined above we synthesised a structurally diverse library of 40 compounds based around 27 molecu-

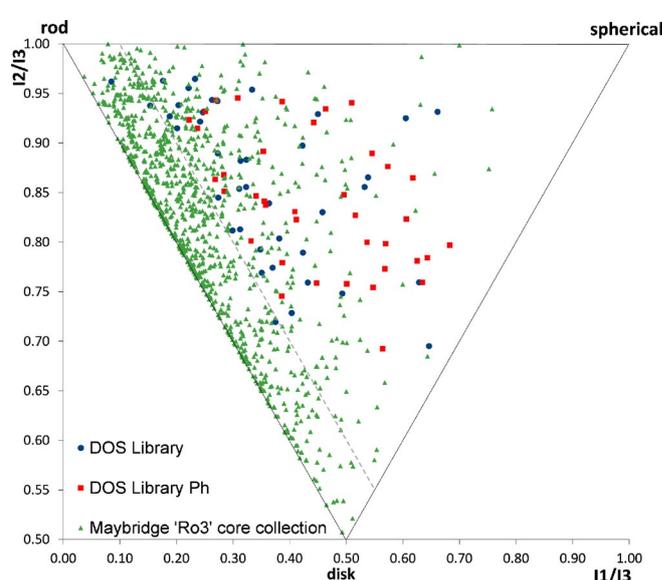


**Figure 3.** (A) Exemplification of the synthesis of other derivatives via alkyne modifications. Reaction conditions: **19 a**, **21 a**, **22 a**, **26 a**: azidobenzene,  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ ,  $t\text{BuOH}:\text{H}_2\text{O}$ , rt, 72–92%; **21 b** and **22 b**:  $\text{PdCl}_2(\text{PPh}_3)_2$ ,  $\text{CuI}$ ,  $\text{Et}_3\text{N}$ , benzyl 2-iodobenzoate, DMF, rt, 16 h, 60% and 26%, respectively; **21 c** and **22 c**:  $\text{Pd/C}$ ,  $\text{H}_2$ , MeOH, rt, 2 h, 69% and 87%, respectively; **23 d**: ethyl 2-azidoacetate,  $\text{Cp}^*\text{RuCl}(\text{COD})$ , PhMe,  $-60^\circ\text{C}$ , 21 h, 42% (B) Synthesis of the phenyl-containing derivative. Reaction conditions: i)  $\text{Boc}_2\text{O}$ , THF,  $70^\circ\text{C}$ , o/n, 84%; ii) ethyl 2-azidoacetate,  $\text{Cp}^*\text{RuCl}(\text{COD})$ , PhMe,  $80^\circ\text{C}$ , 1 h, 87%; iii) TFA,  $\text{CH}_2\text{Cl}_2$ , 2 h then  $\text{NaHCO}_3$ , (quantitative yield); iv) PhMe,  $150^\circ\text{C}$ , o/n, 76%. (C) Synthesis of an enantiopure component of the library. Reaction conditions: i)  $\text{Boc}_2\text{O}$ , THF,  $70^\circ\text{C}$ , o/n, 90%; ii) ethyl 2-azidoacetate,  $\text{Cp}^*\text{RuCl}(\text{COD})$ , PhMe,  $50^\circ\text{C}$ , 1 h, 76%; iii) TFA,  $\text{CH}_2\text{Cl}_2$ , 2 h then  $\text{NaHCO}_3$ , 93%; iv) PhMe,  $150^\circ\text{C}$ , o/n, 79%.

lar frameworks relevant for drug discovery (see Figure S1) featuring a N-containing quaternary stereocenter. Each compound was obtained in no more than five synthetic steps (with an average of only three steps) starting from **1**, which demonstrates the efficiency of this process. Importantly, the flexibility of this methodology was further exemplified via the synthesis of the phenyl-derivative **10'** (Figure 3B) following the same reaction conditions used to generate the methyl derivative **10**. The ability to synthesise the asymmetric version of any compound from the library was also validated through the synthesis of the optically pure triazolodiazepine (*R*)-**10** (Figure 3C) from (*R*)-**1** (see Supporting Information).

### Cheminformatic assessment

In order to assess the molecular shape distribution of the resulting library, a principal moments of inertia (PMI) analysis was conducted (Figure 4).<sup>[35]</sup> Pleasingly, analysis of the DOS library showed a broad distribution of molecular shape space, with 93% of the library out of “flatland”<sup>[36]</sup> and possessing considerable 3D character. In order to investigate the influence of the substituent at the quaternary centre on the shape, a virtual collection based on the phenyl-derivatives (DOS Library Ph, see Supporting Information) was plotted in the same graph (Figure 4). Comparative analysis between the two DOS libraries highlighted a shift to the right-hand side of the plot and more 3D character when modifying the quaternary substituent. These results suggest that installation and substituent modification of the N-containing quaternary  $\text{sp}^3$  stereocenter represents a useful strategy for “tuning” the 3D molecular shape (and thus target binding profile) of a given molecule. Finally, the commercially available Maybridge “Rule of three” (Ro3) core fragment library comprising 1000 screening compounds (see Supporting Information and Figure 4) was analysed and compared to the DOS library. Importantly, only 28% of the Maybridge library appeared out of “flatland” with a high proportion of “disk” and “rod” type compounds, which highlights the increased diversity and more 3D character displayed by the DOS library. Interestingly, only two compounds within the



**Figure 4.** Comparative PMI plot analysis of DOS library (blue circles), Maybridge core 1000-member “Ro3” Fragment library (green triangles) and the DOS Ph virtual library (red squares). Compounds within “flatland” (represented by  $npr1 + npr2 < 1.1$ , dashed line) could also be identified; the further from this area the molecules move the more they extend into three-dimensional space.<sup>[36]</sup>

Maybridge collection featured a N-containing quaternary carbon, exemplifying the limited inclusion of this moiety within traditional screening collections.

In addition to PMI analysis, calculation of the mean values for physicochemical properties related to the “rule of three” guidelines<sup>[37]</sup> commonly adopted within fragment library generation (Table 1) was conducted. Notably the DOS library featured a low mean SlogP (1.37), higher mean number of chiral centres (1.1) and low fraction of aromatic atoms per molecule (0.29), compared to commercial libraries (see Supporting Information), demonstrating the amenability of the DOS library to fragment-based drug discovery approaches.<sup>[38]</sup>

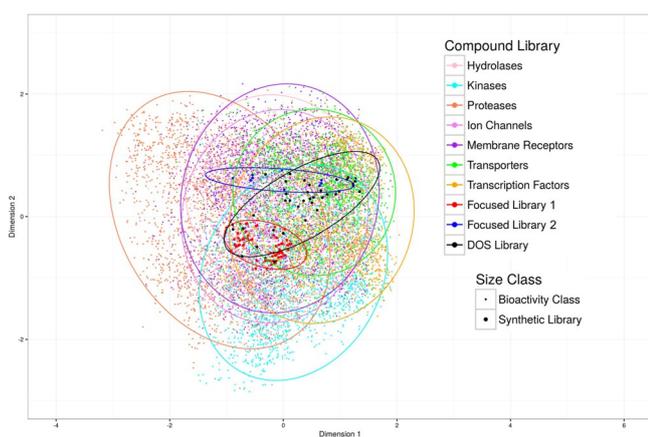
Whilst PMIs are a common way of assessing shape diversity, the relationship to bioactivity coverage appears to be more

Property <sup>[a]</sup>	Ideal range <sup>[b]</sup>	This work	Maybridge "Ro3" core collection	Chembridge
SlogP	0–3	1.37	1.92	1.31
$M_w$	< 300	237	182	222
HBA	< 3	2.63	1.83	1.81
HBD	< 3	0.78	1.01	1.04
chiral centres	–	1.10	0.14	0.27
fraction Ar	–	0.29	0.52	0.42

[a]  $M_w$  = molecular weight, HBA = number of hydrogen bond acceptors, HBD = number of hydrogen bond donors. [b] Ideal range based on guidelines of Fragment "rule of three".<sup>[37]</sup> Green = inside ideal range. Yellow = extreme of ideal range.

complex.<sup>[39]</sup> To assess the coverage of our DOS library in bioactivity space, we compared it to two previously reported focused synthetic libraries,<sup>[40,41]</sup> as well as bioactive compounds retrieved from ChEMBL, a database of binding and functional data for a large number of drug-like bioactive compounds. The molecular weight range of the bioactive compounds selected for the analysis was between 200 and 700, and the compounds covered different target classes.<sup>[42]</sup> Then, their Morgan fingerprints<sup>[43]</sup> were generated for multi-dimensional scaling (MDS). Our DOS library (Figure 5, shown in black) covers a wider area of biologically relevant chemical space compared to the targeted libraries, overlapping with areas occupied by kinase inhibitors, ion channels blockers, membrane receptors ligands and ligands of other protein classes.

Currently, the DOS library is being screened via high-throughput X-ray crystallography methods in collaboration with the XChem fragment screening facility at the Diamond Light Source synchrotron. Hits against three different target classes including hydrolases, the TGF-beta family of growth factors, and peptidases have been successfully identified. Further biological results from these initial campaigns will be published in due course.



**Figure 5.** Bioactive space coverage of the DOS library, compared to two targeted compound libraries and seven different bioactivity classes of common protein targets obtained from ChEMBL (based on Morgan fingerprints and Multi-Dimensional Scaling, MDS; ellipses cover 90% of the data in each class. It can be seen that the DOS library displays significantly wider coverage in bioactivity space than the two targeted libraries, covering all common classes of drug targets used for comparison.

## Conclusion

In summary, we have developed a new, efficient reagent-based DOS approach using  $\alpha$ -methyl- $\alpha$ -propargyl amino esters to generate a library of 40 structurally diverse small molecules. Importantly, throughout this  $sp^3$ -rich library we have installed a N-containing quaternary stereocenter, polar functionalities and N-based heterocycles, answering key calls from within the field. The resultant library also adheres to the fragment rule of three guidelines, highlighting the amenability to this method of screening. Furthermore, we have demonstrated the scope of our library through modifications to the alkyne-containing scaffolds, the synthesis of a phenyl-containing derivative and the asymmetric version of an analogue, all of which can be applied to library enumeration. Finally, cheminformatic analysis has shown a broad molecular shape distribution, the importance of the quaternary  $sp^3$  carbon in modifying the molecular shape and the increased diversity and 3D character of the library when compared to a commercial library. The analysis of the resulting bioactivity space coverage showcases that the target library represents an attractive screening collection of biologically diverse small molecules.

## Experimental Section

### General remarks

All reagents and solvents were purchased from commercial sources and used without further purification unless otherwise stated. All the experiments were carried out under a nitrogen atmosphere unless otherwise stated. Melting points were measured using a Büchi B545 melting point apparatus and are uncorrected. Thin layer chromatography (TLC) was performed on precoated Merck silica gel GF<sub>254</sub> plates. Flash column chromatography was performed on silica gel (230–400 mesh). <sup>1</sup>H NMR and <sup>13</sup>C NMR were recorded on a Bruker Avance 500 MHz instrument in CDCl<sub>3</sub> and MeOD. HRMS was recorded on a Micromass Q-TOF mass spectrometer or a Waters LCT Premier Time of Flight mass spectrometer.

**Methyl 3-methyl-7-phenyl-1,2,3,4-tetrahydro-2,6-naphthyridine-3-carboxylate (2) and methyl 3-methyl-6-phenyl-1,2,3,4-tetrahydro-2,7-naphthyridine-3-carboxylate (3):** A solution of methyl 2-((*tert*-butoxycarbonyl)amino)-2-methylpent-4-ynoate (S2) (150 mg, 0.62 mmol) in DMF (3.0 mL) was added to a suspension of sodium hydride (60% in mineral oil, 30 mg, 0.74 mmol) in DMF (3.2 mL) cooled to 0 °C. After 10 minutes stirring, propargyl bromide (80 wt.% in toluene, 0.138 mL, 1.24 mmol) was added and the mixture was warmed to room temperature and stirred for 4 hours. Then, the mixture was diluted with saturated aqueous solution of NH<sub>4</sub>Cl and the aqueous layer was extracted with EtOAc (3×). The combined organic layers were washed with brine, dried with Na<sub>2</sub>SO<sub>4</sub>, filtered, and the solvents were evaporated. The crude product was purified by flash column chromatography (silica gel petroleum ether/EtOAc, 5:1) to yield 168 mg of methyl 2-((*tert*-butoxycarbonyl)(prop-2-yn-1-yl)amino)-2-methylpent-4-ynoate (S3) (97% yield) as a colourless oil.  $R_f$  = 0.40 (petroleum ether/EtOAc; 5:1); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  = 1.43 (s, 9H), 1.70 (s, 3H), 1.99 (t,  $J$  = 2.6 Hz, 1H), 2.20 (t,  $J$  = 2.4 Hz, 1H), 2.73 (dd,  $J$  = 17.1, 2.5 Hz, 1H), 3.12 (d,  $J$  = 16.6 Hz, 1H), 3.69 (s, 3H), 4.02 (d,  $J$  = 18.5 Hz, 1H), 4.26 ppm (brs, 1H); <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>):  $\delta$  = 22.0, 27.0, 28.3, 34.1, 52.4,

62.5, 70.9, 71.1, 80.1, 80.9, 81.7, 154.3, 174.0 ppm; HRMS (ESI):  $m/z$  calcd for  $C_{15}H_{22}NO_4$   $[M+H]^+$ : 280.1543; found: 280.1534.

$CpCo(CO)_2$  (6.4 mg, 0.036 mmol) and benzonitrile (0.05 mL, 0.48 mmol) were added to a solution of **53** (68 mg, 0.24 mmol) in toluene (2.4 mL), previously degassed with argon for 15 minutes, and the mixture was heated at 110 °C for 36 hours. Then, the solvent was removed under reduced pressure and the crude product was purified by column chromatography (silica gel; petroleum ether/EtOAc, gradient from 5:1 to 3:2). The resulting inseparable mixture of regioisomers (ca. 40:60 ratio as determined by  $^1H$  NMR) was dissolved in  $CH_2Cl_2$  (1 mL) and TFA (0.5 mL) was then added to the solution. After 1 hour stirring the solvent was removed under reduced pressure. The crude product was dissolved in EtOAc (5 mL) and a saturated aqueous solution of  $NaHCO_3$  (5 mL) was added and the mixture was stirred for 10 minutes. Then, the organic layer was separated, dried over  $Na_2SO_4$ , filtered and concentrated in vacuo. The crude product was purified by flash column chromatography (silica gel; gradient from EtOAc:MeOH 1:0 to 9:1) to afford 20 mg of **2** (30% yield) and 18 mg of **3** (26% yield) both as colourless oils.

**Data of minor regioisomer (2):**  $R_f=0.21$  (EtOAc):  $^1H$  NMR (400 MHz,  $CDCl_3$ ):  $\delta=1.48$  (s, 3H), 2.37 (brs, 1H), 2.83 (d,  $J=16.7$  Hz, 1H), 3.35 (d,  $J=16.7$  Hz, 1H), 3.69 (s, 3H), 4.05–4.18 (m, 2H), 7.39 (ddd,  $J=7.3, 3.7, 1.3$  Hz, 1H), 7.41–7.49 (m, 3H), 7.90–7.98 (m, 2H), 8.37 ppm (s, 1H);  $^{13}C$  NMR (101 MHz,  $CDCl_3$ ):  $\delta=26.6, 37.3, 42.6, 52.5, 57.9, 120.6, 126.9, 128.3, 128.8, 128.8, 139.4, 143.2, 147.8, 155.3, 175.9$  ppm; HRMS (ESI):  $m/z$  calcd for  $C_{17}H_{19}N_2O_2$   $[M+H]^+$ : 283.1441; found: 283.1431

**Data of major regioisomer (3):**  $R_f=0.32$  (EtOAc):  $^1H$  NMR (400 MHz,  $CDCl_3$ ):  $\delta=1.48$  (s, 3H), 2.03 (brs, 1H), 2.80 (d,  $J=16.2$  Hz, 1H), 3.33 (d,  $J=16.2$  Hz, 1H), 3.68 (s, 3H), 4.08 (d,  $J=17.0$  Hz, 1H), 4.19 (d,  $J=17.1$  Hz, 1H), 7.34–7.41 (m, 2H), 7.41–7.48 (m, 2H), 7.86–7.97 (m, 2H), 8.43 ppm (s, 1H);  $^{13}C$  NMR (101 MHz,  $CDCl_3$ ):  $\delta=26.4, 34.7, 44.7, 52.6, 58.1, 117.8, 126.9, 127.8, 128.8, 129.0, 139.4, 143.3, 150.2, 155.1, 175.8$  ppm; HRMS (ESI):  $m/z$  calcd for  $C_{17}H_{19}N_2O_2$   $[M+H]^+$ : 283.1441; found: 283.1431.

**Methyl 2-methyl-2-(1H-pyrrol-1-yl)pent-4-ynoate (19):** NaOAc (97 mg, 1.19 mmol) was added to a stirred solution of **1** (140 mg, 0.99 mmol) in a mixture 5:3:1 of DCE:H<sub>2</sub>O:AcOH (2.24 mL) and the reaction was stirred at 90 °C. After 5 minutes, 2,5-dimethoxytetrahydrofuran (0.13 mL, 0.99 mmol) was added and the mixture was stirred at the same temperature overnight. Then, the reaction was cooled to room temperature, diluted with EtOAc and washed with saturated aqueous solution of NaCl. The organic layer was dried over  $Na_2SO_4$ , filtered and the solvents were concentrated in vacuo. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 9:1) to yield 115 mg of **19** (61% yield) as a white solid.  $R_f=0.28$  (petroleum ether/EtOAc, 9:1); m.p. 55–57 °C;  $^1H$  NMR (400 MHz,  $CDCl_3$ ):  $\delta=1.94$  (s, 3H), 2.05 (t,  $J=2.6$  Hz, 1H), 3.00 (dd,  $J=16.8, 2.6$  Hz, 1H), 3.07 (dd,  $J=16.8, 2.6$  Hz, 1H), 3.74 (s, 3H), 6.21 (t,  $J=2.2$  Hz, 2H), 6.83 ppm (t,  $J=2.2$  Hz, 2H);  $^{13}C$  NMR (101 MHz,  $CDCl_3$ ):  $\delta=22.8, 30.0, 53.1, 63.1, 72.3, 78.5, 108.8, 118.7, 172.0$  ppm; HRMS (ESI):  $m/z$  calcd for  $C_{11}H_{14}NO_2$   $[M+H]^+$ : 192.1019; found: 192.1013.

**Methyl 2-(2-cyano-1H-pyrrol-1-yl)-2-methylpent-4-ynoate (20):** Chlorosulfonyl isocyanate (0.047 mL, 0.55 mmol) was added dropwise to a solution of **19** (70 mg, 0.37 mmol) in MeCN (3.6 mL) at 0 °C. After 1 hour stirring, DMF (0.14 mL, 1.85 mmol) was added and the mixture was stirred for 2 hours. Then, saturated aqueous solution of  $NaHCO_3$  was added and the reaction mixture was extracted with EtOAc (3x). The combined organic layers were dried over  $Na_2SO_4$ , filtered and the solvents were concentrated in vacuo. The crude product was purified by flash column chromatography

(silica gel; petroleum ether/EtOAc, 4:1) to yield 57 mg of **20** (71% yield) as colourless oil.  $R_f=0.29$  (petroleum ether/EtOAc; 5:1); mp 64–66 °C;  $^1H$  NMR (400 MHz,  $CDCl_3$ ):  $\delta=1.88$ –2.08 (m, 4H), 3.09 (dd,  $J=17.5, 2.7$  Hz, 1H), 3.31 (dd,  $J=17.5, 2.1$  Hz, 1H), 3.84 (s, 3H), 6.19 (dd,  $J=3.8, 3.0$  Hz, 1H), 6.93 (dd,  $J=3.9, 1.5$  Hz, 1H), 7.07 ppm (dd,  $J=2.8, 1.6$  Hz, 1H);  $^{13}C$  NMR (101 MHz,  $CDCl_3$ ):  $\delta=23.8, 29.5, 53.7, 64.5, 72.8, 77.2, 102.9, 108.5, 114.1, 123.2, 125.5, 171.5$  ppm; HRMS (ESI):  $m/z$  calcd for  $C_{12}H_{12}N_2NaO_2$   $[M+Na]^+$ : 239.0791; found: 239.0783.

**3-Methyl-5-methylene-3-(1H-pyrrol-1-yl)dihydrofuran-2(3H)-one (27):** NaOAc (26 mg, 0.32 mmol) was added to a stirred solution of **19** (30 mg, 0.16 mmol) in a mixture 5:3:1 of DCE:H<sub>2</sub>O:AcOH (1.6 mL) and the reaction was stirred at 90 °C for 24 hours. Then, the reaction was cooled to room temperature, diluted with EtOAc and washed with saturated aqueous solution of NaCl. The organic layer was dried over  $Na_2SO_4$ , filtered and the solvents were concentrated in vacuo. The crude product was purified by flash column chromatography to yield 12 mg of **27** (42% yield) as a colourless oil.  $R_f=0.22$  (petroleum ether/EtOAc, 9:1);  $^1H$  NMR (400 MHz,  $CDCl_3$ ):  $\delta=1.85$  (s, 3H), 3.11 (dt,  $J=16.2, 1.5$  Hz, 1H), 3.39 (dt,  $J=16.2, 1.8$  Hz, 1H), 4.51 (dt,  $J=3.2, 1.7$  Hz, 1H), 4.91 (dt,  $J=2.8, 2.1$  Hz, 1H), 6.24 (t,  $J=2.2$  Hz, 2H), 6.86 ppm (t,  $J=2.1$  Hz, 2H);  $^{13}C$  NMR (101 MHz,  $CDCl_3$ ):  $\delta=24.2, 41.5, 61.6, 91.2, 109.7, 118.4, 151.3, 173.5$  ppm; HRMS (ESI):  $m/z$  calcd for  $C_{10}H_{12}NO_2$   $[M+H]^+$ : 178.0861; found: 178.0863.

**Methyl 5-methyl-5,6-dihydroindolizine-5-carboxylate (28):** A solution of AuSPhos(MeCN)SbF<sub>6</sub> (4.8 mg, 5 mol%) in DCE (0.1 mL) was added dropwise to a solution of **19** (21 mg, 0.11 mmol) and ethanol (0.032 mL, 0.55 mmol) in DCE (1.0 mL). The resulting solution was heated for 2 hours at 40 °C and then the solvent was removed under reduced pressure. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 4:1) to yield 17 mg of **28** (81% yield) as a colourless oil.  $R_f=0.24$  (petroleum ether/EtOAc, 5:1);  $^1H$  NMR (400 MHz,  $CDCl_3$ ):  $\delta=1.80$  (s, 3H), 2.49 (dt,  $J=16.9, 2.7$  Hz, 1H), 2.95 (dd,  $J=16.9, 6.1$  Hz, 1H), 3.64 (s, 3H), 5.58–5.69 (m, 1H), 6.12 (d,  $J=2.5$  Hz, 1H), 6.22 (t,  $J=3.2$  Hz, 1H), 6.46 (dd,  $J=9.7, 2.8$  Hz, 1H), 6.85 ppm (s, 1H);  $^{13}C$  NMR (101 MHz,  $CDCl_3$ ):  $\delta=24.0, 35.8, 52.9, 61.3, 107.4, 108.9, 117.1, 118.9, 120.8, 130.3, 174.0$  ppm; HRMS (ESI):  $m/z$  calcd for  $C_{11}H_{14}NO_2$   $[M+H]^+$ : 192.1019; found: 192.1012.

**Methyl 2-cyclopropyl-6-methyl-5,6-dihydropyrrolo[1,2-h][1,7]naphthyridine-6-carboxylate (29) and methyl 3-cyclopropyl-6-methyl-5,6-dihydropyrrolo[1,2-h][1,7]naphthyridine-6-carboxylate (30):**  $CpCo(CO)_2$  (13.3 mg, 0.038 mmol) and cyclopropylacetylene (0.048 mL, 0.57 mmol) were added to a solution of **20** (40 mg, 0.19 mmol) in toluene (3.5 mL), previously degassed with argon for 15 minutes, and the mixture was heated at 110 °C for 36 hours. Then, the solvent was removed under reduced pressure and the crude product was purified by column chromatography (silica gel; gradient from petroleum ether/EtOAc, gradient from 4:1 to 2:1) to afford 14 mg of **29** (27% yield) and 5 mg of **30** (10% yield) both as white solids.

**Data of major regioisomer (29):**  $R_f=0.34$  (petroleum ether/EtOAc; 4:1); mp 108–110 °C;  $^1H$  NMR (400 MHz,  $CDCl_3$ ):  $\delta=0.94$  (dd,  $J=8.2, 3.3$  Hz, 2H), 0.99–1.12 (m, 2H), 1.88 (s, 3H), 1.95–2.07 (m, 1H), 3.06 (d,  $J=15.4$  Hz, 1H), 3.43 (d,  $J=15.5$  Hz, 1H), 3.54 (s, 3H), 6.33 (t,  $J=3.2$  Hz, 1H), 6.79 (d,  $J=7.8$  Hz, 1H), 6.90–6.99 (m, 2H), 7.28 ppm (d,  $J=7.8$  Hz, 1H);  $^{13}C$  NMR (101 MHz,  $CDCl_3$ ):  $\delta=9.7, 17.2, 24.3, 39.4, 53.0, 61.2, 107.9, 110.1, 117.8, 120.0, 120.2, 131.5, 135.1, 146.9, 161.7, 173.1$  ppm; HRMS (ESI):  $m/z$  calcd for  $C_{17}H_{19}N_2O_2$   $[M+H]^+$ : 283.1441; found: 283.1435.

**Data of minor regioisomer (30):**  $R_f=0.23$  (petroleum ether/EtOAc; 2:1);  $^1H$  NMR (400 MHz,  $CDCl_3$ ):  $\delta=0.67$ –0.75 (m, 2H), 0.97–1.03 (m,

2H), 1.82–1.92 (m, 4H), 3.08 (d,  $J=15.5$  Hz, 1H), 3.43 (d,  $J=15.5$  Hz, 1H), 3.54 (s, 3H), 6.35 (t,  $J=3.0$  Hz, 1H), 6.94 (s, 2H), 7.07 (s, 1H), 8.25 ppm (s, 1H);  $^{13}\text{C}$  NMR (101 MHz,  $\text{CDCl}_3$ ):  $\delta=8.9, 9.0, 13.1, 24.3, 39.8, 53.1, 61.0, 107.6, 110.3, 120.0, 123.5, 131.0, 132.2, 136.4, 145.3, 147.2, 173.0$  ppm; HRMS (ESI):  $m/z$  calcd for  $\text{C}_{17}\text{H}_{19}\text{N}_2\text{O}_2$  [ $M+H$ ] $^+$ : 283.1441; found: 283.1440.

**4-Methyl-4-propyl-1,2-dihydropyrrrolo[1,2-a]pyrazin-3(4H)-one (31):**  $\text{PtO}_2$  (34 mg, 0.15 mmol) was added to a solution of **20** (35 mg, 0.15 mmol) in MeOH (3 mL) and the reaction mixture was stirred under  $\text{H}_2$  atmosphere (1 atm) for 16 h. Then, the resulting suspension was filtered and the filtrate was concentrated under reduced pressure. The crude product was purified by flash column chromatography (silica gel; petroleum ether/EtOAc, 1:1) to yield 10 mg of **31** (35% yield) as colourless oil.  $R_f=0.24$  (petroleum ether/EtOAc; 1:1);  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ ):  $\delta=0.82$  (t,  $J=7.2$  Hz, 3H), 0.88–1.01 (m, 1H), 1.11–1.23 (m, 1H), 1.72 (d,  $J=4.5$  Hz, 3H), 1.82 (ddd,  $J=13.6, 10.3, 4.3$  Hz, 1H), 2.16 (ddd,  $J=13.7, 12.1, 4.6$  Hz, 1H), 4.52–4.63 (m, 2H), 5.93 (d,  $J=2.0$  Hz, 1H), 6.25 (t,  $J=3.2$  Hz, 1H), 6.35 (s, 1H), 6.69 ppm (dd,  $J=2.4, 1.8$  Hz, 1H);  $^{13}\text{C}$  NMR (126 MHz,  $\text{CDCl}_3$ ):  $\delta=14.0, 17.4, 26.4, 39.8, 43.7, 62.9, 102.8, 109.8, 115.9, 122.1, 172.2$  ppm; HRMS (ESI):  $m/z$  calcd for  $\text{C}_{11}\text{H}_{17}\text{N}_2\text{O}$  [ $M+H$ ] $^+$ : 193.1335; found: 193.1328.

## Acknowledgements

Our research is supported by the EPSRC, BBSRC, MRC, Wellcome Trust, and ERC (FP7/2007–2013; 279337/DOS). N.M. thanks the EU for a Marie Curie Fellowship (2013-IEF-626191). S.L.K. thanks AstraZeneca for funding. L.K. thanks the IDB Cambridge International Scholarship. The authors would like to thank Prof. Frank von Delft, Dr. Romain Talon and Dr. Anthony Aimon at XChem for enabling library screening using their facility and collaborators Dr. Marko Hyvönen and Prof. Christopher Dowson.

## Conflict of interest

The authors declare no conflict of interest.

**Keywords:** diversity-oriented synthesis • drug discovery • medicinal chemistry • molecular diversity • quaternary stereocenters

- www.who.int/medicines/publications/essentialmedicines/20th\_EML2017.pdf?ua=1.
- <https://clinicaltrials.gov/ct2/show/NCT00927914>.
- <https://clinicaltrials.gov/ct2/show/NCT02470585>.
- <https://clinicaltrials.gov/ct2/show/NCT01953601>.
- P. Zanos, R. Moaddel, P. J. Morris, P. Georgiou, J. Fischell, G. I. Elmer, M. Alkondon, P. Yuan, H. J. Pribut, N. S. Singh, *Nature* **2016**, *533*, 481–486.
- K. Hashimoto, *Clin. Psychopharmacol. Neurosci.* **2014**, *12*, 72–73.
- A. Treiber, R. de Kanter, C. Roch, J. Gatfield, C. Boss, M. von Raumer, B. Schindelholz, C. Muhlan, J. van Gerven, F. Jenck, *J. Pharmacol. Exp. Ther.* **2017**, *362*, 489–503.
- T. B. Durham, J. Marimuthu, J. L. Toth, C. Liu, L. Adams, D. R. Mudra, C. Swearingen, C. Lin, M. G. Chambers, K. Thirunavukkarasu, *J. Med. Chem.* **2017**, *60*, 5933–5939.
- M. S. Malamas, J. Erdei, I. Gunawan, J. Turner, Y. Hu, E. Wagner, K. Fan, R. Chopra, A. Olland, J. Bard, *J. Med. Chem.* **2010**, *53*, 1146–1158.
- M. S. Malamas, J. Erdei, I. Gunawan, K. Barnes, M. Johnson, H. Yu, J. Turner, H. Yun, E. Wagner, K. Fan, *J. Med. Chem.* **2009**, *52*, 6314–6323.

- K. J. Hodgetts, in *Blood–Brain Barrier Drug Discovery*, Wiley, Hoboken, **2015**, pp. 425–445.
- Z. Fang, Y. Song, P. Zhan, Q. Zhang, X. Liu, *Future Med. Chem.* **2014**, *6*, 885–901.
- F. Lovering, J. Bikker, C. Humblet, *J. Med. Chem.* **2009**, *52*, 6752–6756.
- P. A. Clemons, J. A. Wilson, V. Dančik, S. Muller, H. A. Carrinski, B. K. Wagner, A. N. Koehler, S. L. Schreiber, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 6817–6822.
- J. Clayden, M. Donnard, J. Lefranc, D. J. Tetlow, *Chem. Commun.* **2011**, *47*, 4624–4639.
- X. Yang, F. D. Toste, *J. Am. Chem. Soc.* **2015**, *137*, 3205–3208.
- T. Vilaivan, W. Bhanthumnavin, *Molecules* **2010**, *15*, 917–958.
- A. M. R. Smith, K. Kuok, M. Hii, *Chem. Rev.* **2011**, *111*, 1637–1656.
- F. Zhou, F. M. Liao, J. S. Yu, J. Zhou, *Synthesis* **2014**, *46*, 2983–3003.
- Y. Ohfuné, T. Shinada, *Eur. J. Org. Chem.* **2005**, 5127–5143.
- D. Foley, R. Doveston, I. Churcher, A. Nelson, S. P. Marsden, *Chem. Commun.* **2015**, *51*, 11174–11177.
- A. W. Hung, A. Ramek, Y. Wang, T. Kaya, J. A. Wilson, P. A. Clemons, D. W. Young, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 6799–6804.
- D. C. Blakemore, L. Castro, I. Churcher, D. C. Rees, A. W. Thomas, D. M. Wilson, A. Wood, *Nat. Chem.* **2018**, *10*, 383–394.
- C. W. Murray, D. C. Rees, *Angew. Chem. Int. Ed.* **2016**, *55*, 488–492; *Angew. Chem.* **2016**, *128*, 498–503.
- T. E. Storr, S. J. Cully, M. J. Rawling, W. Lewis, D. Hamza, G. Jones, R. A. Stockman, *Bioorg. Med. Chem.* **2015**, *23*, 2621–2628.
- M. J. Rawling, T. E. Storr, W. A. Bawazir, S. J. Cully, W. Lewis, M. S. I. T. Makki, I. R. Strutt, G. Jones, D. Hamza, R. A. Stockman, *Chem. Commun.* **2015**, *51*, 12867–12870.
- K. F. Morgan, I. A. Hollingsworth, J. A. Bull, *Chem. Commun.* **2014**, *50*, 5203–5205.
- J. Mayol-Llinàs, W. Farnaby, A. Nelson, *Chem. Commun.* **2017**, *53*, 12345–12348.
- J. T. R. Liddon, M. J. James, A. K. Clarke, P. O'Brien, R. J. K. Taylor, W. P. Unsworth, *Chem. Eur. J.* **2016**, *22*, 8777–8780.
- M. J. James, P. O'Brien, R. J. K. Taylor, W. P. Unsworth, *Angew. Chem. Int. Ed.* **2016**, *55*, 9671–9675; *Angew. Chem.* **2016**, *128*, 9823–9827.
- D. J. Foley, P. G. E. Craven, P. M. Collins, R. G. Doveston, A. Aimon, R. Talon, I. Churcher, F. von Delft, S. P. Marsden, A. Nelson, *Chem. Eur. J.* **2017**, *23*, 15227–15232.
- O. A. Davis, R. A. Croft, J. A. Bull, *Chem. Commun.* **2015**, *51*, 15446–15449.
- J. Miró, M. Sánchez-Roselló, J. González, C. Del Pozo, S. Fustero, *Chem. Eur. J.* **2015**, *21*, 5459–5466.
- S. Nicolai, C. Piemontesi, J. Waser, *Angew. Chem. Int. Ed.* **2011**, *50*, 4680–4683; *Angew. Chem.* **2011**, *123*, 4776–4779.
- W. H. B. Sauer, M. K. Schwarz, *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 987–1003.
- A. D. Morley, A. Pugliese, K. Birchall, J. Bower, P. Brennan, N. Brown, T. Chapman, M. Drysdale, I. H. Gilbert, S. Hoelder, *Drug Discovery Today* **2013**, *18*, 1221–1227.
- M. Congreve, R. Carr, C. Murray, H. Jhoti, *Drug Discovery Today* **2003**, *8*, 876–877.
- D. G. Twigg, N. Kondo, S. L. Mitchell, W. R. J. D. Galloway, H. F. Sore, A. Madin, D. R. Spring, *Angew. Chem. Int. Ed.* **2016**, *55*, 12479–12483; *Angew. Chem.* **2016**, *128*, 12667–12671.
- H. J. Roth, *Curr. Opin. Chem. Biol.* **2005**, *9*, 293–295.
- M. C. Schuster, D. A. Mann, T. J. Buchholz, K. M. Johnson, W. D. Thomas, L. L. Kiessling, *Org. Lett.* **2003**, *5*, 1407–1410.
- V. Nesterenko, K. S. Putt, P. J. Hergenrother, *J. Am. Chem. Soc.* **2003**, *125*, 14672–14673.
- A. Gaulton, L. J. Bellis, A. P. Bento, J. Chambers, M. Davies, A. Hersey, Y. Light, S. McGlinchey, D. Michalovich, B. Al-Lazikani, *Nucleic Acids Res.* **2012**, *40*, D1100–1107.
- <http://www.rdkit.org>.

Manuscript received: June 19, 2018

Revised manuscript received: July 11, 2018

Accepted manuscript online: July 16, 2018

Version of record online: August 13, 2018