UNIVERSITY OF CAMBRIDGE

Studies within Fragment-Based Drug Discovery: Library Synthesis and Hit-to-Lead Optimisation

Attila Sveiczer

Selwyn College

University of Cambridge

September 2020

Supervised by Professor David R. Spring

This dissertation is submitted for the degree of Doctor of Philosophy

Declaration

This dissertation is submitted in fulfilment of the requirements for the degree of Doctor of Philosophy. It describes work carried out in the Departments of Chemistry and Biochemistry, University of Cambridge, between October 2016 and March 2020 under the supervision of Professor David R. Spring and Dr. Martin Welch. 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 or is being submitted for any other degree. It does not exceed the prescribed word limit for the Physics and Chemistry Degree Committee.
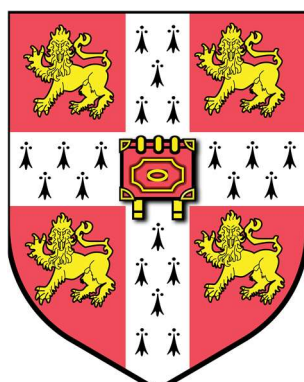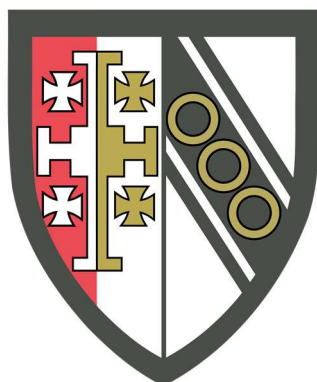
Signed   Date

Attila Sveiczer   02/01/2021

Selwyn College, University of Cambridge

Abstract

Studies within Fragment-Based Drug Discovery: Library Synthesis and Hit-to-Lead Optimisation
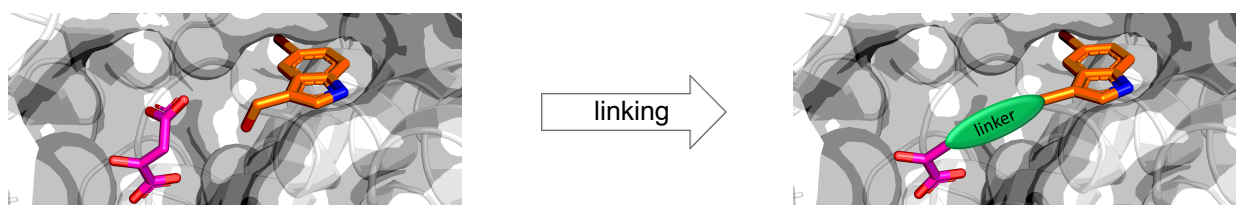
Attila Sveiczer

This thesis reports two projects aimed at addressing challenges within fragment-based drug discovery.

The first project describes efforts towards utilising synthetic methodology to address deficiencies within fragment screening collections. This involved the development of a modular, robust and scalable route to access α,α-disubstituted amino ester building blocks, which in turn were derivatised to allow the rapid assembly of five (a total of eight in collaboration) spirocyclic scaffolds. Importantly, this library was structurally diverse, comprising three (a total of six in collaboration) pharmacophore-like heterocycles and carbocycles. Moreover, numerous three-dimensional exit vectors were incorporated within each core spirocycle, and the ability of these handles to effect a diverse set of chemical modifications was exemplified through the generation of 16 (a total of 21 in collaboration) examples. Computational studies highlighted the excellent physicochemical and 3D properties of the library, as well as the broad coverage of underexplored chemical space that was achieved. This library was also screened for antibacterial activity in a phenotypic assay against the clinically relevant bacterial strains, *Pseudomonas aeruginosa* and *Staphylococcus aureus*.
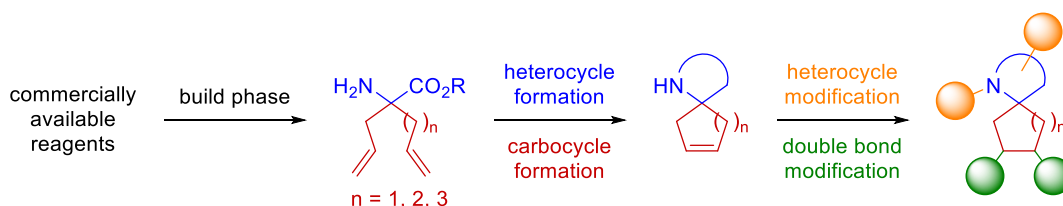
Generation of a diverse spirocyclic fragment library.

The second project examined the inhibition of propionate detoxification mechanisms in bacteria as an attractive strategy for the development of antibacterial agents. A fragment screening campaign against 2-methylcitrate synthase (PrpC) from *Pseudomonas aeruginosa* identified several hit compounds based on an indole unit. Synthetic efforts were undertaken to elaborate these fragment hits to increase potency. The adopted strategy focused on growing the indole fragment towards the nearby oxaloacetate binding pocket and occupying it with a fragment mimicking its natural substrate. This approach yielded a compound with an *in vitro* half maximal inhibitory concentration of 130 μM against the enzymatic activity of PrpC.

Fragment linking approach to the inhibition of PrpC.

Acknowledgements

Abbreviations

°C	Degree Celsius
2D	Two-dimensional
2-MC	2-Methylcitrate
3D	Three-dimensional
AAC	Azide-alkyne cycloaddition
Ac	Acetyl
acac	acetylacetonyl
AcnB	Aconitase B
AcnD	Aconitase D
AMP	antimicrobial peptides
AMR	antimicrobial resistance
Ar	Aryl
ATCC	American Type Culture Collection
ATP	Adenosine triphosphate
ATR	Attenuated total reflection
Aux	Auxiliary
aq	Aqueous
BACE	Aspartyl protease β-secretase
BIOS	Biology-oriented synthesis
Bn	Benzyl
Boc	*tert*-Butoxycarbonyl
br	Broad
BQ	1,4-Benzoquinone
calcd	Calculated
CAN	Ceric ammonium nitrate
CatD	Cathepsin D
CFI_{25}	Cleavage factor 25 kD
CINV	Chemotherapy-induced nausea and vomiting
clogP	Decimal logarithm of the calculated partition coefficient
CML	Chronic myeloid leukemia
CoA	Coenzyme A
conc.	Concentrated
COSY	Correlation spectroscopy
d	Doublet
δ	Chemical shift
Da	Dalton, g mol^{-1}
DCE	1,2-Dichloroethane
DDQ	2,3-Dichloro-5,6-dicyano-1,4-benzoquinone
DEL	DNA-encoded library
DEPT	Distortionless enhancement by polarization transfer
DFP	Deviation from planarity

DHFR	Dihydrofolate reductase
DHP	7,8-Dihydropteroate
DHPPP	6-Hydroxymethyl-7,8-dihydropterin pyrophosphate
DHPS	Dihydropteroate synthase
DIPEA	*N,N*-Diisopropylethylamine
DMA	*N,N*-Dimethylacetamide
DMAP	4-Dimethylaminopyridine
DMF	*N,N*-Dimethylformamide
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DOS	Diversity-oriented synthesis
DPM	Diphenylmethylene
dr	diastereomeric ratio
DSF	Differential scanning fluorimetry
DTNB	5,5'-Dithiobis-(2-nitrobenzoic acid)
E	Linking coefficient
EC_{50}	Half maximal effective concentration
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
ESI	Electrospray ionisation
Et	Ethyl
eq	Equivalent(s)
FabI	Enoyl-acyl carrier protein reductase
Far	Fraction of aromatic atoms
FBDD	Fragment-based drug discovery
FDA	United States of America Food and Drug Administration
Fflat	Fraction of 'flat' compounds
FRET	Fluorescence resonance energy transfer
Fsp^3	Fraction of sp^3-hybridized heavy atoms
G	Gibbs free energy
Grubbs II	Second-generation Grubbs catalyst: (1,3-bis(2,4,6-trimethylphenyl)-2-imidazolidinylidene)dichloro(phenylmethylene)(tricyclohexylphosphine)ruthenium
GSK	GlaxoSmithKline
GSK-3β	Glycogen synthase kinase 3β
h	Hour(s)
H	Enthalpy
HAC	Heavy atom count
HATU	1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate
HBA	Hydrogen-bond acceptor
HBD	Hydrogen-bond donor
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
HIV	Human immunodeficiency virus
HMBC	Heteronuclear multiple-bond correlation spectroscopy
HOBt	1-Hydroxybenzotriazole
HRMS	High resolution mass spectrometry

HSQC	Heteronuclear single-quantum correlation spectroscopy
HTS	High-throughput screening
Hz	Hertz
IC_{50}	Half maximal inhibitory concentration
IR	Infra-red
J	Coupling constant
K_d	Dissociation constant
KHMDS	Potassium bis(trimethylsilyl)amide
LCMS	Liquid chromatography–mass spectrometry
LDA	Lithium diisopropylamide
LE	Ligand efficiency
LLE	Lipophilic ligand efficiency
LLE_{AT}	Heavy atom adjusted lipophilic ligand efficiency
LpxC	Uridine-diphosphate-3-*O*-(*R*-3-hydroxymyristoyl)-*N*-acetylglucosamine deacetylase
λ_{Max}	Wavelength of maximum absorption
M	Molar; general metal or metalloid
m	Multiple; medium
*m*CPBA	*meta*-Chloroperoxybenzoic acid
MDM2	Mouse double minute 2
Me	Methyl
min	Minute(s)
ML	Machine learning
MRSA	Methicillin-resistant *Staphylococcus aureus*
MS	Molecular sieves
MW	Molecular weight
m.p.	Melting point
m/z	Mass to charge ratio
μw	Microwave
NAD	Nicotinamide adenine dinucleotide
NADP	Nicotinamide adenine dinucleotide phosphate
NMO	*N*-Methylmorpholine *N*-oxide
NMR	Nuclear magnetic resonance
NPR	Normalised principal moments of inertia ratio
Nu	Nucleophile
v_{Max}	Frequency of maximum absorption
OAA	Oxaloacetic acid
ORF	Open reading frame
o/n	Overnight
P	Partition coefficient
PABA	*para*-Aminobenzoic acid
PAINS	Pan assay interference compounds
PBP3	Penicillin binding protein 3
PDD	Phenotypic drug discovery
PE	Petroleum ether, 40-60 °C fraction, unless otherwise specified
PG	Protecting group

pH	Decimal logarithm of the reciprocal of the hydrogen ion activity
Ph	Phenyl
PMB	*para*-Methoxy benzyl
PMHS	Polymethylhydrosiloxane
PMI	Principal moments of Inertia
PMP	*para*-Methoxy phenyl
PPI	Protein-protein interaction
ppm	Parts per million
Pr	Propyl
PrpB	2-Methylisocitrate lyase
PrpC	2-Methylcitrate synthase
PrpD	Methylcitrate dehydratase
PrpE	Propionyl coenzyme A synthase
PrpF	Methylaconitate isomerase
Py	Pyridine
R	General group, usually alkyl, aryl or acyl, unless otherwise specified
RBC	Rotatable bond count
RCM	Ring-closing metathesis
R_f	Retention factor
Ro3	'Rule of Three'
rr	regioisomeric ratio
R&D	Research and development
q	Quartet
qui	Quintet
rt	Room temperature
s	Singlet; strong
S	Entropy
SAR	Structure-activity relationship
sat.	Saturated
SCD	Stearoyl-coenzyme A desaturase
siRNA	Small interference ribonucleic acid
SlogP	See clogP
SNP	Single nucleotide polymorphism
SPR	Surface plasmon resonance
t	Triplet
T	Temperature
TBA	Tetrabutylammonium
TBS	*tert*-Butyldimethylsilyl
*t*Bu	*tert*-Butyl
TCA	Tricarboxylic acid
TDD	Target-based drug discovery
Tf	Trifluoromethanesulfonyl
TFA	Trifluoroacetic acid
TFAA	Trifluoroacetic anhydride
TFE	2,2,2-Trifluoroethanol

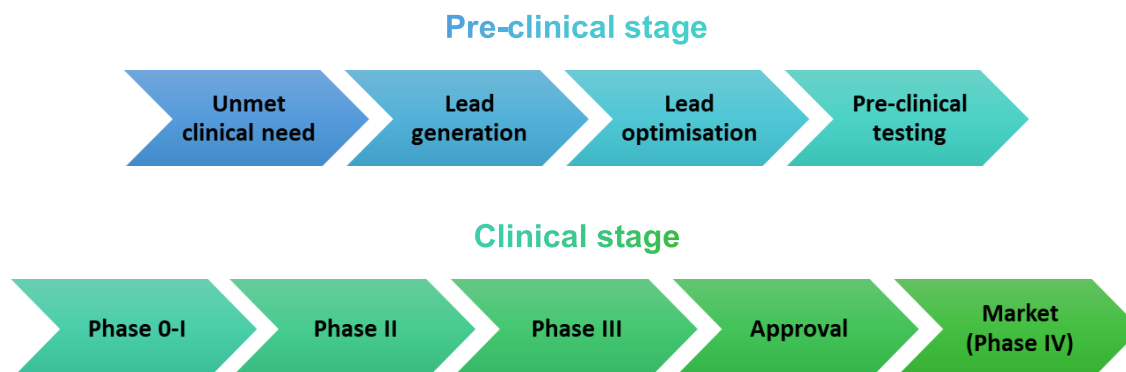|	
THF	Tetrahydrofuran
THFo	Tetrahydrofolate
TLC	Thin-layer chromatography
TMS	Trimethylsilyl
TOS	Target-oriented synthesis
TPSA	Topological polar surface area
Tris	Tris(hydroxymethyl)aminomethane
Ts	Tosyl, para-Toluenesulfonyl
uHTS	Ultra-high-throughput screening
US	United States of America
UV	Ultra-violet
v/v	By volume
w	Weak
w/w	By weight
X	General group, usually heteroatom, unless otherwise specified
XIAP	X-linked inhibitor of apoptosis protein

Table of Contents

1 Introduction

1.1 Small Molecules in Drug Discovery

Global healthcare suffers from a lack of effective treatments and cures for countless communicable and non-communicable diseases. Despite the astounding achievements of the pharmaceutical industry in developing therapeutics against contagious diseases through antibiotic and vaccine development, many challenges remain to be overcome—such as malaria, cancer and the growing antibiotic resistance.[1–3]

Therapeutics—chemical entities used to treat diseases—can be divided into two main categories: small molecules and biopharmaceuticals (also referred to as 'biologics').[4] Within the fields of molecular biology and pharmacology, the term 'small molecule' typically refers to organic compounds possessing a molecular weight (MW) less than 1500 Da, the potential for oral bioavailability, and a structure distinct to those found in biological macromolecules.[5,6] On the other hand, biologics are large MW substances comprising structures isolated from biological systems such as nucleic acids and proteins. In recent years, a number of biologics have proved immensely successful in the treatment of several disease classes, including inflammatory and cardiovascular disease as well as cancer, mainly due to their excellent specificity.[7] Although biologics account for a continually increasing proportion of therapeutics approved by the United States of America Food and Drug Administration (FDA) year on year (from 10% in the 1990's to 23% in the 2010's),[8] it remains the case that the majority of approved therapeutics are small molecules.[8] This continued success of small molecules can be attributed to several advantages over biologics such as: superior chemical stability, are easier to characterise, good homogeneity, better (oral) bioavailability, their non-immunogenic nature and often better pharmacokinetic profiles. They are also cheaper to develop and produce, and generics are typically approved faster providing more cost efficient alternatives after the expiration of the original patents.[7,9]

1.2 The Drug Discovery Process

Small molecule drugs are brought to the market in a well-established pharmaceutical pipeline, which can be separated into the pre-clinical and clinical stages (Figure 1.1).

Pre-clinical stage

| Unmet clinical need | Lead generation | Lead optimisation | Pre-clinical testing |

Clinical stage

| Phase 0-I | Phase II | Phase III | Approval | Market (Phase IV) |

Figure 1.1 The drug discovery process. The pre-clinical stage involves the development of drug candidates and their initial testing. During the clinical stage the drug candidates are tested in humans to assess their safety and efficacy in order to gain approval by regulatory agencies.

The innovative pre-clinical stage involves the identification of a target disease with no or suboptimal treatment(s) available. This is followed by the development of a potent lead compound, which is further optimised to achieve good pharmacokinetic and pharmacodynamic properties. Subsequently, optimised leads are assessed by pre-clinical *in vitro* and *in vivo* animal testing and drug candidates are selected.[10,11]

The heavily regulated clinical stage consists of several phases. The drug candidate is first given to a few healthy humans (30 – 100) during phase 0 and I to evaluate its safety. Upon successful completion of this stage, the drug is introduced to people suffering from the target disease (100 – 300) in phase II to obtain initial information about the efficacy and acute side effects. Finally, in phase III, a much larger cohort of patients (1000 – 3000) are supplied with the drug candidate to fully assess its efficacy and possible side effects. The relevant regulatory agencies are then presented with the collected data and make decision if the compound can be marketed as a drug for the treatment of the specified disease. Marketed drugs are continued to be monitored (phase IV), and if deemed necessary they can be withdrawn by the licensing authorities.[12]

1.2.1 Duration and cost of drug discovery

Estimates on the length and cost of this process are varied within the literature, but most studies examining drugs that received FDA approval after 1990 agree that on average it takes more than 10 years and in excess of US$1 billion to develop a single marketed drug.[13–16] Furthermore, an influential study by Scannell *et al*. in 2012 showed that the number of new drugs approved by FDA per US$1 billion spent on R&D has exponentially decreased (termed 'Eroom's law') since 1950, and halving approximately every 9 years.[17] Moreover, a recent analysis by DiMasi *et al*. in 2016 exposed that the costs associated with the pre-clinical stage (including the pre-clinical

testing) account for 38%, the clinical stage (Phases 0-III) for 51%, and the post-approval (Phase IV) for 11% of the total capitalised expenditure of $2.8 billion per approved new drug.[13] This study also showed that lead generation and optimisation takes about two and a half years—the same as the pre-clinical tests—whereas the clinical stage takes 8 years on average, and the approval process an additional 16 months—amounting to more than 14 years for a drug to reach the market from inception. Thus, the development of new drugs is more costly both in terms of money and time than ever before.

Analyses have shown that the clinical stage contributes tremendously to both the cost and length of the drug development process, partially due to high attrition rates (40% for phase 0-I, 64% for phase II, 38% for phase III and 10% for the FDA approval process, yielding an overall success rate of 12% for drug candidates reaching the market).[13] This is highlighted by the fact that—although the direct cost of a single candidate to transition through the clinical phases is estimated to be around US$460 million—when the huge expense of failed compounds are considered, the total cost of the clinical stage rises to US$1.46 billion for each successful drug candidate.[13]

Analysis of clinical success rates by Smietana *et al*. demonstrated an overall decrease from 16% in the late 1990's to 7.5% around 2010, which can be attributed partially to the increasingly demanding requirements of regulatory agencies.[18] The decline in clinical success rates consequently resulted in a steady decrease in the number of new therapeutics approved for drug use by FDA from the late 1990's until 2007 (Figure 1.2).[8] These findings suggest that the major scientific and technological advances of the past few decades (combinatorial chemistry, DNA sequencing, high-throughput screening, biotechnology, *etc*.) have failed to deliver the anticipated increase in the efficiency of commercial drug research and development (R&D). In response, the pharmaceutical and biotechnological industry has been investing an ever-increasing amount of money into R&D (upwards of US$140 billion worldwide in 2015), producing many new drugs over the last decade (Figure 1.2). As a result however, many companies now question the sustainability of their R&D programmes.[19]
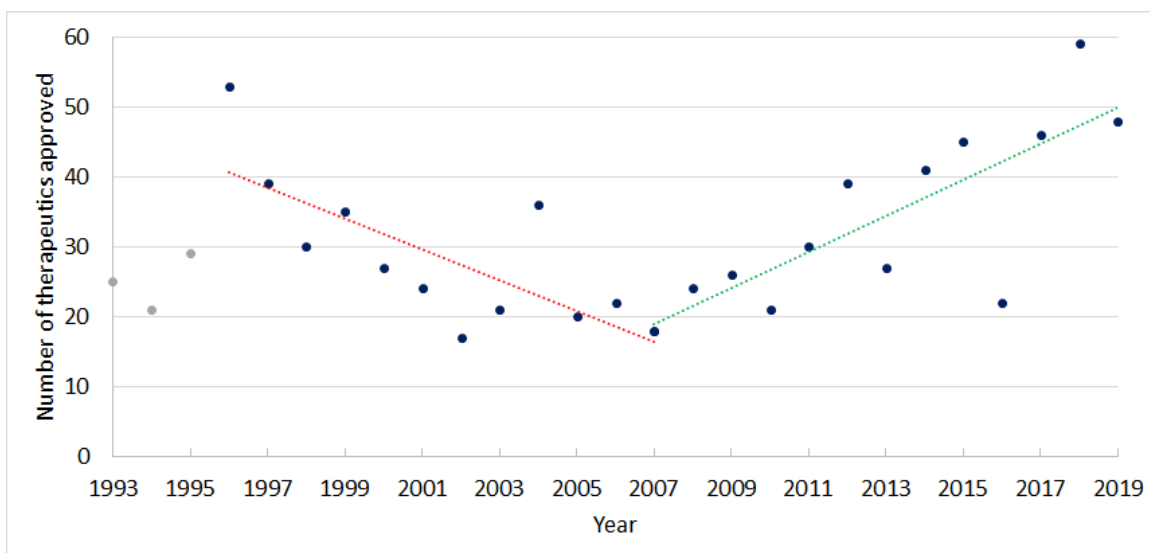
Figure 1.2 Total number of new therapeutics (small molecules and biologics) approved by FDA. A drastic decline can be observed from 1996 to 2007, which is followed by a similarly significant increase from 2007 to present day. Data obtained from Asher *et al*.[8]

1.2.2 Approaches to improve clinical success rates

One strategy that has been investigated to improve the efficiency of R&D campaigns is by filtering out compounds that are likely to fail expensive clinical trials as early as possible in the development process. Several studies have looked at the connections between the success of drug candidates throughout the clinical phases and their molecular properties in order to identify 'drug-like' properties. One of the most well-cited pieces of analysis by Lipinski *et al*. in 1997 proposed a 'Rule of Five', linking a large number of hydrogen-bond donors and acceptors, MW, and high lipophilicity to poor absorbance and permeability, ultimately causing insufficient oral bioavailability and therefore a lack of efficacy and candidate attrition (Table 1.1).[20] Contrastingly, a later study by Veber *et al*. in 2002 showed that the MW cut off is not a reliable descriptor of bioavailability and instead a rotatable bond count of less than 10 and a total polar surface area of less than 140 $Å^2$ gives a much better correlation to available research data (Table 1.1).[21]

4

Table 1.1 Guidelines on physicochemical properties for drug candidates commonly adopted within the field.

Property	Rule of Five
MW[a]	< 500 Da
HBD[a]	≤ 5
HBA[a]	≤ 10
clogP[a]	≤ 5
RBC[b]	≤ 10
TPSA[b]	≤ 140 Å^2

MW = molecular weight, HBD = number of hydrogen-bond donors, HBA = number of hydrogen-bond acceptors, clogP = partition coefficient, RBC = rotatable bond count, TPSA = topological polar surface area. [a]Original properties by Lipinski *et al.*[20] [b]Properties added by Veber *et al.*[21]

Whilst the exact definition of 'drug-like' properties remains under debate, controlling the physicochemical properties of lead candidates continues to be a well-employed strategy within the pharmaceutical industry.[22] The fact that the upper 90th percentile values of clogP (~5) and number of hydrogen-bond donors (4.0) in FDA-approved drugs have not changed significantly over the years suggests that these cut offs may indeed be validated criteria for successful drug candidates.[22] Perhaps most importantly, high lipophilicity has been linked to an increasing likelihood of multiple target binding and thus pharmacological toxicity—another major contributor to clinical attrition.[23,24] Accounting for the long time lag between compound synthesis and clinical trial evaluation, the increase in the clinical success rates of drug candidates from around 2010 can be rationalised by the beneficial application of such guidelines.[18,22]

Another interesting study by Lovering *et al.* observed a significant increase in the average saturation—measured as the fraction of sp^3-hybridised carbon atoms (Fsp^3)—of drug candidates as they transition through the stages of clinical trials. This analysis suggested that saturation increases the chance of success; the authors explain this trend partially by the positive correlation between saturation and aqueous solubility resulting in increased bioavailability.[25] Whilst saturation seems to provide some advantages, the extremely broad statistical distribution of Fsp^3 in approved drugs (10th percentile 0.08 and 90th percentile 0.83 for drugs approved between 1900 and 1997) exposes that saturation is not a requirement for a successful drug candidate.[22]

Aside from the control of physicochemical properties of drug candidates, in recent years, alternative strategies have been employed to improve clinical success. Advancements within information technology and the availability of large data sets have made the use of machine-learning (ML) methodologies possible for analysing clinical failures and successes in great detail.[26] A recently developed supervised ML model using a random forest approach was able to predict

the success of drug candidates during phase II and III clinical trials with accuracies up to 85%.[27] Therefore, such algorithms are demonstrated to be extremely valuable for companies facing difficult decisions about whether a certain candidate should enter these high risk and cost trials.

1.3 Established Lead Generation Methods

As discussed earlier, the development of high-quality drug candidates is the principal objective of all drug discovery programs, and to that end, the generation of multiple lead series is paramount. The lead generation phase includes the whole process of identifying hits and developing them into compounds with sufficient *in vitro* activity alongside the elimination of liabilities related to undesirable physicochemical properties or chemical motifs that would likely lead to failure during later stage tests such as *in vivo* trials. Assessment of these lead series should provide researchers with enough information for decision making on the future of the project, and when high-content lead compounds are identified, they can be further optimised into more promising drug candidates.[28]

The two, fundamentally different lead generation approaches—target-based and phenotypic drug discovery, detailed below—can often be seen as complementary techniques used concurrently or sequentially to form integrated methods combining the information and productivity from both strategies.[29] Furthermore, although computational analyses and ML alone are incapable of generating leads, they can substantially increase the efficiency of all stages of any drug design program.[11,26]

1.3.1 Phenotypic drug discovery

Phenotypic drug discovery (PDD, also known as phenotypic-based, physiology-based or systems-based drug discovery) is a paradigm originating from the earliest drug discovery endeavours prior to the era of genetics and molecular biology.[28,30] A 'pure' PDD programme does not require any information on the identity or structure of the therapeutic target, or understanding of the biology of the disease and mechanism of action—rather it treats the biological system as a 'black box'. Hits, leads and drug candidates are assessed based on their physiological effects in disease-relevant models such as animal models, isolated organ systems and natural cells. The major difficulties with this approach include the development of accurate disease models, the ability to measure and analyse complex and often non-trivial outputs, as well as the construction of a high-quality screening library (further discussed in section 1.4) consisting of diverse 'lead-like' compounds. Furthermore, structure-activity relationship (SAR) studies are generally much more difficult to conduct in the absence of structural information.[28–31]

Aside from these challenges, PDD has several significant advantages and accordingly continues to gain popularity within pharmaceutical research. Firstly, it has the ability to provide instantaneous data on both toxicity and efficacy throughout the whole programme. Another benefit of PDD is that the use of living cells allows for genes, proteins, and pathways to interact providing a much greater number of possible targets and recognition sites, thus increasing the biological space for therapy. As a result, although often technologically difficult, phenotypic screening can identify both valuable novel biological targets and hit compounds simultaneously, and thus increase the efficiency of R&D campaigns. Therefore it is no surprise that the majority

of first-in-class small molecule drugs approved by FDA between 1999 and 2008 were developed in PDD programs.[28–31]
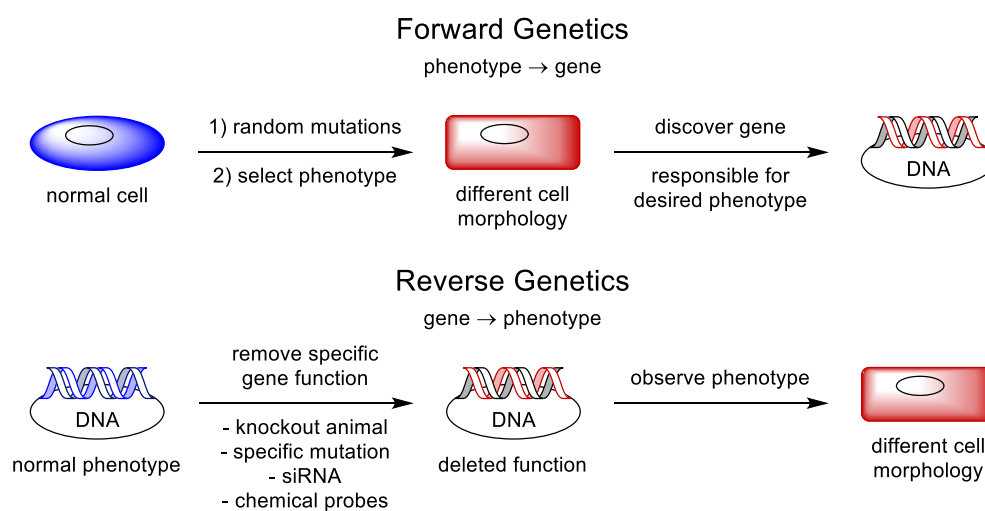
1.3.2 Target-based drug discovery

Target-based drug discovery (TDD) can be viewed as a reductionist approach, where it is assumed that the cause of a disease is a single gene product, and therefore, the most efficient and least ambiguous method for developing a therapeutic is by investigating the isolated gene product that is suspected of causing the physiological defect. Although often criticised for its limited set of poorly validated targets and high clinical attrition rates due to off-site toxicity of candidates, TDD is arguably the most effective approach for developing potent binders of single well-characterised targets due to gaining full advantage from structural information in SAR studies and rational design during hit elaboration. However, the underlying concept renders TDD highly susceptible to the quality and physiological relevance of the target.[28–30] One area within the pharmaceutical industry where targets are sufficiently characterised and validated such that TDD can thrive is the generation of follower drugs.[28] A recent analysis found that 73% of FDA approved small molecule follower drugs originated from TDD methods.[30] They are also well suited to develop first-in-class therapeutics for diseases caused by the inheritance of a single gene mutation such as cystic fibrosis and sickle cell anemia.[28]

Unlike in PDD where the overall lead generation process consists of only the hit generation and hit-to-lead development stages, within TDD the same process requires four stages: 1) target identification and 2) validation, 3) hit generation and 4) hit-to-lead development.[10]

1.3.3 Target identification and validation

The overall success of TDD campaigns is primarily determined by the identity of the target and therefore the stages of target identification and validation are critical.[28–30] To that end, biological systems have been extensively studied to identify target genes responsible for various physiological conditions. Forward genetics applies the screening of randomly mutated cells based on their phenotypes to identify unusual morphologies and ultimately the gene(s) responsible for the phenotype(s) of interest. On the other hand, reverse genetics involves the phenotypic characterisation of model organisms with loss of specific gene functions, usually using gene knockout animals or small interference ribonucleic acid (siRNA) techniques. If available, chemical probes can also be used to modulate the activity of genes either by effecting their expression or the coded function. Both approaches are commonly applied to aid the identification of desirable biological targets (Scheme 1.1). Typically, reverse genetic studies are employed to validate the chosen target, which is then used in initial *in vitro* and *in silico* screening assays to identify hit compounds.[10,28,32]

Forward Genetics
phenotype → gene

normal cell — 1) random mutations, 2) select phenotype → different cell morphology — discover gene responsible for desired phenotype → DNA

Reverse Genetics
gene → phenotype

DNA, normal phenotype — remove specific gene function (- knockout animal, - specific mutation, - siRNA, - chemical probes) → DNA, deleted function — observe phenotype → different cell morphology

Scheme 1.1 Forward genetics utilises phenotypic screening to identify the gene responsible, whereas reverse genetics uses specific removal of gene functions to identify genes responsible for certain phenotypes.

1.3.4 Hit generation methods

Following the identification and validation of a suitable biological target, the next crucial objective of any TDD programme is to find hit compounds possessing the desired activity against the target. In contrast, in PDD programmes hit identification is the first stage following the development of a suitable assay. The apparent hits identified during the initial screening must first undergo hit disqualification processes to remove compounds with promiscuous binding or interference with the assay itself, as well as orthogonal assays to produce validated hits providing chemical starting points for further development of drug candidates.[11,31,33]

Generally speaking, methods can be divided into two major categories based on the size of the compounds within the screening libraries: lead-like molecules (discussed in this section) and fragments (discussed in section 1.5).

1.3.4.1 Screening lead-like compounds

Conceptually the most straightforward method to generate leads is to screen collections of compounds with similar physicochemical properties. Hits are required to be relatively potent, usually possessing low or even submicromolar activities, in order to facilitate both their identification and development into sufficiently active leads. Approaches for this include phenotypic and high-throughput screening.[28,33]

1.3.4.2 Phenotypic screening

Phenotypic screening is analogous to a forward genetic study as it is based on complex model organisms, with the exception that the physiology of the cells is altered using random small molecules rather than random mutations. Compounds causing a desired phenotype are considered to be hits. Although this technique is primarily used in PDD rather than TDD, targets

are often identified, validated and characterised following the initial screening, and as a result, the hit compounds can then enter target-based rational hit-to-lead designs. In this regard, phenotypic screenings can provide both novel targets and hit compounds for TDD programs. It is hypothesised that such an integrated drug discovery approach might combine the advantages of both PDD and TDD whilst concomitantly overcome some of their weaknesses, overall substantially increasing the efficiency of the programme.[29]

1.3.4.3 High-throughput screening

High-throughput screening (HTS) is not a well-defined method, rather it is used collectively within the context of drug discovery for technologies that enable rapid assaying.[34] Robotics for accurately handling small volumes of liquids and imaging such as fluorescence resonance energy transfer (FRET) for the rapid and automated reading of the outputs are examples of essential techniques that were developed in the period leading up to 1990's enabling the faster and more efficient screening of the rapidly expanding lead-like libraries owned by pharmaceutical companies.[34,35] Further innovation and miniaturisation yielded ultra-high-throughput screening (uHTS) methods capable of screening 100,000 compounds per day using microplates, or performing and screening 100 million enzyme reactions in 10 hours using drop-based microfluidics.[35,36]

Methods can be divided into phenotypic and target-based HTS as the assays used can involve either living cells or purified single proteins as targets respectively; therefore, HTS has utility in both PDD and TDD programmes. Furthermore, the screening libraries may consist of various chemical entities such as nucleic acids and analogues, peptides, natural products and synthetic small molecules.[34,37] As a result, HTS is one of the most versatile methods available for hit generation. The overall success rates of HTS campaigns producing hits that could be developed into drug candidates is found to be about 50% for most target classes, and the applicability of this approach is demonstrated by the fact that one third of the drugs with known starting leads approved by FDA between 1991 and 2008 originated from HTS.[37] However, target-based HTS has been subject to numerous criticisms, most of which are either related to intrinsic limitations of TDD programs or the quality of the commonly used combinatorial libraries.[37]

1.3.4.4 Virtual screening

The field of artificial intelligence has grown exponentially over the past few years due to the advances of powerful computer hardware such as graphical processing units making numerically intensive parallel processing faster. ML algorithms, such as those designed for deep learning, can be used to build valuable models based on available research data, that can be used to predict biologically active compounds. Ligand-based approaches use information on known active and inactive compounds to screen virtual libraries by similarity searches and pharmacophore modelling to identify likely hit molecules. Since this approach does not require any information on the target, it can be used in both virtual target-based and phenotypic screenings. Structure-based methods on the other hand require structural information on the target elicited from

techniques such as X-ray crystallography or nuclear magnetic resonance (NMR) and use sophisticated and computationally demanding docking algorithms to predict ligand binding affinities. As a result, when both ligands and target structure are available, hybrid approaches are often employed whereby a huge virtual library is first filtered using a fast ligand-based method and then the resultant manageable compound set is docked against the target to further narrow it down. Compounds identified in a virtual screen are treated as apparent hits and can be experimentally validated.[26,31,38]

1.3.4.5 Miscellaneous methods

The newly emerged DNA-encoded library (DEL) technology can be regarded as a powerful alternative to target-based HTS. From hundreds of millions to even trillions of compounds are synthesised in a combinatorial fashion and conjugated to unique DNA sequences used as bar codes; the library is then screened as a mixture for affinities for immobilised protein targets, and conjugates with the desired affinity are decoded to reveal the hit structures.[39,40] Another alternative is to conjugate compounds to specific locations on the surface of a glass chip forming small molecule microarrays. These can then be incubated with protein solutions, either purified or complex mixtures (e.g. cell lysates), and the bound target can be detected with a fluorescent-labelled antibody.[41,42] Although both approaches can provide cost efficient HTS technologies, there are several major limitations: the conjugation chemistry can interfere with target binding; the screens only give information on affinities toward targets, but not on the function; hit validation requires the synthesis of new compounds (with no conjugation); and the target must be known.[31]

On the other hand, chemical biology approaches can take full advantage of identifying both new ligands and targets at the same time in living cells, and as such, they are closely related to phenotypic screening. However, the use of fully functionalised fragments allows the detection of weakly binding fragments by covalently modifying the target using either an electrophilic or a photoreactive warhead, and easy identification of the target via introducing a fluorescent label using a bioorthogonal 'click' reaction on the alkyne moiety.[31,43]

## 1.4 Screening Libraries

1.4.1 Library design

Individual compounds must possess certain properties to be suitable for screening collections, but their exact nature is dependent on the applications and subject to debate in the literature.[44] The basic physicochemical properties should roughly adhere to the Ro5 to ensure sufficient water solubility, reduce the risk of protein aggregation and promiscuous binding due to excessively high lipophilicity, and limit the size of the hit compounds making them acceptable starting points for drug development.[28] Reactive functional groups such as alkyl halides, peroxides, thiols, epoxides, aziridines, aldehydes, reactive Michael acceptors, *etc*. should be avoided to minimise chemical reactions and covalent modifications of the targets often resulting in aspecific binding.[45] Furthermore, several structural motifs have been identified to cause promiscuous binding or false positives known as pan assay interference compounds (PAINS), including alkylidene barbiturates and heterocycles, hydroxyphenylhydrazones, phenolic Mannich bases, catechols, quinones, dialkylanilines and diazo compounds.[45,46] Synthetic tractability, or the ability to synthesise analogues efficiently, is also an important aspect when considering the hit-to-lead optimisation of any potential hit compound.[28,47]

The characteristics of the library as a whole is also extremely important and should match the requirements of the drug discovery programme in which they are employed. In general we can conclude that although screening a larger library creates more data, in terms of identifying tractable hits, quality is more important than quantity.[28]

1.4.2 Chemical space

The concept of the whole chemical space describes all the possible chemical entities from single atoms through organic and inorganic molecules to polymers and nanostructures. This can be further limited to the chemical space of small organic molecules by using the 'definition' from section 1.1; however, specifying the more relevant 'drug-like', 'lead-like' or 'biologically relevant' chemical spaces is a lot more challenging.[48] An early estimate by Bohacek *et al*. hypothesised that the 'universe of organic molecules' might consist of more than 10^{60} compounds with no more than 30 C, N, O and S atoms.[38] The 'biologically relevant' chemical space can also be approximated by considering chemically stable small molecules made of C, H, N, O, S and halogen atoms disregarding any constraints about physicochemical properties. Ruddigkeit *et al*. enumerated a virtual library of 166 billion such molecules comprising up to 17 non-hydrogen atoms.[49] This computational study also showed that the number of possible molecules is exponentially growing with the heavy atom count (HAC), and each added atom increases the number almost seven-fold. Therefore, the total number of 'drug-sized' molecules (HAC up to 35) can be estimated to be in the order of 10^{27}; although significantly smaller than the previous estimate, chemical space can still be considered as almost infinite in terms of any practical considerations, and no screening library can contain all possible molecules. However, compounds

closely related to one another can be assumed to have similar biological activity, and so, areas within chemical space containing molecules with a certain activity against a target could be expected.[38] Having even just a single compound in a screening library that falls into such a desired subspace can be enough to identify a hit, and then a focused library of analogues could be synthesised and screened to identify more potent compounds and eventually a lead. Therefore, the sampling of chemical space, whether the whole space or just an area of interest, is an absolutely crucial property of any screening library (Figure 1.3).[48]
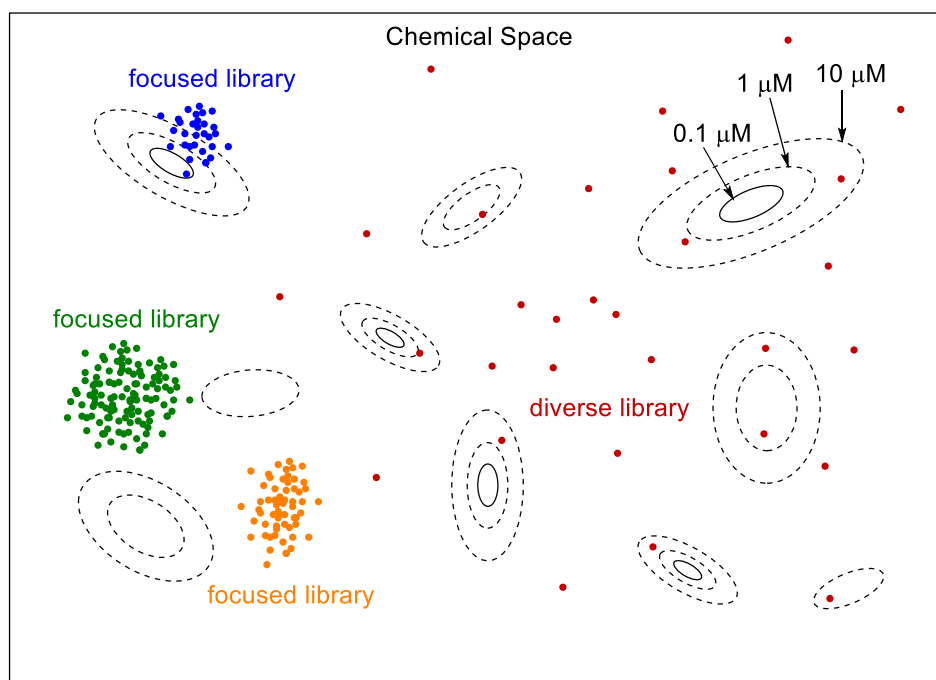
Figure 1.3 Graphical representation of chemical space. Contours show areas of compounds with activities against specific (hypothetical) targets. A diverse library gives a broad coverage of chemical space, and more likely to feature compounds with activities against different targets. Focused libraries give dense coverage of small areas of chemical space.

1.4.3 Complexity

The concept of complexity is often regarded as a crucial element for library design; however, it is extremely difficult to define or quantify. According to the intuitive notion of molecular complexity and assuming complementary protein-ligand binding, a less complex molecule should be more likely to 'match' any target (potentially in more than one way). Thus, ignoring limitations on binding affinities, a lower complexity should result in both higher hit rates and promiscuity.

A very simple model by Hann *et al.* treating both ligands and receptors as one-dimensional strings of binary data has demonstrated that the probability of binding—defined as a perfect complementary match between the ligand and the receptor—decreases dramatically with the increasing complexity (in this case, simply the length) of the ligand.[50] On the other hand, the probability of measuring a binding is highly dependent on its strength, which in this model is proportional to the number of matching features and so the size or complexity of the ligand. As

a result, the probability of a 'useful event'—defined as a measured binding—plotted against the ligand complexity (or length) goes through a maximum value (Figure 1.4).
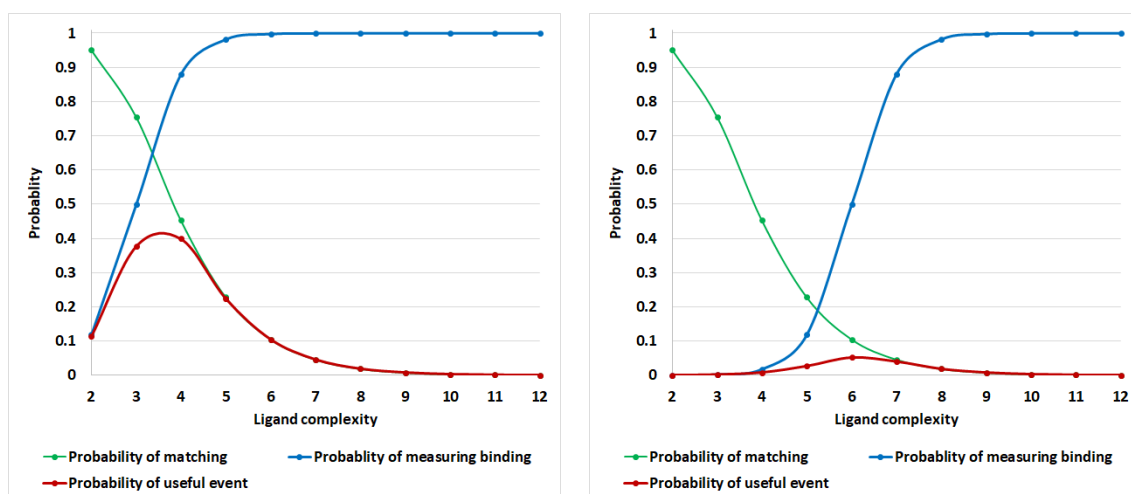
Figure 1.4 The probability of useful event (measured binding) has a maximum value depending on the sensitivity when using a simple complexity model. Values of 'probability of matching' were reproduced from Hann *et al*.[50] Probability of measuring binding is calculated by $P = \frac{1+\tanh(M-S)}{2}$, where M is the complexity of the ligand and S is the sensitivity of the assay. Probability of useful event is the product of the probabilities of matching and measuring binding. The binding site complexity is 12; the sensitivities are 3 (left) and 6 (right).

This further suggests that, depending on the sensitivity of the assay, there should be an optimal ligand complexity that maximises the probability of a useful event, which can be correlated to a hit in a real compound screening. This idea prompts the general paradigm of drug discovery whereby 'less complex' molecules should be screened first to produce more hits, that could then be developed into 'more complex' leads with decreased promiscuity providing high selectivity and low off-target toxicity.

From a mechanistic point of view, the complexity of a compound could be related to the number and types of constituent atoms or even the connectivities between them. Therefore, HAC is often used to estimate the complexity of molecules – the larger the molecule, the more complex it is. This concept is easily related to the arbitrary complexity used by Hann's model.[50] Indeed, whilst hit rates of 0.001–0.151% for the identification of lead-like ligands in HTS programs with micromolar potency thresholds were observed, hit rates of 3% or more were achieved by screening small fragment-sized compounds using more sensitive methods to detect millimolar binders.[51] A number of studies tested the validity of Hann's original complexity model by collecting promiscuity data from numerous screenings, and, although the result vary considerably, there is more evidence suggesting that contrary to the model's prediction, promiscuity increases with MW.[52,53] This observation was rationalised by the fact that, unlike as was assumed in the original model, a perfect match is not required for observing binding; if instead a modified model is used where only a certain minimum number of complementary

features is required for a match, then the probability of a match, and therefore of a useful event or hit, is actually increasing with the size or complexity of the ligand.[53] The true nature of ligand bindings is probably somewhere in between these two extremes of the model, and an attempt has been made to create a more realistic model with better correlation to experimental data.[54]

Apart from its size, other properties of a molecule are also thought to contribute towards its complexity. The concepts of shape and stereochemical complexity, defined by Fsp^3 and the number of chiral centres respectively, are often used to describe the size-independent complexity of compounds by arguing that a complicated three-dimensional (3D) object is less likely to match different binding pockets resulting in lower promiscuity,[55] which might offer an alternative explanation for the observed increase of these properties throughout clinical trials.[25,56]

Even more importantly, it has been observed that interactions between targets and ligands with higher information content, such as directionality or charge gradient, are more difficult to position correctly compared to low information content interactions such as most hydrophobic interactions.[57] As a result, compounds displaying several polar groups, especially hydrogen-bond donors, acceptors and charged groups, tend to be more selective and therefore can be regarded as more complex. This is also in agreement with other studies relating high lipophilicity to promiscuity and toxicity.[23,24]

1.4.4 Diversity and three-dimensionality

Structural diversity is used to describe the variation in the 3D shapes of molecules within a library and is typically divided into four main components: 1) appendage diversity, 2) functional group diversity, 3) stereochemical diversity, and 4) scaffold diversity. The key contributor towards structural diversity however, is the scaffold diversity since the scaffold of a molecule determines its overall 3D shape.[47]

A comprehensive scaffold analysis on more than 2.4 million compounds from 17 different commercially available screening collections highlighted that framework-distributions are generally extremely uneven. Large combinatorial libraries, generated in industrial settings, greatly overrepresent very few and simple scaffolds, many of which overlap between separate suppliers. On the contrary, much smaller but more diverse libraries do exist based on more complex scaffolds, and often originate from academic laboratories or natural sources.[47] A similar study of over 24 million different organic compounds in the CAS registry found that only 0.25% of all known scaffolds are found in almost half of all molecules in the registry, displaying the very low scaffold diversity.[58]

Furthermore, it is often found that current screening libraries overrepresent 'flat' sp^2-rich (hetero)aromatic molecules limiting the covered areas of chemical space.[25,59] A comparative shape analysis demonstrated that whilst most compounds within a commercial collection are indeed quite 'flat', the majority of possible lead-like molecules show high levels of three-

dimensionality.[49] Therefore, in order to obtain better coverage of the potentially biologically relevant chemical space, it is important to develop collections of compounds featuring chiral centres and high Fsp^3, both of which can increase the 3D nature of small molecules.[25,59,60]

1.4.5 Types of libraries

Libraries can be characterised according to their chemical space coverage. Whilst focused libraries give dense coverage of a small area, diverse libraries aim to sample as large an area, and ideally as evenly, as possible (Figure 1.3). Furthermore, they can be divided into two categories based on the type of small molecules employed: either natural or synthetic.

1.4.5.1 Focused libraries

When a chemical starting point for a project is known or assumed (e.g. native substrates for enzymes or ligands for receptors, previous hit or lead compounds for the target, binders for structurally related targets, hits obtained from virtual screenings or developing follower drugs), focused libraries of compounds similar to the starting point are often used. These approaches are anticipated to give higher hit rates, and are especially suitable for establishing SAR studies during hit-to-lead development.[47,60,61]

1.4.5.2 Diverse libraries

If no reliable structural information is available on the target, binding pocket(s) or ligand(s), or if the identification of a novel target, binding mode or compound class is desired, then screening a structurally and thus functionally diverse collection increases the chance of finding desired hits.[6,47,60,62–66]

1.4.5.3 Natural products

One extremely important source of small molecules is nature itself. Secondary metabolites or natural products are produced by all living organisms and generally feature high levels of complexity and diversity.[67] As a result, natural products show activity against a wide range of biological targets and have been used in medicine for many centuries. Modern drug discovery still benefits from natural products and around half of all FDA approved drugs are derived from natural products.[68] However, there are major challenges associated with using natural products to generate screening libraries including sourcing, isolation, purification and identification of the active compounds as well as often poor synthetic tractability due to the lack of efficient synthetic methodologies enabling the construction of such complex molecules.[69]

1.4.5.4 Synthetic compounds

Deliberate chemical synthesis is arguably considered the most efficient approach for the generation of small molecule collections.[63] In fact, synthetic libraries have been widely used in pharmaceutical research for over a century and provided a very important basis for identifying hit compounds in biological screenings. It is well established that the quality of the screening

collections is crucial for discovering novel biologically active molecules, and yet, the synthesis of structurally diverse and biologically relevant small molecule libraries remains a huge challenge for synthetic chemists.[70]

1.4.6 Synthetic approaches

1.4.6.1 Biology-oriented synthesis

Due to the vastness of chemical space, designing individual compounds for screening collections that maximises the success rates of identifying tractable hits in any assay is very difficult. Based on the success of natural products in drug discovery and the theory that proteins and their cognate ligands and substrates have co-evolved, it is assumed that related compounds are more likely to fall into areas of chemical space with biological activity. Therefore, biology-oriented synthesis (BIOS) aims to generate compounds based on scaffolds found within natural products and thus create arguably diverse libraries of molecules with anticipated higher biological relevance or hit rates in screens.[71,72]

1.4.6.2 Target-oriented synthesis

Target oriented synthesis (TOS) utilises retrosynthetic analysis to devise synthetic routes to access compounds from simple starting materials and reagents. Therefore, it is often the best or only available approach for the synthesis of specific complex molecules like late-stage lead-like compounds and drug candidates, and it is also well suited for the generation of focused libraries of such molecules. However, since most synthetic routes are quite long, and diversity is usually only created in the appendages during final steps, this method is often considered to be inefficient for generating diverse libraries for initial screening since the resulting compounds occupy a small and confined region of chemical space.[59,73,74]

1.4.6.3 Combinatorial synthesis

The application of combinatorial synthesis using highly automated split-and-pool techniques in the 1990s enabled the formation of millions of small molecules for biological screenings to populate commercially available libraries and pharmaceutical compound collections.[75] However, as an inherent feature of the synthetic approaches employed, the resulting libraries possessed enormous numbers of molecules that were based on limited number of distinct molecular scaffolds, resulting in libraries where most compounds cluster in a few small areas of chemical space.[66] Moreover, as a direct result of most compatible methodologies utilised for combinatorial synthesis as well as the available starting materials, the resultant libraries tend to be dominated by 'flat' (hetero)aromatic scaffolds. These libraries were revealed to be much less successful in most biological screenings than initially anticipated, something which is often attributed to the lack of structural diversity.[76] However, it was later realised that these techniques can indeed be very effective for the synthesis of focused libraries.[47,60,61]

1.4.6.4 Diversity-oriented synthesis

An important approach emerged in the late 1990s which utilised a forward synthetic analysis to construct divergent synthetic pathways that generate several small molecules bearing different scaffolds from common starting materials. In 2000, Stuart L. Schreiber described this new design as diversity-oriented synthesis (DOS), with the aim of producing structurally diverse compound collections, to be used especially for identifying biological targets and their modulators simultaneously in phenotypic screenings.[73] This strategy benefits from synthetic efficiency since complex and structurally diverse small molecules are generated in typically 3 to 5 synthetic steps and introducing high levels of scaffold diversity yielding powerful screening collections.[66,74]

Within DOS, two main methods are commonly adopted to achieve scaffold diversity.[74] The branching or reagent-based approach uses different reagents to form diverse scaffolds from a single multifunctional or pluripotent intermediate in a divergent manner (Scheme 1.2, A).[77] Alternatively, the folding or substrate-based approach utilizes subtle differences in the structures of diverse intermediates to form different scaffolds under the same reaction conditions (Scheme 1.2, B).[6] Often, however, DOS strategies use elements of both approaches.[66]
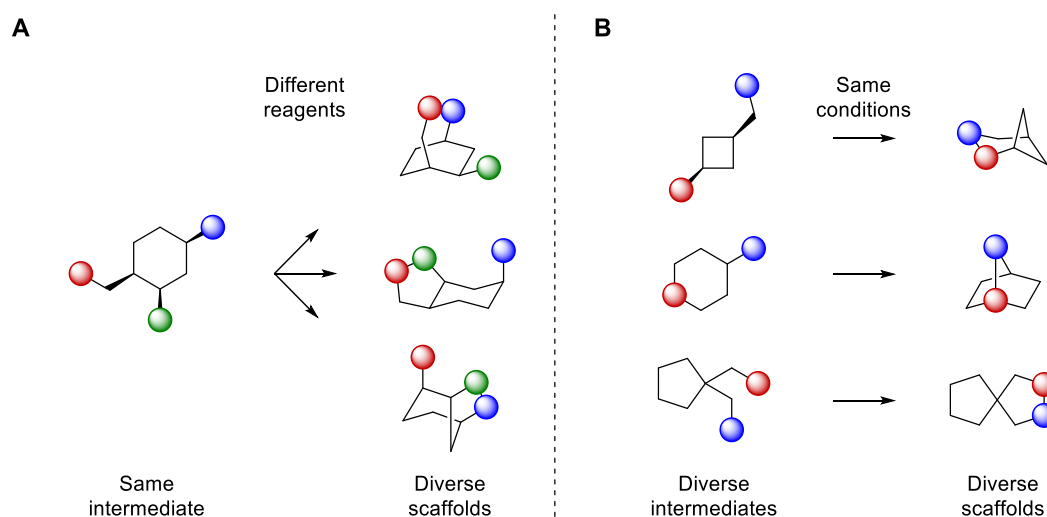
Scheme 1.2 Reagent-based DOS approaches generate complex and diverse scaffold collections from single common intermediates using different reagents (A) and substrate-based approaches achieve similarly diverse and complex small molecule libraries from diverse sets of highly functionalised intermediates under common reaction conditions (B).

1.5 Fragment-Based Drug Discovery

Fragment-based drug discovery (FBDD) is a well-established target-based lead generation strategy used within both industry and academia. Instead of screening lead-like compounds, FBDD focuses on identifying low MW hits, that can then be grown into potent drug candidates or chemical probes.[32,78]

The first report of obtaining SAR and developing a potent nanomolar binder based on fragments utilising NMR studies was published by Shuker *et al*. in 1996.[79] Although the approach faced a significant amount of internal resistance within most companies, following several other successes and major technological advancements in the field, FBDD is now recognised as a validated and popular alternative to other lead generation methods within industry with overall success rates competitive with those of HTS.[80–83]

The success of FBDD is showcased by the development of numerous clinical candidates and four FDA-approved drugs: Vemurafenib, Venetoclax, Erdafitinib and Pexidartinib for the treatment of late-stage melanoma, chronic lymphocytic leukemia, urothelial carcinoma and tenosynovial giant cell tumor respectively (Figure 1.5).[81,84–87] Screening fragment libraries has two main advantages over other approaches: better chemical space coverage of libraries and more optimal interactions with the protein targets (discussed in more detail below).

Venetoclax
anticancer

Vemurafenib
anticancer
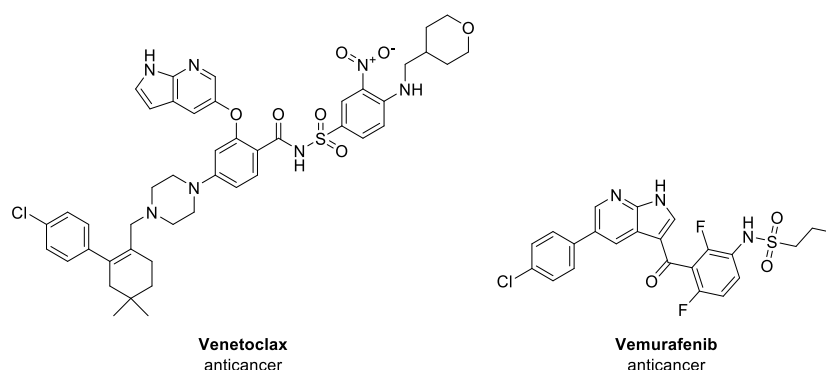
Figure 1.5 Examples of the first two FBDD-derived FDA-approved drugs: Venetoclax and Vemurafenib.

1.5.1 Chemical space coverage

As described in section 1.4.3, the number of possible chemically stable organic molecules grows exponentially with size, and the fragment-like chemical space (HAC up to 16) has been evaluated to contain about 57 billion compounds.[49] Therefore, a relatively small library of a few thousand fragments can give exceedingly better coverage of the relevant chemical space (10^3 vs 10^{10}) than HTS libraries (10^6 vs 10^{27}) or even the largest DELs (10^{12} vs 10^{27}). As a result, fragment screens often yield several hits against targets, even when no hits could be identified using more traditional techniques such as HTS. This is especially beneficial when targeting traditionally 'undruggable' targets, such as protein-protein interactions (PPIs) with large and shallow binding pockets, where HTS is often unable to produce tractable hits.[88,89] Indeed, Venetoclax (Figure 1.5)

was developed from fragment hits against the oncogenic B-cell-lymphoma-2 protein inhibiting a key PPI that results in the activation of an intrinsic apoptosis pathway.[90]

1.5.2   Interactions and hit-to-lead development

Due to their small size, fragments inherently form very few interactions with the protein targets resulting in weak binding with affinities usually in the range of high micromolar to low millimolar. As a result, more sensitive techniques such as NMR spectroscopy, X-ray crystallography and surface plasmon resonance (SPR) or the recently emerging microscale thermophoresis, thermal shift assay and weak affinity chromatography need to be employed when screening fragment libraries.[81] Whilst the commonly used NMR and X-ray methods provide crucial structural information on the binding mode of the initial hits enabling rational design, other methods such as SPR or *in vitro* inhibition assays are also crucial to measure binding affinities or activities in order to establish SAR for further development.[78]

The Gibbs free energy of binding (ΔG) determines the affinity of a ligand toward its target expressed as the dissociation constant (K_d) according to Equation 1.1:

$$\Delta G = RT \ln(K_d) \tag{1.1}$$

where R is the universal gas constant and T is the absolute temperature. Assuming that binding results in inhibition, K_d can be equated to the half maximal inhibitory concentration (IC_{50}) obtained *in vitro*. Assuming standard conditions, Equation 1.1 can be reformulated as Equation 1.2 to give ΔG in kcal mol^{-1} calculated from IC_{50} values given in units of M:

$$\Delta G = 1.37 \text{ kcal mol}^{-1} \log\left(\frac{IC_{50}}{M}\right) = -1.37 \text{ kcal mol}^{-1} \text{ pIC}_{50} \tag{1.2}$$

Therefore, it follows that a weakly binding fragment hit with an IC_{50} of 1 mM would have a free energy of -4.1 kcal mol^{-1}, whereas a strongly binding drug candidate with an IC_{50} of 1 nM would have a free energy of -12.3 kcal mol^{-1} which is three-times greater.

Using the definition of Gibbs free energy, ΔG can be further divided into entropic (ΔS) and enthalpic (ΔH) contributions according to Equation 1.3:

$$\Delta G = \Delta H - T\Delta S \tag{1.3}$$

The rigid-body entropy loss of binding ($T\Delta S$) under standard conditions has been estimated to be in the range of 3.6 – 4.8 kcal mol^{-1} for a small molecule.[91] Thus, in order for a fragment to off-set the entropy loss and give a detectable binding affinity, the sum of interactions must give an enthalpy of at least -8 kcal mol^{-1}, while the aforementioned drug candidate needs about -16 kcal mol^{-1}. This shows that if a fragment hit can be elaborated to twice its size whilst maintaining the average binding enthalpy per size, a highly potent lead compound could be developed.

Therefore, to help researchers focus on progressing the most appropriate compounds, ligand efficiency (LE) has been introduced giving a measure of the average binding free energy

contribution of each heavy atom (Equation 1.4); although LE has units of kcal mol^{-1}, it is most often reported as a unitless number.[92] If the fragment mentioned above had a HAC of e.g. 10, this would give a LE of 0.41, keeping LE constant a HAC of 30 would be required to reach the 1 nM potency, which is a reasonable size for a drug candidate. Since increasing lipophilicity often results in promiscuous binding as well as increased potency against the target, the lipophilic LE (LLE) can be used instead of IC$_{50}$ to compare the utilities of compounds with different lipophilicities (Equation 1.5), whereas the heavy atom adjusted LLE (LLE$_{AT}$) is the alternative of LE (Equation 1.6).[57]

$$LE = -\frac{\Delta G}{HAC} = 1.37\frac{pIC_{50}}{HAC} \qquad (1.4)$$

$$LLE = pIC_{50} - \log P \qquad (1.5)$$

$$LLE_{AT} = 0.111 + 1.37\frac{LLE}{HAC} \qquad (1.6)$$

Lead-like hits, such as those identified by HTS, possess relatively high potency arising from multiple suboptimal interactions resulting in much lower LEs making further optimisation rather difficult. A high-quality fragment hit, however, forms few optimal interactions with the target in order to bind with a sufficient affinity, and these key interactions are often conserved throughout the hit-to-lead optimisation. These fragment elaboration strategies usually involve fragment growing, merging and linking (Figure 1.6). Fragment growing involves the synthetic expansion of the molecule in different direction where further interactions with the target are expected to form. These analogues are tested for activity or affinity, and the ones with the highest LEs are selected for further optimisation, ensuring that the newly added atoms also form near-optimal interactions. If two fragments bind within the same pocket, and there is a significant spatial overlap between them, it can be possible to synthesise a hybrid structure displaying features from both original molecules and keeping all the favourable interactions. If the two fragments do not overlap, there is still a good chance that they can be linked via an additional chemical linker region, even if the two fragments bind in different pockets within proximity.[91]

HTS hit optimisation
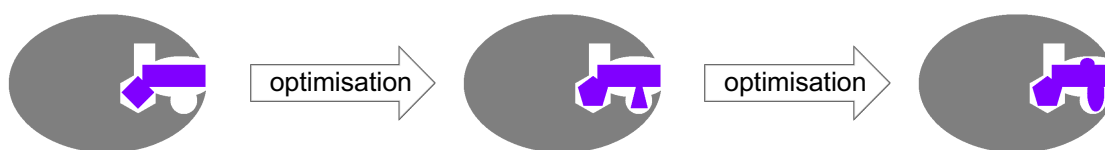
Fragment growing

Fragment merging

Fragment linking
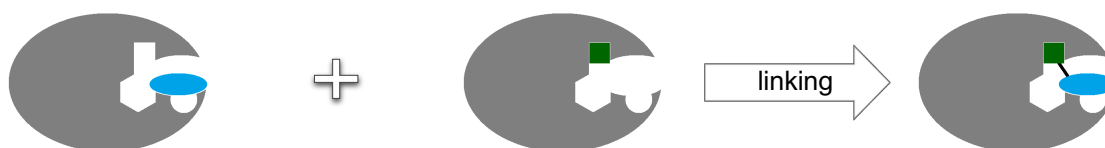
Figure 1.6 Graphical representation of different hit development strategies. HTS hits possess many suboptimal interactions and can be difficult to optimise. Fragment hits tend to form few optimal interactions and can be elaborated into more potent compounds conserving these interactions by growing into nearby areas of the pocket, or if multiple fragments bind to different regions, they can be merged or linked.

In theory, fragment linking can yield compounds with affinities much higher than the sum of the original hits. When two fragments (A and B) are linked together (forming compound AB), the efficiency of the linkage can be expressed by using the linking coefficient (*E*) defined by Equation 1.7:

$$E = \frac{K_{\mathrm{AB}}}{K_{\mathrm{A}} K_{\mathrm{B}}} \tag{1.7}$$

where K_{A}, K_{B} and K_{AB} are the dissociation constants for molecules A, B and AB respectively.[93] Using the expressions from Equations 1.1 and 1.3, Equation 1.7 can be reformulated as Equation 1.8:

$$RT \ln(E) = \Delta G_{\mathrm{AB}} - \Delta G_{\mathrm{A}} - \Delta G_{\mathrm{B}} = \Delta\Delta G = \Delta\Delta H - T\Delta\Delta S \tag{1.8}$$

where $\Delta\Delta G$, $\Delta\Delta H$ and $\Delta\Delta S$ are the changes in the binding free energy, enthalpy and entropy respectively as fragments A and B are linked to compound AB. An ideal linker allows for both fragments A and B to bind in their optimal mode without introducing any strain. Therefore, all the binding enthalpies add up resulting in $\Delta\Delta H \approx 0$, whereas the rigid body entropy of a small molecule is assumed to be independent of size, so $T\Delta S_{AB} \approx T\Delta S_A \approx T\Delta S_B \approx 4.1$ kcal mol^{-1} and $T\Delta\Delta S \approx 4.1$ kcal mol^{-1}. Substituting these values into Equation 1.8 gives a theoretical value of $E = 10^{-3}$.

In practice, ideal linking of fragments is not often achieved successfully; the difficulty is that a suboptimal linker can introduce unfavourable interactions with the target and strain either within the linker itself or by forcing fragments in suboptimal binding modes, as well as flexible linkers increasing the entropy loss on binding. As a result, linking coefficients in the range of 10^{-2} – 10^2 are often considered to be suitable, although values as low as 4.6×10^{-5} have been reported.[93] A typical example from the literature is the development of a potent inhibitor of β-secretase 1 (BACE1) by Jordan *et al*. via linking two fragments.[94] A number of linkers have been trialled and although the best linker still has a linking coefficient of 2.5, the resultant compound has an IC$_{50}$ of 0.8 µM and >2000-fold selectivity over the related aspartic protease cathepsin D (CatD) (Figure 1.7).

$K_{d,BACE1}$ = 114 µM

$K_{d,BACE1}$ = 140 nM

IC$_{50,CatD}$ = 510 nM

IC$_{50,BACE1}$ = 0.8 nM

IC$_{50,CatD}$ = 1.9 µM
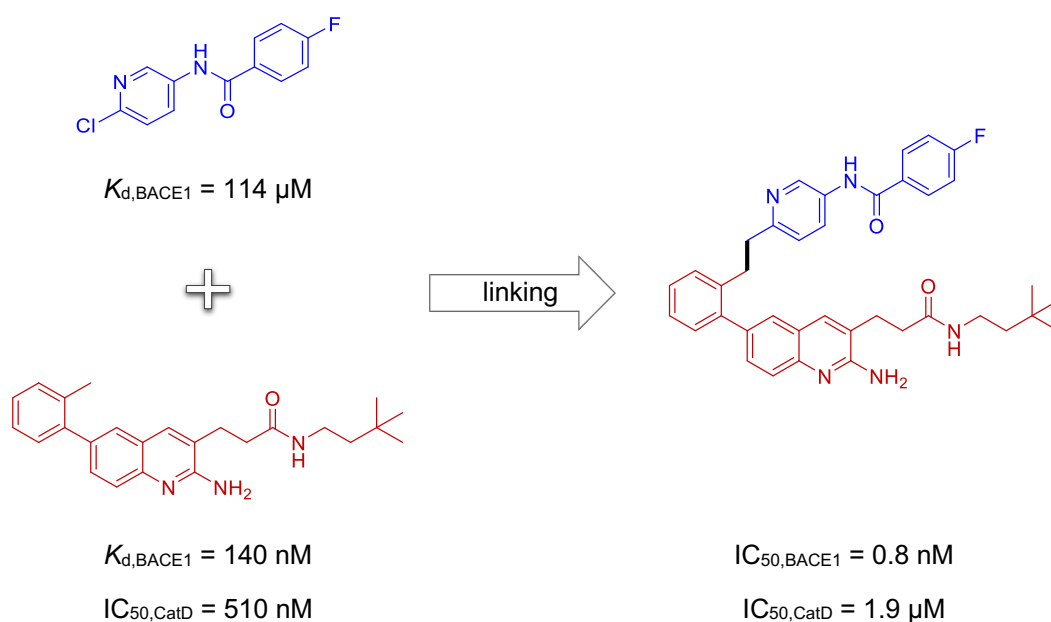
Figure 1.7 Example of fragment linking resulting in a potent and selective β-secretase 1 (BACE1) inhibitor.

1.5.3 Fragment library design

Given that fragment libraries form a special category within general screening libraries, all design criteria described in section 1.4 are relevant, however, there are several further considerations for designing suitable fragment collections.

1.5.3.1 Size and physicochemical properties

In order to take full advantage of the underlying principles of FBDD, the compounds in a screening collection must be of small size. Furthermore, due to the weak affinities, these libraries need to be screened at high concentrations to measure binding, therefore, fragments must be suitably water-soluble.

The first generally accepted guideline on the physicochemical properties came from an early analysis of fragment hits observed by researchers at Astex Pharmaceuticals and found that most hits obeyed a so called 'Rule of Three' (Ro3) (Table 1.2).[95] Later, a study on a much larger set of hits observed by X-ray crystallography showed that the highest hit rates were observed for fragments with HACs of 10 to 14 and clogP values between 0 and 2,[96] therefore, a new guideline with slightly narrower ideal ranges has been adapted by Astex with added emphasis on aqueous solubility (Table 1.2).[97]

Table 1.2 Guidelines on physicochemical properties for fragment libraries set by Astex.

Property	Rule of Three[95]	Astex Guidelines[97]
MW	< 300 Da	140 – 230 Da
HAC	–	10 – 16
HBD	≤ 3	–
HBA	≤ 3	–
clogP	≤ 3	0 – 2
RBC	≤ 3	≤ 3
TPSA	≤ 60 Å^2	–
chiral centres	–	≤ 2
aqueous solubility	–	≥ 5 mM

MW = molecular weight, HAC = heavy atom count, HBD = number of hydrogen-bond donors, HBA = number of hydrogen-bond acceptors, clogP = partition coefficient, RBC = rotatable bond count, TPSA = topological polar surface area.

1.5.3.2 Synthetic tractability and exit vectors

Once fragment hits are identified, the ability to obtain meaningful SAR information quickly and efficiently is paramount for a successful hit-to-lead optimisation. To that end, it is desirable that analogues of all fragments within the library are either commercially available or synthesisable in as few steps as possible. This can be facilitated if the scaffolds can be formed using robust chemistry that allows for the incorporation of different building blocks and reagents. An alternative approach is the utilisation of synthetic exit vectors for the generation of numerous

derivatives in short numbers of steps from the initial branching point hit compound, which can be synthesised on a larger scale. Such exit vectors can be especially useful for exploring the binding pocket further by growing the fragment in different directions.[80–83,98]

1.5.3.3 Shape and three-dimensionality

There is a considerable amount of debate in the literature about the ideal shapes of fragments.[81] The previously mentioned analysis of fragment hits obtained by Astex over a decade showed that two-dimensional (2D) fragments—defined by a mean deviation from planarity (DFP) of less than 0.05 Å—have an approximately 1.5 times higher hit rate when compared to more 3D ones.[96] This observation can be interpreted by the differences in molecular complexity, since a more 3D shape is expected to have more directional features resulting in higher complexity and a lower probability of matching the binding pocket. The higher hit rates and promiscuity associated with flat (hetero)aromatic compounds can also be explained by their intrinsic polarisability, which results in better adaptation to different environments.[57]

While a similar study by AstraZeneca also observed the same trend in the hit rates of 2D *versus* 3D fragments, more importantly, a detailed analysis showed that 3D fragments filled a significantly larger volume of the binding pockets examined, highlighting the importance of 3D fragments.[82] 3D fragments are also more capable of projecting exit vectors in all directions of 3D space allowing for much better exploration of binding pockets. Furthermore, large areas of chemical space can only be covered by molecules with 3D shapes. As a result, the consensus of the field appears to be that a good balance between 2D and 3D fragments should be maintained in screening libraries.[98]

Since there are many fewer possible 2D compounds than 3D ones within the same size limit, and given that synthetic methodologies for generating flat molecules are much better developed by medicinal chemists, a good coverage of 2D fragment space has been achieved early on—whereas the coverage of 3D fragment space is sparse in comparison.[99] To address this issue, the '3D Fragment Consortium' has been established by various academic groups and research organisations to create fragment screening libraries with enhanced 3D characteristics and encourage the development of new synthetic strategies for accessing 3D fragments with novel or underrepresented scaffolds.[100] Although synthetic efforts focused on generating such 3D fragment collections,[101–106] they remain underexplored.

A library of 38 highly saturated 3D fragments was recently published by Hanby *et al*. utilising diastereomeric α,α-dialkyl β-hydroxy cyclopentatone as a versatile building block (Scheme 1.3).[107] A DOS approach was efficiently used by the authors to afford 20 unique and structurally diverse molecular frameworks, each displaying several exit vectors for fragment growth. Importantly, the library showed optimal calculated physicochemical properties, a broad coverage of 3D chemical space and excellent natural product-likeness score[108]. However, the utility of these complex scaffolds featuring multiple chiral centres remains questionable for FBDD purposes.
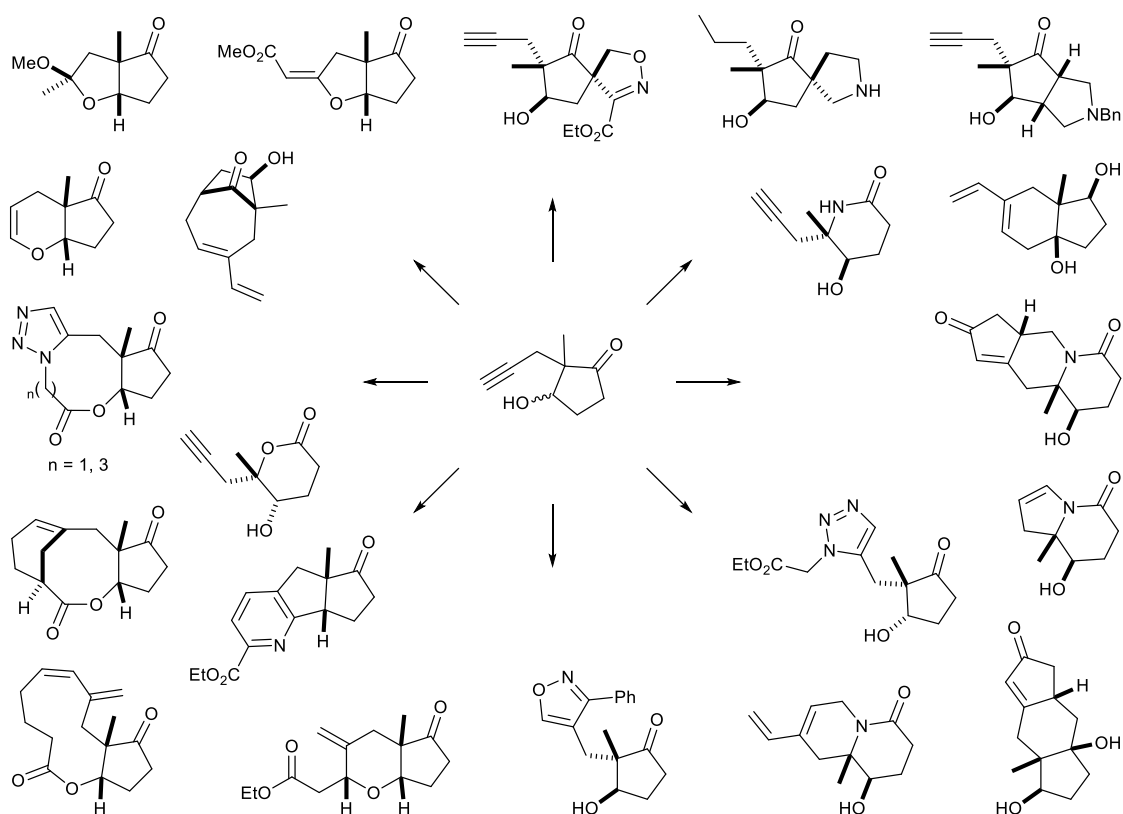
Scheme 1.3 Synthesis of a natural product inspired fragment library by Hanby *et al.*[107]

1.6 Overview

FBDD has been proven a powerful tool for developing new drug candidates in TDD campaigns, especially against 'undruggable' targets where traditional techniques such as HTS fails to provide tractable hit compounds. However, most commercially available libraries mainly consist of 'flat' (hetero)aromatic fragments, therefore, methodologies allowing for the generation of more saturated and 3D fragments could improve the chemical space coverage of screening collections and thus the overall success rate of FBDD campaigns. Hence, the first project described in this thesis (section 2) focuses on the synthesis of highly 3D fragments with spirocyclic scaffolds for potential use in FBDD.

Antibiotic resistance is one of the greatest challenges of modern healthcare. The most straightforward strategy to address this crisis is by developing new antibiotics, ideally, against novel targets to assure bacterial susceptibility. As such, the bacterial propionate mechanism has been studied and its inhibition trialled to attenuate the growth of bacteria in propionate-containing media. The second project within this thesis (section 3) describes synthetic efforts to elaborate simple fragment hits into a more potent chemical probe against a key enzyme involved in the propionate metabolism of *Pseudomonas aeruginosa*.

2 Spirocyclic Fragment Library

2.1 Introduction

2.1.1 sp³-hybridised quaternary centres

Within the literature, many different strategies have been developed to introduce 3D character within a molecule. Amongst these, the introduction of sp³-hybridised quaternary carbon centres into scaffolds has grown in interest. The three-dimensionality of these carbon centres arises from the presence of four groups pointing in different directions spanning the 3D space (Figure 2.1).[109]
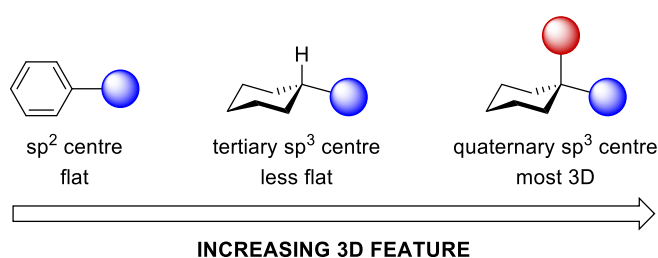
Figure 2.1 Introduction of sp³-hybridised carbon centres increases the complexity and 3D shape of the molecule, therefore increasing structural diversity of the library. The effect is even more dominant for quaternary sp³ centres compared to tertiary ones due to the projection of structural moieties in four different directions.

Indeed, there is strong presence of this structural feature within small molecule FDA-approved drugs. Using DrugBank Version 5.0,[110] it was shown that a significant proportion of these entities contained at least one quaternary sp³-hybridised carbon centre. The most common motif was found to be the all-carbon quaternary centre found in almost 14% of all marketed drugs. The most prevalent heteroatom-containing quaternary centres were those featuring oxygen (10.6%), followed by the less explored, but nonetheless significant class of aza-quaternary centres (3.2%) (Table 2.1).

Table 2.1 Analysis of quaternary sp^3-hybridised carbon centres in FDA approved drugs.

$$R_2 - \underset{\underset{R_1}{|}}{\overset{R_3}{C}} R_4$$

Type of centre	R_1	R_2	R_3	R_4	Drugs featuring the centre	
All-carbon quaternary	C	C	C	C	360	13.6%
Tertiary alkoxy	C	C	C	O	280	10.6%
Aza-quaternary	C	C	C	N	84	3.2%
Trifluoromethyl (ether)	C/O	F	F	F	82	3.1%
Ketal	C	C	O	O	42	1.6%
Tertiary thio	C	C	C	S	31	1.2%
Tertiary fluoride	C	C	C	F	26	1.0%

R groups show the type of atom attached to the quaternary centre and tabulated above. The number of drug molecules containing each type of quaternary centre are given as well as in percentage of the total number (2648) of FDA approved drugs in the study.

Within this sub-section, interesting examples of FDA-approved drugs displaying the aza-quaternary centre include: Amrubicin, an anthracycline for small cell lung cancer treatment;[111] Alfentanil, a potent short-acting opioid anaesthetic and analgesic; Deflazacort, a glucocorticoid anti-inflammatory and immunosuppressant agent;[112] Ketamine, a commonly used anaesthetic;[113] and Methyldopa, for treatment of high blood pressure (Figure 2.2).[114] In addition, this feature is not only present in chemically synthesised small molecules, but is also found in natural products, such as Erythratidine (Figure 2.2).[115]
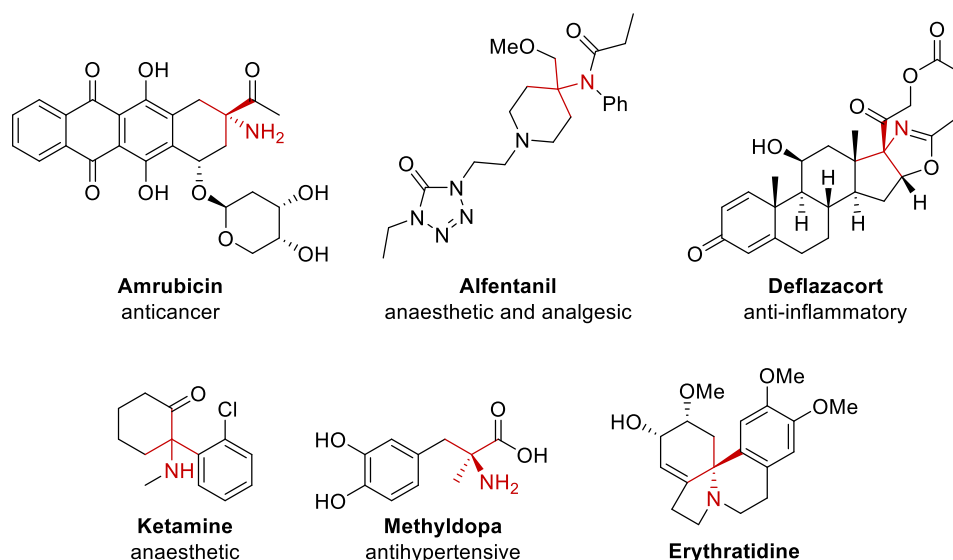
Figure 2.2 Examples of marketed drugs and an alkaloid with aza-quaternary centres highlighted in red.

Often, biologically active molecules that contain the aza-quaternary-centre are those directly relating to or derived from α,α-disubstituted amino acids. Compared to mono-substituted proteogenic amino acids the additional substituent in the α-position sterically constrains the motif, resulting in distinct chemical properties. This has been exploited in drug discovery for many therapeutic applications including enzyme inhibitors, ion-channel blockers, neuronal receptor agonists and antagonists, and antibiotic natural products.[116] The interest in these moieties within the synthetic community has grown, with examples of incorporation of this valuable centre into small molecule libraries being reported. [103,117–123] As a result, a number of different strategies have been developed for the construction of such challenging quaternary amino acids and derivatives,[116] including nucleophilic additions to ketimines (Strecker synthesis),[124] enolate Claisen rearrengements[125] and enolate alkylations.[126] However, there is still an unmet need for approaches to construct small molecule libraries featuring this motif, and new methodologies are urgently required.[102]

Recently, a successful reagent-based DOS approach was applied by Mateu *et al*. to generate 40 structurally different fragments from a common quaternary amino ester building block in no more than five synthetic steps (Scheme 2.1).[102] Importantly, the library demonstrated high levels of 3D character and a broad coverage of chemical space as well as ideal physicochemical properties. The resultant library was screened against three protein targets from different protein families—penicillin binding protein 3 (PBP3), cleavage factor 25 kD (CFI$_{25}$) and Activin A—by X-ray crystallography identifying hit fragments for each target.[127] To validate all hits, a minimum of four synthetic exit vectors were utilised to quickly access 10 to 14 analogues of each, also providing initial SAR information.
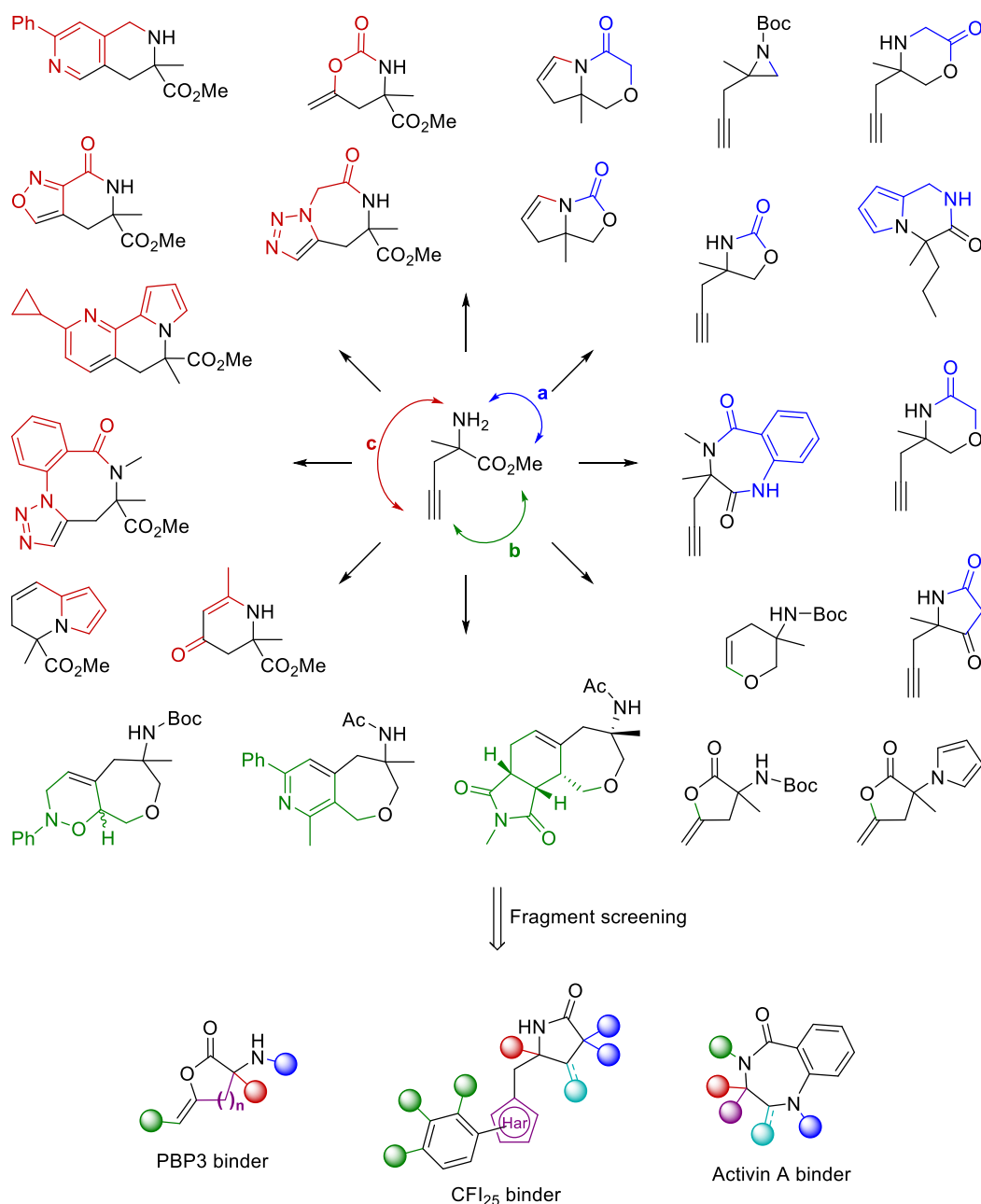
Scheme 2.1 Diversity-oriented synthesis of aza-quaternary carbon containing small molecules by Mateu *et al*.[102] Following chemoselective functionalisations around the versatile amino ester building block, various rings were formed by pairing the a) amino and ester (blue), b) ester and alkyne (green) and c) amino and alkyne (red) groups. X-ray crystallography was used to screen the library against penicillin binding protein 3 (PBP3), cleavage factor 25 kD (CFI$_{25}$) and Activin A identifying hits against all targets.[127] Each hit compound displayed a number of exit vectors that were used to quickly build libraries of 10 to 14 analogues to validate each hit.

2.1.2 Spirocycles

In addition to generating 3D features, it is often desirable to rigidify the structures of drug-like molecules *via* the introduction of cyclic systems. This can reduce both the conformational entropy penalty of target binding and the number of possible conformations (distinct 3D shapes) the molecule can adopt leading to higher potency and selectivity.[128] Examples of ring systems

employed by nature and chemical synthesis range in size from small rings (3 to 6 atoms) to macrocycles (12 or more atoms), and in complexity from simple individual cycles to more complex spiro, fused and bridged systems or combinations of these. In particular, spirocyclic motifs possess interesting structural features, displaying an inherently 3D character due to the presence of a quaternary spiro atom. The conformations adopted by each ring enables the projection of different structural elements present on the spirocycle in different directions (Figure 2.3).[129] Many different classes of spirocycle are known, however, those consisting of two small rings have been shown to offer the advantage of being either conformational locked or possessing only a small number of well-defined conformations that might contribute towards potency and selectivity.[130]
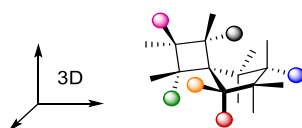
Figure 2.3 Illustration of a spirocycle projecting appendages in all directions of the 3D space.

Spirocyclic motifs are present in many bioactive molecules of both natural and synthetic origins. The spiroketal antibiotic polyketide Oligomycin[131] and Griseofulvin with antifungal activity[132] highlight the relevance of this motif within nature. The first spirocyclic drug to reach the market was the competitive mineralocorticoid (aldosterone) receptor antagonist Spironolactone, over 50 years ago.[133] With the development of new methodologies for the introduction of spirocyclic scaffolds,[134–138] the number of FDA approved drugs containing this motif has increased significantly over the last few decades. Amongst these include the selective and extremely potent κ-opioid agonist Enadoline[139] and Buspirone used to treat generalized anxiety disorder (Figure 2.4).[140]
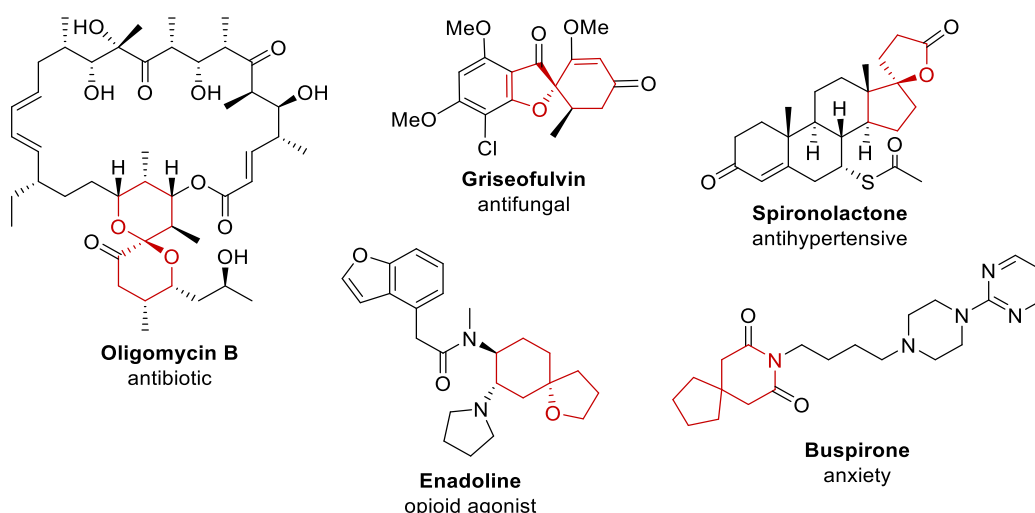
Figure 2.4 Examples of biologically active compounds containing spirocyclic motifs highlighted in red.

More specifically, an interesting and under-represented class of spirocycles are those which incorporate the aza-quaternary centre. Structures of this nature can be found in natural products like Histrionicotoxin found in the skin of poison frogs.[141] Additionally, four FDA approved drugs

display this motif: Irbesartan, a potent angiotensin II receptor antagonist used for the treatment of hypertension;[142] Fluspirilene, an antipsychotic used to treat schizophrenia;[143] Rolapitant used for the treatment of chemotherapy-induced nausea and vomiting (CINV);[144] and Omacetaxine mepesuccinate used for the treatment of chronic myeloid leukemia (CML) (Figure 2.5).[145] Importantly, these molecules exhibit conformational restriction as a result of the spirocyclic nature and feature the biorelevant aza-quaternary centre with four different substituents. Despite the biological importance of aza-quaternary spirocycles, there remains a need to deliver novel scaffolds that feature this particular motif—however, synthesis of compounds containing the motif remains challenging due to the inherent steric congestion of the atomic arrangement.
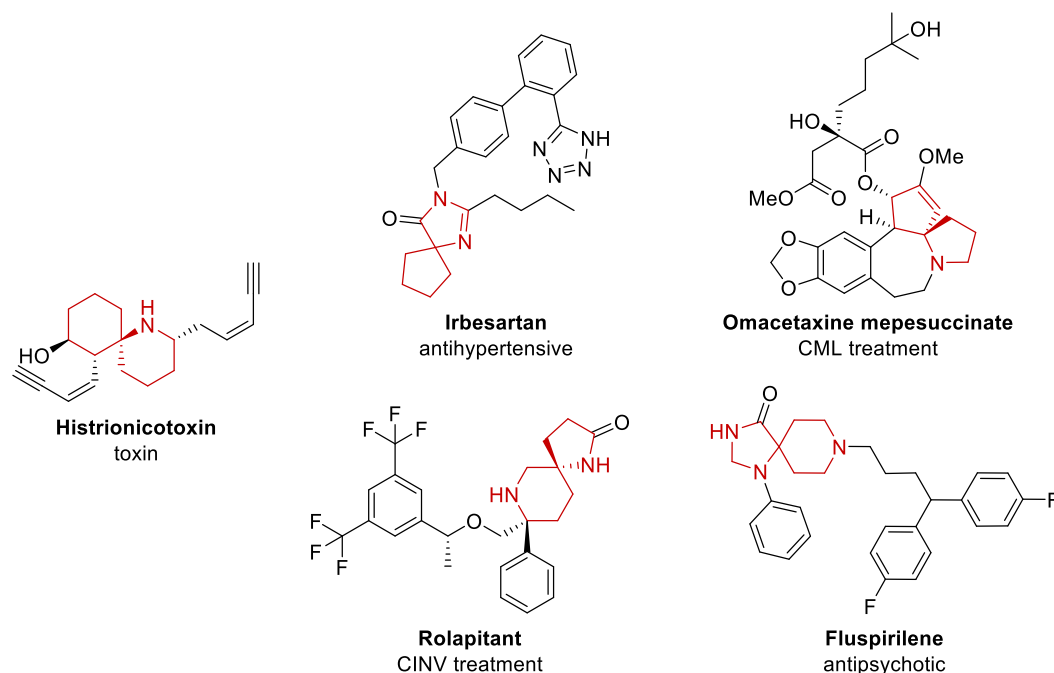
Figure 2.5 Examples of biologically active molecules featuring the aza-quaternary spirocyclic motifs highlighted in red.

The syntheses of spirocyclic α,α-disubstituted amino acid derivatives have previously been reported in the literature.[146–149] Hung *et al*. adopted a DOS strategy to generate a fragment library containing several spirocycles of this type, among others, from cyclic α,α-disubstituted amino acid derivatives.[109] This approach utilised a common build/couple/pair DOS algorithm;[150] first synthesising the multifunctional cyclic allyl amino ester building blocks, followed by intermolecular functionalization, finally pairing the terminal alkene functional groups intramolecularly to yield structurally diverse and complex cyclic molecules. Alternatively, by functionalising the carboxyl group of the quaternary cyclic amino ester building block with an additional unsaturated moiety, two spirocyclic scaffolds could be generated using a ring closing metathesis (RCM) reaction (Scheme 2.2).[109]
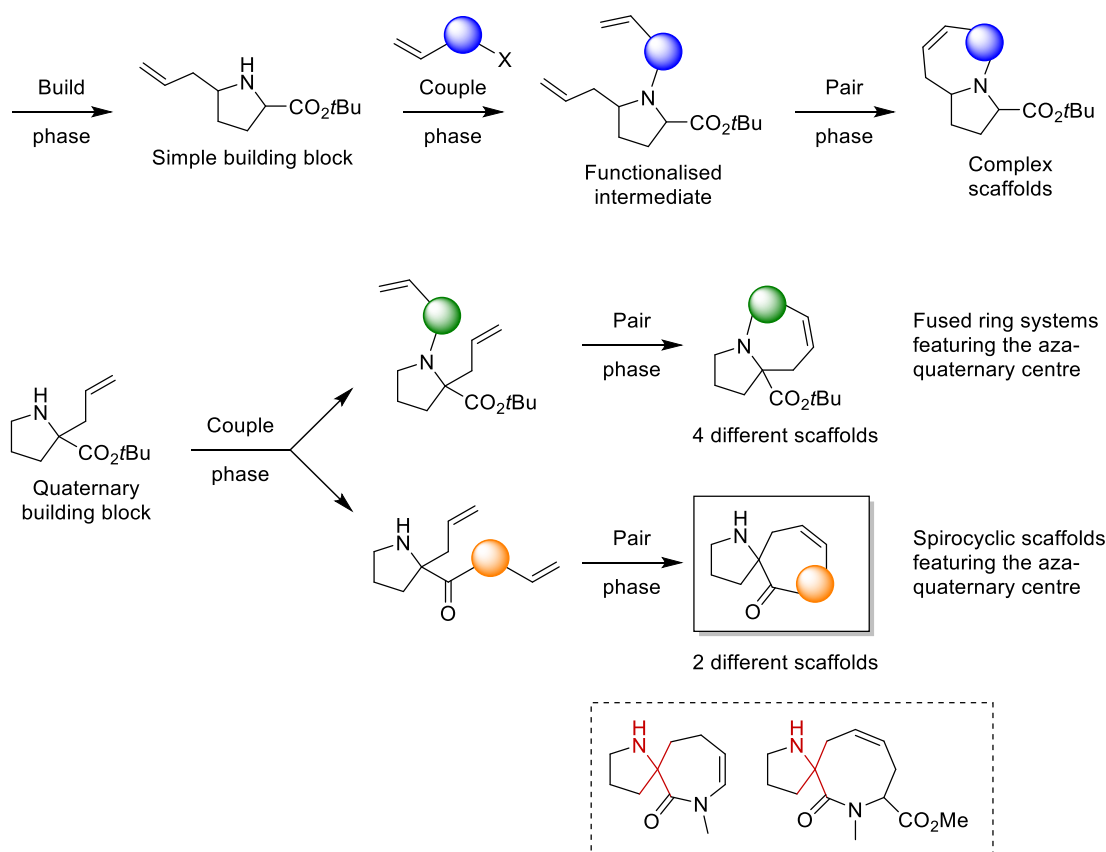
Scheme 2.2 Synthetic strategy to a diverse fragment library including aza-quaternary spirocycles (highlighted in red) by Hung *et al*.[109]

Another interesting synthetic approach was established by Stotani *et al*. for the generation of a bis-spirocyclic library.[151] This strategy utilised a key condensation between cyclic ketones and either a cyclic α-amino amide forming bis-spiro-imidazolinones (Scheme 2.3) or a cyclic α-amino alcohol forming bis-spiro-oxazolidines (not shown). Even though interesting and underrepresented bis-spirocycles were formed, the majority of the final products did not meet the size criterium for a fragment library.
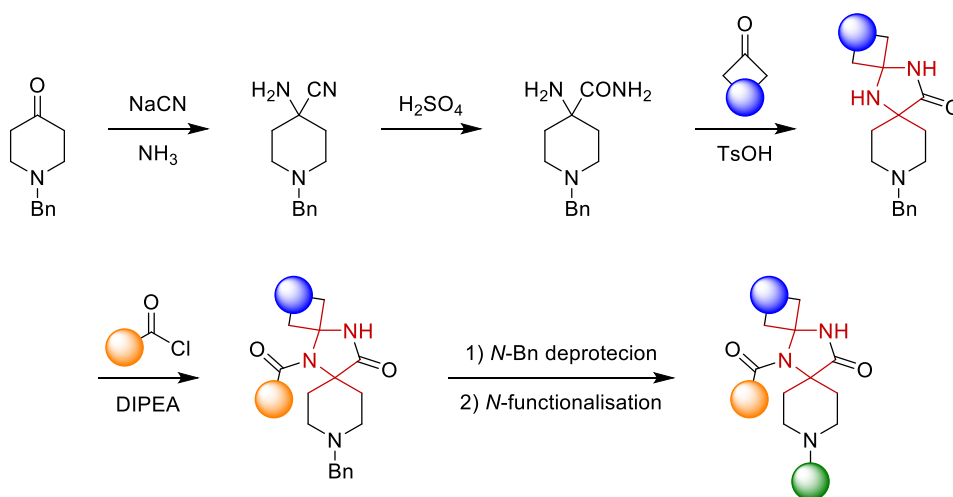
Scheme 2.3 Synthetic combinatorial strategy to a bis-spirocyclec library featuring aza-quaternary centres (highlighted in red) by Stotani *et al*.[151]

Recently, a visible-light-mediated methodology was reported for the rapid generation of aza-quaternary spirocyclic fragments from cyclic ketones and alkenyl amines by Flodén *et al*. (Scheme 2.4).[152] This methodology was used to access 20 different spirocyclic scaffolds, however, the fragments only display a very limited number of synthetic exit vectors—most of which are secondary amines. Therefore, there is still an unmet need for a strategy that would incorporate exit vectors into spirocyclic fragments to allow for rapid analogue synthesis once a hit compound is identified.
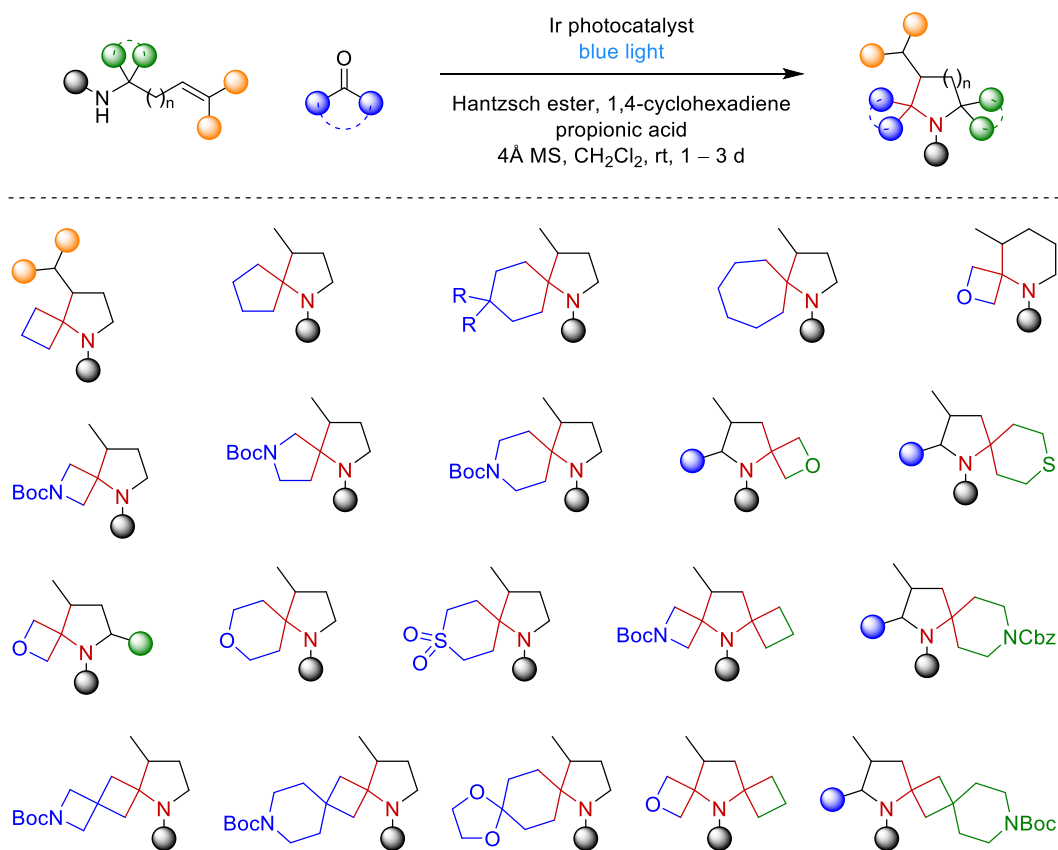
Scheme 2.4 Photocatalytic synthesis of aza-quaternary spirocycles by Flodén *et al*.[152]

34

2.2 Project Outline

In order to address the lack of sp^3-rich fragments in general[100] and especially spirocyclic scaffolds[130] within commercially available screening libraries, a project was devised to develop a novel synthetic strategy to access such highly desired fragments, whilst maintaining optimal physicochemical properties to be attractive for FBDD purposes. Moreover, modular routes to the finalised scaffolds would facilitate introduction of structural diversity and thus a broad coverage of underexplored 3D chemical space. In addition, a crucial consideration during the design of the library was to incorporate invaluable 3D exit vectors that could be synthetically utilised to access analogues once a hit is identified.

Following the previous success of the use of α,α-disubstituted amino esters in DOS strategies,[102] it was envisaged that this moiety could serve as a suitable building block to access a series of fragments containing an aza-quaternary spirocentre (Scheme 2.5). Accordingly, it was hypothesised that the incorporation of different alkyl groups bearing terminal alkenes could be achieved in either a racemic fashion (Scheme 2.5, step A1) or enantioselectively (Scheme 2.5, step A2). Using these motifs, a substrate-based approach was proposed to form carbocyclic rings of various sizes in ring-closing metathesis (RCM) reactions (Scheme 2.5, step C). Moreover, it was hoped that the presence of a key versatile α-amino ester moiety would facilitate the formation of diverse heterocycles (Scheme 2.5, step B) using a reagent-based method. Finally, it was also envisioned that strategic 3D exit vectors could be introduced into all fragments at positions on both the heterocycles (Scheme 2.5, step D) and the double bond within the carbocycles (Scheme 2.5, step E).
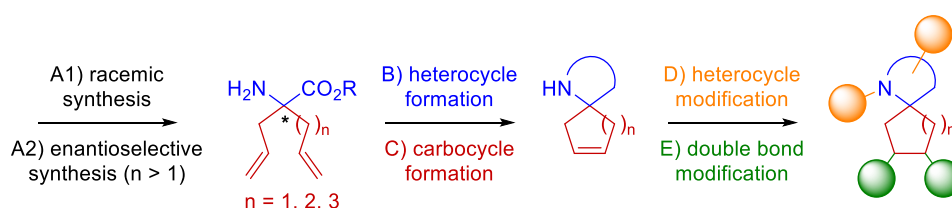
Scheme 2.5 Synthetic strategy for a diverse spirocyclic fragment library. The α,α-disubstituted amino ester building blocks can be synthesised both racemicaly (A1) and enantioselectively (A2) during the build phase. The α-amino ester motif highlighted in blue can be used to form numerous heterocycles (B), whereas the terminal alkenes highlighted in red can be paired in RCM reactions to form cycloalkene rings (C). The different 3D exit vectors can be showcased by various modifications of the heterocycles (D) and the double bond (E).

2.2.1 Scaffold scope

2.2.1.1 Ring size

Most of the spirocyclic natural products, FDA approved drugs and other candidates under clinical trials feature five- and six-membered rings.[153] Cyclopentane- and cyclohexane-containing spirocyclic motifs are found in Spironolactone[133], Buspirone[140], Irbesartan[142], Griseofulvin[132], Enadoline[139] and Histrionicotoxin[141] amongst others (Figure 2.4, Figure 2.5). Therefore, these rings can be regarded as attractive moieties for incorporation into a spirocyclic fragment library. Spirocycles containing a 7-membered ring, however, are currently underrepresented in drug discovery. As of 2014 there were no approved pharmaceuticals featuring this motif and only four compounds were reported with promising bioactivities.[153] Two of them are stearoyl-coenzyme A desaturase-1 (SCD1) inhibitor candidates for the treatment of metabolic disorders.[154] In addition, a potent and selective antagonist of the X-linked inhibitor of apoptosis protein (XIAP), a key regulator of apoptosis often overexpressed in cancer cells, has been identified by Donnell *et al.*[155] The last example is an orally available glycogen synthase kinase 3β (GSK-3β) inhibitor for the treatment of type 2 diabetes[156] (Figure 2.6). The underrepresentation of these moieties may be a result of the lack of robust and flexible methodologies to construct these ring systems and therefore development of our methodology to facilitate the formation of such scaffolds was highly desirable.

SCD1 inhibitors

XIAP inhibitor

GSK-3β inhibitor
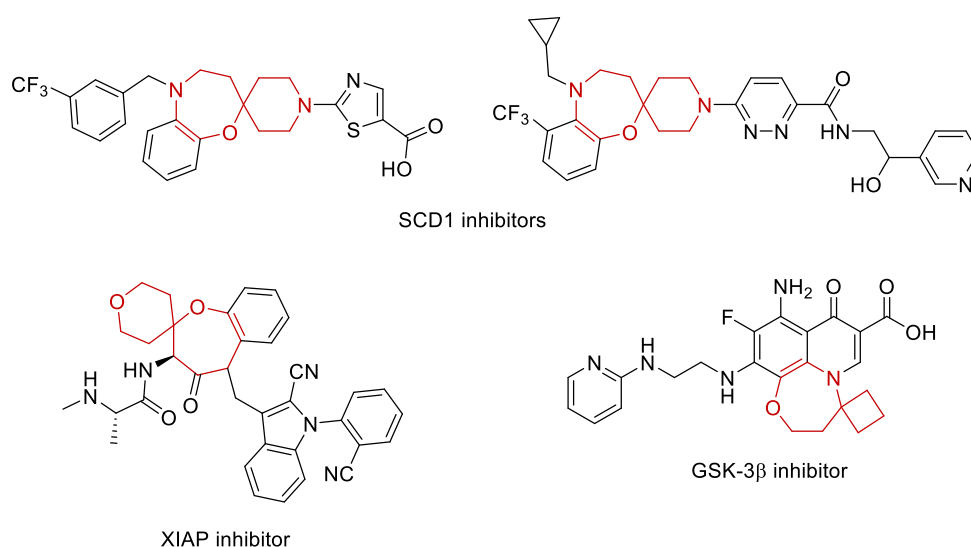
Figure 2.6 Examples of biologically active compounds containing spiro[6.n] scaffolds highlighted in red.

2.2.1.2 Heterocycles

Heterocycles are extremely important motifs in drug-like small molecules which is well demonstrated by the results of the recent analysis by Vitaku *et al.* that found that more than half (59%) of all FDA approved unique small molecule drugs contain at least one nitrogen heterocycle.[157] Therefore, it was envisaged that incorporation of some of these heterocyclic

moieties into our spirocyclic library would significantly increase its biological relevance and expand the scaffold diversity.

Pyrrolidines represent interesting biorelevant motifs and are in fact the most common five-membered non-aromatic heterocycle unit, present in around one hundred different FDA-approved drugs.[157] One example of a marketed drug featuring this motif is Eletriptan, used to treat migraines.[158] In addition, the 2-pyrrolidone and succinimide motifs are found in around ten and five small molecule drugs respectively,[157] and the tetramic acid motif (pyrrolidin-2,4-dione) can be found in the natural product, Streptolydigin, with antibacterial properties[159] (Figure 2.7).

Eletriptan
migraine treatment

Streptolydigin
antibacterial
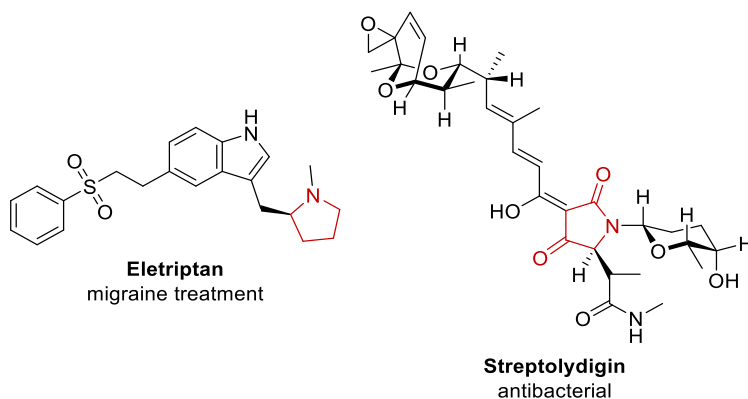
Figure 2.7 Examples of biologically active compounds containing pyrrolidine and tetramic acid moieties highlighted in red.

Morpholine represents a common motif within drug discovery being one of the five most common non-aromatic six-membered heterocycles and found in over thirty different marketed drugs.[157] The 3-morpholinone motif is present in the important anticoagulant Rivaroxaban[160] as well as in CP-271485 having interesting anti-inflammatory properties[161] (Figure 2.8).

Rivaroxaban
anticoagulant

CP-271485
anti-inflamatory
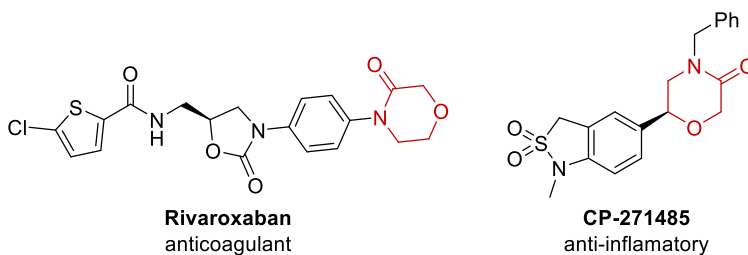
Figure 2.8 Examples of biologically active compounds containing 3-morpholinone moieties highlighted in red.

Oxazolidines are one of the five most common non-aromatic five-membered heterocycles[157] and can be found in the anti-anxiety drug Cloxazolam[162]. In contrast, the oxazolidone unit is much more commonly used and present in the oxazolidinone class of 'last resort' antibiotics, such as Linezolid[163] (Figure 2.9).
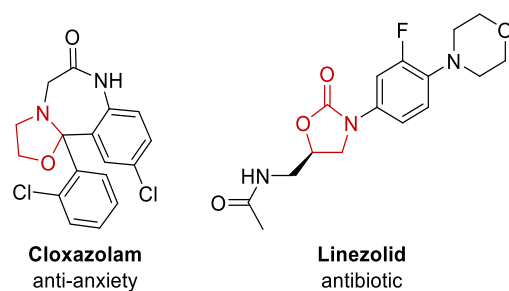
Figure 2.9 Examples of biologically active compounds containing oxazolidine and oxazolidone moieties highlighted in red.

Finally, aziridines are found in a variety of natural alkaloids[164] such as the FDA-approved antibiotic Mitomycin[165]. Some synthetic therapeutics also feature this important motif like the anticancer drug ThioTEPA[166] (Figure 2.10).
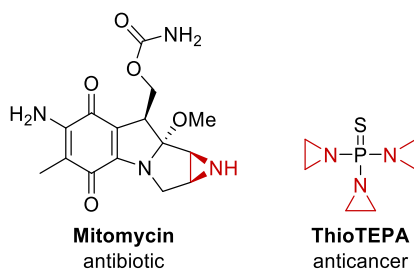
Figure 2.10 Examples of biologically active compounds containing aziridine moieties highlighted in red.

2.3   Results and Discussion

2.3.1   Building block synthesis

Initial investigations involved the development of an efficient and robust synthetic route to the quaternary amino ester building block **1**. It was envisaged that an enolate alkylation approach would serve as a reliable method for introducing different alkyl chains, facilitating the formation of a variety of building blocks from commercial regents. In the literature both isonitrile[167–170] and Schiff base[169,171,172] derivatives of glycine esters have been used as starting materials for various enolate alkylations, recently also extended to ethyl nitroacetate[173]. However, in general better yields, easier purifications and deprotections have been associated with the use of Schiff bases (Scheme 2.6).

$$H_2N \quad CO_2Et$$

alkylations

$$XN \quad CO_2Et$$

X = O$_2$
X = C
X = CR$_1$R$_2$

1
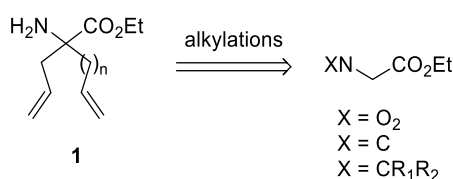
Scheme 2.6 Retrosynthetic analysis for dialkyl amino ester building blocks **1**.

As such, initial investigations began by forming the diphenylmethylene (DPM)-protected amino ester **2** *via* condensation of the inexpensive and commercially available ethyl glycinate hydrochloride and benzophenone as described by Li *et al*. (Scheme 2.7).[174] Pleasingly, this proceeded smoothly, yielding **2** in 75% yield.
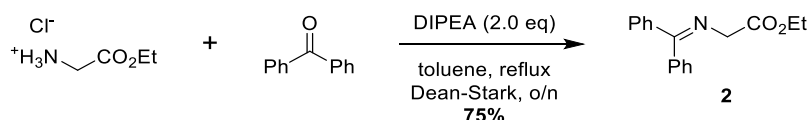
Scheme 2.7 Synthesis of the protected common starting material **2**.

With **2** in hand, investigations into the alkylation steps were then carried out. Selective step-wise double alkylations of substrate **2** have previously been described,[169] however, reproduction of these results proved to be challenging. Initially, following a literature procedure described by Andrei *et al*.,[175] alkylation of **2** with homoallyl bromide in the presence of potassium carbonate base proved unsuccessful with no conversion detectable by TLC, despite the use of higher temperatures and phase-transfer conditions (Table 2.2, Entry 1). Promisingly, when the more reactive allyl bromide was used under similar conditions, the monoallyl intermediate **3a** was isolated in moderate yield (Table 2.2, Entry 2).

Table 2.2 Attempted first alkylation steps.

3a: n = 1
3b: n = 2

Entry	n (eq.)[a]	Base (eq.)	Solvent	Temp.[b]	Time	Quench[c]	Yield[d]
1[e]	2 (1.5)	K_2CO_3 (5.0)	MeCN	reflux	3 days	none	0%[f]
2	1 (1.5)	K_2CO_3 (5.0)	MeCN	reflux	3 days	water	47%
3	2 (3.0)	tBuOK (1.1)	THF	-78 °C to RT	20 h	sat. NH_4Cl	27%
4	2 (2.8)	tBuOK (1.1)	THF	0 °C to RT	20 h	sat. $NaHCO_3$	67%

[a]Alkylating agent. [b]Reaction temperature. [c]Quenching conditions. [d]Isolated yields of products contaminated with benzophenone. [e]0.5 eq. TBAI was also added as phase-transfer catalyst. [f]Only unreacted starting material was observed by TLC.

Disappointingly, however, upon subjecting **3a** to the conditions described by Ilies *et al.*[126] using homoallyl bromide, no desired di-alkylated product was observed, only starting material (Table 2.3, Entry 1). This was suspected to be a result of the lower reactivity of the homoallyl bromide in comparison to that of the allyl bromide. Instead, it was then decided to investigate the reverse procedure, first installing the homoallyl and then allyl moieties. Pleasingly, this route led to the isolation of the di-alkylated species **4b** (Table 2.2, Entry 3 and Table 2.3, Entry 2), but with poor to moderate yields for both steps.

Table 2.3 Attempted second alkylation steps.

3a: n = 1
3b: n = 2

4b: n = 2, m = 1, X = DPM
1b: n = 2, m = 1, X = H$_2$
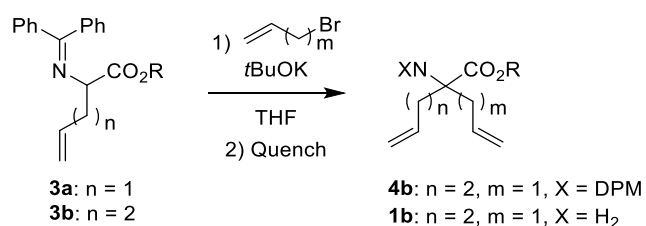
Entry	n	m (eq.)[a]	*t*BuOK eq.	Temp.[b]	Time	Quench[c]	Yield[d]
1	1	2 (3.0)	1.1	-78 °C to RT	2 days	none	–[e]
2	2	1 (3.0)	1.1 + 0.5	-78 °C to RT	2 days	sat. NH$_4$Cl	40%[f]
3	2	1 (3.0)	1.5	0 °C to RT	20 h	3M aq HCl	80%[g]

[a]Alkylating agent. [b]Reaction temperature. [c]Quenching conditions. [d]Isolated yield of **4b**. [e]Only unreacted starting material was observed by TLC. [f]Contaminated with benzophenone. [g]Pure free amine **1b** isolated after acidic deprotection.

Further optimisations at a larger scale were then conducted. One interesting observation throughout the alkylation reactions was the increased presence of benzophenone in the reaction mixture and purified product, suspected to be a result of *in situ* hydrolysis of the imine, affecting both the overall conversion and yield of each step. To minimise this, the mono-alkylation reaction was instead quenched with a basic medium, which significantly improved the yield to 67% (Table 2.2, Entry 4). Notably, the increase in temperature and use of solid base also resulted in more practical reaction conditions. Next, it was decided to directly isolate the free amine **1b** from the reaction mixture using *in situ* hydrolysis of **4b** once full conversion from **3b** was observed. Neutralisation by solid sodium carbonate led to the isolation of α,α-disubstituted quaternary amino ester **1b** with both good yield (80%) and purity, with no further purification. Finally, an even more practical 'one-pot' route was established, where the first alkylation step was permitted to reach completion, followed by the addition of more solid base and allyl bromide. The resultant solution of **4b** was deprotected under acidic aqueous conditions and the previously described aqueous work-up used to yield sufficiently pure **1b**, eliminating the need for any chromatographic purifications from the building block synthesis (Scheme 2.8). The scalability of this synthetic procedure was demonstrated through the multi-gram (4.7 g) preparation of **1b**. However, it is important to mention, that it was essential to monitor the conversion to **3b** as premature addition of allyl bromide lead to the formation of impurity **4a**, which resulted in inseparable contaminations of cyclopentene containing spirocycles in later steps.
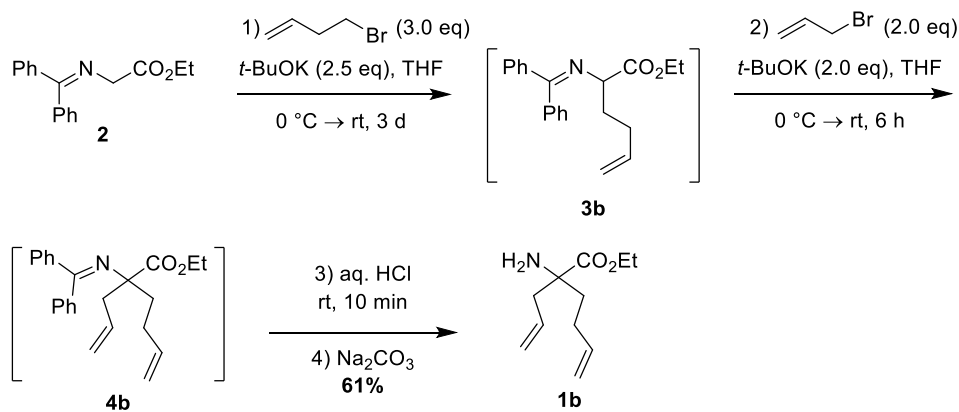
Scheme 2.8 'One-pot' synthesis of the building block **1b**.

Due to the substrate-based nature of the carbocycle formation, it was necessary to synthesise building blocks with variable chain lengths in their alkenyl groups. To achieve this, it was envisaged that synthesis of dialkenyl amino ester **1b** would enable the generation of a cyclohexene ring, whereas the formation of a cyclopentene ring required a building block with one less methylene group. To attain this, the diallyl amino ester **1a** was readily synthesised from the same DPM-protected amino ester **2**. This was achieved using a single one-pot reaction with excess allyl bromide and base. Again, the protected intermediate **4a** was not isolated, but instead directly hydrolysed then neutralised to produce the desired diallyl amino ester **1a** in moderate yield and sufficient purity without chromatography (Scheme 2.9).
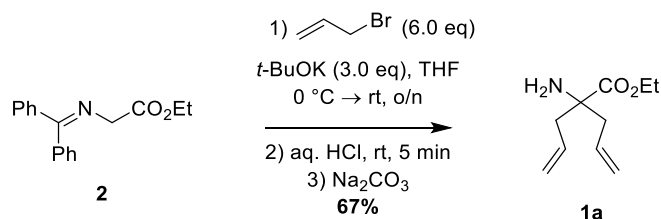
Scheme 2.9 'One-pot' synthesis of the diallyl building block **1a**.

In order to expand the methodology to cycloheptene rings, the synthesis of an additional quaternary amino ester bearing one more methylene group was essential. It was envisaged that following the successful reaction path to **1b** described previously, it would be possible to synthesise the desired **1c** from DPM-protected starting material **2** in the same manner. In the first alkylation step, 5-bromo-1-pentene was used to form intermediate **3c**, which led to the analogous synthesis of the quaternary amino ester **1c** with a low overall yield over the three steps resulting from the partial hydrolysis of the DPM group throughout the work-up and chromatographic purification steps (Scheme 2.10). However, sufficient quantities of **1c** were obtained and thus the optimisation of this route was not pursued.
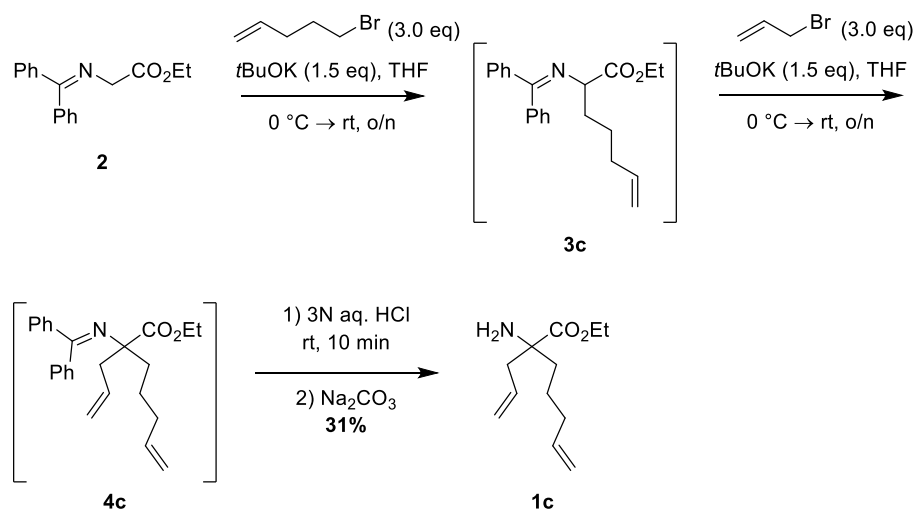
Scheme 2.10 Step-wise synthesis of the allyl pentenyl building block **1c**. Intermediates **3c** and **4c** were contaminated with benzophenone.

2.3.2 Different carbocycle and tetramic acid containing spirocycles

With the amino ester building blocks **1a-c** in hand, a suitable heterocycle could be constructed pairing the amino and ester functionalities, followed by the RCM reaction, to exemplify the different spiro[4.n] ring systems. Tetramic acid was chosen as the default heterocycle due to the ease of synthesis and potential for future functionalisations. Introduction of this heterocycle began with the direct acylation of amino esters **1a-c** by ethyl malonyl chloride in the presence of triethylamine in excellent yields (Scheme 2.11). Subjecting **5a-c** to basic conditions mediated deprotonation of the acidic α-carbonyl protons, followed by intramolecular nucleophilic substitutions upon the second ester to generate the tricarbonyl intermediates **6a-c**. Due to the instability of intermediates **6a-c**, they were immediately subjected to hydrolysis and decarboxylation by heating in wet solvent under either neutral or acidic conditions to produce the quaternary pyrrolidinones **7a-c**. Finally, the cycloalkene rings were formed in RCM reactions to complete the spirocycle formations. The more stable and versatile second-generation Grubbs catalyst was used in higher dilutions to overcome competing oligomerisations and/or polymerisation of the starting materials. Under these conditions, the tetramic acid containing spirocycles **8a-c** were successfully formed in good to excellent yields (Scheme 2.11). However, unlike the other RCM reactions, compound **7a** was not soluble in CH_2Cl_2 at reflux; instead toluene and heating the reaction to 70 °C was required. The scalability of this route was demonstrated with a gram-scale synthesis of **8b** with an overall yield of 75% from **1b**; and with a low 1% catalyst loading in the final RCM reaction with an excellent 99% yield.
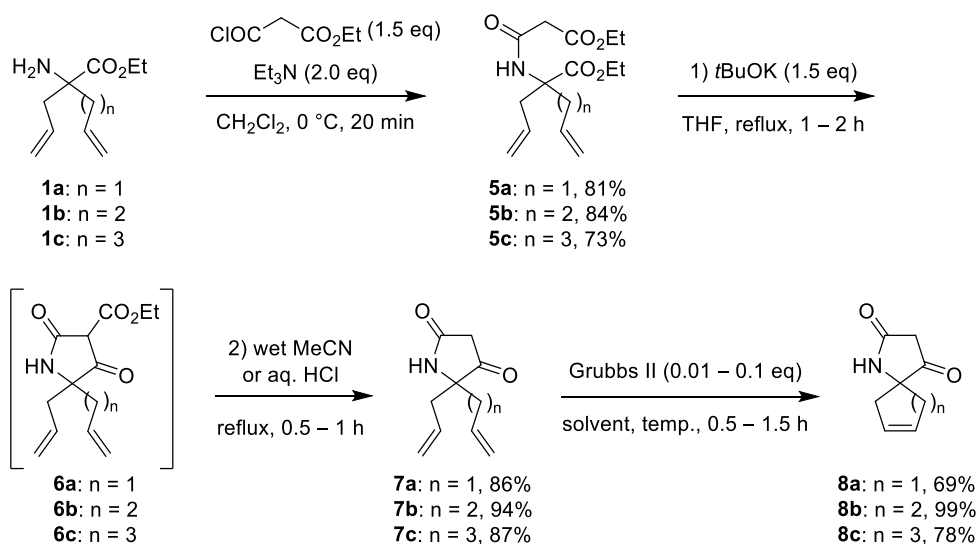
Scheme 2.11 Synthesis of spirocycles containing tetramic acid and 5-, 6- and 7-membered carbocycles.

2.3.3 Enantiopure spirocycle synthesis

A vast number of biologically active compounds and approved drugs contain at least one chiral centre and most of the chiral drugs have been approved as single enantiomers during the last few decades.[176] Additionally, due to the inherently chiral nature of biological targets, often different enantiomers of the same compound can have very distinctive bioactivity profiles. Examples of the use of racemic entities are known, however often only one enantiomer is responsible for the therapeutic effect and the other can also contribute to side-effects or even become the main source of them.[177] The tragic incidence of Thalidomide shows how important it is to evaluate the pharmacokinetic properties of both enantiomers of a drug candidate.[178] In this manner, demonstration of an enantiopure route to both the amine and one final spirocycle was an important objective within the project.

2.3.3.1 Enantioselective building block synthesis

The work in this subsection (2.3.3.1) was devised and carried out by Dr. Natalia Mateu as part of the Spring group's collaborative research into spirocyclic fragments.[179]

This strategy involved the use of a chiral auxiliary for the formation of chiral iminolactone moieties, which are known to give excellent stereocontrol over nucleophilic additions to the imine.[180–183] Importantly, both enantiomers of the chiral auxiliary are readily available, thus if successful this strategy would prove a powerful approach to synthesising the enantiopure building blocks. In the first step, the ketoester **9** was synthesised from commercially available starting materials using a Grignard reaction.[184] The chiral iminolactone **10** was then formed in a double condensation reaction between the ketoester **9** and inexpensive commercially available (*R*)-2-(−)-2-phenylglycinol.[183] The allyl substituent was introduced in a Barbier-type reaction to form the quaternary aminolactone **11** in a good yield. Auxiliary cleavage could then be performed *via* a transesterification induced by thionyl chloride followed by oxidative cleavage of the

resultant 1,2-aminoalcohol **12** by lead tetraacetate to produce the desired enantiopure amino ester **(R)-1d** (Scheme 2.12).
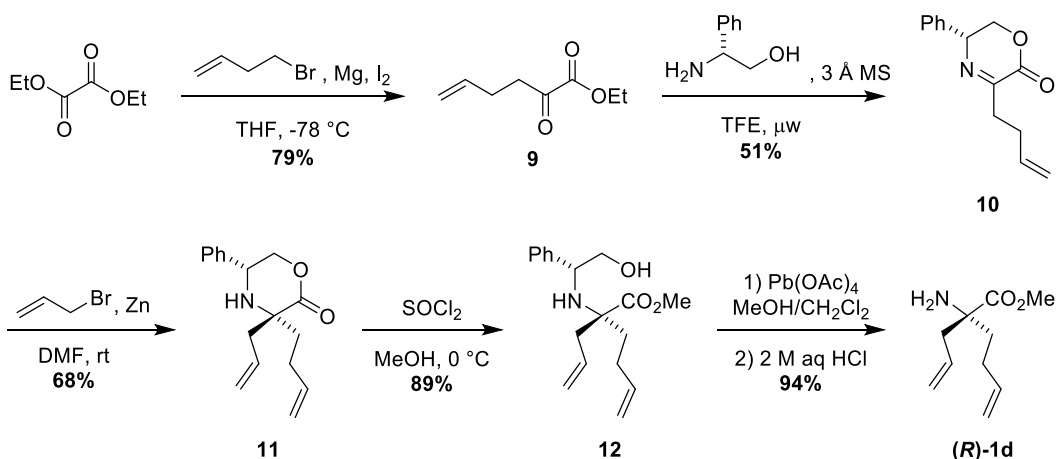
Scheme 2.12 Synthesis of the enantiopure amino ester building block **(R)-1d**. All reactions were carried out by Dr. N. Mateu. For full details see reference 179 or Appendix 6.4.

2.3.3.2 Preparation of the enantiopure tetramic acid containing spirocycle

With enantiopure amino ester **(R)-1d** in hand, the optically pure *R*-enantiomer of **8b** was synthesised as proof of principle, following the previously developed route with an overall yield of 41% over four steps (Scheme 2.13).
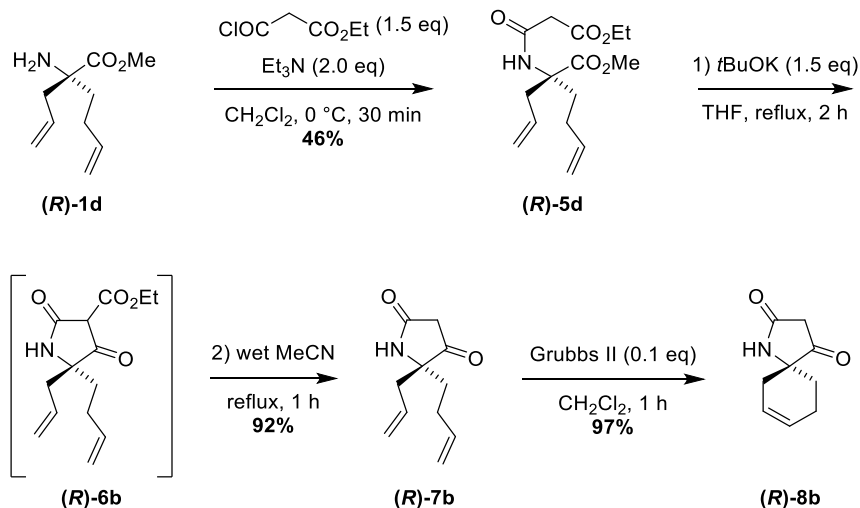
Scheme 2.13 Synthesis of the enantiopure spirocycle **(R)-8b**.

2.3.4 Different heterocycle and cyclohexene containing spirocycles

The next stage of the project involved the formation of a range of heterocycles *via* pairing of the amino and ester functional groups within **1b**.

2.3.4.1 Synthesis of the morpholinone-containing spirocycle

It was envisaged that installation of a chloroacetyl functionality to the amino group could facilitate the formation of a morpholinone-containing spirocycle (Scheme 2.16). Thus, this route towards the key *N*-chloroacetyl alcohol intermediate **13** was next investigated. Initial investigations to form **13** began with the direct acylation of the unprotected amino ester **1b** to give the *N*-chloroacetal ester **14** (Scheme 2.14). This proceeded successfully yielding **14** in good yield. However, suitable conditions for the reduction of **14** to form the *N*-chloroacetyl alcohol **13**, could not be identified. It was found that when using lithium borohydride, no product could be identified. Instead, a highly polar side product (observed by TLC) was formed but which could not be isolated. Alternatively, lithium aluminium hydride was also trialled however the ^{1}H NMR spectrum of the crude product suggested that the chloroacetamide had also been reduced.
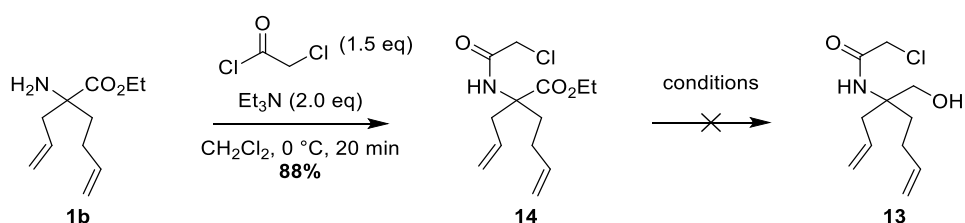
Scheme 2.14 Attempted first route to the *N*-chloroacetyl alcohol **13**.

Following the failure of this first route, it was decided to first protect the amino functionality within **1b** using the *tert*-butoxy carbonyl (Boc) group and then reduce the ester prior to the *N*-acetyl chloride introduction. The Boc protection afforded intermediate **15** in a good yield (Scheme 2.15). Unfortunately, the reduction of **15** with lithium borohydride proved sluggish, forming the desired alcohol **16** as only a minor product with low yield (Scheme 2.15), potentially due to the steric hindrance arising from the neighbouring quaternary centre. Similarly to the reduction of **14** with lithium borohydride, the formation of very polar side products was once more observed. On the basis of literature reports, this was hypothesised to be the result of hydroboration of the terminal alkenes.[185] However, sufficient quantities of **16** were isolated, allowing progression through the reaction sequence.
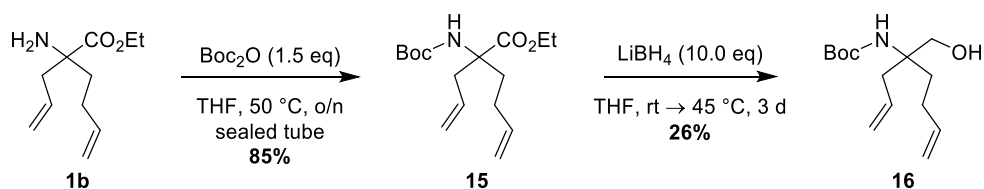
Scheme 2.15 Synthesis of the Boc-protected amino alcohol intermediate **16**.

Next, removal of the Boc protecting group using hydrochloric acid followed by selective acylation of the amino group led to the isolation of chloroacetamide **13** in a moderate yield over two steps (Scheme 2.16). Pleasingly, in the presence of potassium *tert*-butoxide, the cyclisation between the alcohol and the *N*-chloroacetyl group afforded the quaternary morpholinone **17** in an

excellent 98% yield. The cyclohexene ring was successfully formed in the final RCM reaction to afford the morpholinone-containing spirocycle **18** with a modest yield (Scheme 2.16).
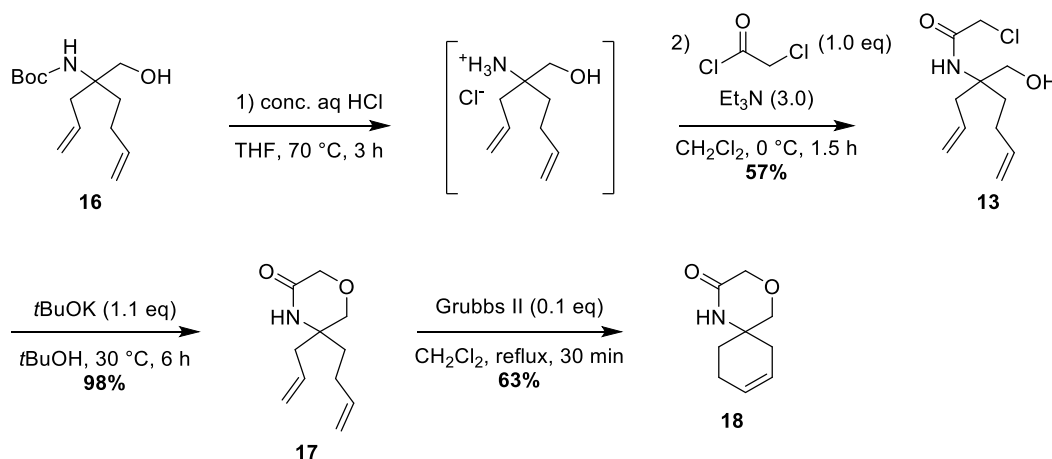
Scheme 2.16 Synthesis of the morpholinone containing spirocycle **18**.

2.3.4.2 Synthesis of the oxazolidone containing spirocycle

Previous work in the group found that treatment of quaternary *N*-Boc amino alcohols with a suitable base induced the cyclisation of the alcohol onto the carbamate, replacing the *tert*-butoxide to form an oxazolidone ring (Scheme 2.17, A).[102] Alternatively, reacting the same starting material with potassium hydroxide in the presence of a stoichiometric amount of tosyl chloride (TsCl) resulted in the construction of the aziridine moiety *via* a tosylate intermediate (Scheme 2.17, B).[102]
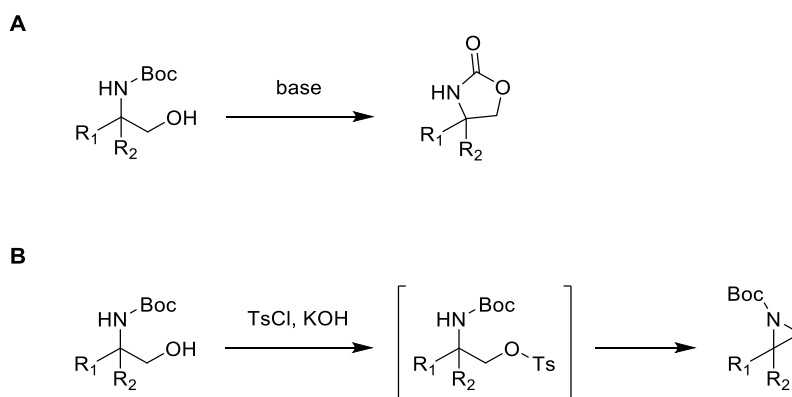
Scheme 2.17 Formation of oxazolidone (A) and aziridine (B) rings.[102]

Firstly, it was decided to synthesise the three-membered aziridine following the procedure previously developed in the group. Although a small scale test reaction showed promising results, in contrast to expectations, treatment of the *N*-Boc alcohol **16** with TsCl and KOH on a larger scale yielded oxazolidone **19** as the major and the aziridine **20** as the minor products (1.7 : 1 ratio), with an overall yield of 76% (Scheme 2.18). A plausible explanation for this observation may be due to partial hydrolysis of the TsCl prior to reacting with alcohol **16** to form the tosylate intermediate. Alternatively, the sterically hindered alkoxide is more likely to react with the much

less electrophilic Boc group intramolecularly than with TsCl intermolecularly. Finally, the carbamate could have acted as an oxygen nucleophile rather than a nitrogen nucleophile displacing the tosylate as a leaving group to form the oxazolidone instead of the expected aziridine. Despite this set back, the two quaternary heterocycles were isolated in sufficient quantities to proceed to the final RCM reactions to form the spirocycles.
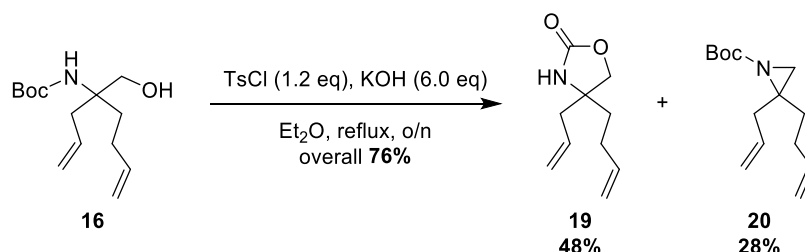
Scheme 2.18 Synthesis of precursors **19** and **20**.

Oxazolidone **19** was then subjected to the same RCM reaction conditions described previously to form the cyclohexene-oxazolidone spirocycle **21** in a good yield (Scheme 2.19). However, the cyclisation of the aziridine intermediate **20** proved to be non-trivial.
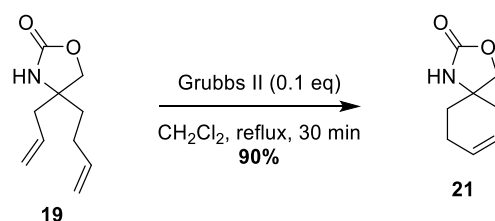
Scheme 2.19 Synthesis of the oxazolidone containing spirocycle **21**.

Initially, cyclisation of the quaternary aziridine **20** was also attempted under the same conditions (Scheme 2.20). Upon subjecting **20** to RCM reaction conditions, total conversion of the starting material was observed after 30 minutes. Formation of the desired product was supported by both ^{1}H NMR and HRMS spectra of the crude product. However, disappointingly a pure sample of **22** could not be isolated after chromatographic purification, potentially due to decomposition on the silica gel, a hypothesis that was supported by two-dimensional TLC studies. Whilst intermediate **20** was a stable compound under standard conditions used for purification and characterisation, it was hypothesised that decreased stability of the spirocycle could have been a result of the increased ring strain in **22**.
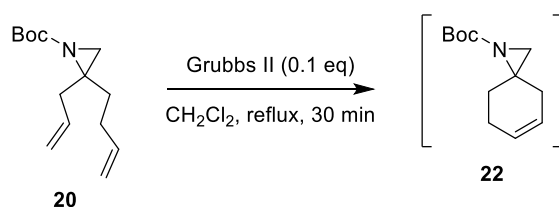
Scheme 2.20 Attempted synthesis of the aziridine containing spirocycle **22**.

Further efforts to synthesise **22** were then attempted using an alternative route *via* first forming the cyclohexene ring and then the aziridine spirocycle **22**. Indeed, the acyclic *N*-Boc ester **15** could be cyclised to give cyclohexene *N*-Boc ester **23** in an excellent yield, with a low 0.2 mol% catalyst loading (Scheme 2.21). Pleasingly, the reduction of the less hindered ester group in the cyclic intermediate **23** proceeded much more smoothly than previously experienced with **15**, producing the cyclohexene-containing alcohol **24** in very good yield (Scheme 2.21). This provides further evidence for the previous hypothesis of hydroboration of the terminal alkenes within **15**, as this was not observed with **23**.
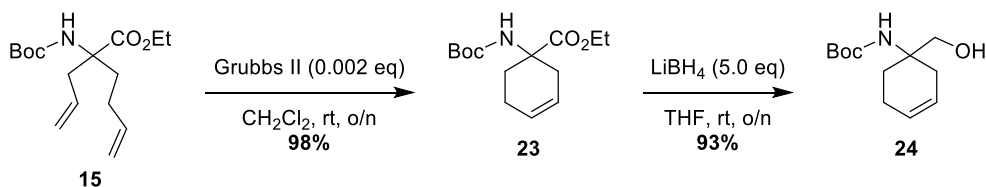
Scheme 2.21 Synthesis of the Boc-protected cyclohexene amino alcohol **24**.

With the cyclohexene *N*-Boc alcohol **24** acquired, the formation of the aziridine ring was then trialled again using TsCl and KOH, but in greater excess to suppress oxazolidone formation (Scheme 2.22). Unfortunately, the aziridine spirocycle **22** was not formed, instead only the tosylate intermediate **25** was isolated in low yield. It could be hypothesised that the lack of intramolecular cyclisation may be a result of the increased ring strain, leading to a very high energy transition state as well as reducing the Thorpe-Ingold effect. At this point it was concluded that the aziridine spirocycle **22** was most likely not stable enough for any utility in a fragment screening library, and therefore further attempts on its synthesis and purification were not carried out.
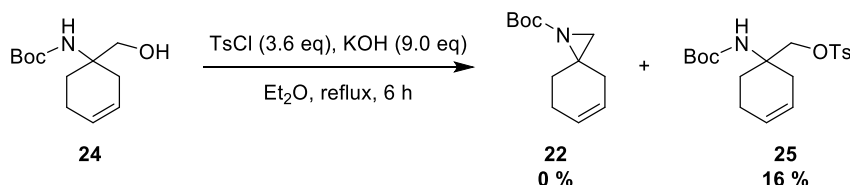
Scheme 2.22 Attempted aziridine formation using the cyclic Boc alcohol intermediate **24**.

Finally, in order to efficiently synthesise larger quantities of the oxazolidone spirocycle **21**, an improved route was established utilising the *N*-Boc cyclohexene intermediates. As expected, the base mediated cyclisation of **24** afforded spirocycle **21** in excellent yield on 2-gram scale.
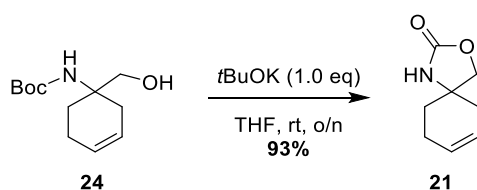
Scheme 2.23 Improved synthesis of **21**.

2.3.5 Double bond modification

With a diverse set of core spirocyclic scaffolds in hand, the utility of the alkene as a functional handle was then investigated. This double bond was strategically incorporated into the carbocycles of each member of the library, to facilitate fragment growth in all directions. Modifications to the cycloalkene rings also increases the number of chiral centres resulting in greater stereochemical diversity within the library and polar functional groups capable of forming hydrogen-bonding interactions with protein targets. Notably, although the objective of this project was to exemplify single step reactions for conciseness, many of these transformations can create further multifunctional handles. To explore this, the oxazolidone heterocycle was selected for its ease of synthesis and chemical stability. The cyclohexene ring was chosen in order to observe the diastereo- and regioselectivity of certain reactions.

Initially, unprotected spirocycle **21** was explored as an appropriate starting material for the envisaged chemical modifications (

Scheme **2.24**). Firstly, hydrogenation of the double bond was investigated and 10% (w/w) palladium on charcoal was found to be a suitable catalyst for the reaction under hydrogen atmosphere, yielding cyclohexane spirocycle **26** in good yield. More pleasingly, epoxidation of unprotected **21** with *m*CPBA also proved moderately successful, however two chromatographic purifications were required before both diastereomers of epoxide **27** could be isolated in a modest combined yield of 52% (dr 8:1). Single crystal X-ray crystallography analysis revealed that in the major diastereomer, **27a**, the epoxide is *cis* to the carbamate nitrogen (Figure 2.11), suggesting that the oxidation by *m*CPBA is directed by hydrogen bonding, further supported by a study on related scaffolds by O'Brien *et al*.[186] However, further synthetic efforts made on the non-protected starting material **21** were unsuccessful. The products of these reactions were often highly polar rendering the work-up procedures very difficult, and the lack of UV-active or easily oxidisable groups made their identification and purification even more challenging. Furthermore, it was hypothesised that the unprotected oxazolidone in fact interfered with some of the attempted reactions.
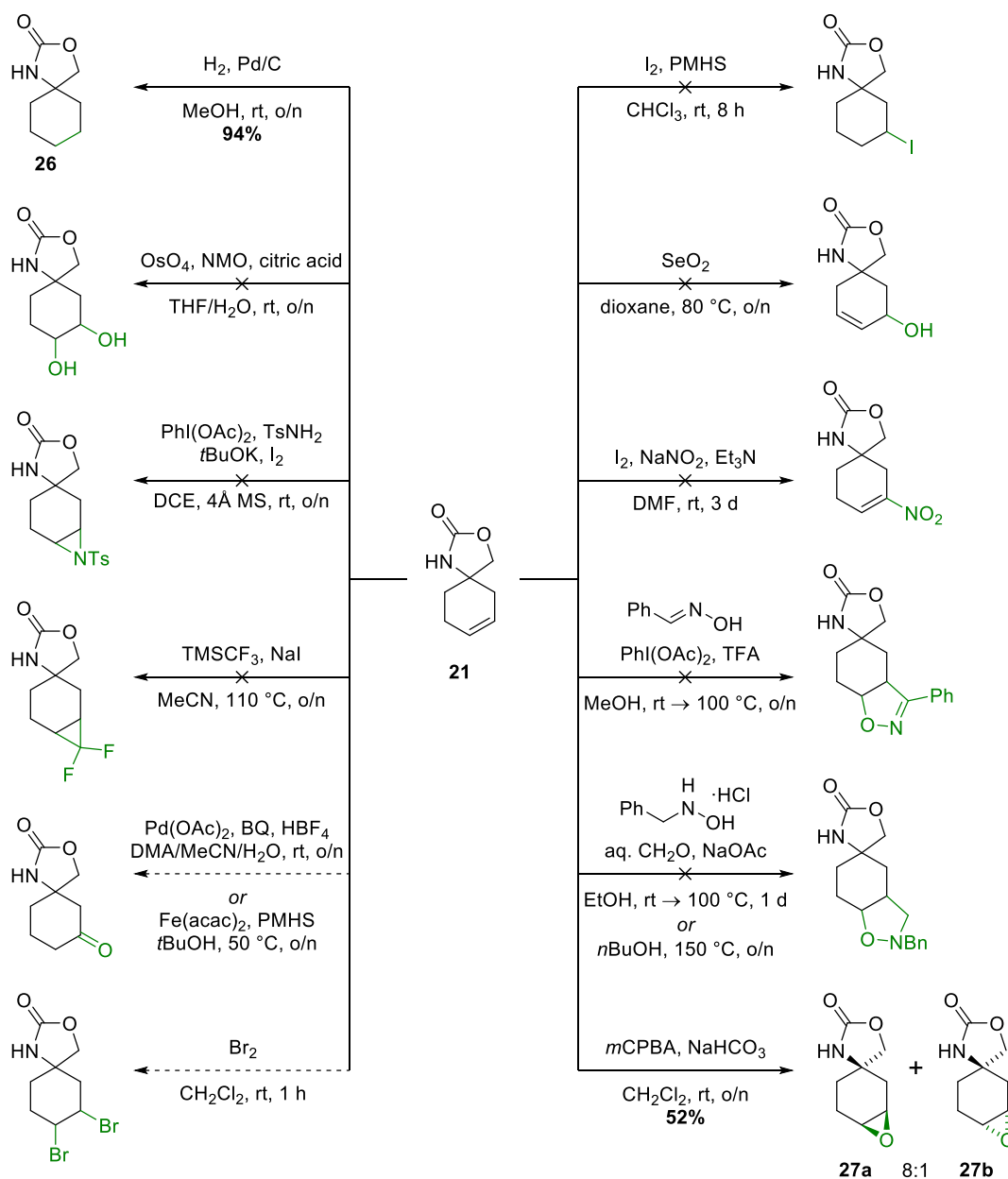
Scheme 2.24 Attempted double bond modifications on non-protected **21**. All products are racemic. Crossed arrow: no indication of desired reaction taking place based on TLC and/or crude ^{1}H NMR analysis; dashed arrow: formation of desired products is evidenced by both TLC and crude ^{1}H NMR analysis, but attempted chromatographic purifications failed to yield products of sufficient purity.
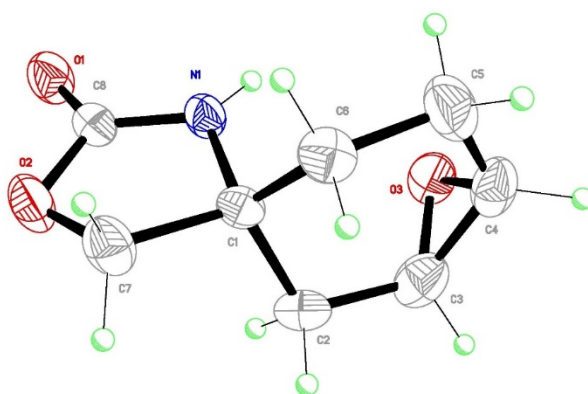
Figure 2.11 X-ray crystal structure of the major epoxide **27a**.

Following the limited progress made on unprotected **21**, the use of a suitable protecting group on the amine was considered. The *para*-methoxy benzyl (PMB) group was identified as the most promising candidate due to the ease of incorporation, versatile deprotection conditions and UV-activity. It was also envisaged that this functionality would render the compounds less polar and more soluble in organic solvents. The unprotected spirocycle **21** was thus treated with sodium hydride and PMBCl to give the PMB-protected oxazolidone **28** in excellent yield (Scheme 2.25).
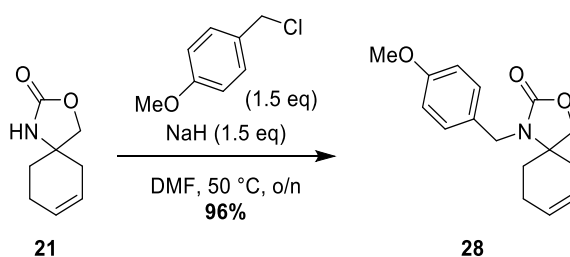
Scheme 2.25 PMB-protection of the oxazolidone ring.

With the PMB-protected cyclohexene **28** in hand, the feasibility of the chemical functionalisation of the key alkene exit vector was once more investigated. Dihydroxylation of **28** in the presence of 1% OsO$_4$ catalyst proceeded smoothly and yielded two diastereomers of **29** in quantitative yield (dr 2.5:1) (Scheme 2.26, step a). Dibromination by phenyltrimethylammonium tribromide (PTAB) proved to be similarly successful yielding **30** in 96% (dr 18:1) (Scheme 2.26, step b); whereas using the more reactive bromine itself gave partial oxidative deprotection of the PMB group and subsequently resulted in lower yields. Formation of the *N*-tosyl aziridine derivative **31** was achieved in 70% (dr 1.4:1) following the conditions described by Jeong *et al.* using Chloramine-T as a practical nitrogen source and PTAB as catalyst (Scheme 2.26, step c).[187] Synthesis of the *gem*-difluorocyclopropane **32** was carried out under the conditions described by Wang *et al.*,[188] however, the reaction did not go to completion and the majority of **28** was recovered unchanged, whilst products were isolated in a low 26% yield (62% based on recovered starting material; dr 20:1) (Scheme 2.26, step d). Following the limited success of the difluorocarbene insertion, a more reactive variant of the traditional Simmons-Smith reaction described by Yang *et al.*[189] and Bojase *et al.*[190] was attempted, however the desired product could not be isolated. Treatment of **28** with chloroform and base did not yield the analogous

dichlorocyclopropane either. Finally, the palladium-free Wacker-type oxidation described by Liu *et al.*[191] was employed to form regioisomers of ketone **33** in a moderate 43% yield (rr 1.9:1) as well as regio- and diastereoisomers of alcohol side products **34** (combined 33%) and unreacted **28** (8%) (Scheme 2.26, step e).
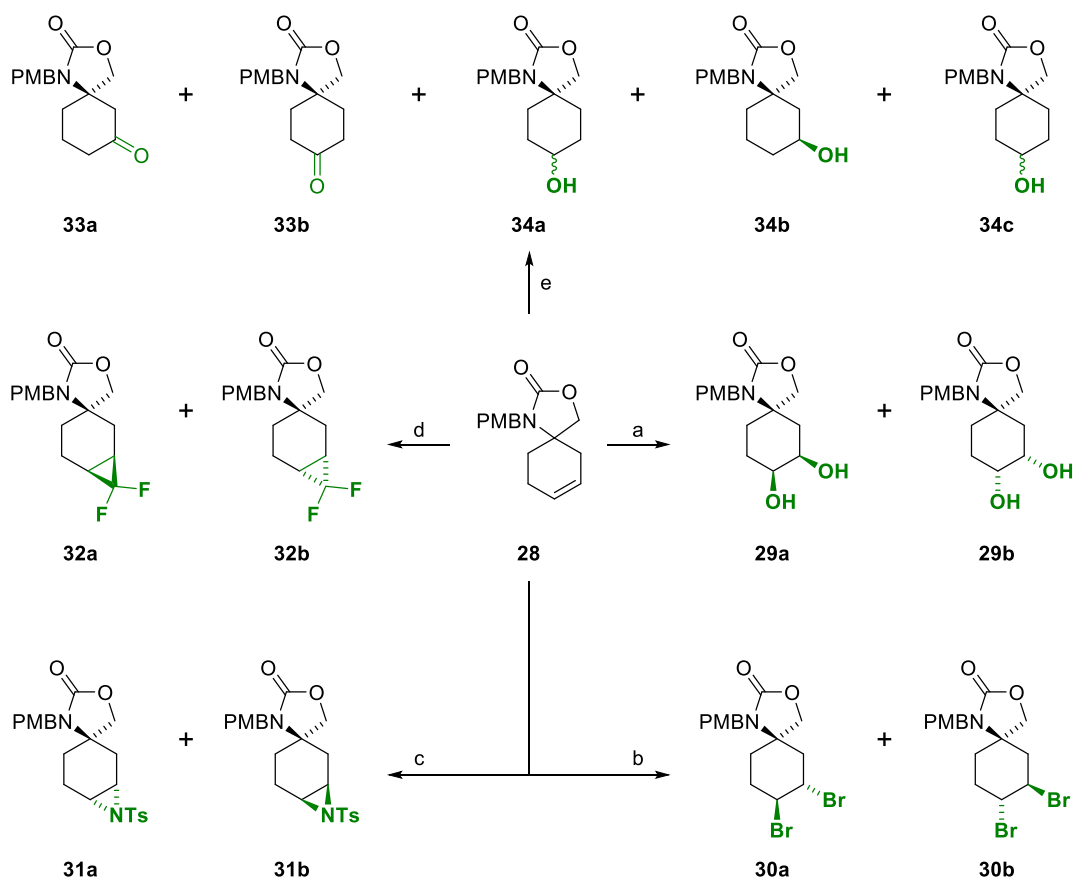
Scheme 2.26 Double bond modifications on the PMB-protected spirocycle **28**. All products are racemic. Reactions condition (combined yields (based on recovered **28** where applicable) with ratios **a/b** are given): a) OsO$_4$, NMO, citric acid, H$_2$O/THF, 99%, 2.5:1; b) PhNMe$_3$Br$_3$, CH$_2$Cl$_2$, 96%, 18:1; c) TsNClNa·3H$_2$O, PhNMe$_3$Br$_3$, 4 Å MS, 70%, 1.4:1; d) TMSCF$_3$, NaI, THF, 62%, 20:1; e) Fe(acac)$_2$, *t*BuOH, air, 47%, 1.9:1 (**33**) + 36% mixture of isomers (**34**).

Assignments of the diastereomers are based on single crystal X-ray crystallography analysis (Figure 2.12), as NMR studies were not conclusive for these compounds. The regioisomers of ketone **33** (and alcohol **34**) were easily distinguished by NMR analysis. Diastero- and regioisomers of these compounds are numbered in general as Xa, Xb, where Xa corresponds to the major isomer, diastereomeric and regioisomeric ratios are given as Xa:Xb. The relative stereochemical outcomes of these reactions can be rationalised by assuming that the initial electrophilic attack on the double bond is preferred on the sterically less hindered face *cis* to the PMB-protected nitrogen occupying the more favourable *pseudo*-equatorial position in the half-chair conformation of **28**. This directly results in the observed major diastereomers of dihydroxy derivative **29** and difluorocyclopropane **32**. The major diastereomer of the dibromo species **30** results from the *trans*-diaxial opening of the bromonium intermediate by bromide. Similarly, the bromonium intermediate is attacked by a sulfonamide anion followed by a second intramolecular

nucleophilic substitution to form the major product **31a** with the aziridine *trans* to the carbamate nitrogen.

29b

30a

31b

32a

34b

35
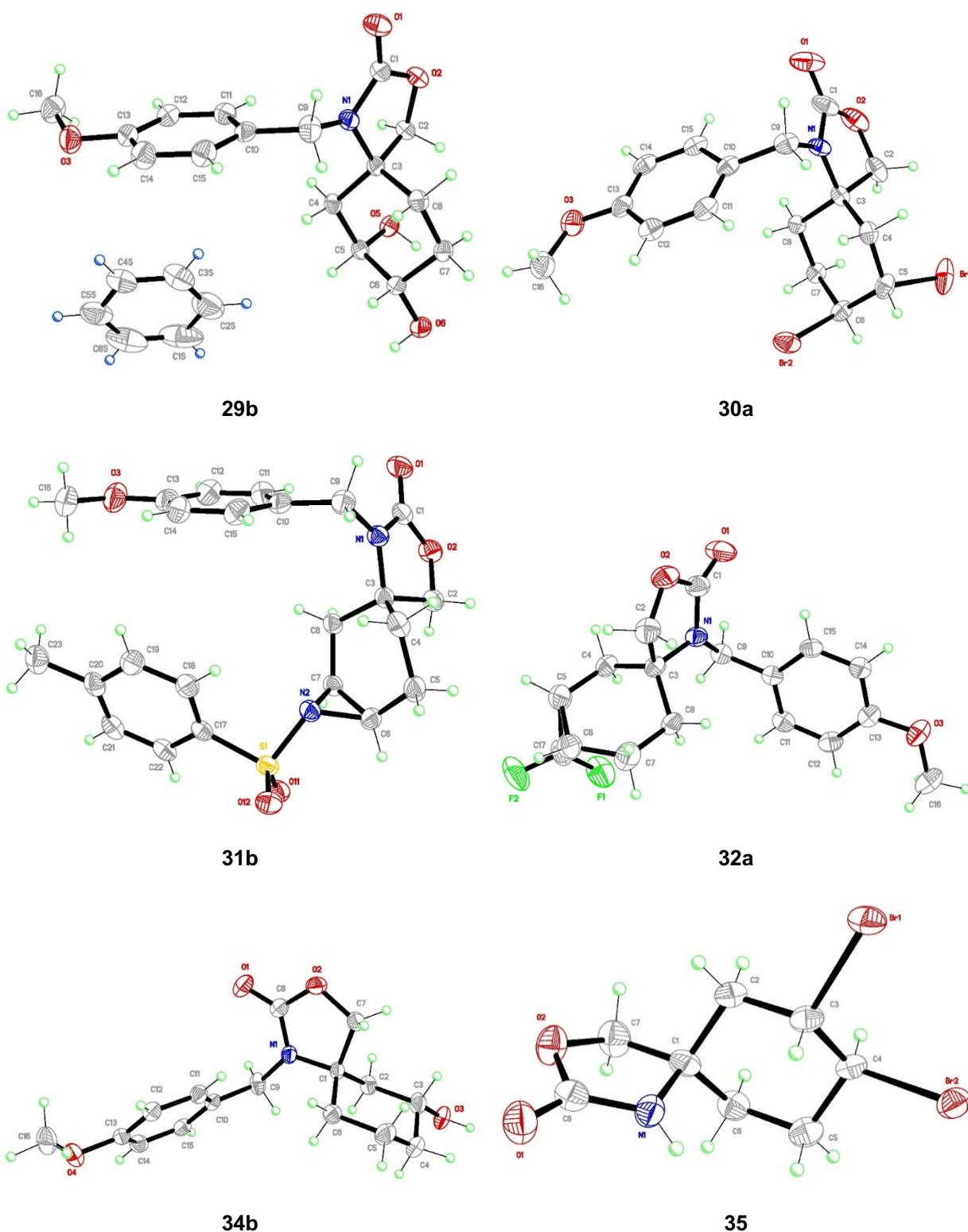
Figure 2.12 X-ray crystal structures of double-bond-modified spirocycles.

It was also crucial to demonstrate that the PMB protecting group could easily be removed to enable use of the final fragments. The three most common reaction types found in literature for the removal of PMB groups from amides and carbamates are: 1) hydrogenation using

heterogeneous palladium catalysts and hydrogen gas, 2) oxidative cleavage by 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ) or ceric ammonium nitrate (CAN) and 3) acidic cleavage by strong acids such as trifluoroacetic acid or hydrochloric acid at elevated temperatures. Disappointingly, the attempted hydrogenolysis of the aziridine **31b** under hydrogen atmosphere in the presence of 10% (w/w) palladium on charcoal returned unchanged starting material. Refluxing dibromide **30a** in TFA overnight, however, did yield the desired deprotected dibromo-spirocycle **35** in a good 88% yield, although these conditions are quite harsh and presumably incompatible with many desirable functional groups. Unfortunately, the deprotection could not be achieved using the milder oxidant DDQ. Pleasingly, however, oxidation of **30a** by CAN revealed **35** in an improved near-quantitative yield of 96% under reasonably mild conditions (Scheme 2.27).
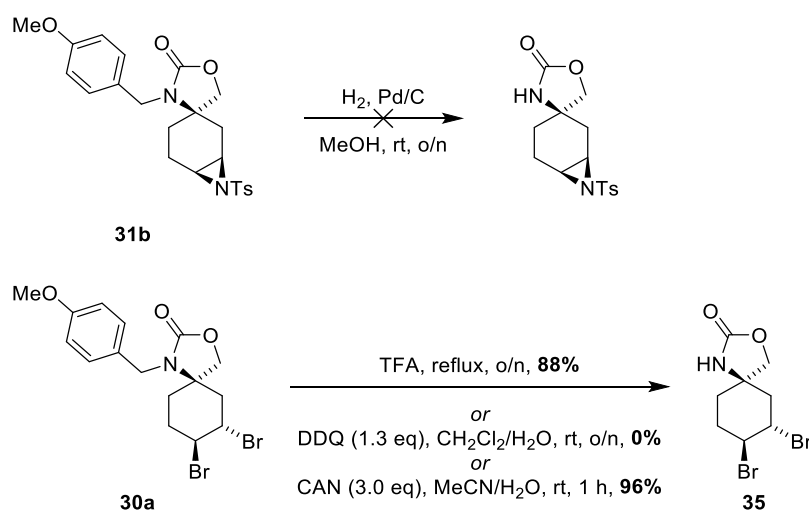
Scheme 2.27 Removal of the PMB protecting group.

2.3.6   Further heterocycles and their modifications

The work in this section (2.3.6) was carried out by Dr. Andrew J. P. North as part of the Spring group's collaborative research into spirocyclic fragments.[179]

In order to increase the diversity of the spirocyclic library and incorporate attractive pharmacophore-like heterocycles, several further pairing reactions between the amino and alcohol groups were also investigated. To achieve this, the more practical route utilising cyclohexene intermediate **24** as the branchpoint was employed. This allowed for the large-scale synthesis of **24**, which was then deprotected by treatment with hydrochloric acid prior to the pairing reactions, forming the amine intermediate **36**. Spirocycles featuring both the morpholin-3-one **18** and morpholin-2-one **37** were synthesised using α-halo acid derivatives. Additionally, substituted oxazolines **38** and **39** were formed by treating **36** with cyanogen bromide and ethyl benzimidate, respectively.

In addition to the expansion of the core fragment scaffold set, the exploration of alternative 3D exit vectors was also expanded to address the heterocyclic component of the compounds.

Substitution at the *N*-3 position of the oxazoline ring has been shown by alkylation (incorporating the PMB protecting group in **28**), whereas at the *C*-2 position by changing the electrophilic reagent used for the heterocycle formation (as shown by **38** and **39**). Oxidation of the morpholin-2-one ring of **37** by lead tetraacetate yielded iminolactone **40**, which, similarly to the chiral iminolactone **10** (described in section 2.3.3.1), should readily undergo nucleophilic additions at the imine centre. The tetramic acid (in **8**) proved to be the most versatile heterocycle, allowing for several chemical modifications to grow this fragment from the *C*-4 position. Chemoselective *O*-alkylation yielded dihydropyrrolone ether **41**, whereas reacting **8** with triflic anhydride, the enol triflate **42** could be isolated. The latter proved to be a suitable substrate for Suzuki-Miyaura cross-coupling with *para*-methoxyphenyl boronic acid affording the aryl functionalised dihydropyrrolone **43** in good yield. Furthermore, a simple reduction of the ketone functional group by sodium borohydride yielded the more saturated alcohol functionalised pyrrolone **44**. Treatment of **44** with trifluoroacetic anhydride followed by base mediated elimination and hydrolysis gave access to the unfunctionalised dihydropyrrolone **45**. Scheme 2.28 details these transformations.
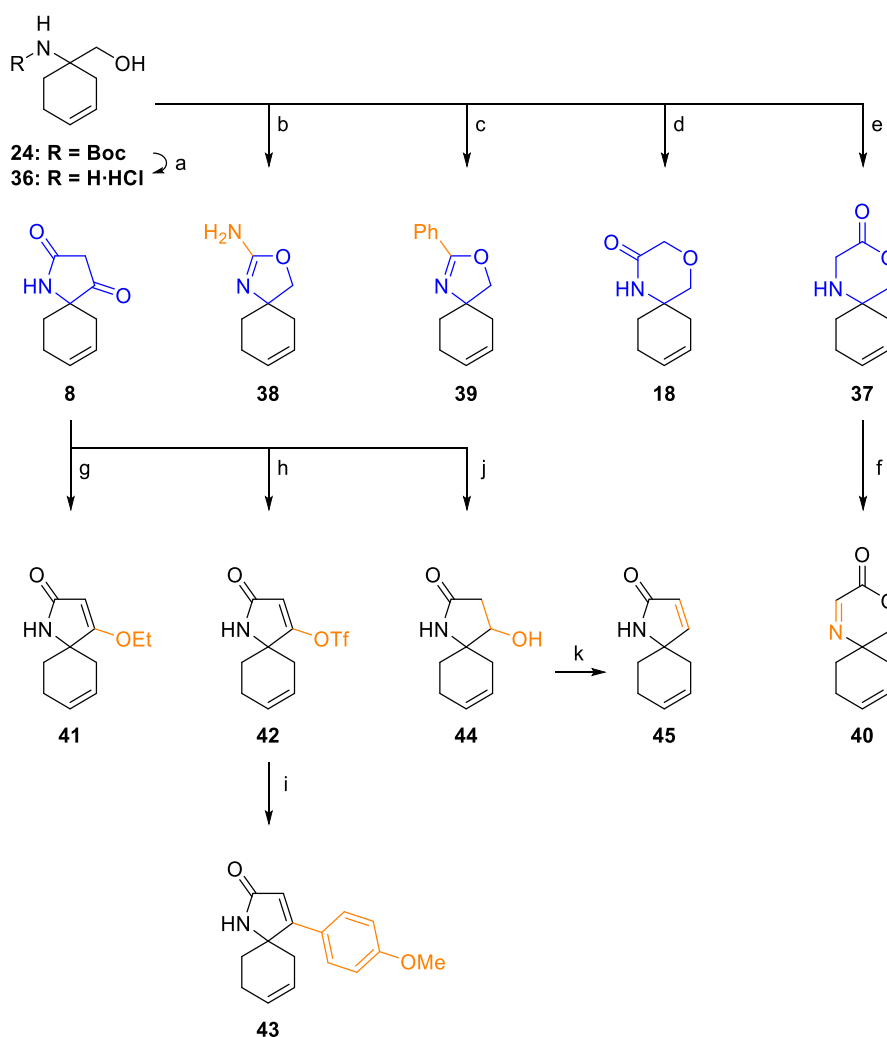
Scheme 2.28 Synthesis of further heterocycles and their chemical modifications. All reactions were carried out by Dr. A. J. P. North. Reaction conditions: a) HCl, dioxane, quant.; b) BrCN, Et$_3$N, EtOH, 58%; c) ethyl benzimidate hydrochloride, Et$_3$N, DCE, 57%; d) (i) chloroacetyl chloride, Et$_3$N, CH$_2$Cl$_2$, 57%; (ii) tBuOK, tBuOH, 99%; e) phenyl bromoacetate, iPr$_2$NEt, MeCN, 43%; f) Pb(OAc)$_4$, MeCN, 92%; g) KHMDS, EtBr, THF, 54%; h) Tf$_2$O, Et$_3$N, CH$_2$Cl$_2$, 59%; i) PMPB(OH)$_2$, Pd(PPh$_3$)$_4$, Na$_2$CO$_3$, H$_2$O/THF, 73%; j) NaBH$_4$, MeOH, 23%; k) TFAA, Et$_3$N, CH$_2$Cl$_2$; then KHCO$_3$, MeOH, 30%. For full details see reference 179 or Appendix 6.4.

2.3.7 Computational analysis

As discussed in section 1.4, obtaining optimal physicochemical properties is crucial for any fragment library and thus was an important aim of this project. Fragments considered in this section include 1) the unmodified core spirocycles described in sections 2.3.2 and 2.3.3; 2) the non-protected modified spirocycles **27a,b**, the synthetically deprotected **35** and the virtual collection obtained by removing the protecting groups from the rest of the modified compounds described in section 2.3.5; 3) molecules synthesised by Dr. A. J. P. North (section 2.3.6). The resultant library of 28 spirocycles was computationally assessed for various physicochemical properties and compared to both the commercially available state of the art Maybridge core fragment collection[192] (further details in Appendix 6.1.2) of 1000 fragments and the widely

accepted guidelines set by Astex Pharmaceuticals[95,97]. The most widely used properties are listed in Table 2.4 for both libraries; whereas their distributions are shown as histograms in Figure 6.1 (see Appendix).

Table 2.4 Calculated physicochemical properties of fragment libraries compared to the ideal range based on the Astex guidelines.

Property[a]	Spirocycles[b]	Maybridge[c]	Ideal Range[d]
MW	186 ± 41	182 ± 42	140 – 230
HBD	1.3 ± 0.7	1.0 ± 0.8	≤ 3
HBA	1.8 ± 0.6	1.8 ± 0.7	≤ 3
SlogP	0.9 ± 0.9	1.9 ± 0.8	0 – 2
RBC	0.2 ± 0.5	2.0 ± 1.5	≤ 3
TPSA	48 ± 13	39 ± 14	≤ 60

[a]MW = molecular weight (Da), HBD = number of hydrogen-bond donors, HBA = number of hydrogen-bond acceptors, SlogP = partition coefficient, RBC = rotatable bond count, TPSA = topological polar surface area ($Å^2$). [b]Protecting groups virtually removed from the library. [c]Maybridge core fragment collection of 1000 fragments. [d]Guidelines set by Astex Pharmaceuticals.[95,97] All values are presented as the mean value followed by its standard deviation. Green = within ideal range, yellow = mean value along with the standard deviation indicates that many fragments fall outside of the ideal range. Table reproduced from reference 179; for further details see appendix 6.1.

The physicochemical properties in Table 2.4 were calculated by the computer program Molecular Operating Environment[193] (see section 6.1 for further details) and the results analysed by Microsoft Excel, which revealed that the spirocyclic library adheres well to the aforementioned guidelines. The molecular weight (MW) and the number of hydrogen-bond donors (HBD) and acceptors (HBA) are all comparable across both libraries and lie well within the ideal range. The slightly higher value of the topological polar surface area (TPSA) shows that the spirocycles incorporate more polar functional groups, however, this is still on the optimal spectrum. Whilst the calculated lipophilicity (SlogP) of the Maybridge library complies well with the Ro3 (logP ≤ 3), it is just on the upper limit of the more stringent guidelines used in this comparison. On the other hand, the spirocyclic library shows superior lipophilicity, which could be inferred to impart good water solubility. Finally, a noteworthy difference in the number of rotatable bonds (RBC) between the two libraries was identified. Importantly, while both values fall within the ideal range (RBC ≤ 3), the higher rigidity of the spirocyclic library should in fact be more desirable for both fragment screening and elaboration purposes as discussed in section 2.1.2.

Next, the three-dimensionality and shape diversity of the libraries were assessed by a principal moments of inertia (PMI) analysis. First, the lowest energy conformer of each molecule was computed using MOE, followed by the corresponding moments of inertia around the

perpendicular principal axes (PMI1 ≤ PMI2 ≤ PMI3). Then the normalised PMI ratios (NPR; npr1 = PMI1 / PMI3, npr2 = PMI2 / PMI3) were calculated and plotted; the geometrical restrictions on the NPR values result in a triangular representation of all possible shapes, the PMI plot. The vertices of this triangle correspond to the points (npr1, npr2) of the extremes of rod-like (0,1), disk-like (0.5, 0.5) and sphere-like (1,1) features. Objects represented by points alongside the rod-disk axis of a PMI plot have all their mass distributed within a single plane, *i.e.* they are 2D. Furthermore, the analysis of commercially available fragments by Morley *et al.* also showed that compounds lying close to the rod-disk axis (npr1 + npr2 ≤ 1.1) are sp^2-rich and considered to be 'flat' or 2D by medicinal chemists,[100] thus the term 'flatland' is often used to describe this area within a PMI plot.

The spirocyclic library (blue diamonds) was first compared to the whole Maybridge core collection of 1000 fragments (red dots) (Figure 2.13, left). This comparison shows that the conventional fragments aggregate towards the rod-disk axis, while the spirocycles are more evenly spread out over the relatively underrepresented area outside of 'flatland'. However, this is visually not necessarily clear when a small set of 28 spirocycles are compared to the much larger set of 1000 Maybridge fragments. Therefore, to give a more discernible representation, the 147 fragments most closely resembling the spirocyclic library based on the heavy and heteroatom counts were selected (Figure 2.13, right).
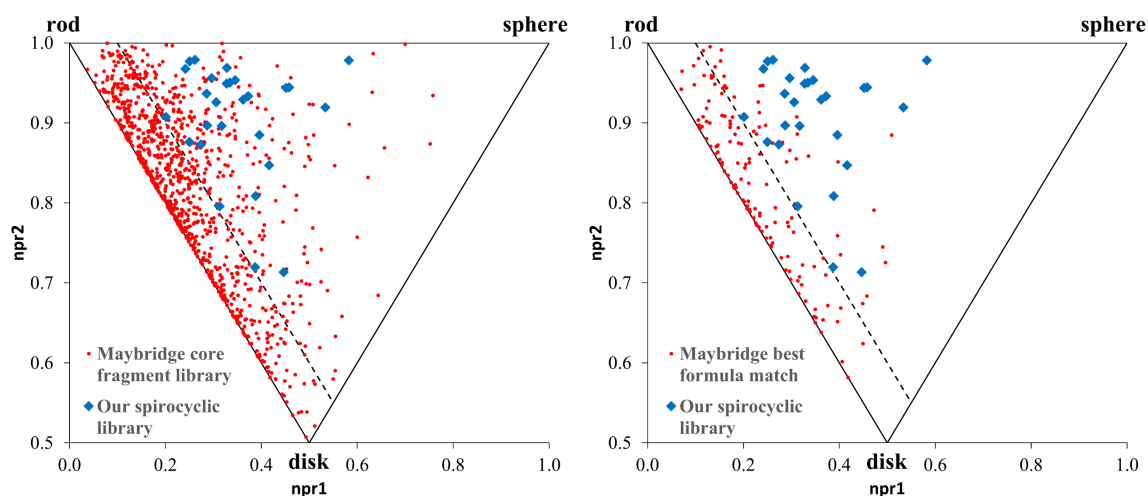
Figure 2.13 PMI plots representing shape diversity. Each corner corresponds to one of the unique rod-, disk- and sphere-like features. The dashed line represents the boundary of 'flatland' defined as npr1 + npr2 ≤ 1.1.[100] The spirocyclic library (blue) is compared to all 1000 fragments of the Maybridge collection (red, left) and the best-matched subset of 147 fragments (red, right). Figure reproduced from reference 179; for further details see appendix 6.1.

In addition to the qualitative results above, the striking difference between the two libraries in term of the 3D properties can be further highlighted by quantitative comparisons. As a direct result of the library design, a remarkable level of saturation of the spirocyclic library was achieved evidenced by the high fraction of sp^3 atoms (Fsp^3) and conversely low fraction of aromatic atoms (Far). Similarly, the more chiral centres in our library create additional stereochemical diversity.

However, the superior 3D nature of the spirocycles is most evident from its much greater sum of NPR (npr1 + npr2) and further highlighted by the fact that whereas over 70% of the whole Maybridge library and 75% of the best-matched subset falls within 'flatland' (Fflat), none of the spirocycles do. The 3D properties of the libraries are listed in Table 2.5.

Table 2.5 Physicochemical properties of fragment libraries describing their 3D properties.

Property[a]	Spirocycles[b]	Maybridge[c]	Best-match[d]
Fsp^3	0.60 ± 0.14	0.29 ± 0.22	0.28 ± 0.19
Far	0.02 ± 0.09	0.52 ± 0.25	0.49 ± 0.22
chiral	1.6 ± 1.1	0.1 ± 0.5	0.2 ± 0.5
npr1+npr2	1.25 ± 0.10	1.08 ± 0.09	1.06 ± 0.07
Fflat	0.00	0.71	0.76

[a]Fsp^3 = fraction of sp^3 atoms, Far = fraction of aromatic atoms, chiral = number of chiral centres, npr = normalised PMI ratio, Fflat = fraction of molecules lying below the 'flatland line'[100]. [b]Protecting groups virtually removed from the library. [c]Maybridge core fragment collection of 1000 fragments. [d]147 best-matched fragments from the Maybridge collection. All values, except Fflat, are presented as the mean value followed by its standard deviation. Blue = highly 3D, red = highly 2D. Table reproduced from reference 179; for further details see appendix 6.1.

2.3.8 Assessing the library's biological activity

With a diverse collection of 28 spirocycles in hand, a collaboration with Dr. Martin Welch from The Department of Biochemistry, University of Cambridge was initiated to evaluate the library for antibacterial activity. Although fragment sized molecules often lack the binding strength required for a detectable activity in phenotypic assays, many such compounds do exhibit outstanding potency—highlighted by the fact that almost 15% of all FDA-approved small molecule drugs have a MW less than 230 Da[110]. With this in mind, the spirocyclic library was screened by phenotypic growth inhibition assays—following the guidelines for minimal inhibitory concentration testing by the European Committee on Antimicrobial Susceptibility Testing[194]— against both the clinically relevant Gram-positive *Staphylococcus aureus* and Gram-negative *Pseudomonas aeruginosa*. However, no bacterial growth inhibition was observed at 256 mg/mL concentration for any of the spirocycles tested against the wild type *S. aureus* (ATCC 25923) and *P. aeruginosa* (PAO1) cell lines. In a last attempt, the compounds were also screened against the efflux pump deficient *P. aeruginosa* YM64 cell line to identify any potential hit compounds rendered inactive by quick transportation from the bacterial cytoplasm and/or periplasm (see section 3.1.2 for further details on bacterial resistance to antibiotics), but again, no activity was observed. We hypothesise that these results can majorly be attributed to the weak undetectable binding of the small fragments to antibacterial targets.

Further biological characterisation of these compounds remains underway with the fragment collection now incorporated within a fragment screening collection based at XChem at the Diamond Light Source Synchrotron, Oxford. The XChem facility represents a world-class high-throughput method for fragment-based drug discovery using crystallography. Our spirocyclic screening collection is offered to all academic users of the facility and thus additional biological activities for the described fragments is expected within the near future and will be reported in due course.

2.4 Conclusions and Future Work

Herein, this chapter has described the establishment of a novel divergent strategy for the formation of an sp^3-rich spirocycle-based fragment library. Initially, a robust and modular glycine alkylation methodology was developed to access the racemic α,α-disubstituted amino ester building blocks **1**. The amine and ester functional groups were paired in various reactions to form three (a total of six in collaboration) pharmacophore-like heterocycles, whilst the alkene moieties were used to generate three different cycloalkenes in RCM reactions, thus allowing for the rapid assembly of five (a total of eight in collaboration) spirocyclic scaffolds. Furthermore, the enantioselective building block synthesis utilising a commercially available chiral auxiliary by Dr. N. Mateu allowed the formation of the optically pure pyrrolidinone containing spirocycle **(R)-8b** as proof-of-concept to show that a single enantiomer of any member of the library could be synthesised to be used in SAR studies. Notably, all core scaffolds display a number of exit vectors to facilitate library expansion and fragment growth including various functional groups of the heterocycles as well as the strategically placed alkene moiety within the carbocycles.

The subsequent library of 28 diverse fragment-sized spirocycles was computationally assessed and showed optimal predicted physicochemical properties with much improved lipophilicity and rigidity compared to the commercially available Maybridge core fragment collection. The spirocycle library also displayed remarkable 3D properties and shape diversity as evidenced by the PMI analysis.

Finally, the spirocycle collection was screened in phenotypic assays for antibiotic activity against *S. aureus* and *P. aeruginosa* with limited success. However, future work in the Spring Group will focus on establishing new collaborations for identifying broader potential biological applications for these interesting fragments as part of the Spring Group Compound Collection.

3 Fragment-Based Approaches Towards Inhibiting PrpC

3.1 Introduction

3.1.1 Bacterial infections and treatments

Humanity has struggled against infectious diseases since prehistoric times. Indeed, many of the most devastating pandemics such as cholera, typhus, tuberculosis and plague were in fact caused by extremely virulent[†] bacteria. As an example, the most likely causative agent of the Black Death in the mid-14th century was found to be *Yersinia pestis*, killing one third of Europe's population at the time.[195,196] Even in the 21st century, opportunistic bacteria, like *Streptococcus pneumoniae*, *Pseudomonas aeruginosa* and *Staphylococcus aureus* amongst many others, are responsible for numerous severe illnesses. These include the often fatal pneumonia and sepsis for those with compromised immune systems including infants, elders, and patients undergoing immunosuppression. In addition, numerous viral infections—like human immunodeficiency virus (HIV), common cold and influenza— increase the patients' liability towards these bacteria, posing a serious healthcare challenge.[197–201] The 'Spanish flu' was one of the deadliest pandemics in history killing an estimated 50 – 100 million people worldwide between 1918 – 1919, and although it was caused by an influenza A virus of the H1N1 subtype, most of the deaths resulted from bacterial superinfections leading to pneumonia.[202]

Historically, ancient remedies, like the use of mouldy bread for the treatment of open wounds, have been shown to possess antimicrobial properties and a number of antiseptics and disinfectants, such as alcohols, aldehydes, phenols and halogen-releasing agents have been used for centuries. Strikingly, however, mankind had no effective systemic treatment for bacterial infections until the 20th century.[2,203,204] The isolation of *Mycobacterium tuberculosis* in 1882 was followed by many advancements in microbiology and medicine making the development of effective antibacterial therapies possible for the first time in human history.[205]

Arguably, conventional antibiotics (discussed further in section 3.1.2) represent the most important therapeutics, however, antimicrobial resistance (AMR) becomes more and more dominant and threatens to negate our ability to treat these infections. As a result, in addition to a drive in the development of novel small molecule antibiotics, several alternative approaches have begun to be investigated, with many now undergoing clinical trials.[206–218]

3.1.2 Antibiotics and resistance

Antibiotics were originally defined as secondary metabolites produced by either bacteria or fungi with antimicrobial properties, and later extended to synthetic and semi-synthetic small

---

[†] Virulence describes the ability of a pathogen to infect and/or damage a host organism.

molecules. Nowadays, antibacterial agents are further divided into bacteriostatic—growth inhibitory—and bactericidal—killing—entities.[205] Although the antibacterial mycophenolic acid has been isolated from *Penicillium glaucum* from as early as 1893,[219] the first clinically used antibiotic, arsphenamine (Salvarsan, synthesised in 1907), only reached the market in 1910, and was mainly used to treat syphilis until replaced by penicillin.[220] Following the landmark discovery by Alexander Fleming in 1928,[221] penicillin was not available to the public until after World War II due to its very limited supply. Therefore, sulfamidochrysoidin (Prontosil, developed in 1932 and marketed from 1935) became the first truly effective broad-spectrum antibiotic used in the clinic.[222] The introduction of penicillin to the public in 1945 marked the beginning of the Golden Age of antibiotic discovery, with many new classes discovered within two decades, saving the lives of millions around the world (Figure 3.1).[203,204]
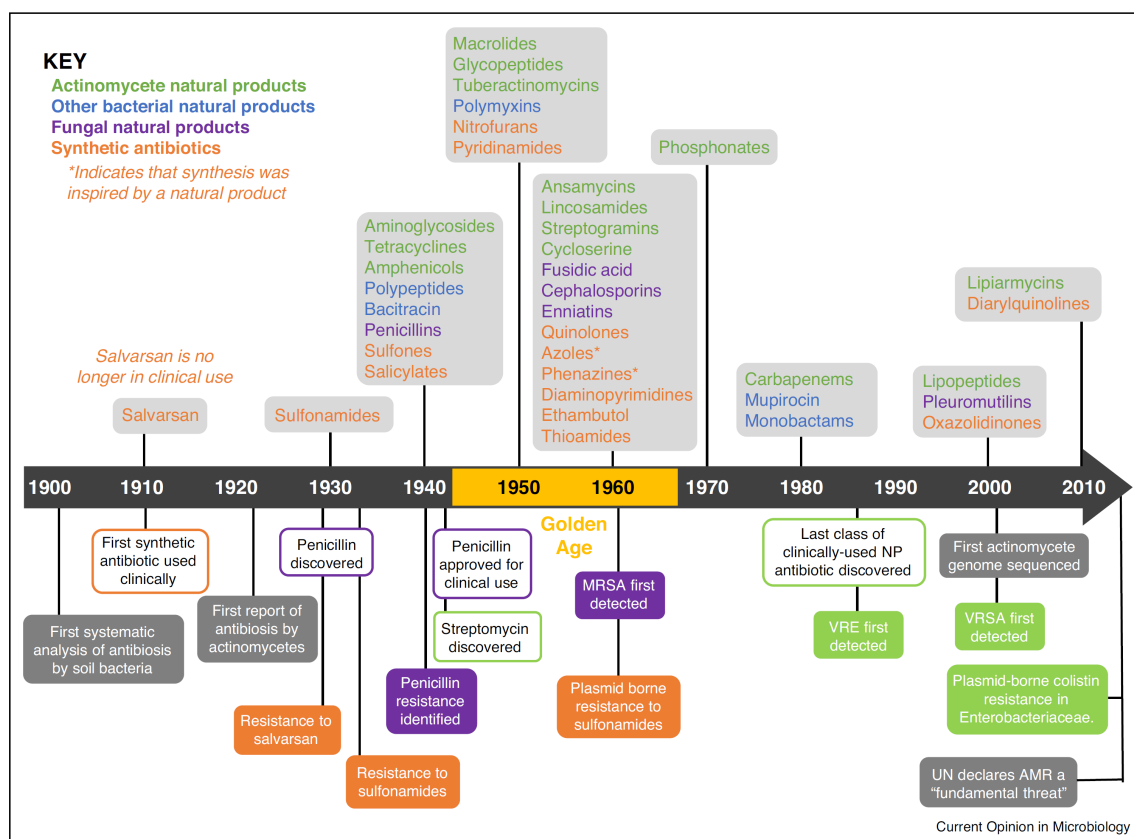
Timeline showing the decade new classes of antibiotic reached the clinic. The antibiotics are coloured per their source: green = actinomycetes, blue = other bacteria, purple = fungi and orange = synthetic. At the bottom of the timeline are key dates relating to antibiotic discovery and antimicrobial resistance, including the first reports of drug resistant strains methicillin-resistant *S. aureus* (MRSA), vancomycin-resistant enterococci (VRE), vancomycin-resistant *S. aureus* (VRSA) and plasmid-borne colistin resistance in Enterobacteriaceae.

Figure 3.1 Timeline of antibiotics. Figure reproduced from Hutchings *et al*.[203]

However, due to the emergence of AMR, the efficacy of most antibiotics has been compromised and many rendered essentially ineffective. Worryingly, AMR is currently attributed to the deaths of 700 thousand patients every year globally, and this number has been estimated to increase to 10 million by 2050.[223] Thus, the development of new antibiotics is absolutely essential to combat this crisis.[208] In addition, it is also paramount to develop our understanding of the processes by which bacteria obtain resistance in order to develop new drugs and generate protocols to minimise casualties or the chance of epidemics breaking out.[224]

64

It is now known that bacteria have developed a range of resistance mechanisms to reduce drug efficacy. Firstly, bacteria can spatially exclude drugs by either preventing their entry to the cytoplasm and/or periplasm—by changing the chemical constitution of the cell walls and membranes, and thickness of the membrane envelopes, resulting in decreased drug permeability—whist efflux pumps serve to actively transport drugs outside the cell.[225,226] Small molecules can also be chemically modified by specific enzymes to render them inactive, including substitutions and degradations or hydrolysis.[227] On the other hand, the target itself can also be modified to prevent drug binding, including residue substitutions and binding by protective factors.[228] Alternatively, the expression level of the target can be modified to minimise the effect of the drug.[229] Finally, the toxicity of the drug can be alleviated by either employing an alternative enzyme for the same function that is not susceptible to the drug or altering the metabolic pathways to remove the need for the target's function (metabolic shunt).[230] These resistance mechanisms can be acquired as a result of several genetic changes including *de novo* point mutations causing single nucleotide polymorphism (SNP) in either the promoter region or the open reading frame (ORF) of the gene, structural rearrangements of the genome or horizontal gene transfer from other organisms.[231,232] Whilst point mutations and genomic shuffling are the most frequent causes in laboratory evolution experiments, horizontal gene transfer is often responsible for the quick emergence of resistance in natural environments and the clinic (Figure 3.2).[225,227,233,234] Thus, these both varied and complex resistance mechanisms pose a significant challenge for the development of novel antibiotic small molecules, making this quest non-trivial. In this manner, the selection of novel and effective biological targets to disrupt is of utmost importance.[230,235]
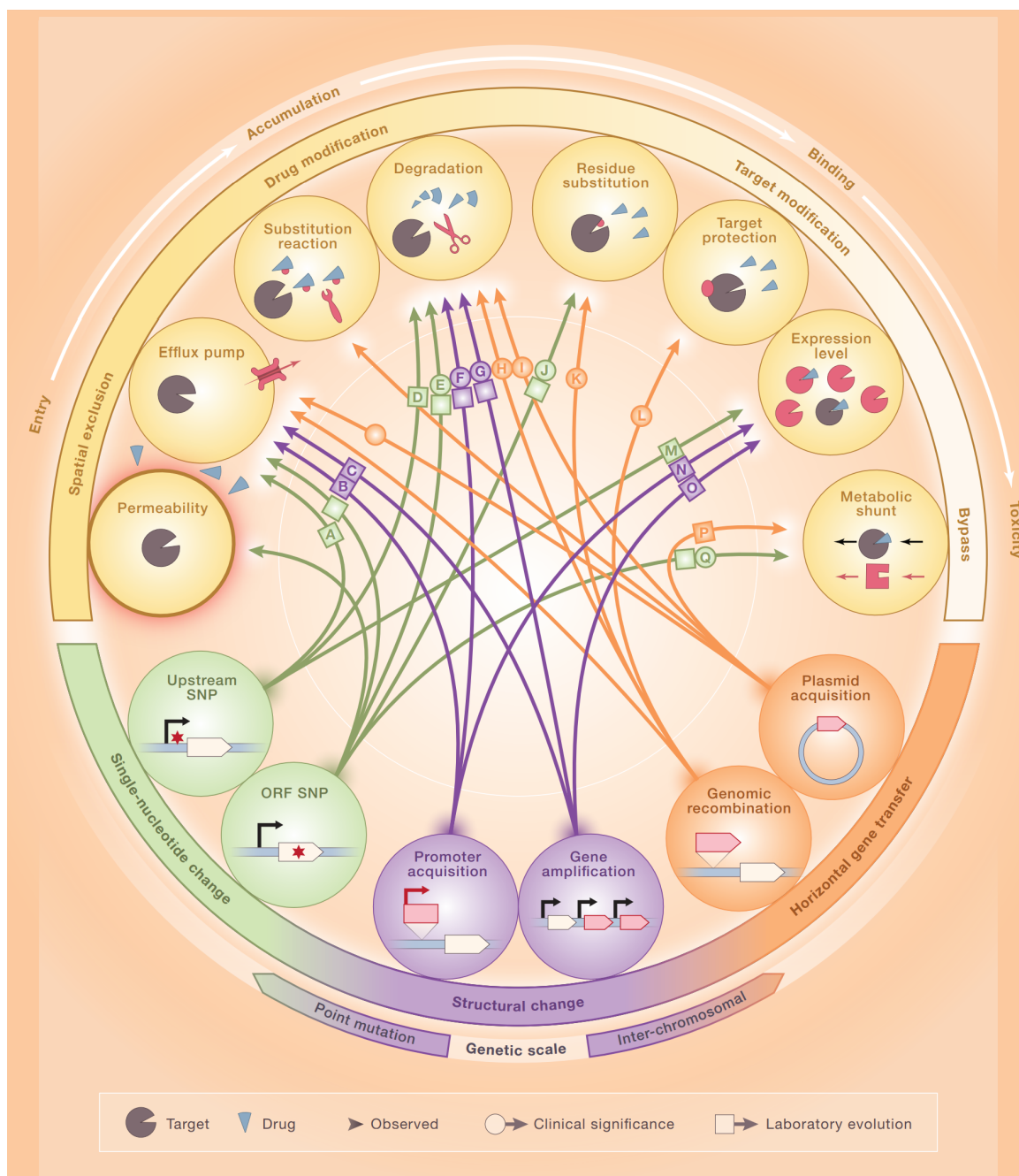
Figure 3.2 Antimicrobial resistance mechanisms and acquisitions. Mechanisms (yellow) can prevent the entry, accumulation, binding or toxicity of the drug, and can result from single-nucleotide changes (green), structural changes within the genome (purple) or horizontal gene transfer (orange) between cells. Figure reproduced from Yelin *et al*.[225] with permission.

3.1.3 Antibacterial targets

A clinically valuable antibiotic needs to meet two criteria: 1) inhibition or disruption of crucial bacterial cell function to efficiently attenuate or kill the pathogen, and 2) low toxicity towards the host for safe administration.[236] Accordingly, most known antibiotics either target functions that are only present in prokaryotes and not in eukaryotes like cell wall synthesis, or biological entities that differ significantly between these organisms such as the 70S *versus* 80S ribosomes.

The clinically relevant antibiotic classes are listed according to their target functions in Table 3.1.[203,204]

Table 3.1 Target functions of all antibiotic classes. Most antibiotic classes were first marketed during the Golden Age (1940 – 1970), and target one of the four major functions: folate, cell wall, protein and nucleic acid synthesis. Data from Hutchings *et al.*[203]

Target Function	Antibiotic Classes	
Folate synthesis	Sulfonamides	Sulfones
	Salicylates	Diaminopyrimidines
Cell wall synthesis	Penicillins	Cephalosporins
	Bacitracin	Thioamides
	Pyridinamides	*Phosphonates*
	Glycopeptides	*Carbapenems*
	Ethambutol	*Monobactams*
	Cycloserines	
Protein synthesis	Aminoglycosides	Lincosamides
	Tetracyclines	Streptogramins
	Amphenicols	*Mupirocin*
	Macrolides	***Oxazolidinones***
	Tuberactinomycins	***Pleuromutilins***
	Fusidic acid	
Nucleic acid synthesis	Azoles (DNA)	Ansamycins (RNA)
	(Fluoro)quinolones (DNA)	***Lipiarmycins*** (RNA)
	Phenazines (DNA)	
ATP synthesis	***Diarylquinolines***	
Multiple	Nitrofurans	
Unknown	Arsphenamines	

Antibiotic classes are coloured according to their source; green: actinomycetes; blue: other bacteria; purple: fungi; orange: synthetic. New antibiotic classes introduced to the clinic 1970 – 1999 are in italic, 2000 – 2019 are in bold italic.

Natural products have historically presented an incredibly rich source of antibiotics; however, these entities only target a small number of functions limited to enzymes involved with cell wall synthesis, protein synthesis (mainly by binding to the 70S ribosome), and RNA polymerase. Even though these targets comply with the principles above, these antibiotic classes have been employed in the microbial arms race for millions of years and as a result, resistance mechanisms have also evolved in nature. Moreover, whilst semi-synthetic analogues can retain activity against resistant strains, bacteria usually adopt to these incremental changes quickly, rendering these analogues equally ineffective. In contrast, many of the synthetic antibiotics were found to target alternative functions such as folate, DNA and adenosine triphosphate (ATP) synthesis.

However, resistance was observed soon after their first use in the clinic, proving that bacteria can develop resistance even to completely new antibiotic classes. The quickly spreading AMR and lacking development of new antibiotics since the end of the Golden Age of antibiotic discovery, resulted in an urgent need for finding novel targets.

The completion of the first bacterial whole genome sequencing from *Haemophilus influenzae* in 1995 sparked huge interest in the pharmaceutical community and enabled hundreds of prospective genes to be explored as potential antibacterial targets.[237] However, sadly this did not deliver and since the 1980s no new antibiotics have been developed. This period, termed the 40-year discovery void, is particularly unnerving.

Our struggles to develop novel and effective antibiotics are best exemplified with case studies from within the pharmaceutical industry. One of the largest scale research efforts was made by GlaxoSmithKline (GSK), spanning over 7 years (1995 – 2001).[238] This included in-depth study of the genomes of several pathogenic bacteria, both Gram-positive and Gram-negative, to identify highly conserved genes that occurred as single copies. Then the 358 genes that had no closely related human homologues were designated as potential broad-spectrum targets. These genes had been evaluated for essentiality using allelic-replacement mutagenesis, and 127 genes were identified as *in vitro* essential in at least one of the bacterial strains tested; almost half of which had known and assayable functions. Across this 7-year period, a total of 70 HTS campaigns were run against these individual targets as well as whole biosynthetic pathways and living cells, screening more than half a million compounds. However, hits were only identified from 16 HTSs and only 5 of these were successfully developed into lead compounds with *in vivo* activity. Strikingly, none ultimately reached the clinic. Similar difficulties were faced by many other companies investing in similar efforts, with more than 125 HTS campaigns against over 60 novel targets by 34 different companies published during 1995 – 2004. Only 2 candidates reached the phase I clinical trials.[239] These efforts highlight the extreme challenge of developing novel and effective broad-spectrum antibiotics, further evidencing why other strategies besides HTS may need to be employed, whilst the search for novel antibiotic targets must continue.

3.1.4   Metabolic pathways

Most of the clinically used antibiotics target the synthesis of macromolecules, including the cell-wall, proteins and nucleic acids. The inhibition of crucial metabolic pathways, however, poses an important alternative approach achieving lethality or growth attenuation.[236]

3.1.4.1   *Folate synthesis*

Tetrahydrofolate (THFo) is a key co-factor used by all living organisms for the transfer of methyl and formyl groups required for the synthesis of pyrimidines, purines, amino acids, *S*-adenosylmethionine and formyl-methionine. Animals are unable to synthesise folate and must therefore obtain it from their diet. Plants, fungi and most prokaryotes however synthesise THFo *de novo* and are incapable of its uptake from the environment. Within these organisms, however,

variations of the essentially same biosynthetic pathway exist, which renders many of the involved enzymes unattractive targets for broad-spectrum antibiotics and also suggest resistance could also develop very easily.[240–242]

One highly conserved enzyme, and therefore valuable target, is dihydropteroate synthase (DHPS), which catalyses the conversion of 6-hydroxymethyl-7,8-dihydropterin pyrophosphate (DHPPP) and *para*-aminobenzoic acid (PABA) into 7,8-dihydropteroate (DHP). The first synthetic antibiotics including sulfonamides, sulfones and salicylates bind to the active site by mimicking PABA and either inhibit DHPS or get incorporated into DHP analogues and as a result inhibit downstream enzymes in the folate biosynthesis (Figure 3.3).[243,244]
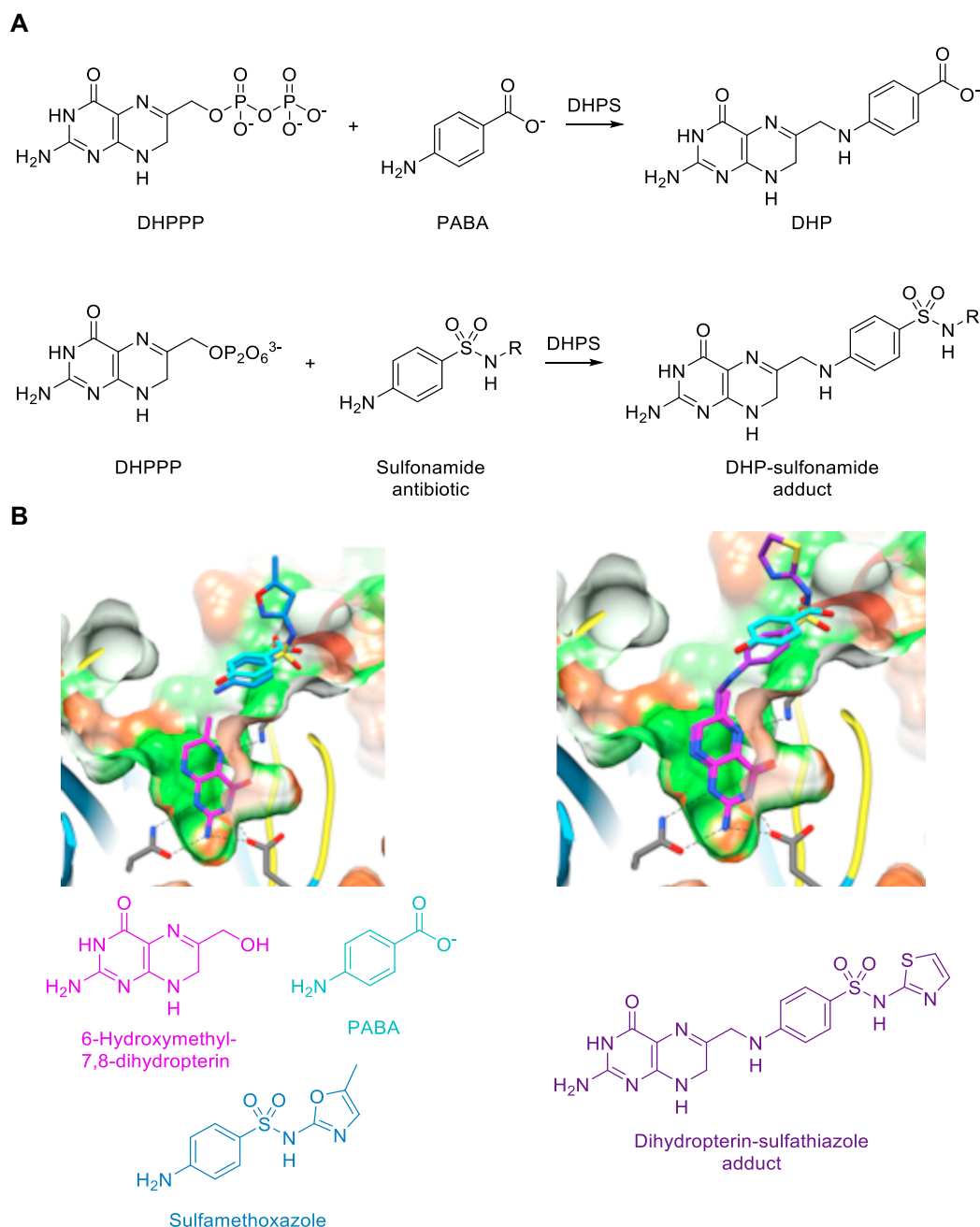
Figure 3.3 A) Reactions catalysed by dihydropteroate synthase (DHPS). B) X-ray structures of DHPS in complex with substrate analogues (left) and products (right). The overlaid structures show how sulfonamides can mimic the natural substrate *para*-aminobenzoic acid (left) and get incorporated into DHP analogues. X-ray images reproduced from Bourne *et al.*[245].

Another well-established target is the upstream dihydrofolate reductase (DHFR), which directly forms THFo. Although DHFR is present in both prokaryotes and eukaryotes, structural differences make selective targeting possible; thus, diaminopyrimidines such as trimethoprim can be used as safe antibiotics, whilst pyrimethamine is an anti-protozoal and methotrexate is an anti-cancer agent (Figure 3.4).[245–247]
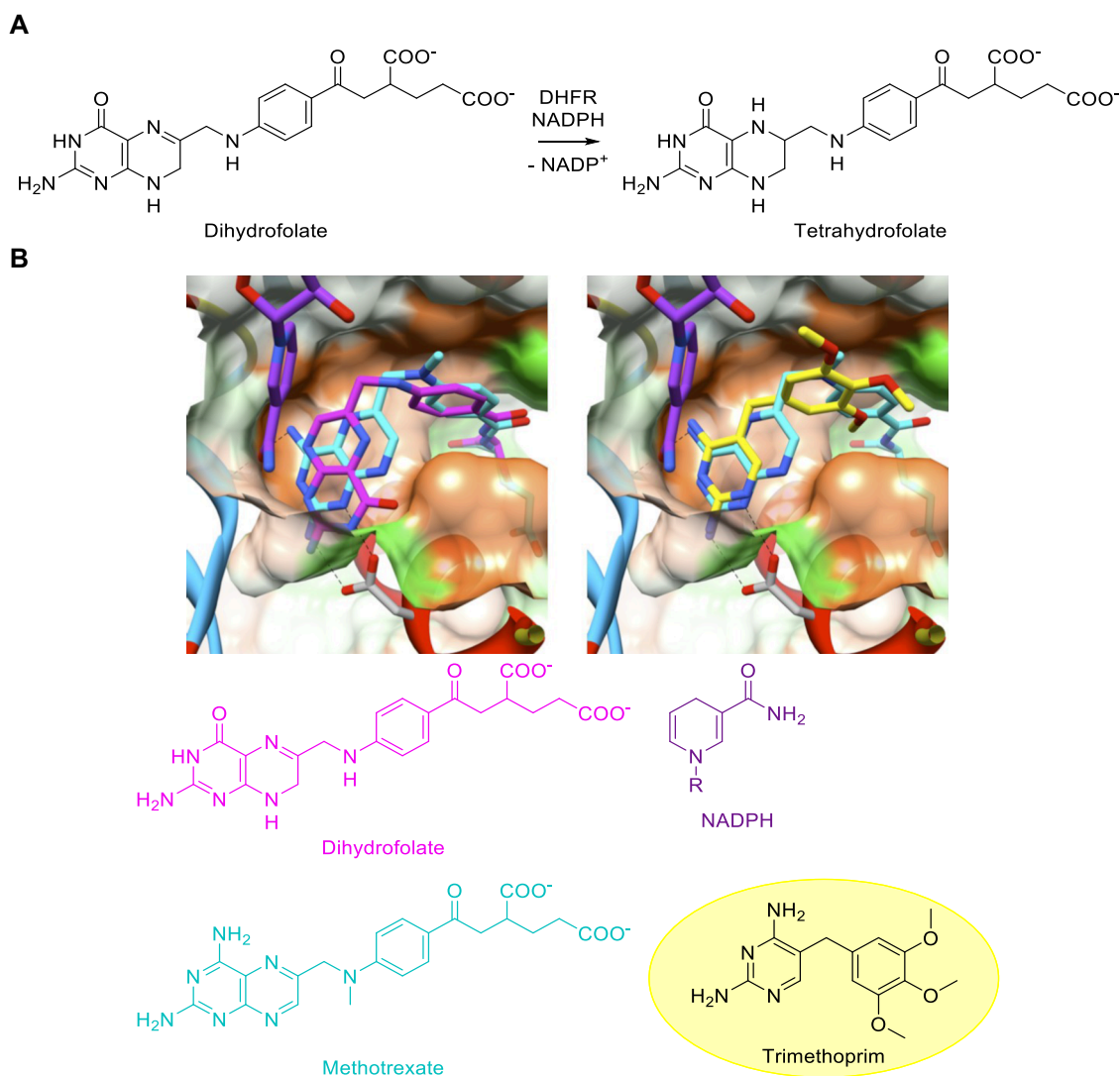
Figure 3.4 A) Reaction catalysed by dihydrofolate reductase (DHFR). B) X-ray structures of DHFR in complex with NADPH and dihydrofolate substrates as well as two clinically used inhibitors: Methotrexate and Trimethoprim. X-ray images reproduced from Bourne *et al*.[245]

3.1.4.2   Fatty acid synthesis

Fatty acid biosynthesis is crucial for membrane and cell-wall generation in all bacteria.[248] Accordingly, all genes involved are essential and therefore potential antibacterial targets.[249]. Both the biotin carboxylase and carboxyltransferase domains of the acetyl coenzyme A carboxylase—the first dedicated enzyme in fatty acid biosynthesis—have been utilised for developing novel small molecules with antibacterial activities.[250,251] In particular, enoyl-acyl carrier protein reductase (FabI) has been identified as a particularly attractive target.[252] Indeed, it was found that both the pyridineamides like Isoniazid and diphenyl ethers like Triclosan primarily target FabI (Figure 3.4, A).[253,254] Isoniazid is a prodrug that gets activated by the catalase-peroxidase enzyme KatG forming Isonicotinoyl radical which reacts with nicotinamide adenine dinucleotide (NAD^+) forming an Isoniazid–NAD adduct that binds to FabI.[255] In contrast, Triclosan—as well as other inhibitors within this category, including the structurally dissimilar CG400549—binds to the FabI–NAD(P) complex (Figure 3.5, B).[256,257] Both Isoniazid and Triclosan

show slow on/off kinetics increasing the likelihood of bacteria acquiring resistance *via* missense mutations.[258] Promisingly, high affinity inhibitors with fast on/off kinetics have also been developed against the FabI–NAD(P)H complex, including AFN-1252 which has completed phase II clinical trials (Figure 3.5, C).[259,260]
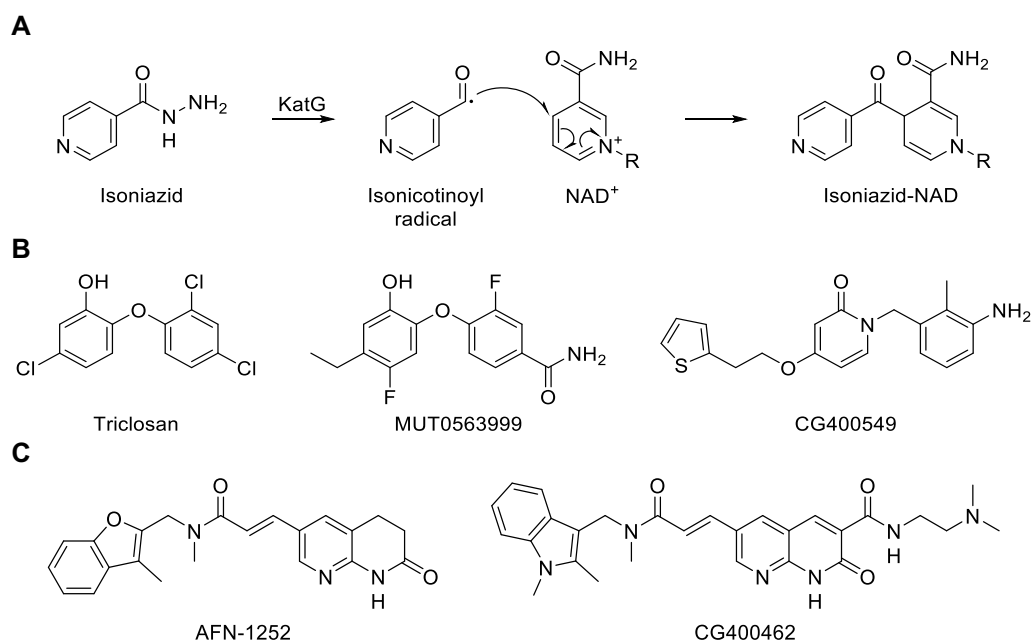
Figure 3.5 Different classes of antibiotics targeting enoyl-acyl carrier protein reductase (FabI). A) Isoniazid prodrug gets activated by the catalase-peroxidase enzyme KatG and reacts with NAD+ forming the inhibitor isoniazid–NAD adduct. B) Examples of substituted diphenyl ethers and CG400549 as inhibitors of the FabI–NAD(P) complex. C) AFN-1252 and CG400462 are inhibitors of the FabI–NAD(P)H complex. Figure adapted from Yao *et al.*[258].

In addition to FabI, other enzymes involved in fatty acid biosynthesis have been investigated such as acetyl coenzyme A carboxylase. In this case both the biotin carboxylase and carboxyltransferase domains have been utilised for developing novel small molecules with antibacterial activities.[250,251] However, none of them were tested in clinical trials.

3.1.4.3 Lipid A synthesis

Lipopolysaccharides are vital components of the outer membrane in Gram-negative bacteria consisting of the O-antigen and core domains, as well as lipid A which anchors the polysaccharide domains to the lipid membrane. Since the zinc-dependent metalloamidase uridine-diphosphate-3-*O*-(*R*-3-hydroxymyristoyl)-*N*-acetylglucosamine deacetylase (LpxC) is a highly conserved enzyme crucial for the biosynthesis of lipid A and therefore the generation of the outer membrane, it has been identified as a promising potent broad-spectrum antibacterial target. So far two inhibitors of LpxC have reached the phase I clinical trials, ACHN-975 (Figure 3.6) and RC-01 (structure undisclosed), but both failed due to safety reasons.[261,262]
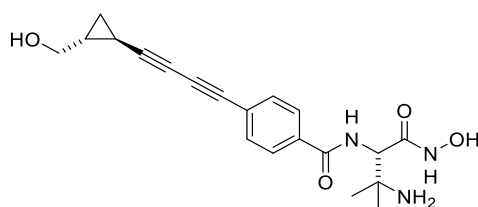
Figure 3.6 Structure of the antibacterial ACHN-975, a potent inhibitor of LpxC.

3.1.5 Propionate metabolism

Many bacteria generate substantial amounts of short-chain fatty acids as by-products of fermentation and in the mammalian gastrointestinal tract these entities play crucial roles in sustaining healthy microbiota.[263] Propionic acid can be converted to propionyl coenzyme A (CoA), a valuable building block for lipid biosynthesis and carbon source, also produced during the degradation of odd- and branched-chain fatty acids, branched-chain amino acids and cholesterol.[264] However, propionate and its metabolic derivatives have been found to be highly toxic if accumulated within cells.[265] Although the mechanism is not fully understood, evidence shows that there are several ways by which they might cause lethality, including the inhibition of pyruvate dehydrogenase and fructose-1,6-bisphosphatase by propionyl-CoA and 2-methylcitrate, respectively.[266,267] Nevertheless, the growth of most microorganisms has been shown to be attenuated by the presence of propionate in the growth medium.[265] As a result, propionic acid and its salts are widely used as preservatives in the food industry, whilst their use as feed additives have been shown to reduce the risk of *Salmonella* infection in farm animals.[265] Several metabolic pathways have evolved to alleviate the toxicity and facilitate the metabolic utility of propionate. Whilst mammalian cells mainly rely on the vitamin B_{12}-dependent methylmalonyl pathway, most bacteria and fungi use the 2-methylcitrate (2-MC) cycle.[268] Since propionate is present in adequate concentrations in many infection sites, the 2-MC cycle has been identified as an appealing antibacterial target.[264] Indeed, deletion of genes involved with the 2-MC cycle was shown to inhibit the growth of many bacterial strains on rich growth media containing propionate, and *Mycobacterium tuberculosis* growth and survival in *ex vivo* infected murine bone-marrow-derived macrophages.[269]

The tricarboxylic acid (TCA) cycle is used by all aerobic organisms to generate energy by oxidising the acetyl group (from acetyl-CoA) completely to CO_2, and it is the most important central pathway connecting almost all other individual metabolic pathways.[270,271] In the glyoxylate cycle the oxidative decarboxylation steps of the TCA cycle are bypassed by two alternative enzymes (isocitrate lyase and malate synthase) allowing the conversion of acetyl groups into four-carbon dicarboxylic acids thus rendering the original catabolic TCA cycle into an anabolic variant.[272,273] The 2-MC cycle is closely related to the TCA and glyoxylate cycles but utilises propionyl-CoA in an aldol reaction with oxaloacetic acid (OAA) forming 2-MC rather than citrate from acetyl-CoA (Scheme 3.1). This first step is catalysed by a methylcitrate synthase (encoded by *prpC*) that has some sequence analogy to citrate synthases, however, due to the often relaxed substrate specificities of the enzymes involved in both the TCA and the 2-MC cycles, citrate synthases can

also catalyse this step leading to lethal accumulation of 2-MC in cells with high propionyl-CoA levels and non-functional 2-MC cycle. The dehydration of 2-MC to 2-methyl-*cis*-aconitate can either be catalysed by the methylcitrate dehydratase (PrpD) directly, or *via* the formation of 4-methyl-*cis*-aconitate by the alternative aconitase (AcnD) followed by isomerisation mediated by a specific isomerase (PrpF). Next, 2-methyl-*cis*-aconitate is hydrated by the aconitase (AcnB) to form 2-methylisocitrate, which undergoes a retro aldol reaction catalysed by 2-methylisocitrate lyase (PrpB) generating pyruvate and succinate. The succinate then undergoes the same oxidative steps as in the TCA and glyoxylate cycles reforming oxalate, whereas pyruvate is an extremely versatile building block for anabolic processes such as amino acid and fatty acid synthesis, and it can also undergo decarboxylative oxidation forming acetyl-CoA for energy generation. Both propionyl- and acetyl-CoA can be generated from the corresponding acids by acyl-CoA synthases such as the propionyl-CoA synthase PrpE, or they can be formed during catabolic processes such as the degradation of fatty acids, carbohydrates and amino acids.[264,268]
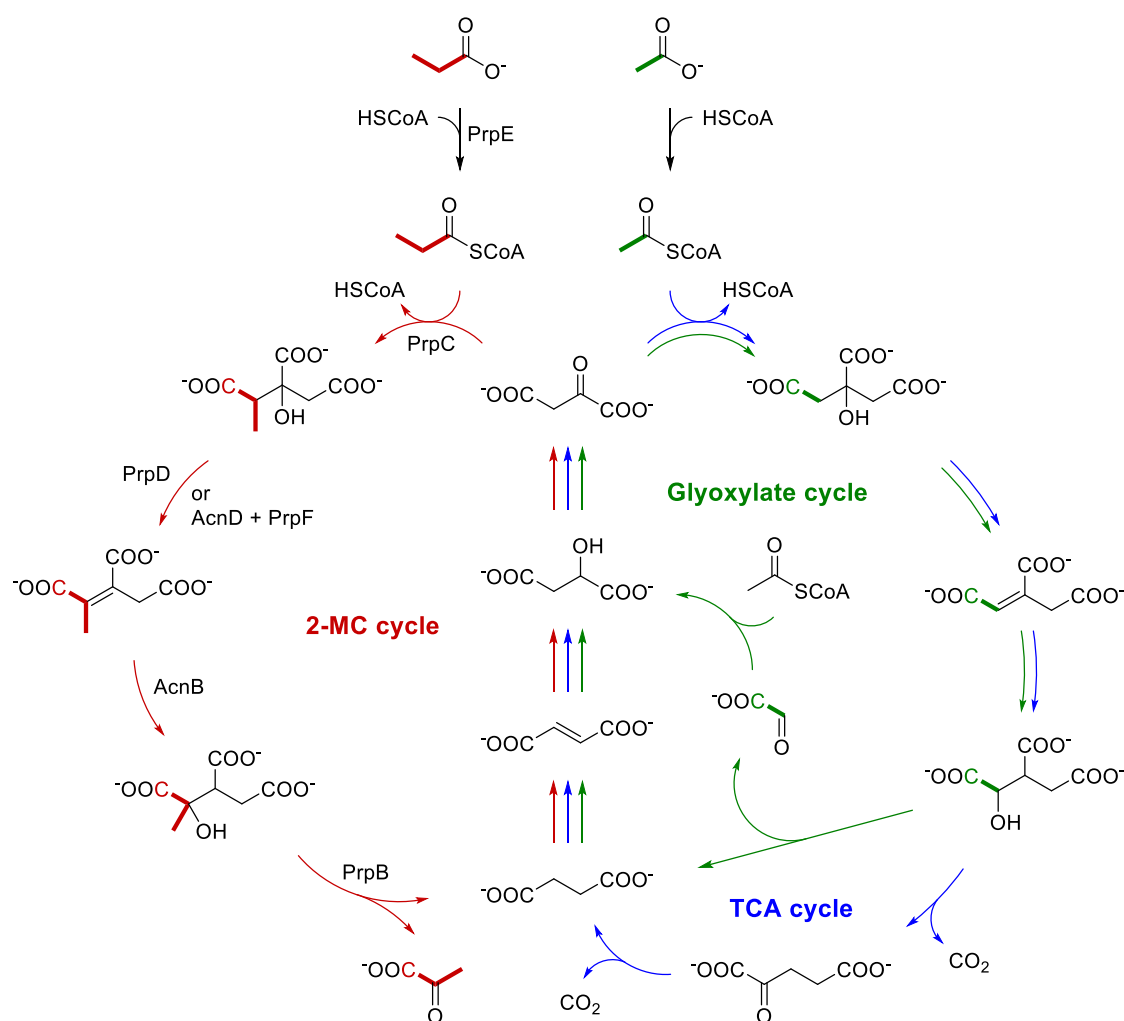
Scheme 3.1 Bacterial metabolism of acyl coenzyme A (CoA) units. The tricarboxylic acid (TCA) cycle, highlighted in blue, is used to oxidise the acetyl group to CO_2 and generate energy under aerobic condition. The glyoxylate cycle, highlighted in green, is an anabolic modification of the TCA cycle whereby two acetyl groups are converted into four-carbon dicarboxylic acids. The 2-methylcitrate (2-MC) cycle, highlighted in red, is an analogous process to the glyoxylate cycle adapted to metabolise propionyl-CoA, the propionyl group, highlighted in red, is oxidised to pyruvate, a versatile metabolite. Stereochemistry is not shown.

Recently, a whole cell phenotypic HTS campaign identified several hits inhibiting *Mycobacterium tuberculosis* growth in macrophages; compound V-13–009920 was found to be an inhibitor of PrpC with an *in vitro* IC_{50} of 4.0 µM and *in vivo* IC_{50}'s of 3.0 µM and 0.3 µM in macrophages and cholesterol media respectively (Figure 3.7).[274] Even though promising hits have been identified in this HTS campaign, the development of inhibitors with much higher potency is required for clinical utility.
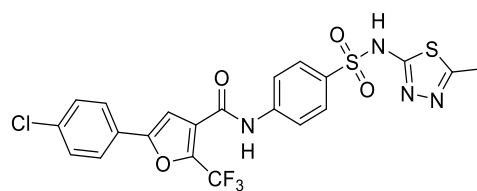
Figure 3.7 Structure of V-13–009920, an inhibitor of *Mycobacterium tuberculosis* 2-methylcitrate synthase (PrpC) and growth in macrophages.

3.2 Previous Work

The 2-MC cycle has been studied extensively as part of the ongoing research into developing our understanding of the physiology and virulence of *Pseudomonas aeruginosa*—a common clinical pathogen with outstanding resistance to antibiotic treatments due to its ability to form biofilms—within the Welch Group. Due to the challenges and high costs associated with HTS campaigns as well as the need for developing new chemotypes for the inhibition of this promising target pathway, a FBDD approach proved an attractive alternative (see section 1.5). Recently, the *P. aeruginosa* PrpC enzyme was successfully crystallised by Andre J. Wijaya in complex with OAA to facilitate structural analysis by X-ray crystallography. Furthermore, a screening—by soaking the PrpC–OAA co-crystals in 10 mM solutions of 580 fragments according to Collins *et al*.[275]—was conducted at the Diamond Light Source, identifying 11 hits (Figure 3.8). This revealed that PrpC has a dimeric structure with most of the residues forming the OAA binding pocket contributed by a single polypeptide chain. On the other hand, most of the fragment hits bind at the interface between the two subunits, and a considerable amount of overlap between these fragments is also evident.
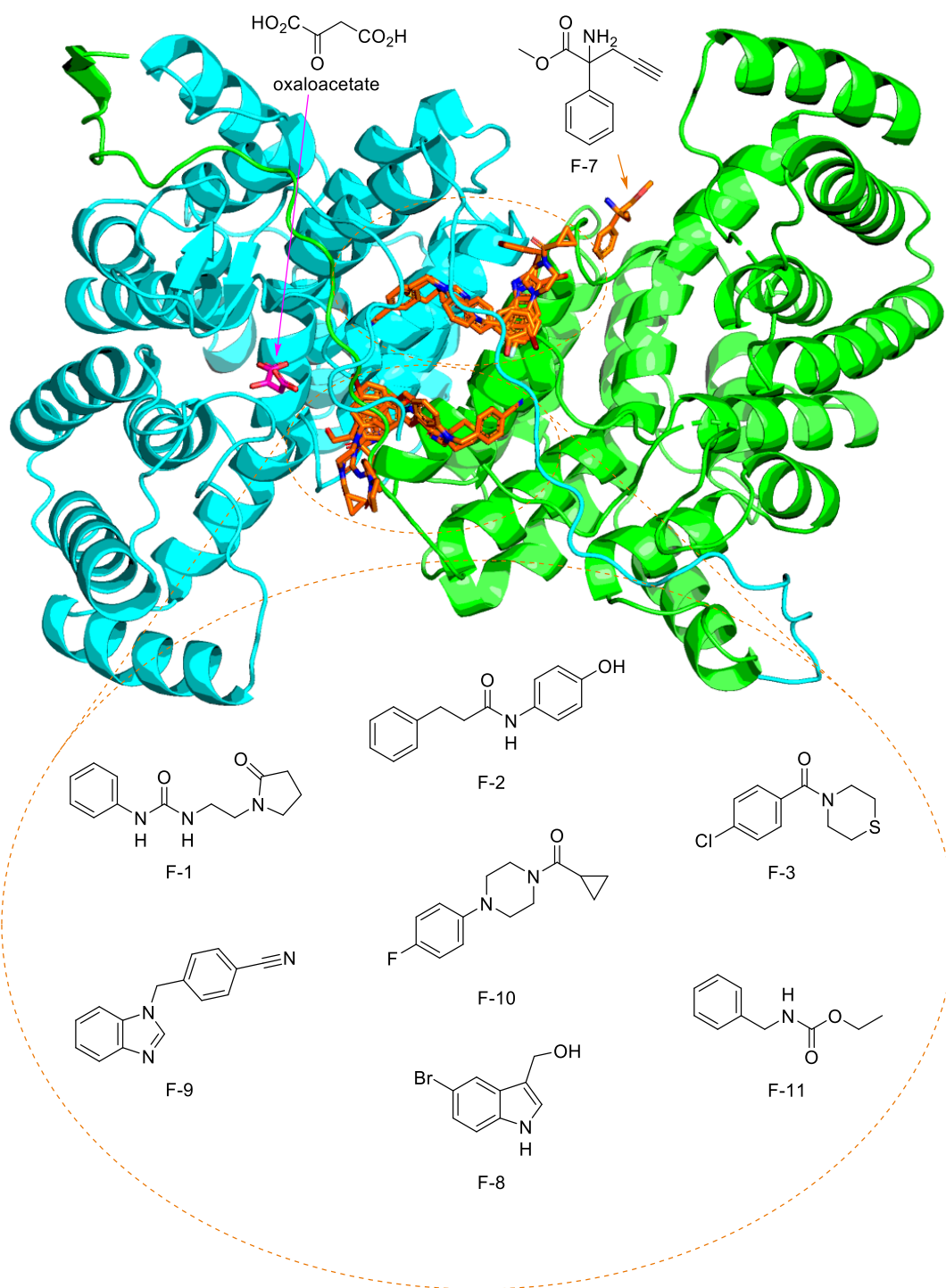
Figure 3.8 X-ray crystal structure of the dimeric *P. aeruginosa* PrpC enzyme (the two chains are coloured in green and cyan) in complex with OAA (co-crystallised, highlighted in magenta); 1.65 Å resolution, PDB6S87. A selection of the initial X-ray hits (soaked) is overlaid in orange, with their structures shown.

Subsequently, two orthogonal assay methods were utilised to validate the binding of these initial X-ray hits. Firstly, a thermal shift assay using differential scanning fluorimetry (DSF)[‡] was used, where a higher melting temperature correlates to a more stable protein–ligand complex.

[‡]Ligands were screened at a concentration of 1 mM against PrpC at a concentration of 5 µM in a pH 7.5 buffer solution of 100 mM NaCl, 50 mM Tris·HCl and 5% (v/v) glycerol using 5X SYPRO Orange Fluorescent Dye.[282]

Therefore, this analysis identified F-2 as the strongest binder (with a corresponding melting temperature of 50.5 °C) followed by F-8 (49.5 °C), both with an increase in melting temperature of approximately 7 °C compared to the DMSO control. Conversely, an *in vitro* enzymatic assay[§] measuring the release of free CoA—following the PrpC catalysed aldol condensation between propionyl-CoA and OAA—by UV-VIS spectroscopy was used. In this assay, fragment F-7 was in fact found to be the most active with a 78% inhibition at 1 mM concentration. Importantly, however, F-8 and F-9 also displayed good activity, whilst the remaining fragments displayed only moderate activity (Figure 3.9).

Both X-ray crystallography and thermal shift assays are prone to give false positives—compounds that bind to the target strongly enough for detection, but without causing significant changes in its function or activity. On the other hand, although the enzymatic assay gives a direct measure of compound activity, as all techniques, it still suffers from the possibility of identifying false positives (see section 1.4.1). Therefore, indole fragment F-8, with good activities in both assays, was identified as a validated fragment hit with the highest confidence and thus the most promising starting point for the development of more potent inhibitors of PrpC.
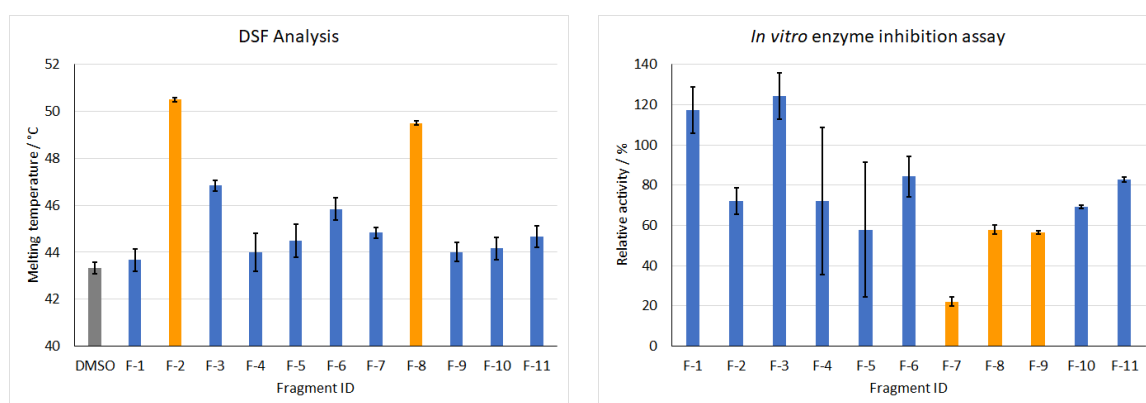
Figure 3.9 Hit validation by differential scanning fluorimetry (DSF, left) and *in vitro* enzyme inhibition (right) assays, compounds were used at 1 mM concentration. The most active fragments are highlighted in orange.

To gain insights into potential avenues for fragment elaboration it was crucial to firstly review the structural information. Comparison of the X-ray crystal structures of PrpC in complex with OAA and F-8 by A. J. Wijaya revealed that the indole fragment inserts into a hidden pocket displacing a loop consisting of residues 352 – 358. As a result, the sidechain of Arg356—one of the key residues forming ionic and hydrogen-bonding interactions with OAA within the active site of PrpC—was moved away from OAA, whereas the rest of the pocket is unchanged (Figure 3.10). Therefore, a dual mechanism of PrpC inhibition by F-8 was hypothesised: 1) displacement of Arg356 reduces the affinity and specificity of PrpC to OAA; and 2) the induced conformational

---

[§]Ligands were screened at a concentration of 1 mM against PrpC at a concentration of 10 µg/mL in a pH 7.5 buffer solution of 100 mM KCl, 50 mM HEPES and 4% (v/v) glycerol using 0.15 mM DTNB; absorbance was read at 420 nm.[283]

change of the protein as well as the presence of F-8 sterically prevents propionyl-CoA from binding to the active site. It was also envisaged that the core fragment could be grown in either the OAA pocket or the propionyl-CoA channel, with the most promising exit vector being the *C*-3 position of the indole core.

A

B
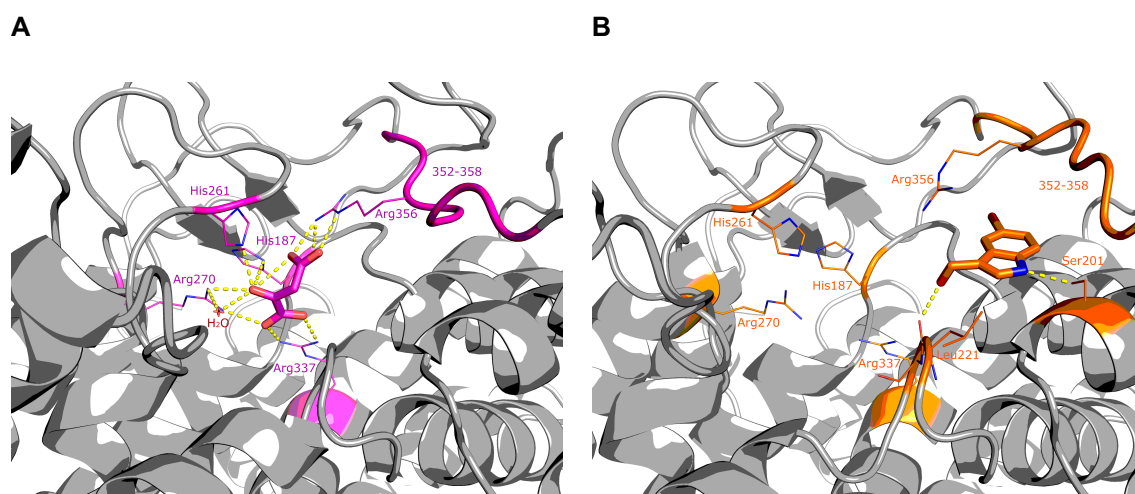
Figure 3.10 A) Structure of the OAA binding pocket, key residues and OAA are highlighted in magenta. B) X-ray crystal structure of PrpC in complex with F-8 (2.0 Å resolution), key residues and F-8 are highlighted in orange.

A common initial hit validation strategy within FBDD involves conducting an SAR-by catalogue approach using commercially available analogues to explore a binding pocket. Crucially, this avoids lengthy and costly synthetic investment during the early stages of a project. Accordingly, 34 commercially available analogues of F-8—containing either the initial monobromoindole or a dibromoindole core and featuring variable groups in the *C*-3 position—were first obtained and screened by DSF, leading to the identification of further potent fragments (Figure 3.11, A and B). The binding modes of these fragments were also resolved using X-ray crystallography showing consistency with that of F-8—seen from the overlapping bromoindole units—whereas the *C*-3 appendages were found to point in different directions (Figure 3.11, C). However, no IC$_{50}$ values could be determined due to lack of sufficient quantities of these compounds. With these results in hand, initial hit validation has proven successful.
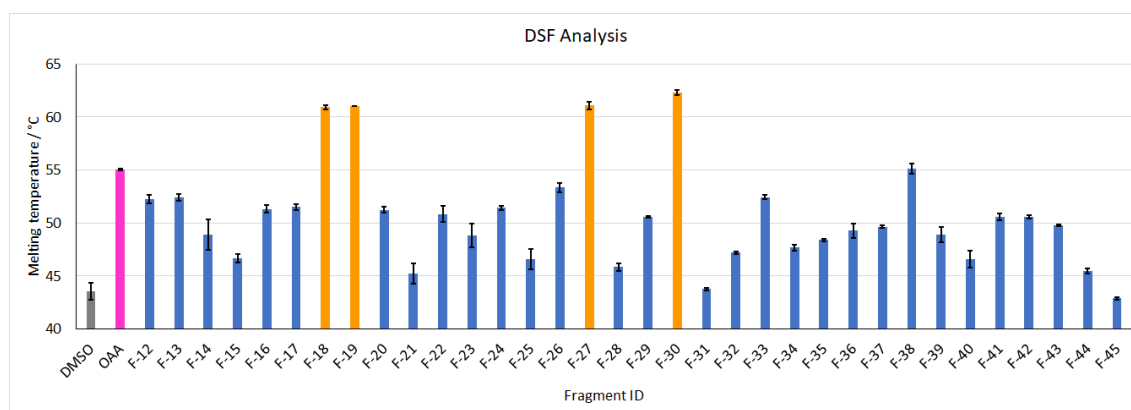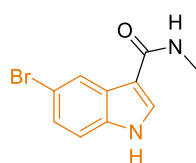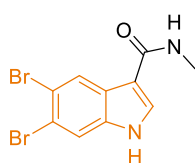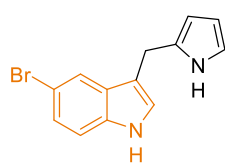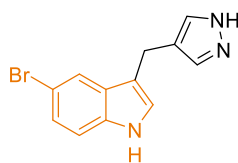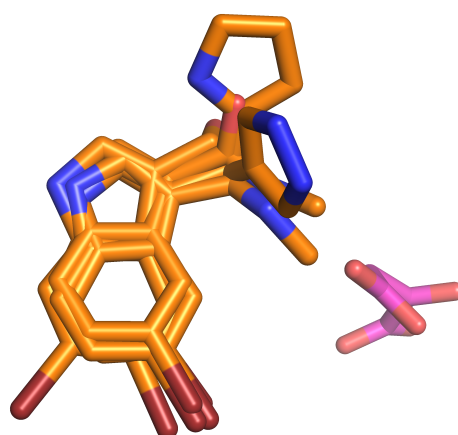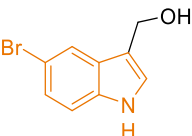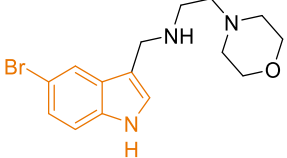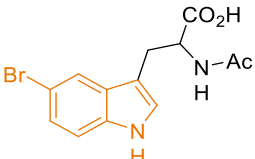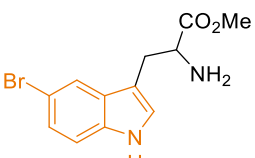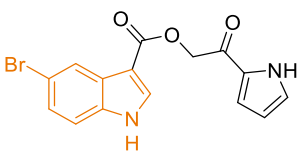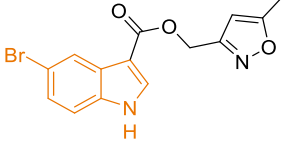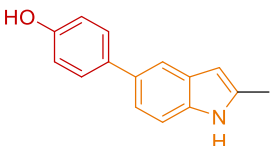
Figure 3.11 A) Screening results of F-8 analogues using differential scanning fluorimetry (DSF), OAA is highlighted in magenta, the most potent analogues are highlighted in orange. B) The structures for the most potent compounds F-18, F-19, F-27 and F-30 are shown with the bromoindole units highlighted in orange. C) An overlay of F-8 and the most potent analogues in orange and OAA in magenta, atomic coordinates taken from X-ray crystal structures.

Inspired by these results, a further five additional analogues featuring modifications to the C-3 appendages were purchased from Enamine in slightly larger quantities enabling the determination of IC_{50}'s. An additional analogue was also synthesised by Dr. Elaine Fowler in the Spring Group to investigate the hydrophobic pocket further by replacing the bromine with the 4-hydroxyphenyl group (Table 3.2). The ester-linked isoxazole F-50 was not adequately soluble for the assay, whereas the bromotyrosine derivatives F-47 and F-48 showed much lower potency compared to the original fragment hit demonstrating that certain C-3 appendages are not suitable for binding. On the other hand, the amine-linked morpholine F-46 and the ester-linked pyrrole F-49 were found to be thirty and seven times more potent than F-8 respectively, however, a decrease in LE can be observed in both cases—most prominent for F-49—indicating that the C-3 appendages do not bind to PrpC as strongly as the bromoindole core itself. The relatively high potency of F-51 indicates that the hydrophobic pocket is flexible enough to accommodate a

larger hydroxyphenyl group resulting in only a slight decrease in LE. These results show that optimising both the *C*-3 appendage and the hydrophobic core can prove to be viable approaches for developing more potent inhibitors for PrpC.

Table 3.2 Half-maximal inhibitory concentrations and ligand efficiencies of compounds assessed by *in vitro* PrpC enzyme inhibition assay.

Compound ID	Structure[d]	IC$_{50}$[e]	HAC[h]	LE[i]
F-8[a]		825 ± 370 µM	12	0.35
F-46[b]		27 ± 2 µM	20	0.31
F-47[b]		>1 mM	19	N/A
F-48[b]		>1 mM	17	N/A
F-49[b]		118 ± 19 µM	21	0.26
F-50[b]		N/A[f]	20	N/A
F-51[c]		97 ± 24 µM[g]	17	0.32

[a]Original fragment hit. [b]Purchased from Enamine. [c]Synthesised by Dr. E. Fowler. [d]Bromoindole units highlighted in orange. [e]Half-maximal inhibitory concentration against PrpC enzyme; data given as best-fit parameter followed by standard error. [f]Compound was not soluble enough for assay. [g]Precipitation was observed during the assay, which may have compromised the result. [h]Heavy atom count. [i]Ligand efficiency in units of kcal mol^{-1}.

3.3 Project Outline

Following these promising preliminary results, a synthetic project was proposed in order to elaborate the fragment hits into more potent inhibitors for PrpC. As shown in Figure 3.11, C) the crystal structure data highlighted that the varying groups from the *C*-3 position of the indole ring were found to project towards the OAA binding pocket. Moreover, as depicted in Figure 3.10, the residues interacting with the α-keto acid unit of OAA remain unaffected by the indole fragment binding, whereas Arg356, which is capping the other carboxylate group, is displaced, allowing for chemical modifications at that position. In light of these findings (Figure 3.12, A), a fragment linking approach (see section 1.5.2) was proposed—utilising varied linkages between the aforementioned *C*-3 position of the indole core and an OAA mimicking warhead—as an ideal strategy to rapidly grow the initial fragment and develop a small molecule binder for PrpC with increased potency (Figure 3.12, B).

A
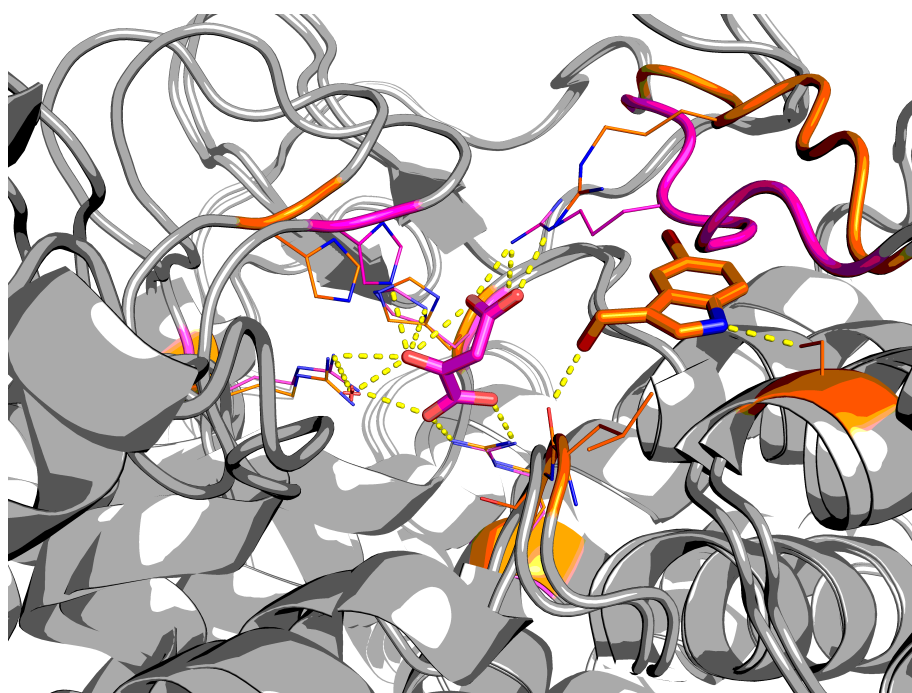
B
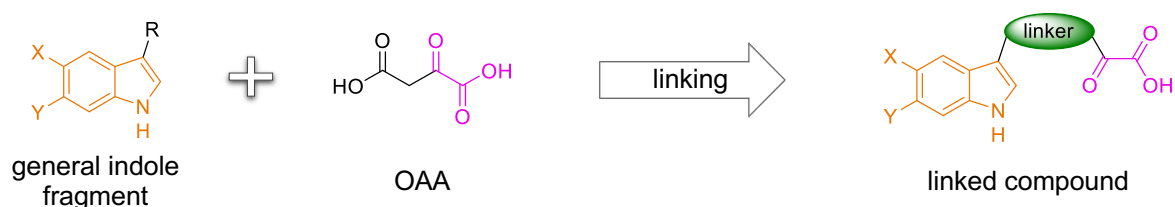
Figure 3.12 A) An overlay of PrpC in complex with OAA (fragment and key residues highlighted in magenta) and F-8 (fragment and key residues highlighted in orange). Atomic coordinates were taken from X-ray crystal structures. B) Proposed fragment linking approach to develop a more potent binder for PrpC. The indole and α-keto acid fragments are highlighted in orange and magenta respectively, while the linker is shown in green.

3.4 Results and Discussion

3.4.1 Bis-amide linked fragments

In accordance with the project proposal, different linkers to connect the two binding portions were firstly considered. Initially, bis-amide linked compounds **46** were identified due to their resemblance to known fragment binders—F18 and OAA—and hypothesised synthetic tractability as well as chemical stability (Scheme 3.2). The optimal length was assumed to correspond to either the ethylenediamine-based linker (n = 1) or its homologue (n = 2) based on X-ray crystal structures (Figure 3.12, A).

F-18   OAA   **46a**: n = 1
                                                              46b: n = 2
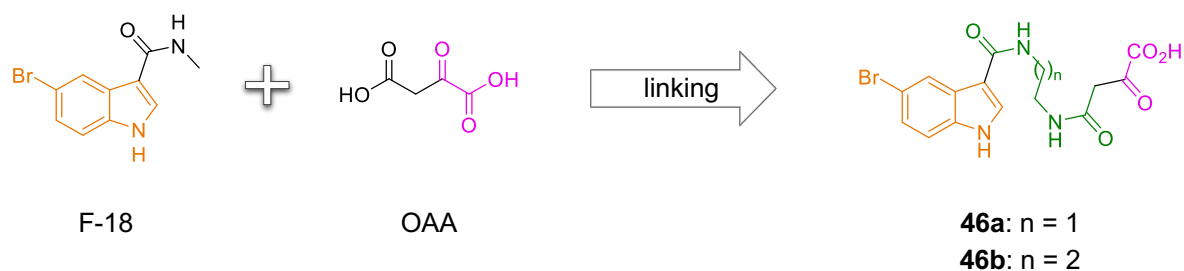
Scheme 3.2 Fragment linking approach for bis-amide linked compounds **46a,b**.

Initially, to access the Boc-protected intermediates **47a,b**, it was proposed to utilise amide couplings to attach varied linker units to the commercially available 5-bromo-1*H*-indole-3-carboxylic acid **48** (Scheme 3.3, A). Upon subjection of acid **48** to EDC-mediated amide formations, **47a,b** were isolated in moderate yields (Scheme 3.3, B). It was hypothesised that the activated ester formed has a low reactivity due to conjugation to the indole ring resulting in a sluggish coupling reaction with the desired amine. On the other hand, the same conjugation increases the acidity of the indole *NH*, making it possible for a weak base to deprotonate and thus promote the undesired acylation of the indole ring.

A

47a: n=1
47b: n=2

48

B

48

H$_2$N～(～)$_n$～NHBoc (1.5 eq)

EDC•HCl (1.5 eq), HOBt•xH$_2$O (1.5 eq), Et$_3$N (1.5 eq)

DMF, 0 °C to rt, 2 d

47a: n=1, 35%
47b: n=2, 56%*
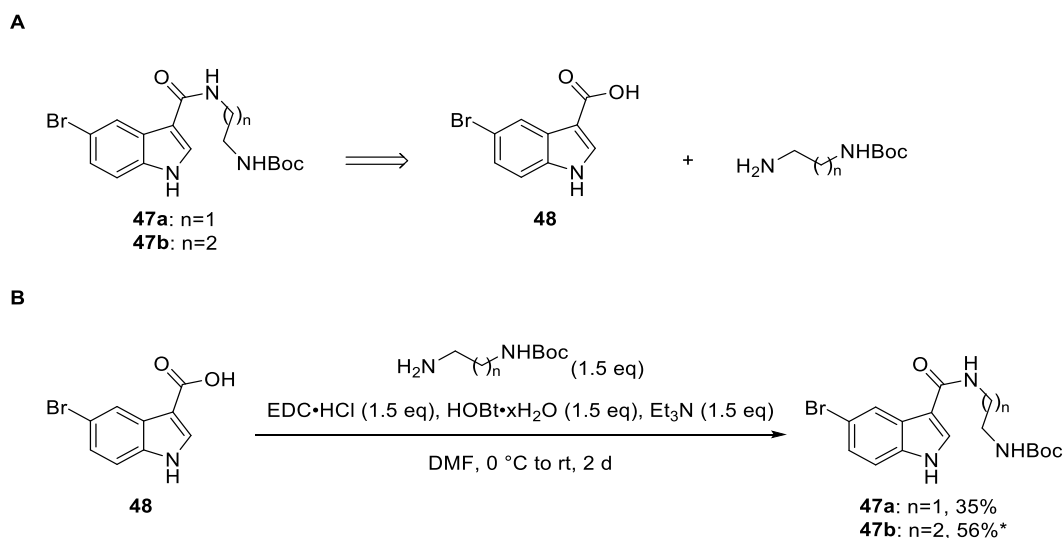
Scheme 3.3 Proposed route to (A) and initial synthesis of (B) the Boc-protected bromoindole-coupled amines **47a,b**. *Product is contaminated with an inseparable impurity.

Due to the low observed yields of **47a,b**, an alternative strategy to access these molecules was proposed, instead proceeding via the acyl chloride intermediate **49**. Indeed, treatment of **48** with thionyl chloride to form **49**, followed by amide formation enabled isolation of **47a** in an improved 78% yield from a one-pot reaction (Scheme 3.4).

48

SOCl$_2$ (2.0 eq)

CH$_2$Cl$_2$
reflux, o/n

49

H$_2$N～～NHBoc (2.0 eq)

DIPEA (10.0 eq)

THF, 0 °C → rt, 2 h
78%

47a
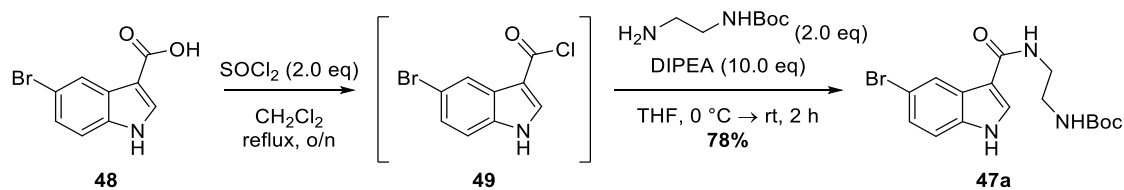
Scheme 3.4 Improved synthesis of the Boc-protected bromoindole-coupled amine **47a**.

Next, to enable downstream coupling of the warhead portion, the removal of the protecting group was first explored. Treatment of the Boc-protected intermediates with hydrochloric acid gave the corresponding crude amine hydrochlorides **50a,b** in quantitative yields (Scheme 3.5). The direct coupling of amine **50a** with OAA was first attempted under the previously used coupling conditions (Scheme 3.5). However, in this instance a complex mixture was obtained, and the mass of the desired product could not be identified by liquid chromatography–mass spectrometry (LCMS) analysis.
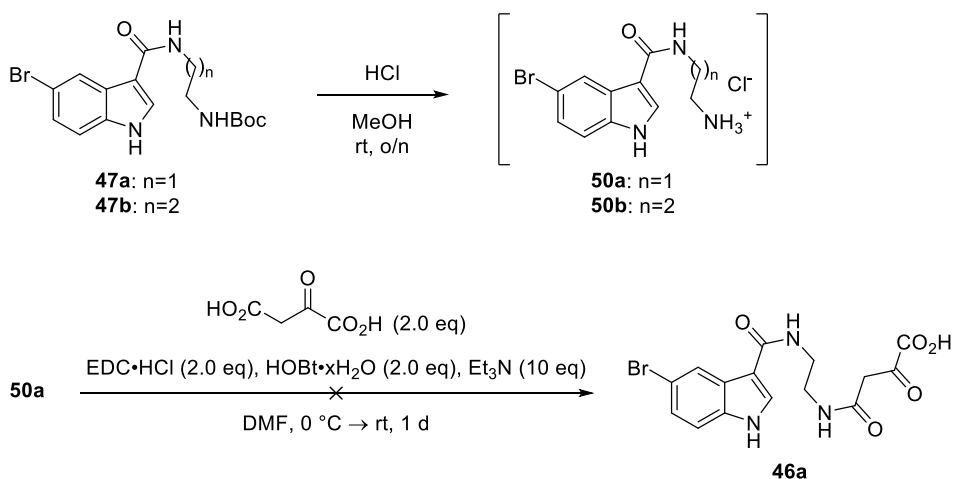
Scheme 3.5 Attempted synthesis of oxaloacetamide **46a**.

Due to the difficulties associated with the use of OAA in the amide coupling reaction, the synthesis of a suitably protected OAA derivative was investigated. First, a simple enolate alkylation approach was trialled, but once again resulted in complex mixtures (Scheme 3.6, A). In analogy to other α-keto acids described by Flores *et al.*[276] the annulation of OAA with oximes (derived from cyclohexanone and cyclohexane carboxaldehyde) was next attempted, however, in these efforts only protected pyruvic acids were formed in tandem decarboxylations (Scheme 3.6, B). Formation of the presumably highly unstable enol acetonide also proved unsuccessful (Scheme 3.6, C).
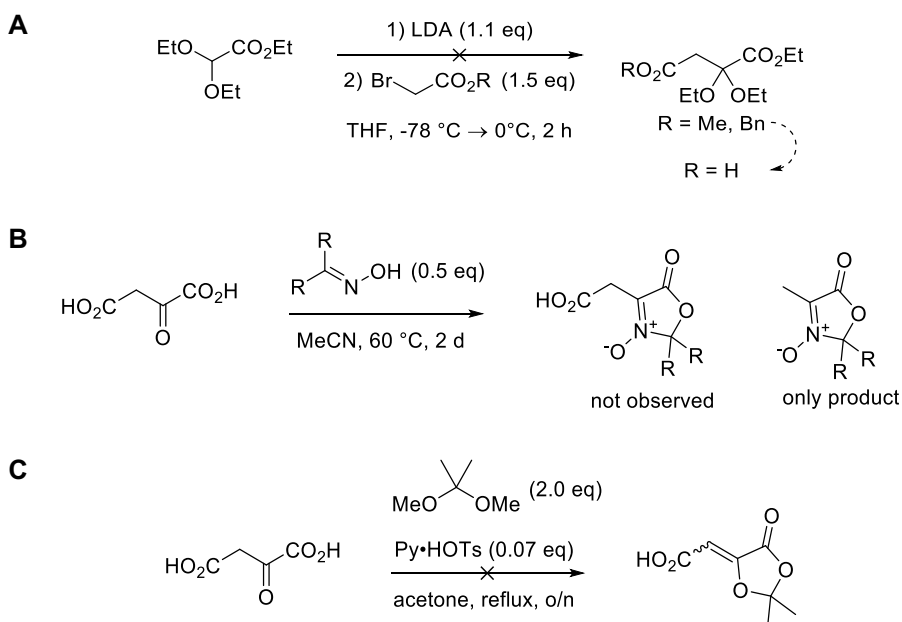
Scheme 3.6 Attempted syntheses of protected OAA derivatives: A) enolate alkylation, B) nitrone formation and C) enol acetonide formation.

As a result of the apparent challenges associated with the synthesis of oxaloacetamides, simplifying the design of the warhead was assumed to be the most viable strategy. Therefore, compound **51a** with succinamide warhead devoid of the α-keto functionality was next pursued.

To access this analogue, the corresponding amine **50a** was reacted with succinic anhydride forming **51a** in good yield. The alternative glutaramide **51b** was also formed in an analogues fashion using glutaric anhydride as the acylating agent. Finally, analogue **51c** was also synthesised in a similar fashion from **50b** and succinic anhydride, however, following the tedious and often low-yielding purifications associated with these highly polar compounds, **51c** was isolated and tested as a crude mixture (Scheme 3.7).
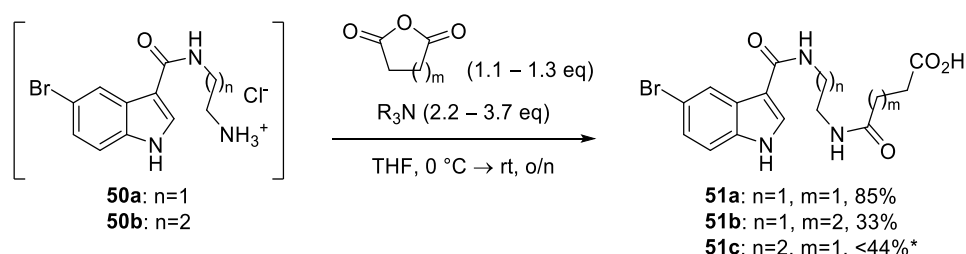
Scheme 3.7 Synthesis of bis-amide-linked fragments **51a-c** with simplified warheads. Either Et$_3$N or DIPEA was used as base (R$_3$N). The structures of **51b,c** were tentatively assigned based on HPLC and LCMS data and analogy to **51a**. *Crude mixture was tested without further purification.

Bis-amides **51a-c** were then tested by both X-ray crystallography for binding and *in vitro* enzyme inhibition assay by A. J. Wijaya to investigate their activity and binding towards PrpC. These results showed that compound **51a** (with a shorter linker) bound to PrpC in X-ray studies (Figure 3.13), and additionally demonstrated inhibitory activity (130 µM). However, analogues **51b,c** (with longer linkers) were found to not bind and also showed no inhibitory activity (**Error! Reference source not found.**). Intermediates **47a,b**, **50a,b** were also tested and found to be inactive in both instances (data not shown).

Table 3.3 *In vitro* biological activity of bis-amide linked fragments **51a-c**.

Compound	Linker-warhead length (number of atoms)	X-ray result	IC$_{50}$
51a	9	good e⁻ density	130 ± 20 µM
51b	10	no e⁻ density	>1 mM
51c	10	no e⁻ density	>1 mM
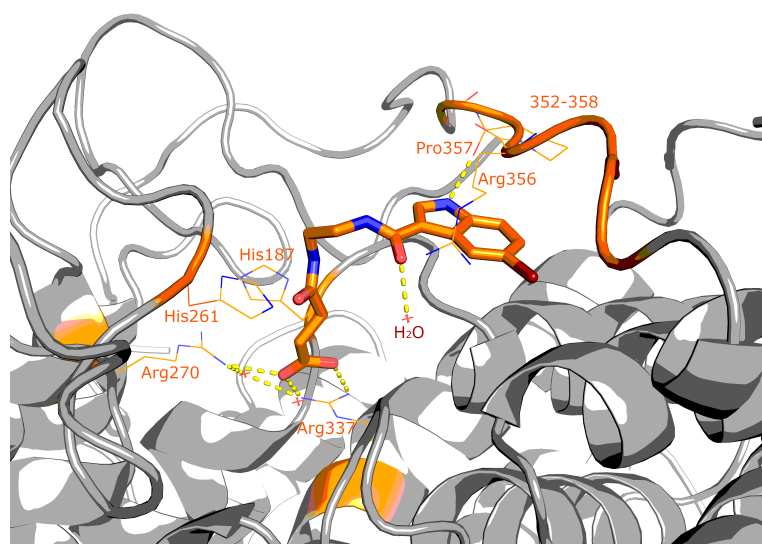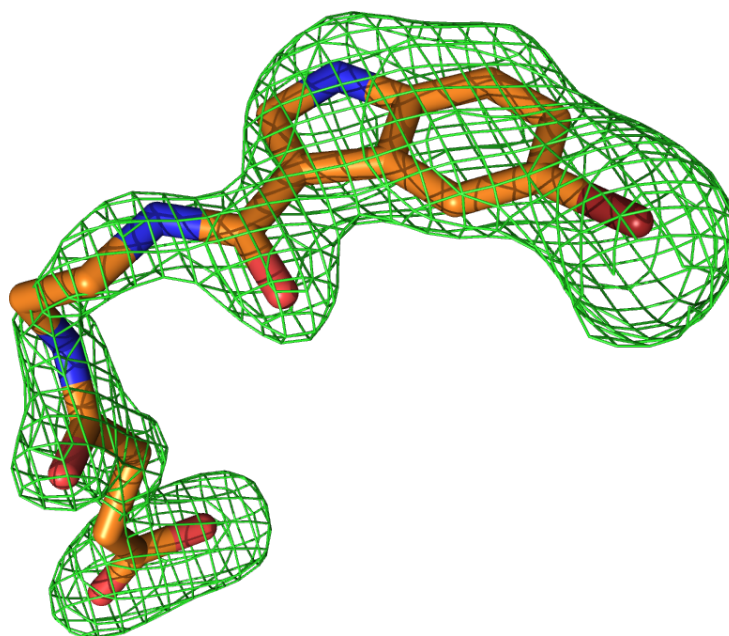
Figure 3.13 A) X-ray structure of PrpC in complex with succinamide **51a**, the key residues are highlighted in orange. B) Fo-Fc map (displated at 3σ) shown in green around **51a** without the ligand being placed.

With these results in hand, compound **51a** was additionally tested by A. J. Wijaya in *P. aeruginosa* growth inhibition assay at a concentration of 1 mM to observe *in vivo* potency. However, no activity was observed against the wild type PAO1 cell line. Next, the efflux pump deficient YM64 cell line was used to inspect if the lack of potency resulted from the quick export of the compound from cytoplasm, but still no activity was detected.

In light of these findings, it was hypothesised that the negatively charged **51a** was unable to penetrate the cell membranes. To mitigate this issue, it was proposed that the neutral ethyl ester prodrug **52** could be in fact readily synthesised and tested in the cellular assay. Accordingly, compound **52** was obtained via a HATU-mediated amide coupling reaction between amine **50a** and the commercially available mono-ethyl succinate in a moderate yield (Scheme 3.8).
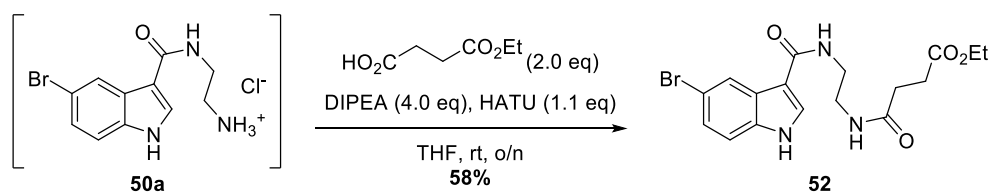
Scheme 3.8 Synthesis of ethyl ester prodrug **52**.

Once again, however, compound **52** was found to have no effect on the growth of either PAO1 or YM64 cells. Whilst cell permeability could still be attributed as the cause, there is no evidence suggesting that the hypothesised hydrolytic release of the active parent compound **51a** took place *in vivo*. Furthermore, another limiting factor was postulated to be the low *in vitro* potency associated with compound **51a**. Accordingly, focus was put on optimising the binding interactions and developing a more potent inhibitor for PrpC.

3.4.2 Studies towards optimising the linker

The X-ray crystallographic data of compound **51a** revealed that the succinamide warhead bound to the OAA pocket, whilst the indole portion bound in a novel mode to the previous indole fragments (Figure 3.14).
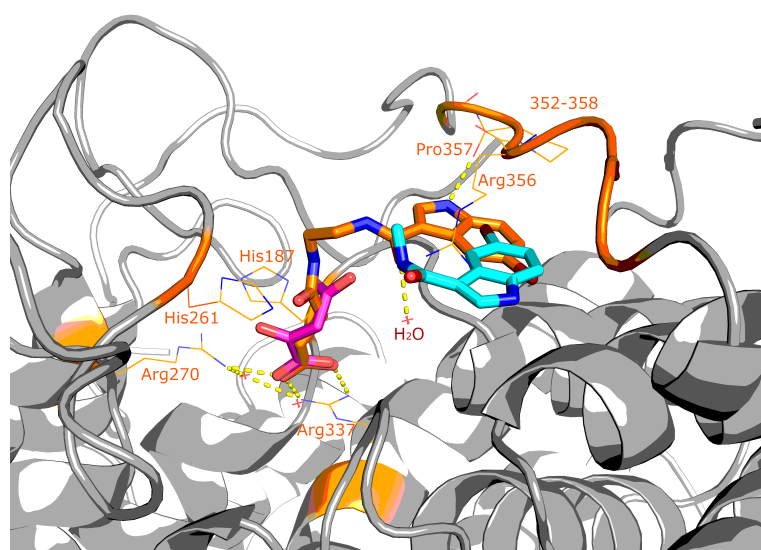
Figure 3.14 X-ray structure of PrpC in complex with succinamide **51a** (orange), overlaid with OAA (magenta) and F-18 (cyan). The key residues are also highlighted in orange.

This data indicated that the ethylenediamine-based bis-amide linker, in fact was not flexible enough to allow for the warhead to reach into the OAA pocket from the original binding mode, instead forcing the indole fragment in a suboptimal position. In order to optimise the binding mode, strategies for introducing more flexible linkers of varied lengths were proposed (Scheme 3.9).

89
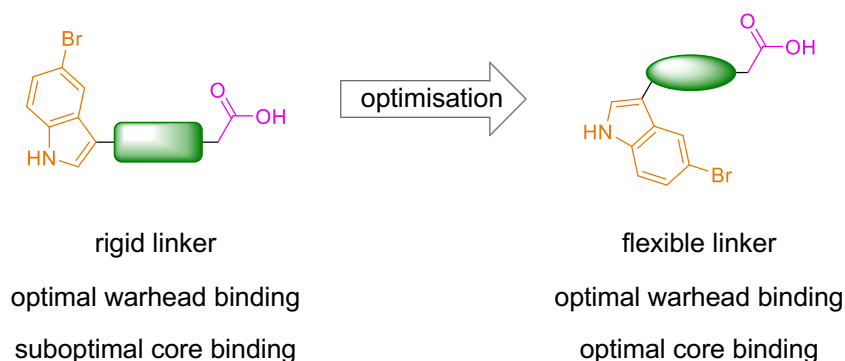
Scheme 3.9 Proposed linker optimisation strategy: a suitable flexible linker would allow for the optimal binding of both the warhead and the indole core fragments.

Following the success of the succinamide warhead strategy utilised previously in the formation of **51a**, it was envisioned that alternative amines could be employed to generate analogues to provide further insight to the effect of linker length on binding as well as identify novel interactions. In the absence of suitable commercially available indole-containing amines, homologues **53a-c** with varied alkyl chain length between the succinamide warhead and the hydrophobic phenyl ring were designed to enable systematic investigations into the ideal linker length. On the other hand, **53d** was designed to incorporate the indole unit—although without the bromine moiety—of **51a** but featured a shorter linker-warhead length of 7 atoms. Following the previously used strategy, compounds **53a-d** were synthesised in a single step from commercially available amines and succinic anhydride (Scheme 3.10). To enable rapid binding readout, the crude products were tested by X-ray crystallography without further purification,[277] and their structures were tentatively assigned based on crude ^{1}H NMR data and analogy to known compounds.
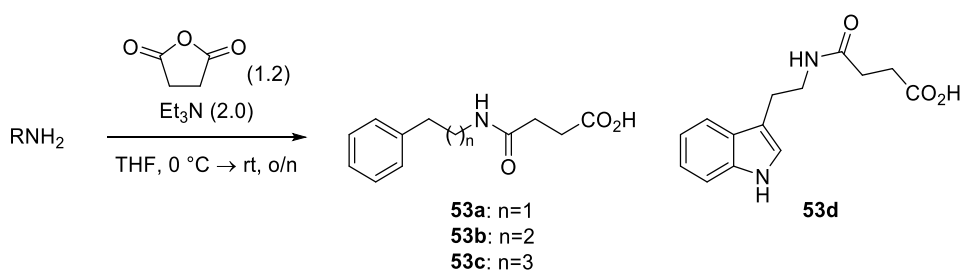
Scheme 3.10 Synthesis of succinamides **53a-d**. Products are tentatively assigned based on crude ^{1}H NMR data, and were tested without further purification.

However, no electron density was observed for any of these fragments, suggesting that the presence of a bromoindole core is crucial. Therefore, synthetic efforts were returned to compounds based on this bromoindole core but having a more flexible linker. Compounds of type **54** with fully saturated ether linkers of variable lengths were hypothesised to be both chemically stable and synthetically tractable as they could be formed by alkylating the commercially available bromotryptophol **55** by suitable haloesters or haloacids (Scheme 3.11).
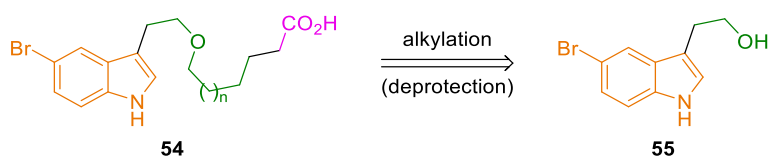
Scheme 3.11 Retrosynthetic analysis for flexible ethers **54**.

To access these compounds, firstly, the *N*-Boc alcohol **56** was synthesised according to conditions described by Liu *et al.*[278] This procedure involved the chemoselective silylation of tryptophol **55** forming the *O*-TBS derivative **57**. Crude **57** was then reacted with Boc$_2$O to form the doubly protected intermediate **58** in an excellent yield over two steps. Finally, removal of the TBS group by TBAF afforded *N*-Boc tryptophol **56** in good yield (Scheme 3.12).
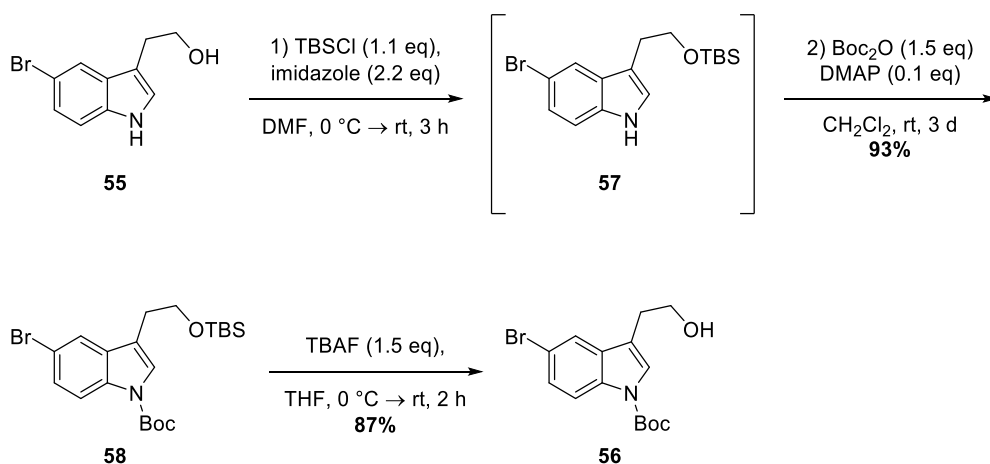
Scheme 3.12 Synthesis of *N*-Boc bromotryptophol **56**.

However, when *N*-Boc alcohol **56** was subjected to strongly basic alkylation conditions, the desired *N*-Boc *tert*-butyl ester **59** was not obtained, instead, the isomer **60** was formed in an intriguing protecting group shift followed by alkylation on the nitrogen (evidenced by 2D NMR analysis, Scheme 3.13). Therefore, the use of weaker bases and higher temperatures was attempted, however, despite our efforts, no product could be observed by TLC analysis under these conditions.
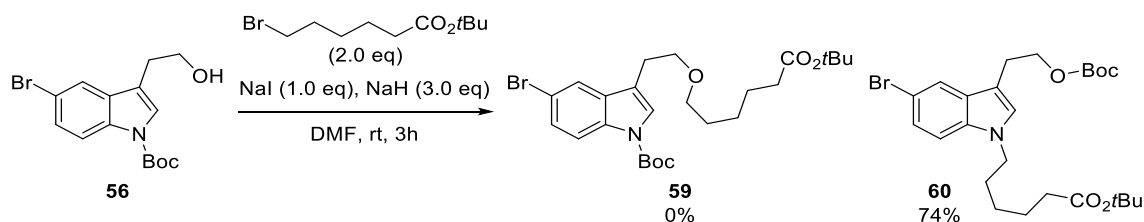
Scheme 3.13 Synthesis of *N*-alkyl *O*-Boc tryptophol **60**.

Following the unsuccessful Boc-protection strategy, it was envisaged that the same approach could be employed successfully by utilising a more suitable protecting group that is stable under basic conditions. Accordingly, the *N*-PMB *O*-TBS tryptophol **61** was obtained in good yield by treating crude intermediate **57** with sodium hydride and PMBCl (Scheme 3.14). This route should

in principle allow for the successful synthesis of analogues with various flexible linkers, like those of type **54**, in order to further explore the binding site and ultimately develop more potent binders for PrpC. However, due to lack of time the work was not continued.
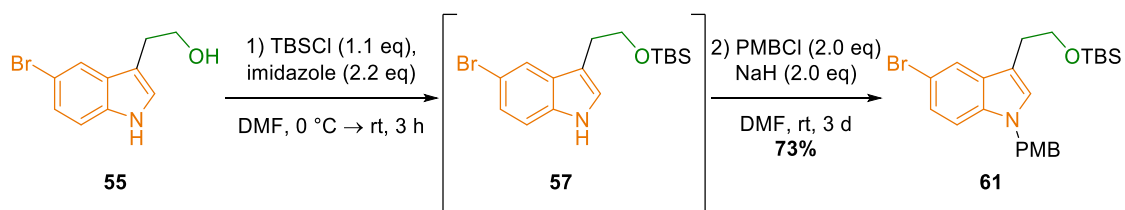
Scheme 3.14 Synthesis of *N*-PMB *O*-TBS tryptophol **61**.

3.5 Conclusions and Future Work

Following the interest in identifying novel binders for *P. aeruginosa* 2-methylcitrate synthase using FBDD, succinamide **51a** was successfully designed, synthesised and confirmed as a novel binder of PrpC with an *in vitro* IC_{50} of 130 μM. Unfortunately, no *in vivo* activity was attained, even when the neutral ester prodrug was trialled, suggesting that higher potency and better cell permeability are key to biological utility. Since X-ray crystallography revealed a suboptimal binding mode of the indole portion, the synthesis of further analogues with more flexible linkers were pursued in order to improve potency. Although the initial Boc-protection strategy proved fruitless, the alternative PMB-protection should provide a functioning route to such analogues.

Future work will focus on further exploring the synthesis of several compounds with flexible ether linkers, including compound **54**. Development of further analogues such as bis-ether linked **62**, sulfonamide **63** and benzimidazolinone **64** could also be considered for their potentially improved water-solubility, polar interactions with the target and chemical stability (Figure 3.15). Once the length and conformation of the linker is optimised, it will be also crucial to rigidify it (reduce the rotatable bond count) by introducing further rings and unsaturated moieties in order to obtain drug-like lead structures.
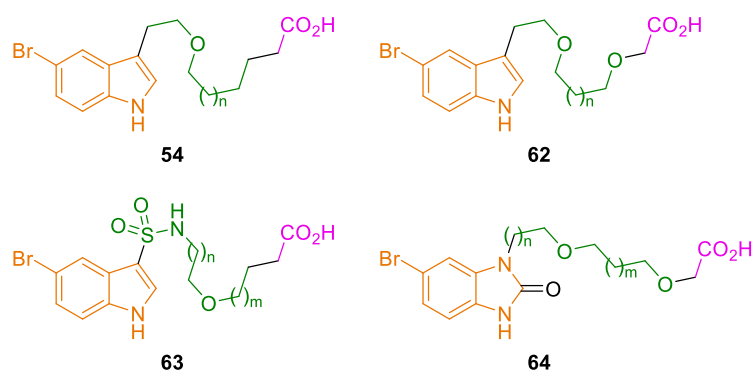
Figure 3.15 Structures of proposed compounds with flexible ether linkers.

Further strategies, such as the incorporation of the α-keto acid warhead could also be explored to investigate how this region of the molecules affect activity. Based on the number of hydrogen-bonding interactions directed at the α-carbonyl group (Figure 3.10, A), the presence of a hydrogen-bond acceptor at this position should increase the potency of the compounds considerably. Possible approaches include the formation of (thio)ether linked pyruvic acids. Alternatively, bioisosteric replacements of the ketone and/or acid functional groups of the warhead could be employed to form chemically and metabolically more stable analogues, such as those shown in Figure 3.16.
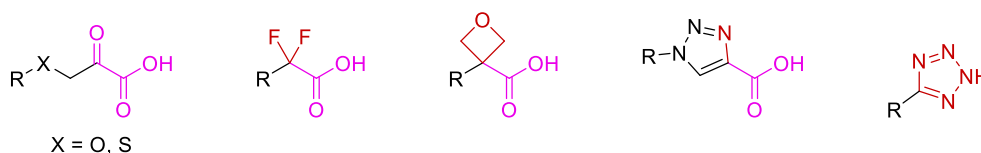
Figure 3.16 Proposed warheads for improving the binding within the OAA pocket. The R group represents an optimised indole fragment attached via a suitable linker to the warhead. The intact α-keto acid unit is highlighted in magenta, bioisosteres are highlighted in red.

This work described the exploration of the *C*-3 indole position. However, SAR studies on the effects of substituents on the *C*-5 and *C*-6 positions of the indole ring could also be conducted. To that end, the bromine can serve as a multifunctional exit vector facilitating Suzuki cross-couplings as well as Sonogashira couplings and substitution to azide allowing for potential azide-alkyne cycloadditions (AAC), thus forming a variety of different (hetero)aromatic groups at this position (Scheme 3.15).
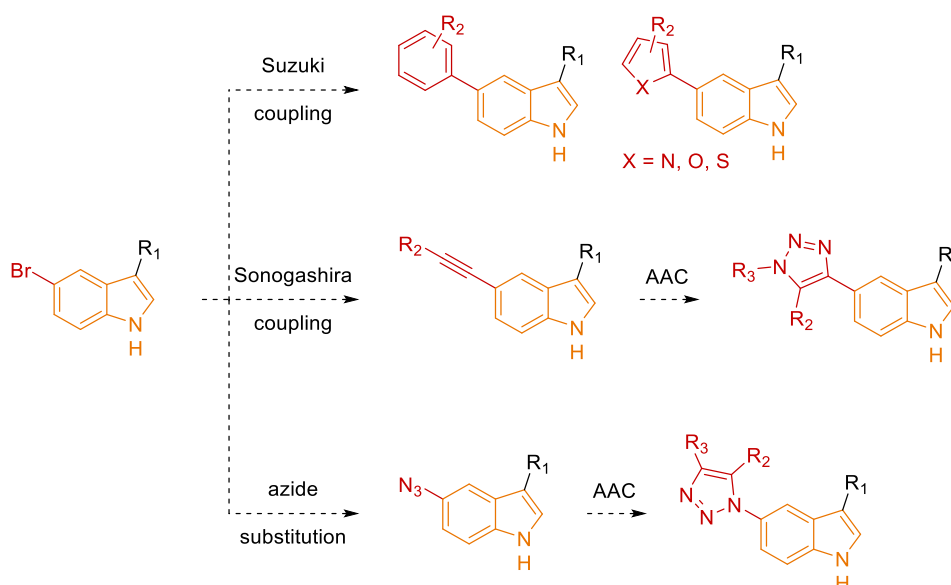
Scheme 3.15 Proposed modifications of the indole ring to further explore the binding pocket. R_1 resembles an appropriate construct of a linker and a warhead, whereas R_2 and R_3 denote general groups introduced.

Finally, given that the target organism is a Gram-negative bacterium, it could be desirable to modify any lead compounds with suitable *in vitro* potency according to the 'eNTRy rules'—stating that compounds with sterically non-hindered basic amines, low three-dimensionality and high rigidity are most likely to accumulate inside Gram-negative bacteria.[279] Assuming that the hypothetical lead compound has a rigidified linker, the 2D and rigid nature are easily attainable, leaving the presence of a negatively charged carboxylate (or bioisostere) instead of a positively charged amine the likely cause for poor cell permeability. Therefore, the introduction of amine functionalities and the use of ester prodrugs to mask the negative charge would be attractive strategies to improve the pharmacokinetic properties of potential lead series (Scheme 3.16).
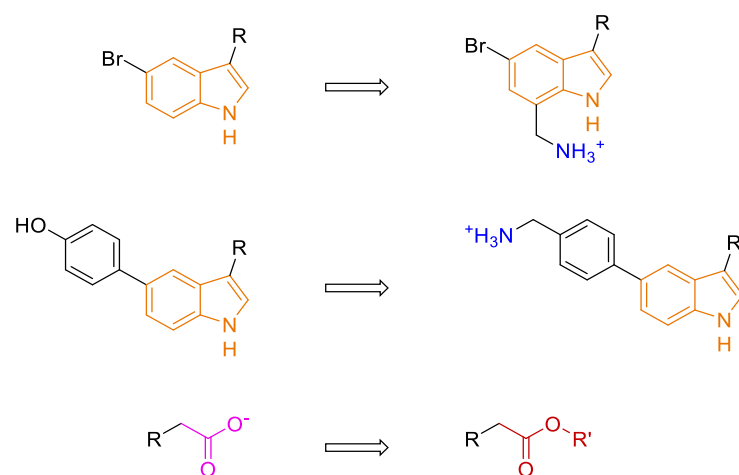
Scheme 3.16 Examples of chemical modifications for increasing the cell permeability of hypothetical lead compounds.

4 Experimental

4.1 General Remarks

All reactions were carried out under argon or nitrogen atmosphere using oven-dried glassware at room temperature unless otherwise stated. Temperatures of -78 °C were maintained using a dry ice acetone bath. Temperatures of 0 °C were maintained using an ice-water bath. Room temperature (rt) refers to ambient temperatures. All reagents were used as received from commercial sources or prepared as described in the literature unless otherwise stated. Acetonitrile (MeCN), dichloromethane (CH_2Cl_2), methanol (MeOH) and toluene were distilled from calcium hydride. Tetrahydrofuran (THF) was dried using sodium wire and distilled from a mixture of calcium hydride and lithium aluminium hydride with triphenylmethane as indicator. Diethyl ether (Et_2O) was distilled from a mixture of calcium hydride and lithium aluminium hydride. Ethyl acetate (EtOAc) was distilled before use; petroleum ether (PE) was distilled before use and refers to the fraction between 40-60 °C. Anhydrous dimethylformamide (DMF), 1,2-dichloroethane (DCE), *tert*-butyl alcohol (*t*BuOH) and pentane were purchased from commercial sources and used without further purification. Reactions were monitored by thin layer chromatography (TLC) using pre-coated Merck glass backed silica gel 60 F_{254} plates and visualised by quenching of UV fluorescence (λ_{Max} = 254 nm) or by staining with potassium permanganate. Retention factors (R_f) are quoted to 0.01. Flash column chromatography was carried out using Merck 9385 Kieselgel 60 SiO_2 (230-400 mesh) under a positive pressure of dry nitrogen. Yields refer to chromatographically and spectroscopically pure compounds unless otherwise stated.

Melting points (m.p.) were obtained using a Büchi Melting Point B-545 or Gallenkamp MPD350. BM2. 5 melting point apparatus and are uncorrected. Optical rotations were measured on an Anton Paar MCP 100 Modular Compact Polarimeter. Infrared (IR) spectra were recorded neat on a Perkin-Elmer Spectrum One spectrometer using an ATR sampling accessory either as solids or liquid films. Selected absorptions (v_{Max}) are reported in wavenumbers (cm^{-1}) with the following abbreviations: w, weak; m, medium; s, strong; br, broad.

Proton magnetic resonance spectra were recorded using an internal deuterium lock at ambient temperatures on Bruker Avance III HD (400 MHz; Smart probe), Bruker Avance III (400 MHz; QNP Cryoprobe) or Bruker Avance III (500 MHz, DUL Cryoprobe) spectrometers. Chemical shifts (δ) are quoted in ppm to the nearest 0.01 ppm and are referenced to the residual non-deuterated solvent peak ($CDCl_3$: 7.26, MeOD: 3.31, DMSO-d_6: 2.50, acetone-d_6: 2.05). Discernible coupling constants (J) are reported as measured values in Hertz, rounded to the nearest 0.1 Hz. Data are reported as: chemical shift, number of nuclei, multiplicity, coupling constant(s) and assignment.

Carbon magnetic resonance spectra were recorded using an internal deuterium lock at ambient temperatures on Bruker Avance III HD (101 MHz), Bruker Avance III (101 MHz) or Bruker Avance

500 (126 MHz) spectrometers with broadband proton decoupling. Chemical shifts (δ) are quoted in ppm to the nearest 0.1 ppm and are referenced to the deuterated solvent peak (CDCl$_3$: 77.16, MeOD: 49.00, DMSO-d$_6$: 39.52, acetone-d$_6$: 29.84). Discernible coupling constants (J) to ^{19}F nuclei are reported as measured values in Hertz, rounded to the nearest 0.1 Hz. Data are reported as: chemical shift, multiplicity (if not singlet), coupling constant(s) (if any) and assignment.

Fluorine magnetic resonance spectra were recorded using an internal deuterium lock at ambient temperatures on Bruker Avance Neo Prodigy (376 MHz, Cryoprobe) spectrometer. Chemical shifts (δ) are quoted in ppm to the nearest 0.1 ppm. Data are reported as: chemical shift, number of nuclei, multiplicity and coupling constant(s).

High resolution mass spectrometry (HRMS) measurements were recorded with a Micromass Q-TOF, Waters Vion IMS Qtof or a Waters LCT Premier TOF mass spectrometer using Electrospray ionisation (ESI) techniques. Mass values are reported within the ±5 ppm error limit.

4.2 Procedures and Analytical Data

4.2.1 Building block synthesis

Ethyl 2-allyl-2-aminopent-4-enoate **1a**
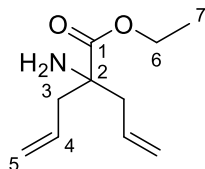
To a solution of **2** (500 mg, 1.87 mmol) in THF (20 mL) at 0 °C was added *t*BuOK (629 mg, 5.61 mmol) and the reaction stirred for 10 min, followed by the dropwise addition of allyl bromide (970 µL, 11.2 mmol) at 0 °C. The reaction mixture was warmed to rt and stirred overnight. Upon completion, HCl (3 M aq, 10 mL) was added and the reaction stirred for 10 min before diluting with H_2O (20 mL). The reaction mixture was extracted with CH_2Cl_2 (3 × 20 mL). The aqueous phase was basified with Na_2CO_3 (pH ≈ 12). The basic aqueous layer was then extracted with EtOAc (3 × 30 mL), and the combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo* to yield the crude product **1a** (229 mg, 1.25 mmol, 67%) as a colourless oil. The crude product **1a** was taken on to the next step without further purification.

R_f = 0.41 (EtOAc). IR (ATR) v_{Max} 3379 (w, N-H), 3078 (w, C-H), 2980 (w, C-H), 1729 (s, C=O), 1640 (m, C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 5.75 – 5.64 (2H, m, H4), 5.16 – 5.10 (4H, m, H5), 4.17 (2H, q, *J* = 7.1 Hz, H6), 2.55 (2H, br dd, *J* = 13.5, 6.5 Hz, H3a), 2.26 (2H, br dd, *J* = 13.5, 8.3 Hz, H3b), 1.67 (2H, br s, NH$_2$), 1.27 (3H, t, *J* = 7.1 Hz, H7) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 176.4 (C1), 132.7 (C4), 119.6 (C5), 61.2 (C6), 60.4 (C2), 44.2 (C3), 14.5 (C7) ppm. HRMS (ESI) calcd for [C$_{10}$H$_{17}$NO$_2$Na]$^+$: 206.1151, found 206.1147.

Ethyl 2-allyl-2-aminohex-5-enoate **1b**
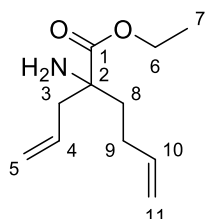
To a solution of **2** (10.35 g, 38.7 mmol) in THF (250 mL) was added *t*BuOK (10.9 g, 96.7 mmol) and 4-bromo-1-butene (11.8 mL, 116 mmol) in three batches over a period of 64 h. Upon completion, the reaction was cooled to 0 °C, *t*BuOK (6.52 g, 58.1 mmol) was added and stirred for 10 min, followed by the dropwise addition of allyl bromide (5.03 mL, 58.1 mmol). The reaction mixture was warmed to rt and stirred for 5 h. A further amount of *t*BuOK (2.17 g, 19.3 mmol) and allyl bromide (1.68 mL, 19.3 mmol) were added and the mixture stirred for 1 h. Upon completion HCl (3 M aq, 50 mL) was added and the reaction stirred for 10 min before removing the organic

solvent *in vacuo*. The aqueous residue was washed with Et$_2$O (3 × 50 mL). The aqueous phase was basified with Na$_2$CO$_3$ (pH ≈ 12). The basic aqueous layer was then extracted with EtOAc (3 × 50 mL), and the combined organic layers were dried over MgSO$_4$, filtered and concentrated *in vacuo* to yield the crude product **1b** (4.67 g, 23.8 mmol, 61%) as a pale orange oil. The crude product **1b** was taken on to the following steps without further purification.

R_f = 0.10 (PE/EtOAc, 4:1). IR (ATR) v_{Max} 3374 (w, N-H), 3077 (w, C-H), 2979 (w, C-H), 2925 (w, C-H), 1726 (s, C=O), 1640 (m, C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 5.83 – 5.61 (2H, m, H4, H10), 5.15 – 5.09 (2H, m, H5), 5.00 (1H, dq, *J* = 17.0, 1.6 Hz, H11$_t$), 4.93 (1H, dq, *J* = 10.1, 1.6 Hz, H11$_c$), 4.16 (2H, q, *J* = 7.1 Hz, H6), 2.55 (1H, br dd, *J* = 13.5, 6.4 Hz, H3a), 2.24 (1H, br dd, *J* = 13.5, 8.5 Hz, H3b), 2.17 – 2.06 (1H, m, H9a), 1.99 – 1.88 (1H, m, H9b), 1.88 – 1.80 (1H, m, H8a), 1.67 – 1.58 (3H, m, H8b, NH$_2$), 1.27 (3H, t, *J* = 7.1 Hz, H7) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 176.7 (C1), 138.0 (C10), 132.8 (C4), 119.6 (C5), 115.0 (C11), 61.1 (C6), 60.5 (C2), 44.5 (C3), 39.2 (C8), 28.5 (C9), 14.4 (C7) ppm. HRMS (ESI) calcd for [C$_{11}$H$_{19}$NO$_2$Na]$^+$: 220.1313, found 220.1309.

Ethyl 2-allyl-2-aminohept-6-enoate **1c**
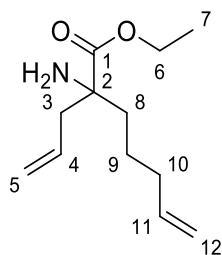
To a solution of **2** (1.00 g, 3.74 mmol) in THF (40 mL) at 0 °C was added *t*BuOK (629 mg, 5.61 mmol) and the reaction stirred for 10 min, followed by the dropwise addition of 5-bromo-1-pentene (1.33 mL, 11.2 mmol). The reaction mixture was warmed to rt and stirred overnight. Upon completion, the reaction was diluted with NH$_4$Cl (sat. aq, 50 mL) and extracted with EtOAc (3 × 50 mL). The combined organic layers were washed with brine (50 mL), dried over MgSO$_4$, filtered and concentrated *in vacuo*. The residue was purified by flash column chromatography (silica gel, PE/EtOAc, 9:1) to give crude **3c** (825 mg) as a colourless oil. To a solution of crude **3c** (550 mg) in THF (20 mL) at 0 °C was added *t*BuOK (276 mg, 2.46 mmol) and the reaction stirred for 10 min, followed by the dropwise addition of allyl bromide (426 μL, 4.92 mmol). The reaction mixture was warmed to rt and stirred overnight. Upon completion, the reaction was diluted with NH$_4$Cl (sat. aq, 25 mL) and extracted with EtOAc (3 × 25 mL). The combined organic layers were washed with brine (25 mL), dried over MgSO$_4$, filtered and concentrated *in vacuo*. The residue was purified by flash column chromatography (silica gel, PE/Et$_2$O, 9:1) to give a crude intermediate (343 mg) as a colourless oil. To a solution of the crude intermediate (300 mg) in THF (8.0 mL) was added HCl (3 M aq, 1.0 mL) and the reaction stirred for 10 min before diluting with H$_2$O (25 mL). The reaction mixture was extracted with CH$_2$Cl$_2$ (3 × 25 mL). The aqueous phase was basified with Na$_2$CO$_3$ (pH ≈ 12). The basic aqueous layer was then extracted with EtOAc (3 × 25 mL), and the combined organic layers were dried over MgSO$_4$, filtered and concentrated *in*

vacuo to yield the crude product **1c** (144 mg, 0.681 mmol, 31%) as a colourless oil. The crude product **1c** was taken on to the next step without further purification.

R_f = 0.10 (PE/EtOAc, 4:1). IR (ATR) ν_{Max} 3376 (w, N-H), 2981 (w, C-H), 2932 (w, C-H), 1728 (s, C=O), 1640 (m, C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 5.80 – 5.60 (2H, m, H4, H11), 5.14 – 5.07 (2H, m, H5), 5.00 – 4.90 (2H, m, H12), 4.15 (2H, q, *J* = 7.2 Hz, H6), 2.53 (1H, br dd, *J* = 13.5, 6.5 Hz, H3a), 2.21 (1H, br dd, *J* = 13.5, 8.4 Hz, H3b), 2.01 (1H, br q, *J* = 7.2 Hz, H10), 1.78 – 1.38 (5H, m, H8, H9a, NH$_2$), 1.28 – 1.15 (4H, m, H9b, H7) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 176.8 (C1), 138.3 (C11), 132.9 (C4), 119.5 (C5), 114.9 (C12), 61.1 (C6), 60.6 (C2), 44.4 (C3), 39.6 (C8), 33.9 (C10), 23.3 (C9), 14.4 (C7) ppm. HRMS (ESI) calcd for [C$_{12}$H$_{22}$NO$_2$]$^+$: 212.1644, found 212.1642.

Ethyl ((diphenylmethylene)amino)acetate **2**
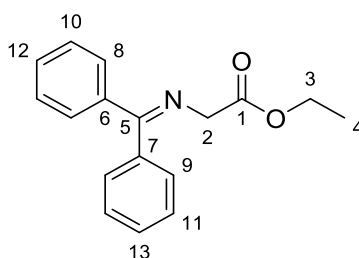
Following a slightly modified version of a reported procedure,[174] ethyl glycinate hydrochloride (33.6 g, 240 mmol), benzophenone (43.8 g, 240 mmol) and DIPEA (42.0 mL, 240 mmol) were added to toluene (150 mL) and the resultant suspension was refluxed for 2 days using Dean-Stark apparatus. The mixture was cooled to room temperature, diluted with EtOAc (300 mL), washed with NaHCO$_3$ (sat. aq, 300 + 4 × 100 mL), brine (100 mL), dried over MgSO$_4$ then filtered through a short silica gel plug eluted with EtOAc and concentrated *in vacuo.* The crude product was recrystallised from Et$_2$O/PE three times to yield pure **2** (10.4 g, 38.9 mmol, 16%) as transparent crystals.

R_f = 0.15 (PE/EtOAc, 9:1). m.p. 52.1 – 52.9 °C(Et$_2$O/PE). IR (ATR) ν_{Max} 3049 (w, C-H), 2980 (w, C-H), 2911 (w, C-H), 1748 (s, C=O), 1619 (m, C=N), 1574 (w, C=C), 1491 (w, C=C), 1476 (w, C=C), 1444 (m, C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.68 – 7.63 (2H, m, H8-13), 7.50 – 7.31 (6H, m, H8-13), 7.21 – 7.16 (2H, m, H8-13), 4.21 (2H, q, *J* = 7.1 Hz, H3), 4.20 (2H, s, H2), 1.27 (3H, t, *J* = 7.1 Hz, H4) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 172.0 (C5), 170.8 (C1), 139.4 (C6/7), 136.2 (C6/7), 130.6 (C12/13), 129.0 (C12/13), 128.9 (C8/9/10/11), 128.8 (C8/9/10/11), 128.2 (C8/9/10/11), 127.8 (C8/9/10/11), 61.0 (C3), 55.9 (C2), 14.3 (C4) ppm. HRMS (ESI) calcd for [C$_{17}$H$_{17}$NO$_2$Na]$^+$: 290.1157, found 290.1170.

These characterization data are in accordance with that previously reported in the literature.[174]

4.2.2 Syntheses of different carbocycles

Ethyl 2-allyl-2-(3-ethoxy-3-oxopropanamido)pent-4-enoate **5a**
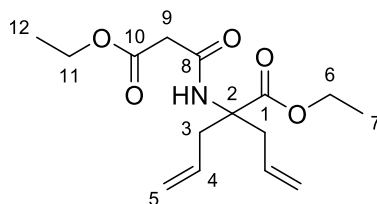
To a solution of **1a** (200 mg, 1.09 mmol) in CH_2Cl_2 (10 mL) was added Et_3N (304 µL, 2.18 mmol) followed by ethyl malonyl chloride (210 µL, 1.64 mmol) at 0 °C and stirred for 20 min. The reaction mixture was diluted with NH_4Cl (sat. aq, 10 mL) and H_2O (5 mL) and stirred for 10 min then extracted with CH_2Cl_2 (4 × 10 mL). The combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel, PE/EtOAc, 4:1) to yield **5a** (264 mg, 0.888 mmol, 81%) as a transparent viscous oil.

R_f = 0.17 (PE/EtOAc, 4:1). IR (ATR) v_{Max} 3310 (w, br, N-H), 3074 (w, C-H), 2981 (w, C-H), 1733 (s, C=O), 1656 (s, C=O and C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.58 (1H, br s, NH), 5.65 − 5.53 (2H, m, H4), 5.09 − 5.02 (4H, m, H5), 4.24 − 4.15 (4H, m, H6, H11), 3.26 (2H, s, H9), 3.15 (2H, br dd, J = 13.9, 7.2 Hz, H3a), 2.52 (2H, br dd, J = 13.9, 7.4 Hz, H3b), 1.29 − 1.24 (6H, m, H7, H12) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 172.6 (C1), 168.8 (C10), 164.0 (C8), 132.2 (C4), 119.1 (C5), 64.4 (C2), 62.0 (C6/11), 61.6 (C6/11), 42.6 (C9), 39.1 (C3), 14.3 (C7/12), 14.1 (C7/12) ppm. HRMS (ESI) calcd for [C$_{15}$H$_{24}$NO$_5$]$^+$: 298.1654, found 298.1644.

Ethyl 2-allyl-2-(3-ethoxy-3-oxopropanamido)hex-5-enoate **5b**
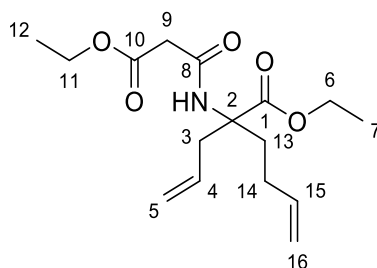
To a solution of **1b** (1.0 g, 5.07 mmol) in CH_2Cl_2 (50 mL) at 0 °C was added Et_3N (1.52 mL, 10.9 mmol), followed by the dropwise addition of ethyl malonyl chloride (1.04 mL, 8.11 mmol) and the reaction stirred for 30 min. The reaction mixture was diluted with NH_4Cl (sat. aq, 25 mL) and stirred for 10 min then extracted with CH_2Cl_2 (3 × 20 mL). The combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel, PE/EtOAc, 4:1) to yield **5b** (1.27 g, 4.27 mmol, 84%) as a pale yellow viscous oil.

R_f = 0.23 (PE/EtOAc, 4:1). IR (ATR) v_{Max} 3337 (w, br, N-H), 2980 (w, C-H), 1732 (s, C=O), 1681 (m, C=O), 1650 (m, C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.67 (1H, br s, NH), 5.79 − 5.67 (1H, m,

H15), 5.65 – 5.51 (1H, m, H4), 5.10 – 5.02 (2H, m, H5), 5.02 – 4.94 (2H, m, H16), 4.27 – 4.18 (4H, m, H6, H11), 3.29 (2H, s, H9), 3.26 – 3.18 (1H, m, H3a), 2.61 – 2.47 (2H, m, H3b, H13a), 2.10 – 1.97 (1H, m, H14a), 1.92 – 1.74 (2H, m, H13b, H14b), 1.32 – 1.26 (6H, m, H7, H12), ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 173.2 (C1), 168.9 (C10), 163.9 (C8), 137.5 (C15), 132.3 (C4), 119.0 (C5), 115.3 (C16), 64.7 (C2), 62.1 (C6), 61.7 (C11), 42.8 (C9), 39.5 (C3), 34.1 (C13), 28.7 (C14), 14.3 (C7/12), 14.2 (C7/12) ppm. HRMS (ESI) calcd for [C$_{16}$H$_{26}$NO$_5$]$^+$: 312.1811, found 312.1820.

Ethyl 2-allyl-2-(3-ethoxy-3-oxopropanamido)hept-6-enoate **5c**
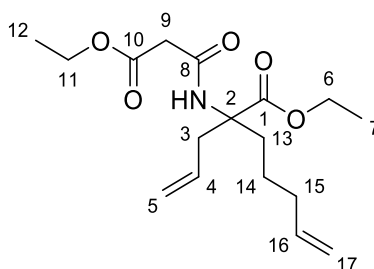
To a solution of **1c** (100 mg, 0.473 mmol) in CH$_2$Cl$_2$ (5.0 mL) at 0 °C was added Et$_3$N (132 µL, 0.946 mmol) followed by ethyl malonyl chloride (91 µL, 0.710 mmol) and the reaction stirred for 20 min. The reaction mixture was diluted with NH$_4$Cl (sat. aq, 10 mL) and H$_2$O (5 mL) and stirred for 10 min then extracted with CH$_2$Cl$_2$ (3 × 10 mL). The combined organic layers were dried over MgSO$_4$, filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel, PE/EtOAc, 4:1) to yield **5c** (113 mg, 0.347 mmol, 73%) as a transparent viscous oil.

R_f = 0.19 (PE/EtOAc, 4:1). IR (ATR) v_{Max} 3326 (w, br, N-H), 3081 (w, C-H), 2982 (w, C-H), 2939 (w, C-H), 1734 (s, C=O), 1682 (s, C=O and C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.63 (1H, br s, NH), 5.78 – 5.66 (1H, m, H16), 5.63 – 5.52 (1H, m, H4), 5.07 – 5.00 (2H, m, H5), 5.00 – 4.90 (2H, m, H17), 4.26 – 4.16 (4H, m, H6, H11), 3.27 (2H, s, H9), 3.18 (1H, br dd, *J* = 14.0, 7.2 Hz, H3a), 2.50 (1H, br dd, *J* = 14.0, 7.5 Hz, H3b), 2.41 (1H, br td, *J* = 13.0, 4.6 Hz, H13a), 2.08 – 1.92 (2H, m, H15), 1.82 – 1.71 (1H, m, H13b), 1.44 – 1.32 (1H, m, H14a), 1.32 – 1.24 (6H, m, H7, H12), 1.14 – 1.01 (1H, m, H14b) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 173.4 (C1), 169.0 (C10), 163.9 (C8), 138.3 (C16), 132.4 (C4), 119.0 (C5), 115.0 (C17), 64.9 (C2), 62.0 (C6/11), 61.7 (C6/11), 42.8 (C9), 39.4 (C3), 34.5 (C13), 33.5 (C15), 23.5 (C14), 14.4 (C7/12), 14.2 (C7/12) ppm. HRMS (ESI) calcd for [C$_{17}$H$_{27}$NO$_5$Na]$^+$: 348.1781, found 348.1770.

5,5-Diallylpyrrolidine-2,4-dione (10a) **7a**
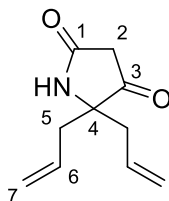
To a solution of **5a** (200 mg, 0.673 mmol) in THF (10 mL) was added *t*BuOK (113 mg, 1.01 mmol) and the reaction heated under reflux for 2 h. The reaction mixture was diluted with EtOAc (20 mL), HCl (3 M aq, 10 mL) and brine (20 mL) and stirred for 10 min. The organic layer was then separated and the aqueous layer was extracted with EtOAc (2 × 20 mL). The combined organic layers were dried over MgSO$_4$, filtered and concentrated *in vacuo*. The residue was dissolved in MeCN/H$_2$O (9:1, 10 mL) and heated under reflux for 1 h, then concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel, PE/EtOAc, 1:1) to yield **7a** (104 mg, 0.580 mmol, 86%) as a white amorphous solid.

R_f = 0.13 (PE/EtOAc, 1:1). IR (ATR) v_{Max} 3212 (w, br, N-H), 2981 (w, C-H), 1768 (m, C=O, ketone), 1698 (s, C=O, amide), 1640 (m, C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 6.64 (H, br s, NH), 5.77 − 5.64 (2H, m, H6), 5.23 − 5.12 (4H, m, H7), 2.89 (2H, s, H2), 2.50 − 2.35 (4H, m, H5) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 209.1 (C3), 170.5 (C1), 130.5 (C6), 121.5 (C7), 71.5 (C4), 41.7 (C2), 41.3 (C5) ppm. HRMS (ESI) calcd for [C$_{10}$H$_{14}$NO$_2$]$^+$: 180.1025, found 180.1021.

5-Allyl-5-(but-3-en-1-yl)pyrrolidine-2,4-dione (10b) **7b**
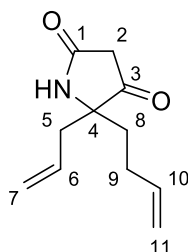
To a solution of **5b** (2.03 g, 6.81 mmol) in THF (70 mL) was added *t*BuOK (1.15 g, 10.2 mmol) and the reaction heated under reflux for 1 h. Upon completion, the reaction was diluted with HCl (3 M aq, 50 mL) and heated under reflux for 30 min. The organic solvent was removed in vacuo and the aqueous layer was extracted with CH$_2$Cl$_2$ (3 × 50 mL). The combined organic layers were dried over MgSO$_4$, filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel, PE/EtOAc, 1:1) to yield **7b** (1.24 g, 6.42 mmol, 94%) as a white amorphous solid.

R_f = 0.10 (PE/EtOAc, 4:1). IR (ATR) v_{Max} 3282 (w, br, N-H), 3081 (w, C-H), 2921 (w, C-H), 2848 (w, C-H), 1639 (s, C=O and C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.90 − 7.62 (1H, m, NH), 5.74 − 5.62 (2H, m, H6, H10), 5.17 (1H, d, J = 10.4 Hz, H7$_c$), 5.13 (1H, d, J = 17.4 Hz, H7$_t$), 5.00 (1H, d, J = 17.4 Hz, H11$_t$), 4.95 (1H, d, J = 10.4 Hz, H11$_c$), 2.91 (1H, d J = 22.3 Hz, H2a), 2.88 (1H, d J = 22.3 Hz, H2b), 2.46 − 2.31 (2H, m, H5), 2.22 − 2.11 (1H, m, H9a), 2.03 − 1.87 (2H, m, H8a, H9b), 1.77 − 1.67

(1H, m, H8b) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 209.9 (C3), 171.6 (C1), 137.0 (C10), 130.5 (C6), 121.3 (C7), 116.1 (C11), 71.6 (C4), 42.8 (C5), 41.8 (C2), 35.8 (C8), 28.4 (C9) ppm. HRMS (ESI) calcd for [C$_{11}$H$_{16}$NO$_2$]$^+$: 194.1181, found 194.1184.

5-Allyl-5-(pent-4-en-1-yl)pyrrolidine-2,4-dione (10c) **7c**
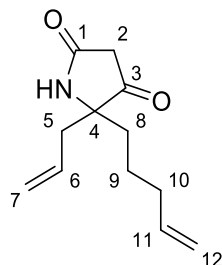
To a solution of **5c** (72 mg, 0.221 mmol) in THF (4.0 mL) was added *t*BuOK (37 mg, 0.332 mmol) and the reaction heated under reflux for 2 h. The reaction mixture was diluted with HCl (3 M aq, 4.0 mL) and brine (10 mL) and stirred for 10 min. The organic layer was then removed and the aqueous layer was extracted with CH$_2$Cl$_2$ (3 × 10 mL). The combined organic layers were dried over MgSO$_4$, filtered and concentrated *in vacuo*. The residue was dissolved in MeCN/H$_2$O (9:1, 4.0 mL) and heated under reflux for 1 h, then concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel, PE/EtOAc, 1:1) to yield **7c** (40 mg, 0.193 mmol, 87%) as a transparent viscous oil.

R_f = 0.22 (PE/EtOAc, 1:1). IR (ATR) v_{Max} 3196 (w, br, N-H), 2943 (w, C-H), 1641 (s, C=O and C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.68 – 7.58 (1H, m, NH), 5.77 – 5.62 (2H, m, H6, H11), 5.19 – 5.09 (2H, m, H7), 5.01 – 4.93 (2H, m, H12), 2.91 (1H, d, *J* = 22.4 Hz, H2a), 2.89 (1H, d, *J* = 22.4 Hz, H2b), 2.46 – 2.31 (2H, m, H5), 2.06 – 1.98 (2H, m, H10), 1.82 – 1.72 (1H, m, H8a), 1.65 – 1.55 (1H, m, H8b), 1.55 – 1.43 (1H, m, H9a), 1.27 – 1.14 (1H, m, H9b) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 210.0 (C3), 171.4 (C1), 137.6 (C11), 130.7 (C6), 121.2 (C7), 115.6 (C12), 71.9 (C4), 41.8 (C5), 41.7 (C2), 36.4 (C8), 33.6 (C10), 23.0 (C9) ppm.

HRMS (ESI) calcd for [C$_{12}$H$_{17}$NO$_2$Na]$^+$: 230.1152, found 230.1146.

1-Azaspiro[4.4]non-7-ene-2,4-dione **8a**
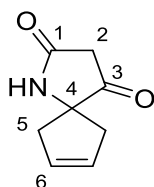
A solution of **7a** (95 mg, 0.530 mmol) in toluene (25 mL) was degassed with argon and heated to 70 °C, followed by the addition of Grubbs II catalyst (17 mg, 20 μmol) and the reaction stirred for 90 min, then concentrated *in vacuo*. The residue was purified by flash column chromatography (silica gel, CH$_2$Cl$_2$ then EtOAc) to yield **8a** (55.0 mg, 0.364 mmol, 69%) as a brown amorphous solid.

R_f = 0.23 (EtOAc). m.p. 154 – 155 °C (Et$_2$O). IR (ATR) v_{Max} 3180 (w, br, N-H), 3083 (w, C-H), 2949 (w, C-H), 2846 (w, C-H), 1764, (s, C=O, ketone), 1702 (s, C=O, amide), 1673 (s, C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 6.89 (1H, bs, NH), 5.69 (2H, s, H6), 3.08 (2H, s, H2), 2.97 (2H, d, J = 15.9 Hz, H5a), 2.53 (2H, d, J = 15.9 Hz, H5b) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 210.1 (C3), 169.9 (C1), 127.6 (C6), 73.4 (C4), 45.5 (C5), 40.5 (C2) ppm. HRMS (ESI) calcd for [C$_8$H$_{10}$NO$_2$]$^+$: 152.0712, found 152.0710.

1-Azaspiro[4.5]dec-7-ene-2,4-dione **8b**
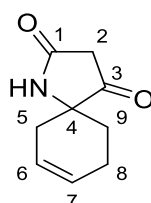
A solution of **7b** (1.12 g, 5.78 mmol) in CH$_2$Cl$_2$ (250 mL) was degassed with argon, followed by the addition of Grubbs II catalyst (49.1 mg, 58 µmol) and the reaction heated under reflux for 1 h, then concentrated *in vacuo*. The residue was purified by flash column chromatography (silica gel, CH$_2$Cl$_2$ then EtOAc) to yield **8b** (950 mg, 5.75 mmol, 99%) as a pale brown amorphous solid.

R_f = 0.23 (EtOAc). IR (ATR) v_{Max} 3177 (w, br, N-H), 3033 (w, C-H), 2921 (w, C-H), 2845 (w, C-H), 1650 (s, C=O and C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.65 (1H, bs, NH), 5.85 – 5.79 (1H, m, H7), 5.72 – 5.66 (1H, m, H6), 3.12 (1H, d, J = 22.1 Hz, H2a), 3.06 (1H, d, J = 22.1 Hz, H2b), 2.46 – 2.31 (2H, m, H5), 2.22 – 2.11 (1H, m, H9a), 2.03 – 1.87 (2H, m, H8a, H9b), 1.77 – 1.67 (1H, m, H8b) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 209.5 (C3), 170.7 (C1), 126.8 (C7), 123.1 (C6), 66.3 (C4), 40.3 (C2), 33.5 (C5), 29.4 (C9), 21.4 (C8) ppm. HRMS (ESI) calcd for [C$_9$H$_{12}$NO$_2$]$^+$: 166.0868, found 166.0869.

1-Azaspiro[4.6]undec-7-ene-2,4-dione **8c**
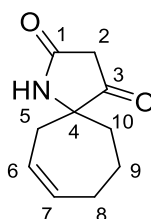

A solution of **7c** (38 mg, 0.183 mmol) in CH$_2$Cl$_2$ (10 mL) was degassed with argon, followed by the addition of Grubbs II catalyst (15 mg, 18 µmol) and the reaction heated under reflux for 30 min, then concentrated *in vacuo*. The residue was purified by flash column chromatography (silica gel, CH$_2$Cl$_2$ then PE/EtOAc, 1:1) to yield **8c** (25.5 mg, 0.142 mmol, 78%) as a pale brown amorphous solid.

R_f = 0.15 (PE/EtOAc, 1:1). IR (ATR) v_{Max} 3199 (w, br, N-H), 2929 (w, C-H), 2836 (w, C-H), 1630 (s, C=O and C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 6.68 (1H, b s, NH), 6.11 – 6.03 (1H, m, H7), 5.67 – 5.58 (1H, m, H6), 3.08 (2H, s, H2), 2.66 – 2.59 (1H, m, H5a), 2.35 – 2.25 (1H, m, H8a), 2.22 – 2.11

(2H, m, H5b, H8b), 2.06 – 1.95 (1H, m, H10a), 1.95 – 1.85 (1H, m, H10b), 1.85 – 1.74 (1H, m, H9a), 1.41 – 1.29 (1H, m, H9b) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 209.0 (C3), 169.5 (C1), 136.2 (C7), 125.2 (C6), 67.4 (C4), 39.7 (C2), 39.0 (C10), 34.5 (C5), 28.1 (C8), 20.7 (C9) ppm. HRMS (ESI) calcd for [C$_{10}$H$_{13}$NO$_2$Na]$^+$: 202.0839, found 202.0832.

4.2.3 Synthesis of enantiopure (*R*)-8b

Methyl (*R*)-2-allyl-2-(3-ethoxy-3-oxopropanamido)hex-5-enoate **(*R*)-5d**
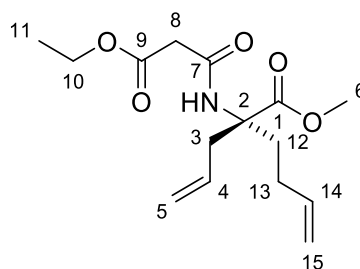
To a solution of **(*R*)-1d** (140 mg, 0.764 mmol) in CH$_2$Cl$_2$ (6.0 mL) at 0 °C was added a solution of Et$_3$N (213 µL, 1.53 mmol) in CH$_2$Cl$_2$ (1.0 mL), followed by the dropwise addition of a solution of ethyl malonyl chloride (147 µL, 1.15 mmol) in CH$_2$Cl$_2$ (1.0 mL) and the reaction stirred for 30 min. The reaction mixture was diluted with NH$_4$Cl (sat. aq, 10 mL) and stirred for 10 min then extracted with CH$_2$Cl$_2$ (3 × 10 mL). The combined organic layers were dried over MgSO$_4$, filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel, PE/EtOAc, 4:1) to yield **(*R*)-5d** (105 mg, 0.353 mmol, 46%) as a transparent viscous oil.

R_f = 0.33 (PE/EtOAc, 2:1). IR (ATR) v_{Max} 3327 (w, br, N-H), 3080 (w, C-H), 2980 (w, C-H), 1735 (s, C=O), 1656 (s, C=O and C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.65 (1H, br s, NH), 5.78 – 5.67 (1H, m, H14), 5.65 – 5.53 (1H, m, H4), 5.10 – 5.03 (2H, m, H5), 4.98 (1H, dq, *J* = 17.2, 1.5 Hz, H15$_t$), 4.93 (1H, br d, *J* = 10.1 Hz, H15$_c$), 4.22 (2H, q, *J* = 7.2 Hz, H10), 3.77 (3H, s, H6), 3.29 (2H, s, H8), 3.20 (1H, br dd, *J* = 13.9, 7.3 Hz, H3a), 2.59 – 2.48 (2H, m, H3b, H12a), 2.09 – 1.99 (1H, m, H13b), 1.93 – 1.75 (2H, m, H12b, H13b), 1.29 (3H, t, *J* = 7.2 Hz, H11) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 173.7 (C1), 169.0 (C9), 164.0 (C7), 137.4 (C14), 132.3 (C4), 119.1 (C5), 115.3 (C15), 64.8 (C2), 61.8 (C10), 52.9 (C6), 42.7 (C8), 39.5 (C3), 34.1 (C12), 28.7 (C13), 14.2 (C11) ppm. HRMS (ESI) calcd for [C$_{15}$H$_{23}$NO$_5$Na]$^+$: 320.1474, found 320.1473. [α]$_D^{20}$ +13.3° (c = 0.120, MeOH).

(*R*)-5-Allyl-5-(but-3-en-1-yl)pyrrolidine-2,4-dione **(*R*)-7b**
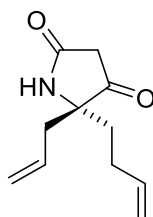
To a solution of **(*R*)-5d** (95 mg, 0.319 mmol) in THF (2.5 mL) was added a solution of *t*BuOK (54 mg, 0.479 mmol) in THF (1.0 mL), and the reaction heated under reflux for 2 h. The reaction

mixture was diluted with EtOAc (10 mL) and HCl (1 M aq, 10 mL) and stirred for 10 min. The organic layer was then separated and the aqueous layer was extracted with EtOAc (3 × 10 mL). The combined organic layers were dried over MgSO$_4$, filtered and concentrated *in vacuo* to give a colourless oil (78 mg). A solution of the oil (70 mg) in MeCN/H$_2$O (9:1, 3.5 mL) was heated under reflux for 1 h, then concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel, PE/EtOAc, 1:1) to yield **(R)-7b** (51 mg, 0.264 mmol, 92%) as a transparent viscous oil.

Analytical data matched that of **7b**. $[\alpha]_D^{20}$ +111.9° (c = 0.176, MeOH).

(*R*)-1-Azaspiro[4.5]dec-7-ene-2,4-dione **(R)-8b**
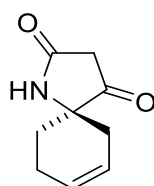
A solution of **(R)-7b** (39 mg, 0.202 mmol) in CH$_2$Cl$_2$ (10 mL) was degassed with argon, followed by the addition of Grubbs II catalyst (17 mg, 20 µmol) and the reaction heated under reflux for 1 h, then concentrated *in vacuo*. The residue was purified by flash column chromatography (silica gel, CH$_2$Cl$_2$ then EtOAc) to yield **(R)-8b** (32.5 mg, 0.197 mmol, 97%) as a brown amorphous solid.

Analytical data matched that of **8b**. $[\alpha]_D^{20}$ +45.4° (c = 0.410, MeOH).

4.2.4 Syntheses of different core heterocycles

2-Chloro-*N*-(4-(hydroxymethyl)octa-1,7-dien-4-yl)acetamide **13**
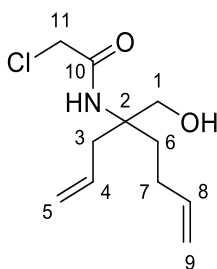
To a solution of **16** (71 mg, 0.278 mmol) in THF (6.0 mL) was added hydrochloric acid (conc. aq, 1.0 mL) and heated under reflux for 3 h, then concentrated *in vacuo*. To a solution of the residue in CH$_2$Cl$_2$ (3.0 mL) was added Et$_3$N (116 µL, 0.834 mmol) followed by chloroacetyl chloride (21.6 µL, 0.278 mmol) at 0 °C and stirred for 90 minutes. The reaction mixture was diluted with NH$_4$Cl (sat. aq, 10 mL) and stirred for 10 minutes then extracted with CH$_2$Cl$_2$ (3 × 10 mL). The combined organic layers were dried over MgSO$_4$, filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel, CH$_2$Cl$_2$ to CH$_2$Cl$_2$/Et$_2$O, 9:1) to yield **13** (37.0 mg, 0.160 mmol, 57%) as a colourless oil.

R_f = 0.38 (PE/EtOAc, 1:1). IR (ATR) v_{Max} 3285 (m, br, O-H and N-H), 3078 (w, C-H), 2936 (w, C-H), 1656 (s, C=O), 1640 (w, C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 6.73 (1H, m, NH), 5.86 – 5.74 (2H, m, H4, H8), 5.24 –5.17 (2H, m, H5), 5.05 (1H, dq, J = 17.1, 1.6 Hz, H9$_t$), 4.97 (1H, dq, J = 10.3, 1.2 Hz, H9$_c$), 4.33 (1H, t, J = 6.3 Hz, OH), 4.02 (2H, s, H11), 3.75 – 3.65 (2H, m, H1), 2.46 (1H, br dd, J = 14.0, 6.8 Hz, H3a), 2.33 (1H, br dd, J = 13.9, 8.2 Hz, H3b), 2.20 – 1.99 (2H, m, H7), 1.91 – 1.79 (1H, m, H6a), 1.74 – 1.65 (1H, m, H6b) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 166.6 (C10), 137.8 (C8), 132.1 (C4), 120.5 (C5), 115.3 (C9), 67.1 (C1), 61.0 (C2), 43.1 (C11), 39.1 (C3), 33.3 (C6), 27.8 (C7) ppm. HRMS (ESI) calcd for [C$_{11}$H$_{18}$NO$_2$NaCl]$^+$: 254.0918, found 254.0914.

Ethyl 2-allyl-2-(2-chloroacetamido)hex-5-enoate **14**
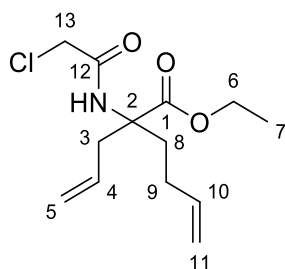
To a solution of **1b** (197 mg, 1.00 mmol) in CH$_2$Cl$_2$ (10 mL) was added Et$_3$N (279 μL, 2.00 mmol) followed by chloroacetyl chloride (117 μL, 1.50 mmol) at 0 °C and stirred for 20 minutes. The reaction mixture was diluted with NH$_4$Cl (sat. aq, 10 mL) and water (5 mL) and stirred for 10 minutes then extracted with CH$_2$Cl$_2$ (4 × 10 mL). The combined organic layers were dried over MgSO$_4$, filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel, PE/EtOAc, 9:1) to yield **14** (241 mg, 0.880 mmol, 88%) as a pale yellow oil.

R_f = 0.21 (PE/EtOAc, 9:1). IR (ATR) v_{Max} 3378 (m, br, N-H), 3079 (w, C-H), 2980 (w, C-H), 1731 (s, C=O, ester), 1677 (s, C=O, amide), 1641 (w, C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 5.54 (1H, m, NH), 5.78 – 5.67 (1H, m, H10), 5.61 – 5.50 (1H, m, H4), 5.10 –5.04 (2H, m, H5), 4.97 (1H, dq, J = 17.2, 1.6 Hz, H11$_t$), 4.94 (1H, dq, J = 10.2, 1.3 Hz, H11$_c$), 4.29 – 4.22 (2H, m, H6), 3.99 (2H, s, H13), 3.23 (1H, br dd, J = 14.1, 7.3 Hz, H3a), 2.63 – 2.49 (2H, m, H3b, H8a), 2.07 – 1.95 (1H, m, H9a), 1.95 – 1.76 (2H, m, H8b, H9b), 1.31 (3H, t, J = 7.2 Hz, H7) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 173.1 (C1), 165.0 (C12), 137.2 (C10), 131.9 (C4), 119.4 (C5), 115.5 (C11), 64.7 (C2), 62.3 (C6), 42.9 (C13), 39.5 (C3), 33.9 (C8), 28.7 (C9), 14.3 (C7) ppm. HRMS (ESI) calcd for [C$_{13}$H$_{21}$NO$_3$Cl]$^+$: 274.1210, found 274.1205.

Ethyl 2-allyl-2-((*tert*-butoxycarbonyl)amino)hex-5-enoate **15**
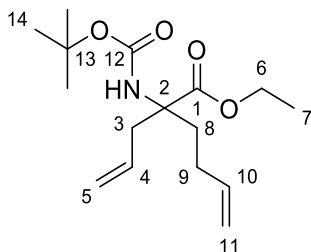
To a solution of **1b** (1.00 g, 5.07 mmol) in THF (35 mL) was added Boc₂O (1.66 g, 7.60 mmol) and the reaction heated to 50 °C in a sealed tube overnight. The reaction mixture was concentrated *in vacuo* and the residue purified by flash column chromatography (silica gel, CH₂Cl₂) to yield **15** (1.28 g, 4.31 mmol, 85%) as a transparent viscous oil.

R_f = 0.37 (PE/EtOAc, 9:1). IR (ATR) v_{Max} 3426 (w, br, N-H), 3080 (w, C-H), 2979 (w, C-H), 1714 (s, C=O), 1641 (w, C=C) cm⁻¹. ¹H NMR (400 MHz, CDCl₃) δ 5.79 – 5.67 (1H, m, H10), 5.67 – 5.53 (1H, m, H4), 5.49 (1H, br s, NH), 5.08 – 5.01 (2H, m, H5), 4.96 (1H, dq, *J* = 17.1, 1.5 Hz, H11ₜ), 4.91 (1H, d, *J* = 10.1 Hz, H11_c), 4.18 (2H, q, *J* = 7.1 Hz, H6), 3.04 (1H, br s, H3a), 2.46 (1H, dd, *J* = 13.9, 7.4 Hz, H3b), 2.42 – 2.28 (1H, m, H8a), 2.11 – 1.96 (1H, m, H9a), 1.90 – 1.72 (2H, m, H8b, H9b), 1.41 (9H, s, H14), 1.26 (3H, t, *J* = 7.1 Hz, H7) ppm. ¹³C NMR (101 MHz, CDCl₃) δ 173.4 (C1), 153.9 (C12), 137.7 (C10), 132.6 (C4), 118.9 (C5), 115.1 (C11), 79.2 (C13), 63.3 (C2), 61.8 (C6), 40.0 (C3), 34.6 (C8), 28.6 (C9), 28.5 (C14), 14.4 (C7) ppm. HRMS (ESI) calcd for [C₁₆H₂₇NO₄Na]⁺: 320.1832, found 320.1822.

tert-Butyl (4-(hydroxymethyl)octa-1,7-dien-4-yl)carbamate **16**
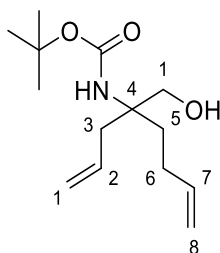
To a solution of **15** (1.00 g, 3.36 mmol) in THF (35 mL) 0 °C was added lithium borohydride (366 mg, 16.8 mmol) and the reaction mixture warmed to room temperature and stirred overnight. A further portion of lithium borohydride (366 mg, 16.8 mmol) was added and stirred for 24 h, followed by heating to 45 °C for 7 h. The reaction mixture was cooled and diluted with NH₄Cl (sat. aq, 20 mL) and water (5 mL) and stirred for 10 minutes. The reaction mixture was extracted with EtOAc (4 × 20 mL) and the combined organic layers washed with NaHCO₃ (sat. aq, 25 mL), brine (25 mL), dried over MgSO₄, filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel, PE/EtOAc, 4:1) to yield **16** (223 mg, 0.873 mmol, 26%) as a white solid.

R_f = 0.20 (PE/EtOAc, 4:1). IR (ATR) v_{Max} 3286 (m, br, O-H and N-H), 3076 (w, C-H), 2979 (w, C-H), 2920 (w, C-H), 2867 (w, C-H), 1678 (s, C=O), 1642 (w, C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 5.85 – 5.74 (2H, m, H4, H8), 5.18 – 5.12 (2H, m, H5), 5.03 (1H, dq, J = 17.1, 1.6 Hz, H9$_t$), 4.96 (1H, dq, J = 10.1, 1.5 Hz, H9$_c$), 4.64 (1H, br s, NH), 4.16 (1H, br s, OH), 3.72 – 3.61 (2H, m, H1), 2.40 – 2.28 (2H, m, H3), 2.17 – 1.97 (2H, m, H7), 1.76 – 1.56 (2H, m, H6), 1.43 (9H, s, H12) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 156.4 (C10), 138.3 (C8), 132.9 (C4), 119.6 (C5), 115.0 (C9), 80.1 (C11), 68.0 (C1), 58.9 (C2), 39.2 (C3), 33.8 (C6), 28.5 (C12), 27.8 (C7) ppm. HRMS (ESI) calcd for [C$_{14}$H$_{25}$NO$_3$Na]$^+$: 278.1727, found 278.1717.

5-Allyl-5-(but-3-en-1-yl)morpholin-3-one **17**
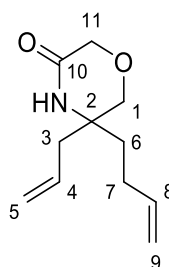
A solution of **13** (35 mg, 0.151 mmol) in *tert*-butanol (3.0 mL) was heated to 30 °C, followed by the addition of potassium *tert*-butoxide (19 mg, 0.166 mmol) and the reaction stirred for 6 h. The reaction mixture was diluted with NH$_4$Cl (sat. aq, 10 mL), stirred for 10 minutes and then extracted with CH$_2$Cl$_2$ (3 × 10 mL). The combined organic layers were dried over MgSO$_4$, filtered and concentrated *in vacuo* to yield **17** (29 mg, 0.149 mmol, 98%) as a pale yellow oil. The crude product **17** was taken on to the following step without further purification.

R_f = 0.26 (PE/EtOAc, 1:1). IR (ATR) v_{Max} 3211 (m, br, N-H), 3074 (w, C-H), 2976 (w, C-H), 2920 (w, C-H), 2867 (w, C-H), 1667 (s, C=O), 1640 (w, C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 6.82 (1H, m, NH), 5.83 – 5.68 (2H, m, H4, H8), 5.22 –5.13 (2H, m, H5), 5.04 (1H, dq, J = 17.1, 1.6 Hz, H9$_t$), 4.98 (1H, dq, J = 10.3, 1.2 Hz, H9$_c$), 4.10 (2H, s, H11), 3.62 (1H, d, J = 11.9 Hz, H1a), 3.58 (1H, d, J = 11.9 Hz, H1b), 2.38 – 2.26 (2H, m, H3), 2.15 – 2.07 (2H, m, H7), 1.68 – 1.60 (2H, m, H6) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 169.2 (C10), 137.6 (C8), 131.5 (C4), 120.5 (C5), 115.4 (C9), 70.8 (C1), 67.6 (C11), 56.3 (C2), 41.4 (C3), 36.1 (C6), 27.7 (C7) ppm. HRMS (ESI) calcd for [C$_{11}$H$_{18}$NO$_2$]$^+$: 196.1338, found 196.1338.

4-Oxa-1-azaspiro[5.5]undec-8-en-2-one **18**
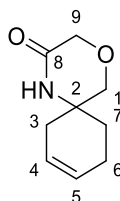
A solution of **17** (26.0 mg, 0.133 mmol) in CH$_2$Cl$_2$ (10 mL) was degassed with argon, followed by the addition of Grubbs II catalyst (11.3 mg, 13 μmol) and the reaction heated under reflux for 30

minutes, then concentrated *in vacuo*. The residue was purified by flash column chromatography (silica gel, PE/EtOAc, 1:1) to yield **18** (14.0 mg, 83.7 µmol, 63%) as a white amorphous solid.

R_f = 0.12 (PE/EtOAc, 1:1). IR (ATR) v_{Max} 3165 (m, br, N-H), 3072 (w, C-H), 2920 (w, C-H), 1664 (s, C=O), 1641 (w, C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 5.76 – 5.70 (1H, m, H5), 5.64 – 5.57 (1H, m, H4), 4.20 (1H, d, *J* = 11.1 Hz, H1a), 4.14 (1H, d, *J* = 11.1 Hz, H1b), 3.69 (2H, s, H9), 2.21 – 1.95 (4H, m, H3, H6), 1.82 – 1.74 (1H, m, H7a), 1.69 – 1.61 (1H, m, H7b) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 169.1 (C8), 126.8 (C5), 123.4 (C4), 76.3 (C1), 48.3 (C2), 44.1 (C9), 33.3 (C3), 28.6 (C7), 22.1 (C6) ppm. HRMS (ESI) calcd for [C$_9$H$_{13}$NO$_2$Na]$^+$: 190.0839, found 190.0833.

4-Allyl-4-(but-3-en-1-yl)oxazolidin-2-one **19** and *tert*-butyl 2-allyl-2-(but-3-en-1-yl)aziridine-1-carboxylate **20**

19   **20**
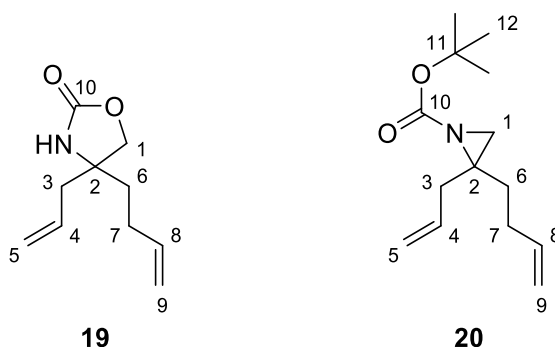
To a solution of **16** (128 mg, 0.500 mmol) in Et$_2$O (10 mL) was added *p*-toluenesulfonyl chloride (114 mg, 0.600 mmol) followed by potassium hydroxide (168 mg, 3.00 mmol), then the reaction was heated under reflux for 20 h. The reaction mixture was stirred for further 2 days at room temperature, then diluted with NaHCO$_3$ (sat. aq, 10 mL) and stirred for 10 minutes. The layers were then separated, and the aqueous layer was extracted with CH$_2$Cl$_2$ (3 × 10 mL). The combined organic layers were dried over MgSO$_4$, filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel, CH$_2$Cl$_2$ to CH$_2$Cl$_2$/EtOAc, 4:1) to yield **19** (43.5 mg, 0.240 mmol, 48%) and **20** (33.7 mg, 0.142 mmol, 28%) both as colourless oils.

Analytical data for **19**:

R_f = 0.29 (PE/EtOAc, 4:1). IR (ATR) v_{Max} 3255 (w, br, N-H), 3080 (w, C-H), 2978 (w, C-H), 2920 (w, C-H), 1738 (s, C=O), 1641 (w, C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 6.70 (1H, br s, NH), 5.83 – 5.68 (2H, m, H4, H8), 5.22 – 5.14 (2H, m, H5), 5.04 (1H, dq, *J* = 17.1, 1.6 Hz, H9$_t$), 4.98 (1H, dq, *J* = 10.1, 1.6 Hz, H9$_c$), 4.14 (1H, d, *J* = 8.7 Hz, H1a), 4.08 (1H, d, *J* = 8.7 Hz, H1b), 2.40 – 2.27 (2H, m, H3), 2.15 – 2.07 (2H, m, H7), 1.75 – 1.60 (2H, m, H6) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 159.7 (C10), 137.3 (C8), 131.4 (C4), 120.5 (C5), 115.6 (C9), 73.5 (C1), 60.0 (C2), 43.1 (C3), 37.6 (C6), 27.9 (C7) ppm. HRMS (ESI) calcd for [C$_{10}$H$_{15}$NO$_2$Na]$^+$: 204.0995, found 204.0989.

Analytical data for **20**:

R_f = 0.84 (PE/EtOAc, 4:1). IR (ATR) v_{Max} 3365 (w, br, N-H), 3080 (w, C-H), 2977 (w, C-H), 2932 (w, C-H), 1693 (s, C=O), 1641 (w, C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 5.88 – 5.74 (2H, m, H4, H8),

5.17 – 5.09 (2H, m, H5), 5.02 (1H, dq, J = 17.1, 1.6 Hz, H9$_t$), 4.96 (1H, dq, J = 10.2, 1.4 Hz, H9$_c$), 2.37 (1H, dd, J = 14.7, 7.0 Hz, H3a), 2.27 – 2.04 (5H, m, H1, H3b, H7), 1.70 – 1.61 (1H, m, H6a), 1.51 – 1.42 (10H, m, H6b, H12) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 161.1 (C10), 137.9 (C8), 133.9 (C4), 118.2 (C5), 115.0 (C9), 80.9 (C11), 45,5 (C2), 38.4 (C3), 36.0 (C1), 33.2 (C6), 29.7 (C7), 28.2 (C12) ppm. HRMS (ESI) calcd for $[C_{14}H_{23}NO_2Na]^+$: 260.1621, found 260.1617.

3-Oxa-1-azaspiro[4.5]dec-7-ene-2-one **21**
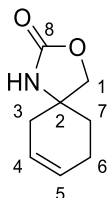
Initial RCM route from **19**:

A solution of **19** (41 mg, 0.226 mmol) in CH$_2$Cl$_2$ (15 mL) was degassed with argon, followed by the addition of Grubbs II catalyst (19 mg, 23 μmol) and the reaction heated under reflux for 30 minutes, then concentrated *in vacuo*. The residue was purified by flash column chromatography (silica gel, CH$_2$Cl$_2$ to CH$_2$Cl$_2$/EtOAc, 9:1) to yield **21** (31.1 mg, 0.203 mmol, 90%) as an off-white solid.

Improved route from cyclohexene **24**:

To a solution of crude **24** (3.28 g, 14.4 mmol) in THF (150 mL) was added *t*BuOK (1.62g, 14.4 mmol) and the reaction stirred for 1 h. The reaction mixture was diluted with NaHCO$_3$ (sat. aq, 150 mL), stirred for 10 min and then extracted with EtOAc (3 × 100 mL). The combined organic layers were washed with brine (100 mL), dried over MgSO$_4$, filtered and concentrated *in vacuo* to yield crude **21** (2.05 g, 13.4 mmol, 93% yield) as a white amorphous solid. The crude product was further purified by recrystallization from Et$_2$O/pentane 1:1 to yield pure **21** (718 mg, 4.69 mmol, 33%) as a white crystalline solid.

Analytical data for **21** (recrystallised product obtained from the improved route):

R_f = 0.21 (PE/EtOAc, 1:1). m.p. 84 – 85 °C (Et$_2$O/Pentane). IR (ATR) v_{Max} 3235 (m, br, N-H), 3039 (w, C-H), 2922 (w, C-H), 2904 (w, C-H), 2845 (w, C-H), 1731 (s, C=O and C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 5.77 – 5.70 (1H, m, H5), 5.66 – 5.59 (1H, m, H4), 5.46 (1H, br s, NH), 4.14 (1H, d, J = 8.5 Hz, H1a), 4.11 (1H, d, J = 8.5 Hz, H1b), 2.34 – 2.15 (4H, m, H3, H6), 1.90 – 1.82 (1H, m, H7a), 1.80 – 1.72 (1H, m, H7b) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 159.4 (C8), 127.1 (C5), 123.6 (C4), 75.6 (C1), 56.2 (C2), 36.9 (C3), 32.3 (C7), 22.7 (C6) ppm. HRMS (ESI) calcd for $[C_8H_{11}NO_2Na]^+$: 176.0682, found 176.0676.

Ethyl 1-((*tert*-butoxycarbonyl)amino)cyclohex-3-ene-1-carboxylate **23**
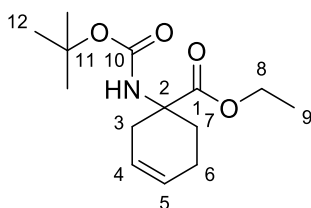
A solution of crude **15** (6.25 g, 21.0 mmol) in CH_2Cl_2 (300 mL) was degassed with argon, followed by the addition of Grubbs II catalyst (18 mg, 21 µmol). The reaction was heated under reflux for 1 h followed by the addition of another portion of Grubbs II catalyst (18 mg, 21 µmol) and the reaction was heated under reflux for further 1 h before being concentrated *in vacuo*. The residue was purified by flash column chromatography (silica gel, PE/EtOAc, 9:1) to yield **23** (3.86 g, 14.3 mmol, 69%) as a transparent viscous oil.

R_f = 0.12 (PE/EtOAc, 9:1). IR (ATR) ν_{Max} 3368 (m, br, N-H), 2977 (w, C-H), 1706 (s, C=O and C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 5.76 – 5.70 (1H, m, H5), 5.61 – 5.55 (1H, m, H4), 4.78 (1H, br s, NH), 4.26 – 4.13 (2H, m, H8), 2.62 – 2.53 (1H, m, H3a), 2.29 – 2.01 (4H, m, H3b, H6, H7a), 1.95 – 1.86 (1H, m, H7b), 1.43 (9H, s, H12), 1.26 (3H, t, J = 7.1 Hz, H7) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 174.2 (C1), 155.0 (C10), 127.2 (C5), 122.6 (C4), 79.9 (C11), 61.2 (C8), 57.0 (C2), 34.2 (C3), 28.4 (C12), 27.7 (C7), 21.9 (C6), 14.3 (C9) ppm. HRMS (ESI) calcd for [C$_{14}$H$_{24}$NO$_4$]$^+$: 270.1705, found 270.1718.

tert-Butyl (1-(hydroxymethyl)cyclohex-3-en-1-yl)carbamate **24**
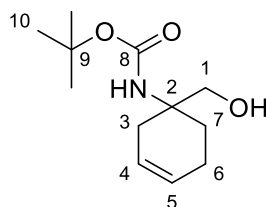
To a solution of **23** (3.86 g, 14.4 mmol) in THF (150 mL) was added LiBH$_4$ (2 M in THF, 14.4 mL, 28.8 mmol), and the reaction stirred overnight. The reaction mixture was diluted with NH$_4$Cl (sat. aq, 150 mL), stirred for 10 min and then extracted with EtOAc (3 × 100 mL). The combined organic layers were washed with NaHCO$_3$ (sat. aq, 100 mL), brine (100 mL), dried over MgSO$_4$, filtered and concentrated *in vacuo* to yield the crude product **24** (3.28 g, 14.4 mmol, 100%) as a white amorphous solid. The crude product **24** was taken on to the following steps without further purification.

R_f = 0.26 (PE/EtOAc, 4:1). IR (ATR) ν_{Max} 3265 (m, br, O-H and N-H), 3076 (w, C-H), 3020 (w, C-H), 2968 (w, C-H), 2933 (w, C-H), 1676 (s, C=O and C=C) cm^{-1}. ^{1}H NMR (400 MHz, DMSO-d_6) δ 5.59 (1H, br d, J = 10.0 Hz, H5), 5.50 (1H, br d, J = 10.0 Hz, H4), 4.62 (1H, t, J = 5.6 Hz, OH), 4.41 (1H, br s, NH), 3.45 – 3.36 (2H, m, H1), 2.27 – 1.86 (5H, m, H3, H6, H7a), 1.57 – 1.47 (1H, m, H7b), 1.36 (9H, s, H10) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 156.7 (C8), 127.6 (C5), 123.3 (C4), 80.1 (C9), 69.3

(C1), 55.2 (C2), 34.1 (C3), 28.5 (C10), 27.4 (C7), 22.2 (C6) ppm. HRMS (ESI) calcd for $[C_{12}H_{22}NO_3]^+$: 228.1600, found 228.1595.

(1-((*tert*-Butoxycarbonyl)amino)cyclohex-3-en-1-yl)methyl 4-methylbenzenesulfonate **25**
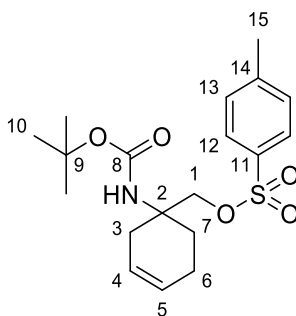
To a solution of **24** (58.0 mg, 0.255 mmol) in Et_2O (4.0 mL) was added *p*-toluenesulfonyl chloride (174 mg, 0.918 mmol) followed by potassium hydroxide (129 mg, 2.30 mmol), and the reaction then heated under reflux for 10 h. The reaction mixture was diluted with water (10 mL), stirred for 10 minutes and then extracted with EtOAc (3 × 10 mL). The combined organic layers were washed with brine (10 mL), dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel, PE/EtOAc, 9:1) to yield **25** (16.0 mg, 41.9 µmol, 16%) as a white solid.

R_f = 0.26 (PE/EtOAc, 4:1). IR (ATR) v_{Max} 3369 (m, N-H), 3030 (w, C-H), 2973 (w, C-H), 2929 (w, C-H), 1697 (s, C=O and C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.77 (2H, d, *J* = 8.2 Hz, H12), 7.32 (2H, d, *J* = 8.2 Hz, H13), 5.71 – 5.64 (1H, m, H5), 5.53 – 5.47 (1H, m, H4), 4.42 (1H, br s, NH), 4.25 (1H, d, *J* = 9.3 Hz, H1a), 4.12 (1H, d, *J* = 9.3 Hz, H1b), 2.43 (3H, s, H15), 2.31 – 1.97 (5H, m, H3, H6, H7a), 1.63 – 1.53 (1H, m, H7b), 1.37 (9H, s, H10) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 154.6 (C8), 144.9 (C11), 133.0 (C14), 130.0 (C13), 128.1 (C12), 127.1 (C5), 122.8 (C4), 71.8 (C1), 52.8 (C2), 33.5 (C3), 28.4 (C10), 27.0 (C7), 21.8 (C15), 21.7 (C6) ppm. HRMS (ESI) calcd for $[C_{19}H_{27}NO_5NaS]^+$: 404.1502, found 404.1494.

4.2.5   Double bond modification

3-Oxa-1-azaspiro[4.5]decan-2-one **26**
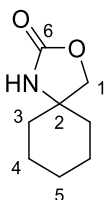
To a solution of pure **21** (15.3 mg, 0.100 mmol) in MeOH (1.0 mL) was added Pd (10 w/w % on charcoal, 10.6 mg, 10 µmol), degassed using hydrogen gas and the reaction was stirred under an atmosphere of hydrogen for 8 h at rt. The reaction mixture was filtered through Celite and concentrated *in vacuo* to yield crude **26** (14.6 mg, 94 µmol, 94%) as a white amorphous solid.

R_f = 0.21 (PE/EtOAc, 1:1). ^{1}H NMR (400 MHz, CDCl$_3$) δ 6.59 (1H, br s, NH), 4.09 (2H, s, H1), 1.80 – 1.30 (10H, m, H3-5) ppm.

These characterization data are in accordance with that previously reported in the literature.[280]

(1R*,3R*,6S*)-7-Oxaspiro[bicyclo[4.1.0]heptane-3,4'-oxazolidin]-2'-one **27a** and (1R*,3S*,6S*)-7-oxaspiro[bicyclo[4.1.0]heptane-3,4'-oxazolidin]-2'-one **27b**

27a   **27b**
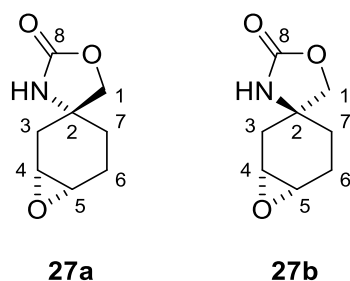
To a solution of **21** (30.6 mg, 0.200 mmol) in CH$_2$Cl$_2$ (2.0 mL) was added mCPBA (69.0 mg, 0.400 mmol) and NaHCO$_3$ (50.4 mg, 0.600 mmol) and the reaction stirred at rt overnight. The reaction mixture was quenched by a mixture of NaHCO$_3$ (sat. aq, 8.0 mL) and Na$_2$SO$_3$ (sat. aq, 2.0 mL), stirred for 10 min then extracted with CH$_2$Cl$_2$ (3 × 10 mL). The combined organic layers were washed with the same aqueous mixture as above (2 × 10 mL), NaCl (sat. aq, 10 mL), dried over MgSO$_4$, filtered and concentrated *in vacuo*. The residue was purified by flash column chromatography (silica gel, 1 to 2% MeOH in CH$_2$Cl$_2$) to yield **27b** (1.9 mg, 11 µmol, 6%) and crude **27a**, both as a white amorphous solids. Crude **27a** was purified by flash column chromatography (silica gel, PE/EtOAc, 1:1 to 0:1) to yield **27a** (15.5 mg, 92 µmol, 46%) as a white solid. **27a** was crystallised from Et$_2$O for the single crystal X-ray crystallography analysis.

Analytical data for **27a**:

R_f = 0.25 (5% MeOH in CH$_2$Cl$_2$). m.p. 110 – 111 °C (crystal decomposition), 116 – 117 °C (Et$_2$O). IR (ATR) v_{Max} 3299 (m, N-H), 3008 (w, C-H), 2910 (w, C-H), 1734 (s, C=O) cm^{-1}. ^{1}H NMR (500 MHz, CDCl$_3$) δ 5.74 (1H, br s, NH), 4.03 (1H, d, J = 9.3 Hz, H1a), 4.02 (1H, d, J = 9.3 Hz, H1b), 3.29 (1H, br s, H4), 3.23 – 3.20 (1H, m, H5), 2.29 (1H, br d, J = 15.0 Hz, H3a), 2.14 – 2.09 (2H, m, H6), 1.94 (1H, br d, J = 15.0 Hz, H3b), 1.78 – 1.72 (1H, m, H7a), 1.45 – 1.37 (1H, m, H7b) ppm. ^{13}C NMR (126 MHz, CDCl$_3$) δ 158.0 (C8), 75.9 (C1), 55.6 (C2), 53.1 (C4), 51.0 (C5), 35.5 (C3), 30.7 (C7), 20.7 (C6) ppm. HRMS (ESI) calcd for [C$_8$H$_{12}$NO$_3$]$^+$: 170.0812, found 170.0810.

Analytical data for **27b**:

R_f = 0.27 (5% MeOH in CH$_2$Cl$_2$). IR (ATR) v_{Max} 3289 (m, N-H), 3000 (w, C-H), 2919 (w, C-H), 2851 (w, C-H), 1737 (s, C=O), 1708 (s, C=O) cm^{-1}. ^{1}H NMR (500 MHz, CDCl$_3$) δ 6.01 (1H, br s, NH), 4.14 (1H, d, J = 9.1 Hz, H1a), 4.09 (1H, d, J = 9.1 Hz, H1b), 3.26 – 3.23 (1H, m, H4), 3.20 – 3.17 (1H, m, H5), 2.22 (1H, br d, J = 15.3 Hz, H3a), 2.18 – 2.02 (3H, m, H3b, H6), 1.69 (1H, dtd, J = 13.1, 6.6, 0.9 Hz, H7a), 1.53 (1H, br dt, J = 13.8, 6.9 Hz, H7b) ppm. ^{13}C NMR (126 MHz, CDCl$_3$) δ 158.9 (C8),

75.1 (C1), 55.6 (C2), 51.8 (C4), 50.7 (C5), 36.6 (C3), 30.1 (C7), 21.1 (C6) ppm. HRMS (ESI) calcd for $[C_8H_{12}NO_3]^+$: 170.0812, found 170.0808.

1-(4-Methoxybenzyl)-3-oxa-1-azaspiro[4.5]dec-7-ene-2-one **28**
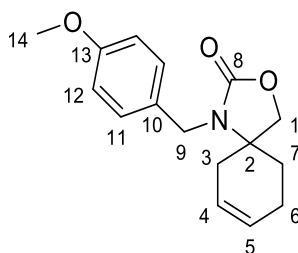
To a solution of pure **21** (718 mg, 4.69 mmol) in DMF (50 mL) was added NaH (60 w/w % dispersion in mineral oil, 281 mg, 7.03 mmol) and the reaction stirred for 1.5 h at 50 °C, followed by the addition of 4-methoxybenzyl chloride (953 μL, 7.03 mmol) and stirred overnight at 50 °C. The reaction mixture was quenched by NaHCO₃ (sat. aq, 50 mL), stirred for 10 min, and then diluted with H_2O (50 mL), brine (100 mL) and extracted with Et₂O (3 × 200 mL). The combined organic layers were concentrated *in vacuo*, the residue dissolved in Et₂O (20 mL), washed with LiCl (10% aq, 3 × 25 mL), dried over MgSO₄, filtered and concentrated *in vacuo*. The residue was purified by flash column chromatography (silica gel, PE/Et₂O, 1:1 to 0:1) to yield **28** (1.23 g, 4.50 mmol, 96%) as a white amorphous solid.

R_f = 0.43 (Et₂O). IR (ATR) v_{Max} 3033 (w, C-H), 2926 (w, C-H), 2902 (w, C-H), 2835 (w, C-H), 1736 (s, C=O), 1613 (w, C=C), 1514 (s, C=C) cm⁻¹. ¹H NMR (400 MHz, CDCl₃) δ 7.28 – 7.22 (2H, m, H11), 6.86 – 6.81 (2H, m, H12), 5.68 – 5.58 (1H, m, H5), 5.58 – 5.48 (1H, m, H4), 4.40 (1H, d, *J* = 15.5 Hz, H9a), 4.30 (1H, d, *J* = 15.5 Hz, H9b), 4.04 (1H, d, *J* = 8.4 Hz, H1a), 4.01 (1H, dd, *J* = 8.4, 0.8 Hz, H1b), 3.79 (3H, s, H14), 2.32 – 2.14 (2H, m, H3a, H6a), 2.14 – 2.00 (1H, m, H6b), 2.00 – 1.90 (1H, m, H3b), 1.86 – 1.75 (1H, m, H7a), 1.58 – 1.50 (1H, m, H7b) ppm. ¹³C NMR (101 MHz, CDCl₃) δ 159.1 (C8), 158.5 (C13), 130.6 (C10), 129.1 (C11), 126.8 (C5), 123.8 (C4), 114.0 (C12), 73.0 (C1), 60.0 (C2), 55.4 (C14), 43.8 (C9), 33.4 (C3), 30.6 (C7), 23.1 (C6) ppm. HRMS (ESI) calcd for $[C_{16}H_{19}NO_3Na]^+$: 296.1257, found 296.1247.

(5*R**,7*R**,8*S**)-7,8-Dihydroxy-1-(4-methoxybenzyl)-3-oxa-1-azaspiro[4.5]decan-2-one **29a** and (5*R**,7*S**,8*R**)-7,8-dihydroxy-1-(4-methoxybenzyl)-3-oxa-1-azaspiro[4.5]-decan-2-one **29b**

29a   29b
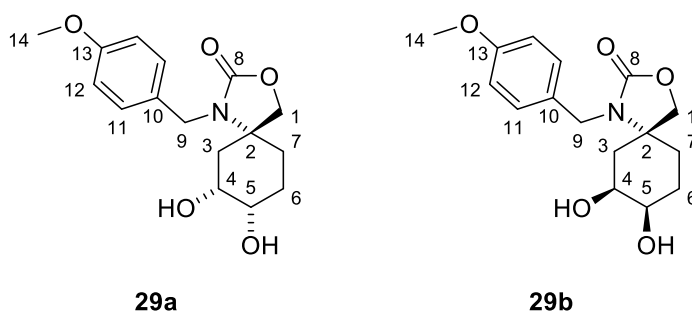
To a solution of **28** (55 mg, 0.20 mmol) in THF (1.0 mL) was added 4-methylmorpholine *N*-oxide (47 mg, 0.40 mmol), citric acid (77 mg, 0.40 mmol), H_2O (1.0 mL) and OsO_4 (2.5 w/w % solution in *t*BuOH, 20 µL, 2.0 µmol) and the reaction stirred for 2 h at rt. The reaction mixture was quenched by Na_2SO_3 (sat. aq, 1.0 mL), stirred for 10 min, and then diluted with brine (1.0 mL) and extracted with EtOAc (3 × 3.0 mL). The combined organic layers were dried over $MgSO_4$, filtered and concentrated *in vacuo* to yield a crude mixture of **29b** and **29a** (63 mg, **29b/29a** = 1:2.5) as a transparent oil. The crude product was purified by flash column chromatography (silica gel, EtOAc to 5% MeOH in Et_2O) to yield **29b** (12.6 mg), **29a** (33.2 mg) and a mixture of **29b** and **29a** (15 mg) all as white solids. Overall yield: **29b** + **29a** (60.8 mg, 0.198 mmol, 99%). **29b** spontaneously crystallised from C_6D_6 and the co-crystals formed were used for the single crystal X-ray crystallography analysis.

Analytical data for **29a**:

R_f = 0.13 (EtOAc). m.p. 120 − 121 °C (Et_2O). IR (ATR) v_{Max} 3422 (w, br, O-H), 3305 (w, br, O-H), 2956 (w, C-H), 2898 (w, C-H), 1717 (s, C=O), 1613 (w, C=C), 1511 (s, C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.25 (2H, dt, *J* = 8.6, 2.5 Hz, H11), 6.83 (2H, dt, *J* = 8.6, 2.5 Hz, H12), .33 (2H, s, H9), 4.07 (1H, d, *J* = 8.6 Hz, H1a), 4.05 (1H, d, *J* = 8.6 Hz, H1b), 3.92 − 3.88 (1H, m, H4), 3.78 (3H, s, H14), 3.56 (1H, br d, *J* = 10.7 Hz, H5), 2.39 (1H, br s, C5OH), 2.31 (1H, br s, C4OH), 2.02 − 1.89 (3H, m, H3a, H6a, H7a), 1.57 (1H, ddd, *J* = 12.2, 4.4, 1.9 Hz, H3b), 1.43 − 1.23 (2H, m, H6b, H7b) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 159.1 (C13), 158.2 (C8), 130.3 (C10), 129.1 (C11), 114.1 (C12), 72.0 (C1), 68.8 (C5), 67.1 (C4), 61.8 (C2), 55.4 (C14), 43.6 (C9), 36.2 (C3), 26.3 (C7), 26.0 (C6) ppm. HRMS (ESI) calcd for [C$_{16}$H$_{22}$NO$_5$]$^+$: 308.1498, found 308.1484.

Analytical data for **29b**:

R_f = 0.20 (EtOAc). m.p. 122 − 123 °C (Et_2O). IR (ATR) v_{Max} 3434 (w, br, O-H), 3356 (w, br, O-H), 2917 (w, C-H), 2851 (w, C-H), 1704 (s, C=O), 1611 (w, C=C), 1511 (s, C=C) cm^{-1}. ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.22 (2H, dt, *J* = 8.7, 2.5 Hz, H11), 6.84 (2H, dt, *J* = 8.7, 2.5 Hz, H12), 4.41 (1H, d, *J* = 15.9 Hz, H9a), 4.33 (1H, dd, *J* = 9.4, 1.2 Hz, H1a), 4.26 (1H, d, *J* = 9.4 Hz, H1b), 4.21 (1H, d, *J* = 15.9 Hz, H9b), 3.99 (1H, br s, H4), 3.79 (3H, s, H14), 3.61 − 3.54 (1H, m, H5), 2.30 (1H, br s, C4OH), 1.91 − 1.82 (2H, m, H3), 1.79 − 1.66 (3H, m, H6, C5OH), 1.66 − 1.52 (2H, m, H7 + H_2O) ppm. ^{13}C NMR (126 MHz, CDCl$_3$) δ 159.1 (C13), 158.6 (C8), 130.5 (C10), 128.8 (C11), 114.2 (C12), 73.4 (C1),

70.5 (C5), 69.5 (C4), 60.4 (C2), 55.4 (C14), 43.5 (C9), 37.9 (C3), 30.9 (C7), 25.1 (C6) ppm. HRMS (ESI) calcd for $[C_{16}H_{22}NO_5]^+$: 308.1498, found 308.1513.

(5*R**,7*S**,8*S**)-7,8-Dibromo-1-(4-methoxybenzyl)-3-oxa-1-azaspiro[4.5]decan-2-one **30a** and (5*R**,7*R**,8*R**)-7,8-dibromo-1-(4-methoxybenzyl)-3-oxa-1-azaspiro[4.5]-decan-2-one **30b**

30a   **30b**
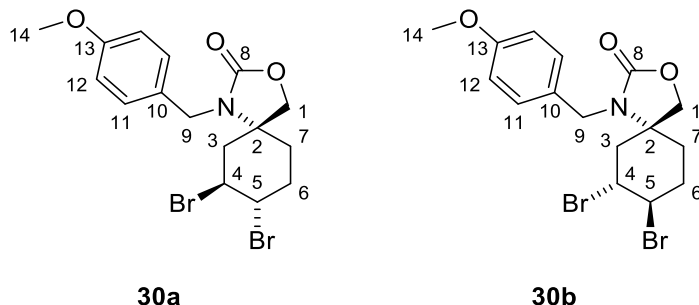
To a solution of **28** (54.7 mg, 0.200 mmol) in CH_2Cl_2 (2.0 mL) at 0 °C was added PTAB (75.2 mg, 0.200 mmol) and the reaction stirred at 0 °C for 2 h, then warmed to rt and stirred overnight. The reaction mixture was filtered and concentrated *in vacuo*, the residue was purified by flash column chromatography (silica gel, PE/Et$_2$O 1:1 to 0:1) to yield **30a** (76.0 mg), a mixture of **30a** and **30b** (3.7 mg, **30a/30b** = 4.9:1) and **30b** (3.4 mg) all as white amorphous solids. Overall yield: **30a** + **30b** (83.1 mg, 0.192 mmol, 96%). **30a** was crystallised from Et$_2$O for the single crystal X-ray crystallography analysis.

Analytical data for **30a**:

R_f = 0.49 (Et$_2$O). m.p. 108 – 109 °C (Et$_2$O). IR (ATR) v_{Max} 2998 (w, C-H), 2960 (w, C-H), 2927 (w, C-H), 2838 (w, C-H), 1737 (s, C=O), 1610 (w, C=C), 1509 (s, C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.28 (2H, br d, *J* = 8.7 Hz, H11), 6.85 (2H, br d, *J* = 8.7 Hz, H12), 4.62 – 4.54 (3H, m, H4, H5, H9a), 4.38 (1H, d, *J* = 9.1 Hz, H1a), 4.33 (1H, dd, *J* = 9.1, 1.4 Hz, H1b), 4.20 (1H, d, *J* = 15.8 Hz, H9b), 3.80 (3H, s, H14), 2.91 (1H, dd, *J* = 15.4, 3.9 Hz, H3a), 2.45 (1H, dddd, *J* = 15.4, 13.2, 3.4, 3.0 Hz, H6a), 2.18 (1H, tdd, *J* = 13.6, 3.9, 1.2 Hz, H7a), 2.04 – 1.96 (1H, m, H6b), 1.89 – 1.82 (1H, m, H3b), 1.64 – 1.57 (1H, m, H7b + H$_2$O) ppm.^{13}C NMR (101 MHz, CDCl$_3$) δ 159.3 (C13), 158.0 (C8), 130.4 (C10), 129.0 (C11), 114.3 (C12), 73.5 (C1), 60.7 (C2), 55.4 (C14), 51.2 (C5), 49.6 (C4), 43.8 (C9), 35.5 (C3), 26.2 (C7), 25.4 (C6) ppm. HRMS (ESI) calcd for $[C_{16}H_{20}NO_3{}^{79}Br_2]^+$: 431.9804, found 431.9800.

Analytical data for **30b**:

R_f = 0.40 (Et$_2$O). m.p. 68 – 69 °C (Et$_2$O). IR (ATR) v_{Max} 2932 (w, C-H), 1737 (s, C=O), 1611 (w, C=C), 1512 (s, C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.23 (2H, d, *J* = 8.7 Hz, H11), 6.86 (2H, d, *J* = 8.7 Hz, H12), 4.48 (1H, d, *J* = 15.8 Hz, H9a), 4.23 (1H, d, *J* = 15.8 Hz, H9b), 4.17 (1H, d, *J* = 8.7 Hz, H1a), 4.10 (1H, d, *J* = 8.7 Hz, H1b), 3.95 – 3.77 (5H, m, H4, H5, H14), 2.47 – 2.33 (2H, m, CH$_2$), 2.29 – 2.13 (1H, m, CH$_2$), 1.86 – 1.72 (1H, m, CH$_2$), 1.64 (1H, td, *J* = 13.6, 3.5 Hz, CH$_2$), 1.60 – 1.52 (1H, m, CH$_2$ + H$_2$O) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 159.4 (C13), 157.6 (C8), 129.7 (C10), 129.0 (C11), 114.4 (C12), 71.0 (C1), 61.7 (C2), 55.5 (C14), 53.6 (C5), 51.4 (C4), 45.0 (C9), 44.0 (C3), 34.5 (C7), 33.5 (C6) ppm. HRMS (ESI) calcd for $[C_{16}H_{20}NO_3{}^{79}Br_2{}^{23}Na]^+$: 453.9624, found 453.9623.

(1*R**,3*S**,6*S**)-3'-(4-Methoxybenzyl)-7-tosyl-7-azaspiro-[bicycle[4.1.0]heptane-3,4'-oxazolidin]-2'-one **31a** and (1*R**,3*R**,6*S**)-3'-(4-methoxybenzyl)-7-tosyl-7-azaspiro[bicyclo[4.1.0]heptane-3,4'-oxazolidin]-2'-one **31b**

31a   **31b**
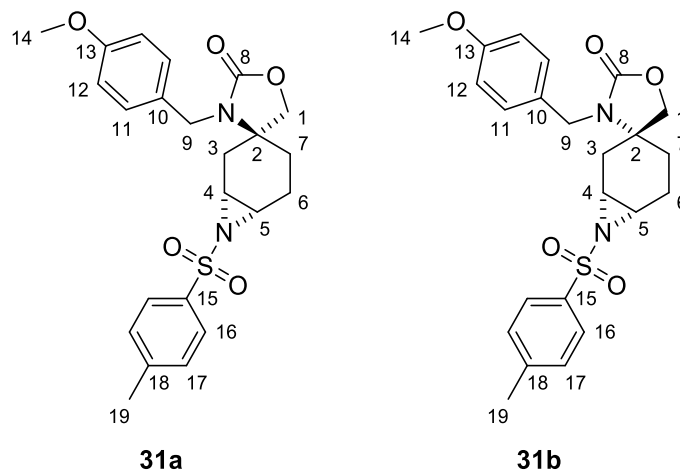
To a solution of **28** (55 mg, 0.20 mmol) in MeCN (1.0 mL) was added chloramine T trihydrate (62 mg, 0.22 mmol) and PTAB (7.5 mg, 20 µmol) and the reaction stirred at rt over 4 Å molecular sieves overnight. The reaction mixture was filtered and concentrated *in vacuo*, the residue was purified by flash column chromatography (silica gel, Et$_2$O) to yield **31a** (34.5 mg, 78 µmol, 39%) and **31b** (25.1 mg, 57 µmol, 28%) both as white amorphous solids. **31b** was crystallised from Et$_2$O for the single crystal X-ray crystallography analysis.

Analytical data for **31a**:

R_f = 0.19 (Et$_2$O). m.p. 177 − 178 °C (Et$_2$O). IR (ATR) v_{Max} 2952 (w, C-H), 2936 (w, C-H), 2921 (w, C-H), 1729 (s, C=O), 1615 (w, C=C), 1513 (s, C=C) cm^{-1}. ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.75 (2H, dt, *J* = 8.3, 1.7 Hz, H16), 7.34 (2H, br d, *J* = 8.3 Hz, H17), 7.19 (2H, dt, *J* = 8.7, 2.5 Hz, H11), 6.83 (2H, dt, *J* = 8.7, 2.5 Hz, H12), 4.37 (1H, d, *J* = 15.8 Hz, H9a), 4.19 (1H, d, *J* = 15.8 Hz, H9b), 4.06 (1H, d, *J* = 9.4 Hz, H1a), 3.96 (1H, dd, *J* = 9.4, 0.7 Hz, H1b), 3.79 (3H, s, H14), 3.02 (1H, ddd, *J* = 6.9, 3.2, 1.7 Hz, H4), 2.95 (1H, t, *J* = 6.7 Hz, H5), 2.46 (3H, s, H19), 2.14 (1H, dtd, *J* = 15.9, 7.1, 1.8 Hz, H6a), 1.93 − 1.77 (3H, m, H3, H6b), 1.50 (1H, td, *J* = 12.7, 7.7 Hz, H7a), 1.39 (1H, ddd, *J* = 13.4, 7.0, 1.8 Hz, H7b) ppm. ^{13}C NMR (126 MHz, CDCl$_3$) δ 159.3 (C13), 158.1 (C8), 145.0 (C15), 135.2 (C18), 130.2 (C10), 130.1 (C17), 128.9 (C11), 127.7 (C16), 114.2 (C12), 72.3 (C1), 60.0 (C2), 55.4 (C14), 43.7 (C9), 40.9 (C4), 36.7 (C5), 32.3 (C3), 28.2 (C7), 21.8 (C19), 20.2 (C6) ppm. HRMS (ESI) calcd for [C$_{23}$H$_{27}$N$_2$O$_5$S]$^+$: 443.1641, found 443.1622.

Analytical data for **31b**:

R_f = 0.12 (Et$_2$O). m.p. 122 − 123 °C (Et$_2$O). IR (ATR) v_{Max} 2962 (w, C-H), 2934 (w, C-H), 1733 (s, C=O), 1615 (w, C=C), 1514 (s, C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.77 (2H, br d, *J* = 8.2 Hz, H16), 7.35 (2H, br d, *J* = 8.2 Hz, H17), 7.18 (2H, dt, *J* = 8.6, 2.4 Hz, H11), 6.78 (2H, dt, *J* = 8.6, 2.4 Hz, H12), 4.34 (1H, d, *J* = 15.5 Hz, H9a), 4.10 (1H, d, *J* = 15.5 Hz, H9b), 4.00 (1H, d, *J* = 8.4 Hz, H1a), 3.87 (1H, dd, *J* = 8.4, 1.4 Hz, H1b), 3.78 (3H, s, H14), 2.97 (1H, br d, *J* = 6.7 Hz, H5), 2.88 (1H, t, *J* =

6.7 Hz, H4), 2.47 (3H, s, H19), 2.10 – 2.02 (1H, m, H6a), 1.92 (1H, ddd, J = 15.2, 6.8, 2.0 Hz, H3a), 1.81 – 1.71 (2H, m, H3b, H7a), 1.67 (1H, ddd, J = 14.5, 4.2, 2.9 Hz, H6b), 1.24 – 1.17 (1H, m, H7b) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 159.2 (C13), 157.9 (C8), 144.9 (C15), 135.1 (C18), 130.0 (C17), 129.9 (C10), 129.3 (C11), 127.8 (C16), 114.1 (C12), 73.0 (C1), 58.8 (C2), 55.4 (C14), 43.7 (C9), 38.4 (C5), 37.4 (C4), 29.5 (C3), 27.2 (C7), 21.8 (C19), 20.4 (C6) ppm. HRMS (ESI) calcd for [C$_{23}$H$_{27}$N$_2$O$_5$S]$^+$: 443.1641, found 443.1629.

(1R^*,3R^*,6S^*)-7,7-Difluoro-3'-(4-methoxybenzyl)spiro-[bicycle[4.1.0]heptane-3,4'-oxazolidin]-2'-one **32a** and (1R^*,3S^*,6S^*)-7,7-difluoro-3'-(4-methoxybenzyl)spiro[bicyclo[4.1.0]heptane-3,4'-oxazolidin]-2'-one **32b**

32a   32b
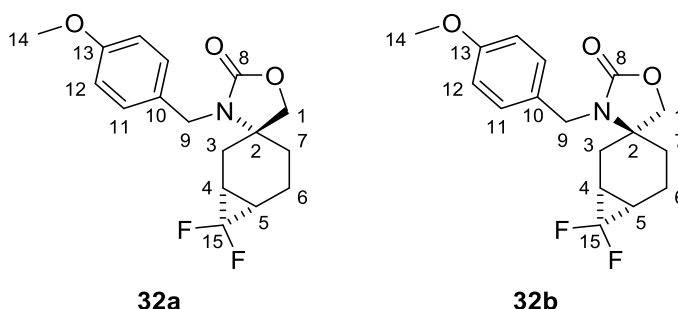
To a solution of **28** (55 mg, 0.20 mmol) in THF (0.30 mL) was added anhydrous NaI (6.0 mg, 40 µmol) and TMSCF$_3$ (74 µL, 0.50 mmol) and the reaction stirred at 65 °C in a sealed tube. After 6 h, the reaction was cooled to rt and opened to air, then more TMSCF$_3$ (74 µL, 0.50 mmol) was added, the tube sealed and heated to 65 °C overnight. The reaction was then cooled to rt again and opened to air followed by the removal of solvent *in vacuo*. The residue was dissolved in Et$_2$O (10 mL) and washed with H$_2$O (10 mL), Na$_2$SO$_3$ (sat. aq, 10 mL), NaHCO$_3$ (sat. aq, 10 mL) and H$_2$O (10 mL), dried over MgSO$_4$, filtered and concentrated *in vacuo*. The residue was purified by flash column chromatography (silica gel, PE/Et$_2$O 1:1 to 1:4) to yield unreacted **28** (31.4 mg, 0.115 mmol, 58%) as a white solid, **32b** (0.8 mg, 2.5 µmol, 1%) as a transparent viscous oil and **32a** (16.2 mg, 50.1 µmol, 25%) as a white solid. Overall yield based on recovered starting material: **32b** + **32a** (17.0 mg, 52.6 µmol, 62%). **32a** was crystallised from Et$_2$O for the single crystal X-ray crystallography analysis.

Analytical data for **32a**:

R_f = 0.04 (PE/Et$_2$O 1:1). m.p. 90 – 91 °C (Et$_2$O). IR (ATR) v_{Max} 2931 (w, C-H), 1736 (s, C=O), 1614 (w, C=C), 1514 (s, C=C) cm^{-1}. ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.23 (2H, dt, J = 8.9, 2.5 Hz, H11), 6.85 (2H, dt, J = 8.9, 2.5 Hz, H12), 4.34 (1H, d, J = 15.6 Hz, H9a), 4.23 (1H, d, J = 15.6 Hz, H9b), 4.16 (1H, d, J = 8.5 Hz, H1a), 3.95 (1H, dd, J = 8.5, 1.7 Hz, H1b), 3.79 (3H, s, H14), 2.10 – 2.03 (1H, m, H6a), 1.87 – 1.72 (2H, m, H3a, H6b), 1.68 – 1.55 (3H, m, H3a, H4, H7a), 1.53 – 1.44 (1H, m, H5), 1.33 – 1.26 (1H, m, H7b) ppm. ^{13}C NMR (126 MHz, CDCl$_3$) δ 159.3 (s, C13), 158.1 (s, C8), 130.2 (s, C10), 129.4 (s, C11), 114.1 (s, C12), 113.9 (dd, J = 287.3, 284.0 Hz, C15), 71.5 (d, J = 1.5 Hz, C1), 59.0 (d, J = 3.6 Hz, C2), 55.4 (s, C14), 43.8 (s, C9), 29.0 (dd, J = 4.6, 1.0 Hz, C7), 23.2 (d, J = 3.1 Hz, C3), 16.2 (t, J = 11.3 Hz, C4/C5), 16.1 (t, J = 11.3 Hz, C4/C5), 15.5 (br s, C6) ppm. ^{19}F NMR (376 MHz, CDCl$_3$) δ -

127.2 (1F, dtd, J = 157.7, 14.1, 1.2 Hz, Fa), -150.4 (1F, d, J = 157.7 Hz, Fb) ppm. HRMS (ESI) calcd for $[C_{17}H_{20}NO_3F_2]^+$: 324.1411, found 324.1418.

Analytical data for **32b**:

R_f = 0.06 (PE/Et$_2$O 1:1). IR (ATR) ν_{Max} 2933 (w, C-H), 1738 (s, C=O), 1612 (w, C=C), 1512 (s, C=C) cm^{-1}. ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.23 (2H, br d, J = 8.7 Hz, H11), 6.85 (2H, br d, J = 8.7 Hz, H12), 4.49 (1H, d, J = 15.9 Hz, H9a), 4.32 (1H, d, J = 15.9 Hz, H9b), 4.03 (1H, dd, J = 8.9, 1.3 Hz, H1a), 3.94 (1H, d, J = 8.9 Hz, H1b), 3.80 (3H, s, H14), 2.06 – 1.97 (1H, m, H6a), 1.96 (1H, dd, J = 15.3, 8.3 Hz, H3a), 1.73 (1H, d, J = 15.3 Hz, H3b), 1.68 – 1.50 (4H, m, H4, H5, H6b, H7a), 1.44 – 1.36 (1H, m, H7b) ppm. ^{13}C NMR (126 MHz, CDCl$_3$) δ 159.2 (s, C13), 158.4 (s, C8), 130.3 (s, C10), 128.6 (s, C11), 114.3 (s, C12), 114.1 (dd, J = 287.3, 284.2 Hz, C15), 73.8 (d, J = 11.5 Hz, C1), 59.1 (d, J = 2.1 Hz, C2), 55.4 (s, C14), 44.2 (s, C9), 29.8 (dd, J = 2.6, 0.6 Hz, C7), 24.9 (dd, J = 2.0, 0.6 Hz, C3), 18.0 (t, J = 11.7 Hz, C4), 16.9 (t, J = 11.5 Hz, C5), 13.8 (d, J = 3.2 Hz, C6) ppm. ^{19}F NMR (376 MHz, CDCl$_3$) δ -127.3 (1F, dt, J = 157.4, 13.8 Hz, Fa), -149.4 (1F, d, J = 157.4 Hz, Fb) ppm. HRMS (ESI) calcd for $[C_{17}H_{20}NO_3F_2]^+$: 324.1411, found 324.1419.

1-(4-Methoxybenzyl)-3-oxa-1-azaspiro[4.5]decane-2,7-dione **33a**, 1-(4-Methoxybenzyl)-3-oxa-1-azaspiro[4.5]decane-2,8-dione **33b**, 8-hydroxy-1-(4-methoxy-benzyl)-3-oxa-1-azaspiro[4.5]decan-2-one **34a**, (5R*,7S*)-7-Hydroxy-1-(4-methoxy-benzyl)-3-oxa-1-azaspiro[4.5]decan-2-one **34b** and 8-hydroxy-1-(4-methoxybenzyl)-3-oxa-1-azaspiro[4.5]decan-2-one **34c**

33a   **33b**   **34a / 34c**   **34b**
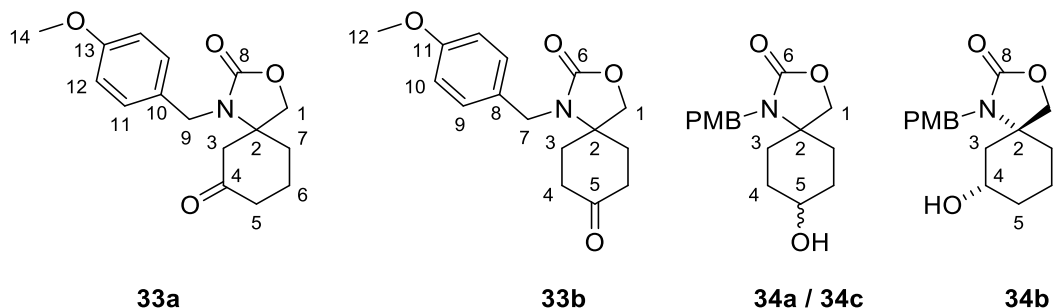
28 (109 mg, 0.400 mmol) was added to a solution of iron(II) acetylacetonate (20.4 mg, 80 μmol) and poly(methylhydrosiloxane) (272 μL) in tBuOH (4.0 mL) and the reaction stirred for 24 h at 50 °C. The reaction mixture was then quenched by silica gel, stirred for 10 min, and concentrated *in vacuo*. The residue was purified by flash column chromatography (silica gel, PE/EtOAc, 9:1 to 0:1) to yield unreacted **28** (8.8 mg, 32 μmol, 8%), ketones **33a** (32.5 mg, 0.112 mmol, 28%) and **33b** (17.0 mg, 59 μmol, 15%) as white solids, and a mixture of alcohols **34a**, **34b** and **34c** (38.5 mg, 0.132 mmol, 33% combined) as a transparent viscous oil. The individual alcohol isomers were separated by preparative TLC (silica gel, eluting with either 5% MeOH in CH$_2$Cl$_2$ or EtOAc). **34a** appeared as a transparent viscous oil, **34b** and **34c** as white amorphous solids. **34b** was crystallised from Et$_2$O for the single crystal X-ray crystallography analysis. **34a** and **34c** gave viscous oils or fibrous materials after each attempted crystallisation, that were not suitable for single crystal X-ray crystallography analysis, therefore their geometry could not be assigned.

Analytical data for **33a**:

m.p. 100 – 101 °C (Et$_2$O). IR (ATR) v_{Max} 2962 (w, C-H), 2930 (w, C-H), 1717 (s, C=O), 1615 (w, C=C), 1514 (s, C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.29 – 7.22 (2H, m, H11), 6.89 – 6.82 (2H, m, H12), 4.43 (1H, d, J = 16.2 Hz, H9a), 4.42 (1H, d, J = 16.2 Hz, H9b), 3.98 (1H, d, J = 8.9 Hz, H1a), 3.96 (1H, d, J = 8.9 Hz, H1b), 3.80 (3H, s, H14), 2.49 (1H, d, J = 13.6 Hz, H3a), 2.39 – 2.29 (2H, m, H3b, H5a), 2.20 (1H, td, J = 14.0, 6.1 Hz, H5b), 2.08 – 1.91 (2H, m, H6a, H7a), 1.81 – 1.72 (1H, m, H7b), 1.44 (1H, qt, J = 13.5, 4.0 Hz, H6b) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 206.9 (C4), 159.4 (C8), 157.7 (C13), 129.7 (C10), 129.1 (C11), 114.3 (C12), 71.7 (C1), 63.7 (C2), 55.4 (C14), 50.3 (C3), 44.0 (C9), 40.2 (C5), 33.6 (C7), 20.0 (C6) ppm. HRMS (ESI) calcd for [C$_{16}$H$_{20}$NO$_4$]$^+$: 290.1392, found 290.1399.

Analytical data for **33b**:

R_f = 0.44 (EtOAc). m.p. 121 – 122 °C (crystal decomposition), 136 – 137 °C (Et$_2$O). IR (ATR) v_{Max} 2906 (w, C-H), 1732 (s, C=O), 1706 (s, C=O), 1616 (w, C=C), 1513 (s, C=C) cm^{-1}. ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.25 – 7.21 (2H, m, H9), 6.86 – 6.82 (2H, m, H10), 4.37 (2H, s, H7), 4.36 (2H, s, H1), 3.79 (3H, s, H12), 2.43 – 2.29 (4H, m, H4), 2.04 (2H, td, J = 13.4, 5.5 Hz, H3a), 1.86 (2H, dqui, J = 13.8, 3.0 Hz, H3b) ppm. ^{13}C NMR (126 MHz, CDCl$_3$) δ 207.3 (C5), 159.4 (C6), 157.9 (C11), 130.1 (C8), 129.0 (C9), 114.3 (C10), 71.2 (C1), 60.4 (C2), 55.4 (C12), 44.0 (C7), 37.2 (C4), 33.0 (C3) ppm. HRMS (ESI) calcd for [C$_{16}$H$_{20}$NO$_4$]$^+$: 290.1392, found 290.1402.

Analytical data for **34a**:

R_f = 0.29 (5% MeOH in CH$_2$Cl$_2$). IR (ATR) v_{Max} 3358 (w, br, O-H), 2921 (w, C-H), 2852 (w, C-H), 1728 (s, C=O), 1660 (w, C=C), 1513 (s, C=C) cm^{-1}. ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.28 (2H, dt, J = 8.7, 2.5 Hz, H9), 6.84 (2H, dt, J = 8.7, 2.5 Hz, H10), 4.37 (2H, s, H7), 4.10 (2H, s, H1), 4.02 (1H, sex, J = 2.5 Hz, H5), 3.79 (3H, s, H12), 2.07 (2H, td, J = 13.5, 4.0 Hz, H3a), 1.81 (2H, br d, J = 16.4 Hz, H4a), 1.47 (2H, tdd, J = 14.3, 3.7, 2.6 Hz, H4b), 1.34 (2H, dqui, J = 13.1, 2.0 Hz, H3b), 1.17 (1H, d, J = 2.4 Hz, OH) ppm. ^{13}C NMR (126 MHz, CDCl$_3$) δ 159.1 (C11), 158.3 (C6), 130.8 (C8), 129.1 (C9), 114.0 (C10), 71.6 (C1), 63.7 (C5), 61.4 (C2), 55.4 (C12), 43.6 (C7), 29.5 (C4), 27.7 (C3) ppm. HRMS (ESI) calcd for [C$_{16}$H$_{22}$NO$_4$]$^+$: 292.1549, found 292.1544.

Analytical data for **34b**:

R_f = 0.23 (5% MeOH in CH$_2$Cl$_2$). m.p. 134 – 135 °C (Et$_2$O). IR (ATR) v_{Max} 3472 (w, br, O-H), 2921 (w, C-H), 2851 (w, C-H), 1728 (s, C=O), 1613 (w, C=C), 1510 (s, C=C) cm^{-1}. ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.25 (2H, dt, J = 9.6, 2.5 Hz, H11), 6.84 (2H, dt, J = 9.6, 2.5 Hz, H12), 4.37 (1H, d, J = 15.9 Hz, H9a), 4.33 (1H, d, J = 15.9 Hz, H9b), 4.05 (1H, d, J = 8.8 Hz, H1a), 4.04 (1H, d, J = 8.8 Hz, H1b), 3.79 (3H, s, H14), 3.58 – 3.48 (1H, m, H4), 1.97 – 1.92 (1H, m, H5a), 1.92 – 1.87 (1H, m, H3a), 1.79 – 1.73 (1H, m, H6a), 1.55 – 1.47 (3H, m, H3b, H7a, OH), 1.45 – 1.37 (1H, m, H7b), 1.27 – 1.16 (1H, m, H6b), 1.14 – 1.04 (1H, m, H5b) ppm. ^{13}C NMR (126 MHz, CDCl$_3$) δ 159.2 (C13), 158.1 (C8), 130.4 (C10), 129.0 (C11), 114.2 (C12), 72.1 (C1), 67.7 (C4), 62.1 (C2), 55.4 (C14), 43.7 (C9), 43.5 (C3), 34.2 (C5), 33.0 (C7), 19.6 (C6) ppm. HRMS (ESI) calcd for [C$_{16}$H$_{22}$NO$_4$]$^+$: 292.1549, found 292.1553.

Analytical data for **34c**:

R_f = 0.23 (EtOAc). IR (ATR) v_{Max} 3441 (w, br, O-H), 2940 (w, C-H), 2861 (w, C-H), 1720 (s, C=O), 1612 (w, C=C), 1511 (s, C=C) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 7.23 (2H, br d, J = 8.6 Hz, H9), 6.84 (2H, br d, J = 8.6 Hz, H10), 4.32 (2H, s, H7), 4.13 (2H, s, H1), 3.79 (3H, s, H12), 3.59 – 3.48 (1H, m, H5), 1.99 – 1.90 (2H, m, H4a), 1.68 – 1.51 (5H, m, H3, OH), 1.32 – 1.19 (2H, m, H4b) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 159.1 (C11), 158.3 (C6), 130.5 (C8), 128.9 (C9), 114.1 (C10), 71.9 (C1), 69.0 (C5), 60.9 (C2), 55.4 (C12), 43.7 (C7), 32.1 (C3), 31.8 (C4) ppm. HRMS (ESI) calcd for [C$_{16}$H$_{22}$NO$_4$]$^+$: 292.1549, found 292.1545.

(5R*,7S*,8S*)-7,8-Dibromo-3-oxa-1-azaspiro[4.5]decan-2-one **35**
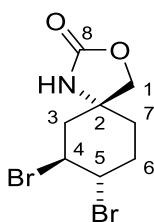
Deprotection with TFA:

A solution of **30a** (10.8 mg, 25.0 µmol) in TFA (250 µL) was heated under reflux overnight, then cooled to rt and the reaction mixture was concentrated *in vacuo*. The residue was purified by flash column chromatography (silica gel, CH$_2$Cl$_2$/Et$_2$O 9:1) to yield **35** (6.9 mg, 22.0 µmol, 88%) as a white amorphous solid. **35** was crystallised from Et$_2$O for the single crystal X-ray crystallography analysis.

Improved deprotection with CAN:

To a solution of **30a** (10.8 mg, 25 µmol) in MeCN (400 µL) and H$_2$O (100 µL) was added CAN (41.1 mg, 75 µmol) and stirred for 1 h at rt. Upon completion, the reaction mixture was quenched with NaHCO$_3$ (sat. aq, 3 mL), diluted with H$_2$O (6 mL) and extracted with CH$_2$Cl$_2$ (4 × 10 mL). The combined organic layers were dried (MgSO$_4$), then filtered through a short silica gel plug, washed with CH$_2$Cl$_2$ to remove the p-anisaldehyde by-product, then eluted with Et$_2$O and concentrated *in vacuo* to yield **35** (7.5 mg, 24 µmol, 96%) as a white amorphous solid.

Analytical data for **35** (recrystallised product obtained from reaction with TFA):

R_f = 0.27 (CH$_2$Cl$_2$/Et$_2$O 4:1). m.p. 173 – 174 °C (decomposition, Et$_2$O). IR (ATR) v_{Max} 3197 (w, N-H), 3122 (w, C-H), 2954 (w, C-H), 1741 (s, C=O) cm^{-1}. ^{1}H NMR (400 MHz, CDCl$_3$) δ 5.91 (1H, br s, NH), 4.55 – 4.46 (1H, m, H4), 4.41 (1H, br s, H5), 4.35 (1H, d, J = 9.0 Hz, H1a), 4.31 (1H, d, J = 9.0 Hz, H1b), 2.78 (1H, dd, J = 14.8, 3.8 Hz, H3a), 2.55 – 2.44 (1H, m, H6a), 2.22 (1H, dd, J = 14.8, 5.7 Hz, H3b), 2.12 – 1.95 (2H, m, H6b, H7a), 1.92 – 1.81 (1H, m, H7b) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 158.3 (C8), 75.1 (C1), 57.5 (C2), 51.8 (C5), 50.4 (C4), 41.0 (C3), 32.8 (C7), 28.4 (C6) ppm. HRMS (ESI) calcd for [C$_8$H$_{12}$NO$_2$^{79}Br$_2$]$^+$: 311.9229, found 311.9219.

4.2.6 Syntheses of bis-amide linked compounds

tert-Butyl (2-(5-bromo-1*H*-indole-3-carboxamido)ethyl)carbamate **47a**
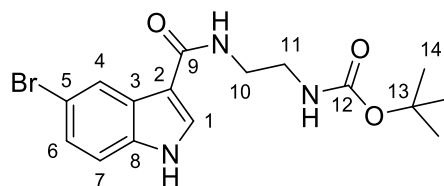
To a suspension of 5-bromo-1*H*-indole-3-carboxilic acid **48** (200 mg, 0.834 mmol) in CH_2Cl_2 (15 mL) was added DMF (6 µL, 80 µmol) and thionyl chloride (121 µL, 1.67 mmol) and heated under reflux overnight. The reaction mixture was concentrated under a flow of dry nitrogen to yield the crude acid chloride intermediate **49** as a pale pink/brown solid. The residue was dissolved in THF (10 mL) and the resultant solution added dropwise to a solution of *tert*-butyl (2-aminoethyl)carbamate (264 µL, 1.67 mmol) and DIPEA (1.45 mL, 8.34 mmol) in THF (10 mL) at 0 °C, the reaction mixture was warmed to rt and stirred for 2 h then concentrated *in vacuo*. The residue was purified by short silica gel plug filtration, washed with CH_2Cl_2, then eluted with EtOAc and concentrated *in vacuo* to yield **47a** (250 mg, 0.654 mmol, 78%) as a pale yellow amorphous solid.

R_f = 0.72 (acetone). IR (ATR) v_{Max} 3361 (m, N-H), 3214 (m, N-H), 2931 (w, C-H), 1684 (s, C=O), 1612 (s, C=O), 1545 (s, C=C), 1493 (m, C=C), 1426 (m, C=C) cm^{-1}. ^{1}H NMR (400 MHz, MeOD) δ 8.29 (1H, d, *J* = 1.9 Hz, H4), 7.88 (1H, s, H1), 7.34 (1H, dd, *J* = 8.6, 0.5 Hz, H7), 7.28 (1H, dd, *J* = 8.6, 1.9 Hz, H6), 3.45 (2H, t, *J* = 6.1 Hz, H10), 3.28 (2H, t, *J* = 6.1 Hz, H11), 1.41 (9H, s, H14) ppm. ^{13}C NMR (101 MHz, MeOD) δ 168.1 (C9), 158.9 (C12), 136.7 (C8), 130.1 (C1), 128.9 (C3), 126.3 (C6), 124.5 (C4), 115.3 (C5), 114.4 (C7), 111.6 (C2), 80.2 (C13), 41.2 (C11), 40.8 (C10), 28.7 (C14) ppm. HRMS (ESI) calcd for $[C_{16}H_{21}N_3O_3{}^{79}Br]^+$: 382.0766, found 382.0753.

tert-Butyl (3-(5-bromo-1*H*-indole-3-carboxamido)propyl)carbamate **47b**
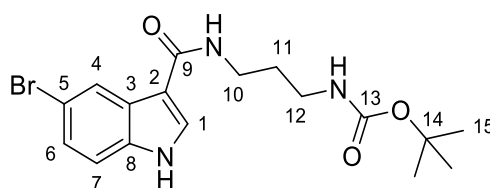
To a solution of 5-bromo-1*H*-indole-3-carboxilic acid **48** (100 mg, 0.417 mmol) in DMF (5.0 mL) was added *tert*-butyl (3-aminopropyl)carbamate (109 mg, 0.626 mmol), EDC·HCl (120 mg, 0.626 mmol), HOBt·xH₂O (~80%, 110 mg, 0.626 mmol) and Et₃N (87 µL, 0.626 mmol) at 0 °C. The reaction mixture was warmed to rt and stirred for 2 days, then concentrated *in vacuo*. The residue was purified by flash column chromatography (silica gel, PE/EtOAc 1:1 to 1:9) to yield impure **47b** (92.5 mg, <0.232 mmol, <56%) as a yellow amorphous solid. Structure is tentatively assigned based on limited analytical data and analogy to **47a**.

R_f = 0.75 (acetone). ^{1}H NMR (400 MHz, MeOD) δ 8.30 (1H, s, H4), 7.89 (1H, s, H1), 7.35 (1H, d, J = 8.6 Hz, H7), 7.28 (1H, d, J = 8.6 Hz, H6), 3.41 (2H, t, J = 6.5 Hz, H10), 3.16 (2H, t, J = 6.5 Hz, H12), 1.76 (2H, qui, J = 6.5 Hz, H11), 1.44 (9H, s, H15) ppm.

(2-(5-Bromo-1H-indole-3-carboxamido)ethan-1-aminium chloride **50a**
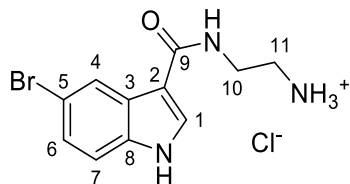
To a solution of **47a** (55.0 mg, 0.144 mmol) in MeOH (10 mL) was added acetyl chloride (2 mL) dropwise at -40 °C. The reaction mixture was warmed to rt and stirred for 4 h, then concentrated *in vacuo* to yield crude amine hydrochloride **50a** (51.6 mg) as a pale brown solid. The crude product **50a** was taken on to the next steps without further purification, and its structure was tentatively assigned based on limited analytical data.

IR (ATR) v_{Max} 3161 (m, N-H), 2872 (m, C-H), 1574 (s, C=O), 1499 (m, C=C), 1444 (s, C=C) cm^{-1}. HRMS (ESI) calcd for [$C_{11}H_{13}N_3O^{79}Br$]$^+$: 282.0242, found 282.0234.

3-(5-Bromo-1H-indole-3-carboxamido)propan-1-aminium chloride **50b**
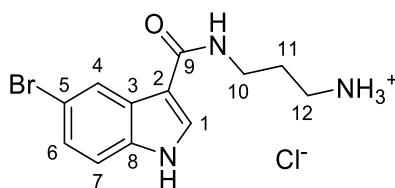
To a solution of **47b** (82.0 mg, 0.207 mmol) in MeOH (2.0 mL) was added HCl (4 M in dioxane, 1.0 mL, 4.0 mmol) at rt and the reaction mixture was stirred overnight, then concentrated *in vacuo* to yield crude amine hydrochloride **50b** (96.6 mg) as a pale brown solid. The crude product **50b** was taken on to the next step without further purification, and its structure was tentatively assigned based on limited analytical data.

IR (ATR) v_{Max} 2888 (m, br, N-H and C-H), 1582 (s, C=O), 1516 (m, C=C), 1438 (s, C=C) cm^{-1}. HRMS (ESI) calcd for [$C_{12}H_{15}N_3O^{79}Br$]$^+$: 296.0399, found 296.0389.

4-((2-(5-Bromo-1H-indole-3-carboxamido)ethyl)amino)-4-oxobutanoic acid **51a**
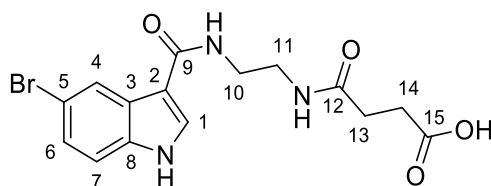
To a suspension of crude **50a** (17.0 mg, 46 μmol) in THF (6 mL) was added DIPEA (17.4 μL, 0.100 mmol) followed by succinic anhydride (5.0 mg, 50 μmol) and stirred at rt for 2 days, then

concentrated *in vacuo*. The residue was purified by flash column chromatography (silica gel, MeCN/H$_2$O 9:1) to yield **51a** (15.0 mg, 39 μmol, 85%) as a white amorphous solid.

R_f = 0.22 (MeCN/H$_2$O 9:1). IR (ATR) v_{Max} 3342 (w, N-H), 3191 (m, O-H), 2918 (w, C-H), 1660 (s, C=O), 1610 (s, C=O), 1545 (s, C=C), 1465 (m, C=C), 1437 (m, C=C), 1408 (m, C=C) cm^{-1}. ^{1}H NMR (400 MHz, MeOD) δ 8.28 (1H, dd, J = 1.9, 0.5 Hz, H4), 7.90 (1H, s, H1), 7.35 (1H, dd, J = 8.7, 0.5 Hz, H7), 7.28 (1H, dd, J = 8.7, 1.9 Hz, H6), 3.52 – 3.47 (2H, m, H10), 3.43 – 3.39 (2H, m, H11), 2.63 – 2.58 (2H, m, H14), 2.51 – 2.46 (2H, m, H13) ppm. ^{13}C NMR (101 MHz, MeOD) δ 176.8 (C15), 175.3 (C12), 168.1 (C9), 136.6 (C8), 130.1 (C1), 128.9 (C3), 126.3 (C6), 124.5 (C4), 115.3 (C5), 114.4 (C7), 111.4 (C2), 40.4 (C11), 40.1 (C10), 31.8 (C13), 30.6 (C14) ppm. HRMS (ESI) calcd for [C$_{15}$H$_{17}$N$_3$O$_4$^{79}Br]$^+$: 382.0402, found 382.0383.

5-((2-(5-Bromo-1*H*-indole-3-carboxamido)ethyl)amino)-5-oxopentanoic acid **51b**
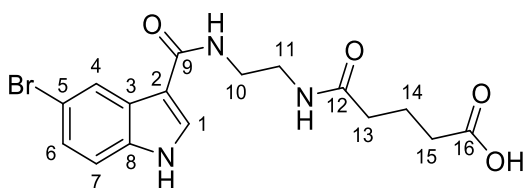
To a suspension of crude **50a** (5.5 mg, 15 μmol) in THF (2 mL) was added DIPEA (5.8 μL, 33 μmol) followed by succinic anhydride (1.8 mg, 18 μmol) and stirred at rt for 2 days, then concentrated *in vacuo*. The residue was purified by flash column chromatography (silica gel, MeCN/H$_2$O 9:1) to yield **51b** (1.9 mg, 4.9 μmol, 33%) as a white amorphous solid. The structure of **51b** was tentatively assigned based on LCMS analysis and analogy to **51a**.

R_f = 0.24 (MeCN/H$_2$O 9:1). LCMS (ESI) calcd for [C$_{16}$H$_{19}$N$_3$O$_4$^{81}Br]$^+$: 398.1, found 398.2; calcd for [C$_{16}$H$_{17}$N$_3$O$_4$^{81}Br]$^-$: 396.0, found 396.1.

4-((3-(5-Bromo-1*H*-indole-3-carboxamido)propyl)amino)-4-oxobutanoic acid **51c**
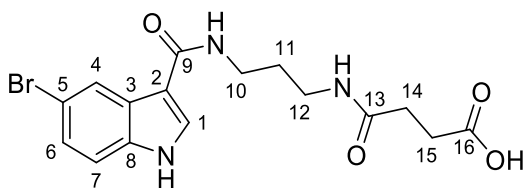
To a suspension of crude **50b** (6.7 mg, 14 μmol) in THF (2.0 mL) was added Et$_3$N (9.8 μL, 70 μmol) followed by succinic anhydride (2.5 mg, 25 μmol) and stirred at rt overnight. The reaction mixture was diluted with HCl (conc. aq, 10 mL) and brine (10 mL), then extracted with EtOAc/*i*PrOH (9:1, 5 × 20 mL), the combined organic layers were dried over Na$_2$SO$_4$, filtered and concentrated *in vacuo* to yield the crude product **51c** (2.3 mg, <5.8 μmol, <41%) as a pale yellow amorphous solid. The crude **51c** was tested for biological activity without further purification, and its structure was tentatively assigned based on crude ^{1}H NMR and LCMS analyses and analogy to **51a**.

^{1}H NMR (400 MHz, MeOD) δ 8.25 (1H, br s, H4), 7.96 (1H, s, H1), 7.38 (1H, br d, J = 8.7 Hz, H7), 7.30 (1H, br d, J = 8.7 Hz, H6), 3.64 – 3.34 (4H, m, H10, H12), 2.72 – 2.50 (4H, m, H14, H15), 1.94 – 1.79 (2H, m, H11) ppm. LCMS (ESI) calcd for $[C_{16}H_{19}N_3O_4{}^{81}Br]^+$: 398.1, found 398.2; calcd for $[C_{16}H_{17}N_3O_4{}^{79}Br]^-$: 394.0, found 394.1.

Ethyl 4-((2-(5-Bromo-1*H*-indole-3-carboxamido)ethyl)amino)-4-oxobutanoate **52**
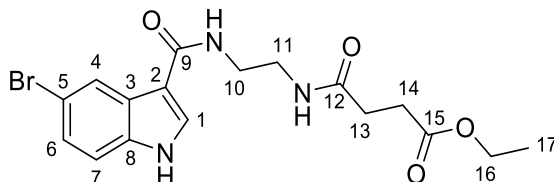
To a suspension of crude **50a** (18.6 mg, 52 μmol) in THF (6 mL) was added DIPEA (20.9 μL, 0.200 mmol) and mono-ethyl succinate (14.6 mg, 0.100 mmol) followed by HATU (21.0 mg, 55 μmol) and stirred at rt overnight. The reaction mixture was purified by flash column chromatography (silica gel, EtOAc to EtOAc/acetone 4:1) to yield **52** (11.8 mg, 29 μmol, 56%) as a white amorphous solid.

R_f = 0.58 (acetone). IR (ATR) v_{Max} 3330 (w, N-H), 3223 (m, N-H), 2982 (w, C-H), 2934 (w, C-H), 1707 (s, C=O, ester), 1652 (s, C=O, amide), 1610 (s, C=O, amide), 1556 (s, C=C), 1469 (m, C=C), 1444 (m, C=C) cm^{-1}. ^{1}H NMR (400 MHz, acetone-d$_6$) δ 8.46 (1H, dd, J = 2.0, 0.5 Hz, H4), 7.97 (1H, s, H1), 7.51 (1H, br t, J = 5.0 Hz, NH), 7.47 (1H, br s, NH), 7.44 (1H, dd, J = 8.6, 0.5 Hz, H7), 7.29 (1H, dd, J = 8.6, 2.0 Hz, H6), 4.04 (2H, q, J = 7.1 Hz, H16), 3.51 – 3.45 (2H, m, H10), 3.43 – 3.37 (2H, m, H11), 2.86 (1H, br s, NH), 2.60 – 2.55 (2H, m, H14), 2.50 – 2.45 (2H, m, H13), 1.17 (3H, t, J = 7.1 Hz, H17) ppm. ^{13}C NMR (101 MHz, acetone-d$_6$) δ 173.3 (C15), 172.7 (C12), 165.6 (C9), 136.0 (C8), 129.1 (C3), 129.1 (C1), 125.8 (C6), 124.7 (C4), 114.6 (C5), 114.3 (C7), 112.2 (C2), 60.7 (C16), 40.3 (C10/C11), 40.3 (C10/C11), 31.2 (C13), 30.0 (C14, under solvent peak), 14.5 (C17) ppm. HRMS (ESI) calcd for $[C_{17}H_{21}N_3O_4{}^{79}Br]^+$: 410.0715, found 410.0702.

4.2.7 Syntheses of further derivatives

General procedure **A**

To a solution of amine (1.00 mmol) in THF (5.0 mL) was added Et$_3$N (280 μL, 2.00 mmol) followed by succinic anhydride (120 mg, 1.20 mmol) at 0 °C. The reaction mixture was warmed to rt and stirred overnight, then diluted with HCl (conc. aq, 10 mL) and brine (10 mL), and extracted with EtOAc (5 × 20 mL). The combined organic layers were dried over Na$_2$SO$_4$, filtered and concentrated *in vacuo* to yield the crude product **53**. The crude mixtures **53** were tested for biological activity without further purification, and their structures were tentatively assigned based on crude ^{1}H NMR analysis.

4-oxo-4-(phenethylamino)butanoic acid **53a**
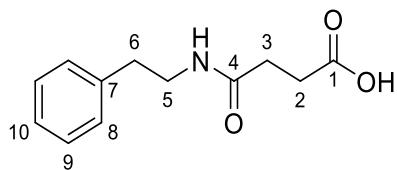
Synthesised according to general procedure **A**.

^{1}H NMR (400 MHz, CDCl$_3$) δ 7.35 – 7.29 (2H, m, H9), 7.27 – 7.21 (1H, m, H10, CHCl$_3$), 7.21 – 7.16 (2H, m, H8), 5.71 (1H, br s, NH), 3.54 (2H, q, *J* = 6.6 Hz, H5), 2.82 (2H, t, *J* = 6.9 Hz, H6), 2.69 (2H, t, *J* = 6.6 Hz, H3), 2.46 (2H, t, *J* = 6.5 Hz, H2) ppm.

These characterization data are in accordance with that previously reported in the literature.[281]

4-oxo-4-((3-phenylpropyl)amino)butanoic acid **53b**
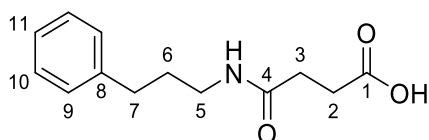
Synthesised according to general procedure **A**.

^{1}H NMR (400 MHz, CDCl$_3$) δ 7.32 – 7.27 (2H, m, H10), 7.23 – 7.15 (3H, m, H9, H11), 5.64 (1H, br s, NH), 3.31 (2H, q, *J* = 6.6 Hz, H5), 2.71 – 2.62 (4H, m, H3, H7), 2.44 (2H, t, *J* = 6.4 Hz, H2), 1.86 (2H, qui, *J* = 7.3 Hz, H2) ppm.

4-oxo-4-((4-phenylbutyl)amino)butanoic acid **53c**
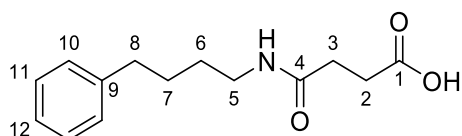
Synthesised according to general procedure **A**.

^{1}H NMR (400 MHz, CDCl$_3$) δ 7.31 – 7.27 (2H, m, H11), 7.22 – 7.14 (3H, m, H10, H12), 5.68 (1H, br s, NH), 3.29 (2H, q, *J* = 6.6 Hz, H5), 2.69 (2H, t, *J* = 6.4 Hz, H3), 2.64 (2H, t, *J* = 7.4 Hz, H8), 2.49 (2H, t, *J* = 6.4 Hz, H2), 1.65 (2H, qui, *J* = 7.5 Hz, H6/H7), 1.54 (2H, qui, *J* = 7.2 Hz, H6/H7) ppm.

4-((2-(1*H*-indol-3-yl)ethyl)amino)-4-oxobutanoic acid **53d**
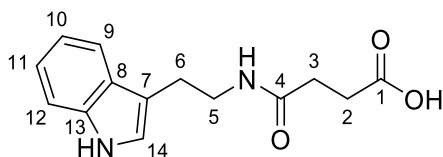
Synthesised according to general procedure **A**.

^{1}H NMR (400 MHz, MeOD) δ 7.55 (1H, d, J = 7.8 Hz, H9/H12), 7.32 (1H, d, J = 8.0 Hz, H9/H12), 7.11 – 7.05 (2H, m, H10/H11, H14), 7.00 (1H, t, J = 7.4 Hz, H10/H11), 3.47 (2H, t, J = 7.3 Hz, H5), 2.94 (2H, t, J = 7.3 Hz, H6), 2.61 – 2.54 (2H, m, H3), 2.49 – 2.42 (2H, m, H2) ppm. LCMS (ESI) calcd for $[C_{14}H_{17}N_2O_3]^+$: 261.1, found 261.3; calcd for $[C_{14}H_{15}N_2O_3]^-$: 259.1, found 259.2.

tert-Butyl 5-bromo-3-(2-hydroxyethyl)-1*H*-indole-1-carboxylate **56**
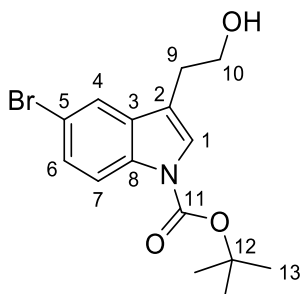
Following a slightly modified version of a reported procedure,[278] to a solution of **58** (1.36 g, 3.00 mmol) in THF (30 mL) was added TBAF (1.0 M in THF, 4.5 mL, 4.5 mmol) at 0 °C. The reaction mixture was warmed to rt and stirred for 1 h, then filtered through a short silica plug, eluted with EtOAc and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel, 5 to 25% EtOAc in PE) to yield **56** (916 mg, 2.60 mmol, 87%) as a transparent viscous oil.

R_f = 0.13 (PE/EtOAc 4:1). ^{1}H NMR (400 MHz, DMSO-d_6) δ 7.97 (1H, d, J = 8.8 Hz, H7), 7.81 (1H, d, J = 1.9 Hz, H4), 7.52 (1H, br s, H1), 7.45 (1H, dd, J = 8.8, 1.9 Hz, H6), 4.68 (1H, t, J = 5.3 Hz, OH), 3.66 (2H, td, J = 6.5, 5.3 Hz, H10), 2.79 (2H, t, J = 6.5 Hz, H9), 1.61 (9H, s, H13) ppm. ^{13}C NMR (101 MHz, CDCl$_3$) δ 149.5 (C11), 134.5 (C8), 132.4 (C3), 127.4 (C6), 124.8 (C1), 121.9 (C4), 116.9 (C7), 116.7 (C2), 116.0 (C5), 84.2 (C12), 62.0 (C10), 28.4 (C9), 28.3 (C13) ppm.

These characterization data are in accordance with that previously reported in the literature.[278]

5-Bromo-3-(2-((*tert*-butyldimethylsilyl)oxy)ethyl)-1*H*-indole **57**
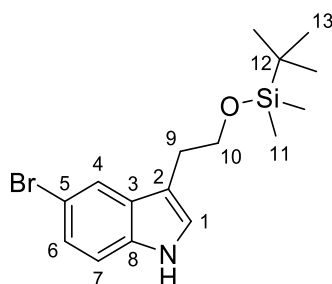
Following a slightly modified version of a reported procedure,[278] to a solution of 5-bromotryptophol **55** (960 mg, 4.00 mmol) and imidazole (599 mg, 8.80 mmol) in DMF (4.0 mL) was added TBSCl (663 mg, 4.40 mmol) at 0 °C. The reaction mixture was warmed to rt and stirred for 3 h, then diluted with EtOAc (20 mL), washed with water (20 mL), NH$_4$Cl (5% aq, 3 × 20 mL), NaHCO$_3$ (sat. aq, 20 mL) and brine (5 × 20 mL), dried over MgSO$_4$, filtered and concentrated *in*

vacuo to yield the crude product **57** (1.40 g) as a yellow oil. The crude product **57** was taken on to the next step without further purification, and its structure was tentatively assigned based on the crude ^{1}H NMR data that were in accordance with that previously reported in the literature[278].

R_f = 0.59 (CH$_2$Cl$_2$). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.01 (1H, br s, NH), 7.74 (1H, d, *J* = 1.8 Hz, H4), 7.26 (1H, dd, *J* = 8.6, 1.8 Hz, H6), 7.21 (1H, d, *J* = 8.6 Hz, H7), 7.04 (1H, d, *J* = 2.2 Hz, H1), 3.86 (2H, t, *J* = 7.0 Hz, H10), 2.93 (2H, t, *J* = 7.0 Hz, H9), 0.89 (9H, s, H13), 0.01 (6H, s, H11) ppm.

tert-Butyl 5-bromo-3-(2-((*tert*-butyldimethylsilyl)oxy)ethyl)-1*H*-indole-1-carboxylate **58**
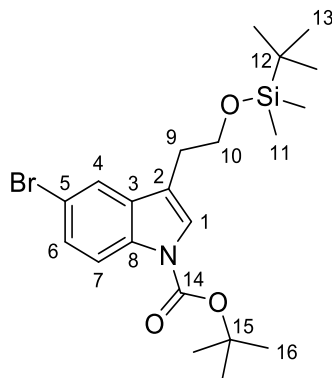
Following a slightly modified version of a reported procedure,[278] to a solution of crude **57** (1.40 g, ~4.0 mmol) and DMAP (52.0 mg, 0.40 mmol) in CH$_2$Cl$_2$ (12.0 mL) was added Boc$_2$O (1.31 g, 6.00 mmol) and stirred for 3 days. The reaction mixture was then concentrated *in vacuo*, and the crude product was purified by flash column chromatography (silica gel, 0 to 3% EtOAc in PE) to yield **58** (1.69 g, 3.71 mmol, 93% over two steps) as a transparent viscous oil.

R_f = 0.26 (5% EtOAc in PE). ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.00 (1H, br d, *J* = 8.8 Hz, H7), 7.66 (1H, d, *J* = 1.8 Hz, H4), 7.42 (1H, br s, H1), 7.38 (1H, dd, *J* = 8.8, 1.8 Hz, H6), 3.86 (2H, t, *J* = 6.6 Hz, H10), 2.85 (2H, t, *J* = 6.6 Hz, H9), 1.65 (9H, s, H16), 0.89 (9H, s, H13), 0.01 (6H, s, H11) ppm.

These characterization data are in accordance with that previously reported in the literature.[278]

tert-Butyl 6-(5-bromo-3-(2-((*tert*-butoxycarbonyl)oxy)ethyl)-1*H*-indol-1-yl)hexanoate **60**
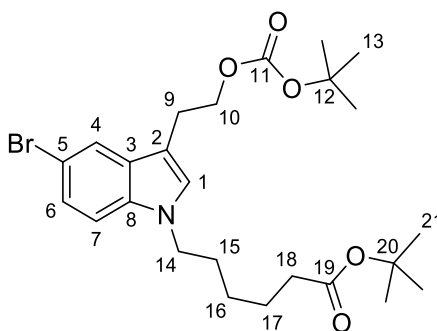
To a solution of **56** (34.0 mg, 0.100 mmol) in DMF (1.0 mL) was added NaH (60 w/w % dispersion in mineral oil, 12.0 mg, 0.300 mmol) and NaI (15.0 mg, 0.100 mmol) and the reaction stirred at rt for 15 min, followed by the addition of *tert*-butyl 6-bromohexanoate (50.2 mg, 0.200 mmol) and

stirred for 2.5 h. The reaction mixture was quenched by $NaHCO_3$ (sat. aq, 10 mL), stirred for 10 min, then extracted with EtOAc (10 mL). The organic layer was washed with brine (6 × 10 mL), dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel, 5 to 15% EtOAc in PE) to yield **60** (38.0 mg, 74.4 μmol, 74%) as a transparent viscous oil.

R_f = 0.20 (15% EtOAc in PE). ^{1}H NMR (400 MHz, $CDCl_3$) δ .71 (1H, dd, J = 1.9, 0.4 Hz, H4), 7.26 (1H, dd, J = 8.7, 1.9 Hz, H6, $CHCl_3$), 7.15 (1H, dd, J = 8.7, 0.4 Hz, H7), 6.94 (1H, s, H1), 4.28 (2H, t, J = 7.3 Hz, H10), 4.04 (2H, t, J = 7.1 Hz, H14), 3.05 (2H, td, J = 7.2, 0.5 Hz, H9), 2.18 (2H, t, J = 7.4 Hz, H18), 1.80 (2H, qui, J = 7.4 Hz, H15), 1.59 (2H, qui, J = 7.6 Hz, H17, H_2O), 1.48 (9H, s, H13/H21), 1.41 (9H, s, H13/H21), 1.34 – 1.24 (2H, m, H16) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 173.0 (C19), 153.6 (C11), 135.0 (C8), 129.8 (C3), 127.2 (C1), 124.5 (C6), 121.8 (C4), 112.4 (C5), 110.9 (C7), 110.1 (C2), 82.1 (C12/C20), 80.3 (C12/C20), 67.0 (C10), 46.3 (C14), 35.4 (C18), 30.1 (C15), 28.2 (C13/C21), 27.9 (C13/C21), 26.5 (C16), 24.9 (C9), 24.7 (C17) ppm.

5-Bromo-3-(2-((*tert*-butyldimethylsilyl)oxy)ethyl)-1-(4-methoxybenzyl)-1*H*-indole **61**
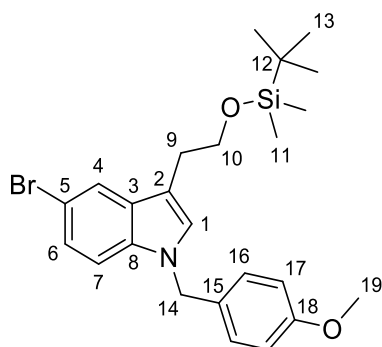
To a solution of crude **57** (700 mg, ~2.0 mmol) in DMF (10.0 mL) was added NaH (60 w/w % dispersion in mineral oil, 160 mg, 4.00 mmol) and the reaction stirred at rt for 10 min, followed by the addition of PMBCl (542 μL, 4.00 mmol) and stirred for 2 days. The reaction mixture was quenched by $NaHCO_3$ (sat. aq, 10 mL), stirred for 10 min, then diluted with H_2O (10 mL) and brine (30 mL), and extracted with EtOAc (50 mL). The organic layer was washed with brine (6 × 50 mL), dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude product was purified by flash column chromatography (silica gel, 0 to 4% EtOAc in PE) to yield **61** (690 mg, 1.45 mmol, 73% over two steps) as a transparent viscous oil.

R_f = 0.20 (5% EtOAc in PE). ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.73 (1H, dd, J = 1.8, 0.4 Hz, H4), 7.22 (1H, dd, J = 8.7, 1.8 Hz, H6), 7.11 (1H, dd, J = 8.7, 0.4 Hz, H7), 7.04 (2H, dm, J = 8.8 Hz, H16), 6.94 (1H, s, H1), 6.83 (2H, dm, J = 8.8 Hz, H17), 5.17 (2H, s, H14), 3.84 (2H, t, J = 6.9 Hz, H10), 3.78 (3H, s, H19), 2.91 (2H, td, J = 6.9, 0.6 Hz, H9), 0.88 (9H, s, H13), 0.00 (6H, s, H11) ppm. ^{13}C NMR (101 MHz, $CDCl_3$) δ 159.3 (C18), 135.2 (C8), 130.3 (C3), 129.3 (C15), 128.3 (C16), 127.5 (C1), 124.4 (C6), 122.0 (C4), 114.3 (C17), 112.4 (C2/C5), 112.3 (C2/C5), 111.2 (C7), 63.9 (C10), 55.4 (C19), 49.7 (C14), 28.9 (C9), 26.1 (C13), 18.5 (C12), -5.2 (C11) ppm.

5 References

(1)   Dougherty, T. J.; Pucci, M. J. *Antibiotic Discovery and Development*; Dougherty, T. J., Pucci, M. J., Eds.; Springer US: Boston, MA, **2012**.

(2)   Murray, C. J. L.; Ortblad, K. F.; Guinovart, C.; Lim, S. S.; Wolock, T. M.; Roberts, D. A.; Dansereau, E. A.; Graetz, N.; Barber, R. M.; Brown, J. C.; et al. *Lancet* **2014**, *384*, 1005–1070.

(3)   Siegel, R. L.; Miller, K. D.; Jemal, A. *CA. Cancer J. Clin.* **2016**, *66*, 7–30.

(4)   Rader, R. A. *Nat. Biotechnol.* **2008**, *26*, 743–751.

(5)   Schreiber, S. L. *Nat. Chem. Biol.* **2005**, *1*, 64–66.

(6)   O'Connell, K. M. G.; Galloway, W. R. J. D.; Spring, D. R. *The Basics of Diversity-Oriented Synthesis*. In *Diversity-Oriented Synthesis: Basics and Applications in Organic Synthesis, Drug Discovery, and Chemical Biology*. 1st ed.; Trabocchi, A., Ed.; John Wiley & Sons, Inc., **2013**, chapter 1, pp 1–26.

(7)   Lacaná, E.; Amur, S.; Mummanneni, P.; Zhao, H.; Frueh, F. W. *Clin. Pharmacol. Ther.* **2007**, *82*, 466–471.

(8)   Asher, M. *Nat. Rev. Drug Discov.* **2020**, *News*.

(9)   Berkowitz, S. A.; Engen, J. R.; Mazzeo, J. R.; Jones, G. B. *Nat. Rev. Drug Discov.* **2012**, *11*, 527–540.

(10)   Sams-Dodd, F. *Drug Discov. Today* **2005**, *10*, 139–147.

(11)   Bleicher, K. H.; Böhm, H. J.; Müller, K.; Alanine, A. I. *Nat. Rev. Drug Discov.* **2003**, *2*, 369–378.

(12)   Van Norman, G. A. *JACC Basic to Transl. Sci.* **2016**, *1*, 170–179.

(13)   DiMasi, J. A.; Grabowski, H. G.; Hansen, R. W. *J. Health Econ.* **2016**, *47*, 20–33.

(14)   Morgan, S.; Grootendorst, P.; Lexchin, J.; Cunningham, C.; Greyson, D. *Health Policy (New. York).* **2011**, *100*, 4–17.

(15)   Sertkaya, A.; Wong, H. H.; Jessup, A.; Beleche, T. *Clin. Trials* **2016**, *13*, 117–126.

(16)   Jayasundara, K.; Hollis, A.; Krahn, M.; Mamdani, M.; Hoch, J. S.; Grootendorst, P. *Orphanet J. Rare Dis.* **2019**, *14*, 1–10.

(17)   Scannell, J. W.; Blanckley, A.; Boldon, H.; Warrington, B. *Nat. Rev. Drug Discov.* **2012**, *11*, 191–200.

(18)   Smietana, K.; Siatkowski, M.; Møller, M. *Nat Rev Drug Discov* **2016**, *15*, 379–380.

(19)   Schuhmacher, A.; Gassmann, O.; Hinder, M. *J. Transl. Med.* **2016**, *14*, 105.

(20)   Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Deliv. Rev.* **1997**, *23*, 3–25.

(21)   Veber, D. F.; Johnson, S. R.; Cheng, H. Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D. *J. Med. Chem.* **2002**, *45*, 2615–2623.

(22)   Shultz, M. D. *J. Med. Chem.* **2019**, *62*, 1701–1714.

(23)   Leeson, P. D.; Springthorpe, B. *Nat. Rev. Drug Discov.* **2007**, *6*, 881–890.

(24)   Hughes, J. D.; Blagg, J.; Price, D. A.; Bailey, S.; DeCrescenzo, G. A.; Devraj, R. V.; Ellsworth, E.; Fobian, Y. M.; Gibbs, M. E.; Gilles, R. W.; et al. *Bioorganic Med. Chem. Lett.* **2008**, *18*, 4872–4875.

(25)   Lovering, F.; Bikker, J.; Humblet, C. *J. Med. Chem.* **2009**, *52*, 6752–6756.

(26)   Vamathevan, J.; Clark, D.; Czodrowski, P.; Dunham, I.; Ferran, E.; Lee, G.; Li, B.; Madabhushi, A.; Shah, P.; Spitzer, M.; et al. *Nat. Rev. Drug Discov.* **2019**, *18*, 463–477.

(27)   Feijoo, F.; Palopoli, M.; Bernstein, J.; Siddiqui, S.; Albright, T. E. *Drug Discov. Today* **2020**, *25*, 414–421.

(28) Kenakin, T. P. *The Drug Discovery Process*. In *A Pharmacology Primer*. 4th ed.; Elsevier, **2014**, chapter 11, pp 281–320.

(29) Sams-Dodd, F. *Drug Discov. Today* **2005**, *10*, 139–147.

(30) Swinney, D. C.; Anthony, J. *Nat. Rev. Drug Discov.* **2011**, *10*, 507–519.

(31) Holenz, J.; Stoy, P. *Bioorganic Med. Chem. Lett.* **2019**, *29*, 517–524.

(32) Hajduk, P. J.; Greer, J. *Nat. Rev. Drug Discov.* **2007**, *6*, 211–219.

(33) Keserű, G. M.; Makara, G. M. *Drug Discov. Today* **2006**, *11*, 741–748.

(34) Persidis, A. *Nat. Biotechnol.* **1998**, *16*, 488–489.

(35) Walters, W. P.; Namchuk, M. *Nat. Rev. Drug Discov.* **2003**, *2*, 259–266.

(36) Agresti, J. J.; Antipov, E.; Abate, A. R.; Ahn, K.; Rowat, A. C.; Baret, J.-C.; Marquez, M.; Klibanov, A. M.; Griffiths, A. D.; Weitz, D. A. *Proc. Natl. Acad. Sci.* **2010**, *107*, 4004–4009.

(37) Macarron, R.; Banks, M. N.; Bojanic, D.; Burns, D. J.; Cirovic, D. A.; Garyantes, T.; Green, D. V. S.; Hertzberg, R. P.; Janzen, W. P.; Paslay, J. W.; et al. *Nat. Rev. Drug Discov.* **2011**, *10*, 188–195.

(38) Bohacek, R. S.; McMartin, C.; Guida, W. C. *Med. Res. Rev.* **1996**, *16*, 3–50.

(39) Franzini, R. M.; Neri, D.; Scheuermann, J. *Acc. Chem. Res.* **2014**, *47*, 1247–1255.

(40) Goodnow, R. A.; Davie, C. P. *DNA-Encoded Library Technology: A Brief Guide to Its Evolution and Impact on Drug Discovery*. In *Annual Reports in Medicinal Chemistry*; Academic Press Inc., **2017**, Vol. 50, chapter 1, pp 1–15.

(41) Duffner, J. L.; Clemons, P. A.; Koehler, A. N. *Curr. Opin. Chem. Biol.* **2007**, *11*, 74–82.

(42) Hong, J. A.; Neel, D. V.; Wassaf, D.; Caballero, F.; Koehler, A. N. *Curr. Opin. Chem. Biol.* **2014**, *18*, 21–28.

(43) Parker, C. G.; Galmozzi, A.; Wang, Y.; Correia, B. E.; Sasaki, K.; Joslyn, C. M.; Kim, A. S.; Cavallaro, C. L.; Lawrence, R. M.; Johnson, S. R.; et al. *Cell* **2017**, *168*, 527-541.e29.

(44) Oprea, T. I.; Allu, T. K.; Fara, D. C.; Rad, R. F.; Ostopovici, L.; Bologa, C. G. *J. Comput. Aided. Mol. Des.* **2007**, *21*, 113–119.

(45) Baell, J. B.; Holloway, G. A. *J. Med. Chem.* **2010**, *53*, 2719–2740.

(46) Baell, J. B.; Nissink, J. W. M. *ACS Chem. Biol.* **2018**, *13*, 36–44.

(47) Krier, M.; Bret, G.; Rognan, D. *J. Chem. Inf. Model.* **2006**, *46*, 512–524.

(48) Koutsoukas, A.; Paricharak, S.; Galloway, W. R. J. D.; Spring, D. R.; IJzerman, A. P.; Glen, R. C.; Marcus, D.; Bender, A. *J. Chem. Inf. Model.* **2014**, *54*, 230–242.

(49) Ruddigkeit, L.; van Deursen, R.; Blum, L. C.; Reymond, J.-L. *J. Chem. Inf. Model.* **2012**, *52*, 2864–2875.

(50) Hann, M. M.; Leach, A. R.; Harper, G. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 856–864.

(51) Schuffenhauer, A.; Ruedisser, S.; Marzinzik, A.; Jahnke, W.; Selzer, P.; Jacoby, E. *Curr. Top. Med. Chem.* **2005**, *5*, 751–762.

(52) Hann, M. M.; Leach, A. R. *Coping with Complexity in Molecular Design*. In *De novo Molecular Design*. 1st ed.; Schneider, G., Ed.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, **2013**, chapter 2, pp 57–77.

(53) Leach, A. R.; Hann, M. M. *Curr. Opin. Chem. Biol.* **2011**, *15*, 489–496.

(54) Nilar, S. H.; Ma, N. L.; Keller, T. H. *J. Comput. Aided. Mol. Des.* **2013**, *27*, 783–792.

(55) Clemons, P. A.; Bodycombe, N. E.; Carrinski, H. A.; Wilson, J. A.; Shamji, A. F.; Wagner, B. K.; Koehler,

A. N.; Schreiber, S. L. *Proc. Natl. Acad. Sci.* **2010**, *107*, 18787–18792.

(56)   Lovering, F. *Med. Chem. Commun.* **2013**, *4*, 515–519.

(57)   Keserű, G. M.; Erlanson, D. A.; Ferenczy, G. G.; Hann, M. M.; Murray, C. W.; Pickett, S. D. *J. Med. Chem.* **2016**, *59*, 8189–8206.

(58)   Lipkus, A. H.; Yuan, Q.; Lucas, K. A.; Funk, S. A.; Bartelt, W. F.; Schenck, R. J.; Trippe, A. J. *J. Org. Chem.* **2008**, *73*, 4443–4451.

(59)   Schreiber, S. L. *Nature* **2009**, *457*, 153–154.

(60)   O' Connor, C. J.; Beckmann, H. S. G.; Spring, D. R. *Chem. Soc. Rev.* **2012**, *41*, 4444–4456.

(61)   Galloway, W. R. J. D.; Spandl, R. J.; Bender, A.; Thomas, G. L.; Diaz-Gavilan, M.; O'connell, K. M. G.; Spring, D. R. *Diversity-Oriented Synthesis*. In *Chemical Genomics*. 1st ed.; Fu, H., Ed.; Cambridge University Press: Cambridge, **2012**, chapter 4, pp 39–59.

(62)   Galloway, W. R. J. D.; Isidro-Llobet, A.; Spring, D. R. *Nat. Commun.* **2010**, *1*, 80.

(63)   Connell, K. M. G. O.; Galloway, W. R. J. D.; Ibbeson, B. M.; Isidro-Llobet, A.; O'Connor, C. J.; Spring, D. R. *Diversity-Oriented Synthesis*. In *Solid-Phase Organic Synthesis: Concepts, Strategies, and Applications*. 1st ed.; Toy, P. H., Lam, Y., Eds.; John Wiley & Sons, Inc., **2012**, chapter 4, pp 131–150.

(64)   Kim, J.; Jung, J.; Koo, J.; Cho, W.; Lee, W. S.; Kim, C.; Park, W.; Park, S. B. *Nat. Commun.* **2016**, *7*, 13196, 1–10.

(65)   Galloway, W. R. J. D.; Wilcke, D.; Nie, F.; Hadje-Georgiou, K.; Laraia, L.; Spring, D. R. *Diversity-Oriented Synthesis: Developing New Chemical Tools to Probe and Modulate Biological Systems*. In *Concepts and Case Studies in Chemical Biology*. 1st ed.; Waldmann, H., Janning, P., Eds.; Wiley-VCH Verlag, **2014**, chapter 26, pp 379–389.

(66)   Galloway, W. R. J. D.; Stokes, J. E.; Spring, D. R. *New Advances in Diversity-Oriented Synthesis*. In *Small Molecule Medicinal Chemistry: Strategies and Technologies*. 1st ed.; Czechtizky, W., Hamley, P., Eds.; John Wiley & Sons, Inc., **2016**, chapter 4, pp 79–101.

(67)   Butler, M. S. *Nat. Prod. Rep.* **2005**, *22*, 162–195.

(68)   Li, J. W.-H.; Vederas, J. C. *Science (80-.).* **2009**, *325*, 161–165.

(69)   Harvey, A. L. *Drug Discov. Today* **2008**, *13*, 894–901.

(70)   Spring, D. R. *Org. Biomol. Chem.* **2003**, *1*, 3867.

(71)   Wetzel, S.; Bon, R. S.; Kumar, K.; Waldmann, H. *Angew. Chemie Int. Ed.* **2011**, *50*, 10800–10826.

(72)   van Hattum, H.; Waldmann, H. *J. Am. Chem. Soc.* **2014**, *136*, 11853–11859.

(73)   Schreiber, S. L. *Science (80-.).* **2000**, *287*, 1964–1969.

(74)   Burke, M. D.; Schreiber, S. L. *Angew. Chemie Int. Ed.* **2004**, *43*, 46–58.

(75)   Tan, D. S.; Foley, M. A.; Shair, M. D.; Schreiber, S. L. *J. Am. Chem. Soc.* **1998**, *120*, 8565–8566.

(76)   Borman, S.; Washington, C. *Chem. Eng. News* **2004**, *4*, 32–40.

(77)   Spandl, R. J.; Bender, A.; Spring, D. R. *Org. Biomol. Chem.* **2008**, *6*, 1149.

(78)   Scott, D. E.; Coyne, A. G.; Hudson, S. A.; Abell, C. *Biochemistry* **2012**, *51*, 4990–5003.

(79)   Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. *Science (80-.).* **1996**, *274*, 1531–1534.

(80)   Hajduk, P. J.; Sheppard, G.; Nettesheim, D. G.; Olejniczak, E. T.; Shuker, S. B.; Meadows, R. P.; Steinman, D. H.; Carrera, G. M.; Marcotte, P. A.; Severin, J.; et al. *J. Am. Chem. Soc.* **1997**, *119*, 5818–5827.

(81)   Erlanson, D. A.; Fesik, S. W.; Hubbard, R. E.; Jahnke, W.; Jhoti, H. *Nat. Rev. Drug Discov.* **2016**, *15*,

605–619.

(82)   Fuller, N.; Spadola, L.; Cowen, S.; Patel, J.; Schönherr, H.; Cao, Q.; McKenzie, A.; Edfeldt, F.; Rabow, A.; Goodnow, R. *Drug Discov. Today* **2016**, *21*, 1272–1283.

(83)   Murray, C. W.; Rees, D. C. *Nat. Chem.* **2009**, *1*, 187–192.

(84)   Bollag, G.; Hirth, P.; Tsai, J.; Zhang, J.; Ibrahim, P. N.; Cho, H.; Spevak, W.; Zhang, C.; Zhang, Y.; Habets, G.; et al. *Nature* **2010**, *467*, 596–599.

(85)   Souers, A. J.; Leverson, J. D.; Boghaert, E. R.; Ackler, S. L.; Catron, N. D.; Chen, J.; Dayton, B. D.; Ding, H.; Enschede, S. H.; Fairbrother, W. J.; et al. *Nat. Med.* **2013**, *19*, 202–208.

(86)   Loriot, Y.; Necchi, A.; Park, S. H.; Garcia-Donas, J.; Huddart, R.; Burgess, E.; Fleming, M.; Rezazadeh, A.; Mellado, B.; Varlamov, S.; et al. *N. Engl. J. Med.* **2019**, *381*, 338–348.

(87)   Gelderblom, H.; Sande, M. van de. *Futur. Oncol.* **2020**, *16*, 2345–2356.

(88)   Dang, C. V.; Reddy, E. P.; Shokat, K. M.; Soucek, L. *Nat. Rev. Cancer* **2017**, *17*, 502–508.

(89)   Toogood, P. L. *J. Med. Chem.* **2002**, *45*, 1543–1558.

(90)   Roberts, A.; Huang, D. *Clin. Pharmacol. Ther.* **2017**, *101*, 89–98.

(91)   Murray, C. W.; Verdonk, M. L. *J. Comput. Aided. Mol. Des.* **2002**, *16*, 741–753.

(92)   Murray, C. W.; Erlanson, D. A.; Hopkins, A. L.; Keserü, G. M.; Leeson, P. D.; Rees, D. C.; Reynolds, C. H.; Richmond, N. J. *ACS Med. Chem. Lett.* **2014**, *5*, 616–618.

(93)   Ichihara, O.; Barker, J.; Law, R. J.; Whittaker, M. *Mol. Inform.* **2011**, *30*, 298–306.

(94)   Jordan, J. B.; Whittington, D. A.; Bartberger, M. D.; Sickmier, E. A.; Chen, K.; Cheng, Y.; Judd, T. *J. Med. Chem.* **2016**, *59*, 3732–3749.

(95)   Congreve, M.; Carr, R.; Murray, C.; Jhoti, H. *Drug Discov. Today* **2003**, *8*, 876–877.

(96)   Hall, R. J.; Mortenson, P. N.; Murray, C. W. *Prog. Biophys. Mol. Biol.* **2014**, *116*, 82–91.

(97)   Palmer, N.; Peakman, T. M.; Norton, D.; Rees, D. C. *Org. Biomol. Chem.* **2016**, *14*, 1599–1610.

(98)   Murray, C. W.; Rees, D. C. *Angew. Chemie Int. Ed.* **2016**, *55*, 488–492.

(99)   Boyd, S. M.; Turnbull, A. P.; Walse, B. *Comput. Mol. Sci.* **2012**, *2*, 868–885.

(100)   Morley, A. D.; Pugliese, A.; Birchall, K.; Bower, J.; Brennan, P.; Brown, N.; Chapman, T.; Drysdale, M.; Gilbert, I. H.; Hoelder, S.; et al. *Drug Discov. Today* **2013**, *18*, 1221–1227.

(101)   Twigg, D. G.; Kondo, N.; Mitchell, S. L.; Galloway, W. R. J. D.; Sore, H. F.; Madin, A.; Spring, D. R. *Angew. Chemie - Int. Ed.* **2016**, *55*, 12479–12483.

(102)   Mateu, N.; Kidd, S. L.; Kalash, L.; Sore, H. F.; Madin, A.; Bender, A.; Spring, D. R. *Chem. - A Eur. J.* **2018**, *24*, 13681–13687.

(103)   Foley, D. J.; Doveston, R. G.; Churcher, I.; Nelson, A.; Marsden, S. P. *Chem. Commun.* **2015**, *51*, 11174–11177.

(104)   Hassan, H.; Marsden, S. P.; Nelson, A. *Bioorganic Med. Chem.* **2018**, *26*, 3030–3033.

(105)   Troelsen, N. S.; Shanina, E.; Gonzalez-Romero, D.; Danková, D.; Jensen, I. S. A.; Śniady, K. J.; Nami, F.; Zhang, H.; Rademacher, C.; Cuenda, A.; et al. *Angew. Chemie Int. Ed.* **2020**, *59*, 2204–2210.

(106)   Downes, T. D.; Jones, S. P.; Klein, H. F.; Wheldon, M. C.; Atobe, M.; Bond, P. S.; Firth, J. D.; Chan, N. S.; Waddelove, L.; Hubbard, R. E.; et al. *Chem. − A Eur. J.* **2020**, *26*, 8969–8975.

(107)   Hanby, A. R.; Troelsen, N. S.; Osberger, T. J.; Kidd, S. L.; Mortensen, K. T.; Spring, D. R. *Chem. Commun.* **2020**, *56*, 2280–2283.

(108)   Jayaseelan, K. V.; Moreno, P.; Truszkowski, A.; Ertl, P.; Steinbeck, C. *BMC Bioinformatics* **2012**, *13*,

106.

(109)   Hung, A. W.; Ramek, A.; Wang, Y.; Kaya, T.; Wilson, J. A.; Clemons, P. A.; Young, D. W. *Proc. Natl. Acad. Sci.* **2011**, *108*, 6799–6804.

(110)   ChemQuery Search by structure - DrugBank https://www.drugbank.ca/structures/search/small_molecule_drugs/structure#results (accessed Jun **2017**).

(111)   Von Pawel, J.; Jotte, R.; Spigel, D. R.; O'Brien, M. E. R.; Socinski, M. A.; Mezger, J.; Steins, M.; Bosquée, L.; Bubis, J.; Nackaerts, K.; et al. *J. Clin. Oncol.* **2014**, *32*, 4012–4018.

(112)   Möllmann, H.; Hochhaus, G.; Rohatagi, S.; Barth, J.; Derendorf, H. *Pharm. Res. An Off. J. Am. Assoc. Pharm. Sci.* **1995**, *12*, 1096–1100.

(113)   White, P. F.; Way, W. L.; Trevor, A. J. *Anesthesiology* **1982**, *56*, 119–136.

(114)   Podymow, T.; August, P. *Semin. Nephrol.* **2011**, *31*, 70–85.

(115)   Hosoi, S.; Nagao, M.; Tsuda, Y.; Isobe, K.; Sano, T.; Ohta, T. *J. Chem. Soc., Perkin Trans. 1* **2000**, 1505–1511.

(116)   Ohfune, Y.; Shinada, T. *European J. Org. Chem.* **2005**, *2005*, 5127–5143.

(117)   Yang, X.; Toste, F. D. *J. Am. Chem. Soc.* **2015**, *137*, 3205–3208.

(118)   Vilaivan, T.; Bhanthumnavin, W. *Molecules* **2010**, *15*, 917–958.

(119)   Smith, A. M. R.; Hii, K. K. (Mimi). *Chem. Rev.* **2011**, *111*, 1637–1656.

(120)   Zhou, F.; Liao, F.-M.; Yu, J.-S.; Zhou, J. *Synthesis (Stuttg).* **2014**, *46*, 2983–3003.

(121)   Sheikh, N. S.; Leonori, D.; Barker, G.; Firth, J. D.; Campos, K. R.; Meijer, A. J. H. M.; O'Brien, P.; Coldham, I. *J. Am. Chem. Soc.* **2012**, *134*, 5300–5308.

(122)   Rayner, P. J.; O'Brien, P.; Horan, R. A. J. *J. Am. Chem. Soc.* **2013**, *135*, 8071–8077.

(123)   O'Brien, P.; D. Firth, J.; Gelardi, G.; J. Rayner, P.; Stead, D. *Heterocycles* **2018**, *97*, 1288.

(124)   Connon, S. J. *Angew. Chemie - Int. Ed.* **2008**, *47*, 1176–1178.

(125)   Plé, K.; Haudrechy, A.; Probst, N. P. *Tetrahedron* **2010**, *66*, 5030–5035.

(126)   Ilies, M.; Di Costanzo, L.; Dowling, D. P.; Thorn, K. J.; Christianson, D. W. *J Med Chem* **2011**, *54*, 5432–5443.

(127)   Kidd, S. L.; Fowler, E.; Reinhardt, T.; Compton, T.; Mateu, N.; Newman, H.; Bellini, D.; Talon, R.; McLoughlin, J.; Krojer, T.; et al. *Chem. Sci.* **2020**.

(128)   Wermuth, C. G.; Villoutreix, B.; Grisoni, S.; Olivier, A.; Rocher, J.-P. *Strategies in the Search for New Lead Compounds or Original Working Hypotheses*. In *The Practice of Medicinal Chemistry*. 4th ed.; Wermuth, C., Aldous, D., Raboisson, P., Rognan, D., Eds.; Academic Press: London, **2015**, chapter 4, pp 73–96.

(129)   Carreira, E. M.; Fessard, T. C. *Chem. Rev.* **2014**, *114*, 8257–8322.

(130)   Zheng, Y.; Tice, C. M. *Expert Opin. Drug Discov.* **2016**, *11*, 831–834.

(131)   Perron, F.; Albizati, K. F. *Chem. Rev.* **1989**, *89*, 1617–1661.

(132)   *Disinfection, Sterilization, and Preservation*, 5th ed.; Block, S. S., Ed.; Lippincott Williams & Wilkins, **2000**.

(133)   Taylor, F. F.; Faloon, W. W. *J. Clin. Endocrinol. Metab.* **1959**, *19*, 1683–1687.

(134)   Mayol-Llinàs, J.; Farnaby, W.; Nelson, A. *Chem. Commun.* **2017**, *53*, 12345–12348.

(135)   Haftchenary, S.; Nelson, S. D.; Furst, L.; Dandapani, S.; Ferrara, S. J.; Bošković, Ž. V.; Figueroa Lazú,

S.; Guerrero, A. M.; Serrano, J. C.; Crews, D. K.; et al. *ACS Comb. Sci.* **2016**, *18*, 569–574.

(136) Foley, D. J.; Craven, P. G. E.; Collins, P. M.; Doveston, R. G.; Aimon, A.; Talon, R.; Churcher, I.; von Delft, F.; Marsden, S. P.; Nelson, A. *Chem. - A Eur. J.* **2017**, *23*, 15227–15232.

(137) King, T. A.; Stewart, H. L.; Mortensen, K. T.; North, A. J. P.; Sore, H. F.; Spring, D. R. *European J. Org. Chem.* **2019**, *2019*, 5219–5229.

(138) Rice, S.; Cox, D. J.; Marsden, S. P.; Nelson, A. *Tetrahedron* **2019**, *75*, 130513.

(139) Walsh, S. L.; Strain, E. C.; Abreu, M. E.; Bigelow, G. E. *Psychopharmacology (Berl).* **2001**, *157*, 151–162.

(140) Mahmood, I.; Sahajwalla, C. *Clin. Pharmacokinet.* **1999**, *36*, 277–287.

(141) Kim, S.; Ko, H.; Lee, T.; Kim, D. *J. Org. Chem.* **2005**, *70*, 5756–5759.

(142) Brunner, H. R. *Am. J. Hypertens.* **1997**, *10*, 311S–317S.

(143) Altamura, A. C.; Sassella, F.; Santini, A.; Montresor, C.; Fumagalli, S.; Mundo, E. *Drugs* **2003**, *63*, 493–512.

(144) Duffy, R. A.; Morgan, C.; Naylor, R.; Higgins, G. A.; Varty, G. B.; Lachowicz, J. E.; Parker, E. M. *Pharmacol. Biochem. Behav.* **2012**, *102*, 95–100.

(145) Quintás-Cardama, A.; Kantarjian, H.; Cortes, J. *Cancer* **2009**, *115*, 5382–5393.

(146) Hammer, K.; Undheim, K. *Tetrahedron* **1997**, *53*, 2309–2322.

(147) Panov, I.; Drabina, P.; Hanusek, J.; Sedlák, M. *Tetrahedron: Asymmetry* **2011**, *22*, 215–221.

(148) Mazaleyrat, J. P.; Wright, K.; Gaucher, A.; Toulemonde, N.; Dutot, L.; Wakselman, M.; Broxterman, Q. B.; Kaptein, B.; Oancea, S.; Peggion, C.; et al. *Chem. - A Eur. J.* **2005**, *11*, 6921–6929.

(149) Levins, C. G.; Schafmeister, C. E. *J. Am. Chem. Soc.* **2003**, *125*, 4702–4703.

(150) Nielsen, T. E.; Schreiber, S. L. *Angew. Chemie - Int. Ed.* **2008**, *47*, 48–56.

(151) Stotani, S.; Lorenz, C.; Winkler, M.; Medda, F.; Picazo, E.; Ortega Martinez, R.; Karawajczyk, A.; Sanchez-Quesada, J.; Giordanetto, F. *ACS Comb. Sci.* **2016**, *18*, 330–336.

(152) Flodén, N. J.; Trowbridge, A.; Willcox, D.; Walton, S. M.; Kim, Y.; Gaunt, M. J. *J. Am. Chem. Soc.* **2019**, *141*, 8426–8430.

(153) Zheng, Y.; Tice, C. M.; Singh, S. B. *Bioorganic Med. Chem. Lett.* **2014**, *24*, 3673–3682.

(154) Zhang, Z.; Dales, N. A.; Winther, M. D. *J. Med. Chem.* **2014**, *57*, 5039–5056.

(155) Donnell, A. F.; Michoud, C.; Rupert, K. C.; Han, X.; Aguilar, D.; Frank, K. B.; Fretland, A. J.; Gao, L.; Goggin, B.; Heather Hogg, J.; et al. *J. Med. Chem.* **2013**, *56*, 7772–7787.

(156) Seto, S.; Yumoto, K.; Okada, K.; Asahina, Y.; Iwane, A.; Iwago, M.; Terasawa, R.; Shreder, K. R.; Murakami, K.; Kohno, Y. *Bioorg. Med. Chem.* **2012**, *20*, 1188–1200.

(157) Vitaku, E.; Smith, D. T.; Njardarson, J. T. *J. Med. Chem.* **2014**, *57*, 10257–10274.

(158) Napier, C.; Stewart, M.; Melrose, H.; Hopkins, B.; McHarg, A.; Wallis, R. *Eur. J. Pharmacol.* **1999**, *368*, 259–268.

(159) Walsh, C. T.; Wencewicz, T. A. *J. Antibiot. (Tokyo).* **2014**, *67*, 7–22.

(160) Patel, M. R.; Mahaffey, K. W.; Garg, J.; Pan, G.; Singer, D. E.; Hacke, W.; Breithardt, G.; Halperin, J. L.; Hankey, G. J.; Piccini, J. P.; et al. *N. Engl. J. Med.* **2011**, *365*, 883–891.

(161) Morales, R.; Perrier, S.; Florent, J.-M.; Beltra, J.; Dufour, S.; De Mendez, I.; Manceau, P.; Tertre, A.; Moreau, F.; Compere, D.; et al. *J. Mol. Biol.* **2004**, *341*, 1063–1076.

(162) Scholze, P.; Ebert, V.; Sieghart, W. *Eur. J. Pharmacol.* **1996**, *304*, 155–162.

(163) Barbachyn, M. R.; Ford, C. W. *Angew. Chemie Int. Ed.* **2003**, *42*, 2010–2023.

(164) Ismail, F. M. D.; Levitsky, D. O.; Dembitsky, V. M. *Eur. J. Med. Chem.* **2009**, *44*, 3373–3387.

(165) Wakaki, S.; Marumo, H.; Tomioka, K.; Shimizu, G.; Kato, E.; Kamada, H.; Kudo, S.; Fujimoto, Y. *Antibiot. Chemother.* **1958**, *5*, 228–235.

(166) Maanen, M. J. van; Smeets, C. J. M.; Beijnen, J. H. *Cancer Treat. Rev.* **2000**, *26*, 257–268.

(167) Kotha, S.; Sreenivasachary, N. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 257–260.

(168) Kotha, S.; Sreenivasachary, N.; Mohanraja, K.; Durani, S. *Bioorganic Med. Chem. Lett.* **2001**, *11*, 1421–1423.

(169) Kotha, S.; Sreenivasachary, N.; Brahmachary, E. *European J. Org. Chem.* **2001**, *2001*, 787–792.

(170) Kotha, S.; Brahmachary, E. *ChemInform* **2002**, *33*, 192.

(171) Ezquerra, J.; Pedregal, C. *Tetrahedron Lett.* **1993**, *34*, 8535–8538.

(172) López, A.; Moreno-Mañas, M.; Pleixats, R.; Roglans, A.; Ezquerra, J.; Pedregal, C. *Tetrahedron* **1996**, *52*, 8365–8386.

(173) Khan, I. U.; Kattela, S.; Hassan, A.; Correia, C. R. D. *Org. Biomol. Chem.* **2016**, *14*, 9476–9480.

(174) Li, K.; Tan, G.; Huang, J.; Song, F.; You, J. *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 12942–12945.

(175) Andrei, D.; Wnuk, S. F. *Org. Lett.* **2006**, *8*, 5093–5096.

(176) Agranat, I.; Caner, H.; Caldwell, J. *Nat. Rev. Drug Discov.* **2002**, *1*, 753–768.

(177) Ariëns, E. J. *Eur. J. Clin. Pharmacol.* **1984**, *26*, 663–668.

(178) Lenz, W. *Ann. N. Y. Acad. Sci.* **1965**, *123*, 228–236.

(179) Sveiczer, A.; North, A. J. P.; Mateu, N.; Kidd, S. L.; Sore, H. F.; Spring, D. R. *Org. Lett.* **2019**, *21*, 4600–4604.

(180) Harwood, L. M.; Vines, K. J.; Drew, M. G. B. *Synlett* **1996**, *1996*, 1051–1053.

(181) Harwood, L. M.; Tyler, S. N. G.; Drew, M. G. B.; Jahans, A.; Macgilp, I. D. *Arkivoc* **2000**, *v*, 820–831.

(182) Fustero, S.; Mateu, N.; Simón-Fuentes, A.; Aceña, J. L. *Org. Lett.* **2010**, *12*, 3014–3017.

(183) Fustero, S.; Mateu, N.; Albert, L.; Aceña, J. L. *J. Org. Chem.* **2009**, *74*, 4429–4433.

(184) Macritchie, J. A.; Silcock, A.; Willis, C. L. *Tetrahedron: Asymmetry* **1997**, *8*, 3895–3902.

(185) Brown, H. C.; Narasimham, S. *J. Org. Chem.* **1984**, *49*, 3891–3898.

(186) O'Brien, P.; Childs, A. C.; Ensor, G. J.; Hill, C. L.; Kirby, J. P.; Dearden, M. J.; Oxenford, S. J.; Rosser, C. M. *Org. Lett.* **2003**, *5*, 4955–4957.

(187) Jeong, J. U.; Tao, B.; Sagasser, I.; Henniges, H.; Sharpless, K. B. *J. Am. Chem. Soc.* **1998**, *120*, 6844–6845.

(188) Wang, F.; Luo, T.; Hu, J.; Wang, Y.; Krishnan, H. S.; Jog, P. V.; Ganesh, S. K.; Prakash, G. K. S.; Olah, G. A. *Angew. Chemie - Int. Ed.* **2011**, *50*, 7153–7157.

(189) Yang, Z.; Lorenz, J. C.; Shi, Y. *Tetrahedron Lett.* **1998**, *39*, 8621–8624.

(190) Bojase, G.; Nguyen, T. V.; Payne, A. D.; Willis, A. C.; Sherburn, M. S. *Chem. Sci.* **2011**, *2*, 229–232.

(191) Liu, B.; Jin, F.; Wang, T.; Yuan, X.; Han, W. *Angew. Chemie - Int. Ed.* **2017**, *56*, 12712–12717.

(192) Maybridge. Maybridge Ro3 Diversity Fragment Library http://www.maybridge.com/images/pdfs/MB_Ro3_fragment_flyer_2011_EUR_v7.pdf (accessed Jun **2020**).

(193) Chemical; Inc, C. G. Molecular Operating Environment (MOE) | MOEsaic | PSILO

https://www.chemcomp.com/Products.htm (accessed Jun **2020**).

(194) EUCAST. *EUCAST Reading Guide for Broth Microdilution*; **2020**; Vol. 1.0.

(195) Prentice, M. B.; Rahalison, L. *Lancet* **2007**, *369*, 1196–1207.

(196) Stenseth, N. C.; Atshabar, B. B.; Begon, M.; Belmain, S. R.; Bertherat, E.; Carniel, E.; Gage, K. L.; Leirs, H.; Rahalison, L. *PLoS Med.* **2008**, *5*, e3.

(197) Riedemann, N. C.; Guo, R.-F.; Ward, P. A. *J. Clin. Invest.* **2003**, *112*, 460–467.

(198) Bogaert, D.; de Groot, R.; Hermans, P. *Lancet Infect. Dis.* **2004**, *4*, 144–154.

(199) Bodey, G. P.; Bolivar, R.; Fainstein, V.; Jadeja, L. *Clin. Infect. Dis.* **1983**, *5*, 279–313.

(200) Garau, J.; Gomez, L. *Curr. Opin. Infect. Dis.* **2003**, *16*, 135–143.

(201) Lowy, F. D. *N. Engl. J. Med.* **1998**, *339*, 520–532.

(202) Morens, D. M.; Fauci, A. S. *J. Infect. Dis.* **2007**, *195*, 1018–1028.

(203) Hutchings, M. I.; Truman, A. W.; Wilkinson, B. *Curr. Opin. Microbiol.* **2019**, *51*, 72–80.

(204) McDonnell, G.; Russell, A. D. *Clin. Microbiol. Rev.* **1999**, *12*, 147–179.

(205) Nicolaou, K. C.; Rigol, S. *J. Antibiot. (Tokyo).* **2018**, *71*, 153–184.

(206) Sockett, R. E.; Lambert, C. *Nat. Rev. Microbiol.* **2004**, *2*, 669–675.

(207) Hill, C.; Mills, S.; Ross, R. P. *Future Microbiol.* **2018**, *13*, 711–726.

(208) Ghosh, C.; Sarkar, P.; Issa, R.; Haldar, J. *Trends Microbiol.* **2019**, *27*, 323–338.

(209) Parrino, B.; Schillaci, D.; Carnevale, I.; Giovannetti, E.; Diana, P.; Cirrincione, G.; Cascioferro, S. *Eur. J. Med. Chem.* **2019**, *161*, 154–178.

(210) Aroniadis, O. C.; Brandt, L. J. *Gastroenterol. Hepatol. (N. Y).* **2014**, *10*, 230–237.

(211) Pires, D. P.; Cleto, S.; Sillankorva, S.; Azeredo, J.; Lu, T. K. *Microbiol. Mol. Biol. Rev.* **2016**, *80*, 523–543.

(212) Saylor, C.; Dadachova, E.; Casadevall, A. *Vaccine* **2009**, *27*, G38–G46.

(213) Nelson, D. C.; Schmelcher, M.; Rodriguez-Rubio, L.; Klumpp, J.; Pritchard, D. G.; Dong, S.; Donovan, D. M. *Endolysins as Antimicrobials*. In *Advances in Virus Research*; Łobocka, M., Szybalski, W., Eds.; Academic Press Inc., **2012**, Vol. 83, chapter 7, pp 299–365.

(214) Lewies, A.; Du Plessis, L. H.; Wentzel, J. F. *Probiotics Antimicrob. Proteins* **2019**, *11*, 370–381.

(215) Scorciapino, M. *Front. Immunol.* **2012**, *3*, 171.

(216) Hancock, R. E. W.; Haney, E. F.; Gill, E. E. *Nat. Rev. Immunol.* **2016**, *16*, 321–334.

(217) Ayhan, D. H.; Tamer, Y. T.; Akbar, M.; Bailey, S. M.; Wong, M.; Daly, S. M.; Greenberg, D. E.; Toprak, E. *PLOS Biol.* **2016**, *14*, e1002552.

(218) Cotter, P. D.; Ross, R. P.; Hill, C. *Nat. Rev. Microbiol.* **2013**, *11*, 95–105.

(219) Gosio, B. *G. Accad. Med. Torino* **1893**, *61*, 464–487.

(220) Williams, K. *J. R. Soc. Med.* **2009**, *102*, 343–348.

(221) Fleming, A. *Br. J. Exp. Pathol.* **1929**, *10*, 226–236.

(222) Domagk, G. *Angew. Chemie* **1935**, *48*, 657–667.

(223) O'Neill, J. *Antimicrobial Resistance: Tackling a Crisis for the Health and Wealth of Nations*; **2014**.

(224) Bell, G.; MacLean, C. *Trends Microbiol.* **2018**, *26*, 471–483.

(225) Yelin, I.; Kishony, R. *Cell* **2018**, *172*, 1136-1136.e1.

(226) Zampieri, M.; Enke, T.; Chubukov, V.; Ricci, V.; Piddock, L.; Sauer, U. *Mol. Syst. Biol.* **2017**, *13*, 917.

(227) Salverda, M. L. M.; De Visser, J. A. G. M.; Barlow, M. *FEMS Microbiol. Rev.* **2010**, *34*, 1015–1036.

(228) Toprak, E.; Veres, A.; Michel, J.-B.; Chait, R.; Hartl, D. L.; Kishony, R. *Nat. Genet.* **2012**, *44*, 101–105.

(229) Palmer, A. C.; Kishony, R. *Nat. Commun.* **2014**, *5*, 4296.

(230) Liu, Y.-Y.; Wang, Y.; Walsh, T. R.; Yi, L.-X.; Zhang, R.; Spencer, J.; Doi, Y.; Tian, G.; Dong, B.; Huang, X.; et al. *Lancet Infect. Dis.* **2016**, *16*, 161–168.

(231) Ameyama, S.; Onodera, S.; Takahata, M.; Minami, S.; Maki, N.; Endo, K.; Goto, H.; Suzuki, H.; Oishi, Y. *Antimicrob. Agents Chemother.* **2002**, *46*, 3744–3749.

(232) Jaurin, B.; Normark, S. *Cell* **1983**, *32*, 809–816.

(233) Chevereau, G.; Dravecká, M.; Batur, T.; Guvenek, A.; Ayhan, D. H.; Toprak, E.; Bollenbach, T. *PLOS Biol.* **2015**, *13*, e1002299.

(234) Weigel, L. M. *Science (80-.).* **2003**, *302*, 1569–1571.

(235) Hiramatsu, K.; Cui, L.; Kuroda, M.; Ito, T. *Trends Microbiol.* **2001**, *9*, 486–493.

(236) Silver, L. L. *Cold Spring Harb. Perspect. Med.* **2016**, *6*, a030239.

(237) Fleischmann, R.; Adams, M.; White, O.; Clayton, R.; Kirkness, E.; Kerlavage, A.; Bult, C.; Tomb, J.; Dougherty, B.; Merrick, J. *Science (80-.).* **1995**, *269*, 496–512.

(238) Payne, D. J.; Gwynn, M. N.; Holmes, D. J.; Pompliano, D. L. *Nat. Rev. Drug Discov.* **2007**, *6*, 29–40.

(239) Chan, P. F.; Holmes, D. J.; Payne, D. J. *Drug Discov. Today Ther. Strateg.* **2004**, *1*, 519–527.

(240) Rossi, M.; Amaretti, A.; Raimondi, S. *Nutrients* **2011**, *3*, 118–134.

(241) Swarbrick, J.; Iliades, P.; Simpson, J. S.; Macreadie, I. *Open Enzym. Inhib. J.* **2008**, *1*, 12–33.

(242) Young, P. G.; Smith, C. A.; Metcalf, P.; Baker, E. N. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2008**, *64*, 745–753.

(243) Levy, C.; Minnis, D.; Derrick, J. P. *Biochem. J.* **2008**, *412*, 379–388.

(244) Yun, M.-K.; Wu, Y.; Li, Z.; Zhao, Y.; Waddell, M. B.; Ferreira, A. M.; Lee, R. E.; Bashford, D.; White, S. W. *Science (80-.).* **2012**, *335*, 1110–1114.

(245) Bourne, C. *Antibiotics* **2014**, *3*, 1–28.

(246) Bermingham, A.; Derrick, J. P. *BioEssays* **2002**, *24*, 637–648.

(247) Schwalbe, C. H.; Cody, V. *Crystallogr. Rev.* **2006**, *12*, 267–300.

(248) Zhang, Y.-M.; Rock, C. O. *Nat. Rev. Microbiol.* **2008**, *6*, 222–233.

(249) Parsons, J. B.; Rock, C. O. *Curr. Opin. Microbiol.* **2011**, *14*, 544–549.

(250) Tong, L. *Cell. Mol. Life Sci.* **2013**, *70*, 863–891.

(251) Polyak, S. W.; Abell, A. D.; Wilce, M. C. J.; Zhang, L.; Booker, G. W. *Appl. Microbiol. Biotechnol.* **2012**, *93*, 983–992.

(252) Parsons, J. B.; Rock, C. O. *Prog. Lipid Res.* **2013**, *52*, 249–276.

(253) Menzies, D. *Infect. Drug Resist.* **2011**, *4*, 129.

(254) Russell, A. D. *J. Antimicrob. Chemother.* **2004**, *53*, 693–695.

(255) Vilchèze, C.; Wang, F.; Arai, M.; Hazbón, M. H.; Colangeli, R.; Kremer, L.; Weisbrod, T. R.; Alland, D.; Sacchettini, J. C.; Jacobs, W. R. *Nat. Med.* **2006**, *12*, 1027–1029.

(256) Heath, R. J.; Rubin, J. R.; Holland, D. R.; Zhang, E.; Snow, M. E.; Rock, C. O. *J. Biol. Chem.* **1999**, *274*, 11110–11114.

(257) Park, H. S.; Yoon, Y. M.; Jung, S. J.; Kim, C. M.; Kim, J. M.; Kwak, J.-H. *J. Antimicrob. Chemother.* **2007**, *60*, 568–574.

(258) Yao, J.; Rock, C. O. *Cold Spring Harb. Perspect. Med.* **2016**, *6*, a027045.

(259) Park, H. S.; Yoon, Y. M.; Jung, S. J.; Yun, I. N. R.; Kim, C. M.; Kim, J. M.; Kwak, J.-H. *Int. J. Antimicrob. Agents* **2007**, *30*, 446–451.

(260) Kaplan, N.; Awrey, D.; Bardouniotis, E.; Berman, J.; Yethon, J.; Pauls, H. W.; Hafkin, B. *J. Chemother.* **2013**, *25*, 18–25.

(261) Erwin, A. L. *Cold Spring Harb. Perspect. Med.* **2016**, *6*, a025304.

(262) Theuretzbacher, U.; Outterson, K.; Engel, A.; Karlén, A. *Nat. Rev. Microbiol.* **2020**, *18*, 275–285.

(263) Macfarlane, G. T.; Macfarlane, S. *Scand. J. Gastroenterol.* **1997**, *32*, 3–9.

(264) Dolan, S. K.; Wijaya, A.; Geddis, S. M.; Spring, D. R.; Silva-Rocha, R.; Welch, M. *Microbiology* **2018**, *164*, 251–259.

(265) Horswill, A. R.; Dudding, A. R.; Escalante-Semerena, J. C. *J. Biol. Chem.* **2001**, *276*, 19094–19101.

(266) Maruyama, K.; Kitamura, H. *J. Biochem.* **1985**, *98*, 819–824.

(267) Rocco, C. J.; Escalante-Semerena, J. C. *J. Bacteriol.* **2010**, *192*, 771–778.

(268) Textor, S.; Wendisch, V. F.; Graaf, A. A. De; Müller, U.; Linder, M. I.; Linder, D.; Buckel, W. *Arch. Microbiol.* **1997**, *168*, 428–436.

(269) Munoz-Elias, E. J.; Upton, A. M.; Cherian, J.; McKinney, J. D. *Mol. Microbiol.* **2006**, *60*, 1109–1122.

(270) Krebs, H. A. *Lancet* **1937**, *230*, 736–738.

(271) Akram, M. *Cell Biochem. Biophys.* **2014**, *68*, 475–478.

(272) Kornberg, H. L.; Krebs, H. A. *Nature* **1957**, *179*, 988–991.

(273) Dolan, S. K.; Welch, M. *Annu. Rev. Microbiol.* **2018**, *72*, 309–330.

(274) VanderVen, B. C.; Fahey, R. J.; Lee, W.; Liu, Y.; Abramovitch, R. B.; Memmott, C.; Crowe, A. M.; Eltis, L. D.; Perola, E.; Deininger, D. D.; et al. *PLoS Pathog.* **2015**, *11*, e1004679.

(275) Collins, P. M.; Ng, J. T.; Talon, R.; Nekrosiute, K.; Krojer, T.; Douangamath, A.; Brandao-Neto, J.; Wright, N.; Pearce, N. M.; von Delft, F. *Acta Crystallogr. Sect. D Struct. Biol.* **2017**, *73*, 246–255.

(276) Flores, M. A.; Bode, J. W. *Org. Lett.* **2010**, *12*, 1924–1927.

(277) Baker, L. M.; Aimon, A.; Murray, J. B.; Surgenor, A. E.; Matassova, N.; Roughley, S. D.; Collins, P. M.; Krojer, T.; von Delft, F.; Hubbard, R. E. *Commun. Chem.* **2020**, *3*, 1–11.

(278) Liu, H.; Jiang, G.; Pan, X.; Wan, X.; Lai, Y.; Ma, D.; Xie, W. *Org. Lett.* **2014**, *16*, 1908–1911.

(279) Richter, M. F.; Hergenrother, P. J. *Ann. N. Y. Acad. Sci.* **2019**, *1435*, 18–38.

(280) Espino, C. G.; Du Bois, J. *Angew. Chemie Int. Ed.* **2001**, *40*, 598–600.

(281) Sun, W. S.; Park, Y. S.; Yoo, J.; Park, K. D.; Kim, S. H.; Kim, J.-H.; Park, H.-J. *J. Med. Chem.* **2003**, *46*, 5619–5627.

(282) Vivoli, M.; Novak, H. R.; Littlechild, J. A.; Harmer, N. J. *J. Vis. Exp.* **2014**, No. 91, 51809.

(283) Srere, P. A.; Brazil, H.; Gonen, L.; Takahashi, M. *Acta Chem. Scand.* **1963**, *17 supl.*, 129–134.

6 Appendix

6.1 Computational Analysis

Calculation of the energy minimised conformations for both libraries were performed with Molecular Operating Environment (MOE) software package version 2012.10 using the search settings summarised in Table 6.1.

Table 6.1 Conformational search settings

Force field	MMFF94x
Solvation	Born
Method	LowModeMD
Rejection Limit	100
RMS Gradient	0.005
Iteration Limit	10000
MM Iteration Limit	500
RMSD Limit	0.15
Energy Window	3
Conformation Limit	100

Structural and physicochemical parameters calculated using MOE 2018.0602 are summarised in Table 6.2.

Table 6.2 Physicochemical parameters calculated by MOE.

Parameter	Description	Property[a]
npr1	Normalised PMI ratio (1) (PMI1 / PMI3)	-
npr2	Normalised PMI ratio (2) (PMI2 / PMI3)	-
a_acc	Number of hydrogen-bond acceptor atoms	HBA
a_aro	Number of aromatic atoms	-
a_don	Number of hydrogen-bond donor atoms	HBD
a_heavy	Number of non-hydrogen heavy atoms	-
b_rotN	Number of rotatable bonds	RBC
chiral	Number of chiral centres	chiral
SlogP	Log octanol/water partition coefficient	SlogP
TPSA	Topological polar surface area (Å^2)	TPSA
weight	Molecular weight (Da)	MW

[a]as appear in Table 2.4

The number of sp^3 atoms (sp3-Atom) were calculated using Osiris Datawarrior version 4.7.3.

Further properties calculated using Microsoft Excel 2010 are summarised in Table 6.3.

Table 6.3 Physicochemical properties calculated by Excel.

Parameter[a]	Description
Fsp^3	Fraction of sp^3 atoms (sp3-Atom / a_heavy)
Far	Fraction of aromatic atoms (a_aro / a_heavy)
npr1 + npr2	Sum of the normalised PMI ratios
Fflat	Fraction of molecules below the 'flatland' line (defined as: npr1 + npr2 ≤ 1.1)

[a]as appear in Table 2.5

6.1.1 Spirocyclic library

The spirocyclic library is based on the fragments described in 2.3.7. When applicable, protecting groups were removed yielding compounds numbered in general as X'. Normalised PMI ratios and molecular formulae of the library are summarised in Table 6.4, whereas distributions of the key physicochemical properties are displayed as histograms in Figure 6.1.

Table 6.4 Normalised PMI ratios and molecular formulae of the spirocyclic library.

Compound	SMILES	npr1	npr2	Formula
8b	O=C1[C@]2(NC(=O)C1)CC=CCC2	0.5336	0.9195	$C_9H_{11}NO_2$
21	O=C1OC[C@]2(N1)CC=CCC2	0.3346	0.9505	$C_8H_{11}NO_2$
37	O=C1OC[C@]2(NC1)CC=CCC2	0.2861	0.9367	$C_9H_{13}NO_2$
39	c1(C=2OC[C@]3(N=2)CC=CCC3)ccccc1	0.2010	0.9075	$C_{14}H_{15}NO$
18	O=C1N[C@@]2(COC1)CC=CCC2	0.3961	0.8850	$C_9H_{13}NO_2$
38	NC=1OC[C@]2(N=1)CC=CCC2	0.3275	0.9491	$C_8H_{12}N_2O$
8a	O=C1C2(NC(=O)C1)CC=CC2	0.3879	0.8086	$C_8H_9NO_2$
8c	O=C1[C@]2(NC(=O)C1)CC=CCCC2	0.5826	0.9781	$C_{10}H_{13}NO_2$
43	O(C)c1ccc(C=2[C@]3(NC(=O)C=2)CC=CCC3)cc1	0.2737	0.8730	$C_{16}H_{17}NO_2$
44	O=C1N[C@@]2(C(O)C1)CC=CCC2	0.4161	0.8470	$C_9H_{13}NO_2$
40	O=C1OC[C@]2(N=C1)CC=CCC2	0.2499	0.8763	$C_9H_{11}NO_2$
41	O(CC)C=1[C@]2(NC(=O)C=1)CC=CCC2	0.4464	0.7134	$C_{11}H_{15}NO_2$
45	O=C1N[C@@]2(C=C1)CC=CCC2	0.3282	0.9689	$C_9H_{11}NO$
33b'	O=C1OCC2(N1)CCC(=O)CC2	0.2504	0.9771	$C_8H_{11}NO_3$
33a'	O=C1OC[C@]2(N1)CC(=O)CCC2	0.2870	0.8968	$C_8H_{11}NO_3$
34b'	O=C1OC[C@]2(N1)C[C@@H](O)CCC2	0.3619	0.9290	$C_8H_{13}NO_3$
34a'	O=C1OCC2(N1)CCC(O)CC2	0.2614	0.9788	$C_8H_{13}NO_3$
34c'	O=C1OCC2(N1)CCC(O)CC2	0.2614	0.9789	$C_8H_{13}NO_3$
29a'	O=C1OC[C@]2(N1)C[C@@H](O)[C@@H](O)CC2	0.2965	0.9561	$C_8H_{13}NO_4$
29b'	O=C1OC[C@]2(N1)C[C@H](O)[C@H](O)CC2	0.3059	0.9261	$C_8H_{13}NO4$
35	Br[C@@H]1[C@@H](Br)CC[C@@]2(NC(=O)OC2)C1	0.3131	0.7957	$C_8H_{11}NO_2Br_2$

145

30b'	Br[C@H]1[C@H](Br)CC[C@@]2(NC(=O)OC2)C1	0.3871	0.7196	$C_8H_{11}NO_2Br_2$
27a	O=C1OC[C@]2(N1)C[C@H]1O[C@H]1CC2	0.3455	0.9534	$C_8H_{11}NO_3$
32a'	FC1(F)[C@H]2[C@@H]1CC[C@@]1(NC(=O)OC1)C2	0.3173	0.8961	$C_9H_{11}NO_2F_2$
27b	O=C1OC[C@]2(N1)C[C@@H]1O[C@@H]1CC2	0.4577	0.9443	$C_8H_{11}NO_3$
32b'	FC1(F)[C@@H]2[C@H]1CC[C@@]1(NC(=O)OC1)C2	0.3725	0.9335	$C_9H_{11}NO_2F_2$
31a'	O=C1OC[C@]2(N1)C[C@H]1N[C@H]1CC2	0.2416	0.9675	$C_8H_{12}N_2O_2$
31b'	O=C1OC[C@]2(N1)C[C@@H]1N[C@@H]1CC2	0.4516	0.9439	$C_8H_{12}N_2O_2$
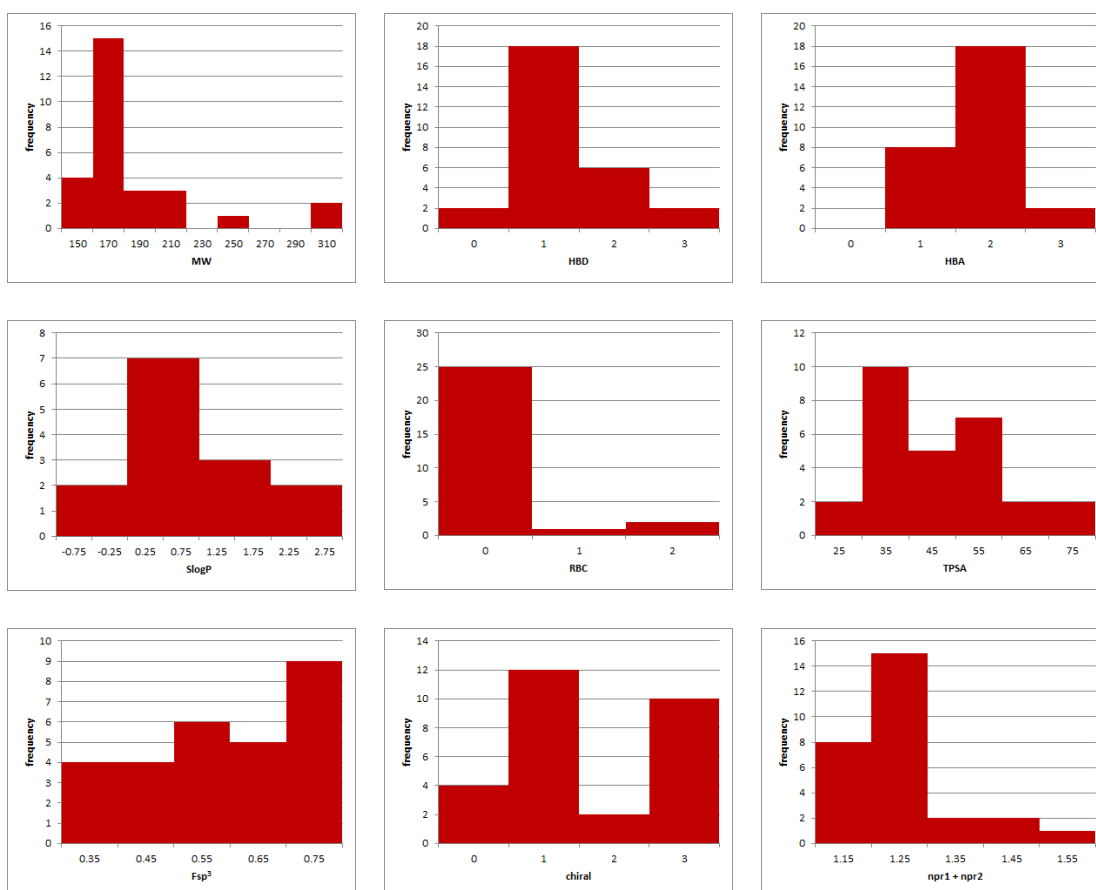
Figure 6.1 Distribution of key physicochemical properties of the spirocyclic library.

146

6.1.2 Maybridge 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'

The best-matched fragments were chosen based on heavy atom and hetero atom counts compared to the spirocyclic library. For heteroatom counts of 2 and 3, only exact heavy atom matches (i.e. same number of *N* and *O* atoms) were used, whereas for heteroatom counts of 4 and 5 no exact matches were found and therefore only the total heteroatom counts were used.

Normalised PMI ratios and molecular formulae of the Maybridge best-matched fragments are summarised in Table 6.5.

Table 6.5 Normalised PMI ratios and formulae of the Maybridge best-matched fragments.

SMILES	npr1	npr2	Formula
OCCNCc1ccccc1	0.1566	0.9438	$C_9H_{13}NO$
Oc1c2c(nccc2)ccc1	0.3618	0.6382	C_9H_7NO
O=C(C)c1cc(C#N)ccc1	0.2475	0.7561	C_9H_7NO
O=C1Nc2c(cccc2)CC1	0.2480	0.7704	C_9H_9NO
OC[C@H](N)Cc1ccccc1	0.2401	0.9471	$C_9H_{13}NO$
Oc1cc2ncccc2cc1	0.2253	0.7747	C_9H_7NO
O(C)c1cc(CC#N)ccc1	0.2339	0.8611	C_9H_9NO
O(C)c1cc2c([nH]cc2)cc1	0.2076	0.7963	C_9H_9NO
NCc1cc2c(OCC2)cc1	0.2047	0.8484	$C_9H_{11}NO$
N#Cc1cc2c(occ2)cc1	0.1741	0.8259	C_9H_5NO
Oc1c(C)cc(C#N)cc1C	0.3298	0.6775	C_9H_9NO
c1(-c2ccccc2)ocnc1	0.1600	0.8400	C_9H_7NO
OCCc1ccc(C#N)cc1	0.1608	0.9466	C_9H_9NO
O=C(N)c1c(C)c(C)ccc1	0.3472	0.7208	$C_9H_{11}NO$
O=C(N)c1cc(C)c(C)cc1	0.2338	0.7872	$C_9H_{11}NO$
OC[C@H]1[C@@H](NCc2ccccc2)CCCC1	0.1717	0.9024	$C_{14}H_{21}NO$
O[C@@H]([C@@H](N)c1ccccc1)c1ccccc1	0.4721	0.7911	$C_{14}H_{15}NO$
O(c1c(CNC)cccc1)c1ccccc1	0.3960	0.7590	$C_{14}H_{15}NO$
O(c1ccc(CNC)cc1)c1ccccc1	0.1063	0.9808	$C_{14}H_{15}NO$
O(Cc1cc(CN)ccc1)c1ccccc1	0.1368	0.9586	$C_{14}H_{15}NO$
O=C1CC2N(Cc3ccccc3)C(C1)CC2	0.2342	0.9468	$C_{14}H_{17}NO$
O=C(OCc1ccccc1)N	0.1520	0.9775	$C_8H_9NO_2$
O=C(N)c1ccc(OC)cc1	0.1434	0.8704	$C_8H_9NO_2$
O=C1NC(=O)[C@@H]2[C@H]1CC=CC2	0.3972	0.8514	$C_8H_9NO_2$
O=C(OCC)[C@H]1[C@@H](N)CCC1	0.2814	0.8759	$C_8H_{15}NO_2$
O=C(C)c1c(O)cc(N)cc1	0.2658	0.7756	$C_8H_9NO_2$
O=C1OCc2c1cc(N)cc2	0.2879	0.7163	$C_8H_7NO_2$
O=C(N(C)C)C1CCOCC1	0.3110	0.9079	$C_8H_{15}NO_2$

O=C(N)Cc1ccc(O)cc1	0.1807	0.9326	$C_8H_9NO_2$
Nc1cc2c(OCOC2)cc1	0.2340	0.7841	$C_8H_9NO_2$
N#Cc1cc2OCOc2cc1	0.1844	0.8194	$C_8H_5NO_2$
O=C(OC)c1c(C)[nH]c(C)c1	0.2785	0.7317	$C_8H_{11}NO_2$
O=C(N)COc1ccccc1	0.1227	0.8802	$C_8H_9NO_2$
O=C(OC)c1cc(N)ccc1	0.2016	0.8018	$C_8H_9NO_2$
O=C(O)c1c(N)cc(C)cc1	0.2578	0.7459	$C_8H_9NO_2$
O(C)c1cc2nc[nH]c2cc1	0.2120	0.7920	$C_8H_8N_2O$
O=C1N(C)N=C(C(C)(C)C)C1	0.2896	0.8515	$C_8H_{14}N_2O$
Oc1nc(C(C)C)nc(C)c1	0.4024	0.7353	$C_8H_{12}N_2O$
OCC1=Cn2c(ncc2)C=C1	0.1881	0.8499	$C_8H_8N_2O$
OCc1cc2nc[nH]c2cc1	0.2025	0.8439	$C_8H_8N_2O$
O=C1NN=Cc2c1cccc2	0.3495	0.6505	$C_8H_6N_2O$
N#CCCNCC1OCCC1	0.0719	0.9648	$C_8H_{14}N_2O$
OC1(C#N)C2CCN(C1)CC2	0.5090	0.8849	$C_8H_{12}N_2O$
Oc1c(C#N)c(C)cc(C)n1	0.3510	0.6569	$C_8H_8N_2O$
OCc1nc(CCCC)[nH]c1	0.1723	0.8988	$C_8H_{14}N_2O$
O=C(CC#N)N1CCCCC1	0.1898	0.8750	$C_8H_{12}N_2O$
O=C(NCc1cnccc1)C	0.2305	0.9262	$C_8H_{10}N_2O$
O=C(O)c1cc2c([nH]cc2)cc1	0.1798	0.8202	$C_9H_7NO_2$
O=C(OCC)c1ccc(N)cc1	0.1503	0.8545	$C_9H_{11}NO_2$
O=C(Nc1ccc(OC)cc1)C	0.1073	0.8974	$C_9H_{11}NO_2$
O=C(OCC)[C@H]1[C@H](N)CC=CC1	0.3119	0.8862	$C_9H_{15}NO_2$
O=C(OCC)c1c(C)cc(C)[nH]1	0.2959	0.7147	$C_9H_{13}NO_2$
O(C(C)(C)C)C(=O)N1CC=CC1	0.2214	0.8897	$C_9H_{15}NO_2$
O=C(OCc1ccccc1)CN	0.1301	0.9955	$C_9H_{11}NO_2$
O=C(OC)c1cc(N)c(C)cc1	0.1827	0.8227	$C_9H_{11}NO_2$
O=C1C(CCC#N)C(=O)CCC1	0.3067	0.7990	$C_9H_{11}NO_2$
O=C(C)N1CCC(C(=O)C)CC1	0.2330	0.8951	$C_9H_{15}NO_2$
O=C(OC)c1cc(CN)ccc1	0.2111	0.8248	$C_9H_{11}NO_2$
O=C(OC)c1cc(C#N)ccc1	0.2563	0.7466	$C_9H_7NO_2$
O=C(OCC)c1c(C)[nH]c(C)c1	0.2602	0.7501	$C_9H_{13}NO_2$
O=C(OC)c1c(C#N)cccc1	0.3960	0.6518	$C_9H_7NO_2$
O=Nc1c(O)ccc2c1cccc2	0.3216	0.6978	$C_{10}H_7NO_2$
O=C(Oc1c2c([nH]c1)cccc2)C	0.2915	0.7757	$C_{10}H_9NO_2$
O=C1O[C@H]([C@@H](C)N1)c1ccccc1	0.2327	0.9131	$C_{10}H_{11}NO_2$
O=C(OC)c1cc2c([nH]cc2)cc1	0.1679	0.8345	$C_{10}H_9NO_2$
O=C1Oc2c(C(C)=C1)ccc(N)c2	0.2982	0.7046	$C_{10}H_9NO_2$
NCc1c2OCCCOc2ccc1	0.4024	0.6688	$C_{10}H_{13}NO_2$
NCc1cc2OCCCOc2cc1	0.2122	0.8505	$C_{10}H_{13}NO_2$
OCc1noc(-c2ccccc2)c1	0.1157	0.9080	$C_{10}H_9NO_2$
OCc1onc(-c2ccccc2)c1	0.1294	0.9000	$C_{10}H_9NO_2$
OCC1N(Cc2occc2)CCC1	0.2605	0.8751	$C_{10}H_{15}NO_2$
O=C1OC[C@H](Cc2ccccc2)N1	0.1333	0.9562	$C_{10}H_{11}NO_2$
OCCN(CCO)c1ccccc1	0.4901	0.7449	$C_{10}H_{15}NO_2$
O(C)c1c(OC)cc2c(c1)CNCC2	0.3130	0.7029	$C_{11}H_{15}NO_2$
O=C1OC(C)(C)[C@@H](c2ccccc2)N1	0.2965	0.8657	$C_{11}H_{13}NO_2$
O=C(NCC(=O)C)Cc1ccccc1	0.2460	0.8973	$C_{11}H_{13}NO_2$

O=C(OC)c1ncc2c(c1)cccc2	0.1409	0.8612	$C_{11}H_9NO_2$
O=C(C)c1c(N)c(CCC)c(O)cc1	0.3040	0.7637	$C_{11}H_{15}NO_2$
O=C(O)CC1(CN(C)C)CCCCC1	0.4960	0.7257	$C_{11}H_{21}NO_2$
OCc1c(C)onc1-c1ccccc1	0.2664	0.7941	$C_{11}H_{11}NO_2$
O(CCN(C)C)c1c(CO)cccc1	0.1946	0.8300	$C_{11}H_{17}NO_2$
O(C(C)(C)C)C(=O)c1cc(N)ccc1	0.1824	0.8858	$C_{11}H_{15}NO_2$
O=C(Nc1ccc(OCC=C)cc1)C	0.0702	0.9422	$C_{11}H_{13}NO_2$
O=C(Nc1cc2c(C(=O)CC2)cc1)C	0.1551	0.8506	$C_{11}H_{11}NO_2$
O=C(N(C)C)c1c-2c(C(=O)c3c-2cccc3)ccc1	0.4483	0.6742	$C_{16}H_{13}NO_2$
O(C)c1ccc(CNCc2ccc(OC)cc2)cc1	0.1178	0.9595	$C_{16}H_{19}NO_2$
O=C(CC12CC3CC(C1)CC(C2)C3)N1CCOCC1	0.1592	0.9586	$C_{16}H_{25}NO_2$
FC(F)(F)c1ccc(CO)cc1	0.1638	0.9439	$C_8H_7OF_3$
Clc1cc(OCC(=O)O)ccc1	0.1310	0.8709	$C_8H_7O_3Cl$
Fc1cc2C(=O)C(=O)Nc2cc1	0.2277	0.7723	$C_8H_4NO_2F$
Fc1c(NC(=O)C)ccc(F)c1	0.2136	0.7891	$C_8H_7NOF_2$
Clc1c(NC(=O)C)c(F)ccc1	0.4058	0.6018	C_8H_7NOClF
Fc1c(OC(=O)C)ccc(F)c1	0.2273	0.8861	$C_8H_6O_2F_2$
S(CC)c1c(C(=O)O)cccn1	0.3754	0.6717	$C_8H_9NO_2S$
ClC1=NNC(=O)c2c1cccc2	0.4180	0.5820	$C_8H_5N_2OCl$
O=C(O)c1sc2ncccc2c1	0.1817	0.8183	$C_8H_5NO_2S$
NCc1nc(-c2sccc2)sc1	0.2239	0.7930	$C_8H_8N_2S_2$
OCc1noc(-c2sccc2)c1	0.1229	0.9019	$C_8H_7NO_2S$
O=C(O)c1n(C)c2c(scc2)c1	0.2212	0.7815	$C_8H_7NO_2S$
OCc1n[nH]c(-c2sccc2)c1	0.1134	0.8956	$C_8H_8N_2OS$
N(Cc1scc2OCCOc12)C	0.3228	0.7224	$C_8H_{11}NO_2S$
S(=O)(=O)(C)c1ccc(C#N)cc1	0.1549	0.9370	$C_8H_7NO_2S$
S(C)c1sc2c(n1)ccc(N)c2	0.1548	0.8474	$C_8H_8N_2S_2$
S=C1NCN(C2CCCC2)CN1	0.1229	0.9184	$C_8H_{15}N_3S$
Clc1nc(-c2occc2)ccn1	0.1873	0.8127	$C_8H_5N_2OCl$
Clc1sc(C2=NN(C)CC2)cc1	0.1058	0.9037	$C_8H_9N_2ClS$
Clc1c(Cl)ccc(NC(=O)C)c1	0.1982	0.8037	$C_8H_7NOCl_2$
Clc1ccc(SCC(=O)O)cc1	0.0971	0.9480	$C_8H_7O_2ClS$
Clc1c(C)c(C#N)c(O)nc1C	0.3474	0.6582	$C_8H_7N_2OCl$
Clc1cc(C(=O)OC)c(O)cc1	0.2323	0.7701	$C_8H_7O_3Cl$
Fc1ccc(CNC(=O)N)cc1	0.1160	0.9905	$C_8H_9N_2OF$
FC(F)(F)c1c(CO)cccc1	0.4573	0.6836	$C_8H_7OF_3$
Clc1c(CO)nc(CCCC)[nH]1	0.2479	0.8232	$C_8H_{13}N_2OCl$
S=C(NN)NC1C2C=CC(C1)C2	0.3102	0.8552	$C_8H_{13}N_3S$
Clc1cc(Cl)cc(OCC#N)c1	0.3151	0.6866	$C_8H_5NOCl_2$
S(CC#N)c1c(F)cc(F)cc1	0.1607	0.9087	$C_8H_5NF_2S$
S=C(Nc1c(OC)cccc1)N	0.2920	0.7332	$C_8H_{10}N_2OS$
Clc1c(C(=O)OC)ccc(F)c1	0.3007	0.7813	$C_8H_6O_2ClF$
O=C1NN=C(c2sccc2)CC1	0.1465	0.8724	$C_8H_8N_2OS$
FC(F)(F)c1cc(N)c(OC)cc1	0.2119	0.8552	$C_8H_8NOF_3$
Fc1cc(F)cc(C(O)C(=O)O)c1	0.3356	0.8586	$C_8H_6O_3F_2$
FC(F)(F)c1nc(C)c(C#N)cc1	0.1955	0.8740	$C_8H_5N_2F_3$
Clc1ccc(CNS(=O)(=O)C)cc1	0.0957	0.9716	$C_8H_{10}NO_2ClS$
O=C(O)c1nc(-c2sccc2)sc1	0.1809	0.8191	$C_8H_5NO_2S_2$

S(=O)(=O)(N)c1cc2c(cc1)COC2	0.1890	0.8939	$C_8H_9NO_3S$
Clc1sc(C(OC(C)(C)C)=O)cn1	0.1276	0.9275	$C_8H_{10}NO_2ClS$
ClC=1C(=O)C(Cl)=CN(CCC#N)C=1	0.3168	0.7079	$C_8H_6N_2OCl_2$
Clc1cc(NC(=S)N)c(OC)cc1	0.3658	0.6540	$C_8H_9N_2OClS$
FC(F)(F)c1cnc(N(C)C)cc1	0.1554	0.9094	$C_8H_9N_2F_3$
FC(F)(F)Oc1ccc(CO)cc1	0.1400	0.9484	$C_8H_7O_2F_3$
FC(F)(F)Oc1ccc(CC#N)cc1	0.1403	0.9705	$C_9H_6NOF_3$
Fc1c(NC(=O)C)c(C#N)cc(F)c1	0.3497	0.6673	$C_9H_6N_2OF_2$
Fc1c(F)ccc(-c2nc(N)sc2)c1	0.1461	0.8637	$C_9H_6N_2F_2S$
O=C(O)c1sc(-c2nc(C)sc2)cc1	0.1087	0.8927	$C_9H_7NO_2S_2$
Fc1c(N(C)C)c(F)cc(C(=O)N)c1	0.2561	0.7567	$C_9H_{10}N_2OF_2$
Clc1c(F)ccc(NC(=O)CSC)c1	0.1720	0.8946	$C_9H_9NOClFS$
O=S1(=O)CCN(Cc2sccc2)CC1	0.1538	0.9913	$C_9H_{13}NO_2S_2$
S(C)c1c(C(=O)C)c(C)c(C(=O)O)s1	0.3663	0.7309	$C_9H_{10}O_3S_2$
Clc1c(F)ccc(N2C(=O)C=CS2)c1	0.2071	0.7929	$C_9H_5NOClFS$
Clc1c(F)c(N2C(=O)C=CS2)ccc1	0.2380	0.7645	$C_9H_5NOClFS$
Clc1sc([SH0](=O)C)c2C(=O)CCCc12	0.4492	0.6241	$C_9H_9O_2ClS_2$
Clc1c(Cl)cccc1NC(=O)N(C)C	0.2101	0.8023	$C_9H_{10}N_2OCl_2$
Clc1ccc(CCNC(=S)NN)cc1	0.1770	0.9074	$C_9H_{12}N_3ClS$
Clc1cc2C(=O)CCS(=O)(=O)c2cc1	0.2936	0.7708	$C_9H_7O_3ClS$
Clc1ccc(S(=O)(=O)CCC#N)cc1	0.1325	0.9591	$C_9H_8NO_2ClS$
Fc1cc2C(=O)CCS(=O)(=O)c2cc1	0.3716	0.7099	$C_9H_7O_3FS$
FC(F)(F)c1cc(OCC#N)ccc1	0.1718	0.8767	$C_9H_6NOF_3$

6.2 Crystallographic Data

(5*R**,7*S**)-7-Hydroxy-1-(4-methoxy-benzyl)-3-oxa-1-azaspiro[4.5]decan-2-one **34b**
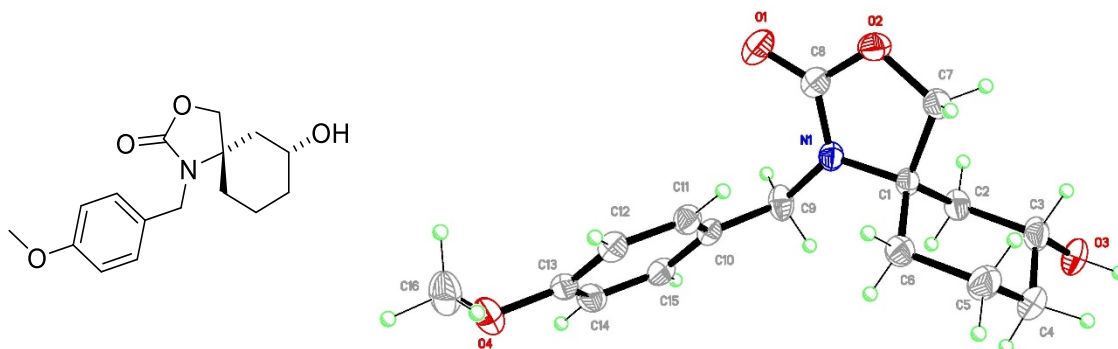
Identification code	DS_B1_0022	CCDC	1912287
Empirical formula	$C_{16}H_{21}NO_4$	Formula weight (Da)	291.34
Temperature (K)	180(2)	Wavelength (Å)	1.54178
Crystal system	Monoclinic	Space group	$P\ 2_1/c$
Unit cell lengths (Å)	a = 9.1715(2) b = 6.6200(2) c = 23.5770(6)	Unit cell angles (°)	α = 90 β = 92.5461(12) γ = 90
Volume (Å^3)	1430.07	Z	4
Density calculated (gcm^{-3})	1.353	Absorption coefficient (mm^{-1})	0.795
F(000)	624	Crystal size (mm^3)	0.220 × 0.100 × 0.040
θ range for data coll. (°)	3.753 – 67.040	Completeness to θ = 67.040°	99.7%
Reflections collected	15488	Independent reflections	2545
Index ranges	–10 ≤ h ≤ 10 –7 ≤ k ≤ 7 –28 ≤ l ≤ 28	Refinement method	Full-matrix least-squares on F^2
Absorption correction	Multi-scan	Max/min. transmission	0.969/0.845
Data/restraints/parameters	2545/2/204		
Goodness of fit F^2	1.306	Largest diff. peak/hole (eÅ^{-3})	0.245/-0.286
Final R indices [I > 2σ(I)]	R1 = 0.0654 wR2 = 0.1474	R indices (all data)	R1 = 0.0692 wR2 = 0.1490

(5*R**,7*S**,8*R**)-7,8-Dihydroxy-1-(4-methoxybenzyl)-3-oxa-1-azaspiro[4.5]-decan-2-one **29b**
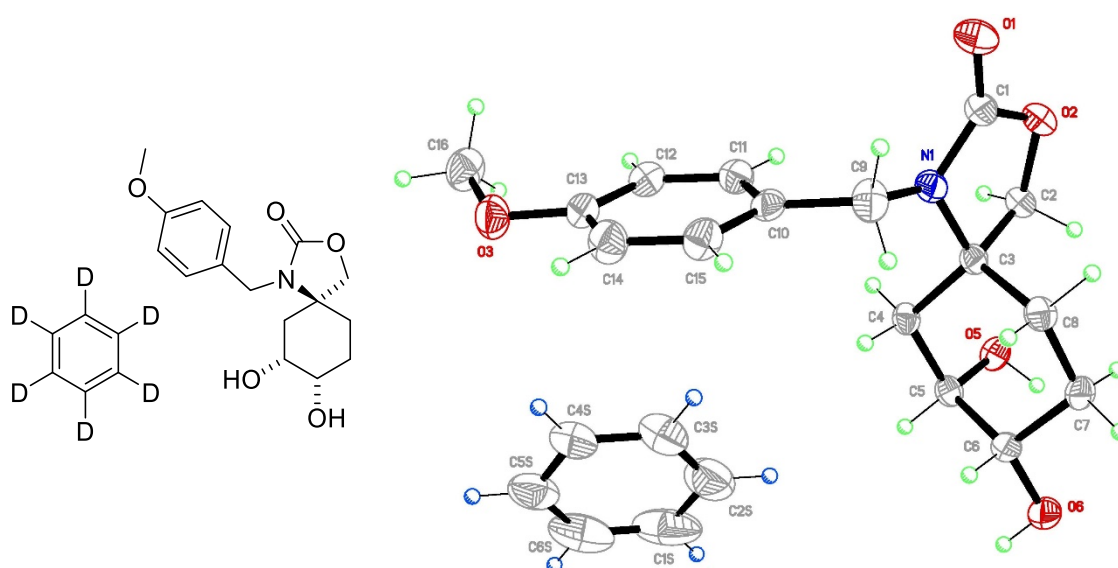
Identification code	DS_B1_0015	CCDC	1912286
Empirical formula	$C_{22}H_{21}D_6NO_5$	Formula weight (Da)	391.48
Temperature (K)	180(2)	Wavelength (Å)	1.54178
Crystal system	Triclinic	Space group	$P\bar{1}$
Unit cell lengths (Å)	a = 6.2880(2) b = 7.9649(3) c = 20.1941(8)	Unit cell angles (°)	α = 82.118(3) β = 87.204(2) γ = 82.302(2)
Volume (Å³)	992.33(6)	Z	2
Density calculated (gcm⁻³)	1.310	Absorption coefficient (mm⁻¹)	0.743
F(000)	412	Crystal size (mm³)	0.120 × 0.120 × 0.020
θ range for data coll. (°)	2.210 – 66.855	Completeness to θ = 66.855°	99.5%
Reflections collected	13031	Independent reflections	3530
Index ranges	–7 ≤ h ≤ 7 –9 ≤ k ≤ 9 –23 ≤ l ≤ 24	Refinement method	Full-matrix least-squares on F^2
Absorption correction	Multi-scan	Max/min. transmission	0.985/0.916
Data/restraints/parameters	3530/0/261		
Goodness of fit F^2	1.037	Largest diff. peak/hole (eÅ⁻³)	0.234/-0.181
Final R indices [I > 2σ(I)]	R1 = 0.0434 wR2 = 0.0890	R indices (all data)	R1 = 0.0681 wR2 = 0.0988

(1*R**,3*R**,6*S**)-7,7-Difluoro-3'-(4-methoxybenzyl)spiro-[bicycle[4.1.0]heptane-3,4'-oxazolidin]-2'-one **32a**
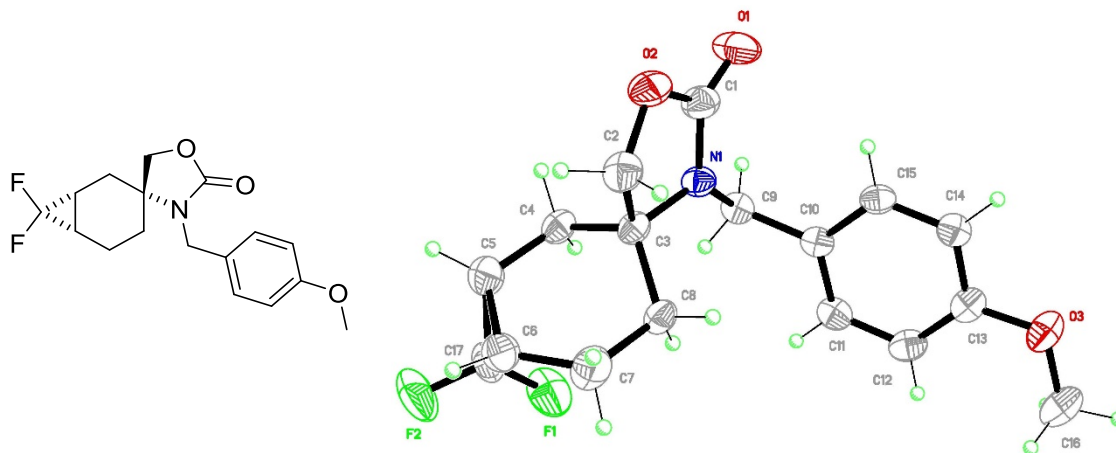
Identification code	DS_B1_0018	CCDC	1912284
Empirical formula	$C_{17}H_{19}F_2NO_3$	Formula weight (Da)	323.33
Temperature (K)	180(2)	Wavelength (Å)	1.54178
Crystal system	Monoclinic	Space group	P 2_1/c
Unit cell lengths (Å)	a = 13.5002(5) b = 12.3831(5) c = 9.9272(4)	Unit cell angles (°)	α = 90 β = 110.910(2) γ = 90
Volume (Å³)	1550.28(11)	Z	4
Density calculated (gcm⁻³)	1.385	Absorption coefficient (mm⁻¹)	0.934
F(000)	680	Crystal size (mm³)	0.300 × 0.180 × 0.120
θ range for data coll. (°)	3.505 – 66.845	Completeness to θ = 66.845°	99.6%
Reflections collected	11811	Independent reflections	2739
Index ranges	−16 ≤ h ≤ 16 −14 ≤ k ≤ 10 −11 ≤ l ≤ 11	Refinement method	Full-matrix least-squares on F²
Absorption correction	Multi-scan	Max/min. transmission	0.896/0.767
Data/restraints/parameters	2739/0/228		
Goodness of fit F²	1.112	Largest diff. peak/hole (eÅ⁻³)	0.239/-0.193
Final R indices [I > 2σ(I)]	R1 = 0.0416 wR2 = 0.0990	R indices (all data)	R1 = 0.0453 wR2 = 0.1012

(1*R**,3*R**,6*S**)-3'-(4-Methoxybenzyl)-7-tosyl-7-azaspiro[bicyclo[4.1.0]heptane-3,4'-oxazolidin]-2'-one **31b**
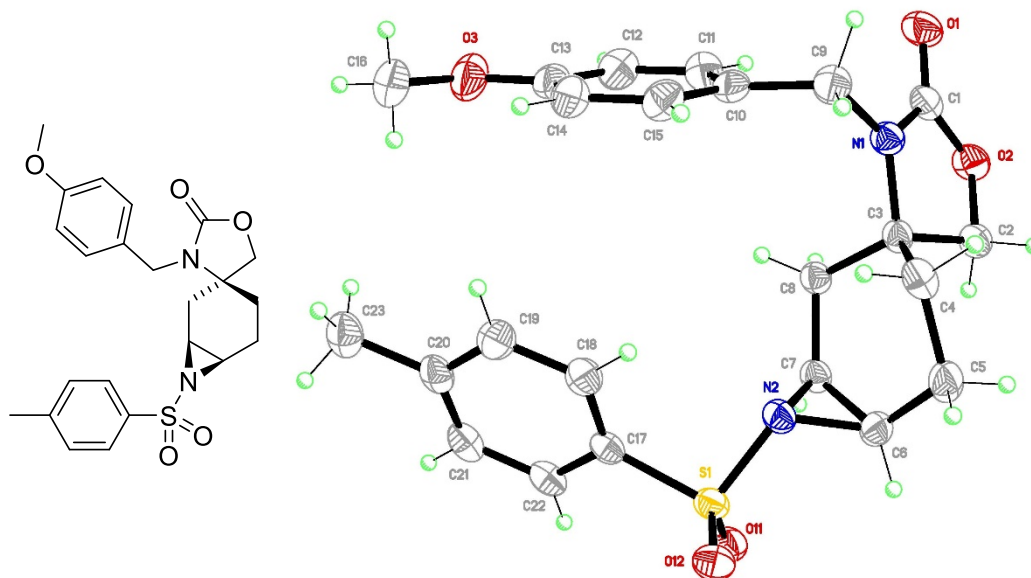
Identification code	DS_B1_0014	CCDC	1912283
Empirical formula	$C_{23}H_{26}N_2O_5S$	Formula weight (Da)	442.52
Temperature (K)	180(2)	Wavelength (Å)	1.54178
Crystal system	Triclinic	Space group	$P\bar{1}$
Unit cell lengths (Å)	a = 7.0012(2) b = 12.6065(4) c = 12.9625(4)	Unit cell angles (°)	α = 80.1740(10) β = 75.6130(10) γ = 76.4970(10)
Volume (Å^3)	1069.83(6)	Z	2
Density calculated (gcm^{-3})	1.374	Absorption coefficient (mm^{-1})	1.668
F(000)	468	Crystal size (mm^3)	0.250 × 0.200 × 0.150
θ range for data coll. (°)	3.545 – 66.774	Completeness to θ = 66.774°	98.7%
Reflections collected	8306	Independent reflections	3741
Index ranges	–7 ≤ h ≤ 8 –15 ≤ k ≤ 13 –14 ≤ l ≤ 15	Refinement method	Full-matrix least-squares on F^2
Absorption correction	Multi-scan	Max/min. transmission	0.7886/0.681
Data/restraints/parameters	3741/0/282		
Goodness of fit F^2	1.038	Largest diff. peak/hole (eÅ^{-3})	0.308/-0.418
Final R indices [I > 2σ(I)]	R1 = 0.0353 wR2 = 0.0894	R indices (all data)	R1 = 0.0403 wR2 = 0.0931

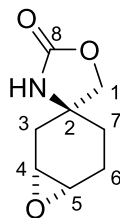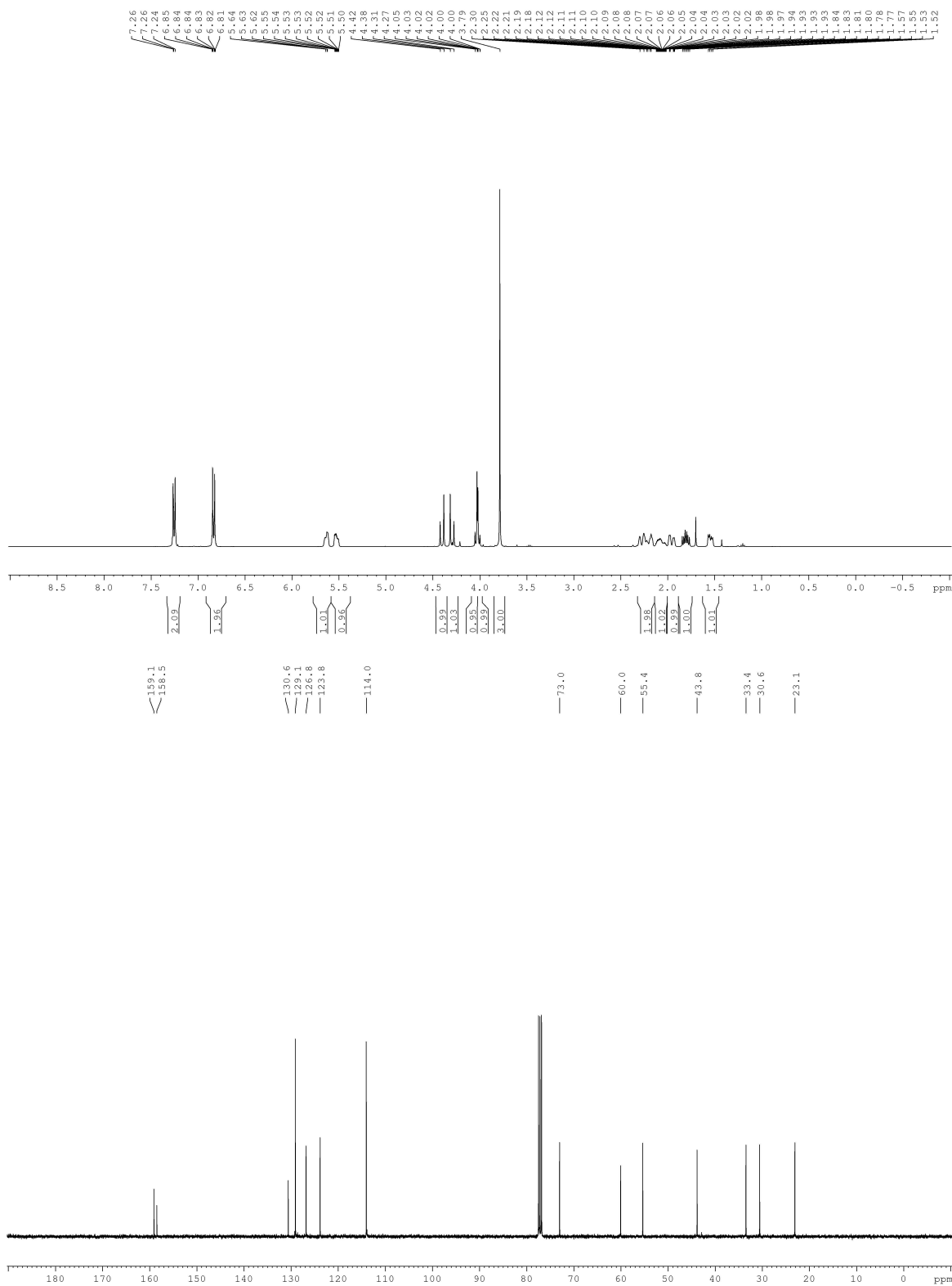(1*R**,3*R**,6*S**)-7-Oxaspiro[bicyclo[4.1.0]heptane-3,4'-oxazolidin]-2'-one **27a**
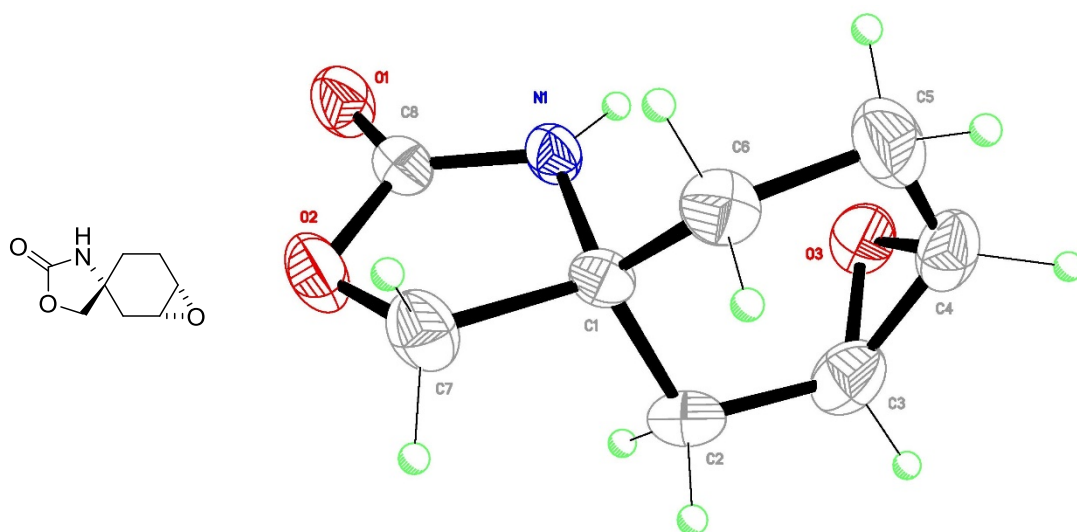
|			
Identification code	DS_B1_0021	CCDC	1912289
Empirical formula	$C_8H_{11}NO_3$	Formula weight (Da)	169.18
Temperature (K)	180(2)	Wavelength (Å)	1.54178
Crystal system	Monoclinic	Space group	$P\,2_1/n$
Unit cell lengths (Å)	a = 5.7530(2) b = 12.8809(4) c = 10.6497(3)	Unit cell angles (°)	α = 90 β = 92.918(2) γ = 90
Volume (Å^3)	788.16(4)	Z	4
Density calculated (gcm^{-3})	1.426	Absorption coefficient (mm^{-1})	0.919
F(000)	360	Crystal size (mm^3)	0.250 × 0.080 × 0.070
θ range for data coll. (°)	5.393 – 66.842	Completeness to θ = 66.842°	99.6%
Reflections collected	5783	Independent reflections	1394
Index ranges	–7 ≤ h ≤ 8 –15 ≤ k ≤ 13 –14 ≤ l ≤ 15	Refinement method	Full-matrix least-squares on F^2
Absorption correction	Multi-scan	Max/min. transmission	0.7886/0.681
Data/restraints/parameters	1394/18/132		
Goodness of fit F^2	1.137	Largest diff. peak/hole (eÅ^{-3})	0.187/-0.199
Final R indices [I > 2σ(I)]	R1 = 0.0415 wR2 = 0.1001	R indices (all data)	R1 = 0.0477 wR2 = 0.1035

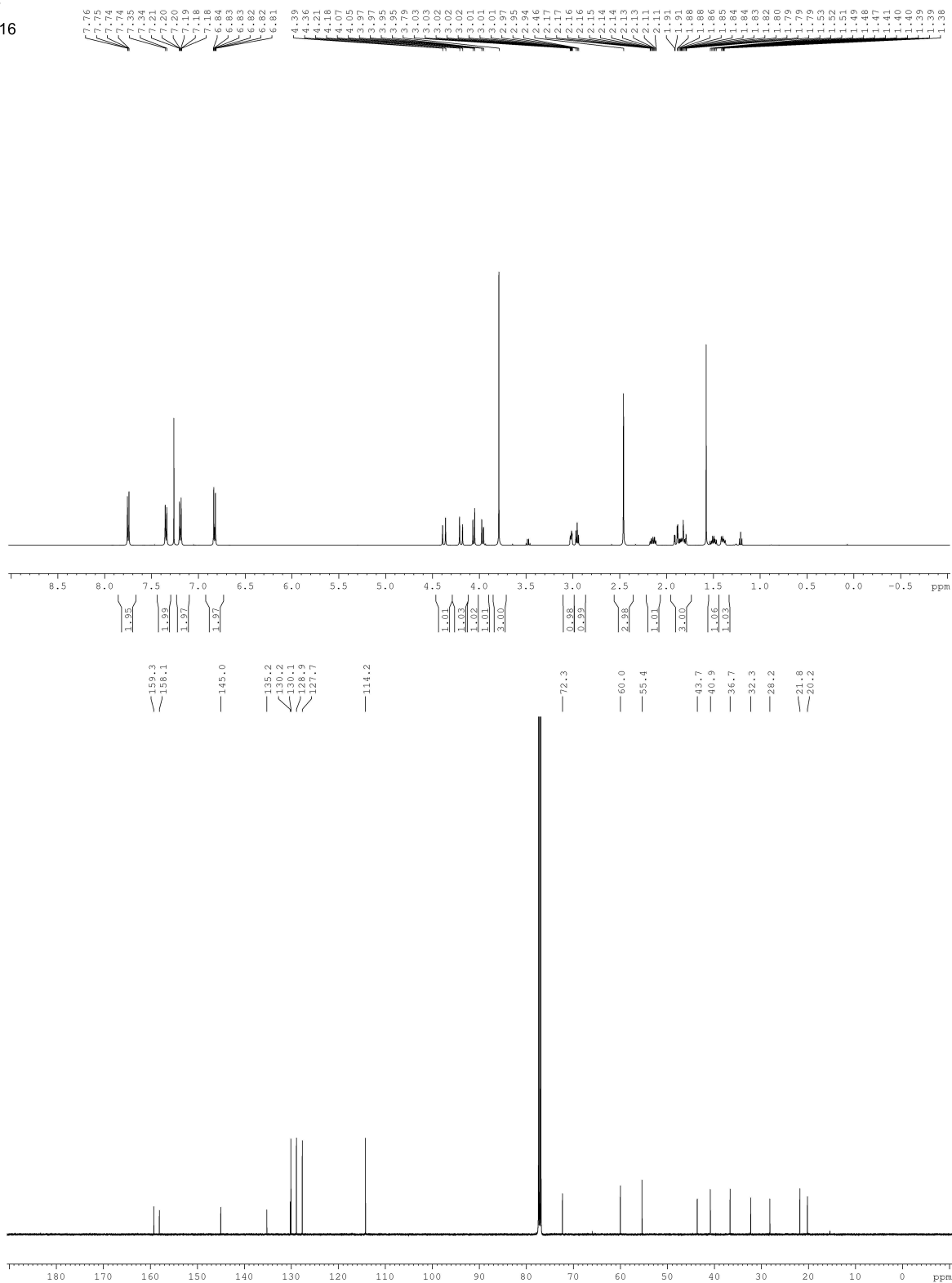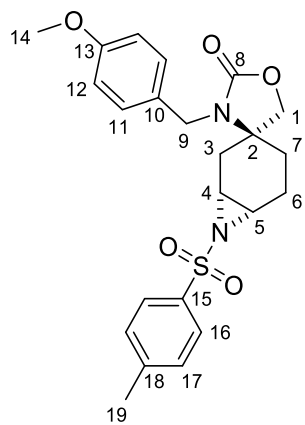(5*R**,7*S**,8*S**)-7,8-Dibromo-1-(4-methoxybenzyl)-3-oxa-1-azaspiro[4.5]decan-2-one **30a**
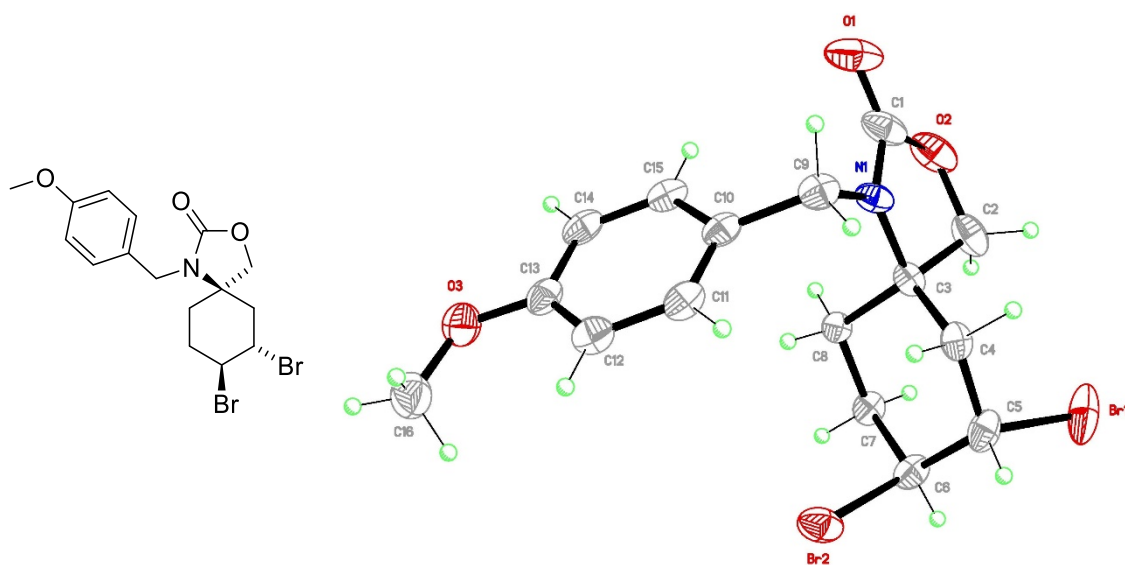
Identification code	DS_B1_0019	CCDC	1912285
Empirical formula	$C_{16}H_{19}Br_2NO_3$	Formula weight (Da)	433.14
Temperature (K)	180(2)	Wavelength (Å)	1.54178
Crystal system	Monoclinic	Space group	P 2_1/n
Unit cell lengths (Å)	a = 6.4606(2) b = 12.5480(3) c = 20.3610(6)	Unit cell angles (°)	α = 90 β = 98.4700(10) γ = 90
Volume (Å^3)	1632.61(8)	Z	4
Density calculated (gcm^{-3})	1.762	Absorption coefficient (mm^{-1})	6.403
F(000)	864	Crystal size (mm^3)	0.140 × 0.140 × 0.140
θ range for data coll. (°)	4.151 – 66.735	Completeness to θ = 66.735°	99.5%
Reflections collected	17954	Independent reflections	2886
Index ranges	–7 ≤ h ≤ 6 –14 ≤ k ≤ 13 –21 ≤ l ≤ 24	Refinement method	Full-matrix least-squares on F^2
Absorption correction	Multi-scan	Max/min. transmission	0.468/0.468
Data/restraints/parameters	2886/0/201		
Goodness of fit F^2	1.092	Largest diff. peak/hole (eÅ^{-3})	0.546/-0.461
Final R indices [I > 2σ(I)]	R1 = 0.0243 wR2 = 0.0573	R indices (all data)	R1 = 0.0264 wR2 = 0.0581

(5*R**,7*S**,8*S**)-7,8-Dibromo-3-oxa-1-azaspiro[4.5]decan-2-one **35**
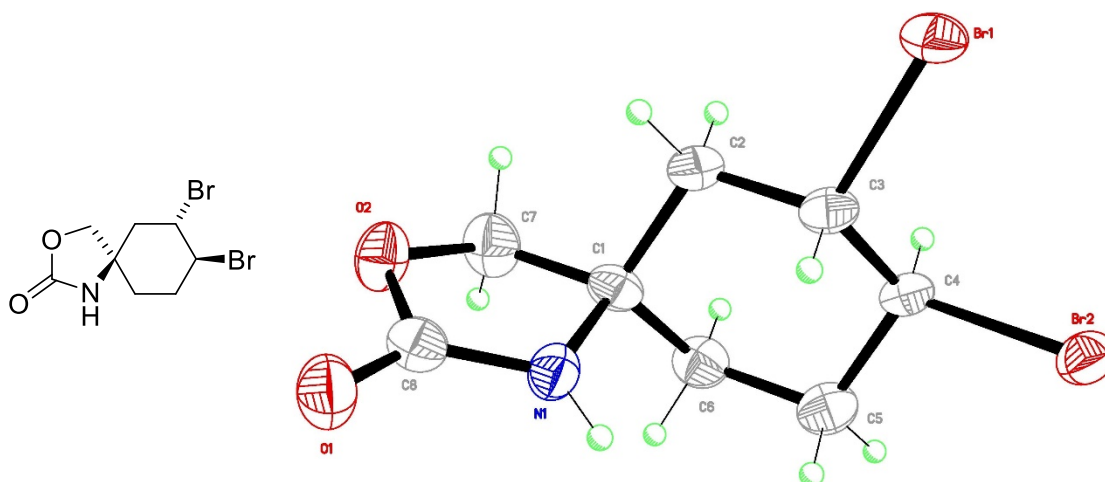
Identification code	DS_B1_0026	CCDC	1912288
Empirical formula	$C_8H_{11}Br_2NO_2$	Formula weight (Da)	313.00
Temperature (K)	180(2)	Wavelength (Å)	1.54178
Crystal system	Monoclinic	Space group	P 2_1/c
Unit cell lengths (Å)	a = 13.0067(12) b = 6.2766(6) c = 12.8006(10)	Unit cell angles (°)	α = 90 β = 103.002(6) γ = 90
Volume (Å^3)	1018.22(16)	Z	4
Density calculated (gcm^{-3})	2.042	Absorption coefficient (mm^{-1})	9.863
F(000)	608	Crystal size (mm^3)	0.300 × 0.040 × 0.010
θ range for data coll. (°)	3.486 – 66.672	Completeness to θ = 66.672°	99.7%
Reflections collected	13072	Independent reflections	1806
Index ranges	–15 ≤ h ≤ 15 –7 ≤ k ≤ 6 –14 ≤ l ≤ 15	Refinement method	Full-matrix least-squares on F^2
Absorption correction	Multi-scan	Max/min. transmission	0.908/0.156
Data/restraints/parameters	1806/0/118		
Goodness of fit F^2	1.046	Largest diff. peak/hole (eÅ^{-3})	0.945/-0.788
Final R indices [I > 2σ(I)]	R1 = 0.0528 wR2 = 0.1174	R indices (all data)	R1 = 0.0884 wR2 = 0.1336

157

6.3 NMR Spectra

Ethyl 2-allyl-2-aminopent-4-enoate **1a**
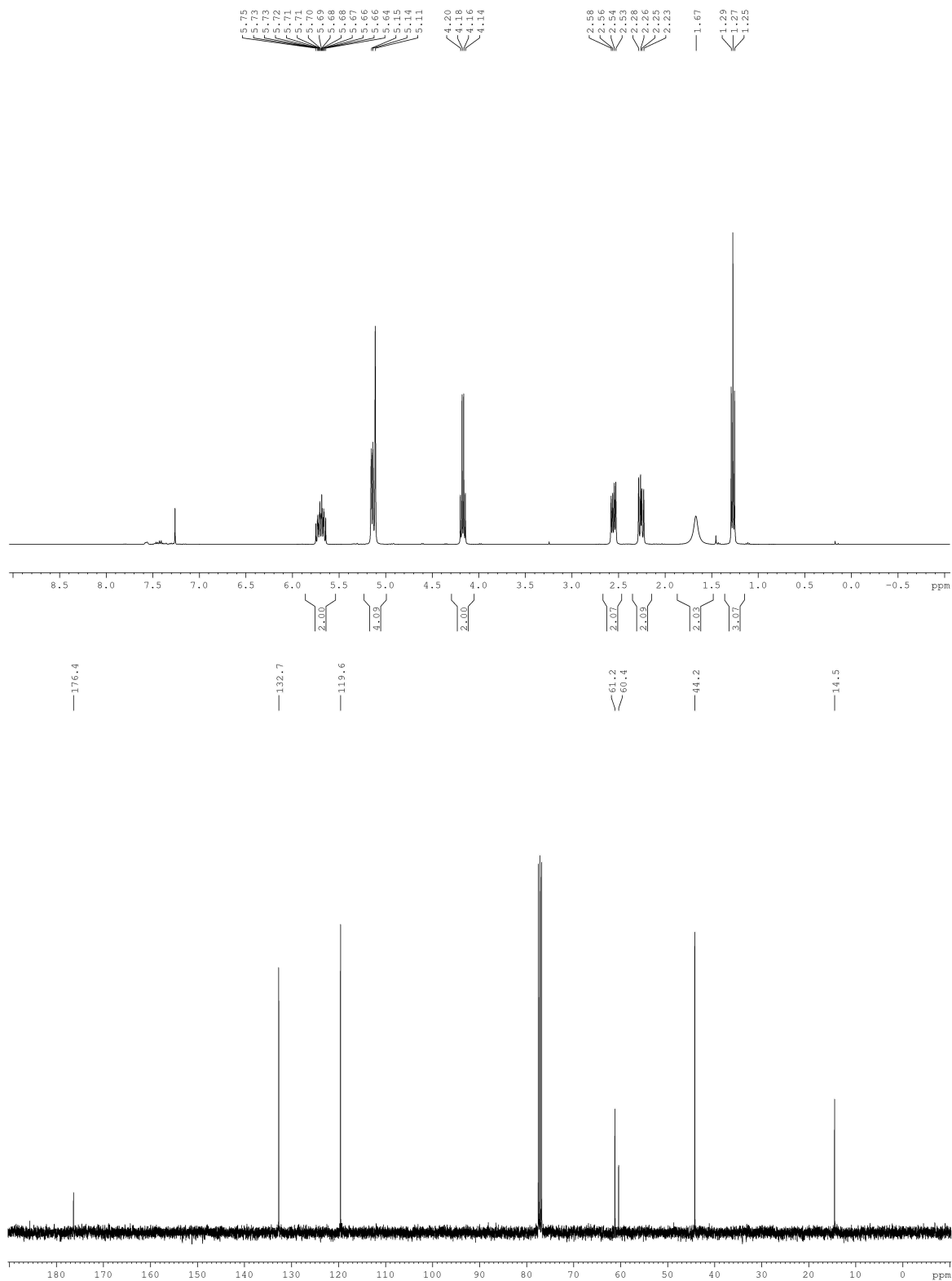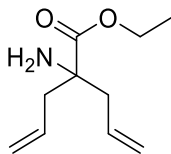
Ethyl 2-allyl-2-aminohex-5-enoate **1b**
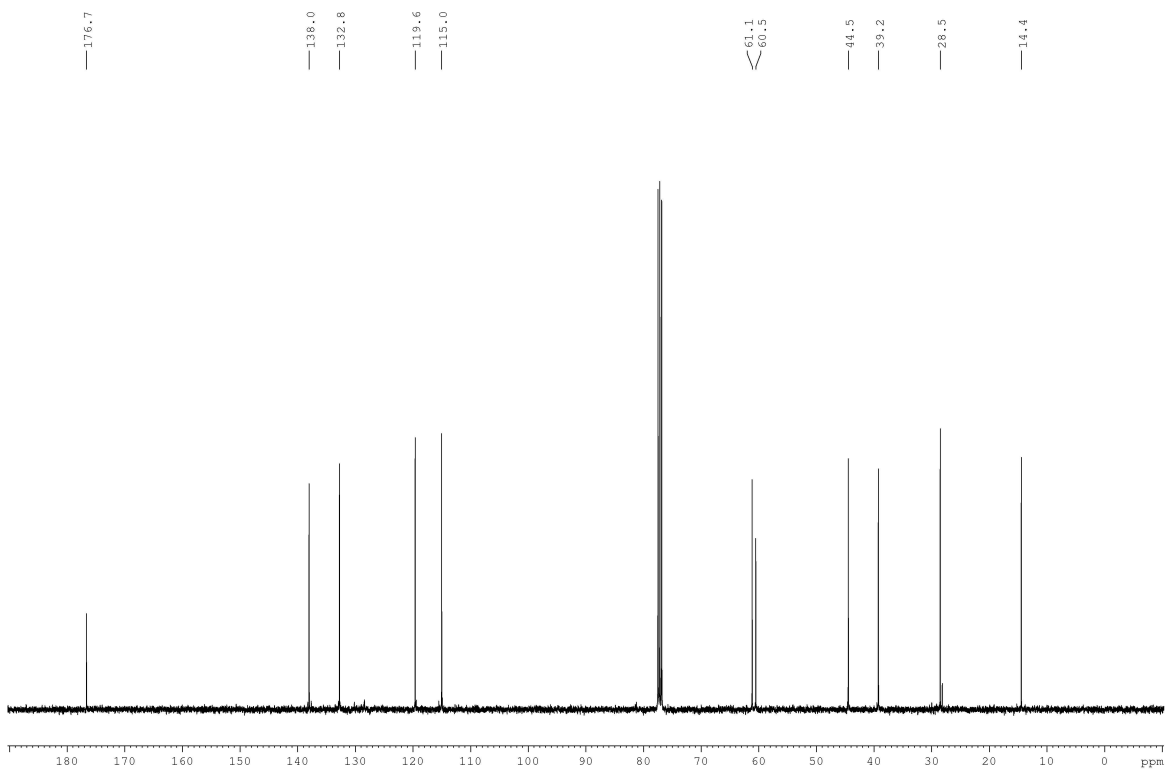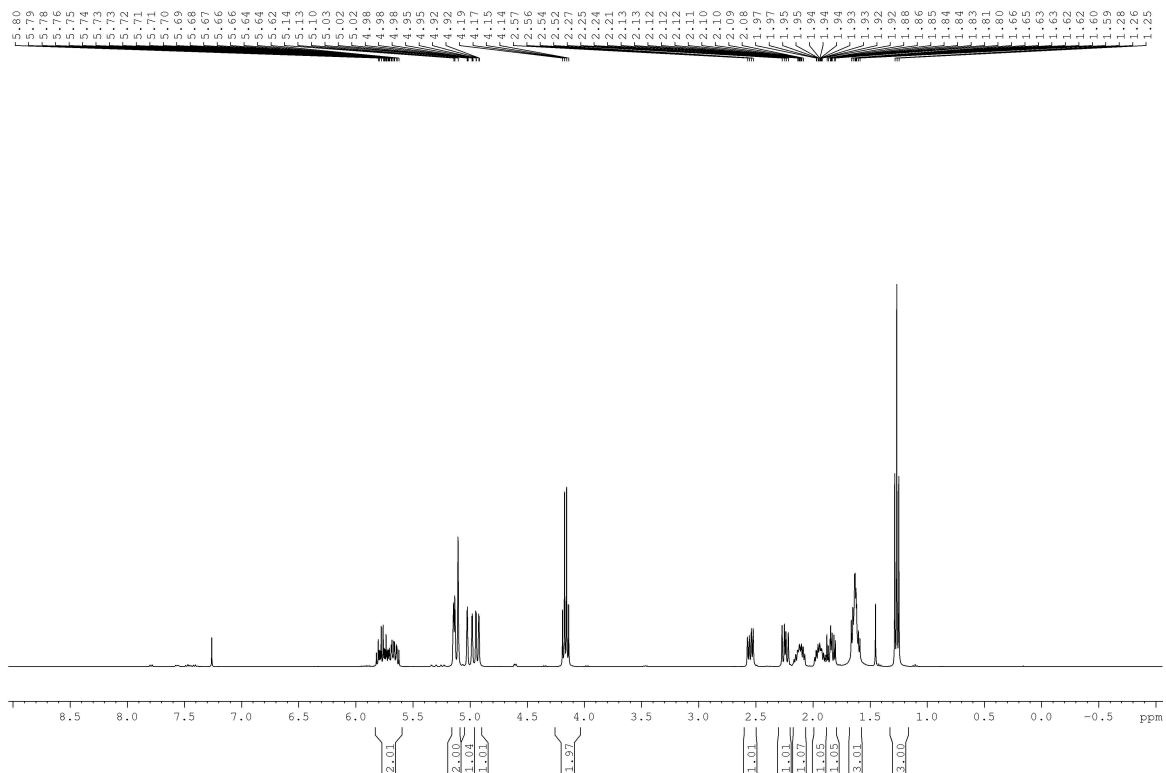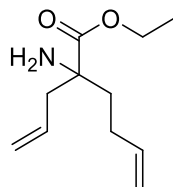
Ethyl 2-allyl-2-aminohept-6-enoate **1c**
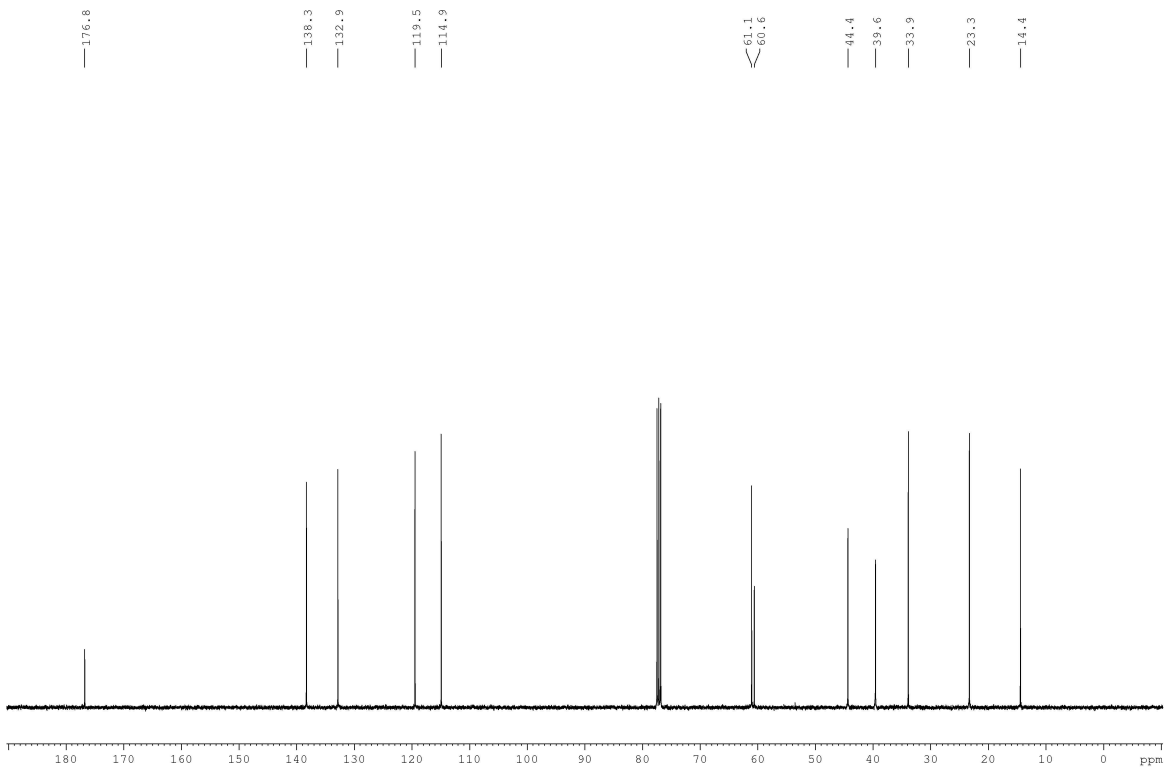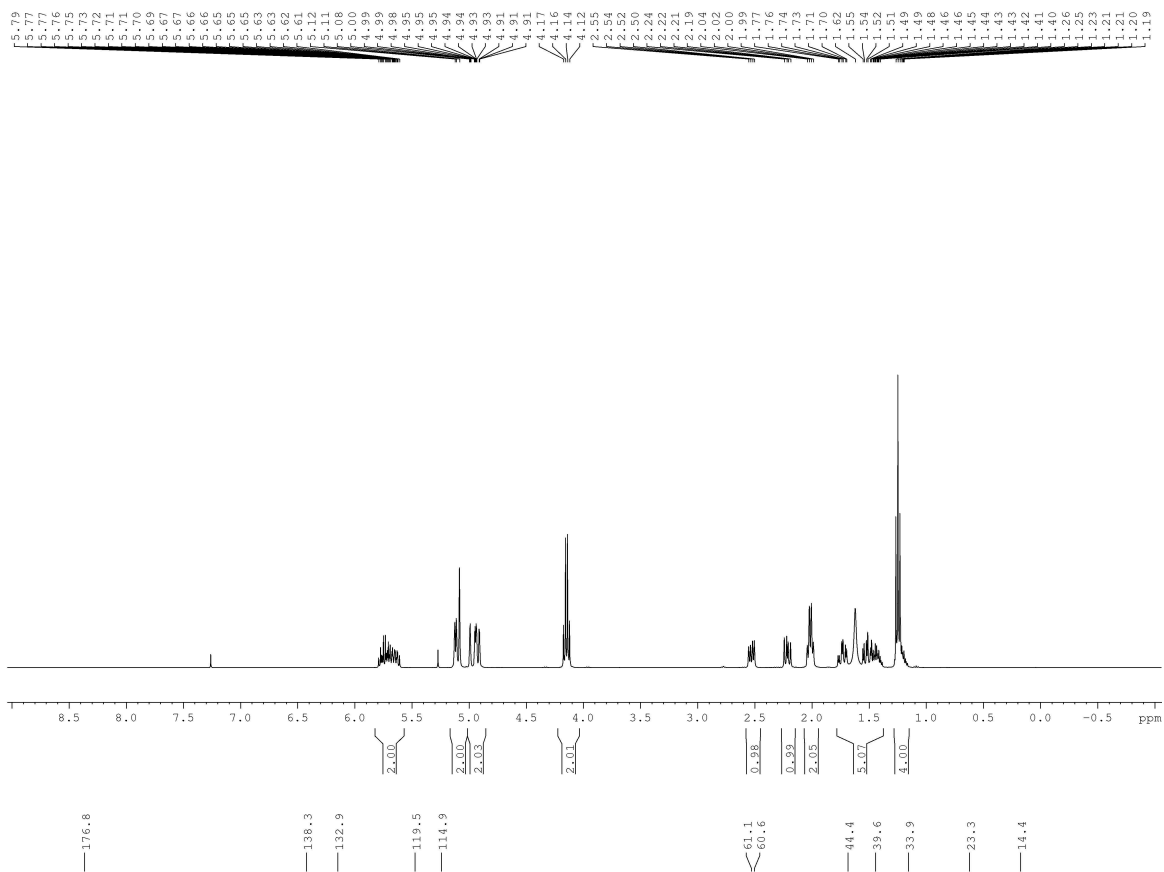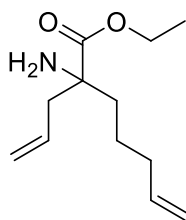
Ethyl ((diphenylmethylene)amino)acetate **2**
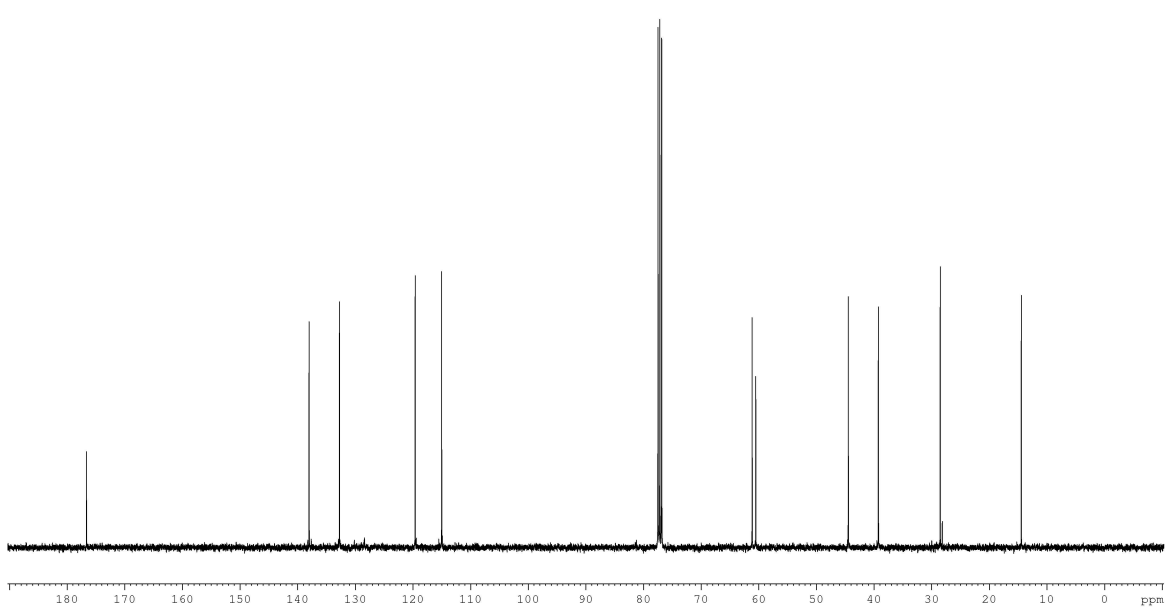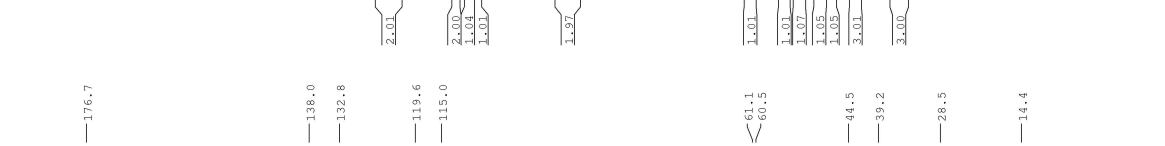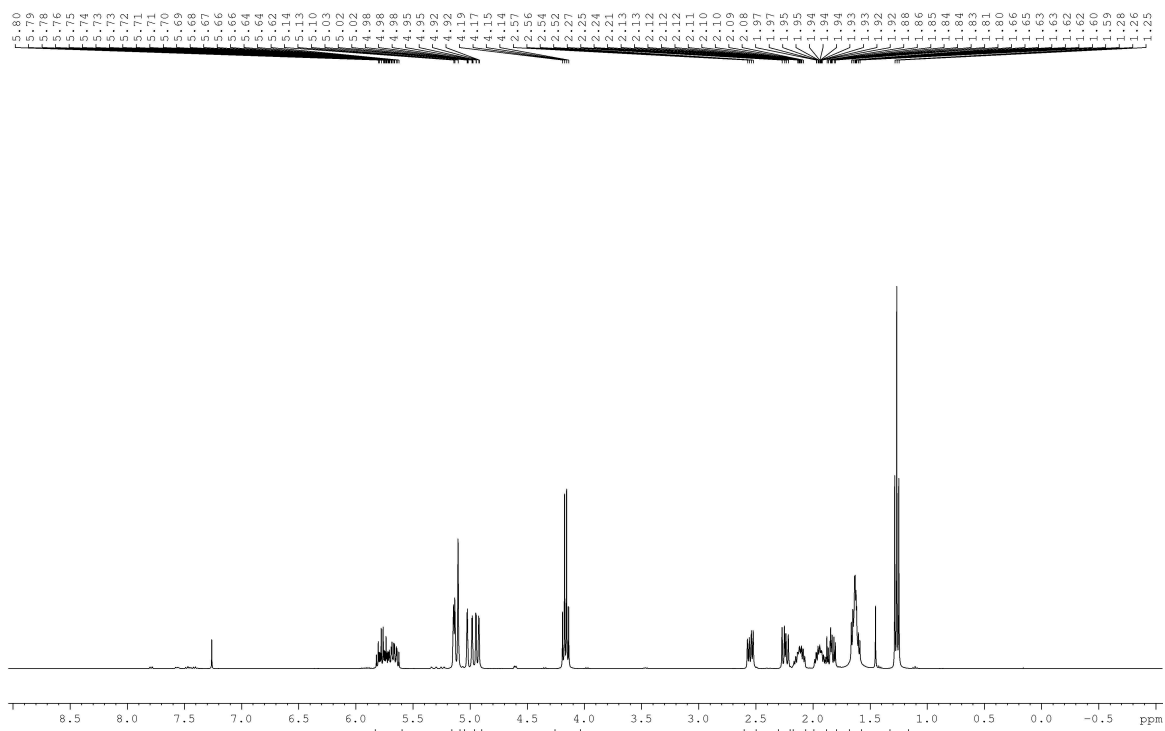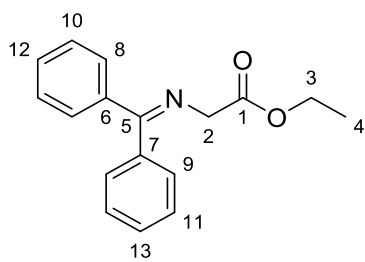
Ethyl 2-allyl-2-(3-ethoxy-3-oxopropanamido)pent-4-enoate **5a**
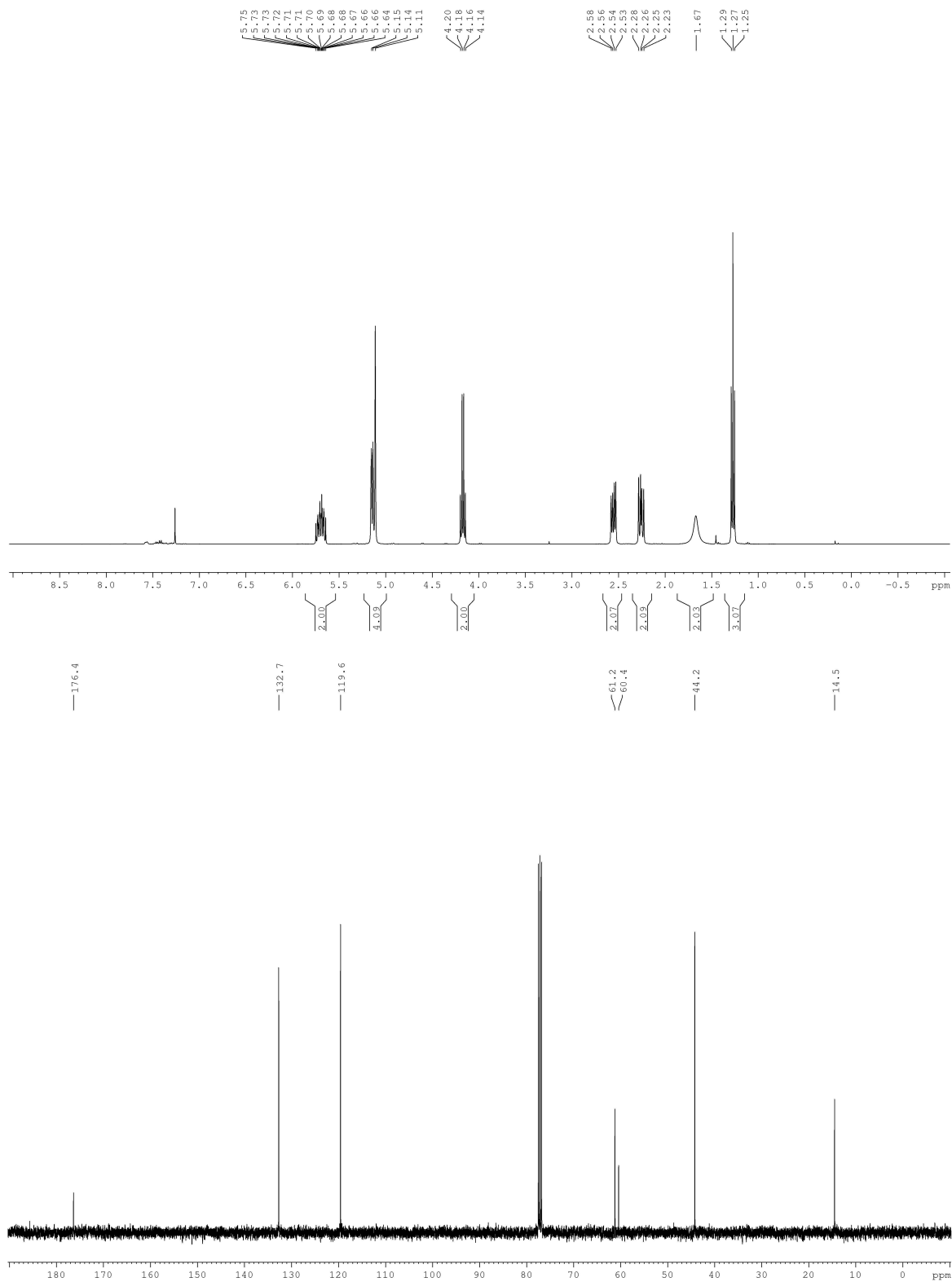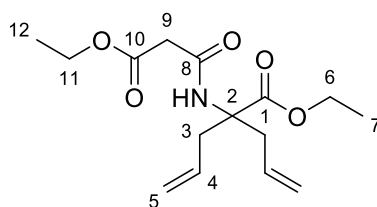
Ethyl 2-allyl-2-(3-ethoxy-3-oxopropanamido)hex-5-enoate **5b**
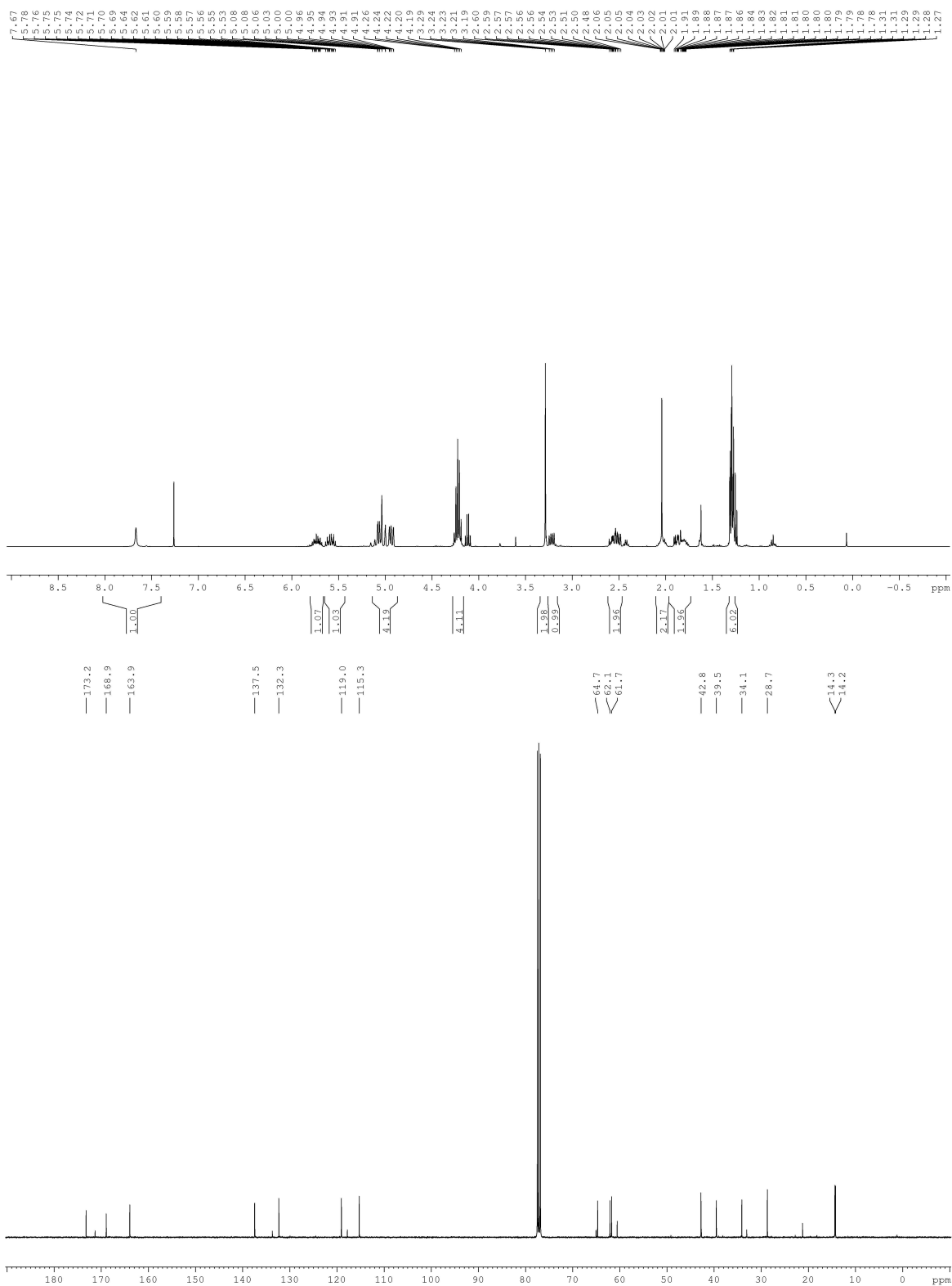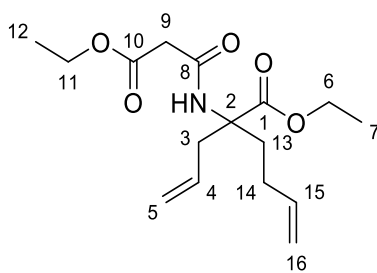
Ethyl 2-allyl-2-(3-ethoxy-3-oxopropanamido)hept-6-enoate **5c**
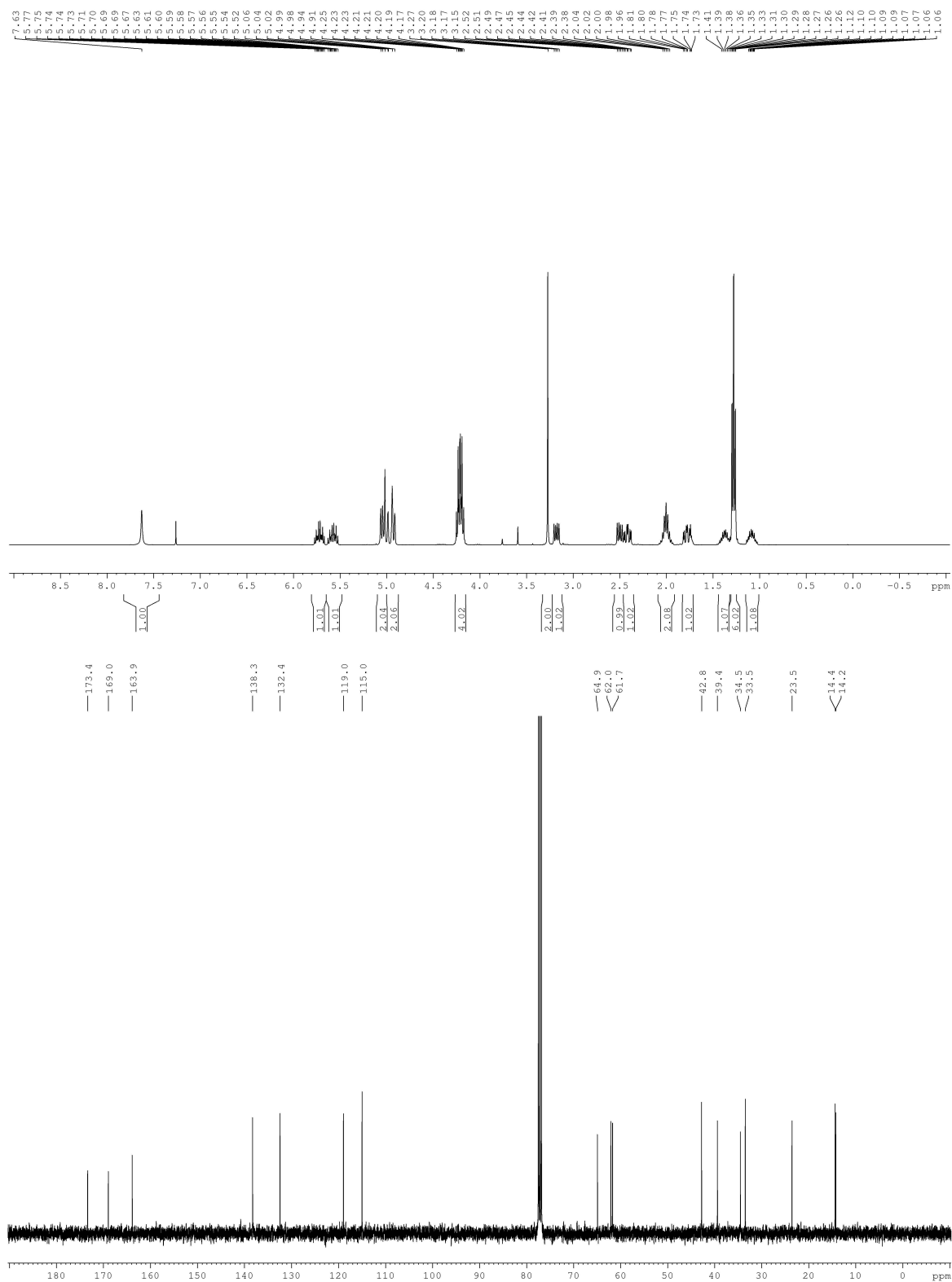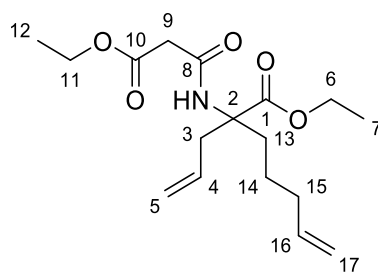
Methyl (*R*)-2-allyl-2-(3-ethoxy-3-oxopropanamido)hex-5-enoate **(*R*)-5d**
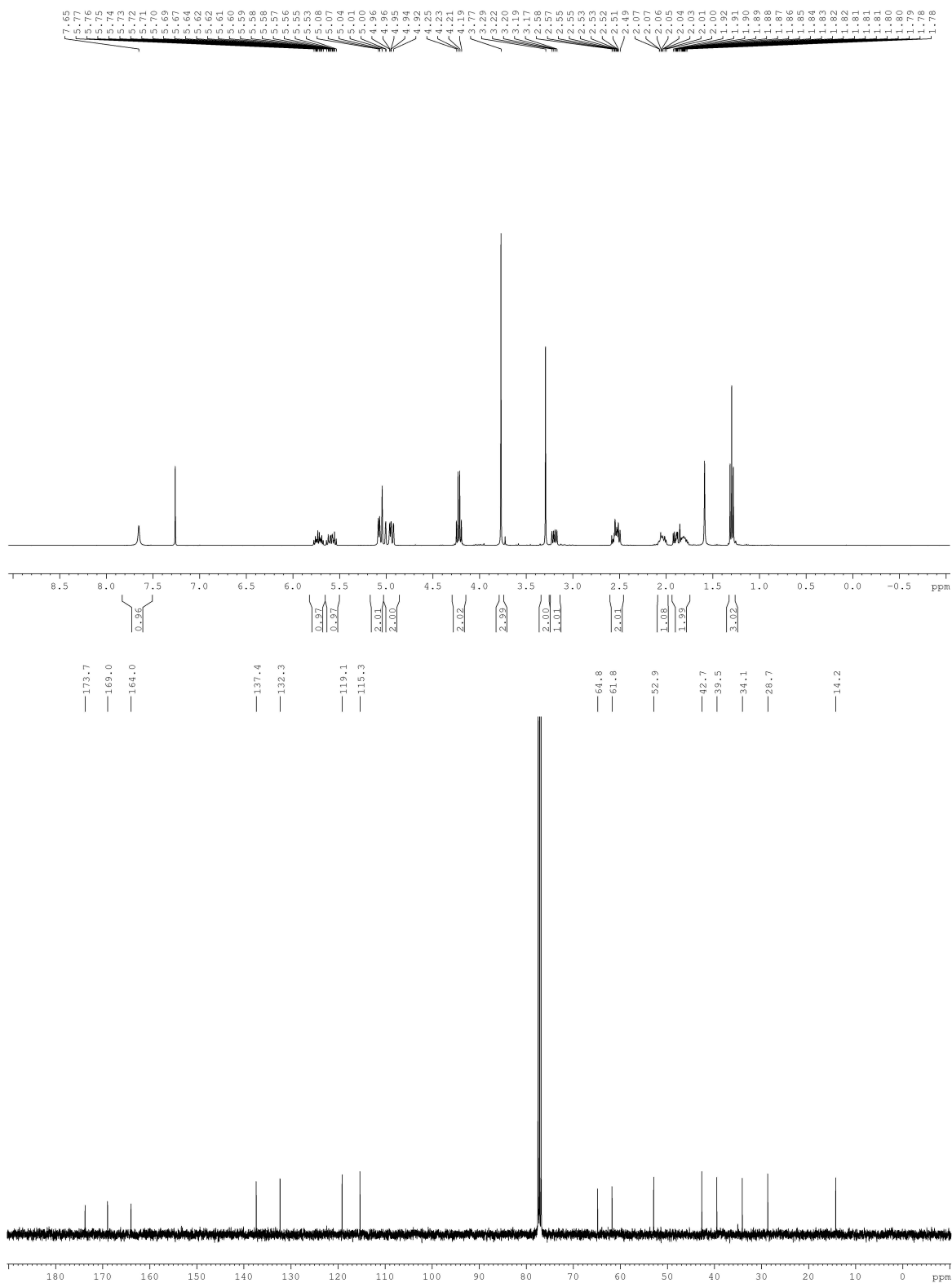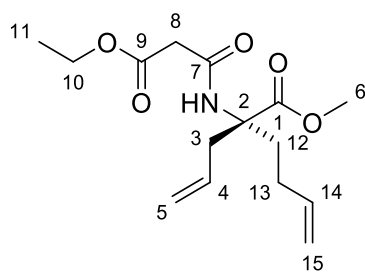
5,5-Diallylpyrrolidine-2,4-dione (10a) **7a**
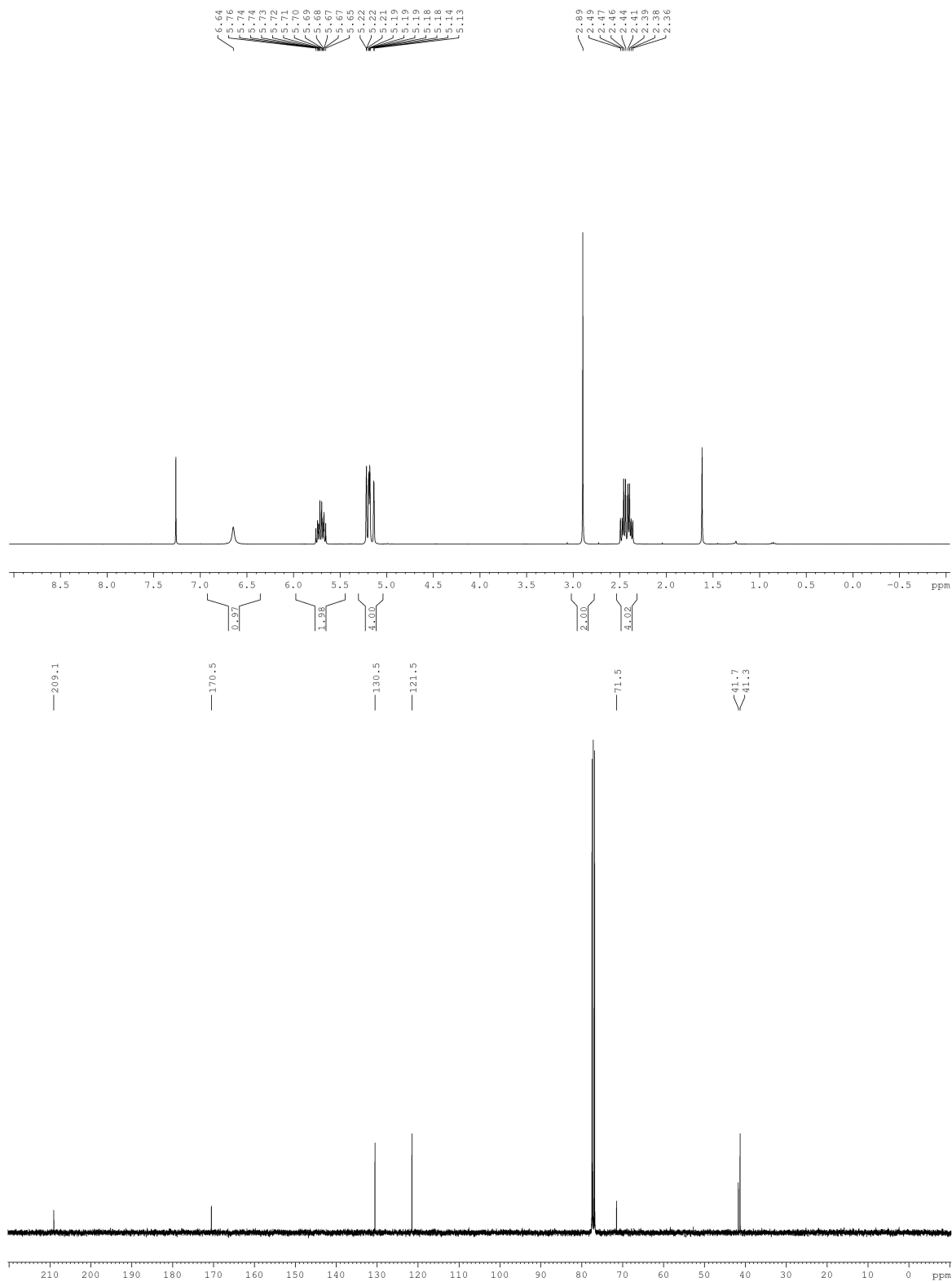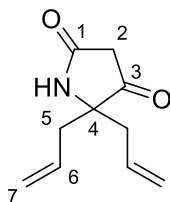
5-Allyl-5-(but-3-en-1-yl)pyrrolidine-2,4-dione (10b) **7b**
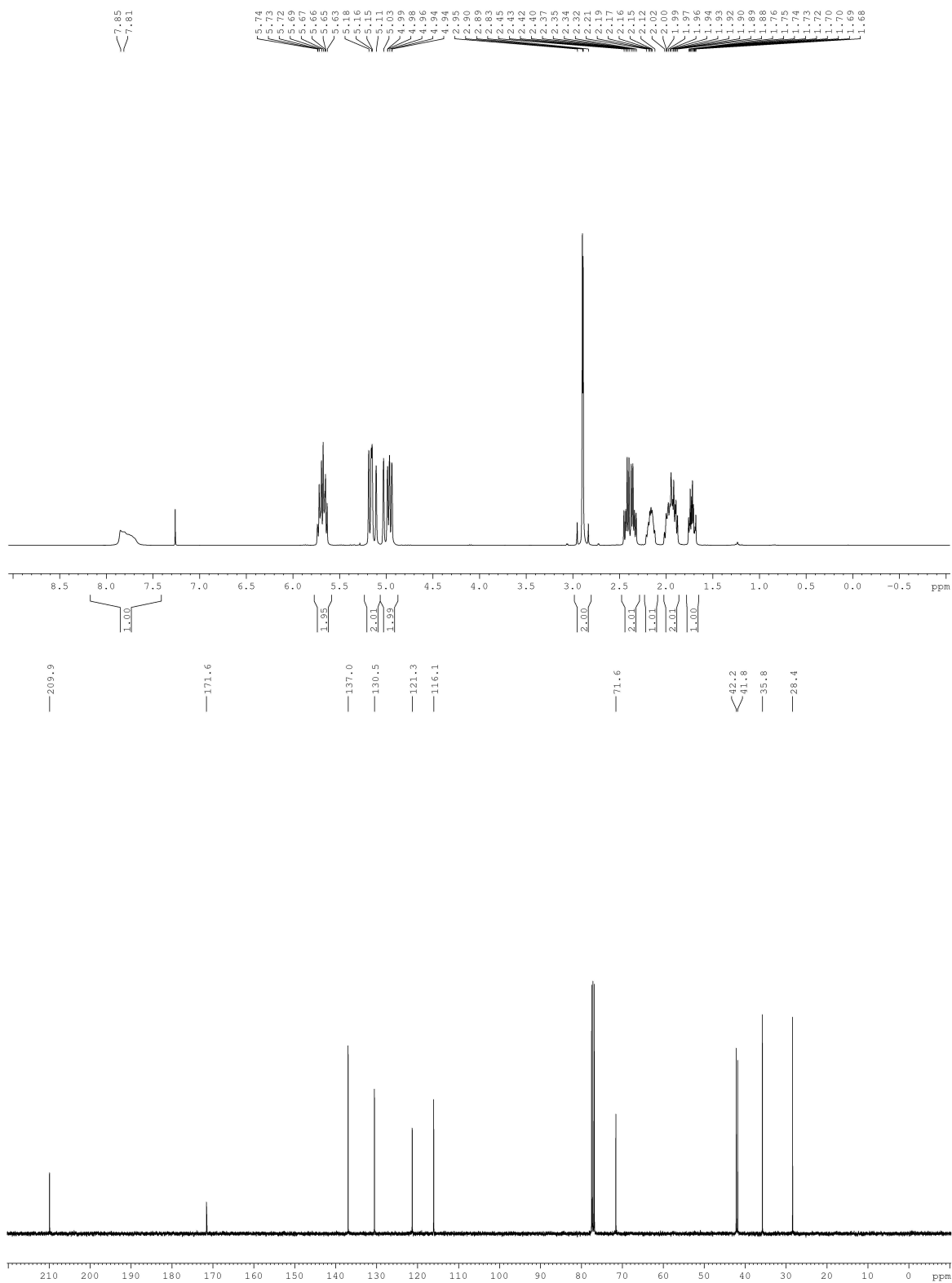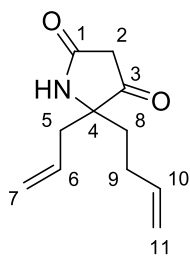
5-Allyl-5-(pent-4-en-1-yl)pyrrolidine-2,4-dione (10c) **7c**
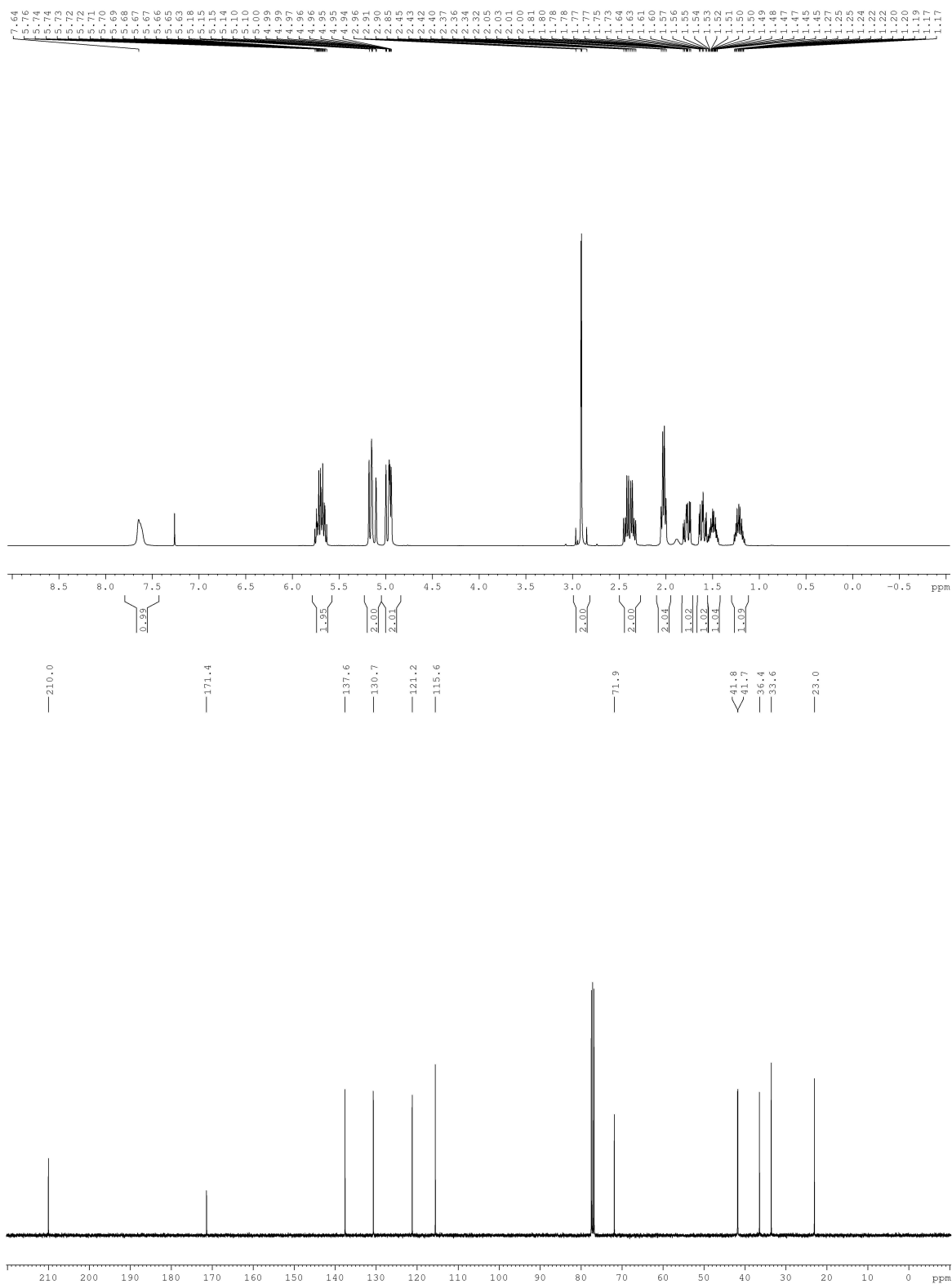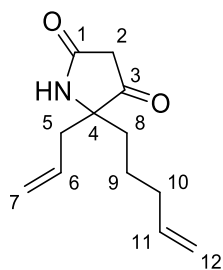
1-Azaspiro[4.4]non-7-ene-2,4-dione **8a**
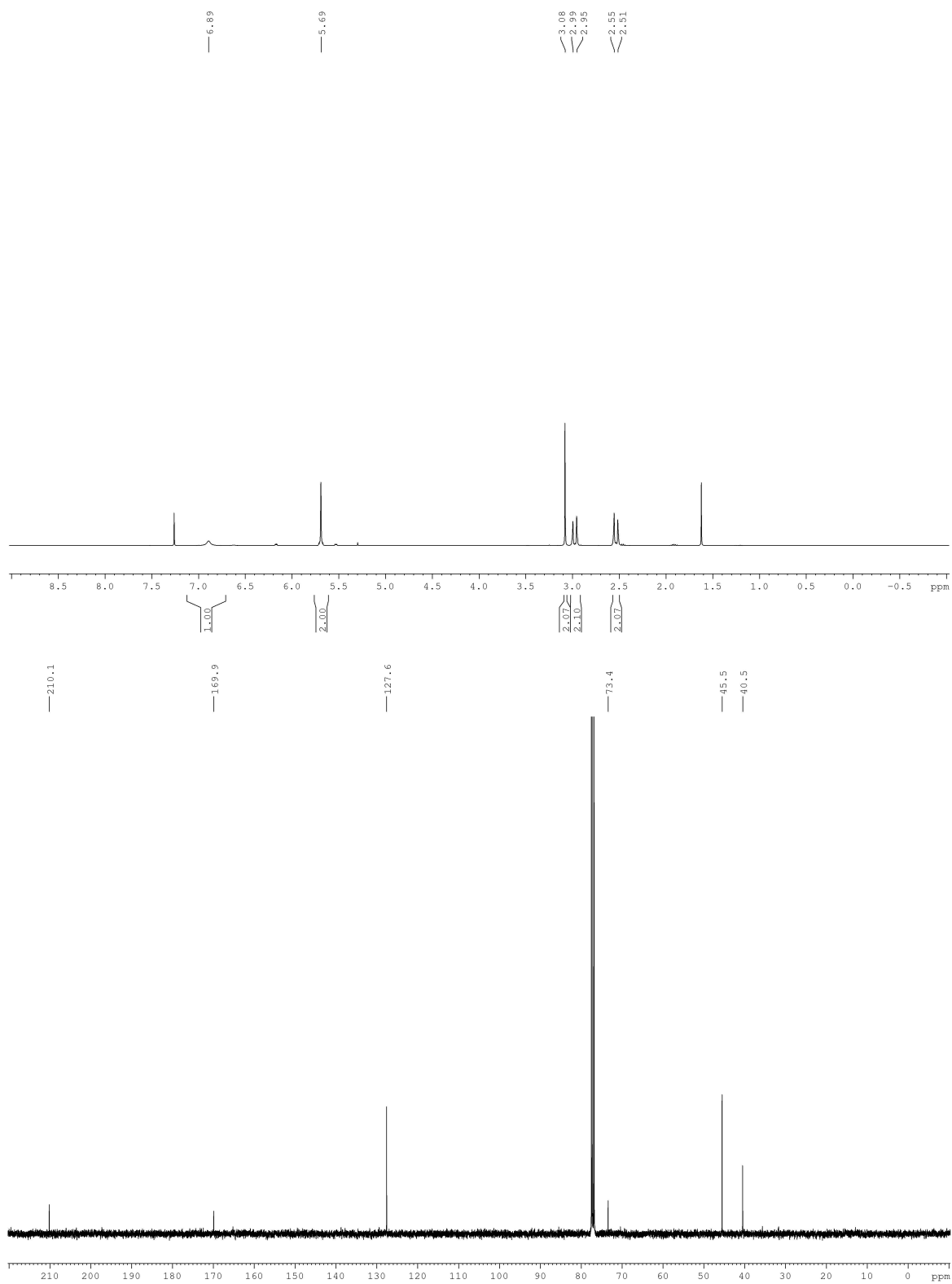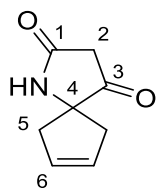
1-Azaspiro[4.5]dec-7-ene-2,4-dione **8b**
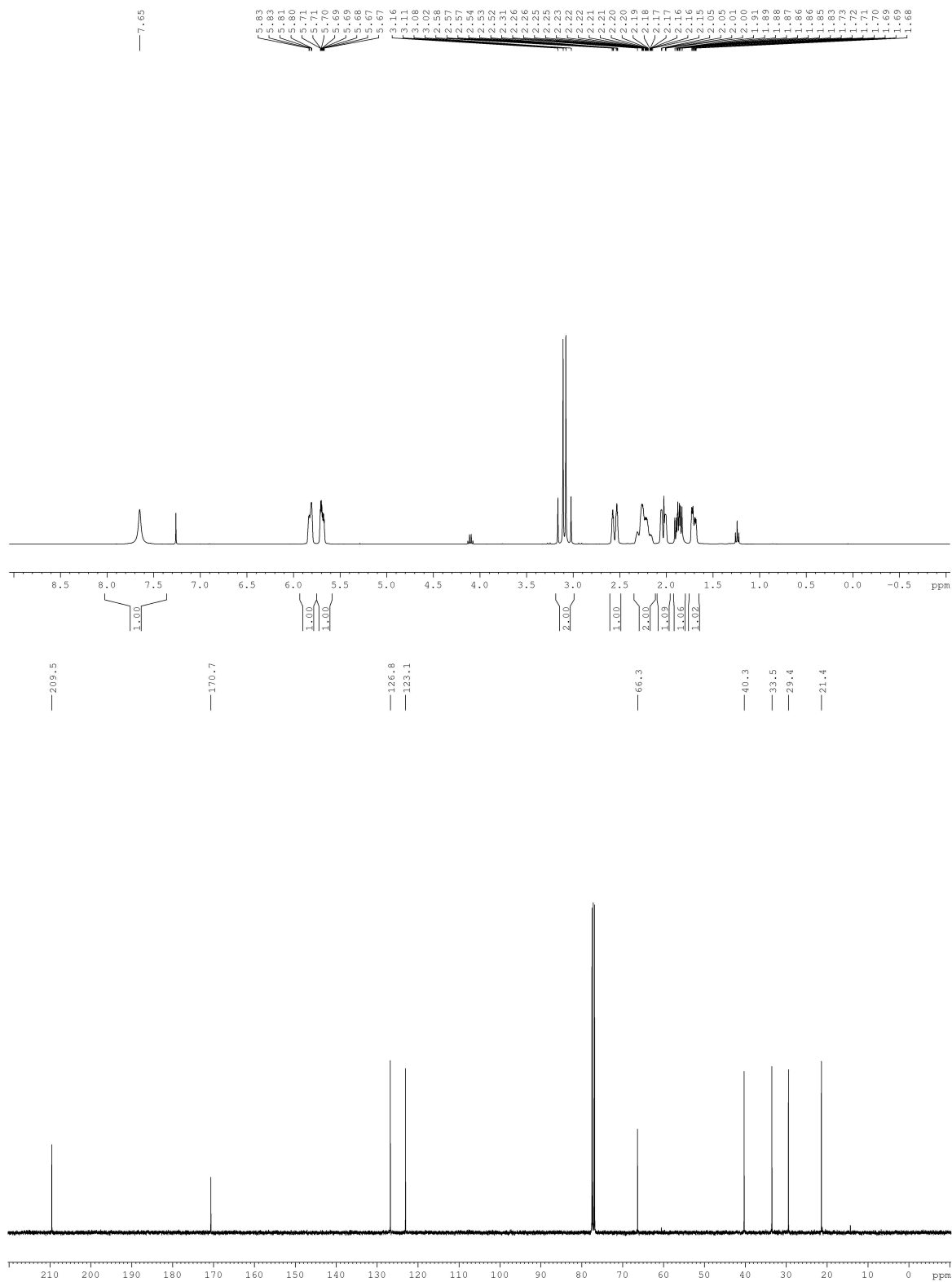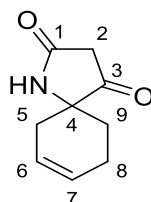
1-Azaspiro[4.6]undec-7-ene-2,4-dione **8c**
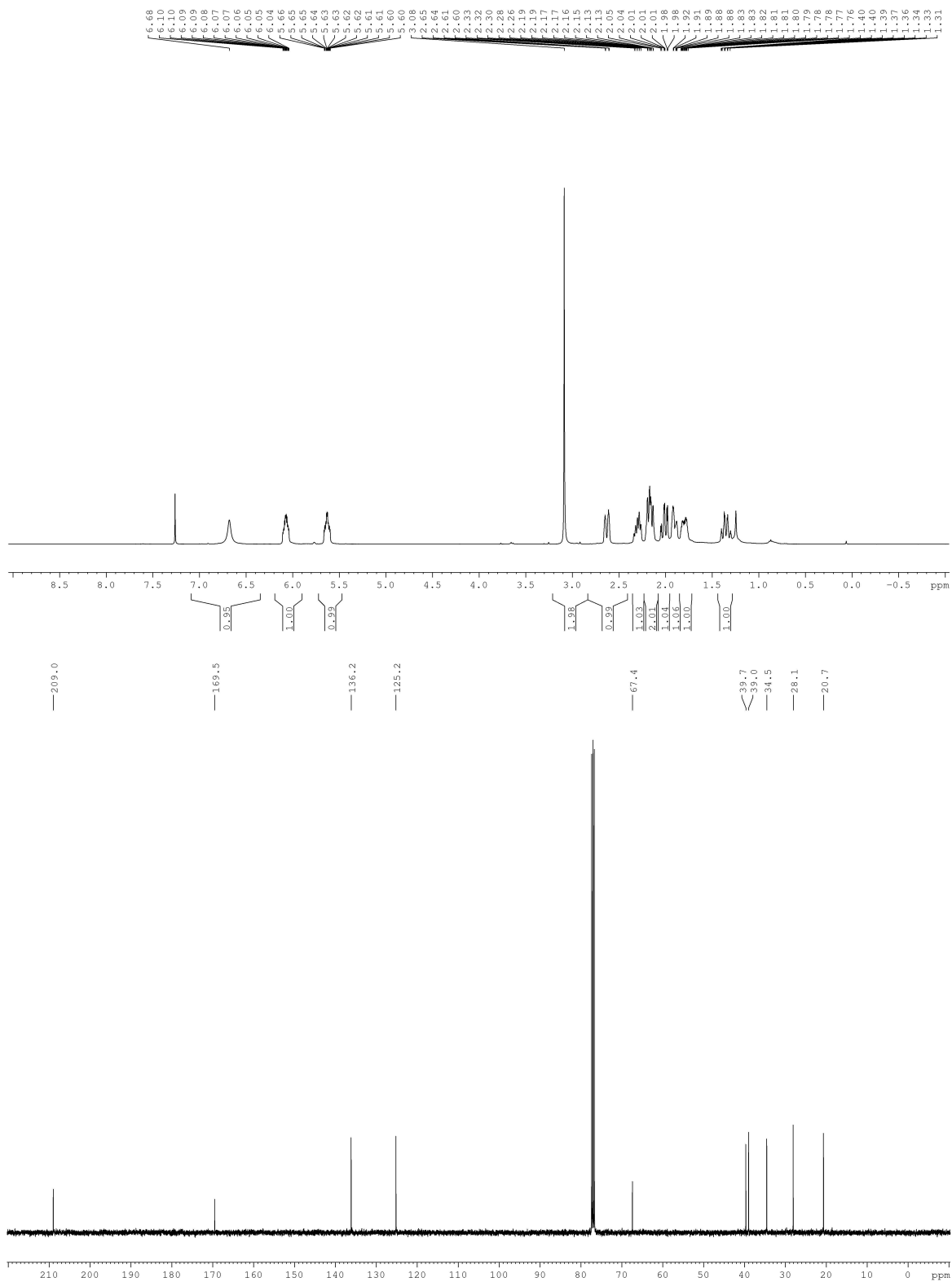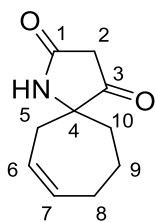
2-Chloro-*N*-(4-(hydroxymethyl)octa-1,7-dien-4-yl)acetamide **13**
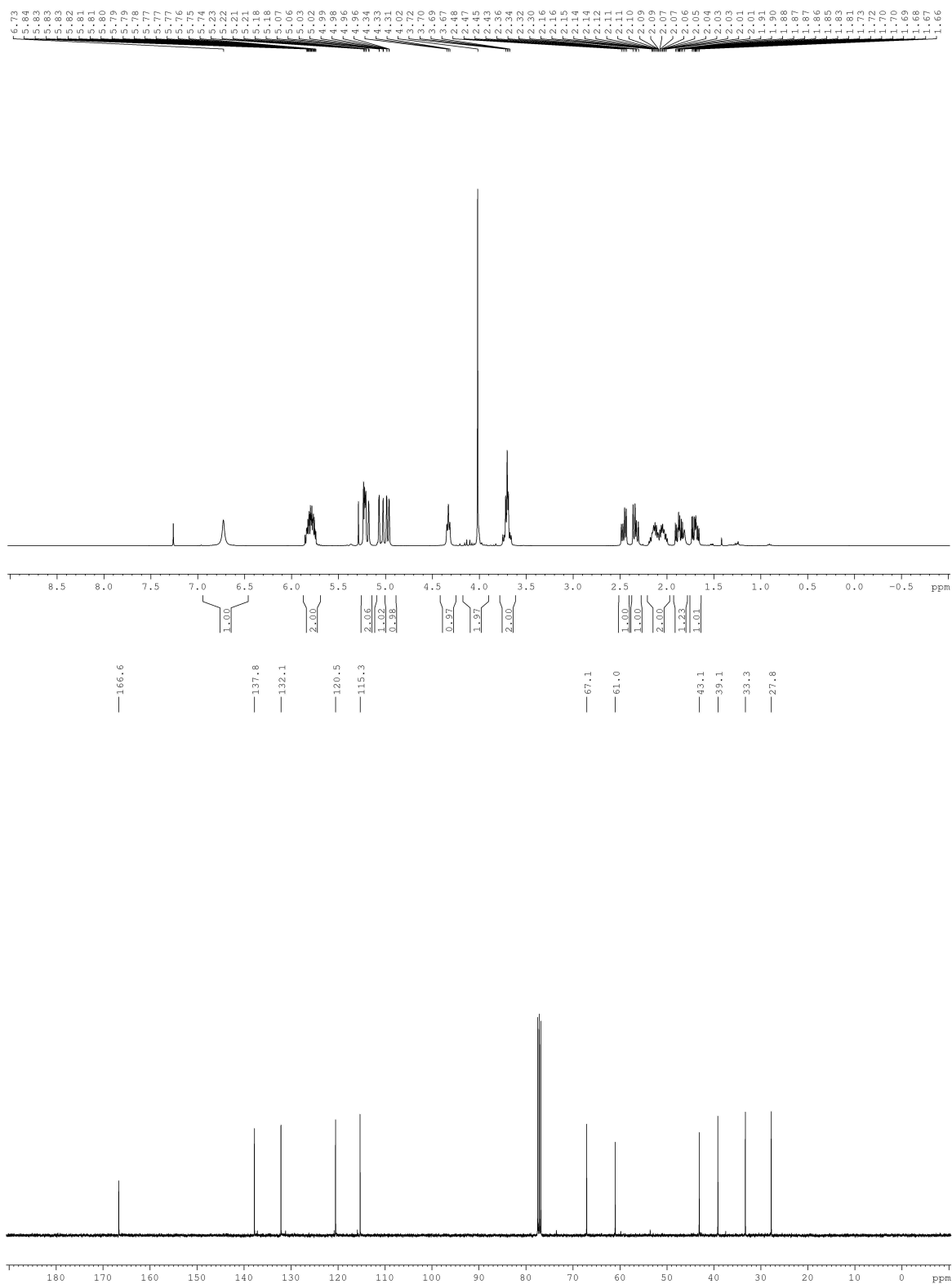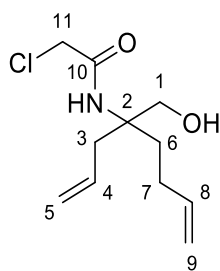
Ethyl 2-allyl-2-(2-chloroacetamido)hex-5-enoate **14**
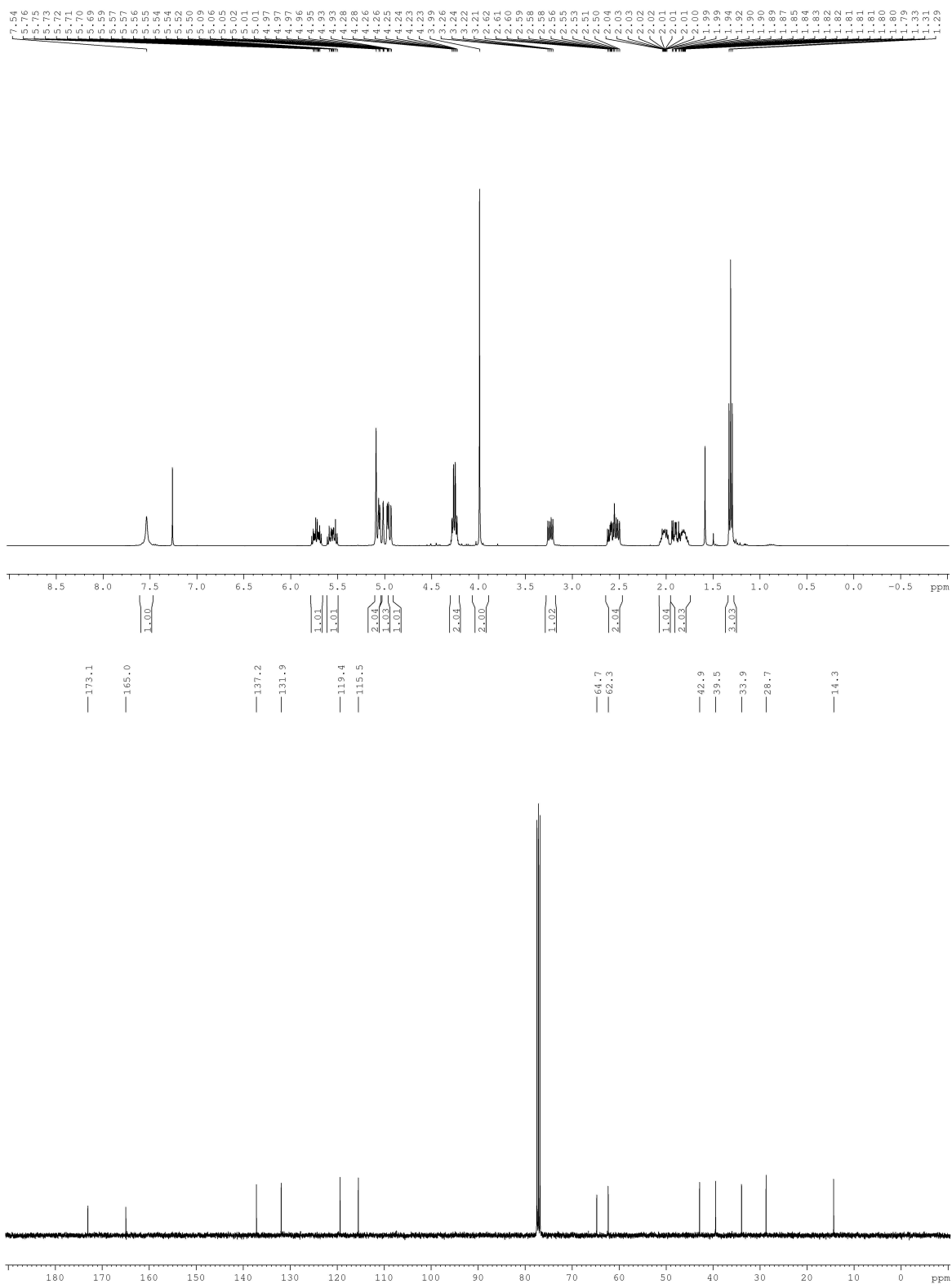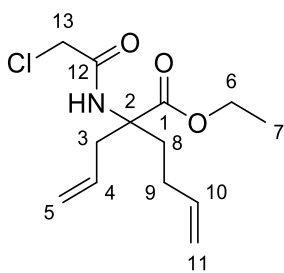
Ethyl 2-allyl-2-((*tert*-butoxycarbonyl)amino)hex-5-enoate **15**
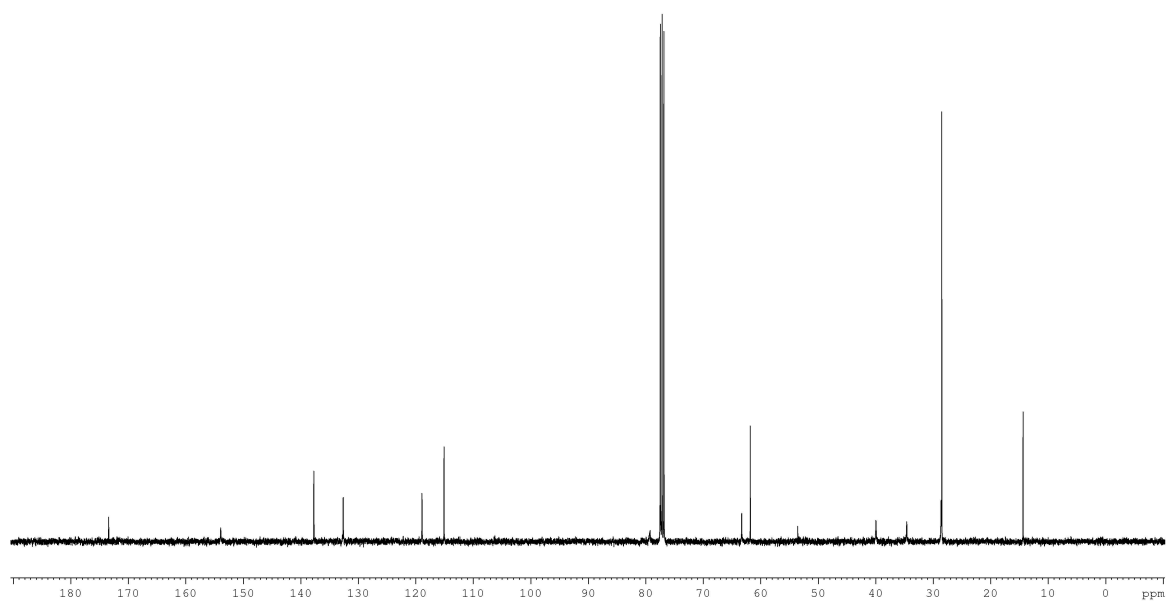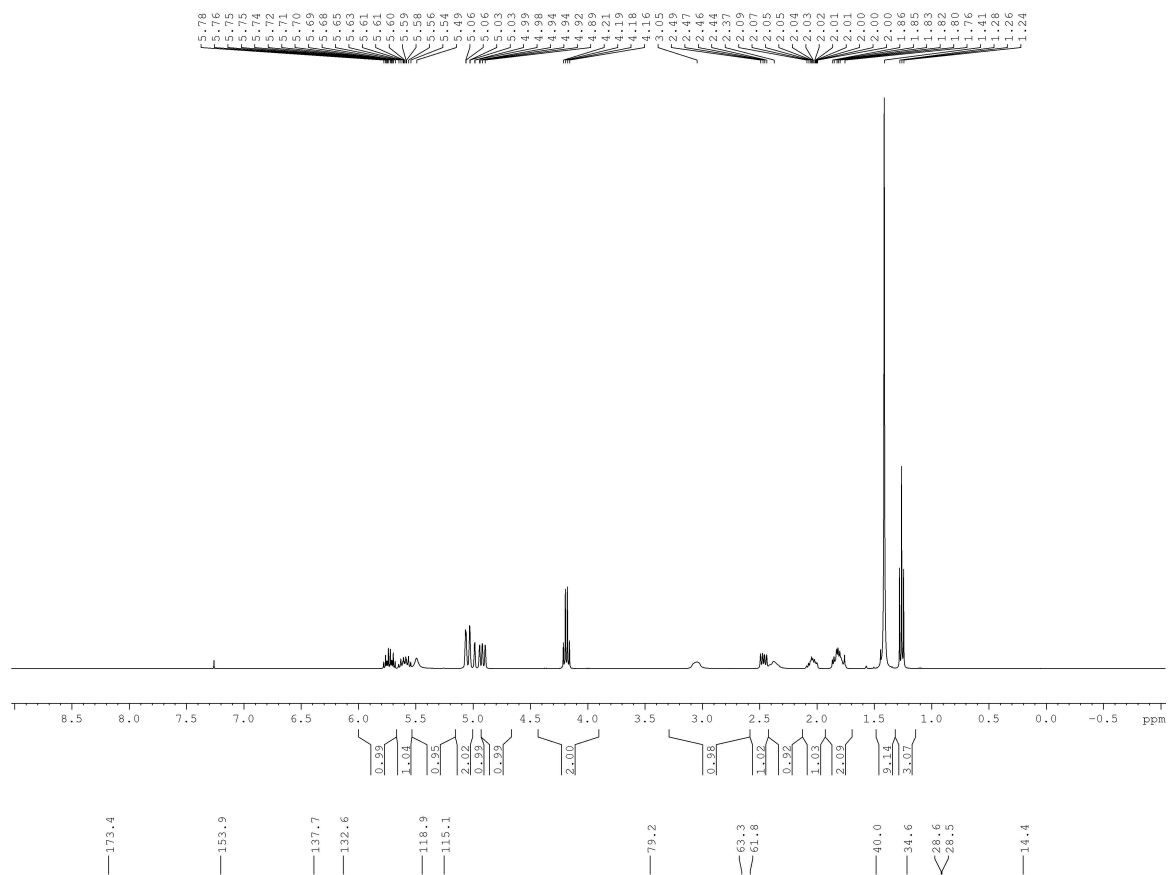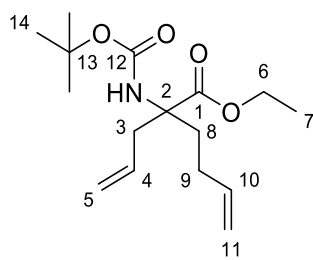
tert-Butyl (4-(hydroxymethyl)octa-1,7-dien-4-yl)carbamate **16**
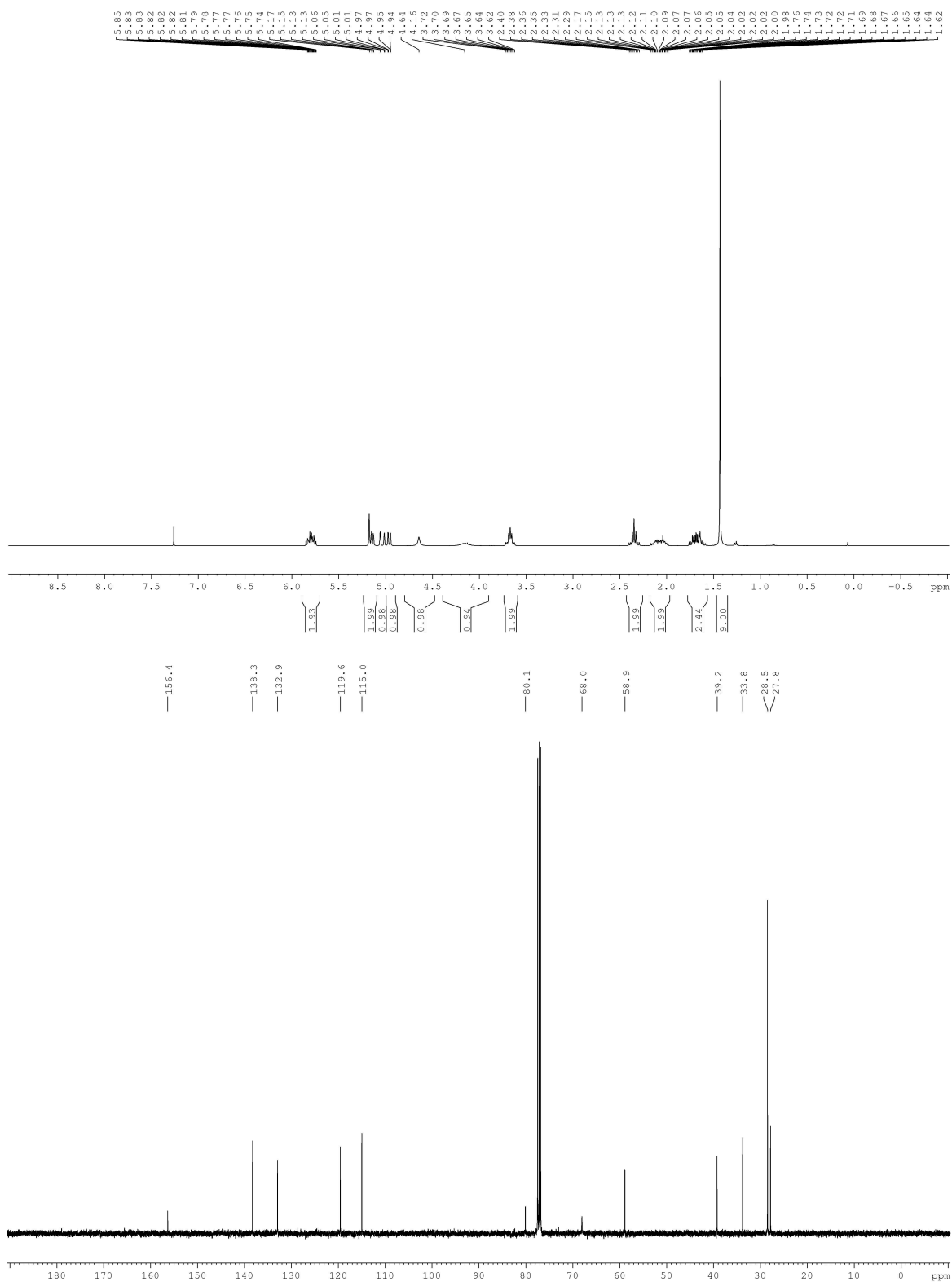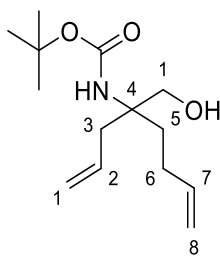
5-Allyl-5-(but-3-en-1-yl)morpholin-3-one **17**
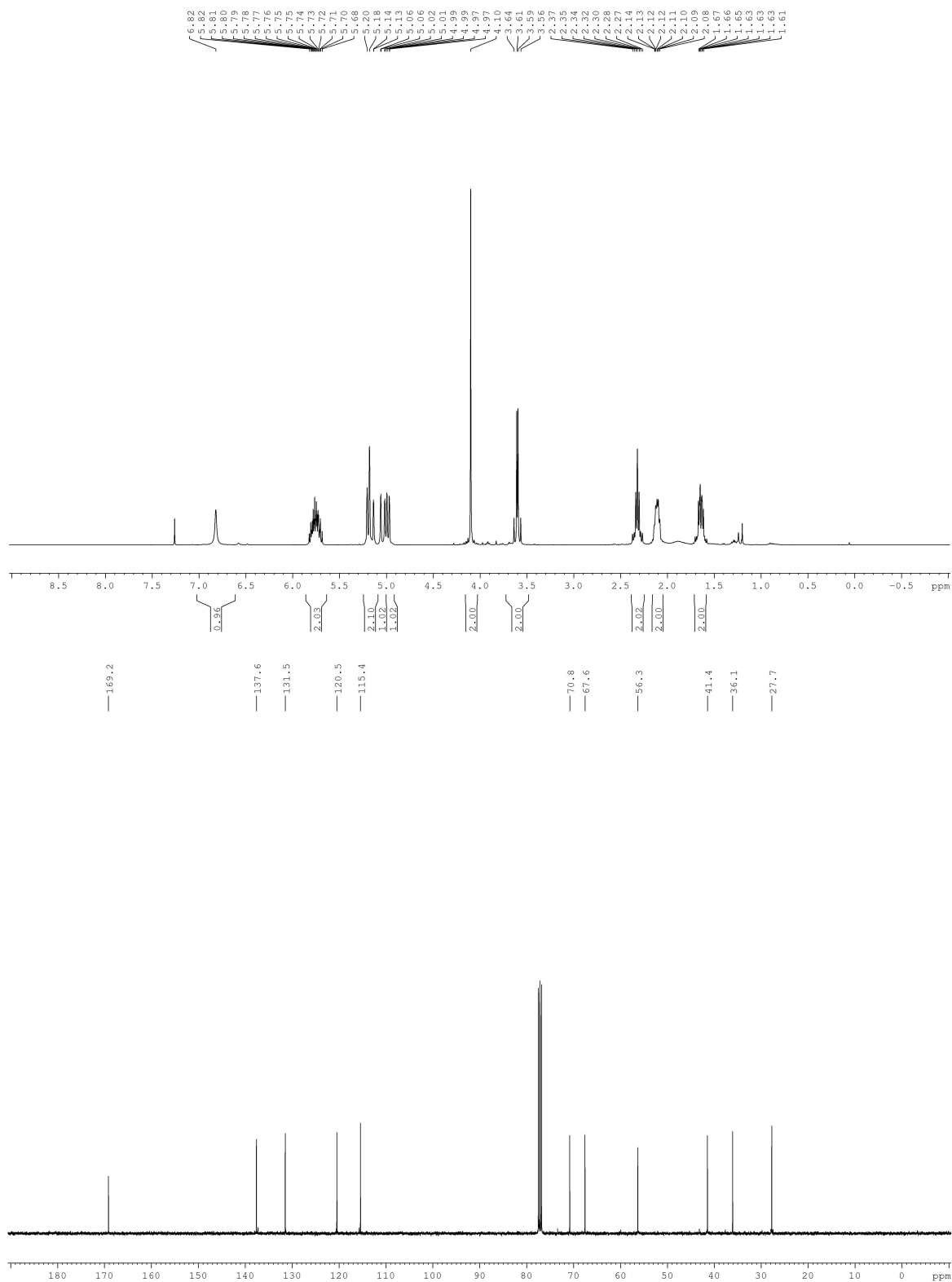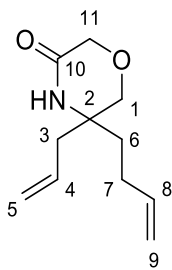
4-Oxa-1-azaspiro[5.5]undec-8-en-2-one **18**
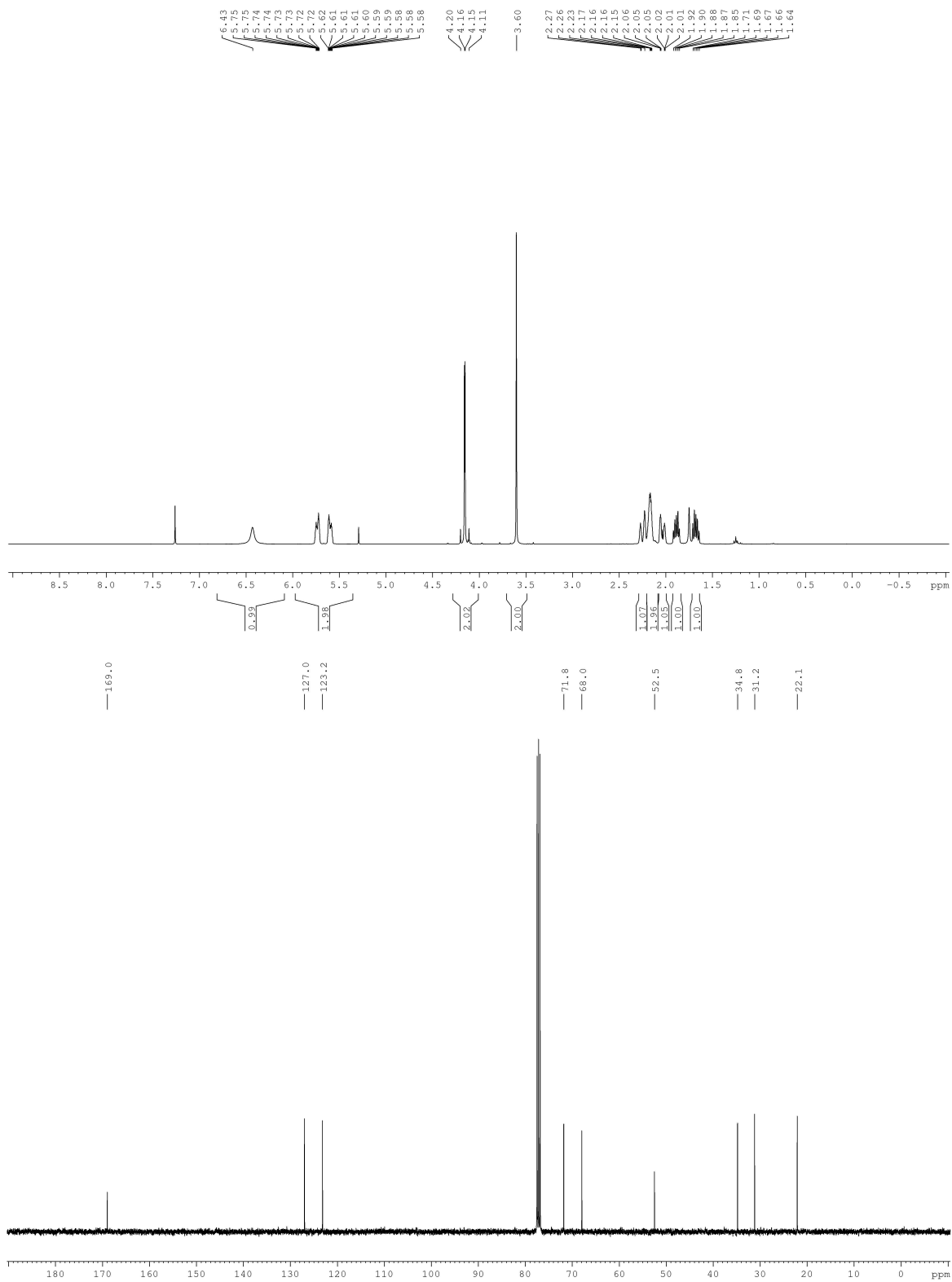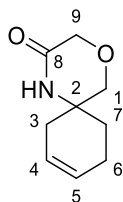
4-Allyl-4-(but-3-en-1-yl)oxazolidin-2-one **19**
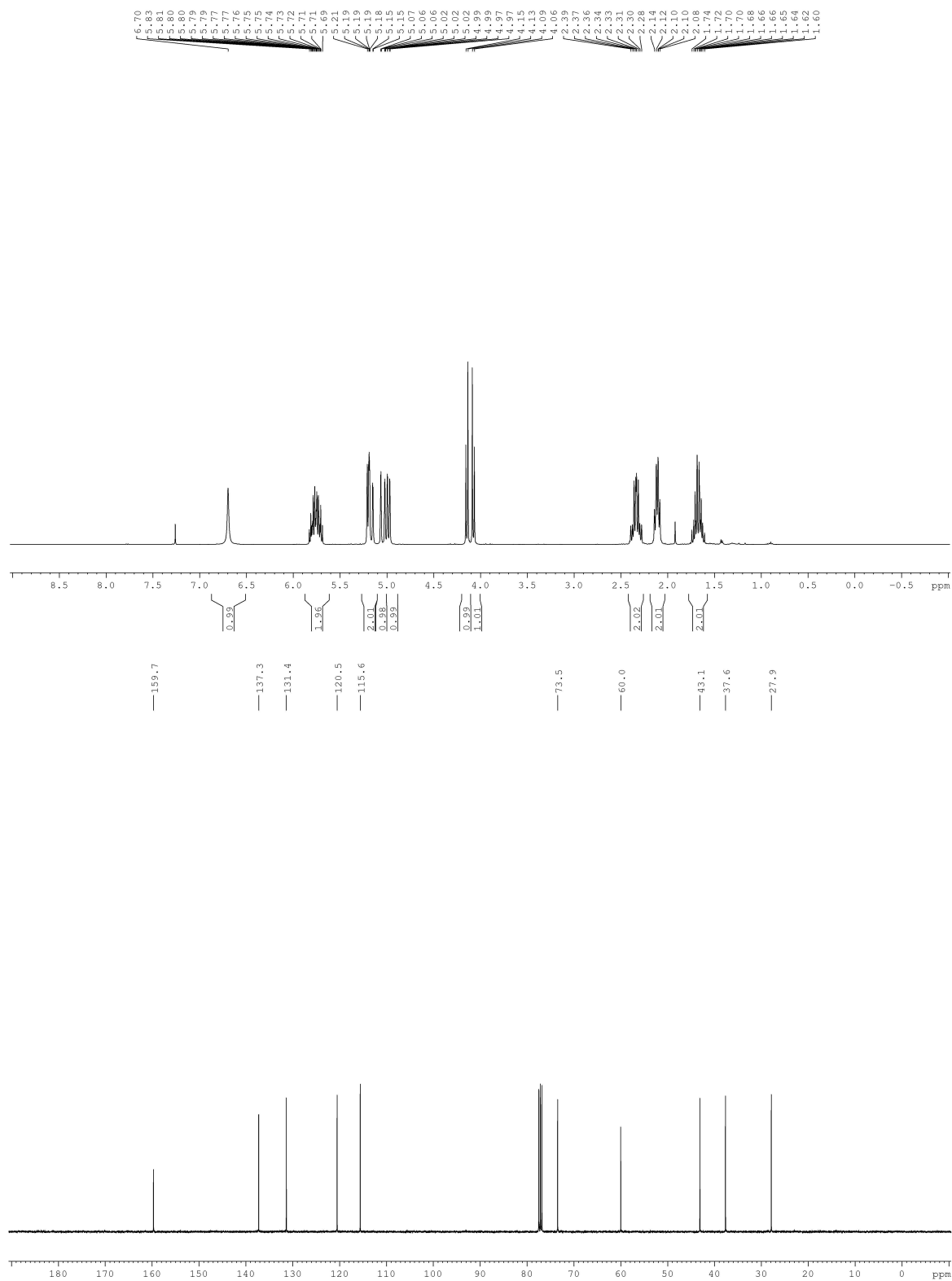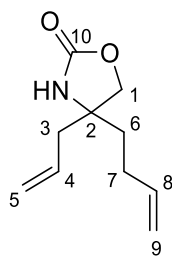
tert-butyl 2-allyl-2-(but-3-en-1-yl)aziridine-1-carboxylate **20**
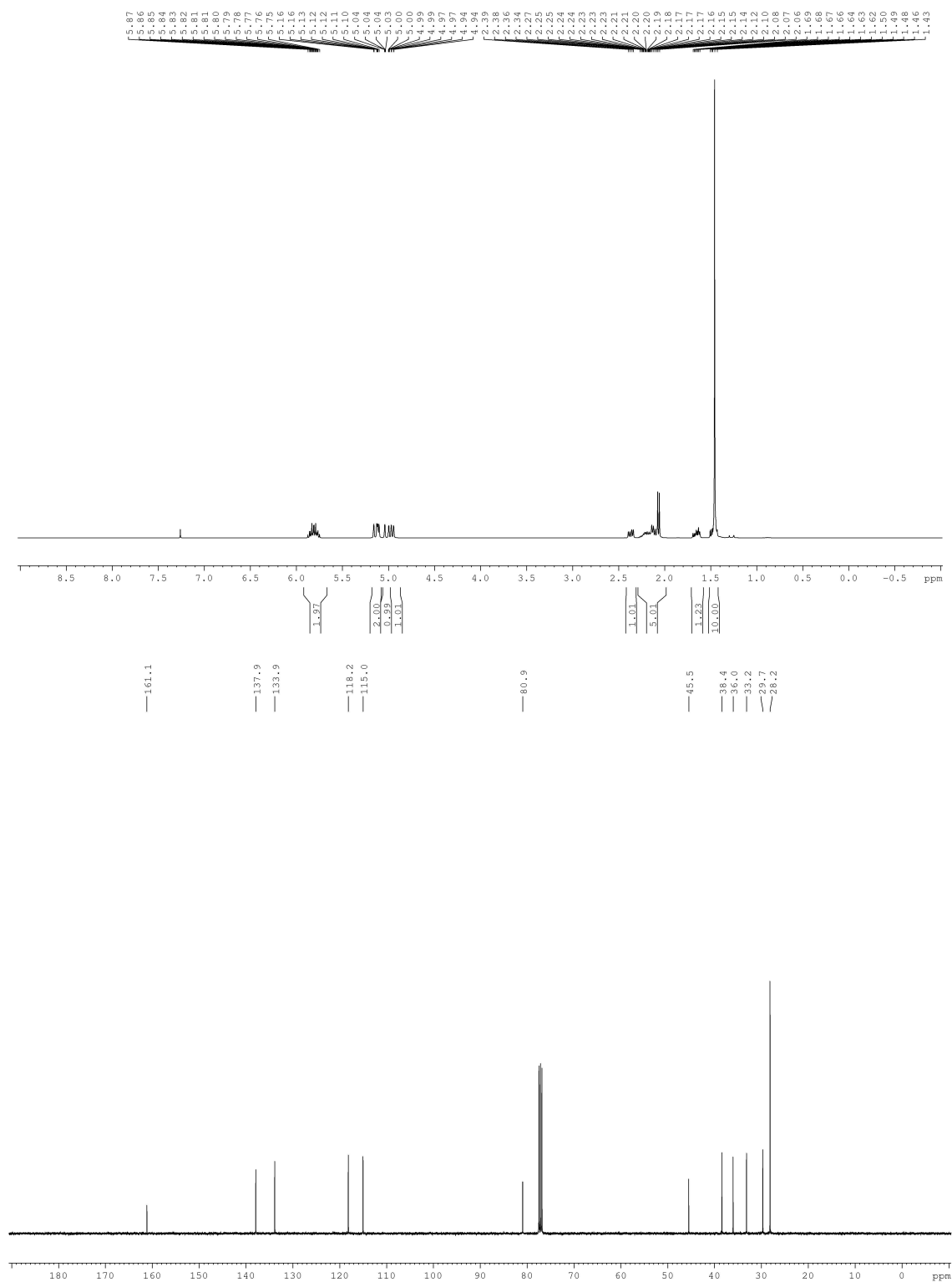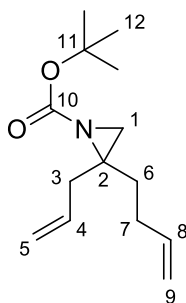
3-Oxa-1-azaspiro[4.5]dec-7-ene-2-one **21**
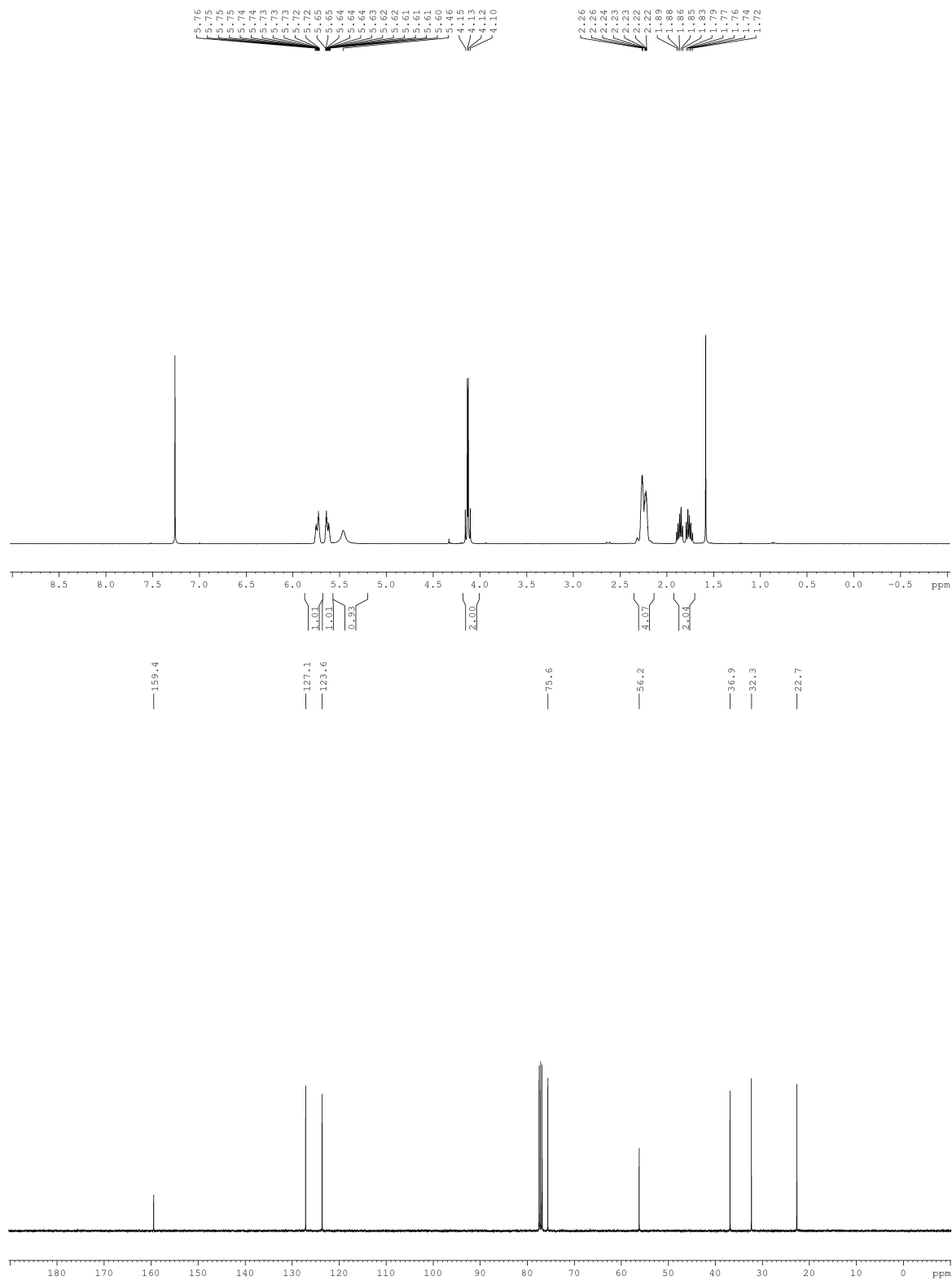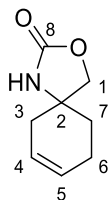
Ethyl 1-((*tert*-butoxycarbonyl)amino)cyclohex-3-ene-1-carboxylate **23**
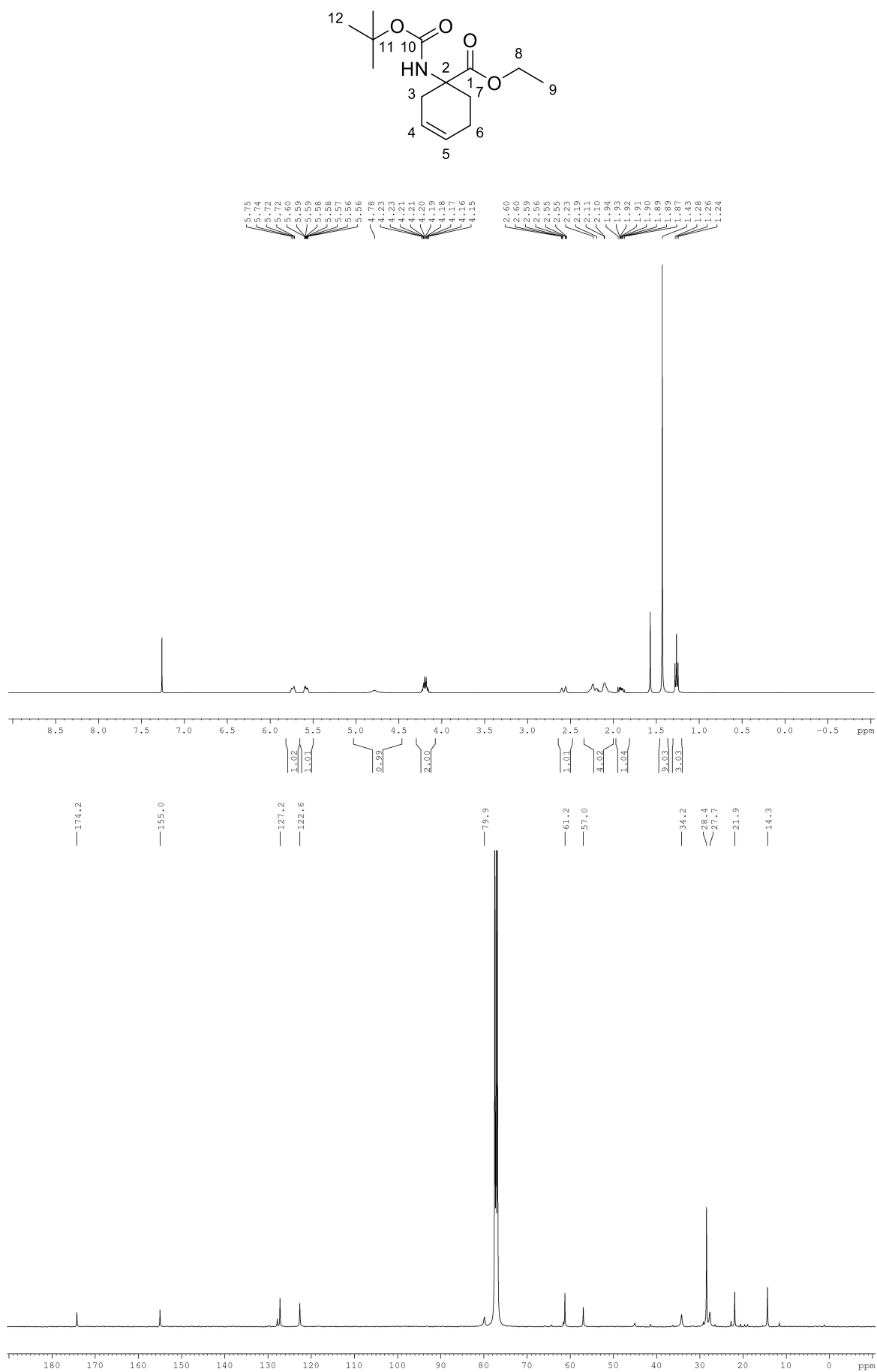
tert-Butyl (1-(hydroxymethyl)cyclohex-3-en-1-yl)carbamate **24**
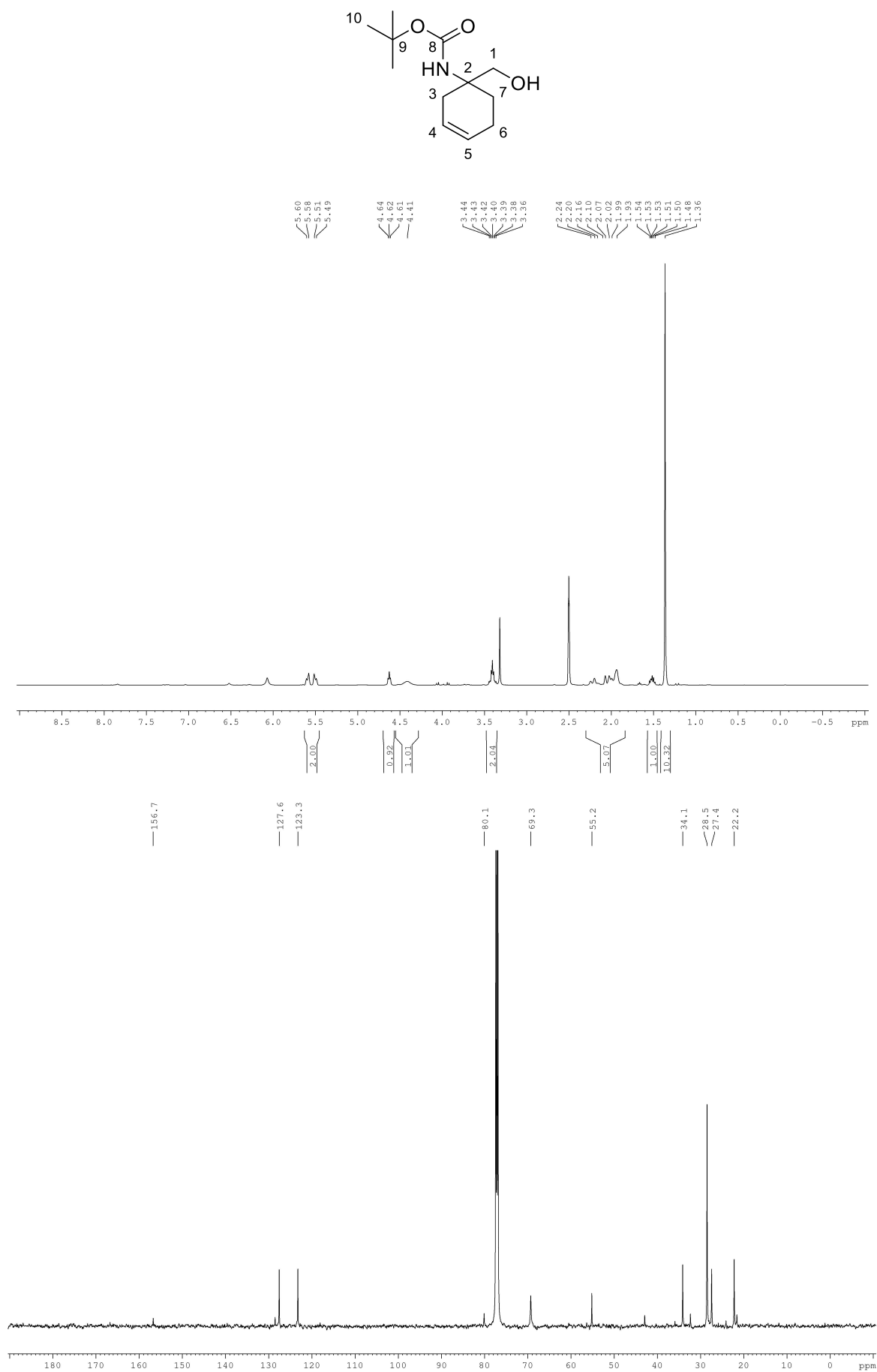
1-((*tert*-Butoxycarbonyl)amino)cyclohex-3-en-1-yl)methyl 4-methylbenzenesulfonate 25
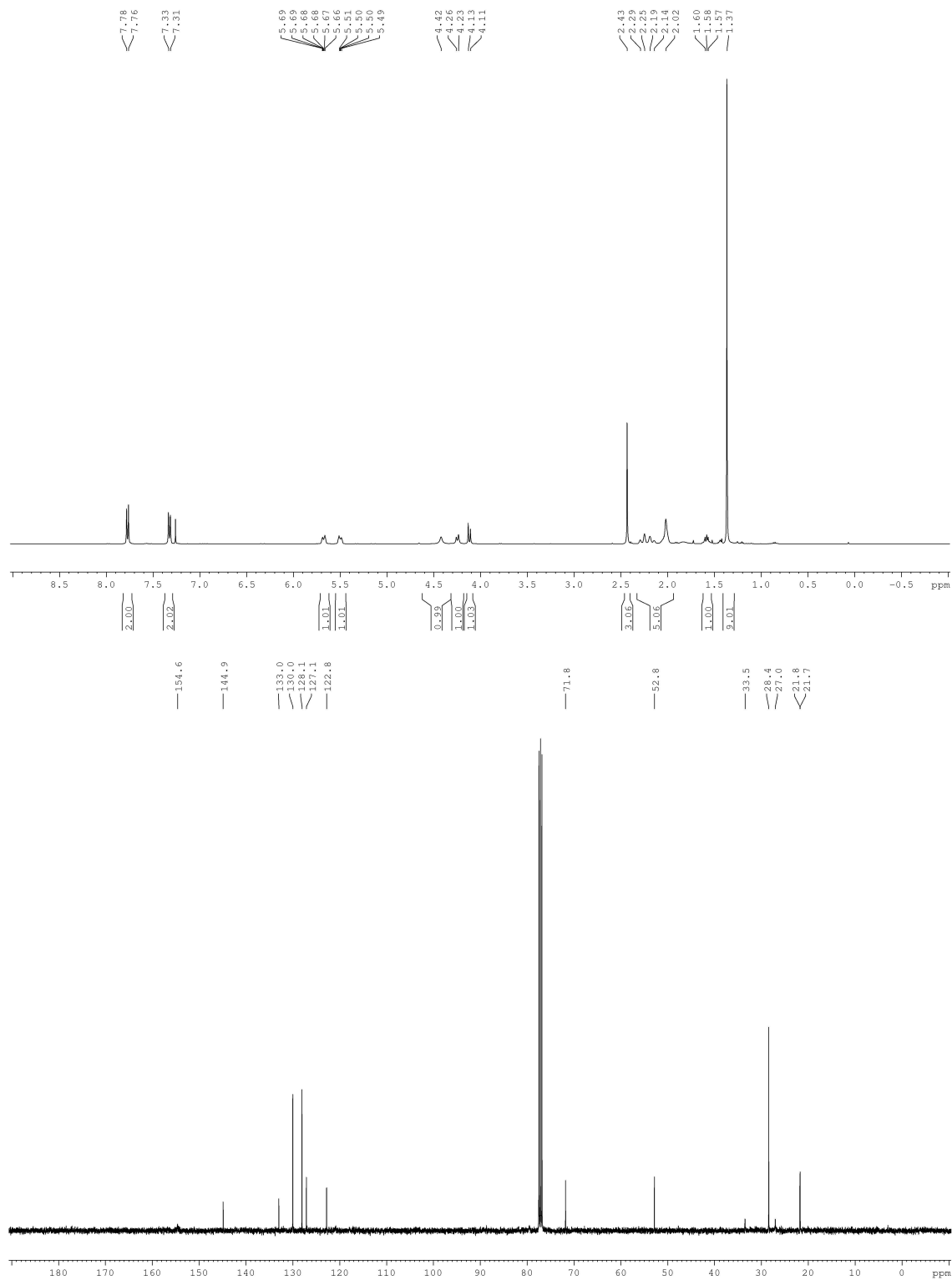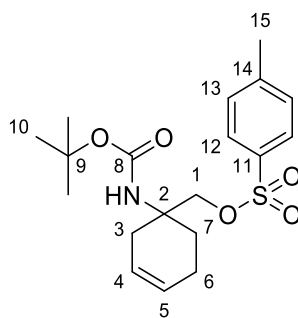
3-Oxa-1-azaspiro[4.5]decan-2-one **26**
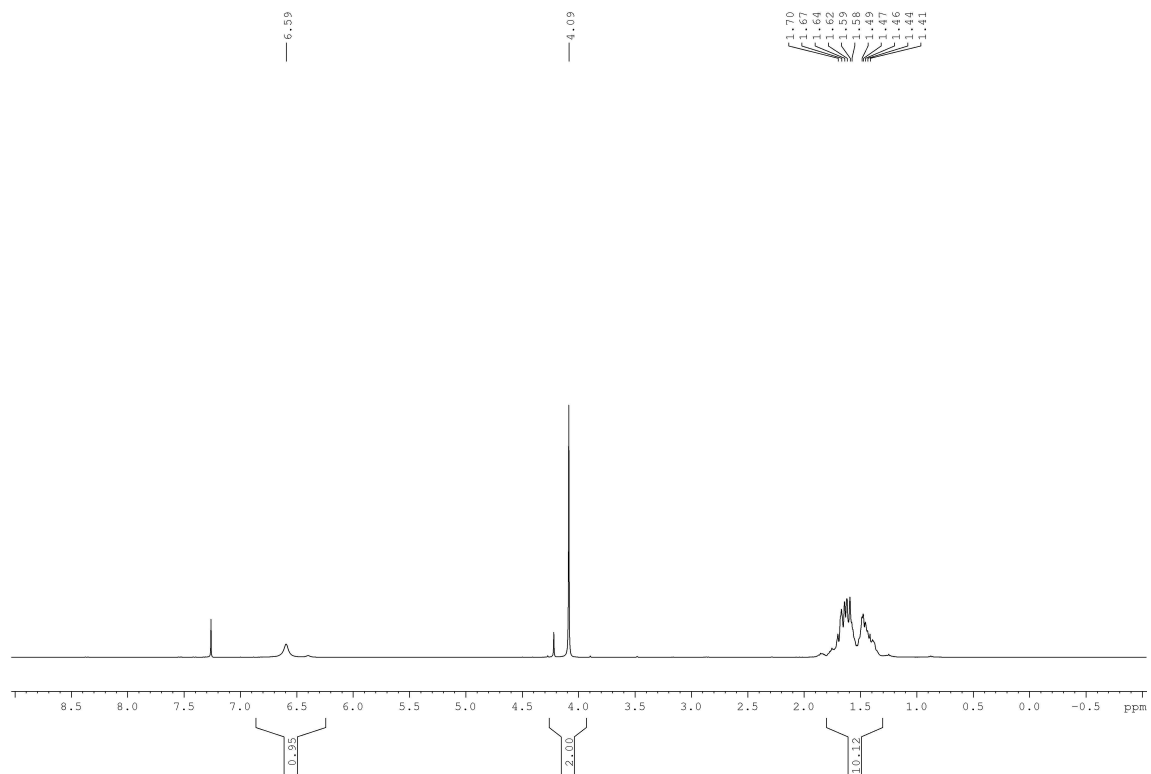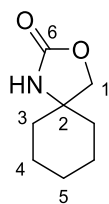
(1*R**,3*R**,6*S**)-7-Oxaspiro[bicyclo[4.1.0]heptane-3,4'-oxazolidin]-2'-one **27a**
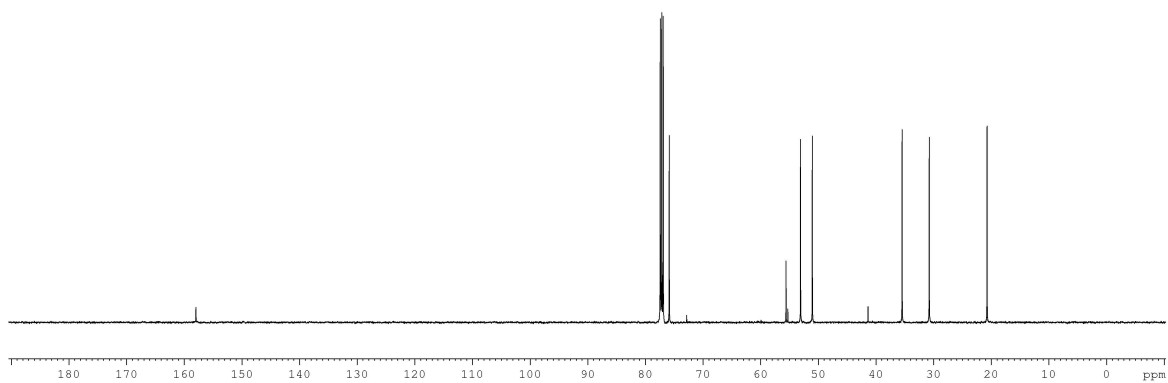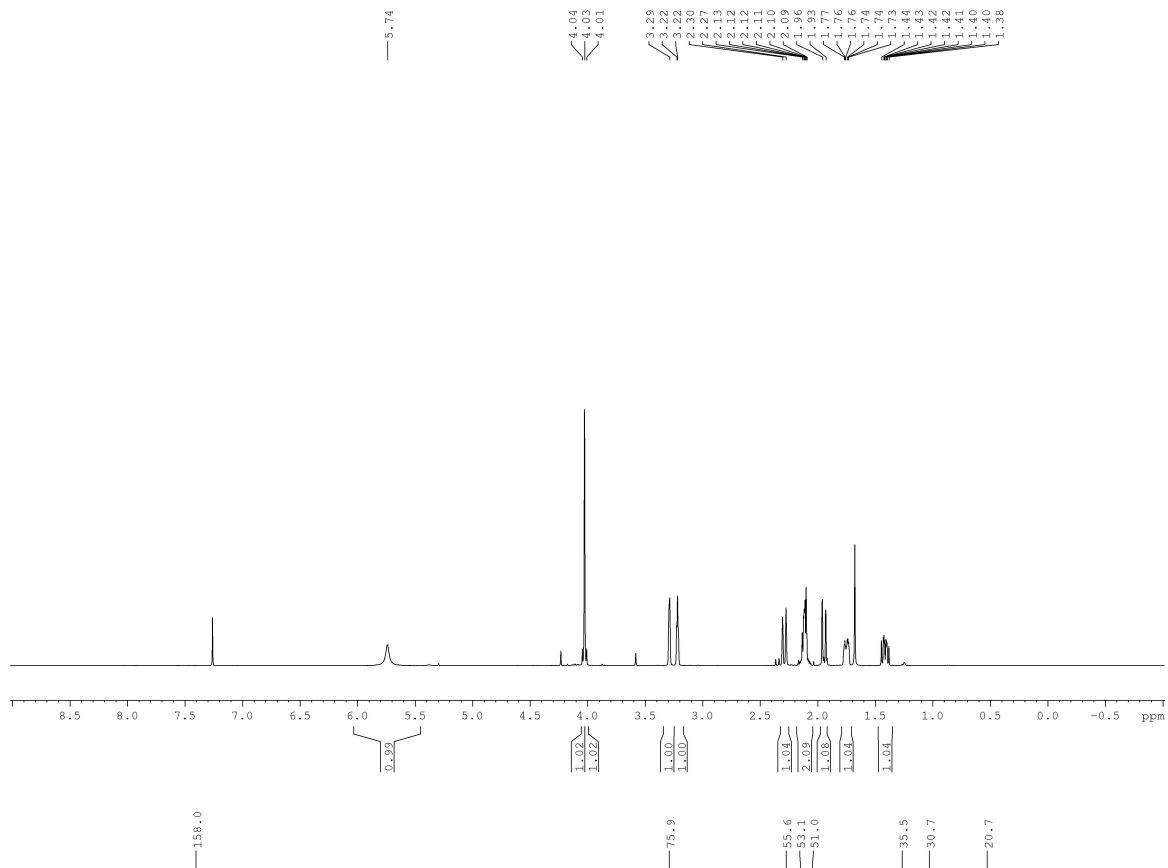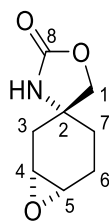
(1*R**,3*S**,6*S**)-7-oxaspiro[bicyclo[4.1.0]heptane-3,4'-oxazolidin]-2'-one **27b**
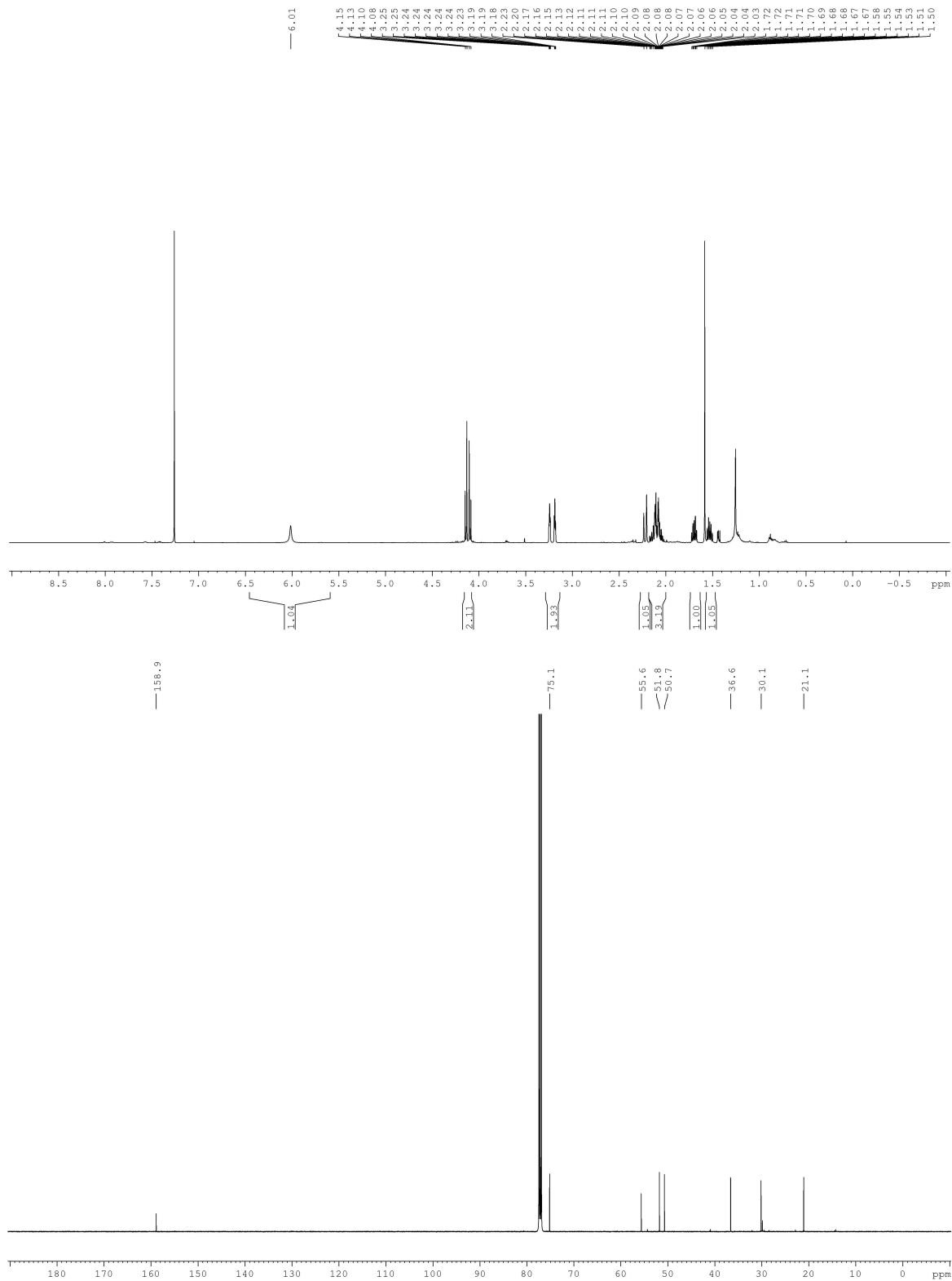
1-(4-Methoxybenzyl)-3-oxa-1-azaspiro[4.5]dec-7-ene-2-one **28**

^{1}H NMR peaks: 7.26, 7.26, 7.24, 6.85, 6.84, 6.83, 6.82, 6.81, 5.64, 5.63, 5.62, 5.55, 5.54, 5.53, 5.53, 5.52, 5.51, 5.50, 4.42, 4.38, 4.31, 4.27, 4.05, 4.03, 4.02, 4.00, 4.00, 3.79, 2.30, 2.25, 2.22, 2.21, 2.19, 2.18, 2.12, 2.12, 2.11, 2.10, 2.09, 2.08, 2.08, 2.07, 2.06, 2.06, 2.05, 2.04, 2.04, 2.03, 2.02, 2.02, 1.98, 1.98, 1.97, 1.94, 1.93, 1.93, 1.84, 1.83, 1.81, 1.80, 1.78, 1.77, 1.57, 1.55, 1.53, 1.52

Integrations: 2.09, 1.96, 1.01, 0.96, 0.99, 1.03, 0.95, 0.99, 3.00, 1.98, 1.02, 0.99, 1.00, 1.01

^{13}C NMR peaks: 159.1, 158.5, 130.6, 129.1, 126.8, 123.8, 114.0, 73.0, 60.0, 55.4, 43.8, 33.4, 30.6, 23.1

187
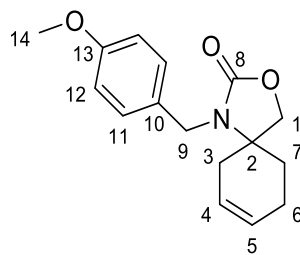
# (5R*,7R*,8S*)-7,8-Dihydroxy-1-(4-methoxybenzyl)-3-oxa-1-azaspiro[4.5]decan-2-one **29a**

(5*R**,7*S**,8*R**)-7,8-dihydroxy-1-(4-methoxybenzyl)-3-oxa-1-azaspiro[4.5]-decan-2-one **29b**

189
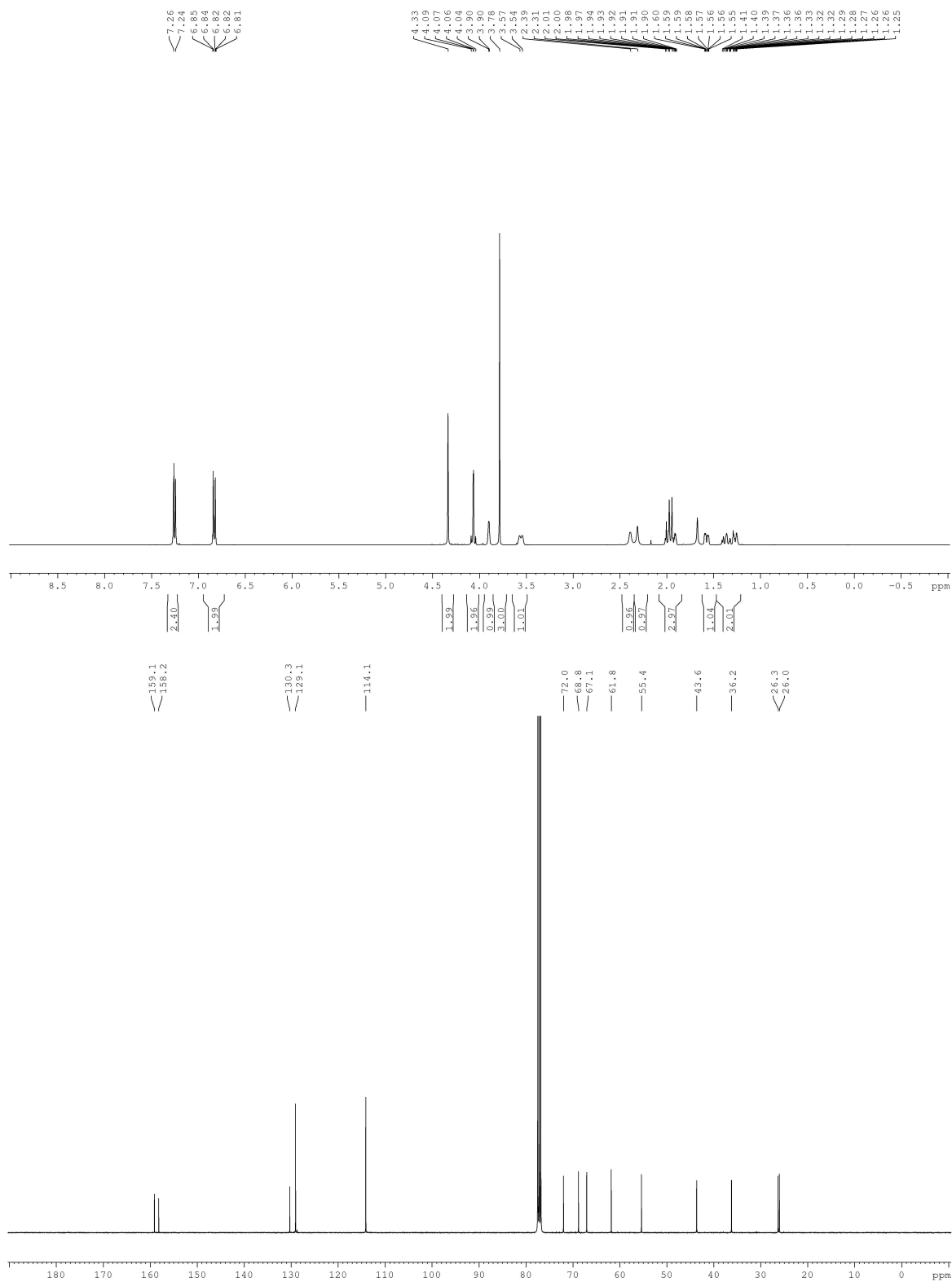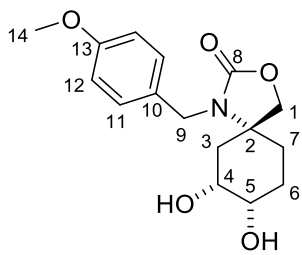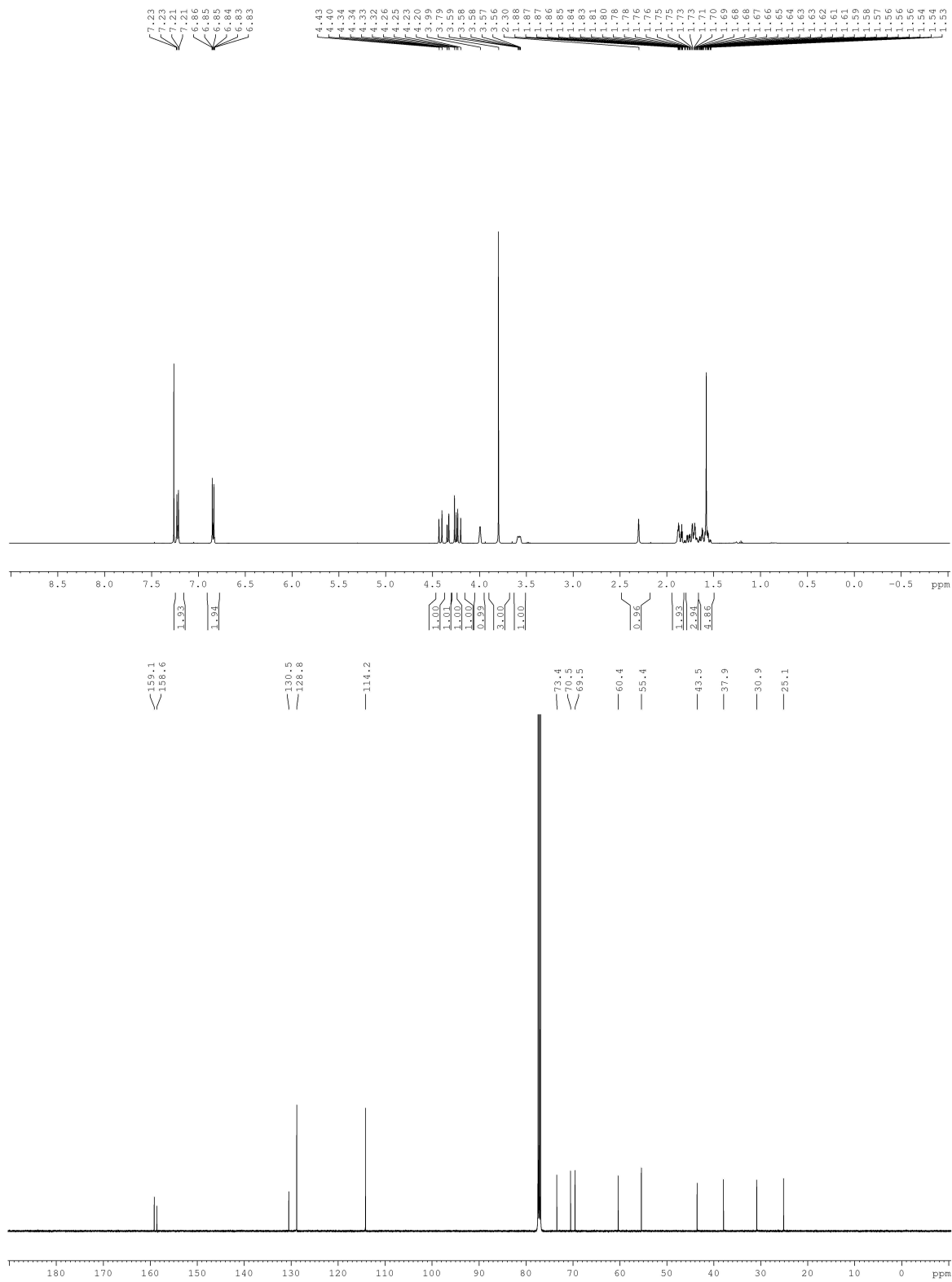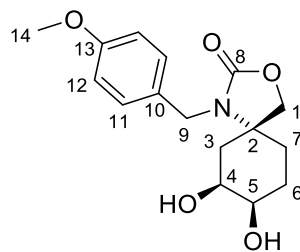
(5*R**,7*S**,8*S**)-7,8-Dibromo-1-(4-methoxybenzyl)-3-oxa-1-azaspiro[4.5]decan-2-one **30a**
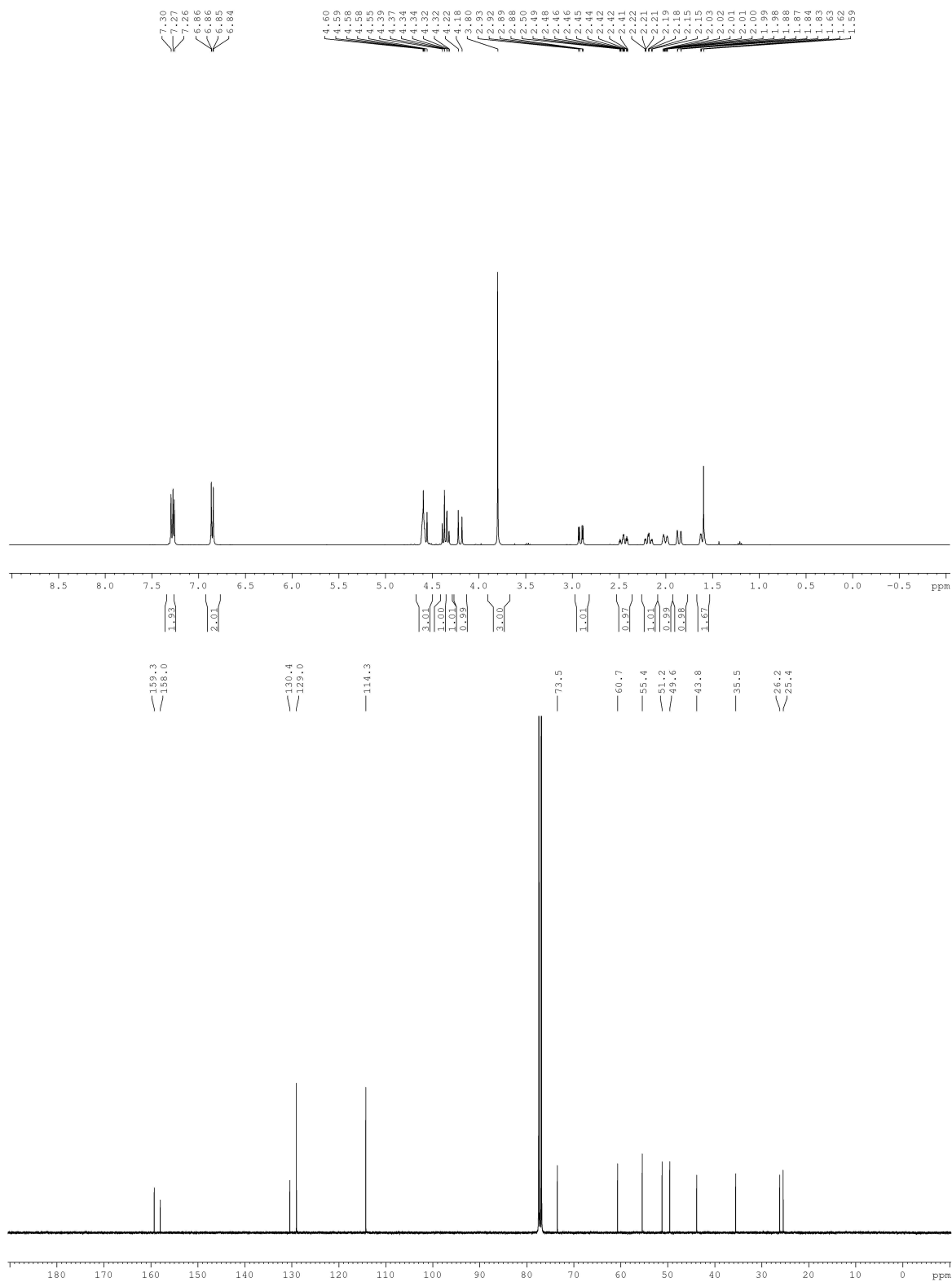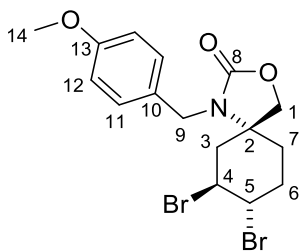
(5R*,7R*,8R*)-7,8-dibromo-1-(4-methoxybenzyl)-3-oxa-1-azaspiro[4.5]-decan-2-one **30b**
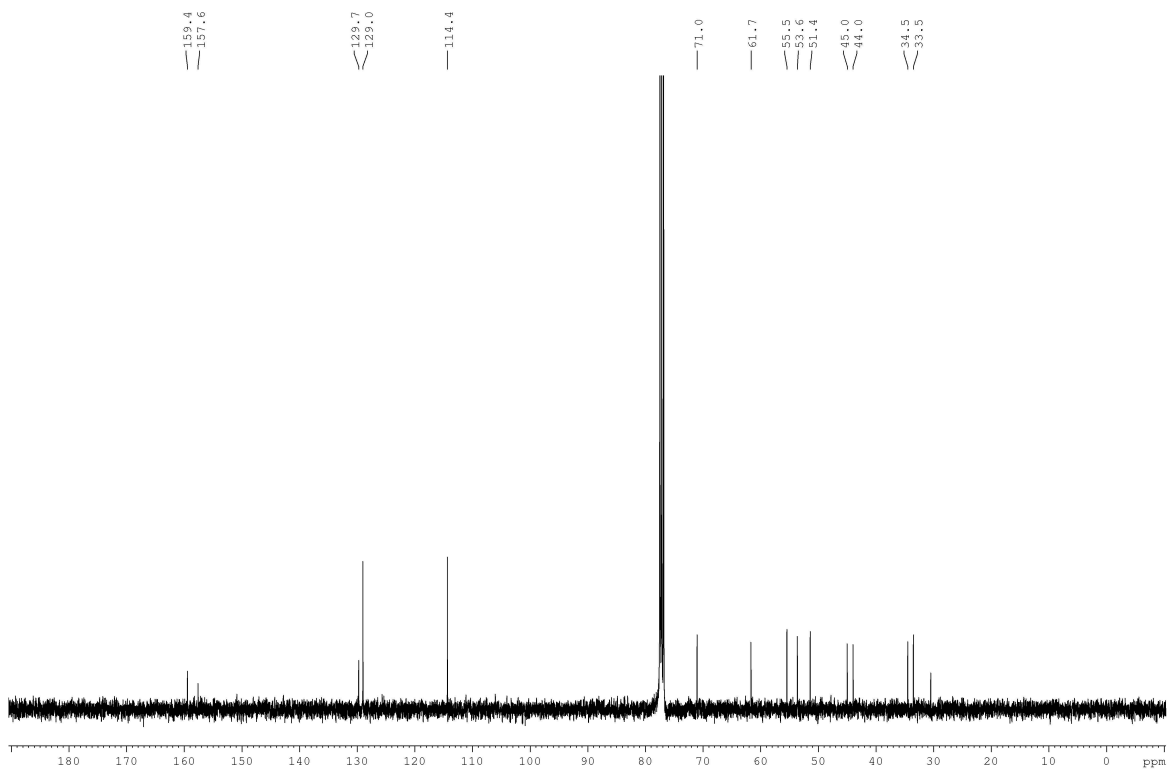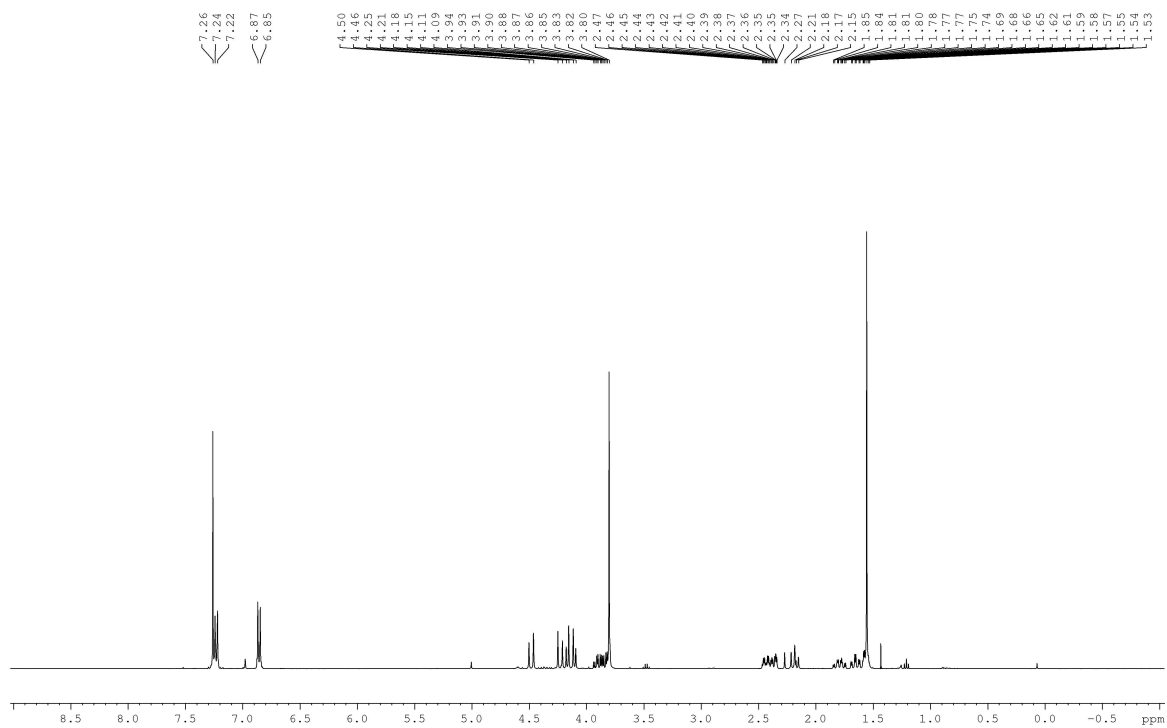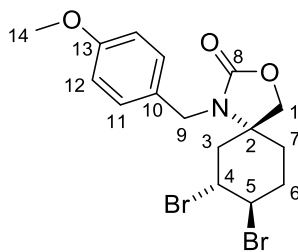
(1*R**,3*S**,6*S**)-3'-(4-Methoxybenzyl)-7-tosyl-7-azaspiro-[bicyclo[4.1.0]heptane-3,4'-oxazolidin]-2'-one **31a**

7.76 7.75 7.74 7.35 7.34 7.21 7.20 7.19 7.18 6.84 6.83 6.82 6.81

4.39 4.36 4.21 4.18 4.07 4.05 3.97 3.95 3.79 3.03 3.02 3.01 2.97 2.95 2.94 2.46 2.17 2.16 2.15 2.14 2.13 2.11 1.91 1.88 1.86 1.85 1.84 1.83 1.82 1.80 1.79 1.53 1.52 1.51 1.49 1.48 1.47 1.41 1.40 1.39 1.38

159.3 158.1 145.0 135.2 130.2 130.1 128.9 127.7 114.2 72.3 60.0 55.4 43.7 40.9 36.7 32.3 28.2 21.8 20.2

192

(1*R**,3*R**,6*S**)-3'-(4-methoxybenzyl)-7-tosyl-7-azaspiro[bicyclo[4.1.0]heptane-3,4'-oxazolidin]-2'-one **31b**
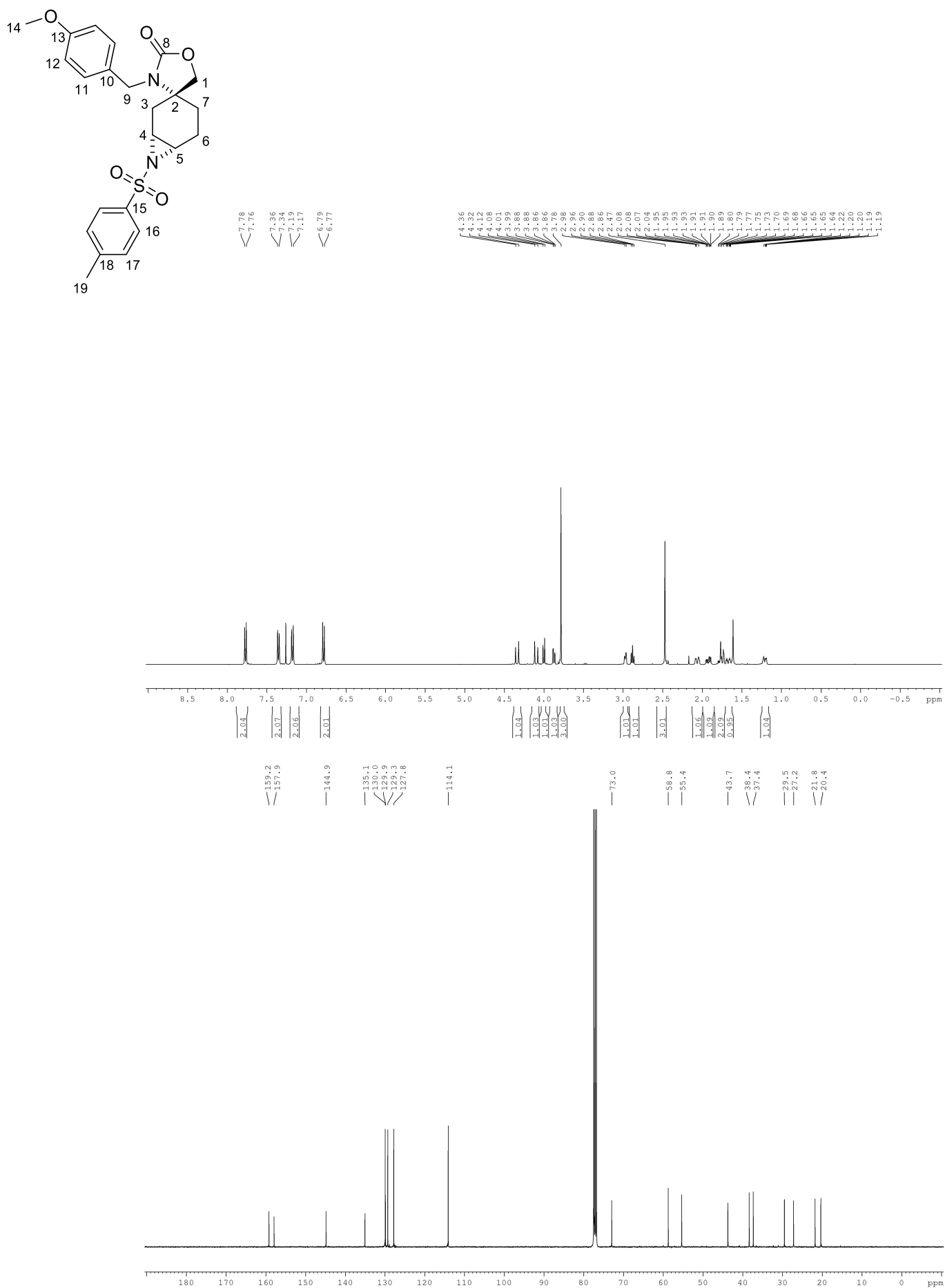
(1*R**,3*R**,6*S**)-7,7-Difluoro-3'-(4-methoxybenzyl)spiro-[bicycle[4.1.0]heptane-3,4'-oxazolidin]-2'-one **32a**
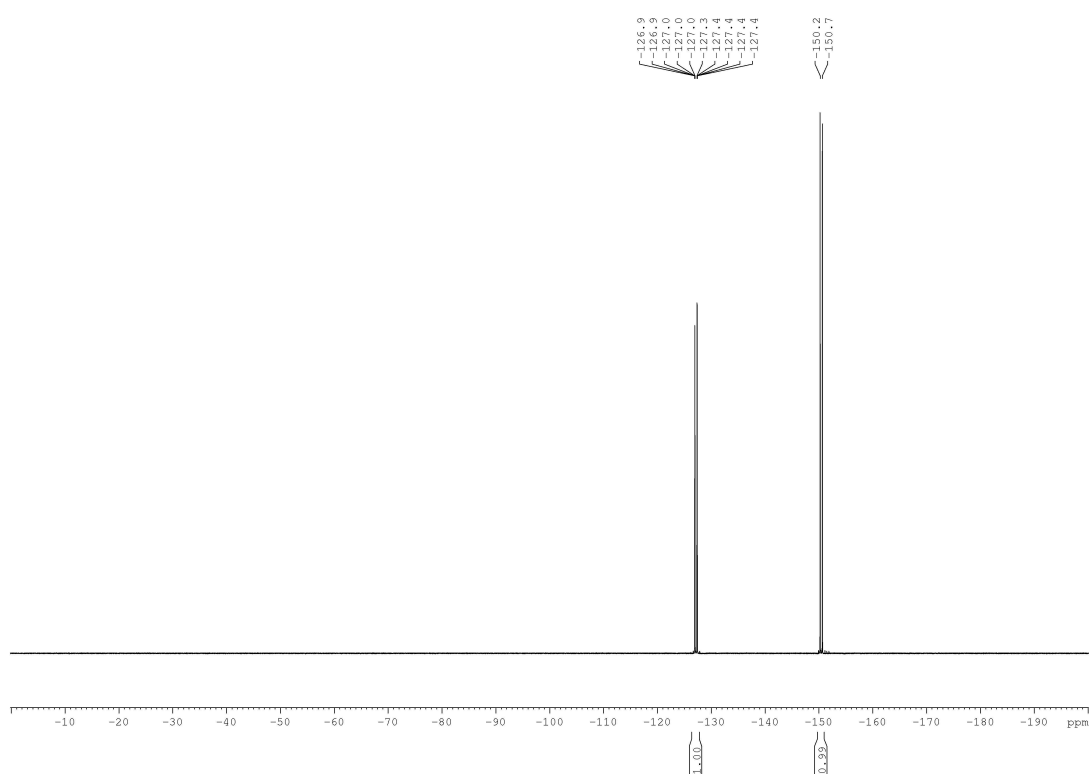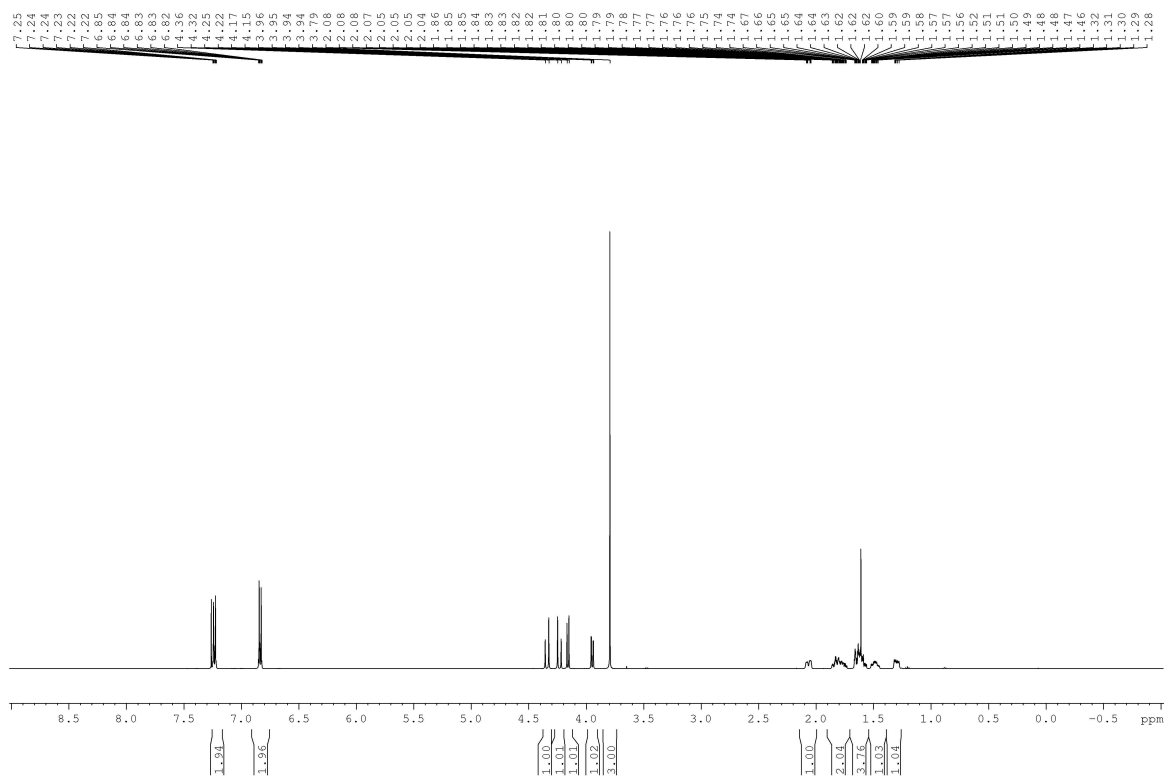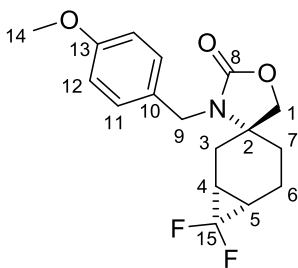
(1R*,3R*,6S*)-7,7-Difluoro-3'-(4-methoxybenzyl)spiro-[bicycle[4.1.0]heptane-3,4'-oxazolidin]-2'-one **32a**

195
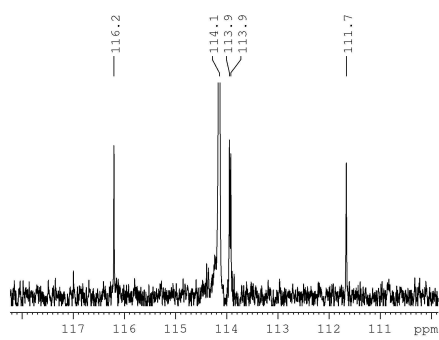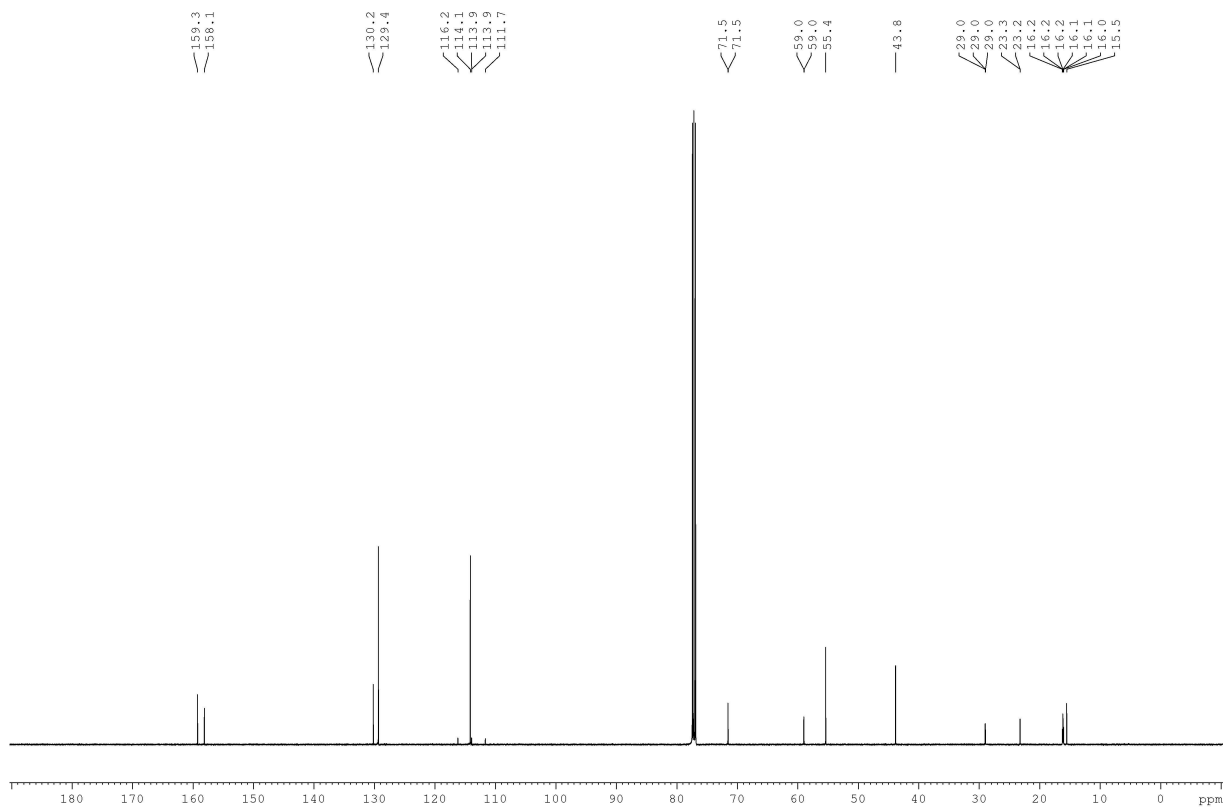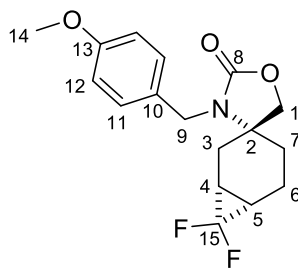
(1R*,3S*,6S*)-7,7-difluoro-3'-(4-methoxybenzyl)spiro[bicyclo[4.1.0]heptane-3,4'-oxazolidin]-2'-one **32b**
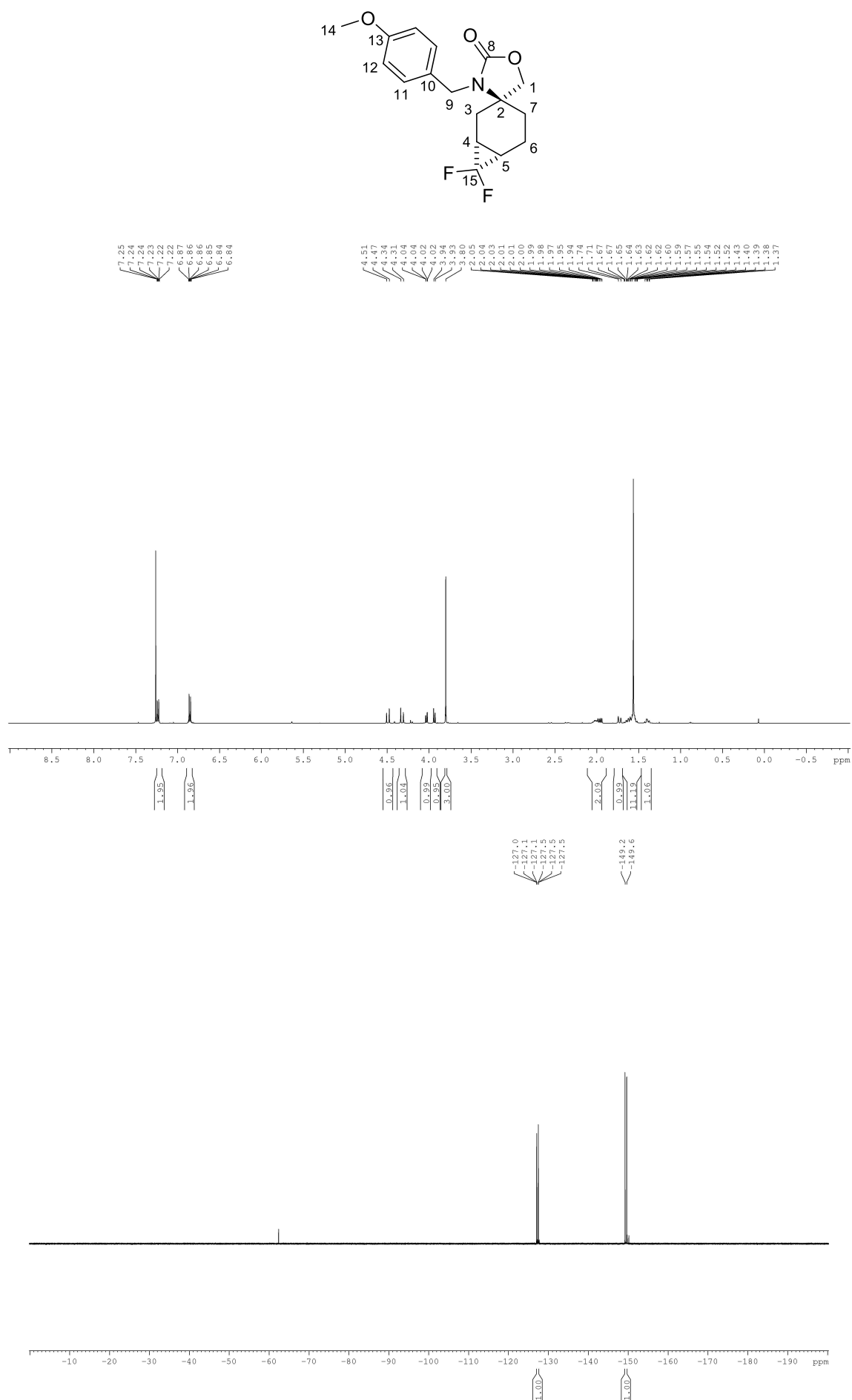
(1*R**,3*S**,6*S**)-7,7-difluoro-3'-(4-methoxybenzyl)spiro[bicyclo[4.1.0]heptane-3,4'-oxazolidin]-2'-one **32b**
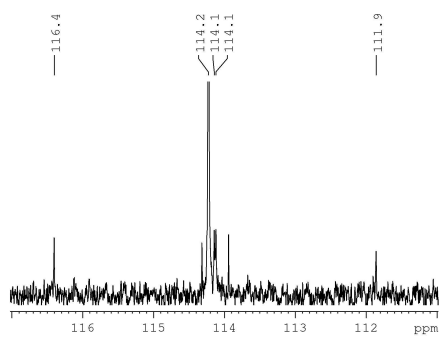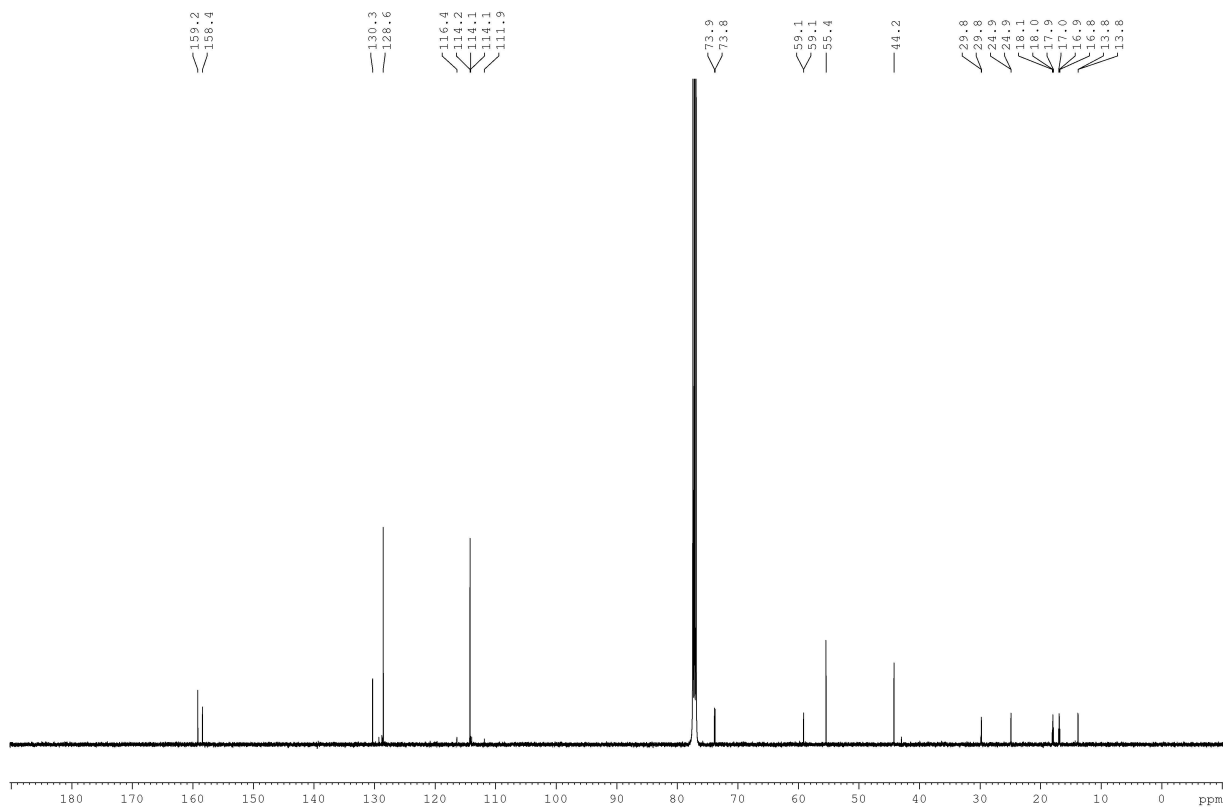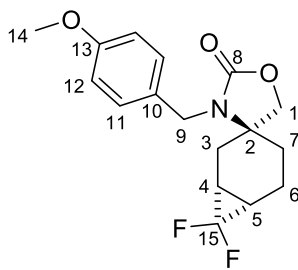
1-(4-Methoxybenzyl)-3-oxa-1-azaspiro[4.5]decane-2,7-dione **33a**
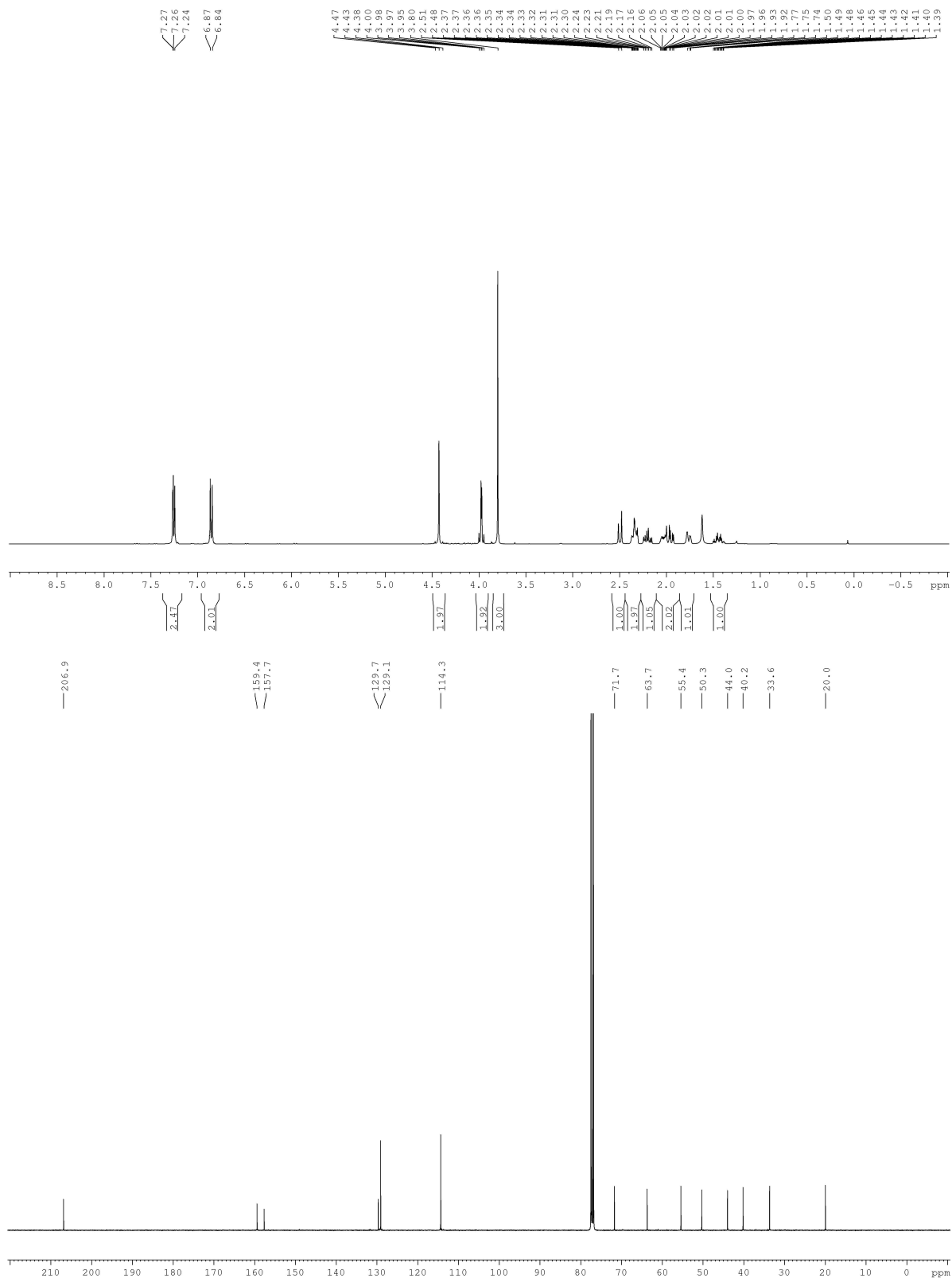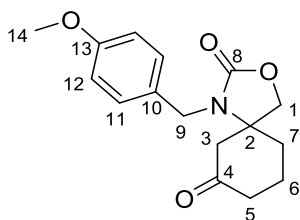
1-(4-Methoxybenzyl)-3-oxa-1-azaspiro[4.5]decane-2,8-dione **33b**
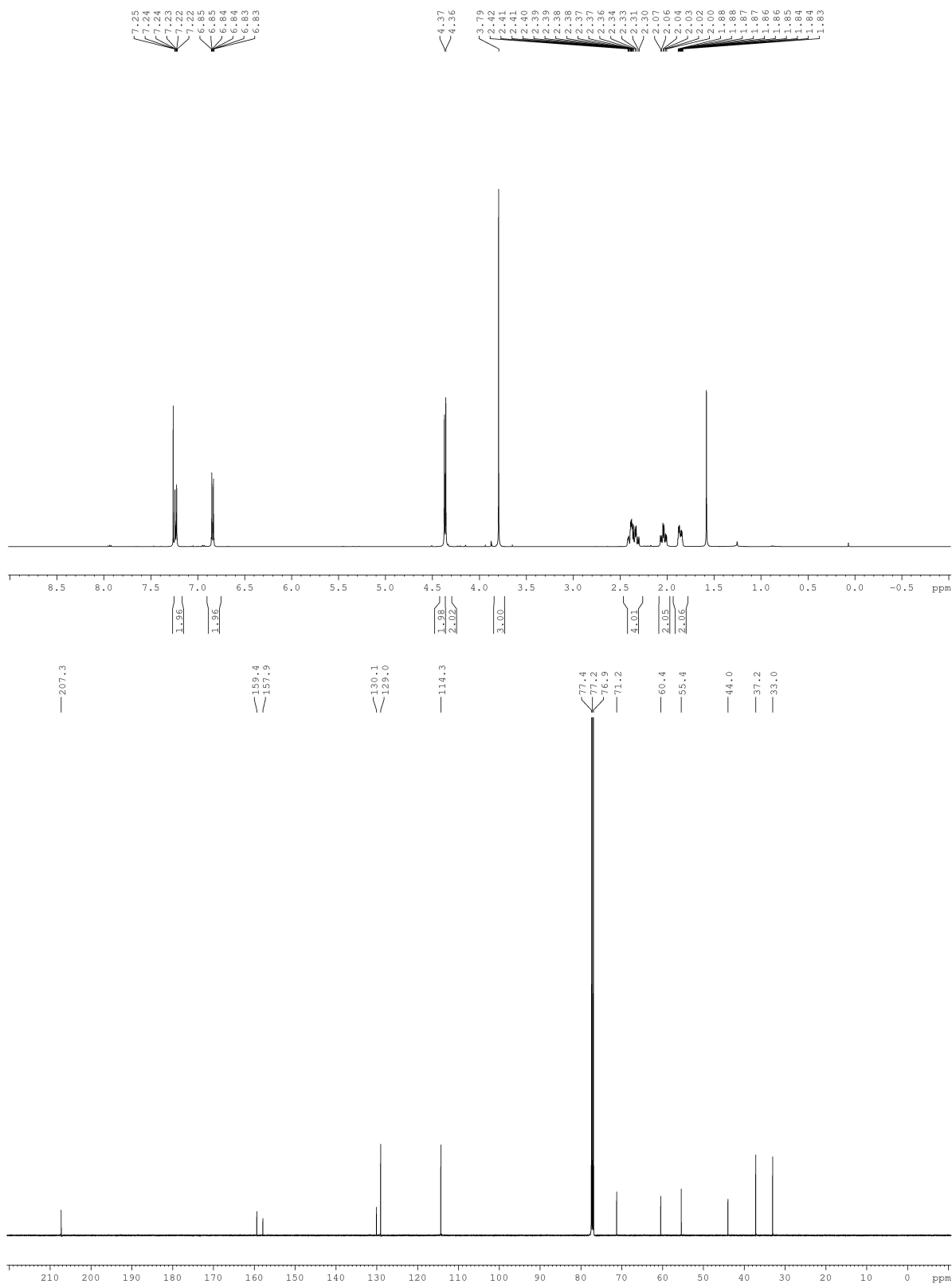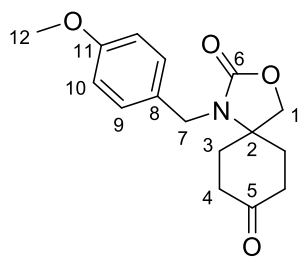
8-hydroxy-1-(4-methoxy-benzyl)-3-oxa-1-azaspiro[4.5]decan-2-one 34a
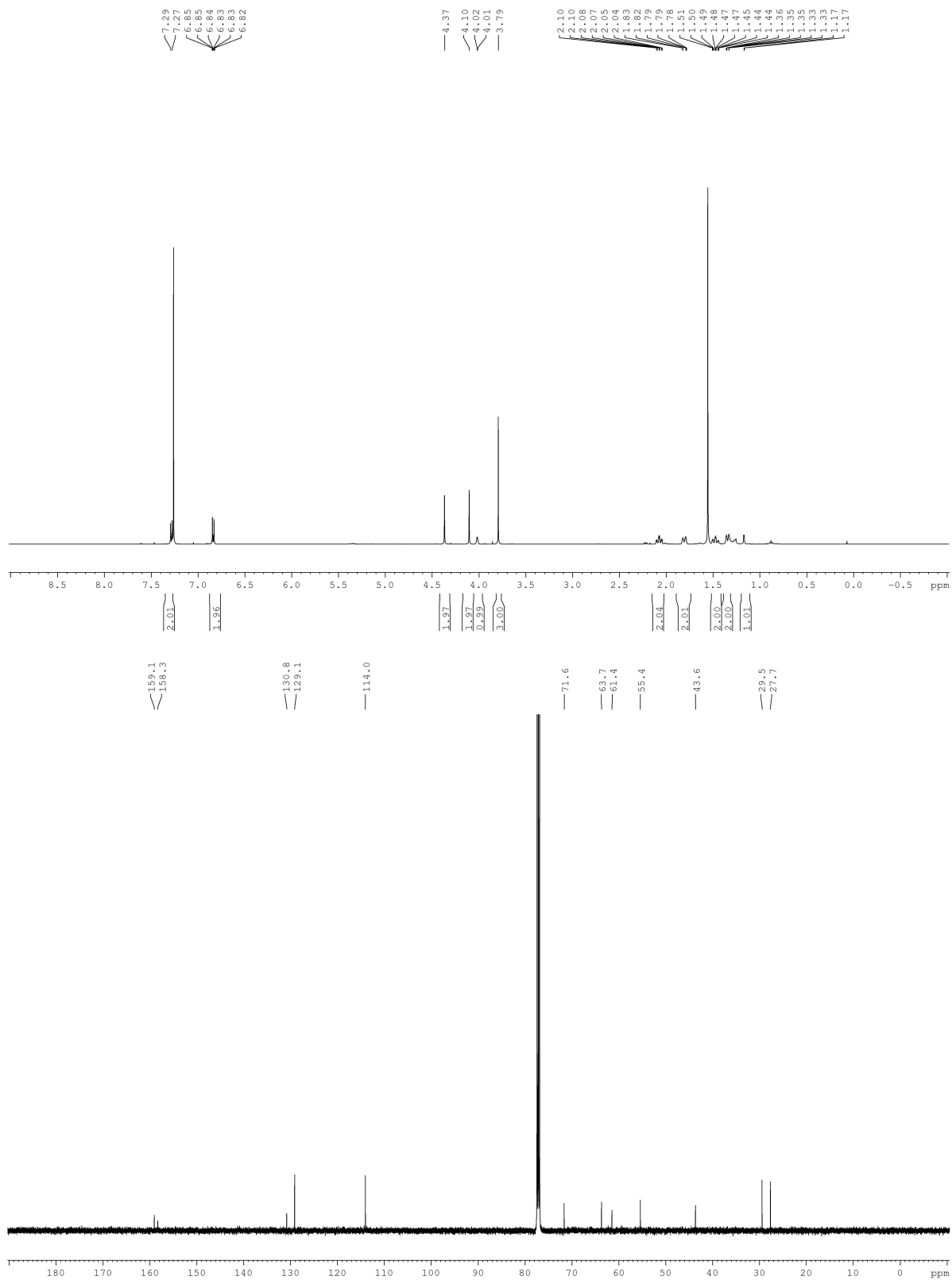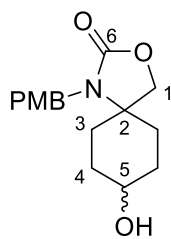
(5*R**,7*S**)-7-Hydroxy-1-(4-methoxy-benzyl)-3-oxa-1-azaspiro[4.5]decan-2-one **34b**
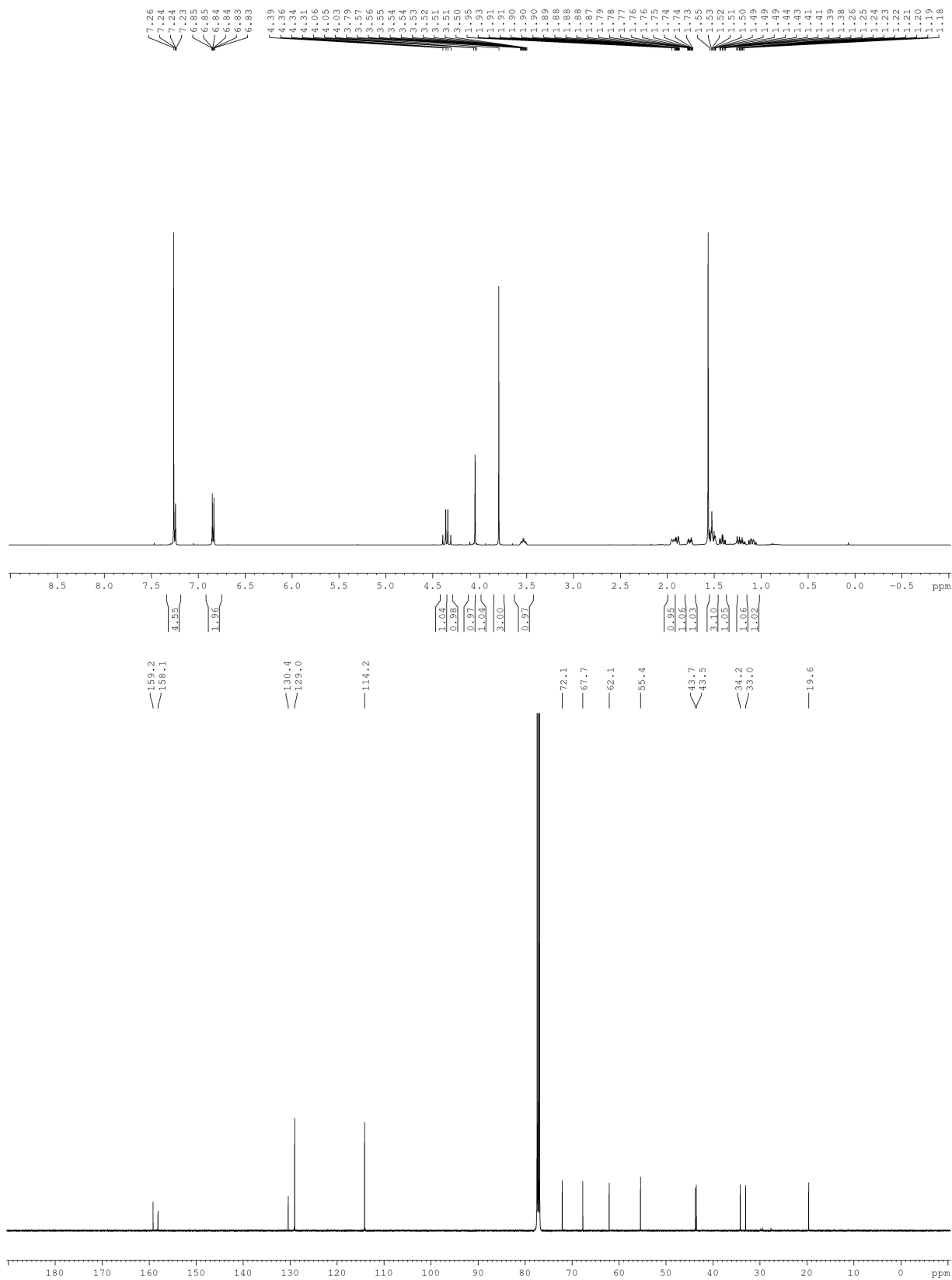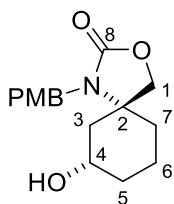
8-hydroxy-1-(4-methoxybenzyl)-3-oxa-1-azaspiro[4.5]decan-2-one **34c**
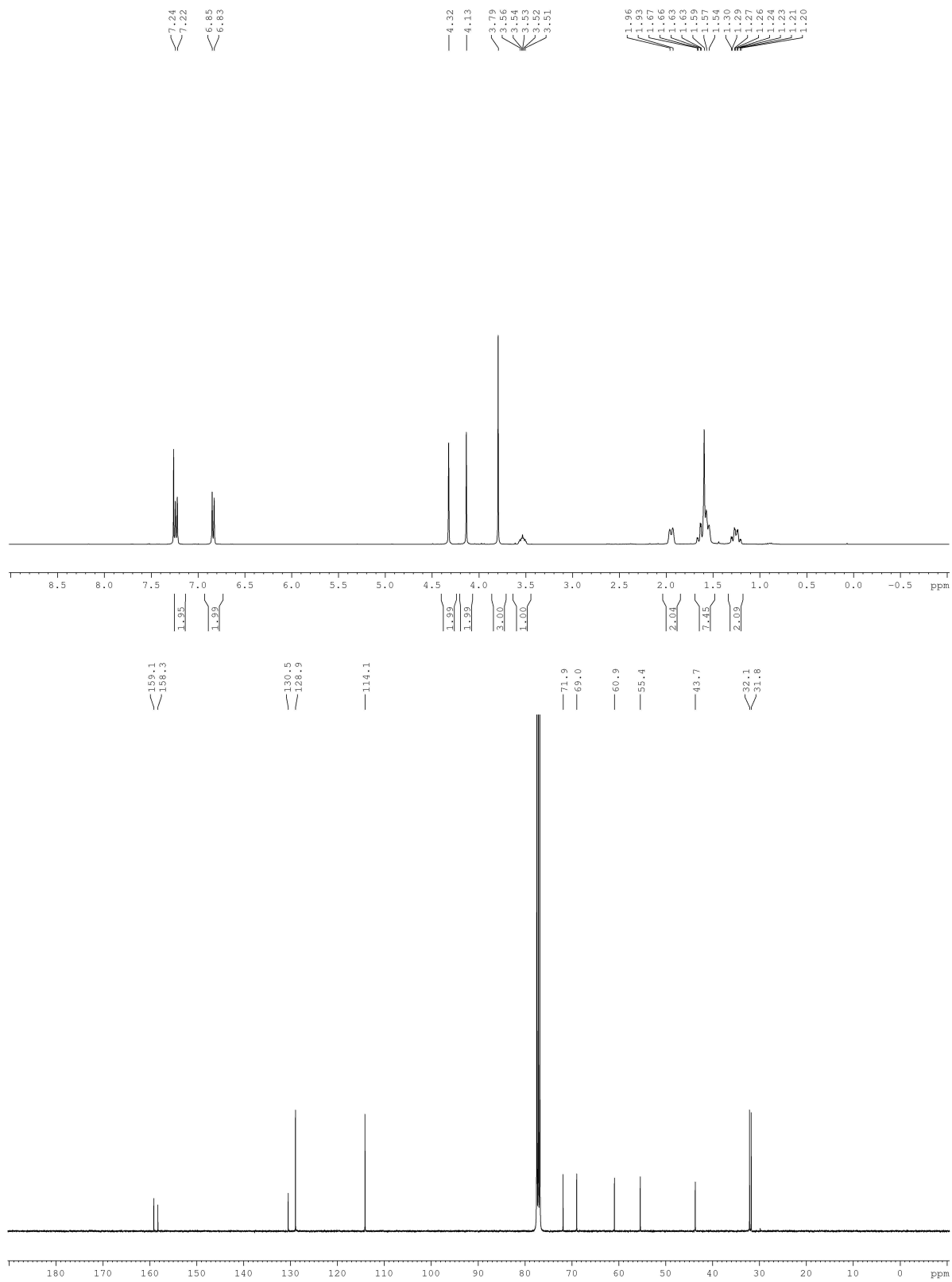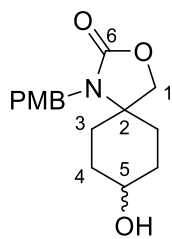
(5R*,7S*,8S*)-7,8-Dibromo-3-oxa-1-azaspiro[4.5]decan-2-one 35
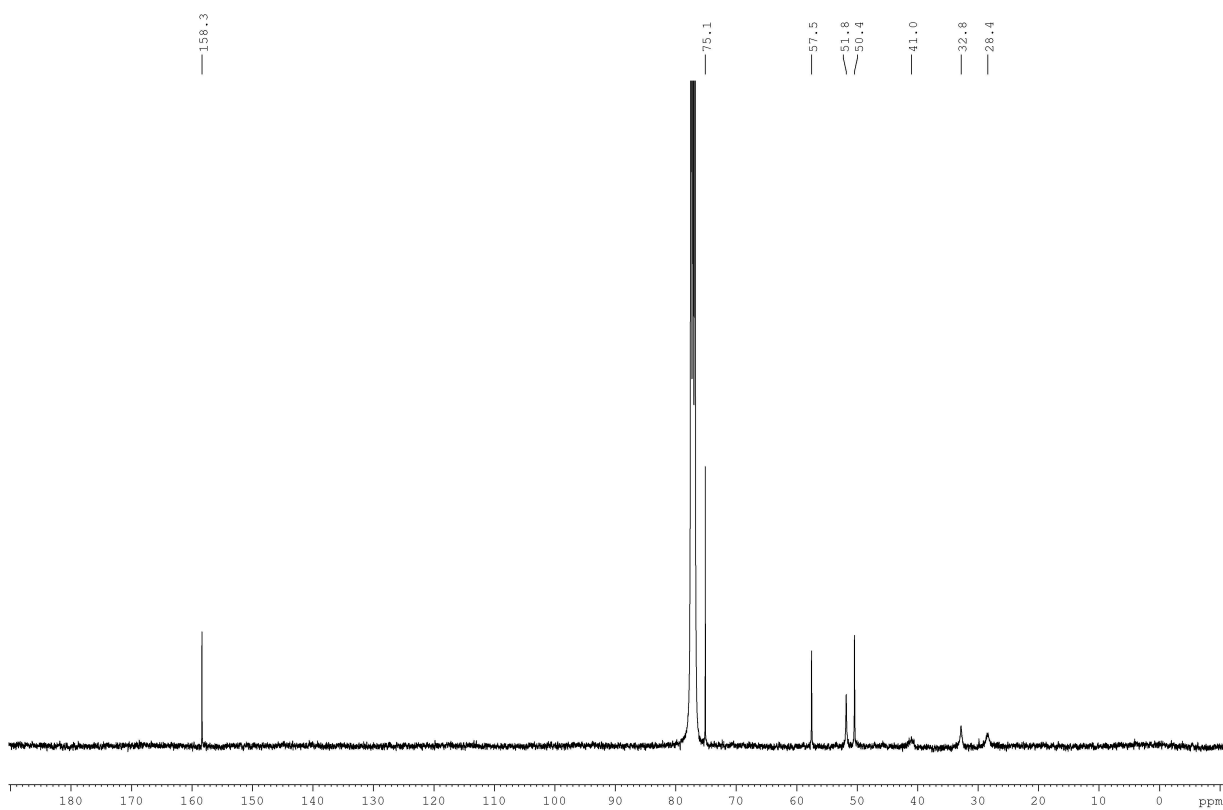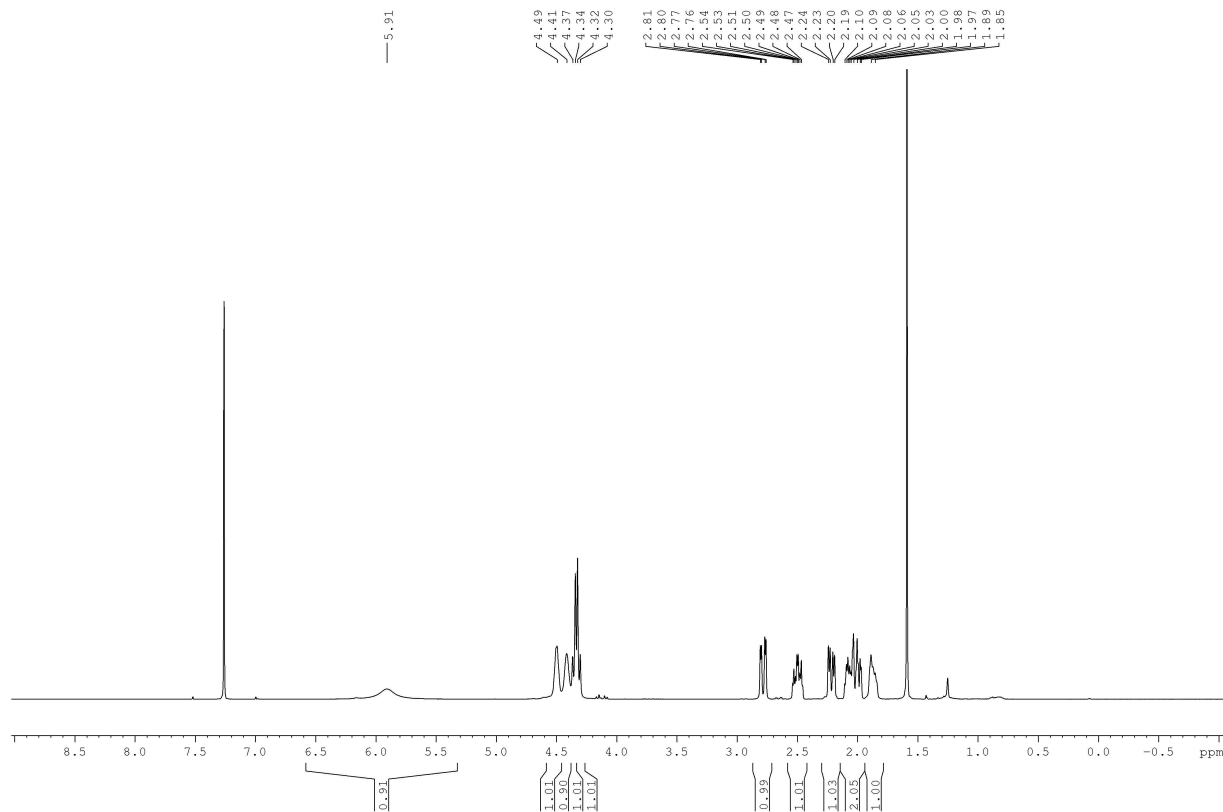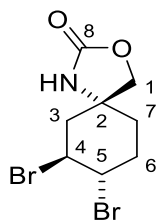
tert-Butyl (2-(5-bromo-1*H*-indole-3-carboxamido)ethyl)carbamate **47a**
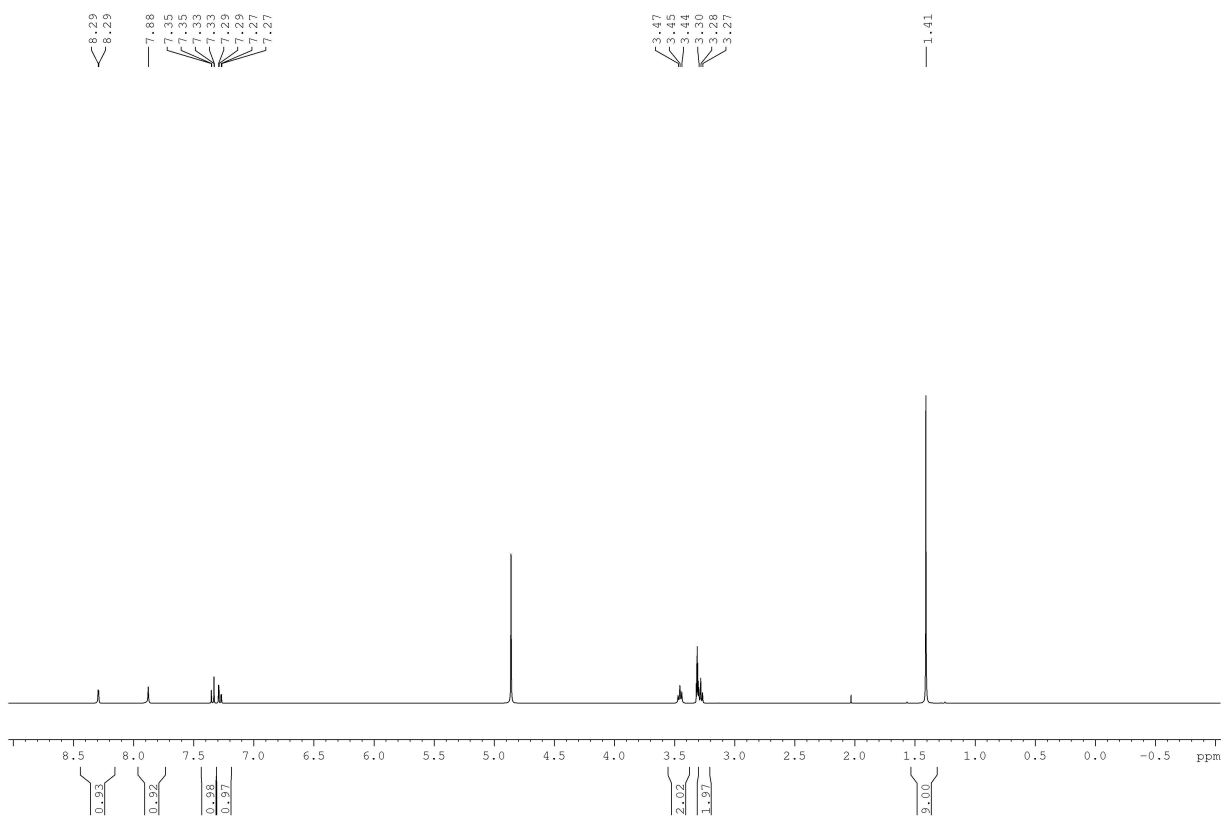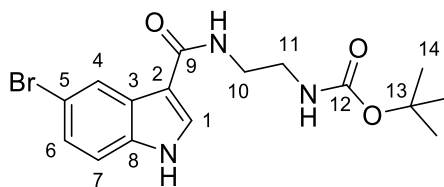
tert-Butyl (3-(5-bromo-1*H*-indole-3-carboxamido)propyl)carbamate **47b**
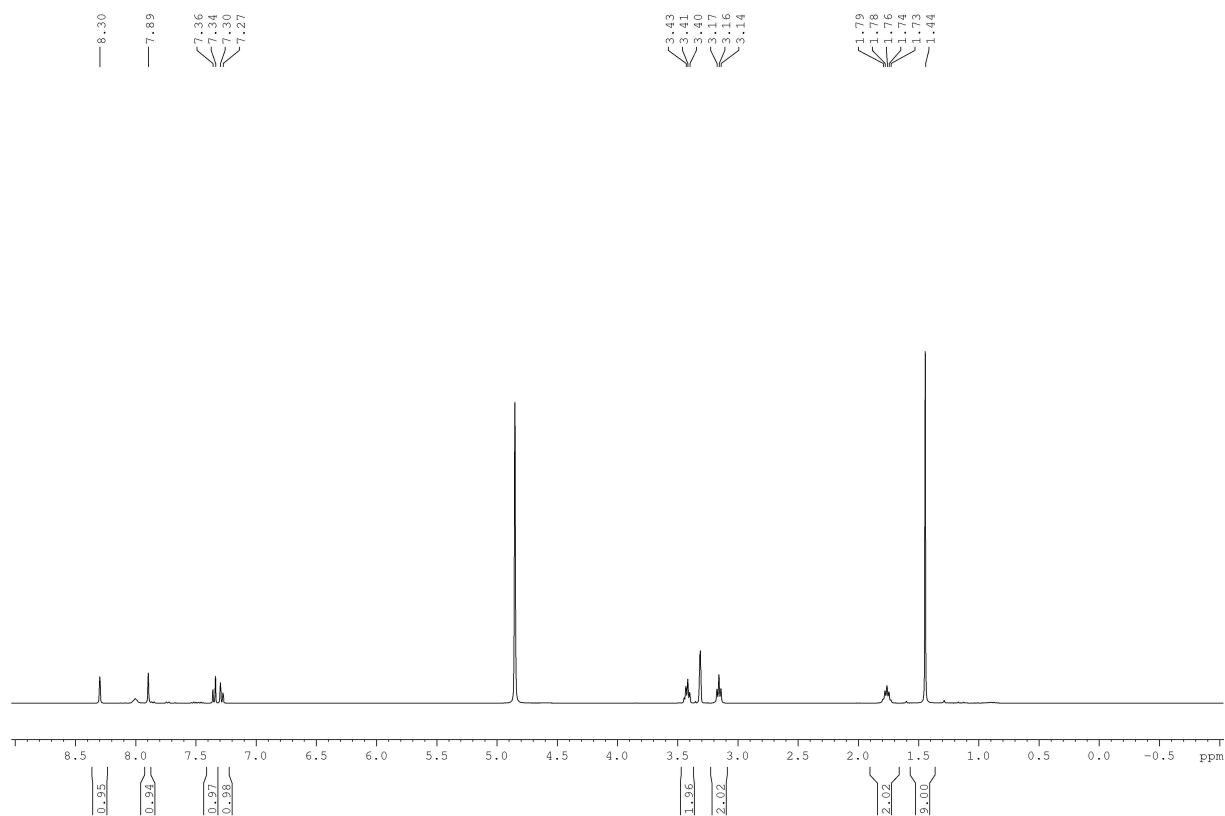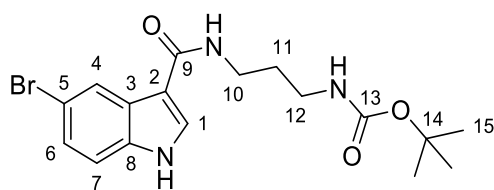
4-((2-(5-Bromo-1*H*-indole-3-carboxamido)ethyl)amino)-4-oxobutanoic acid **51a**
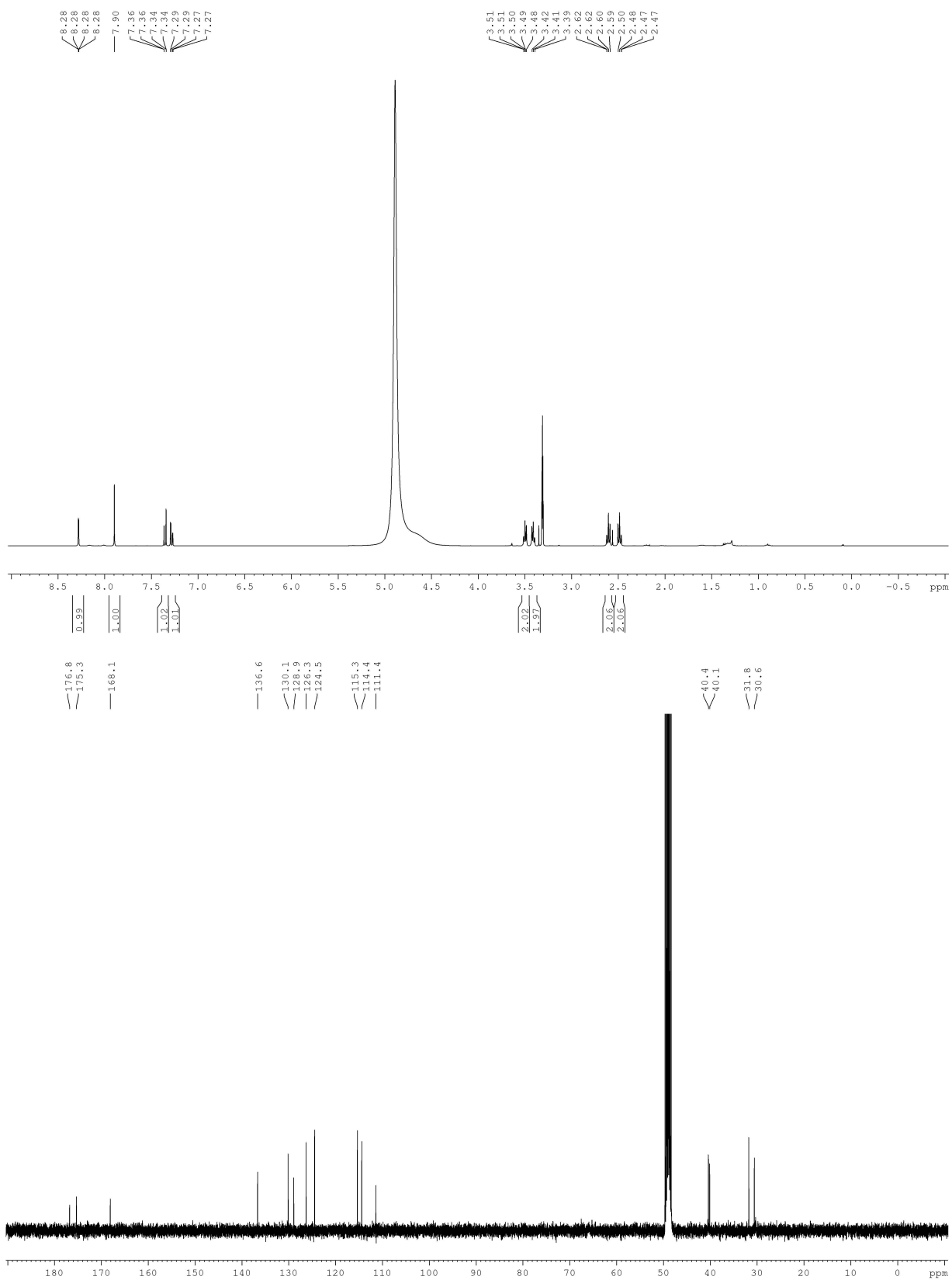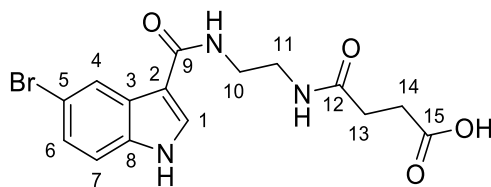
4-((3-(5-Bromo-1*H*-indole-3-carboxamido)propyl)amino)-4-oxobutanoic acid **51c**
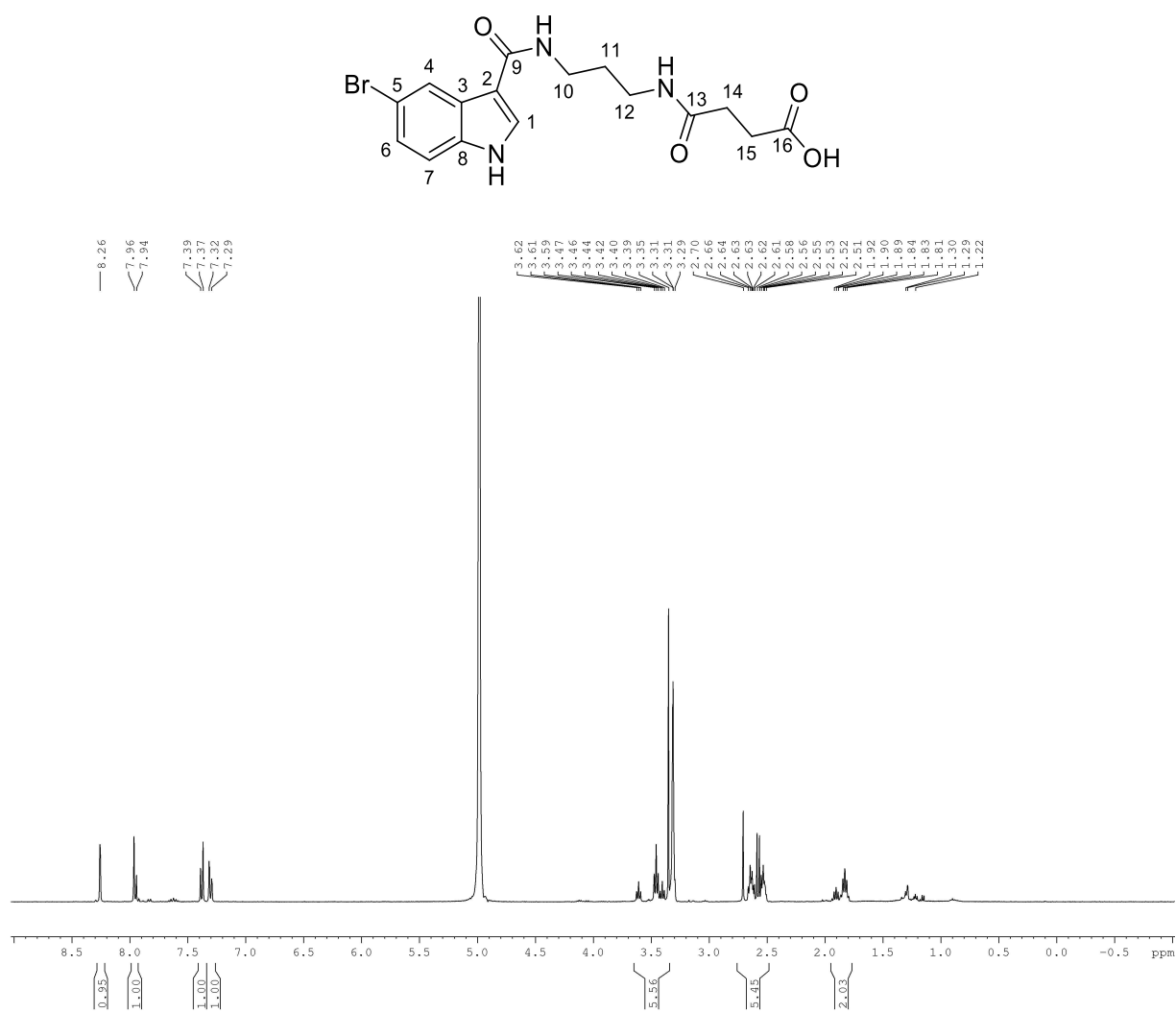
Ethyl 4-((2-(5-Bromo-1*H*-indole-3-carboxamido)ethyl)amino)-4-oxobutanoate **52**
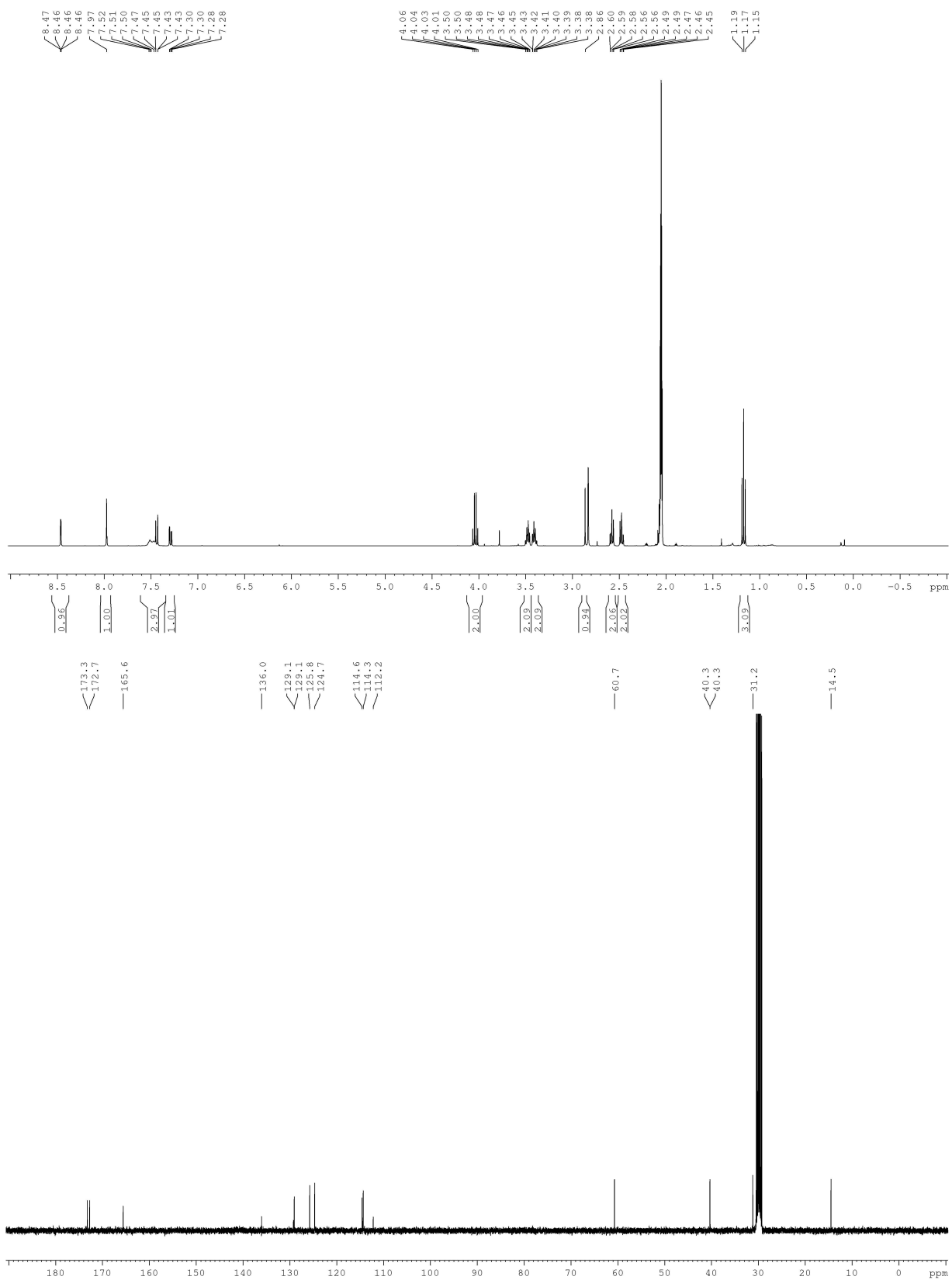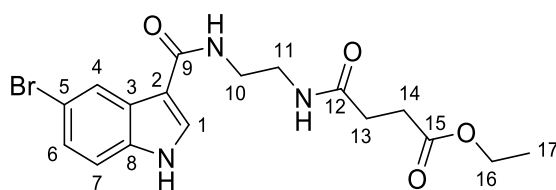
4-oxo-4-(phenethylamino)butanoic acid **53a**
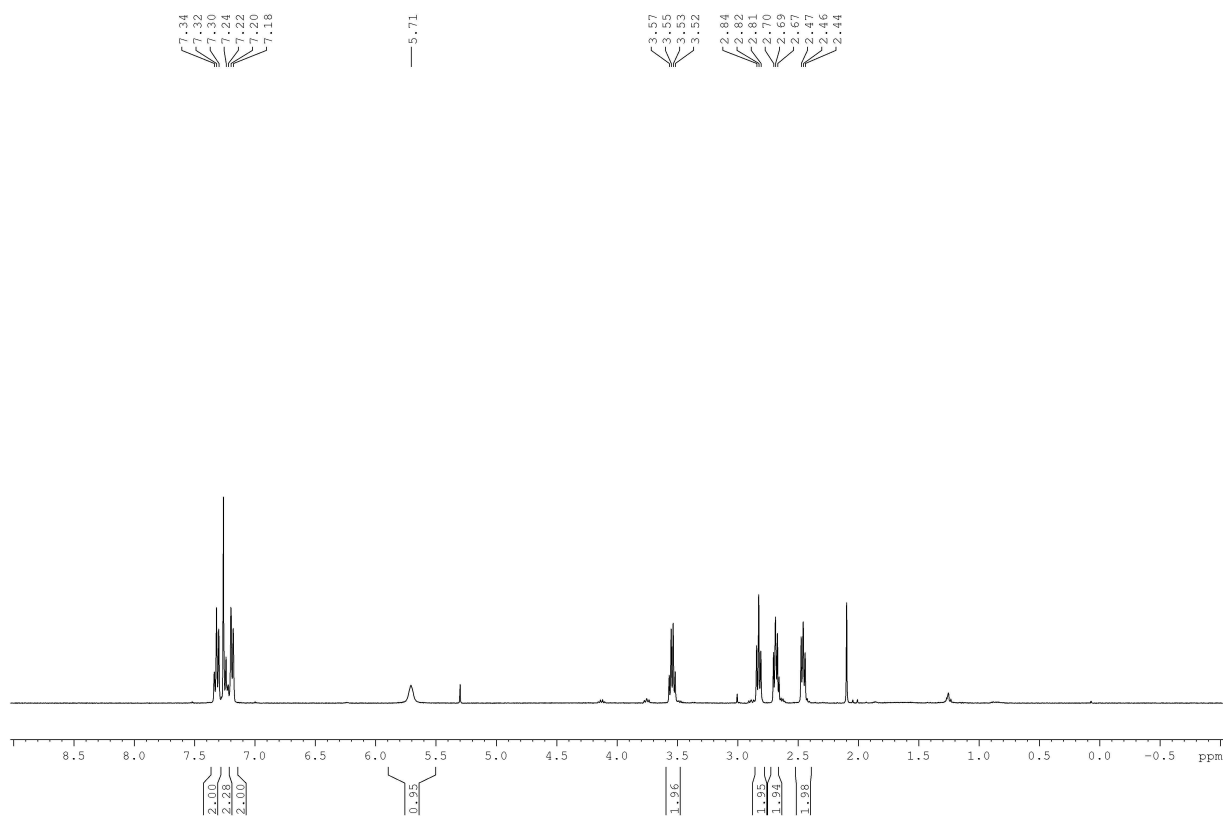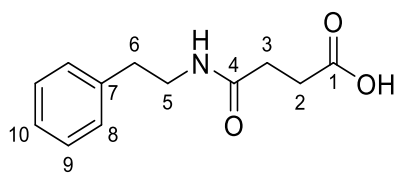
4-oxo-4-((3-phenylpropyl)amino)butanoic acid **53b**
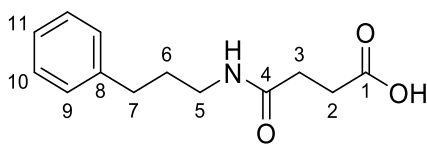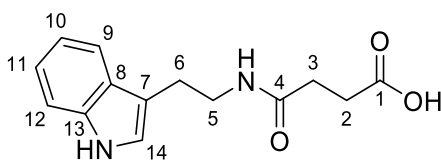

7.31
7.29
7.27
7.22
7.20
7.19
7.17

5.64

3.33
3.32
3.30
3.29

2.69
2.68
2.66
2.64
2.46
2.44
2.42

1.89
1.88
1.86
1.84
1.82

8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0 −0.5 ppm
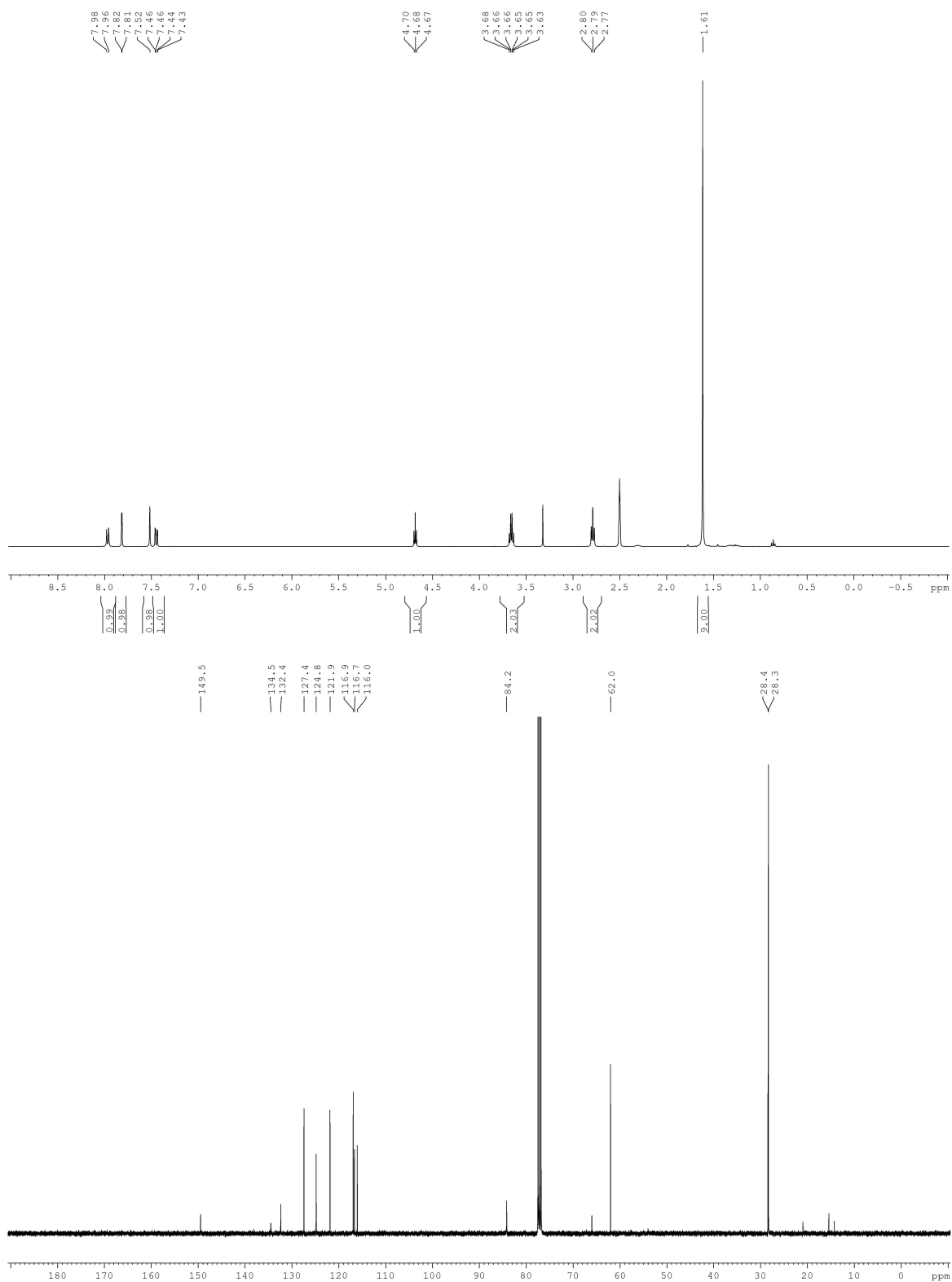
4.21
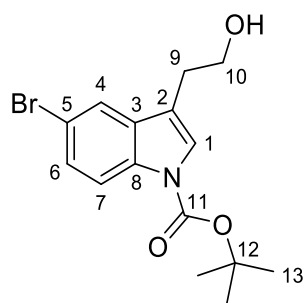3.10

0.92

1.82

4.08
1.86

2.00
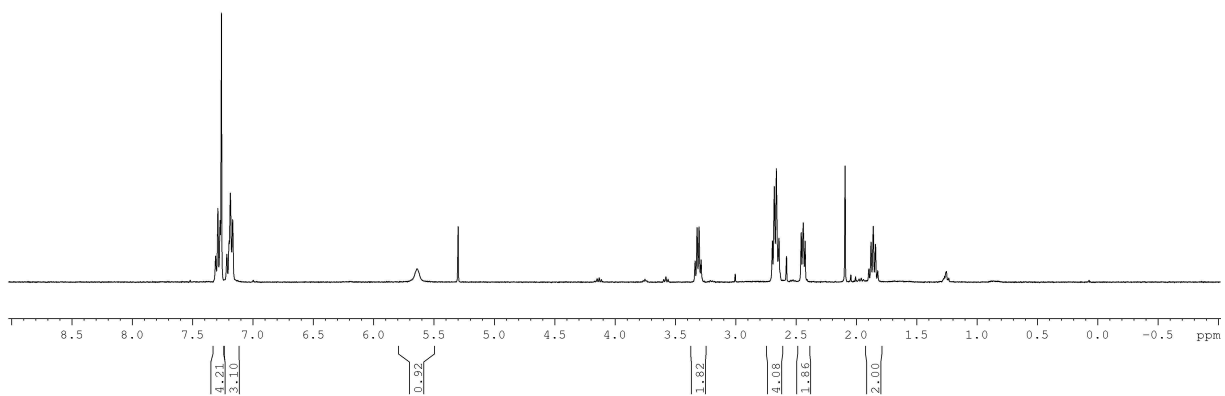
4-oxo-4-((4-phenylbutyl)amino)butanoic acid **53c**
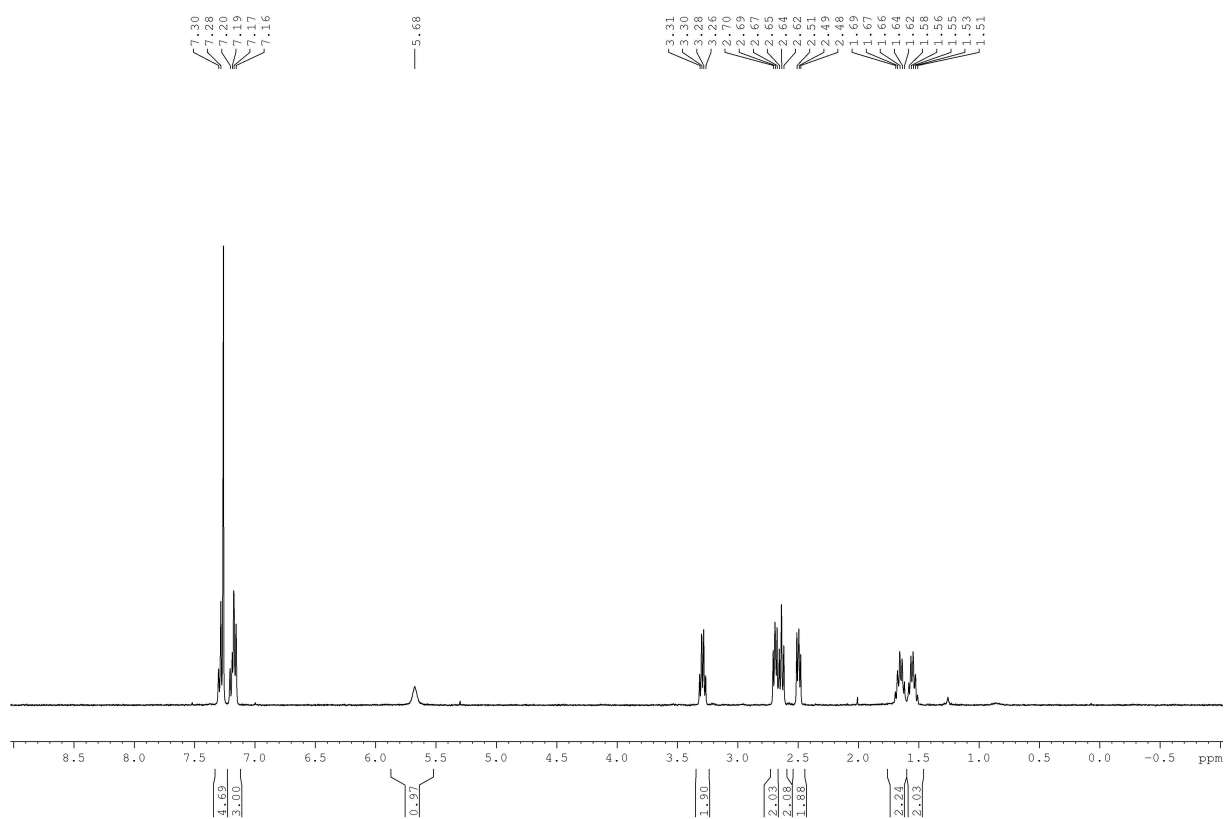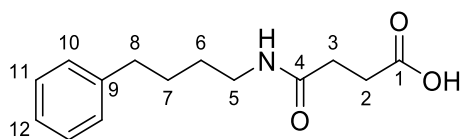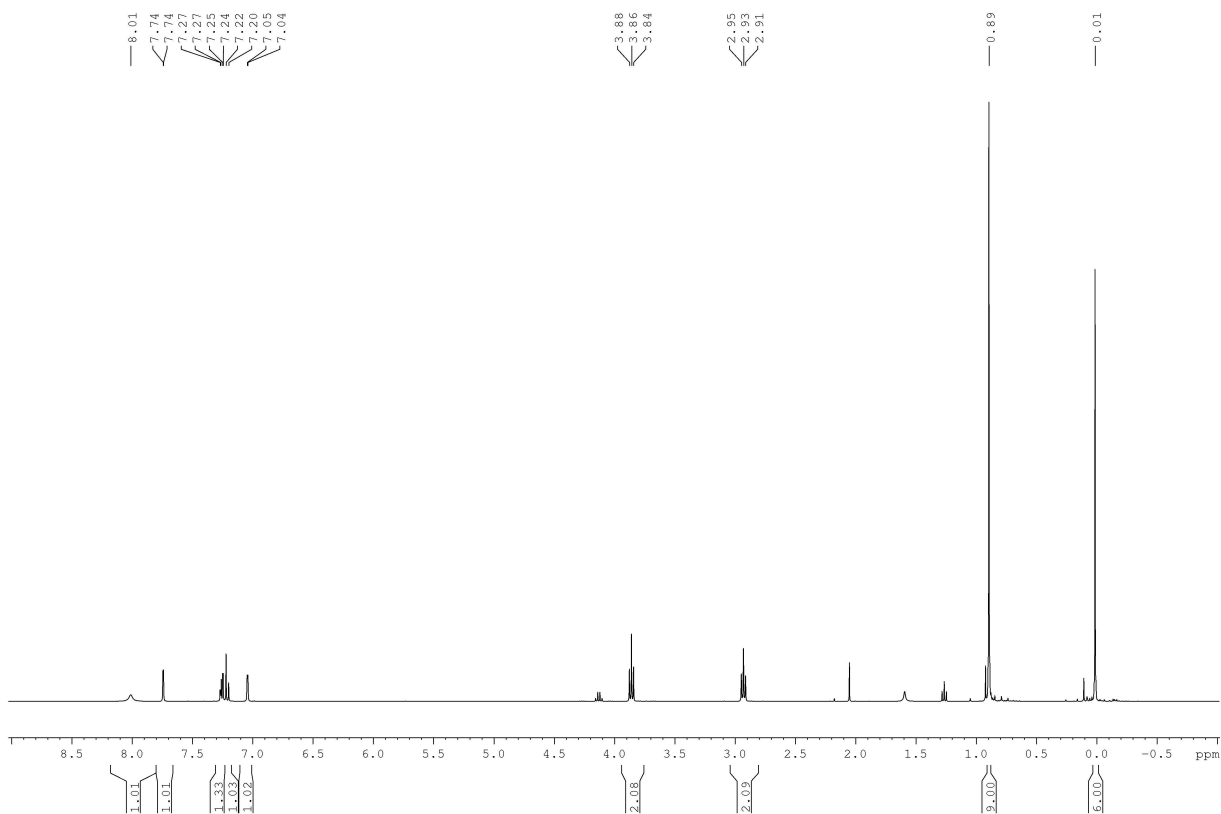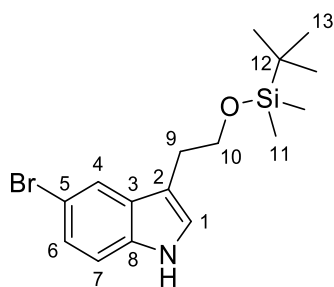
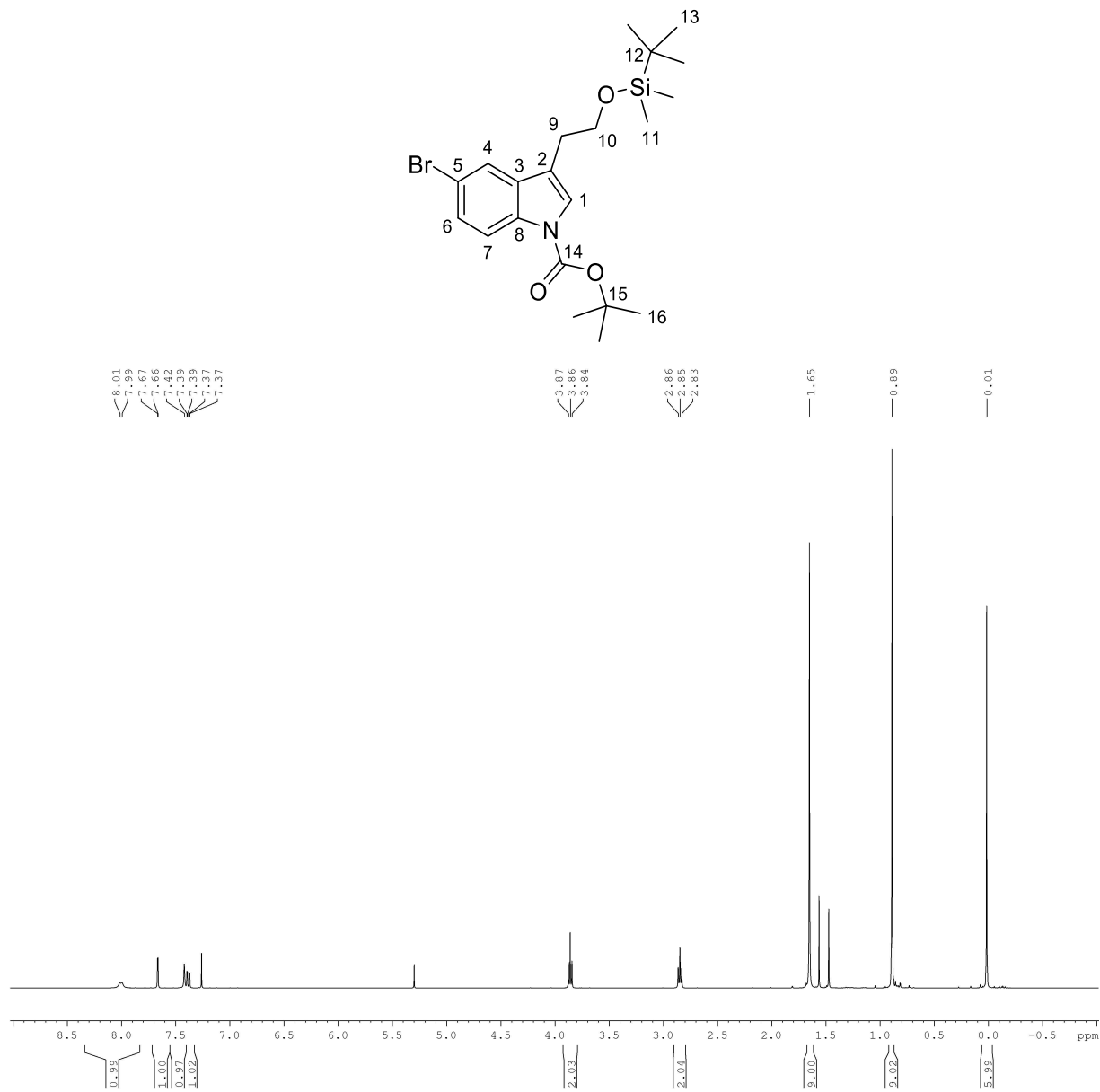# 4-((2-(1H-indol-3-yl)ethyl)amino)-4-oxobutanoic acid **53d**
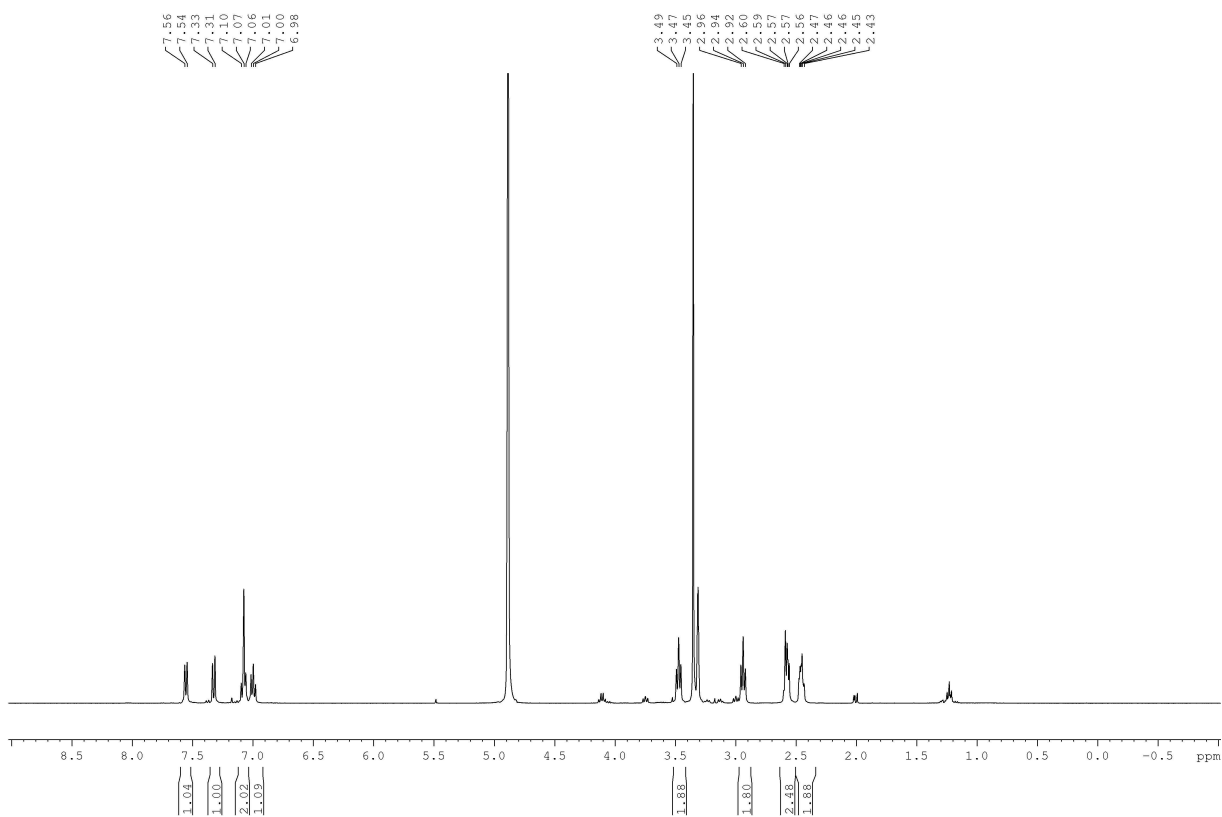
tert-Butyl 5-bromo-3-(2-hydroxyethyl)-1*H*-indole-1-carboxylate **56**
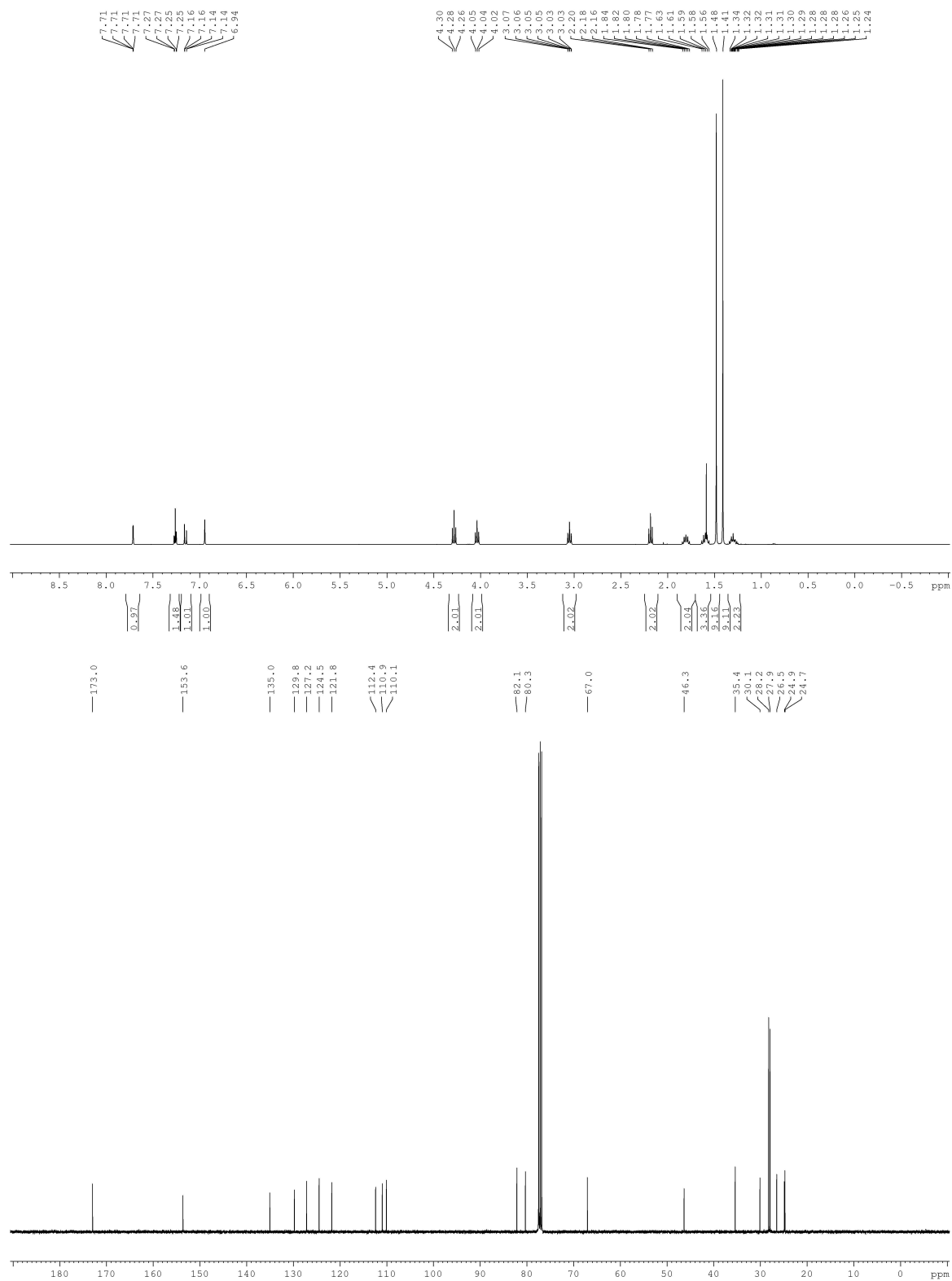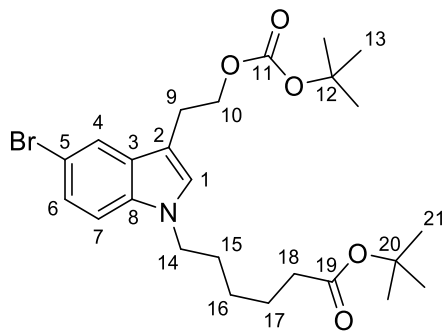
5-Bromo-3-(2-((*tert*-butyldimethylsilyl)oxy)ethyl)-1*H*-indole **57**

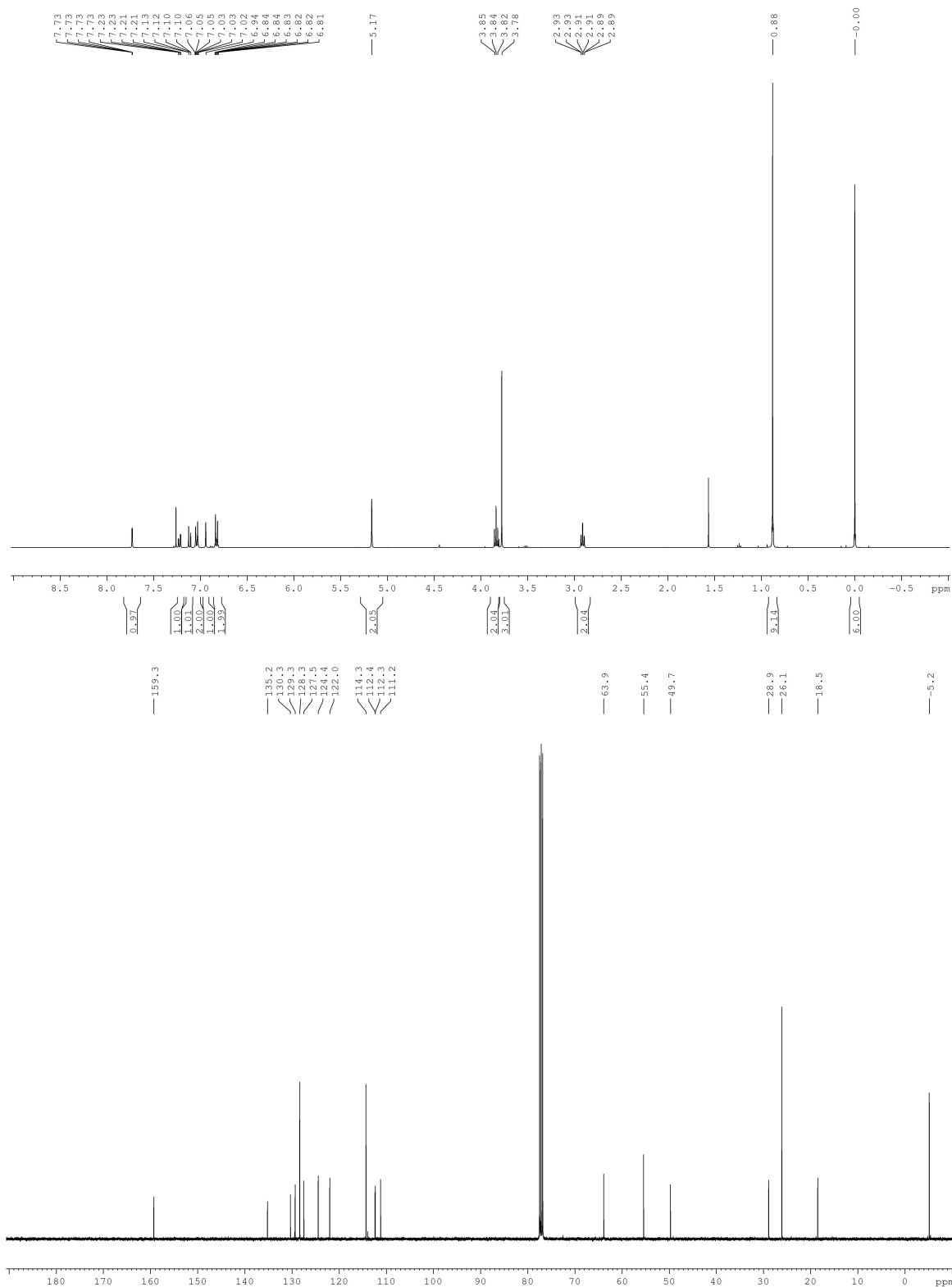*tert*-Butyl 5-bromo-3-(2-((*tert*-butyldimethylsilyl)oxy)ethyl)-1*H*-indole-1-carboxylate **58**

tert-Butyl 6-(5-bromo-3-(2-((*tert*-butoxycarbonyl)oxy)ethyl)-1*H*-indole-1-yl)hexanoate **60**

5-Bromo-3-(2-((*tert*-butyldimethylsilyl)oxy)ethyl)-1-(4-methoxybenzyl)-1*H*-indole **61**
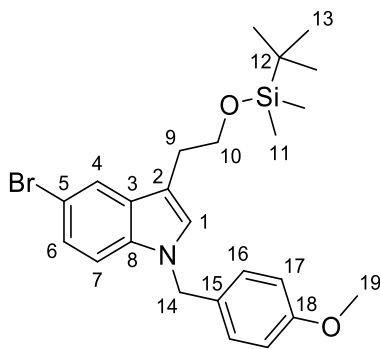
Spirocycles as Rigidified sp³-Rich Scaffolds for a Fragment Collection

Attila Sveiczer, Andrew J. P. North, Natalia Mateu,* Sarah L. Kidd, Hannah F. Sore, and David R. Spring*

This paper contains the Spring group's collaborative research into spirocyclic fragments described in section 2 of this thesis. In addition to the synthetic work described in this thesis, I wrote the entire manuscript and compiled the supplementary information.

# Spirocycles as Rigidified sp[3]-Rich Scaffolds for a Fragment Collection

Attila Sveiczer, Andrew J. P. North, Natalia Mateu,* Sarah L. Kidd, Hannah F. Sore, and David R. Spring*

Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, U.K.

S Supporting Information

ABSTRACT: Novel divergent methodology to access sp[3]-rich spirocyclic fragments is reported. First, a robust modular synthesis of bis-alkene amino ester building blocks was developed. Three different carbocycles and six heterocycles were then constructed to assemble eight spirocycles. Importantly, strategic exit vectors were incorporated within each scaffold to aid fragment growth and were elaborated via chemical modifications. Finally, computational methods demonstrate higher levels of rigidity, three-dimensionality, and structural diversity of the library compared to a commercial collection.
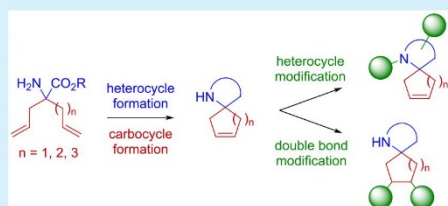
Since the rise of fragment-based drug discovery (FBDD) over two decades ago this strategy has proven particularly effective, producing numerous clinical candidates and two FDA-approved drugs, Vemurafenib[1] and Venetoclax[2] (Figure 1A). The success of this approach can be linked to two main benefits. First, due to the considerably fewer number of possible fragment-sized molecules, the chemical space coverage of a relatively small fragment library is exceedingly more efficient than that of a vast high-throughput screening (HTS) library. Second, fragment hits possess fewer but nonetheless high-quality binding interactions with the protein target, which can be later elaborated to afford highly potent lead compounds.[3−7]

Within this paradigm the generation of a suitable screening library is paramount. However, despite the undoubted success of FBDD, within recent years organic synthesis has been identified as a significant bottleneck within this process, owing to the overrepresentation of predominantly "flat" (hetero)-aromatic fragments lacking three-dimensionality as well as synthetically tractable exit vectors that could be utilized in rapid structure−activity relationship (SAR) studies.[4,5,8] While complexity of fragments remains under debate within the literature,[9,10] importantly, more three-dimensional (3D) fragments displaying exit vectors increase the potential for multidirectional fragment growth and the ability to identify leads for challenging targets such as protein−protein interactions.[8,11] Thus, recent efforts from within the synthetic community have focused on the development of novel strategies to access 3D fragments.[12−15]

Spirocyclic motifs remain an important bioactive substructure appearing within several natural products[16−18] and FDA-approved drugs[19,20] (Figure 1B). Importantly, as a direct result of their architecture, these small molecules often provide several advantages.[21] First, the spiranic center generates an inherently 3D structure that gives rise to higher levels of complexity, a feature which has been linked to improved clinical success.[10,22,23] Moreover, the conformationally restricted nature of spirocycles can reduce both the conformational entropy penalty of target binding and the number of possible conformations (distinct 3D shapes) that a molecule can adopt leading to higher potency and selectivity, respectively.[24,25] However, despite their utility, these motifs remain underrepresented in fragment screening collections. Indeed, very few compounds within the ChemBridge spirocycle library meet the size criteria of FBDD,[26,27] while only
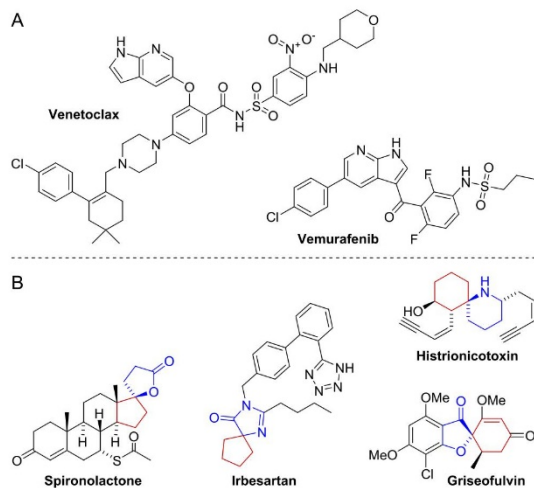
Figure 1. (A) Two FBDD-derived FDA-approved drugs: Venetoclax and Vemurafenib. (B) Examples of spirocyclic natural products (Griseofulvin and Histrionicotoxin) and FDA-approved drugs (Spironolactone and Irbesartan).

three spirocycles feature within the Maybridge core fragment collection. Thus, there is an urgent need for spirocyclic fragments and calls from within the field have encouraged the development of novel strategies to access such motifs.[25]

While strategies to access spirocycles within sp³-rich screening libraries have been reported, they either do not solely seek to construct spirocycles[28−30] or focus on more complex bis-[31] and bridged-[32] spirocycles. In addition, these compounds often lie outside the requirements for FBDD. Thus, we envisaged that a novel approach to access diverse spirocyclic fragments containing a polar heterocycle and a lipophilic carbocycle could give rise to a valuable library complementing already existing screening collections. To achieve this, our efforts were directed at utilizing α,α-disubstituted amino acid derivatives as building blocks, providing the potential to exploit the functional handles to generate fragment-like spirocyclic scaffolds. The incorporation of the two alkene handles enabled the carbocycle formation via ring-closing metathesis (RCM), forming an essential alkene exit vector, allowing us to alter the properties of this portion of the fragments. In addition, the amino and ester functionalities were installed to enable diverse heterocycle formations, increasing the possible polar interactions and the overall water solubility of the fragments. Accordingly, rapid access to varied scaffolds and the potential to exploit the exit vectors for fragment growth and merging to aid hit-to-lead development was envisioned.

First, the building blocks were prepared through the double alkylation of the glycine Schiff base **1** to form building blocks **3a−c** (Scheme 1). Despite similar procedures having been

**Scheme 1. Synthesis of the Building Blocks**
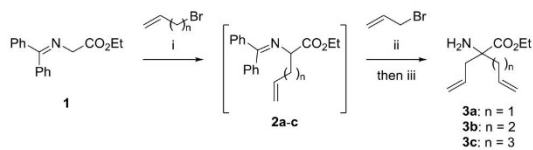
*Reaction conditions: (i) *t*BuOK, THF; (ii) *t*BuOK, THF; (iii) HCl, THF/H₂O; then Na₂CO₃. Overall yields **3a**: 67% (steps i and ii the same, steps i−iii in one pot); **3b**: 61% (steps i−iii in one pot); **3c**: 31% (steps i−iii done stepwise).

reported before,[33−36] herein we describe the straightforward racemic synthesis of α-quaternary amino esters with two different alkyl chains bearing terminal alkenes (**3b,c**). A brief optimization of the related literature procedure[33] allowed us to form **3c** in a simple stepwise process. Further optimization of the analogous route, however, resulted in the development of a one-pot procedure removing the chromatographic steps to access **3b** on large-scale in an improved 61% yield. This approach also enabled the formation of **3a**.

Subsequently, investigations into the formation of the different heterocycles were pursued using the racemic amino ester building block **3b**. First, the amine was Boc-protected to allow the cyclohexene formation in an RCM, followed by the reduction of the ester to the hydroxymethyl group by LiBH₄ to afford the key intermediate **4** in good yield. Intramolecular base-mediated pairing between the alcohol and Boc groups could then be achieved forming the oxazolidone moiety in **5**. The removal of the Boc protecting group under acidic conditions could be followed by pairing reactions incorporat-

ing various reagents. Reaction with cyanogen bromide formed amino oxazoline **6** whereas ethyl benzimidate hydrochloride gave the phenyl-substituted oxazoline **7**. The two morpholinones **8** and **9** were constructed by the chemoselective alkylation/acylation with chloroacetyl chloride and phenyl bromoacetate respectively (Scheme 2).

**Scheme 2. Synthesis of Different Core Heterocycles and Carbocycles**[a]
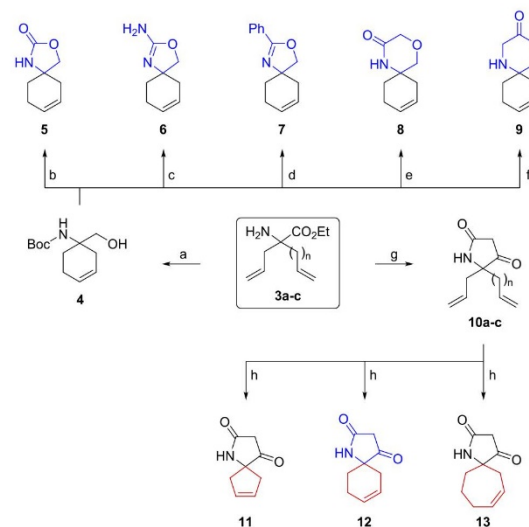
[a]Reaction conditions: (a) (i) Boc₂O, THF, 85%; (ii) Grubbs II, CH₂Cl₂, 81%; (iii) LiBH₄, THF, 89%; (b) *t*BuOK, THF, 90%; (c) HCl, dioxane; then BrCN, Et₃N, EtOH, 58%; (d) HCl, dioxane; then ethyl benzimidate hydrochloride, Et₃N, DCE, 57%; (e) (i) HCl, dioxane; then chloroacetyl chloride, Et₃N, CH₂Cl₂, 57%; (ii) *t*BuOK, *t*BuOH, 99%; (f) HCl, dioxane; then phenyl bromoacetate, iPr₂NEt, MeCN, 43%; (g) (i) ethyl malonyl chloride, Et₃N, CH₂Cl₂, 73−81%; (ii) *t*BuOK, THF; then aq. HCl, THF, 86−92%; (h) Grubbs II, CH₂Cl₂, 69−99%.

Building block **3b** could also be acylated with ethyl malonyl chloride, and then the base-mediated cyclization onto the ester group followed by the acid-catalyzed decarboxylative hydrolysis yielded the tetramic acid intermediate **10b**. Formation of the carbocycle ring in an RCM gave **12** in a good yield. To exemplify the potential to expand the cyclohexene ring, spirocycles also featuring the five- and seven-membered carbocycles were also synthesized from **3a** and **3c**, respectively (Scheme 2). Importantly, all the spirocycles were synthesized in no more than five steps from the building blocks.

Although racemic compounds were sought for our fragment library, the ability to produce optically pure isomers e.g. for SAR studies was also crucial. Thus, a second asymmetric route to the desired building blocks was devised utilizing the well-precedented stereoselective alkylation of iminolactones derived from phenylglycinol.[37,38] In this case, only one diastereomer of the aminolactone (*R*)-**15** was observed, which was successively deprotected to form the optically pure building block (*R*)-**3d**. As proof of concept, the single *R*-enantiomer of **12**, with the spiro[4,5] scaffold, was also synthesized (Scheme 3).

With eight fragment scaffolds in hand, it was next crucial to demonstrate the utility of the exit vectors installed within the

Scheme 3. Enantioselective Synthesis of (R)-12[a]
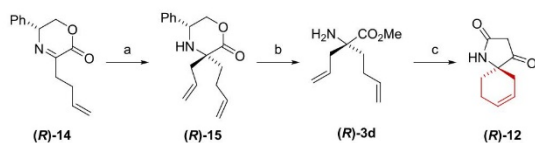
[a]Reaction conditions: (a) allyl bromide, Zn, DMF, 68%; (b) (i) SOCl$_2$, MeOH; (ii) Pb(OAc)$_4$, MeOH/CH$_2$Cl$_2$; then HCl, H$_2$O, 92% over 2 steps; (c) (i) ethyl malonyl chloride, Et$_3$N, CH$_2$Cl$_2$, 46%; (ii) tBuOK, THF; then aq. HCl, THF, 92%; (iii) Grubbs II, CH$_2$Cl$_2$, 97%.

molecules. Thus, N-alkylation of the oxazolidone (**16**), O-alkylation (**17**), and cross-coupling of the tetramic acid (**18**) were demonstrated. In addition, a modified heterocycle with an alcohol functionality (**19**) and two with new double bonds (**20** and **21**) were synthesized, incorporating new exit vectors to the resultant molecules (Scheme 4).

Scheme 4. Heterocycle Modification[a]
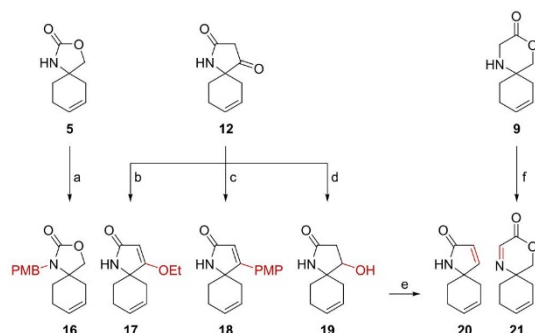
[a]Reaction conditions: (a) PMBCl, NaH, DMF, 96%; (b) KHMDS, EtBr, THF, 54%; (c) (i) Tf$_2$O, Et$_3$N, CH$_2$Cl$_2$, 59%; (ii) PMPB(OH)$_2$, Pd(PPh$_3$)$_4$, Na$_2$CO$_3$, H$_2$O/THF, 73%; (d) NaBH$_4$, MeOH, 23%; (e) TFAA, Et$_3$N, CH$_2$Cl$_2$; then KHCO$_3$, MeOH, 30%; (f) Pb(OAc)$_4$, MeCN, 92%.

Finally, to exhibit the versatility of the double bond as an exit vector, modifications such as Wacker oxidation (**22a,b**), dihydroxylation (**23a,b**), difluoro-cyclopropanation (**24a,b**), aziridination (**25a,b**), epoxidation (**26a,b**), and dibromination (**27a,b**) were explored (Scheme 5). Initial attempts of the epoxidation of **5**, however, proved challenging with respect to the isolation and purification of **26a,b**. Thus, a PMB group was installed (**16**) indeed proving to be compatible with several reaction conditions generating the diversified fragments in good to excellent yields. Removal of the PMB protecting group was also exemplified by the treatment of compound **27a** with CAN to generate the unprotected modified spirocycle **28**.

The physicochemical properties of our library consisting of 28 different nonprotected spirocyclic molecules were then calculated and compared to the commercially available Maybridge core fragment library using the widely accepted guidelines from within the field (Table 1).[26,27] This revealed the spirocyclic library adheres well to the guidelines and was additionally predicted to be significantly less lipophilic (SlogP of 0.9 versus 1.8) and more rigid (rotatable bond count of 0.2 versus 2.0) than the Maybridge core fragment collection.

Scheme 5. Double Bond Modification and PMB Deprotection[a]
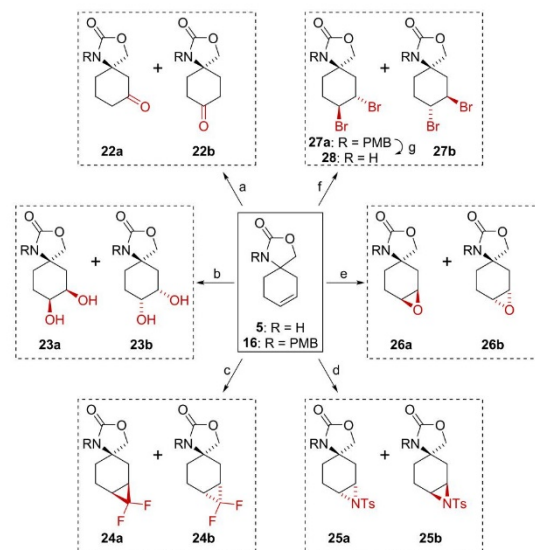
[a]R = PMB unless specified otherwise. All products are racemic. Reaction conditions (combined yields with ratios **a**/**b** are given): (a) Fe(acac)$_2$, tBuOH, air, 47%, 1.9:1; (b) OsO$_4$, NMO, citric acid, H$_2$O/THF, 99%, 2.5:1; (c) TMSCF$_3$, NaI, THF, 62%, 20:1; (d) TsNClNa·3H$_2$O, PhNMe$_3$Br$_3$, 4 Å MS, 70%, 1.4:1; (e) R = H, mCPBA, NaHCO$_3$, CH$_2$Cl$_2$, 52%, 8:1; (f) PhNMe$_3$Br$_3$, CH$_2$Cl$_2$, 96%, 18:1; (g) CAN, MeCN/H$_2$O, 96% (R = H in product).
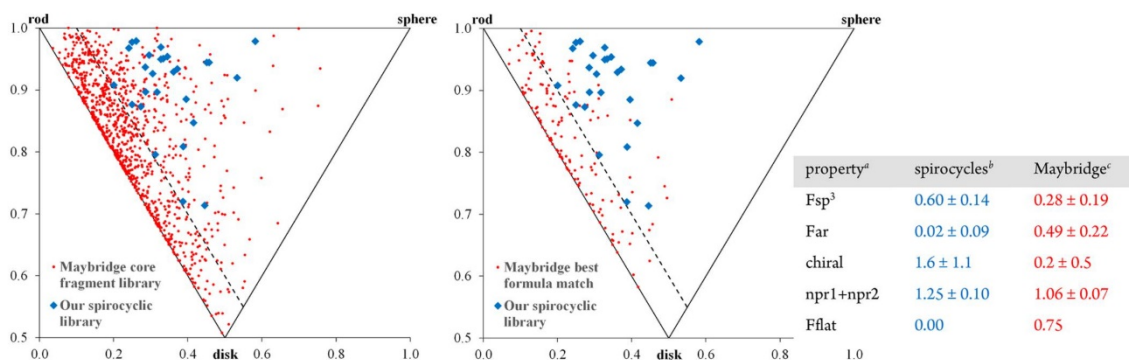
Table 1. Physicochemical Properties of Fragment Libraries Compared to the Ideal Range

property[a]	spirocycles[b]	Maybridge[c]	ideal range[d]
MW	186 ± 41	182 ± 42	140–230
HBD	1.3 ± 0.7	1.0 ± 0.8	≤3
HBA	1.8 ± 0.6	1.9 ± 0.7	≤3
SlogP	0.9 ± 0.9	1.8 ± 0.8	0–2
RBC	0.2 ± 0.5	2.0 ± 1.5	≤3
TPSA	48 ± 13	39 ± 14	≤60

[a]MW = molecular weight (Da), HBD = number of hydrogen-bond donors, HBA = number of hydrogen-bond acceptors, SlogP = partition coefficient, RBC = rotatable bond count, TPSA = topological polar surface area (Å^2). [b]Protecting groups virtually removed from our library. [c]Maybridge core fragment collection of 1000 fragments. [d]Guidelines set by Astex Pharmaceuticals.[26,27] See Supporting Information.

Strikingly, our library also displayed far superior sp^3-content with an average fraction of sp^3 atoms (Fsp^3) of 0.52, which also translates into a much lower fraction of aromatic atoms (Faro) of 0.16. Moreover, the average number of chiral centers (1.6) is also considerably higher, resulting in greater stereochemical diversity achieved by the spirocyclic fragments.

In order to qualitatively assess the shape diversity in our library, a principal moments of inertia (PMI) analysis was carried out. Our library was then compared to the whole Maybridge core fragment collection of 1000 fragments and a

221

property[a]	spirocycles[b]	Maybridge[c]
Fsp^3	0.60 ± 0.14	0.28 ± 0.19
Far	0.02 ± 0.09	0.49 ± 0.22
chiral	1.6 ± 1.1	0.2 ± 0.5
npr1+npr2	1.25 ± 0.10	1.06 ± 0.07
Fflat	0.00	0.75

Figure 2. PMI plots for the visual representation of shape diversity. Each corner of the plot represents a unique shape (rod-, disk-, and sphere-like features). The dashed line represents the boundary of "flat land".[11] Our virtual deprotected library of 28 spirocycles (blue) is compared to 1000 fragments (red, left) and the 147 best-matched fragments based on heavy and heteroatom count (red, right) from the Maybridge collection. The table on the right summarizes the physicochemical properties used to describe the 3D properties of the two libraries. ^{a}Fsp^3 = fraction of sp^3 atoms, Far = fraction of aromatic atoms, chiral = number of chiral centers, npr = normalized PMI ratio, Fflat = fraction of molecules lying below the "flat land line"[11]. ^{b}Virtual deprotected library. ^{c}147 best-matched fragments from the Maybridge collection. See Supporting Information.

representative subset consisting of the 147 best-matched compounds based on heavy atoms (Figure 2). Both plots show that the Maybridge fragments tend to aggregate to the left-hand edge (rod- and disc-like features) or the "flat land", whereas our spirocyclic fragments are more evenly distributed. Analysis showed that more than 70% of the whole Maybridge collection and 75% of the best formula match subset fall within "flat land" (defined as npr1 + npr2 ≤ 1.1).[11] On the other hand, no spirocyclic fragment in our library was found below the "flat land" criteria, suggesting more 3D molecules.

In conclusion, we have developed a robust, scalable, and modular route to racemic α,α-disubstituted amino ester building blocks including an adapted stereoselective alkylation protocol to access the optically pure intermediate. These were utilized in the efficient construction of eight novel spirocyclic scaffolds comprising pharmacophore heterocycles and variable carbocycles. All core scaffolds display an array of 3D exit vectors demonstrated by a number of chemical modifications to both the hetero- and carbocycles. Together with the enantioselective synthesis, rapid SAR studies and binding pocket exploration by fragment growth could therefore be envisioned. Finally, the computational predictions revealed optimal physicochemical properties, higher rigidity increased 3D properties, shape, and stereochemical diversity compared to a commercial fragment library.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.orglett.9b01499.

General Remarks, Procedures and Analytical Data, Computational Analysis, Crystallographic Data, and NMR Spectra (PDF)

Accession Codes

CCDC 1912266−1912268 and 1912283−1912289 contain the supplementary crystallographic data for this paper. These data can be obtained free of charge via www.ccdc.cam.ac.uk/data_request/cif, or by emailing data_request@ccdc.cam.ac.uk, or by contacting The Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, UK; fax: +44 1223 336033.

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: spring@ch.cam.ac.uk.
*E-mail: nm462@ch.cam.ac.uk.

ORCID Ⓘ

David R. Spring: 0000-0001-7355-2824

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Bollag, G.; Hirth, P.; Tsai, J.; Zhang, J.; Ibrahim, P. N.; Cho, H.; Spevak, W.; Zhang, C.; Zhang, Y.; Habets, G.; et al. *Nature* **2010**, 467, 596−599.
(2) Souers, A. J.; Leverson, J. D.; Boghaert, E. R.; Ackler, S. L.; Catron, N. D.; Chen, J.; Dayton, B. D.; Ding, H.; Sampath, D.; Lee, J.; et al. *Nat. Med.* **2013**, 19, 202−208.
(3) Scott, D. E.; Coyne, A. G.; Hudson, S. A.; Abell, C. *Biochemistry* **2012**, 51, 4990−5003.
(4) Boyd, S. M.; Turnbull, A. P.; Walse, B. *Comput. Mol. Sci.* **2012**, 2, 868−885.
(5) Keserü, G. M.; Erlanson, D. A.; Ferenczy, G. G.; Hann, M. M.; Murray, C. W.; Pickett, S. D. *J. Med. Chem.* **2016**, 59, 8189−8206.
(6) Murray, C. W.; Rees, D. C. *Nat. Chem.* **2009**, 1, 187−192.

(7) Hajduk, P. J.; Greer, J. *Nat. Rev. Drug Discovery* **2007**, *6*, 211−219.

(8) Murray, C. W.; Rees, D. C. *Angew. Chem., Int. Ed.* **2016**, *55*, 488−492.

(9) Hann, M. M.; Leach, A. R.; Harper, G. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 856−864.

(10) Lovering, F. *MedChemComm* **2013**, *4*, 515−519.

(11) Morley, A. D.; Pugliese, A.; Birchall, K.; Bower, J.; Brennan, P.; Brown, N.; Chapman, T.; Drysdale, M.; Gilbert, I. H.; Hoelder, S.; et al. *Drug Discovery Today* **2013**, *18*, 1221−1227.

(12) Twigg, D. G.; Kondo, N.; Mitchell, S. L.; Galloway, W. R. J. D.; Sore, H. F.; Madin, A.; Spring, D. R. *Angew. Chem., Int. Ed.* **2016**, *55*, 12479−12483.

(13) Mateu, N.; Kidd, S. L.; Kalash, L.; Sore, H. F.; Madin, A.; Bender, A.; Spring, D. R. *Chem. - Eur. J.* **2018**, *24*, 13681−13687.

(14) Foley, D. J.; Doveston, R. G.; Churcher, I.; Nelson, A.; Marsden, S. P. *Chem. Commun.* **2015**, *51*, 11174−11177.

(15) Hassan, H.; Marsden, S. P.; Nelson, A. *Bioorg. Med. Chem.* **2018**, *26*, 3030−3033.

(16) Müller, G.; Berkenbosch, T.; Benningshof, J. C. J.; Stumpfe, D.; Bajorath, J. *Chem. - Eur. J.* **2017**, *23*, 703−710.

(17) Oxford, A. E.; Raistrick, H.; Simonart, P. *Biochem. J.* **1939**, *33*, 240−248.

(18) Kim, S.; Ko, H.; Lee, T.; Kim, D. *J. Org. Chem.* **2005**, *70*, 5756−5759.

(19) Taylor, F. F.; Faloon, W. W. *J. Clin. Endocrinol. Metab.* **1959**, *19*, 1683−1687.

(20) Brunner, H. R. *Am. J. Hypertens.* **1997**, *10*, 311S−317S.

(21) Carreira, E. M.; Fessard, T. C. *Chem. Rev.* **2014**, *114*, 8257−8322.

(22) Lovering, F.; Bikker, J.; Humblet, C. *J. Med. Chem.* **2009**, *52*, 6752−6756.

(23) Zheng, Y.; Tice, C. M.; Singh, S. B. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 3673−3682.

(24) Wermuth, C. G.; Villoutreix, B.; Grisoni, S.; Olivier, A.; Rocher, J.-P. In *The Practice of Medicinal Chemistry*; Wermuth, C., Aldous, D., Raboisson, P., Rognan, D., Eds.; Academic Press: London, 2015; pp 73−96.

(25) Zheng, Y.-J.; Tice, C. M. *Expert Opin. Drug Discovery* **2016**, *11*, 831−834.

(26) Congreve, M.; Carr, R.; Murray, C.; Jhoti, H. *Drug Discovery Today* **2003**, *8*, 876−877.

(27) Palmer, N.; Peakman, T. M.; Norton, D.; Rees, D. C. *Org. Biomol. Chem.* **2016**, *14*, 1599−1610.

(28) Hung, A. W.; Ramek, A.; Wang, Y.; Kaya, T.; Wilson, J. A.; Clemons, P. A.; Young, D. W. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 6799−6804.

(29) Mayol-Llinàs, J.; Farnaby, W.; Nelson, A. *Chem. Commun.* **2017**, *53*, 12345−12348.

(30) Haftchenary, S.; Nelson, S. D.; Furst, L.; Dandapani, S.; Ferrara, S. J.; Bošković, Ž. V.; Lazú, S. F.; Guerrero, A. M.; Serrano, J. C.; Crews, D. K.; et al. *ACS Comb. Sci.* **2016**, *18*, 569−574.

(31) Stotani, S.; Lorenz, C.; Winkler, M.; Medda, F.; Picazo, E.; Martinez, R. O.; Karawajczyk, A.; Sanchez-Quesada, J.; Giordanetto, F. *ACS Comb. Sci.* **2016**, *18*, 330−336.

(32) Foley, D. J.; Craven, P. G. E.; Collins, P. M.; Doveston, R. G.; Aimon, A.; Talon, R.; Churcher, I.; von Delft, F.; Marsden, S. P.; Nelson, A. *Chem. - Eur. J.* **2017**, *23*, 15227−15232.

(33) López, A.; Moreno-Mañas, M.; Pleixats, R.; Roglans, A.; Ezquerra, J.; Pedregal, C. *Tetrahedron* **1996**, *52*, 8365−8386.

(34) Hammer, K.; Undheim, K. *Tetrahedron* **1997**, *53*, 2309−2322.

(35) Kotha, S.; Sreenivasachary, N. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 257−260.

(36) Khan, I. U.; Kattela, S.; Hassan, A.; Correia, C. R. D. *Org. Biomol. Chem.* **2016**, *14*, 9476−9480.

(37) Harwood, L. M.; Vines, K. J.; Drew, M. G. B. *Synlett* **1996**, *1996*, 1051−1053.

(38) Fustero, S.; Mateu, N.; Albert, L.; Aceña, J. L. *J. Org. Chem.* **2009**, *74*, 4429−4433.

223