Understanding the foliar uptake of amino acids

BETHANY CAROLINE RUTH HENDERSON

Understanding the Foliar Uptake of Amino Acids

by Bethany C. R. Henderson

University of Durham

Department of Chemistry

Van Mildert College

A thesis presented in fulfilment of the requirements

for the degree of Doctor of Philosophy

March 2025

Abstract

Research to enhance crop yields in an environmentally sustainable fashion is of great interest to the agricultural sector to meet the increasing demand in the face of pressure to reduce pesticide use.
Biostimulants, such as naturally occurring amino acids, are environmentally benign plant treatments that potentially enable improved yield, growth, and resistance to external environmental factors. Amino acids can reduce the impact of abiotic stress, stimulate plant defence mechanisms, and improve rate of photosynthesis. The foliar application of amino acids is known to increase the overall yield and quality of some plant species. However, there are several gaps in our understanding of their mode of action that prevent us from realising their full potential.
This project investigates the movement of amino acids in soybeans using isotopically and fluorescently labelled amino acids. By tracing amino acid movement using sensitive detection methods, such as mass spectrometry, a clearer understanding of the timescale of uptake of the amino acids, their movement around the plant, and their eventual break down and metabolism, has been gained.
In summary, this thesis aims to reduce the dependence on agrochemicals, aligning with the key characteristics of sustainable chemistry, by understanding the mechanisms by which amino acids improve crop yields and targeting them more effectively. Results show rapid uptake of the labelled tryptophan treatment within the four hours of application and movement around the plant to all plant parts with metabolism starting around 48 hours after application.

Table of Contents

List of Tables

List of Schemes

List of Figures

Abbreviations

AAT	alcohol acyltransferase
ABA	abscisic acid
ACC	aminocyclopropane-1-carboxylic acid
ACO	ACC oxidase
ACS	ACC synthase
ADH	alcohol dehydrogenase
AI	artificial intelligence
Ala	alanine
ANOVA	analysis of variance
ASAP	atmospheric solids analysis probe
ASE	accelerated solvent extraction
Asn	asparagine
Asp	aspartic acid
AUC	area under the curve
BAP	6-benzylaminopurine
BINAP	(2,2'-bis(diphenylphosphino)-1,1'-binaphthyl)
BP	benzyl proline
BPB	(S)-2-[N-(N'-benzylprolyl)amino]benzophenone
br (in IR or NMR)	broad
CID	collision-induced dissociation
COSY	homonuclear correlation spectroscopy
CSD	Cambridge structural database
CSV	comma separated values
CV	coefficient of variation
Cys	cysteine
d (in NMR)	doublet
DCE	dichloroethane
DCM	dichloromethane
DEA	diethanolamine
DEGs	differentially expressed genes
DEPT-135	distortionless enhancement by polarization transfer
DMA	dimethylamine
DMSO	dimethyl sulfoxide
E2	bimolecular elimination
EAAE	enzyme assisted aqueous extraction
EACP	enzyme assisted cold pressing
EAE	enzyme-assisted extraction
EI	electron ionisation
EIC	extracted ion chromatogram
ESI	electron spray ionisation
FAO	the food and agricultural organisation of the United Nations
FT-IR	fourier transform infrared spectroscopy
GA	gibberellic acid
GABA	gamma-aminobutyric acid
GAP	global agricultural productivity report
GC-MS	gas chromatography - mass spectrometry

Gln	glutamine
Glu	glutamic acid
Gly	glycine
h	hours
HMBC	heteronuclear multiple bond correlation
HPLC	high-performance liquid chromatography
HSQC	heteronuclear single quantum coherence spectroscopy
Hz	hertz
IAA	indole-3-acetic acid
IAM	indole-3-acetamide
IBA	indole-3-butyric acid
Ile	isoleucine
IPT	isopentenyl transferase
IPyA	indole-3-pyruvic acid
IR	infrared spectroscopy
J	scalar coupling constant
JA	jasmonic acid
K	kelvin
KYNA	kynurenic acid
L&MICs	lower- and middle-income countries
$L.ha^{-1}$	litres per hectare
LC-MS	liquid chromatography - mass spectrometry
LED	light-emitting diode
Leu	leucine
LOD	limit of detection
M	molar
m (in IR)	medium
M (in MS)	molecular ion
m (in NMR)	multiplet
m/z	mass to charge ratio
MAE	microwave-assisted extraction
MALDI-TOF	matrix-assisted laser desorption/ionization
MAPK	mitogen-activated protein kinase
MAT	methionine adenosyltransferase
Met	methionine
min	minutes
mp	melting point
MRM	multiple reaction monitoring
MS	mass spectrometry
MS (plants)	apical meristem
MS/MS	tandem mass spectrometry
MTA	methylthioadenosine
NAA	naphthalene acetic acid
NBS	N-bromosuccinimide
NFSI	N-fluorobenzenesulfonimide
ng/mL	nanograms per millilitre
NIR-FT-Raman	near infrared fourier transform Raman
NMR	nuclear magnetic resonance
NOSEY	nuclear Overhauser effect spectroscopy

PAR	photon flux density
Phe	phenyl alanine
PLE	pressurised-liquid extraction
PMA	phosphomolybdic acid
Pro	proline
PSII	photosystem ii
Q	quadrupole
q (in NMR)	quartet
RH	relative humidity
ROS	reactive oxygen species
RT	retention time
s (in IR)	strong
s (in NMR)	singlet
SAM	S-adenosylmethionine
SEM	standard error of the mean
Ser	serine
SFE	supercritical fluid extraction
sh (in IR)	sharp
S_N2	substitution nucleophilic bimolecular
SNR	signal to noise ratio
SPAD	soil plant analysis development
SPS	solvent purification system
SQD	single quadrupole detector
SRM	selected reaction monitoring
SS	sum of squares
T	time
t (in NMR)	triplet
TAA	tryptophan aminotransferases
TBADT	tetrabutylammonium decatungstate
TFA	trifluoroacetic acid
TFP	total factor productivity
THF	tetrahydrofuran
Thr	threonine
TIC	total ion chromatogram
TLC	thin layer chromatography
TPP	thermal proteome profiling
TQD	tandem quadrupole detector
Trp	tryptophan
Tyr	tyrosine
UAE	ultrasonic assisted extraction
UATR (FTIR)	universal attenuated total reflectance
UV	ultraviolet light
v/v	volume to volume
Val	valine
w (in IR)	weak
XRD	X-ray diffraction
YUC	YUCCA flavin monooxygenases

Declaration

I declare that;
This submission is entirely my own work and is based on research carried out within the Sanderson group at Durham University. I confirm that I have read and understood the Department and University regulations on plagiarism. No part of this thesis has been submitted elsewhere for any other degree or qualification. Each contribution to, and quotation in this submission which is taken from the work or works of other people, has been cited correctly.

BETHANY C.R. HENDERSON
March 2025

Copyright

Acknowledgements

This project was funded by the Biotechnology and Biological Sciences Research Council through Croda International Ltd.

A huge thank you goes to my supervisor Dr John Sanderson for his support mentorship, expertise and patience throughout my PhD. Thanks to all in the Sanderson group past and present for their teaching, guidance and friendship. My particular thanks to Dr Gen Duché and Yesh Tanna, for all their friendship, support and shared long hours in the lab.

I am also incredibly grateful to all at the Plant Impact site of Croda Europe Ltd. for their support in this project, particularly my industry supervisors Dr Andy Fowles and Dr Ian Wood, and all the plant scientists who helped me immensely throughout the project: Jahmel Neblett, Laura Bishop, Robert Kempster and Dr Steven Adams.

To the Kitching Group: Dr Matt Kitching, Dr James Barclay, Dr Jinyi Xuan, Alex Lee, Dr Guilia Grelli and Dr Callum Begg, thank you for your kindness, friendship, help and innumerable pub trips. These past years would have been far duller without you all.

I would also like to thank those in the analytical services at Durham University, particularly Mr Peter Stokes in Mass Spectrometry; Dr Juan Aguilar Malavia and Dr Eric Hughes in the NMR service; Dr Aileen Congreve in Chromatography; and Dr Toby Blundell and Dr Natalie Pridmore in the X-ray Crystallography service. Their expertise and support have been invaluable throughout many aspects of this project.

I have been incredibly lucky to make so many wonderful friends during my time in Durham. I would also like to express my gratitude to Van Mildert College, in particular the Middle Common Room, for providing a welcoming and supportive community throughout my PhD. And to the members of the Inclusive Christian Movement Durham (ICMD) for your friendship, support and prayers over the years, they have meant a lot.

Thanks to my amazingly supportive family; your advice, encouragement, and generosity have been invaluable in getting me to the end of this long journey. I am so very grateful.
Finally, my greatest thanks go to my partner, Johanna, for her love, patience and support over the past few years. Thank you for putting up with my stress, panic and endless ramblings about soybeans, I couldn't have done this without you.

I think it is also important to acknowledge that this thesis looks hugely different to how I would have imagined it when I first started in October 2019. The Covid-19 pandemic had a huge impact globally, but also personally, and I have faced challenges I never could have anticipated. Becoming disabled by long Covid during the course of my studies has led to a significant change in my capacity to work, study, and live as I once did. Finishing this PhD under those circumstances has been incredibly difficult, and the completion of this thesis reflects not only my own persistence, resilience and hard work, but also the patience, care, and unwavering support of those around me. I'm deeply grateful to everyone who helped me keep going when things felt impossibly hard.

Chapter 1: Introduction

1.1 Need for improved yields

Plant yield efficacy is an increasingly important issue, especially in the global agricultural industries. Maximising yield size in affordable, easily reproducible, and reliable ways offers a necessary response to the increasing demand for food and other agricultural resources. This project will investigate the use of individual amino acids as foliar biostimulants as one intervention to this long-recognised problem. As the global population continues to increase in size, and so the demand for food also increases, the need to improve plant yield becomes increasingly necessary to combat world hunger. For example, the United Nation's Food and Agricultural Organisation (FAO) estimated that global agricultural output must increase by 60% by the 2050 to combat the deficit between production and need.[1] Their analysis was based on baseline production levels in 2012, while Hunter *et al* indicated in 2017 that this value could be anywhere between 25 - 70%.[2]

To address the global agricultural productivity gap, The Global Harvest Initiative was established in 2009 as a partnership between four major U.S.-based agrochemical companies. Their annual report tracks current agricultural outputs and projects growth based on recent trends, comparing these figures against an ambitious target: doubling global agricultural productivity growth from 2010 to 2050.[3] In 2023, they calculated that the level of agricultural growth needed to increase by 1.91% per year between now and 2050 to meet these targets (Fig. 1).[4] This figure was a revision of their earlier calculation of a 1.73% increase required in 2010, after they identified only an average growth rate of 1.14% during the first 10 years of the Initiative's work (Fig. 1). The revised Global Harvest Initiative targets, and their recognition of a failure to deliver on those targets indicates the need for further significant changes to how agriculture is done globally are required to meet these targets.

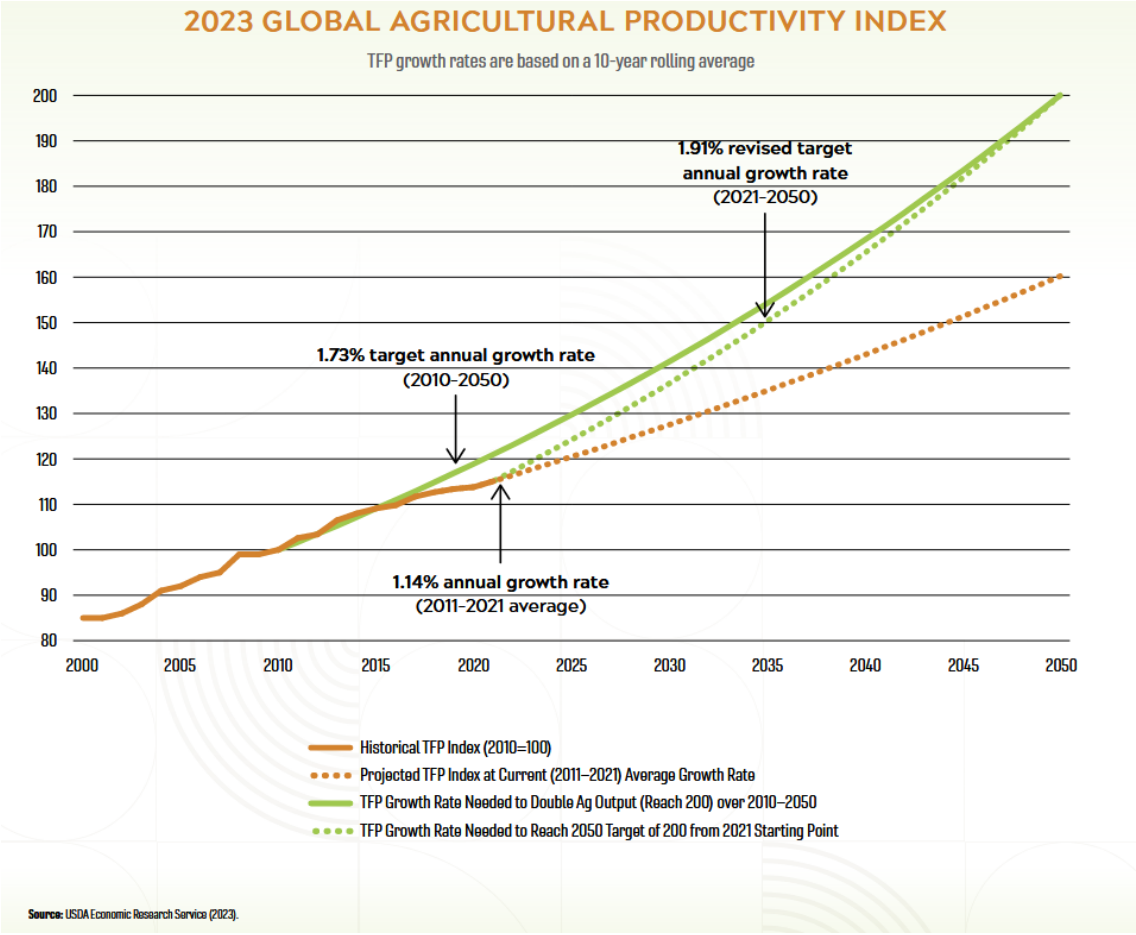


Figure 1: *Observed and projected Total Factor Productivity (TFP) growth from the 2023 Global Productivity Report (GAP), with revised annual growth targets (1.91%) exceeding historical benchmarks (1.14%). Reproduced without modification from J. Agnew and S. Hendery, GAP Report 2023, Virginia Tech College of Agriculture and Life Sciences and Virginia Tech Foundation.*[4] Plant Impact Ltd collated open-sourced data which has shown a significant difference between the average and the highest possible crop yield that could be obtained in favourable conditions for three commercially grown crops: soybeans, wheat, and cocoa (Fig. 2). The data for these agriculturally significant crops demonstrates that there is much room for improvement of yield in commercial growing contexts. If this gap between potential yield and reality could be lessened, it would increase crop supply and so more closely match the amount of food needed to support the population. Plant research and product design is ideally placed to enable yield size to be maximised in line with potential yield projections.

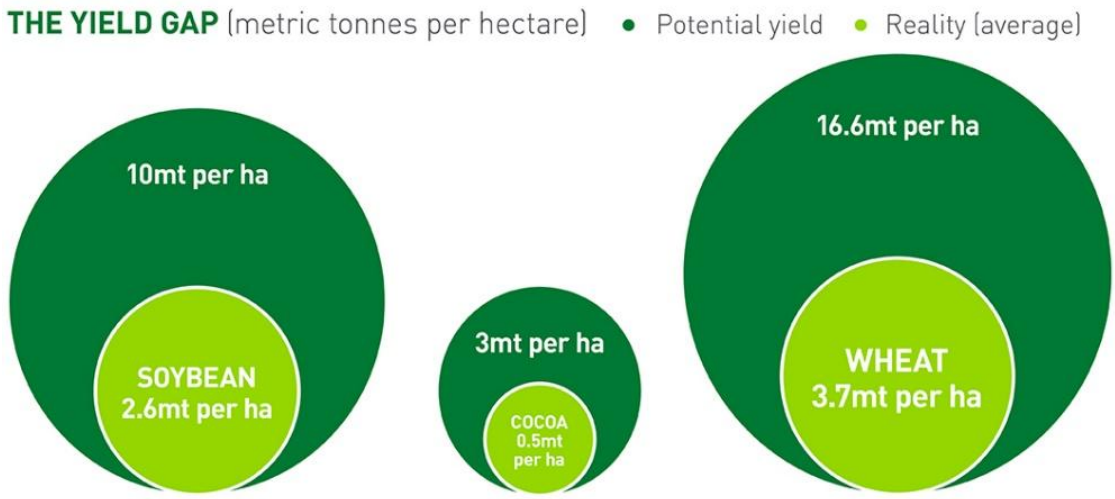


Figure 2: *Yield gap representation for soybeans, cocoa, and wheat. For each crop, the average yield per hectare is shown as a central circle, surrounded by a larger circle indicating the potential yield based on record high values. This visual comparison highlights the disparity between current production levels and attainable benchmarks. Data sourced from FAO statistics; figure reproduced with permission from Plant Impact Ltd.*[5]

At the same time as the global population is increasing, climate change is increasing environmental stresses experienced by crops.[6] For example, the FAO estimated in 2017 that 45 million ha of irrigated land was affected by excessive salt.[7] Intense and unpredictable rainfall patterns across Asia in recent years have contributed to reduced crop yields, with some areas experiencing flooding and others, particularly arid regions, facing more frequent and severe droughts.[8]

To mitigate these effects, the use of biostimulants for the promotion of useful traits in agricultural crops, such as improving yield and resistance to adverse environmental factors, has increased rapidly over the past decade.[9] However, earlier interventions to increase yield size also highlight the need to be intentional and aware of the impact on people and the environment.

1.1.1 Environmental considerations: The legacy of The Green Revolution

One notable development in agricultural techniques came through *The Green Revolution* which occurred between the 1940s and 1960s. Led by American scientist Norman Borlaug, it was heralded as an end to crop poverty, particularly in low and

middle income countries (L&MICs). The Green Revolution marketed high yielding "miracle seeds" including semi-dwarf varieties of wheat and rice to tackle poor crop growth in L&MICs.[10] By shortening the height of the plants, chemicals from fertilisers were not converted into nutrients to increase the height of the plant but were used by the plant primarily for the crop growth. It was the increased use of chemical fertilisers in combination with these seeds that led to such an increase in yield output.[11] However the semi-dwarf varieties of crops were reliant on intense application of nitrate fertilisers to have the intended significant effect on yield. Despite this reliance, the Green Revolution was widely celebrated for its impact on global food poverty, with Borlaug receiving the Nobel Peace Prize in 1970.[12]

Problems soon emerged as there were several negative impacts to the methods used to increase plant yield. Due to the intensive use of these "miracle seeds", genetic diversity in L&MICS - notably Mexico and India - has been significantly reduced.[13] The reliance on intensive chemical pesticides, as well as the nitrate-based fertilisers, also resulted in a loss of biodiversity in these areas.[14] The long-term effects of these interventions continues to be felt, most notably through the toxic effects of the fertilisers on people and animals. The excessive use of nitrate-based fertilisers to promote the growth of the "miracle seeds" led to high nitrate levels in soils.[15] Nitrate leaching from the soil into water led to eutrophication of nearby water sources, causing the consumption of nitrates by people and animals in the area.[14] In the body, nitrates form nitrous oxide which is linked to many health problems including: numerous cancers, infantile methemoglobinemia and Alzheimer's disease.[16] The production process of such fertilisers also creates nitrous dioxide gas, a contributor to the "greenhouse effect".[17] These problems associated with nitrate use have seen a desire to move away from nitrate products towards more environmentally friendly techniques, including use of naturally occurring bio-stimulants such as amino acids. By focussing on using naturally occurring treatments in small quantities, the risk of environmental leaching and health related side effects is reduced.

The ongoing impact of the Green Revolution shows the necessity of considering the environmental impact of compounds used to improve agricultural yields. Focusing

research into the impact and development of naturally occurring biostimulants offers an environmentally benign option for improving plant growth with lower risks of side effects to the wider environment.

1.2 Biostimulants: An introduction

Biostimulants are natural compounds or microorganisms externally applied to crops, enhancing nutritional efficiency, stress resilience, or overall plant quality.[18] These compounds are usually derived from natural sources, are used in small amounts, and can be categorised into several groups. These include seaweed and plant extracts (containing nutrients and plant hormones to stimulate metabolism); humic and fulvic acids (found in soil); inorganic compounds (such as phosphites and silicon); chitosan and other biopolymers (derived from plant/animal sources); and amino acids (found naturally in plants for protein synthesis).

Research into biostimulants has dramatically increased in recent years, with hundreds of papers on the topic being published yearly since 2018.[19] As there are many different types of biostimulant in use commercially, it is necessary to focus on the most prevalent ones. This literature review follows the categories for biostimulants outlined above.

1.2.1 Use of seaweed as biostimulant

A main area of focus in biostimulant research and commercial use are seaweeds and products derived from seaweed to promote phytostimulatory, phytoelicitory and phytohormonal responses.[20] Seaweeds are biodegradable and non-toxic and so make good candidates for use as biostimulants. Extracts of seaweeds have been used for foliar and soil application in various forms.[21] When separated into the individual components that make up seaweed extracts, the growth effects are reduced. This shows that the effects observed are related to multiple components working together.[20]

1.2.2 Use of humic and fulvic acids as biostimulants

Humic and fulvic acids are not individual chemical compounds but heterogeneous mixtures of large organic molecules formed by the decomposition of plant and animal material. They are distinguished operationally by solubility: humic acids are soluble in alkali but precipitate when acidified, whereas fulvic acids remain soluble across the pH range.[22] Both are classed as humic substances, and the chemical components of these mixtures are highly varied. The majority of compounds are aromatic or aliphatic in nature, bearing functional groups such as carboxyl, phenolic, hydroxyl, carbonyl, and amino substituents.[23, 24] Humin is the third fraction of humic substances, defined by its insolubility in both alkali and acid, and is therefore not used as a biostimulant.[25]

Exogenous application of humic substances to soil has been shown to increase shoot and root dry weights over different species, including maize, wheat and tomato.[26] The response changes depending on the nature of the plant species, with dicotyledonous plants experiencing greater effects than monocotyledonous species, and the source of the humic substances. For example, extracts from brown coal are less effective than from peat.[27]

The effect seen by the exogenous application of humic substances comes from interactions with plant membrane transporters and plant hormones rather than from a direct source of nutrients.[26] Humic substances have been linked to increased expression of genes relating to auxin and cytokinin biosynthesis pathways.[28] They are also understood to bind with Fe(III), forming stable humic-Fe complexes.[29] These complexes then also bind to phosphorus, increasing the availability of iron and phosphorus in the soil and so increasing the uptake of these elements by higher plants.

1.2.3 Commercial use of amino acids as biostimulants

There are 23 proteinogenic amino acids, of which 20 are encoded by the nuclear genome in eukaryotes.[30] With the exception of glycine, which is achiral, all proteinogenic amino acids exhibit the L-configuration.[31] These L-α-amino acids share

a general structure consisting of a carboxylic acid group and a primary amine, both bound to a central (α) carbon atom, which also carries a variable side chain.

Many commercially available plant treatments contain mixtures of amino acids. Amino16®, for example, contains an 11.3% mixture of amino acids including all the aliphatic, basic, acidic and hydroxylic amino acids alongside methionine, phenylalanine, and tyrosine. Foliar application of Amino16® has been shown to improve yield in several plant species, including lettuce and olive trees (*Olea europaea var. minor rotunda),* and is already used within industrial agricultural practices.[32, 33]

Studies into the effect of amino acid mixtures have shown an increase in chlorophyll content and antioxidant capacity in golden pothos (*Epipremnum aureum*) and basil (*Ocimum basilicum*) suggesting the improvements in plant health and yield are related to photosynthetic efficiency and improved stress tolerance.[34, 35] While most research into amino acid mixtures as biostimulants has focused on measuring changes to yield and biochemical composition, few have explored the mechanisms behind these effects. Li *et al.* (2024) reported that after the foliar application of an unspecified mixture of water-soluble amino acids, almost 4000 differently expressed genes (DEGs) were up regulated, while 3000 genes were down regulated. These genes were primarily involved in photosynthesis; photosynthesis-antenna proteins; plant-pathogen interactions; phenylpropanoid biosynthesis; glyoxylate and dicarboxylate metabolism; Mitogen-Activated Protein Kinase (MAPK) signalling pathway; and plant hormone signal transduction. However, due to the presence of multiple amino acids in these supplements, it is not possible to determine if all the components were necessary to reach the desired effect and, if so, how each interacted within the plant to cause these effects.

The lack of detailed understanding of how these commercial products work is a significant research gap. Improving this understanding could reduce costs for manufacturers of such products by identifying and removing less effective compounds from mixtures. It could also lower the environmental impact, as reducing synthesis needs would decrease the resources required to produce these products.

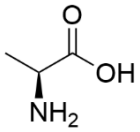
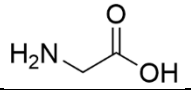
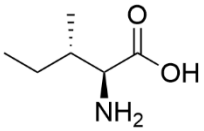
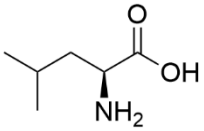
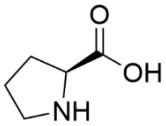
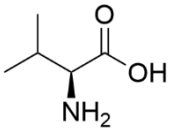
Better understanding of the differences between individual amino acids could also enable targeted treatment for specific crops or stress conditions, where some amino acids may be more effective than others.

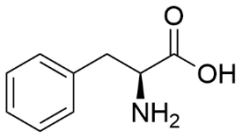
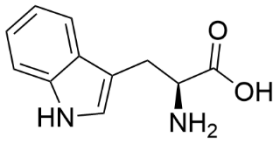
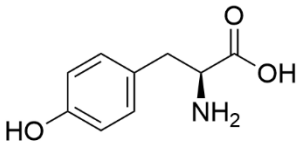
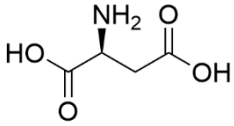
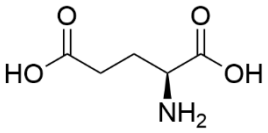
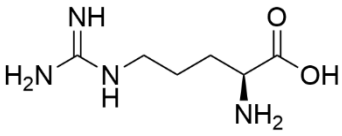
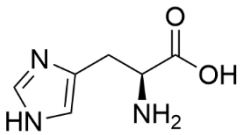
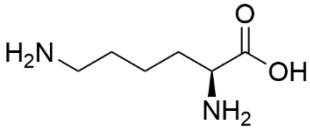
1.2.4 Understanding amino acids as biostimulants

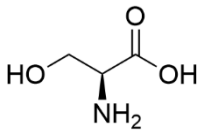
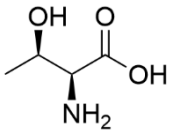
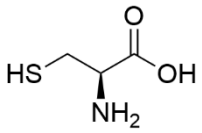
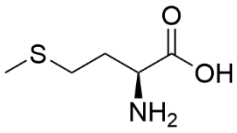
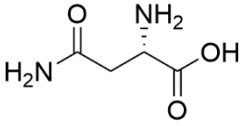
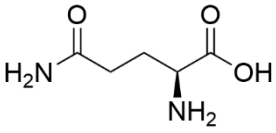
The following section covers the current state of the literature regarding specific amino acids, and their known roles as plant biostimulants, in more detail, focusing on foliar application.

For the purposes of this thesis, the 20 eukaryote nuclear genome encoded proteinogenic amino acids will be classified based on the structure of their side chains into seven groups: aliphatic (having an alkyl chain); aromatic (containing an aromatic ring); acidic; basic; hydroxylic (containing an OH hydroxyl group); sulfur-containing; and amidic (containing an amide group). Due to the variety of functional groups present in these side chains, amino acids exhibit a range of reactivities. The pK_a of the side chain of an amino acid plays a major role in its reactivity and interactions within a plant (Table 1). Those with acidic side chains such as glutamic acid and aspartic acid will be mostly in their dissociated (conjugate base) form at normal physiological pH. Whilst those with more basic side chains (high pK_a) such as arginine and lysine will be predominantly in their protonated form (Table 1).

Table 1: *Structural representations and pKa values of the α-carboxyl, α-amino, and side chain functional groups for the 20 common proteogenic amino acids. Amino acids are grouped according to functional similarity of their side chains (e.g., aliphatic, aromatic, acidic, sulfur-containing), reflecting shared chemical properties relevant to protein structure and reactivity. pKa values are sourced from standard biochemical reference data.*[36]

	Amino Acids	Structures	pK_a Value		
			α-Carboxyl Group	α-Ammonium Group	Side Chain
Aliphatic	Alanine (Ala)		2.34	9.69	-
	Glycine (Gly)		2.34	9.60	-
	Isoleucine (Ile)		2.36	9.68	-
	Leucine (Leu)		2.36	9.63	-
	Proline (Pro)		1.99	10.60	-
	Valine (Val)		2.30	9.74	-

Aromatic	Phenylalanine (Phe)		1.83	9.13	-
	Tryptophan (Trp)		2.38	9.39	-
	Tyrosine (Tyr)		2.20	9.11	10.07
Acidic	Aspartic acid (Asp)		1.92	9.87	3.87
	Glutamic acid (Glu)		2.19	9.67	4.25
Basic	Arginine (Arg)		2.18	9.09	13.2
	Histidine (His)		1.78	8.97	5.97
	Lysine (Lys)		2.18	8.95	10.53

Hydroxy-lic	Serine (Ser)		2.21	9.15	-
	Threonine (Thr)		2.63	10.43	-
Sulfur-contain-ing	Cysteine (Cys)		1.71	8.33	10.78
	Methionine (Met)		2.28	9.21	-
Amidic	Asparagine (Asn)		2.02	8.80	-
	Glutamine (Gln)		2.17	9.13	-

Foliar application of many individual amino acids has been shown to increase the yield of crops and reduce the effects of stresses on plants. This improves growth in less desirable environmental conditions, such as where soils are saline, or where heat or drought stress is present.[37] Understanding how these amino acids interact with plants to cause such effects is necessary for exploiting these beneficial qualities further.

Under normal conditions, amino acids can enter the plant via several pathways: through stomata on the leaf underside, across the waxy leaf cuticle, or via root absorption.[38] Amino acids that are taken up by or synthesised in the roots, then move up the plants via transpiration through the xylem, while those entering through or

synthesised in the leaves then travel down the plant via translocation through the phloem (Fig. 3).[39, 40]

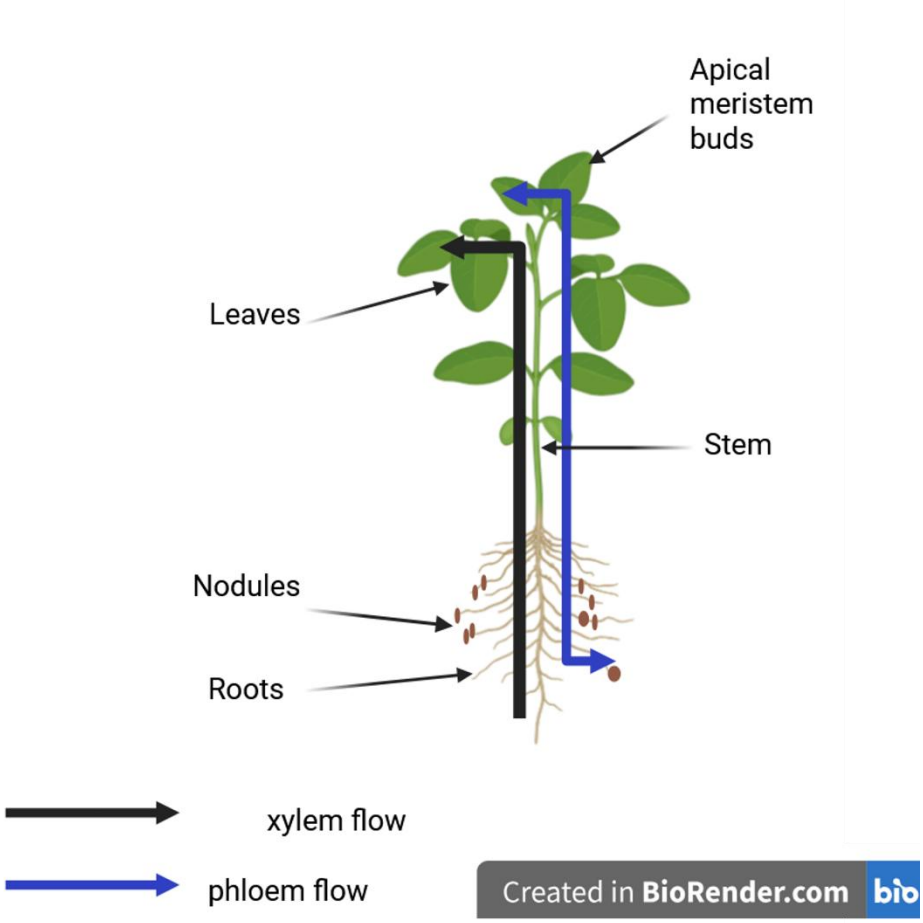


Figure 3: *Directional transport pathways of xylem and phloem in a soybean plant. Xylem flow (black arrow) facilitates unidirectional upward movement of water and mineral nutrients from the roots to the aerial tissues. Phloem flow (blue arrow) enables bidirectional translocation of photoassimilates and signalling molecules between source tissues (e.g. leaves) and sink tissues (e.g. roots, nodules, and growing buds).[40] Created in BioRender.com*

Amino acids serve multiple critical functions within plants, particularly with protein synthesis, nitrogen storage and transport. Plants primarily metabolise amino acids through transamination, converting them into keto acids to assimilate and redistribute nitrogen. Due to this role in nitrogen transport, the concentration of free amino acids in a plant is dependent on the level of nitrogen present.[14] As well as storing nitrogen, specific amino acids play many roles within plants. These affect membrane permeability, antioxidant effects and rate of photosynthesis and are discussed in detail in the subsequent sections.

1.3 Previous research on foliar application of amino acids

In order to consider these factors further a review of the literature relating to foliar application of amino acids was undertaken. An initial Web of Science search returned hundreds of results. From these, 51 papers were selected as most relevant to include in the literature review. A summary of the effects recorded and suggested mechanism for the foliar application of each amino acid can be found in Table 2.

For foliarly applied amino acids to enter the leaf and reach the vascular tissues, they must pass through the waxy cuticle, epidermal and mesophyll cell layers, before being actively loaded into xylem or phloem for redistribution (Fig. 4). If the treatment reaches the underside of the leaf, entry may also occur via stomata, which provide a direct aqueous pathway into the apoplast.

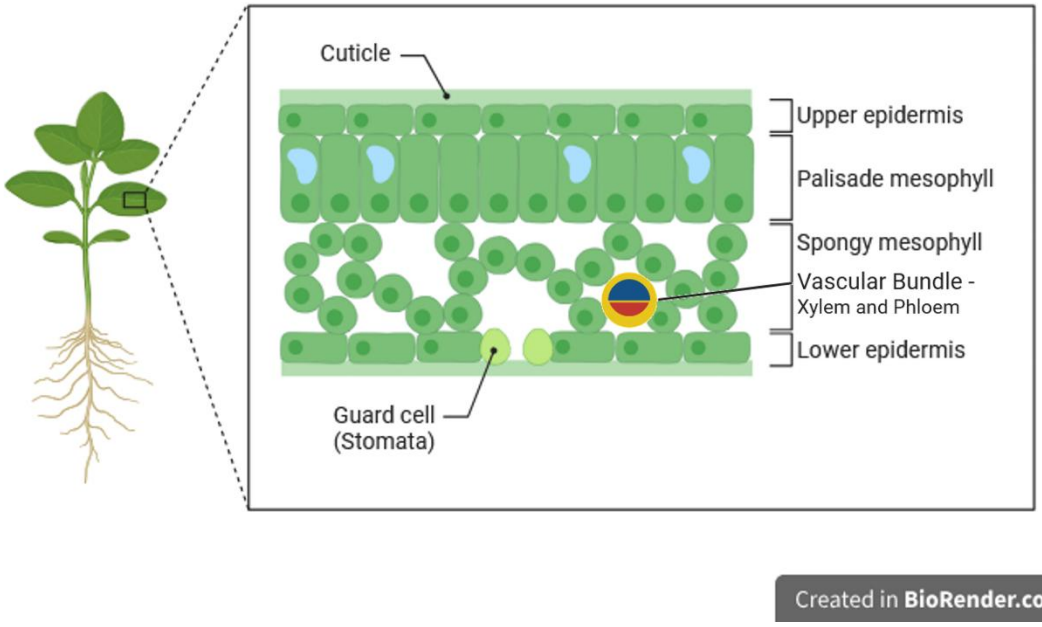


Figure 4: *Cross section of a soybean leaf showing key anatomical features. The cuticle forms a waxy barrier above the epidermis. Beneath this, the palisade mesophyll contains chloroplast rich cells for photosynthesis, followed by spongy mesophyll with air spaces for gas exchange. Vascular bundles (xylem and phloem) transport water, minerals, and assimilates. Stomata in the lower epidermis, bordered by guard cells, regulate gas exchange and transpiration. Created with BioRender.com.*

Foliar applied amino acids enter the leaves by passing through the cuticle. There is some discrepancy regarding the dominant pathway for amino acid uptake across the cuticle. Some studies suggest that all amino acids permeate primarily via aqueous

pores, with uptake rates determined by charge and hydration shell size.[41] In this model, small non-polar amino acids have the fastest uptake, while large or negatively charged amino acids permeate more slowly.

Alternatively, a dual pathway model (Fig. 5) has been proposed in which hydrophobic solutes may diffuse directly through the lipid rich waxy cuticle, whereas hydrophilic and charged solutes rely on aqueous pores.[42] In this case, amino acids with hydrophobic side chains can enter the leaf more readily due to their lipophilicity, as observed in species with thick waxy cuticles such as grapevine (*Vitis vinifera*). Hydrophilic amino acids, by contrast, move through aqueous pores, which are often negatively charged and hydrophilic.[43, 44] At pH values below 9, the ammonium group of most amino acids is positively charged, facilitating electrostatic attraction and movement through these pores. However, the net charge of an amino acid at physiological pH (around 7.4) depends on its side chain properties (Table 1).

Cuticle structure itself also influences uptake: in plants with thinner or less waxy cuticles, such as maize (*Zea mays*), both hydrophilic and hydrophobic amino acids may encounter less resistance during diffusion.

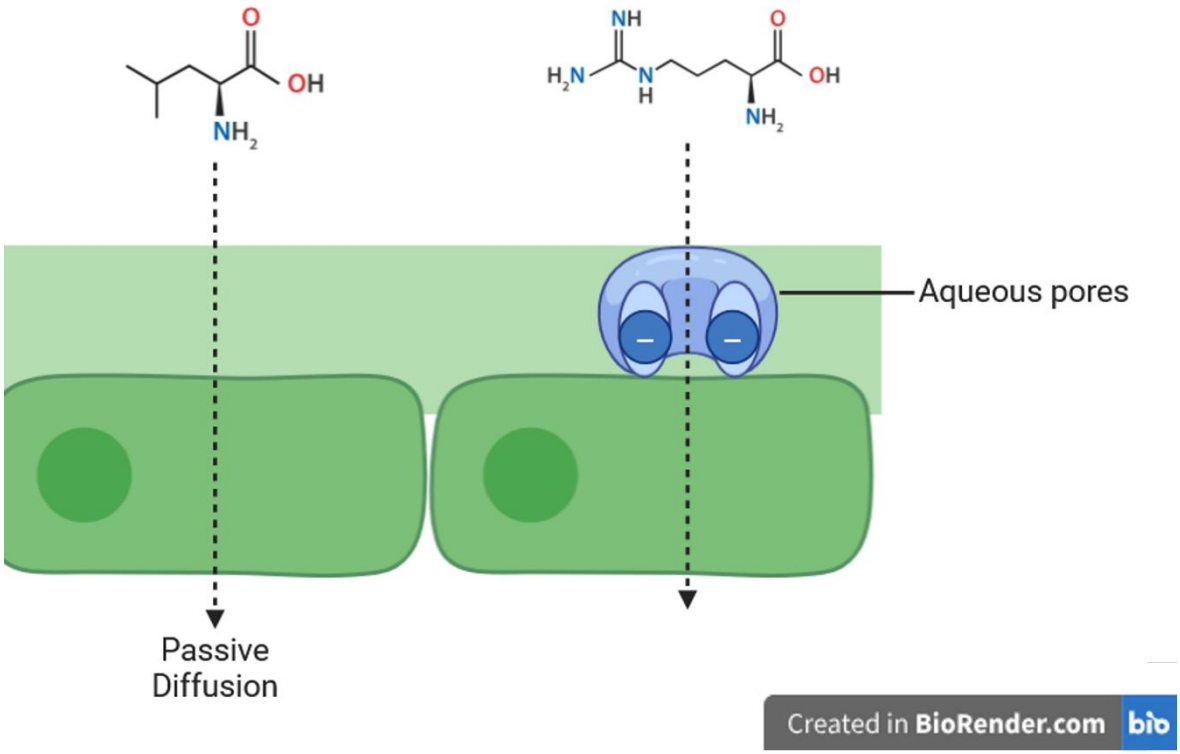


Figure 5: *Two proposed pathways for amino acid movement across the leaf cuticle. Small non-polar solutes may cross directly through the waxy cuticle via passive diffusion (left). Hydrophilic and charged amino acids permeate via aqueous pores (right), where uptake efficiency depends on solute size, hydration shell, and charge interactions with the pore walls. Created with BioRender.com.*

After passing through the cuticle, amino acids must diffuse across the cell wall, a porous hydrophilic matrix composed of cellulose, hemicellulose, and pectin. This layer does not act as a selective barrier, but its structure can influence the rate of diffusion depending on the size and charge of the molecule. Finally, amino acids must cross the plasma membrane of epidermal or mesophyll cells to enter the cytoplasm (Fig. 6). Here, uptake is governed by the electrochemical gradient and mediated by specific transport proteins. Most physiological and molecular evidence indicates that amino acid uptake across the plasma membrane is mediated by specific transporters, with saturable, proton-coupled kinetics.[45-47] While some studies have hypothesised that hydrophobic amino acids may also diffuse across membranes.[48]

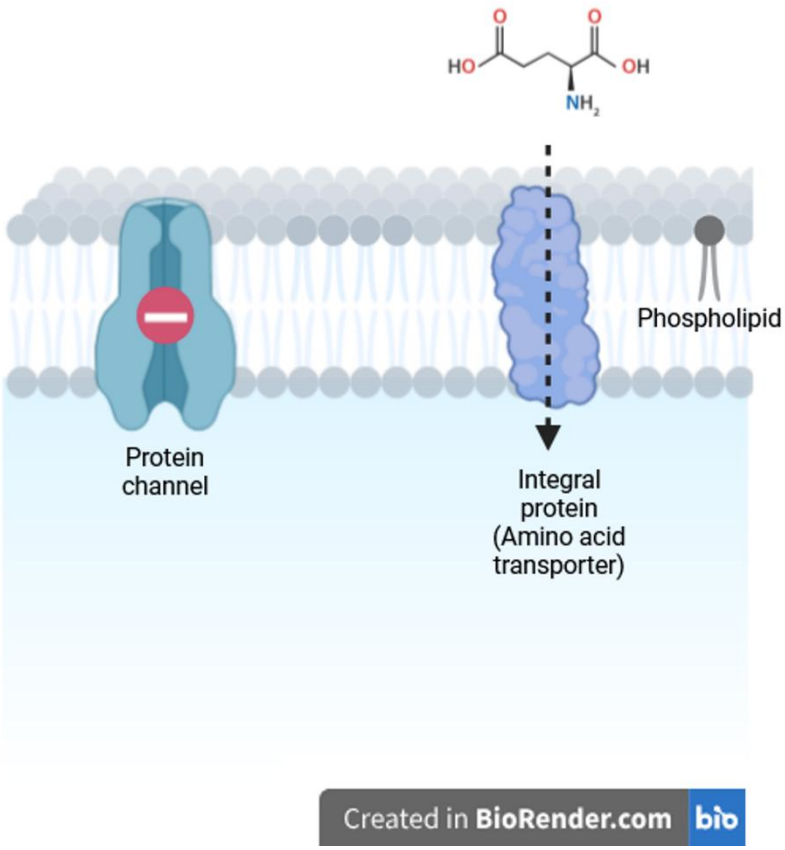


Figure 6: *Amino acid uptake across the plasma membrane is mediated by integral transport proteins rather than by non-specific aqueous channels. Distinct transporter families (e.g. AAPs, LHTs, UMAMITs) selectively recognise different amino acid classes, enabling controlled uptake into the cytoplasm. the phospholipid bilayer provides the structural barrier, while transporters confer specificity and regulation. Created with BioRender.com.*

There are multiple types of amino acid transporters known to mediate the uptake of basic, acidic, or neutral amino acids. Some, such as amino acid permeases (AAPs), have broad substrate specificity and can transport a wide range of amino acids. Others, like lysine-histidine transporters (LHTs), are more selective and were initially characterised for uptake of lysine and histidine, though they also transport other amino acids. These transporters mediate selective, energy-dependent uptake across the plasma membrane.[46] From here the amino acids enter the cytoplasm, where they may be metabolised locally or redistributed around the plant. Redistribution typically involves active loading into vascular tissues such as the phloem via specialised transporters in companion or parenchyma cells.

In some cases, the concentration of an amino acid solution used, and therefore the amount of amino acid applied to the plant, has been found to play a role in the effect on plant yield. An example of this can be seen with the application of L-methionine to lettuce (*Lactuca sativa* L.).[49] At high concentrations an inverse effect on growth was seen.[50] In comparison, the application of lower concentrations saw an increase in yield. This suggests that overstimulation damages photosynthetic apparatus, while moderate supplementation can enhance productivity.

1.3.1 Controls and conditions

To evaluate the correlation between the foliar application of amino acids and the proposed effects on plants, it is crucial to consider the controls employed to support these assertions. Additionally, it is essential to examine the growing conditions utilised in the plant trials. In the papers discussed in this review there are a number of themes in the controls used. Some papers report extensively on their control treatments and the conditions in which their trials occurred, while others provide fewer details.

1.3.1.1 Treatment of control plants

In the papers reviewed, control treatments were either:

a) no treatment given. This provides a comparison that is true to the natural growth conditions without any additional treatments, but still allows for doubt regarding the specific ingredient (e.g. surfactant) in the foliar spray responsible for the effects;
b) water. Distilled water is used in most cases, although tap water is sometimes used, which has the same issue as a) with regard to uncertainty over the active component; and
c) treatment with same composition of spray but with amino acid removed (i.e. still containing surfactants and other adjuvants). This eliminates the possibility that something other than the amino acid is causing the observed effects.

1.3.1.2 Growth conditions

An imperative consideration in the comparative analysis of studies of the efficacy of foliar sprays involves an examination of the environmental parameters governing plant growth. Ideally, studies should document the minimum and maximum temperatures the plants encountered, along with details about soil composition and humidity levels. Additionally, the timing of experiments holds significance, as the season in which trials are conducted can influence the growth of the plants and, therefore, add another level of uncertainty as to the effects of the biostimulant.[51] Of the 51 research papers covered in this review, 13 did not include growing conditions. When growth conditions were given they varied greatly between papers so comparability became more difficult. The problems with comparability should be considered when exploring the effects of foliar application of amino acids, as variations in growth conditions can significantly influence their observed efficacy and outcomes. The reviewed literature covered a variety of amino acid applications, noting both their effects and, where identified, the mechanisms behind them. Full details of the controls and growth conditions covered in each paper in this section can be found in Appendix A.

1.3.2 Foliar effects of aliphatic amino acids

There is no consistent theme for the activities of exogenously applied aliphatic amino acids. However a few similarities between some of them: proline and alanine both contribute to the plant's antioxidant defence system, proline working directly as an antioxidant and alanine promoting an increase in polyphenolic compounds which can act as antioxidants; glycine and proline cause an increase in chlorophyll concentration; the branched chain amino acids appear to promote the response of plant growth regulators, although more research is needed to determine their exact mode of action.[48] Only proline has been extensively studied in isolation and it this amino acid that gives the best evidence for a positive effect on growth. In all cases however, there is a need for better control experiments to rule out effects from other components of the spray.

1.3.2.1 Alanine

Alanine has mostly been studied as part of a mixture with other amino acids, and so the actual influence of the alanine on the plant growth is unknown. A comparative study of foliar application of amino acids to lisianthus (*Eustoma grandiflorum)* showed that foliar application of alanine reduced the final dry mass of the plant in comparison to a control of water spray. It also reduced the number of leaves produced, and had a negative effect on the yield.[52]

However, contrary to these findings, one study has shown a link between the foliar application of alanine and an improved quality of some fruit species, specifically Fuji apple (*Malus domestica 'Fuji'*), by increasing the total amount of compounds responsible for aroma and increasing the activity of alcohol dehydrogenase (ADH) and alcohol acyltransferase (AAT) enzymes involved in the metabolism of amino acids.[53] These effects on fruit yield for alanine were significantly greater than that the effects of foliar application of leucine, isoleucine and valine.[54] However, in both these studies, limited information was provided about the control plants used or the conditions in which the field trials took place, which leaves room for doubt over the significance of these effects.

1.3.2.2 Glycine

The effects of glycine vary greatly between plant species, ranging from no effect in some species such as maize (*Zea mays*),[55] to growth improvements in others such as lettuce (*Lactuca sativa*) and sweet basil (*Ocimum basilicum*). In the latter cases, glycine at low concentration (between 250 $mg.L^{-1}$ and 500 $mg.L^{-1}$) improved plant growth and several characteristics relating to yield, including chlorophyll concentration, fresh and dry mass, and leaf area.[56, 57] It was suggested that this improved yield may be related to an increase in the concentration of chlorophyll and carotenoids in plant leaves. In a comparative study of amino acids in lisianthus (*Eustoma grandiflorum)*,[52] glycine was shown to increase chlorophyll levels, and produced the greatest increase in plant height, despite no significant difference in dry mass being observed. Control plants in these studies were treated with distilled water, so there remains some uncertainty over the active ingredient.

As glycine is a precursor for the synthesis of chlorophyll via chlorophyll A, this pathway may operate to increase chlorophyll levels.[58] Khan *et al.* (2012) also suggested that an increase in chlorophyll concentration caused by foliar application of amino acids, with a concomitant increase in the rate of photosynthesis, leads to improved plant yield.[59]

More extensive research has been done into the application of exogenous glycine betaine (Fig. 7). This molecule has proved to be highly effective when used in a number of different plant species under differing stress conditions, including drought, high salinity, and extreme temperatures.[60, 61]

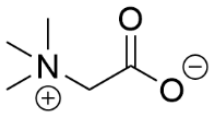


Figure 7: *Chemical structure of trimethyl glycine (glycine betaine), a quaternary ammonium compound involved in plant tolerance to osmotic and environmental stress.*

However, in cotton (*Gossypium herbaceum*), glycine betaine did not show improvement against environmental stresses.[62]

Glycine betaine can reduce oxidative stress as it increases expression of antioxidant enzymes (e.g. ascorbate peroxidase).[63] This effect is also thought to be related to a sequestering of sodium ions in plant roots. Under salt related stress in rice (*Oryza sativa*), application of glycine betaine as a foliar spray yielded lower sodium ion levels in the shoots (a sign of reduced stress) compared to control plants, where individual substrates were omitted from each trial.[64]

1.3.2.3 Branched chain amino acids (Ile, Leu, Val)

The limited studies on the foliar application of these three branched chain amino acids have yielded a similar improvement in growth to the soil application of urea, with increased shoot density but no significant difference in root length or in root and shoot weight.[48] However, there was no significant change in the levels of any growth factors compared to untreated plants when these amino acids were applied individually or in pairs. The improved shoot density has been suggested to be related

to a plant growth regulator, rather than provision of an additional nitrogen source. However, no suggestion as to which plant growth hormone the action was related to or how the amino acid may interact with these growth hormones was given.

1.3.2.4 Proline

The foliar application of proline is a potentially effective tool for improving yield in poor growing environments affected by the changing climate.[65, 66] It remains unclear whether the increased concentration of endogenous L-proline is a direct result of plant stress, or a regulated internal response as part of the plant's antioxidant defence system.[67] It has been observed that in plants which are put under stressed growing conditions there is a greater concentration of proline and a lower yield.[68] However, when L-proline is applied exogenously, positive effects are seen against heat stress, high salinity, and heavy metal induced stress in many plant species, suggesting that the internal response to stress is the key factor responsible.[69-71] The positive effect proline application has on environmentally stressed plants may be linked to antioxidant effects, scavenging reactive oxygen species (ROS) to form adducts of proline derivatives such as 4-hydroxyproline and 3-hydroxyproline as well as the promotion of nitrogen fixation.[72, 73] Mondal *et al.* (2012) suggest that proline also offers a source of carbon, nitrogen, and oxygen to the plant to aid its recovery if exposed to stress.[52] However, as this activity is not seen with other amino acids, and the effect is seen with very low concentrations, it seems unlikely that such a small amount of additional proline could produce such beneficial results if acting purely as a source of nutrients.

It has been suggested that foliar application of L-proline improves nitrogen fixation at the roots of plants, leading to improvements in the synthesis of amino acids and chlorophyll, leading eventually to increased photosynthesis and improved yield.[74] Interestingly, it is notable in these studies that increased chlorophyll content in the plant was correlated to low amounts of nitrogen in the soil. This may be because proline uses a less favourable pathway for nitrogen fixation than the pathway used in the absence of stress and so is only used if the plant is under stress and proline is available.

The studies above all use either water or no treatment on the control plants, so there is room for doubt over whether another component in the mixture is responsible for the effect. However, given the number and diversity of studies reporting growth enhancements, proline clearly has some positive effect. The range of compositions used in these studies are all between 2 - 1600 $mg.L^{-1}$.

1.3.3 Foliar effects of aromatic amino acids

Aromatic amino acids all have an aromatic side chain and are ultraviolet light (UV) active, absorbing in the region 220 - 280 nm. Tryptophan (Trp) and phenylalanine are both classed as hydrophobic; however, the indole side chain of Trp has some amphiphilic character, with the pyrrole ring favouring water solvation. It is due to this property that Trp is often found in the interfacial regions of membrane proteins. Tyrosine is classified as amphiphilic. The phenolic pK_a of tyrosine (9.11) is also above physiological pH (7.4), so the phenolic group of tyrosine is not significantly deprotonated *in vivo* under normal conditions.

Foliar application of aromatic amino acids, especially Trp, induces an increase in the levels of phenolic compounds within the plant.[75] This response is mediated by an increase in the levels of abscisic acid. Increased levels of phenolic compounds can improve the ability of a plant to withstand oxidative stress by providing a sink for ROS.

1.3.3.1 Phenylalanine

Most research into the foliar application of phenylalanine has looked at its ability to increase the levels of phenolic compounds (anthocyanins and stilbenes) in grapes (*Vitis vinifera*), which is beneficial for improving quality in wine production. The application also increases synthesis of some amino acids within the grapes, particularly phenylalanine itself, without changing the total amino acid content, when compared to control plants with phenylalanine omitted.[76-80] Increased levels of phenolic compounds lead to improved antioxidant activity in grape vines grown in low nitrogen environments.[81-83] This increase occurs as the increase in phenylalanine leads to an increase in abscisic acid (ABA) levels. These increased ABA levels

upregulate the expression of some genes (*VvPAL, VvCHS, VvF3H, VvUFGT,* and *VvSTS*) in the phenolic synthesis pathway.[81]

Foliar application of both phenylalanine and tryptophan in combination has also been shown to lead to an increase in total phenolic compounds present within the plant. The ability of aromatic amino acids to improve plant yield and defend against harsh environmental conditions may be related to the antioxidant properties of these phenolic compounds.

1.3.3.2 Tryptophan

L-Tryptophan is known to serve as a precursor for auxins, a group of plant growth hormones, including indole-3-acetic acid (IAA). IAA is a key plant hormone which regulates different aspects of growth and stress response.[84]

Many plant species have shown growth improvements from foliar application of tryptophan including okra, lupines, maize and apples.[54, 85-88] In okra, fruit was produced earlier and in greater volume, and the size of the plant increased compared to untreated control samples.[85] This study also reported a higher level of chlorophyll, stomatal conductance and transpiration and suggested that all these effects are related to L-tryptophan-dependent IAA biosynthesis.

In lupines (*Lupines termis L)*, foliar application of tryptophan increased both overall growth and levels of several key chemical components including alkaloids and phenolic compounds. Physiological measures of growth increased, such as plant height and the mass and yield of seeds, alongside increases in key chemical contents of the plant, such as total nitrogen, crude protein, total soluble sugars, total alkaloids, and total phenolic compounds. Smaller increases were seen in the levels of ascorbic acid.

In soybean plants (*Glycine max*), foliar application of tryptophan enhanced root growth, though to a lesser extent than soil drenching. As IAA is responsible for the promotion of lateral root growth, this study suggests that the application of tryptophan is increasing IAA biosynthesis, which causes an improvement in growth.[89]

In maize, it was observed that foliar application of tryptophan caused significantly higher relative water content, leaf membrane stability index, chlorophyll and potassium content. The treatment reduced the effect of induced drought stress on the plants. The study also suggests that tryptophan can act as an osmolyte, an ion transport regulator, modulate stomatal opening and detoxify the effects of heavy metals.[88]

Similar positive growth traits were seen in foliar application to apple orchards, showing improved fruit weight, size, length, and diameter when applied at 100 ppm. However when combined with glycine the effects were even more pronounced though the abscisic acid level decreased.[90] This is particularly interesting as the mode of action of phenylalanine has been shown to be driven by an increase in abscisic acid and so it is likely that the mode of action is different here.

1.3.3.3 Tyrosine

Few studies have been published looking at the effect on yield of tyrosine as a foliar spray. In one study tyrosine was shown to increase total phenolic content in one species of mint, *M. piperita* 'Swiss', when applied at a concentration of 100 $mg.L^{-1}$, compared to water sprayed controls. However, this was not found to be a significant increase.[75]

Tyrosine has been used as part of protein mixtures applied to leaves, with no indication as to the role that tyrosine may play within the mixture.[33] Tyrosine may have beneficial application as a soil based fertiliser however.[91]

1.3.4 Foliar effects of acidic amino acids

Aspartic acid and glutamic acid are both hydrophilic amino acid due to their acidic side chains permitting hydrogen bonding with, and solvation by, water molecules. The side chains of both aspartic acid and glutamic acid have low pK_a values (3.9 and 4.3 respectively) and so, at physiological pH, they are predominantly deprotonated in the carboxylate form.

Similar to aromatic amino acids, foliar application of aspartic acid improves a plant's antioxidant defence system.[92] However, rather than causing an increase in the levels of other compounds which can act as sinks for ROS, aspartic acid itself acts as a sink for ROS.[93] Similar to the aliphatic amino acids, glutamic acid increases the level of chlorophyll within plants, and improves the efficiency of photosystem II (PSII), increasing the rate of photosynthesis.[94]

1.3.4.1 Aspartic acid

Aspartic acid has been shown to alleviate osmotic stress caused by water salinity in tomato plants, with this effect being linked to antioxidant properties of the amino acid as well as the ability of this amino acid to act as an osmoregulatory compound.[95] It has also been used to reduce the effect of oxidative stress caused by cadmium poisoning in water supplies from contaminated sources. This occurs through a reduction in the uptake of cadmium ions by the plant and combatting the ROS formed by the present cadmium ions.[92] This may be due to the amino acid forming a complex with the Cd^{2+} ion, reducing free mobility within the cell.[93] The same effect is also seen for foliar application of proline.[93]

1.3.4.2 Glutamic acid

Glutamic acid was shown to have a positive impact on growth and fruit quality in grapes, with increased chlorophyll levels, leaf size, and fruit yield. Higher concentrations of nitrogen, potassium, and phosphorus were also found, when compared to control plants treated with water.[96] Similar effects, of increased chlorophyll levels and rates of photosynthesis, have been observed for hawthorn (*Crataegus monogyna*) and Chinese chives (*Allium tuberosum*).[94, 97] It has been suggested that glutamate improves electron transfer within PSII, a protein mediated complex responsible for generating energy for the plant but the mechanism by which this effect occurs is unclear.[94]

Time course studies of isotopically labelled glutamate applied to bentgrass leaves (*Agrostis capillaris*) at a concentration of 0.01 $mol.L^{-1}$ showed that glutamate is utilised in the leaf as a precursor for *gamma*-aminobutyric acid (GABA) and proline

synthesis. These metabolites are known to play a role within plant stress adaptation.[98] This study also concluded that the glutamate enters the plant leaf whole, being metabolised within plant cells, rather than being metabolised to ammonia on the leaf surface. No control experiments were reported for this study, however.

1.3.5 Foliar effects of basic amino acids

Arginine, histidine, and lysine have basic side chains which are protonated at neutral pH. Arginine and histidine have hydrophilic side chains, while lysine is amphipathic due to an alkyl chain separating the alpha-carbon from the amino group. Due to the high pK_a of their side chains (Table 1), these amino acids exist in their protonated form at physiological pH.

The effect of foliar application of arginine is very different between plant species. In some plant species arginine induces an increase chlorophyll levels, but in others it is shown to be a poor biostimulant.[99, 100] Histidine and lysine by contrast, are both able to form complexes with metals to help transport these metals into plants.[101, 102]

1.3.5.1 Arginine

Studies of arginine applied to leaves of shrub willow (S*alix purpurea* L., and *Salix viminalis* L. × (*Salix sachalinensis* F. Schmidt × *Salix miyabeana* Seemen)) showed it to be a poorer growth stimulant than urea, but also clarified that the two are metabolised within plants in different ways.[99] When applied to grapevines, no significant difference was seen from the control for measure of pH or total anthocyanins. However, total flavanol levels and proline concentration both decreased.[76] Decreases in flavanol levels suggests that basic amino acids are metabolised in plants through a different pathway to aromatic amino acids. For aromatic amino acids, the concentrations of polyphenolic compounds, a group of compounds to which flavanols belong, increased following application to leaves (see section 1.3.3).[77-80]

When applied to tomato plants, arginine has been shown to significantly improve plant growth, and had a positive effect on total fruit production and chlorophyll

content, though no significant effect was seen on the average mass of individual fruits.[100]

1.3.5.2 Histidine and Lysine

Histidine and lysine have been most extensively studied as their chelates with metals, rather than as free amino acids. Complexes of histidine and lysine with metal ions have been shown to improve the resistance of plants to increased salinity and modify their nitrogen content.[101]

Application of iron complexes of histidine and arginine both reduced the damage caused to tomato plants by increased salinity.[102] An Fe-glycine complex was also used for comparison. Both arginine and histidine were more effective in reducing the effects of salinity than glycine in one tomato cultivar, whereas all three expressed similar ability in a second tomato cultivar.

Other studies have looked at the effects of foliar application of zinc complexes of amino acids including those previously used with iron and the other basic amino acid, lysine.[101] Zn-Lys was the most effective foliar spray, with several cultivars of onion experiencing improved growth, bulb size, and pyruvic acid content. The Zn-Lys foliar spray also reduced nitrogen content of the bulb. This may be due to this compound producing the highest zinc bulb content for the amino acid complexes tested, causing a greater rate of metabolism of nitrogen in the bulb.[103] Application of the unchelated, free form of lysine did increase both the number and size of leaves, but had no effect on nitrogen concentration in the bulb. This indicates that the reduced nitrogen content was not a result of the amino acid application, but due to the presence of zinc.[104]

1.3.6 Foliar effects of hydroxylic amino acids

The hydroxylic amino acids, serine, and threonine, both contain a hydroxy group in their side chain. They are both hydrophilic due to this group being able to hydrogen bond with water. Serine and threonine have not been extensively examined and do not appear to yield a significant improvement in plant growth. In one study, serine inhibited plant growth to a similar extent to alanine, whereas threonine showed no

significant difference from a control group.[52] Threonine has been used as a ligand for zinc ions in the same way as histidine and lysine. Zn-Thr did increase number of leaves and leaf area in one cultivar but not the other, and to a lesser extent than Zn-Lys.[104]

1.3.7 Foliar effects of sulfur-containing amino acids

The side chain of cysteine is ionisable at physiological pH. Both cysteine and methionine can be metabolised through oxidation, with cysteine forming cystine, sulfinic acids, cysteic acid and sulfocysteine and methionine forming methionine sulfoxide. Both cysteine and methionine can reduce the adverse effects of increased salinity in the plant growth medium but exert their influence through different pathways.[7]

1.3.7.1 Cysteine

Cysteine has been shown to reverse negative effects of drought conditions. Foliar application of cysteine increased fresh weight of roots, root length, and chlorophyll concentration in Malka type maize (*Zea mays* L. var. Malka) under salt stress. In another maize variety (hybrid DTC) shoot length, total free amino acids, total phenolics, and free proline contents were significantly improved under similar environmental conditions.[7] Increased growth was also noted when cysteine was applied to plants growing under normal conditions, with plant height and yield having increased for treated plants in a number of other studies.[100, 105, 106]

1.3.7.2 Methionine

Methionine used in conjunction with proline and glutamine, and when applied with tryptophan, was shown to reverse the negative effects of salinity in tomato plants. What is particularly interesting is that the study showed that unlike with glycine betaine (Section 1.3.2.2) the effect is not thought to be due to changes in Na^+ ion concentrations in the leaf, but instead due to greater accumulation of soluble sugars. These play a role in quenching reactive oxygen species created by the ions more effectively than at control levels.[107, 108]

1.3.8 Foliar Effects of Amidic Amino Acids

The effect of foliar application of the amidic amino acids follows the trend of many other amino acids in causing an increase in chlorophyll concentration, which may increase plant growth and yield. It is also noteworthy that aspartate and glutamate are metabolites of the amidic amino acids asparagine and glutamine and may be responsible for some of the observed effects after foliar application of the corresponding amides.

1.3.8.1 Asparagine

Asparagine, as with many of the other amino acids previously mentioned, produces significant improvements in plant growth after foliar application as a spray to cress, including an increase in chlorophyll concentration (both chlorophyll a and b) and leaf nitrogen and phosphorus levels.[109] In increased salinity environments, asparagine was shown to reduce sodium ion concentration in leaves, while increasing levels of phosphorus, calcium, and potassium ions present in the roots of maize.[110]

1.3.8.2 Glutamine

Glutamine has been shown to increase height, fresh and dry weight, and chlorophyll concentration of a number of plants, but to a lesser extent than glycine. However, its effect when applied to the roots through the soil is more pronounced.[56]

A positive growth effect was also documented for application of glutamate (a metabolite of glutamine) to lettuce and sweet basil over a range of concentrations, with 250 $mg.L^{-1}$ (the lowest concentration) having the greatest improvement of root growth. This improved yield may be related to an increase in the concentration of chlorophyll and carotenoids in plant leaves.[57]

1.4 Plant growth regulators

Several papers covered in section 1.3, such as branched chain amino acids suggest that the effects observed from foliar application of amino acids were related to one of the five classes of plant growth regulators: auxins, gibberellins, cytokinins, abscisic acid and ethylene

(Fig. 8).[49] These groups of compounds work in different ways to regulate the growth of the plant, so when stimulated or inhibited they can change the plant's growth rate and qualities.[111]

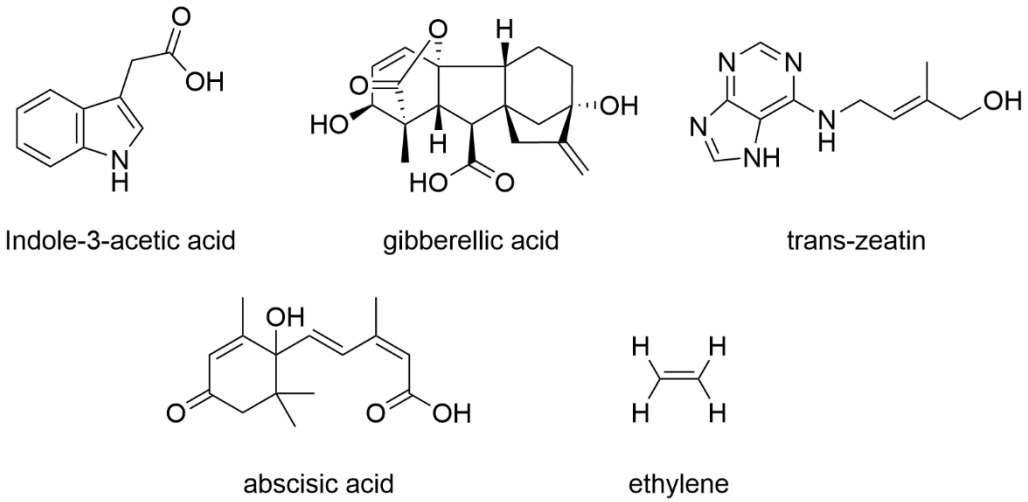
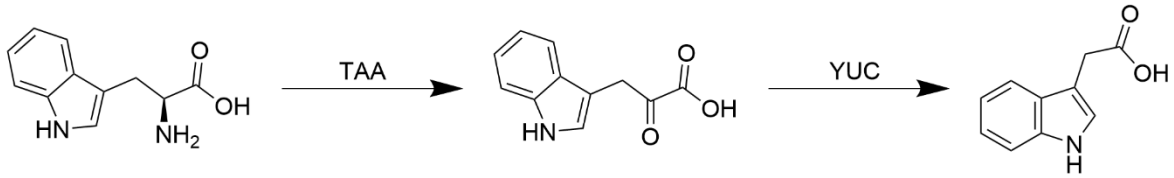


Figure 8: *Example structures of the five major classes of plant growth regulators discussed in this section: indole-3-acetic acid (auxin), gibberellic acid (gibberellin), trans-zeatin (cytokinin), abscisic acid, and ethylene.*

1.4.1 Auxins

Auxins are a group of hormones synthesised from amino acids, including tryptophan, at the plant's apical meristems, in the stem, bud and root tips.[49, 111] They regulate stem growth towards light, inhibiting the growth of lateral buds and promoting apical dominance. At a cellular level, auxins promote cell elongation and division to produce either organ structures or unorganised tissue.[112] Free auxin levels in plant cells are stabilised through conjugation, with excess auxins being stored in the cell to enable the plants to control auxin concentration. The most abundant naturally occurring auxin in plants is indole-3-acetic acid (IAA). This is synthesised from tryptophan via a two-step enzymatic pathway. First, tryptophan aminotransferases (TAA) convert the amino group into a keto acid. This is followed by the action of YUCCA flavin monooxygenases (YUC) flavin monooxygenases which mediate the formation of a shorter chained carboxylic acid (Scheme 1).[84]



Scheme 1: *Conversion of tryptophan to the predominant auxin, indole-3-acetic acid (IAA), mediated by two groups of enzymes: tryptophan aminotransferases (TAA) and YUCCA flavin monooxygenases (YUC).*

Studies have shown that rhizobacteria can metabolise L-tryptophan to produce indole-3-acetic acid in high concentrations, which led to suppression of root growth when taken up through plant roots.[113] In high auxin concentrations, ethylene synthesis is stimulated leading to this inhibiting effect. Under normal conditions, this is used to inhibit axillary bud growth and encourage growth from the apical bud.

1.4.2 Gibberellins

Gibberellins are tetracyclic diterpenoid plant hormones produced in the apical meristems and young tissues of plants, synthesised within the plant by the terpenoid pathway.[114] They promote stem elongation as well as increase leaf growth by stimulating cell division.[115, 116] They have also been used as an externally applied stimulant to reduce flowering and fruit number while increasing the size of the produced fruits.[117] There are around 100 characterised species of gibberellins, which fit into two different classes: C-19 gibberellins and C-20 gibberellins (Fig. 9). The C-19 gibberellins are made up of 19 carbons and feature a lactone bridge, while C-20 gibberellins contain an additional carbon.

GA_{12} GA_9

Figure 9: *Structures of GA_{12} and GA_9, illustrating the two structural classes of gibberellins. GA_{12} is a C-20 gibberellin, while GA_9 is a C-19 gibberellin formed by enzymatic carbon loss and lactone bridge formation during biosynthesis.*

Gibberellins also play critical roles in seed germination, flowering regulation and fruit development and can interact with other plant hormones such as abscisic acid and auxins to coordinate growth responses.[118]

1.4.3 Cytokinins

Cytokinins are a class of hormones which promote cell division. They are produced at meristem at the tip of shoots and roots by isopentenyl transferase (IPT) and lonely guy (LOG) enzymes. They work alongside auxins in the production of organ structures.

Structurally, most naturally occurring cytokinins contain an adenine moiety with side chain modification at N-6.[119] The most common naturally occurring cytokinin is zeatin (Fig. 6). Both the cis and trans stereoisomers of zeatin are present in plants. However, the trans stereoisomer has greater bioactive properties.[120]

Exogenous application of L-methionine is known to act as a growth regulator for cytokinins leading to increased root initiation.[49]

1.4.4 Abscisic acid

Abscisic acid has many roles in the developmental growth and stress response of plants (Fig. 10). These include inhibiting cell growth to keep seeds dormant through the inhibition of gibberellin synthesis, increasing the speed of senescence, and controlling the opening and closing of stomata in wilting leaves to promote abscission of leaves.[111] This ability to control stomata opening also helps to defend the plant from water stress, keeping the stomata closed when water is scarce so that less water is lost from the leaf.[121] The ability of abscisic acid to inhibit growth comes from its effects on cell membranes. Inhibiting proton excretion of the cell curtails the movement of potassium ions in and out of the cell.[122] As previously discussed in section 1.3.5.1, abscisic acid is also known to regulate genes responsible for the promotion of the polyphenol synthetic pathways which help to protect the plant from oxidative stress.[81]

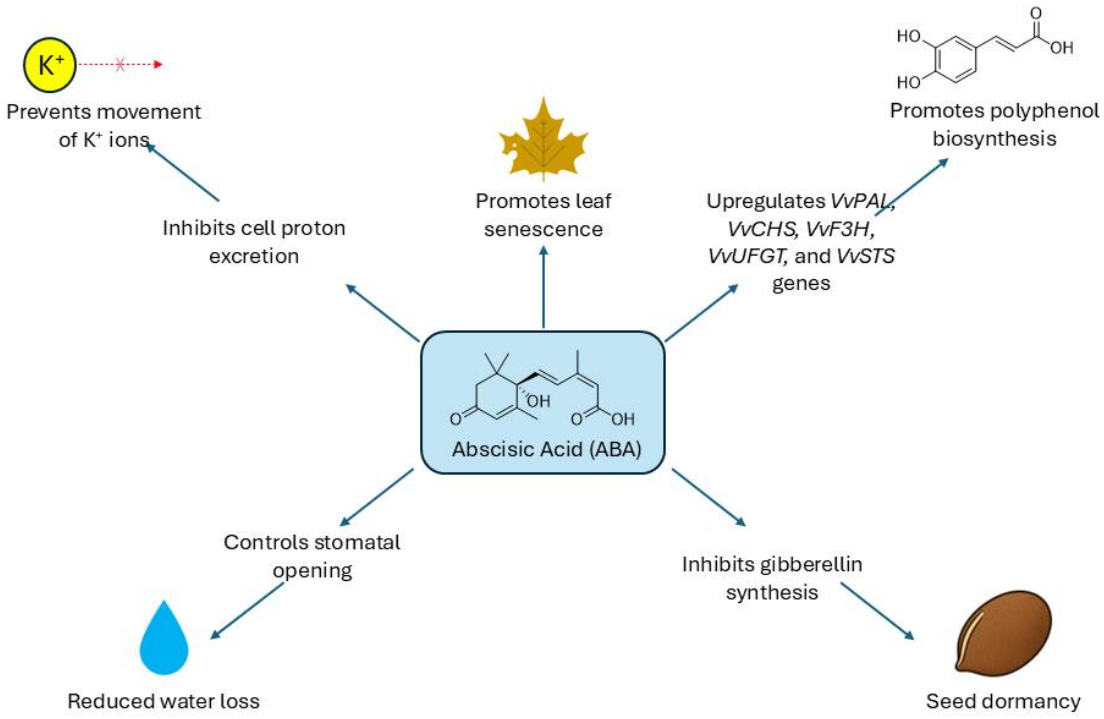


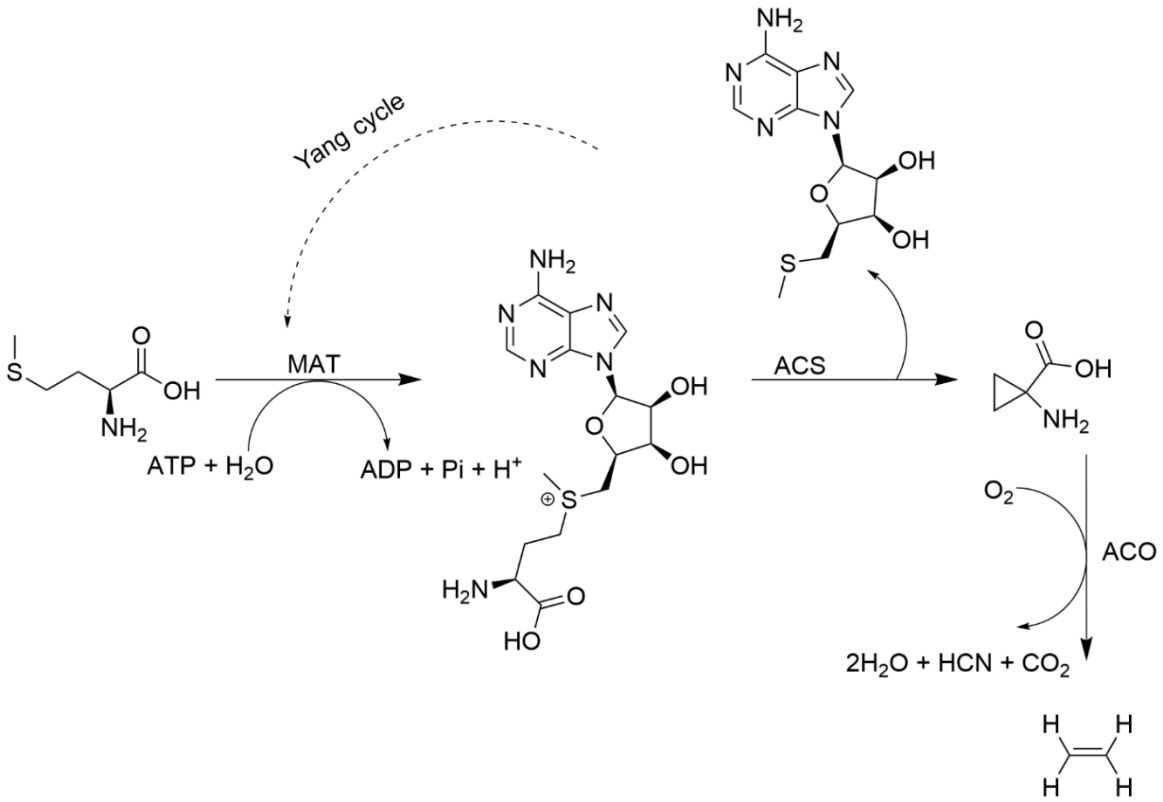
Figure 10: *Roles of abscisic acid (ABA) in plant development and stress adaptation. ABA regulates multiple physiological processes: it inhibits cell proton excretion, preventing K^+ ion movement and reducing cell expansion; controls stomatal opening, limiting water loss under drought stress; promotes leaf senescence; inhibits gibberellin synthesis to maintain seed dormancy; and upregulates VpPAL, WvCHS, WvF3H, WvUFGT and WvSTS genes, promoting polyphenol biosynthesis for oxidative stress protection.*

The increase in ABA levels in the study by Cheng *et al.* (2020) was linked to the foliar application of phenylalanine.[81] This suggests that Phe plays a role within the biosynthesis of ABA. However, this is likely through indirect promotion of biosynthetic pathways, as Phe is not known to act as an ABA precursor, which are typically carotenoids such as violaxanthin and neoxanthin.[123]

1.4.5 Ethylene

Methionine can be converted to ethylene gas by plant cells, with ethylene coming from the C4 and C5 positions on methionine.[111] The enzyme methionine adenosyltransferase (MAT) catalyses the conversion of methionine into S-adenosylmethionine (SAM). SAM is then converted into 1-aminocyclopropane-1-carboxylic acid (ACC) by ACC synthase (ACS), with a by-product of methylthioadenosine (MTA) being produced. Finally, ACC is oxidised by ACC oxidase

(ACO) to produce ethylene, alongside by-products including carbon dioxide, hydrogen cyanide, and water (Scheme 2). The MTA by-product is converted back to methionine via the Yang cycle.[124] This process allows the plant to regulate its ethylene concentration, even with low levels of methionine present.



Scheme 2: *Biosynthetic pathway of ethylene from methionine. Methionine is converted to S-adenosylmethionine (SAM) by methionine adenosyltransferase (MAT), then to 1-aminocyclopropane-1-carboxylic acid (ACC) by ACC synthase (ACS), and finally to ethylene by ACC oxidase (ACO), with CO_2, HCN, and water as by-products.*

Within the plant, ethylene contributes to leaf growth and controls degradation associated with aging of the plant.[125] It is also involved in various physiological processes including responses to stresses, fruit ripening and flower senescence. At low concentrations, ethylene has been shown to promote plant growth, while when present in higher concentrations, this growth was inhibited.[126] This suggests that different pathways are activated depending on concentration of ethylene.

1.5 Conclusions

The foliar application of certain, but not all, of the twenty main proteinogenic amino acids has been shown to benefit the growth of a range of plant species. While some mechanisms underlying these effects have genetic explanations, many remain poorly understood (Table 2). Given the increased demand for global agricultural productivity, it is clear that research into new methods of improving growth and resistance to stress in crops is necessary. When looking at previous advancements in agricultural techniques such as The Green Revolution, considerations should be made about the environmental and social impacts of any new treatments or technologies. Due to this growing demand for environmentally stable solutions, biostimulants offer an important area of research, especially given their potential to enhance agricultural productivity while minimising the risk of potentially harmful side effects.

Table 2: *Summary of growth effects and proposed mechanisms associated with foliar application of amino acids, based on a literature review of 51 peer-reviewed studies*.

	Amino Acids	Summary of Effect on Growth	Proposed Mechanisms
Aliphatic	Alanine (Ala)	Mostly studied in mixtures. • Negative effect on *Eustoma grandilorum* growth • Positive effect on quality of Fuji apples	Contributes to plant's antioxidant defence system through promotion of polyphenolic compounds - mechanism for this not given
	Glycine (Gly)	Significant variance in effects based on plant species • No effect on maize • At low concentrations, improves growth of lettuce, sweet basil and *Eustoma grandilorum*	Chlorophyll and carotenoid concentrations increase in leaves. Glycine is a precursor for chlorophyll A synthesis, which may be related to its mode of action.
	Isoleucine (Ile)	Limited studies on the branched-chain amino acids	Improved shoot density is suggested to be related to

	Leucine (Leu) Valine (Val)	• Increased root density without change to root length or shoot weight • Similar growth improvements to urea soil application	plant growth regulators, not just the presence of additional nitrogen.
	Proline (Pro)	• Positive effect on plant growth • Mitigates the effects of heat stress, high salinity, and heavy metal-induced stress	An increase in Chlorophyll content was observed. An increase in polyphenolic compounds was observed. Proline may improve nitrogen fixation in the roots, leading to increased amino acid and chlorophyll synthesis - no specific mechanism for this is given.
Aromatic	Phenylalanine (Phe)	• Improves quality of wine grapes through increased levels of phenolic compounds	Increased phenolic compound concentration improves antioxidant activity. The increase in phenolic compounds is due to phenylalanine metabolism leading to increased ABA levels, which upregulate genes in the phenolic synthesis pathway.
	Tryptophan (Trp)	• Increased overall growth in lupines • Increased plant height, mass, yield of seeds, total nitrogen, crude protein, total soluble sugars, alkaloids and phenolic compounds • Small increase in ABA	The small increase in ABA suggests a different mechanism to Phe, but its application has also been related to reducing oxidative stress through the production of compounds which can act as antioxidants.
	Tyrosine (Tyr)	Limited studies on Tyrosine • Increased total phenolic content in M. piperita 'Swiss' mint	Contributes to plant's antioxidant defence system through promotion of polyphenolic compounds -

			mechanism for this not given.
Acidic	Aspartic acid (Asp)	• Mitigates osmotic stress in tomato plants caused by water salinity • Also reduces oxidative stress from cadmium poisoning • Does not cause an increase in polyphenolic compounds	Acts as an osmoregulatory compound and directly as an antioxidant.
	Glutamic acid (Glu)	• Improved chlorophyll levels, leaf size, and fruit yield in grapes • Increased chlorophyll levels and photosynthesis rates in hawthorn and Chinese chives	May improve electron transfer within photosystem II. Shown to act as a precursor for GABA in proline synthesis. These are both involved in plant stress adaptation.
Basic	Arginine (Arg)	• Increased chlorophyll levels • Improved plant growth and fruit quality in tomatoes • Decrease in total flavanol levels and proline concentration	Different pathway to aromatic amino acids and to urea. No specific mechanism has been proposed.
	Histidine (His)	Mainly studied as chelates with metals rather than free amino acids. • Improved salinity resistance in some tomato cultivars	No indication of mechanism found.
	Lysine (Lys)		
Hydroxylic	Serine (Ser)	Very few studies on these compounds.	No indication of mechanism found.

	Threonine (Thr)	One study showed: • Serine inhibited plant growth similarly to alanine • Threonine showed no significant difference from the control group.	
Sulfur-containing	Cysteine (Cys)	• Improved salinity and drought resistance • Increased root and shoot growth, chlorophyll concentration, and phenolic compounds in maize	No indication of mechanism, though may be related to the mechanism of others that increase chlorophyll concentration.
	Methionine (Met)	• Improved salinity resistance when used in conjunction with other amino acids	Mechanism believed to be related to the accumulation of soluble sugars.
Amidic	Asparagine (Asn)	• Increased chlorophyll concentration, leaf nitrogen and phosphorus levels • Can reduce the effects of high-salinity environments	In a high-salinity environment it reduces sodium ion concentrations in maize leaves while increasing phosphorus, calcium, and potassium ion levels in the roots. No mechanism for chlorophyll increase given.
	Glutamine (Gln)	• Increased height, fresh and dry weight, chlorophyll concentration (not as well as glycine) • More pronounced effect when applied as soil treatment	No indication of mechanism found.

The literature review showed the diverse ways amino acids interact with plants when applied as foliar biostimulants, influencing both metabolic pathways and plant

hormone regulation. These benefits are typically attributed to increasing resistance to stress conditions or promoting the synthesis of chlorophyll (Fig. 11).[72, 73, 92, 93] However, in some cases, the control experiments are either not documented or insufficient to isolate the amino acid as the beneficial agent. When large errors of measurement for some metrics (such as dry mass) are factored in, it becomes clear that more research is needed to verify these hypotheses. In addition, other routes by which beneficial effects may be mediated have been proposed. These include amino acids providing a direct supply of "building blocks" for protein synthesis, thereby reducing the metabolic energy required for nascent amino acid synthesis and the interaction of the applied treatments with plant hormone regulators.[52]

Both chlorophyll synthesis and resistance to stress involve the production of beneficial compounds to the plant. In the latter case, the generation of polyphenolic compounds after foliar application of phenylalanine is accompanied by the upregulation of specific genes. Whether other amino acids that have been reported to elevate polyphenol concentrations also act through transcriptional regulation remains unknown, and testing this link would be a valuable direction for future research.

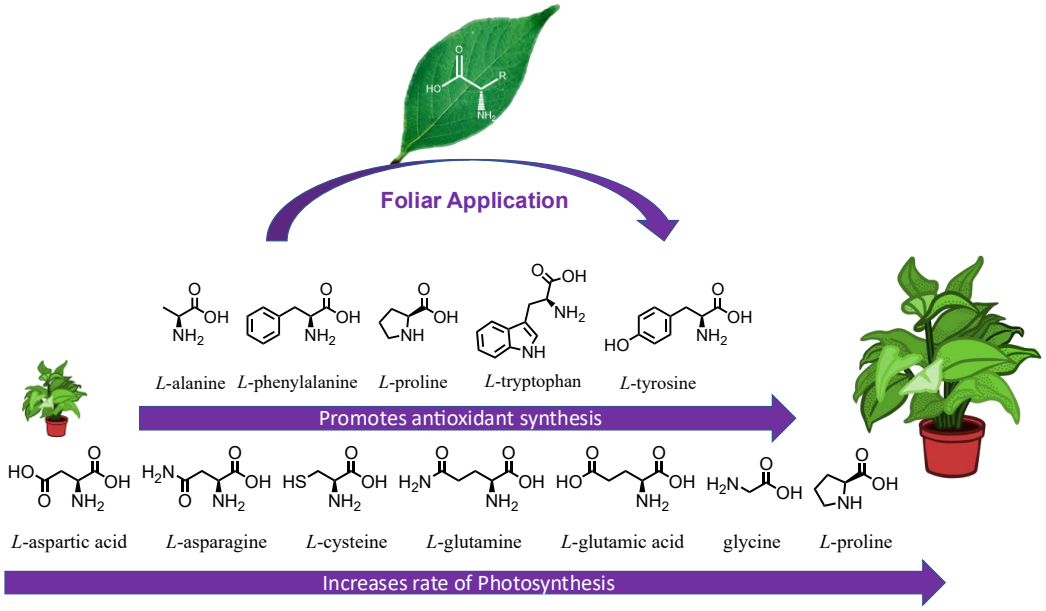


Figure 11: *Summary of growth improvement trends following foliar amino acid application. Consistently reported effects include enhanced antioxidant activity and increased photosynthetic rate.*

Adapting the use of amino acids for the agricultural market requires more research into the workings of individual amino acids across a wider range of plant species. For several of the major amino acids there is a paucity of available data. In some cases, this may be due to poor results from initial studies, but without comparative data it cannot be confirmed what effect specific amino acids have when applied as foliar spray. Where comparative studies have been done, such as with lisianthus (*Eustoma grandiflorum)*, whilst the differences in dry weight of seedlings between treatments are seen, the error bars are often too large for any significant effect to be ascertained (Fig. 12).[52]

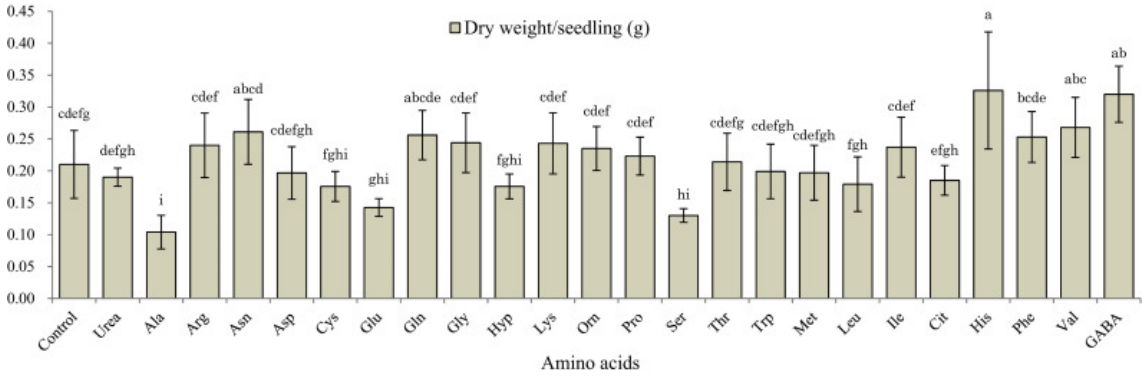


Figure 12: *Comparative study of foliar application of 23 amino acids and urea to lisianthus (Eustoma grandiflorum). Bars represent mean dry weight of seedlings (g) per treatment, with statistical groupings indicated by letters above each bar. Reprinted from Sci. Hortic., Vol. 192, Mondal MF, Asaduzzaman M, Tanaka H, Asao T. (2015), Effects of amino acids on the growth and flowering of Eustoma grandiflorum under autotoxicity in closed hydroponic culture, p. 455, with permission from Elsevier.*

The level of a foliar application that proves most effective is dependent on both the specific amino acid(s) used and the species or genotype of the plant to which it is applied.[91] However, the specific amount of amino acid given to each plant is often challenging to assess from the concentration of the treatment solution, as reported in many studies. The actual dose depends not only on the output speed and the surface area sprayed, but also on spray deposition and retention, which vary widely with leaf surface properties, run-off, and environmental conditions.[127] A more useful measurement to use for those wishing to apply this data in an agricultural setting would be litres per hectare (L ha^{-1}). Unlike concentration alone, this metric integrates the actual spray volume delivered per unit area, which reflects nozzle output, travel speed and canopy coverage. By expressing application rates in L ha^{-1} results from experimental studies can be directly compared with actual field practice.

The modes of action of some amino acids may be complementary. For example, aromatic amino acids increase the levels of polyphenols, whereas proline and aspartic acid quench reactive oxygen species themselves. Both increase resistance to oxidative stress and promote growth and yield but through different modes of action.[72, 128]

1.6 Summary and aims

Through the literature review of the use of amino acids as foliar applied biostimulants, it is evident that further research into the modes of action of these treatments is needed to fully understand how the application of different amino acids can cause the observed effects. This thesis attempts to further knowledge in this area with a focus on tryptophan through the tracking of labelled amino acids within soybeans.

The primary objectives of this project were to:

- Understand the metabolic pathways of foliar-applied tryptophan within the plant system.
- Investigate the mechanisms of foliar uptake of tryptophan, including factors affecting absorption efficiency.
- Develop and validate analytical methods for tracking labelled tryptophan within plant parts using mass spectrometry.
- Determine the uptake, transport, and conversion of tryptophan and its derivatives in different plant tissues.
- Assess the impact of tryptophan application on plant growth, yield, and stress response.
- Explore the hypothesis that amino acid interactions with plant hormones drive specific physiological effects, with different amino acid influencing different hormonal pathways.

The following chapters are organised to lead from the synthesis of labelled amino acids through to their application in plants and the subsequent analytical studies used to investigate their uptake and metabolism. In addition, the thesis addresses the development of R scripts for the streamlined processing of large-scale mass spectrometry datasets, the exploration of unforeseen synthetic challenges encountered during the project, and the methodology development undertaken in response.

Chapter 2: Design and synthesis of labelled amino acids

To facilitate monitoring of the uptake and movement of amino acids within the plant labelled amino acids were synthesised. Based on the existing literature, tryptophan was chosen as the primary amino acid to investigate for this thesis project. This was due to it showing significant promise for its foliar application even within the limited research to date. Additionally, previous work in the Sanderson group looked at the synthesis of tryptophan analogues, so there was precedence to start here.[129] To expand the scope, leucine and valine were additionally investigated, since both had demonstrated activity after foliar application (Section 1.3.2.3) and, being hydrophobic like tryptophan, offered the potential to probe whether hydrophobicity influences uptake and effects.

2.1 Labelling strategy

Three different forms of labelling were proposed for tracing through multiple methods (Fig. 13) to better understand the movement of the amino acids. All have the benefit of requiring minimal changes to the structure of the amino acid.

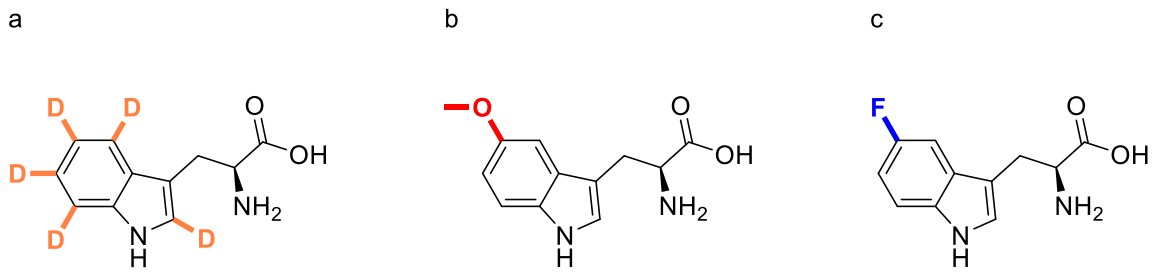


Figure 13: *Proposed labelling techniques for tracing amino acid movement. (a) Isotopic labelling with ^{2}H for mass spectrometry detection, (b) fluorescent labelling with a methoxy group for fluorescence and MS detection, and (c) fluorine labelling for ^{19}F NMR detection with minimal natural background.*

2.1.1 Isotopic labelling

Labelling with stable isotopes was used to enable tracing using mass spectrometry. Isotopic labelling of amino acids (with deuterium (^{2}H) or nitrogen (^{15}N)) allows clear differentiation from naturally occurring amino acids due to a shift in the mass and isotope pattern. Applying isotopically labelled amino acids also offers the potential

for detecting metabolites of foliar-applied amino acids that share the same distinctive isotope pattern.

2.1.2 Fluorescent labelling

Fluorescent labelling, using a fluorescently active methoxy-derivatised indole, potentially enables tracing of the labelled compound by fluorescence and mass spectrometry. The additional functional group again shifts the mass of the amino acid; however, as methoxy tryptophan can be present in nature as a metabolite of tryptophan, this is not a unique molecule, making it more challenging to see in the plant.

2.1.3 Fluorine labelling

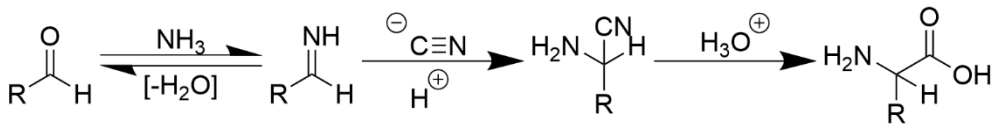
Fluorine labelled amino acids can be detected with good sensitivity by fluorine nuclear magnetic resonance (^{19}F NMR), as this is the only stable isotope of fluorine. Fluorine labelling allows for minimal natural background noise, as fluorine is not commonly found in nature. Therefore, observing only the exogenously applied material without other compounds being mistakenly identified is easy. As the carbon-fluorine bond is a polar covalent bond, it is unlikely to be lost from the compound easily within the plant. Fluorine's small size reduces the chance of unwanted additional biological activity within the plant.

2.2 Synthetic routes

Synthetic routes for labelled amino acids have been widely explored, ranging from classical chemical syntheses to modern biocatalytic approaches. Each method offers different advantages in terms of cost, scalability, environmental impact, and suitability for isotopic labelling. For this project, the ideal route needed to balance low cost and risk with environmental considerations, while being adaptable to both milligram-scale trials and gram-scale preparations.

2.2.1 Strecker synthesis

The Strecker synthesis, named after Adolph Strecker, is a well-known classical method used for preparation of α-aminonitriles from aldehydes or ketones, which are used as an intermediate to produce amino acids (Scheme 3).[130]



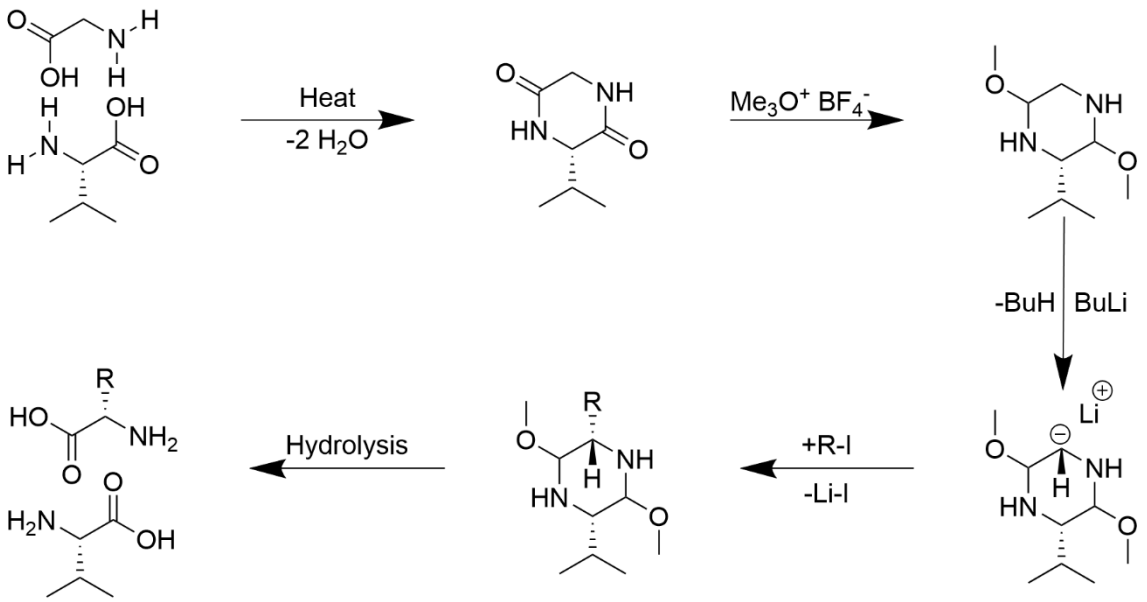
Scheme 3: *Amino acid synthesis using the Strecker method. The reaction proceeds through imine formation from an aldehyde and ammonia, followed by nucleophilic addition of cyanide to yield an α-aminonitrile, which is then hydrolysed with aqueous acid (H_3O^+) to produce the corresponding α-amino acid.*

This synthesis involves the formation of an imine from an aldehyde or ketone, followed by a cyanide addition, and then acid hydrolysis to give the amino acid. It can produce a broad range of amino acids and uses relatively inexpensive reagents. The Strecker method produces a racemic mixture of the L- and D- amino acids which is undesirable. However, more recently, the Strecker synthesis has been adapted to use catalysts, such as the metal complex Al^{III}-N,N'-bis(salicylidene)-ethylenediamine dianion, to produce asymmetric α-amino acids.[131]

Due to the use of hydrogen cyanide, which is highly toxic, and the poor enantioselectivity of this method without the use of expensive catalysts, Strecker synthesis is not a suitable method for this project.

2.2.2 Schöllkopf bis-lactim ether synthesis

In the Schöllkopf bis-lactim ether synthesis method for amino acid synthesis, chiral bis-lactim ethers derived from diketopiperazines and alkylating agents are used to synthesise enantiomerically pure amino acids (Scheme 4).[132]



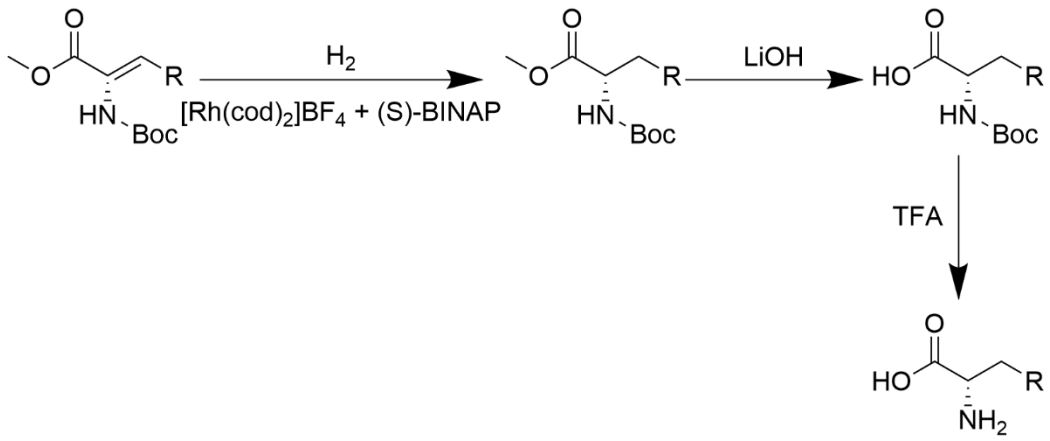
Scheme 4: *Overview of the Schöllkopf bis-lactim ether synthesis for enantiopure amino acids. Reaction steps include cyclisation, methylation, alkylation, and hydrolysis.*This method involves the forming of a diketopiperazine from two amino acids (glycine and valine), followed by methylation to make a bis-lactim. This is then alkylated using butyl lithium to deprotonate and the desired side chain added. Acid is then used to promote hydrolysis of the ring to give a chiral auxiliary and the desired amino acid. The Schöllkopf bis-lactim ether synthesis method can, therefore, produce high purity enantiopure amino acids of varying types and with minimal purification required.

However, diketopiperazine synthesis can be expensive due to the need for high-purity amino acids to start with and the very controlled conditions required for the cyclisation. There are also added steps as the chiral auxiliary must be removed at the end and is not regenerated.

2.2.3 Enantioselective hydrogenation

This method uses metal catalysts such as rhodium or ruthenium to hydrogenate dehydroamino acids to form α-amino acids with high enantiomeric purity. Chiral ligands, such as 2,2'-Bis(diphenylphosphino)-1,1'-binaphthyl (BINAP) coordinate with the metal catalyst to control stereoselectivity. This method is very versatile, having been used to synthesise a range of natural and non-natural α-amino acids.[133]

This route (Scheme 5) begins with an N-protected α,β-dehydroamino acid methyl ester, such as Cbz-, Boc- or benzyl ether-dehydrotryptophan methyl ester. In the major step, hydrogenation is carried out in the presence of a rhodium or ruthenium catalyst ligated with BINAP or a related chiral phosphine ligand. Coordination of the dehydroamino acid ester to the complex generates two possible diastereomeric intermediates. The thermodynamically favoured major diastereomer binds more tightly but undergoes hydrogenation more slowly, while the sterically hindered minor diastereomer binds less strongly but hydrogenates more rapidly. As a result, the faster-reacting minor diastereomer dictates the stereochemical outcome, leading predominantly to one enantiomer of the amino acid. The product is an N-protected amino acid ester, which can then be deprotected and hydrolysed to afford the free amino acid in enantiopure form.



Scheme 5: *Enantioselective synthesis of an L-amino acid via hydrogenation of an N-protected α,β-dehydroamino acid methyl ester using a chiral BINAP-ligated rhodium or ruthenium catalyst.*[134]

Although the versatility and stereoselectivity of this method make it a strong contender for use in this project, it does have several drawbacks. The cost of using these precious metals is high, and synthesis of the necessary ligands can be complicated. Additionally, hydrogenation reactions require very specialist equipment and present safety risks. Therefore, overall, this method is less favourable for the project than other more accessible synthetic routes.

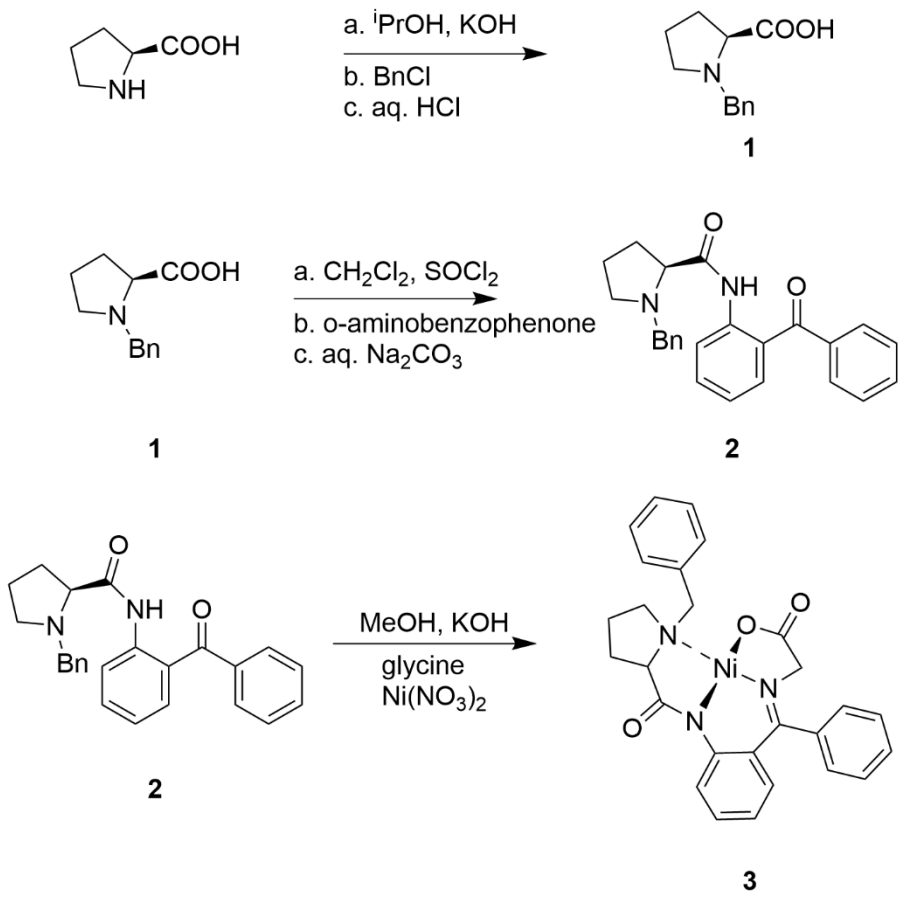
2.2.4 Tryptophan synthase

Tryptophan synthase works via a two-step reaction involving indole-3-glycerol phosphate and L-serine.[129] This method limits the positions and groups which can be used in the tryptophan analogues due to larger indoles being unable to enter the enzyme's active site. This method allows the synthesis of many different tryptophan analogues but is limited to tryptophan derivatives rather than being able to produce many amino acid analogues by the same method as the starting material for the reaction is an indole. More enzymes would have to be explored to make a range of amino acids, increasing the cost of synthesis and labelling may become more difficult.

2.2.5 Schiff base

The Schiff base method first proposed by Belokon *et al*. in 1988 is effective in achieving chiral purity in L-amino acids, thanks to the steric hindrance provided by the Nickel-complexed ligand.[135] By forming a complex around a Ni(II) ion only one face of the glycine moiety is available for attack at the glycine moiety so only the L amino acid is produced.[136] This method is versatile and suitable for synthesising a wide variety of amino acids, including both proteinogenic and non-proteinogenic amino acids.

The nickel(II)-Schiff base complex can be synthesised in three steps beginning from L-proline (Scheme 6). The reaction conditions are relatively mild, making it safer to use this method than those with more toxic reagents, such as Strecker synthesis. In this method, the formation of stable intermediates (Schiff bases) typically reduces the likelihood of side reactions, leading to cleaner reactions with fewer by-products.



Scheme 6: *Synthetic route for the formation of a nickel(II) Schiff base complex with glycine. The multistep sequence includes N-benzylation, acyl chloride formation, Schiff base condensation, and metal complexation with* Ni^{2+} *to yield the tetradentate ligand system.*

Another advantage of this technique is that the (S)-2-[N-(N'-benzyl prolyl)amino] benzophenone (BPB) ligand is recoverable at the end of the process. This means that the ligand can be recycled for the production of multiple amino acid species. This method was also suitable for the milligram scale required of the project, with compounds being applied to plants at a maximum concentration of 60 mM.

2.2.6 Summary of synthesis techniques

Several potential synthetic routes were considered for the synthesis of labelled tryptophan. Strecker synthesis lacked enantioselectivity and required toxic reagents so was deemed unsuitable. Schöllkopf bis-lactim ether synthesis was not used due to its complexity and the cost of starting material being very high. Despite the good stereoselectivity for enantioselective hydrogenation, high cost and safety risks made

this method less suitable for this project. Enzyme-mediated synthesis did not have the desired versatility to be applied to other amino acids without increasing costs considerably. Schiff base synthesis, using Belokon's methods proved the most suitable for the project as it was versatile, relatively safe, and scalable.

2.3 Synthesis of tryptophan analogues

Tryptophan (Trp) formed the main target for synthetic labelling due to its pivotal role in plant metabolism and its potential to act as a tracer for metabolic pathways. Labelling Trp enables the investigation of its uptake and distribution within the plant. The selection of Trp was also influenced by its structural versatility, which allowed different labelling techniques to be employed. Building on the methods of Belokon, this study aimed to adapt the synthetic strategy to produce isotopically and structurally modified derivatives suitable for analytical studies within plants.[137]

2.3.1 Tryptophan synthesis via the Belokon Schiff base

The synthesis of 5-L-methoxy tryptophan and 5-fluoro tryptophan was based on the general method for the synthesis of amino acids by Belokon *et al.* (1998) through the reaction of a trimethylammonium iodide with a chiral Schiff base of glycine, using modifications used by Xu *et al*. (2017).[135-139]

Synthesis of the nickel Schiff base complex used for amino acid synthesis followed a three-step pathway. The first of these steps was the S_N2 addition of a benzyl group to the secondary amine of L-proline using benzyl chloride to form benzyl proline (BP, compound **1**). The initial protocol followed for this called for the reagents to be stirred with potassium hydroxide in propan-2-ol and heated at 40 °C for four hours followed by the addition of chloroform and overnight stirring. At this stage, a precipitate was expected to form, with further benzyl proline recovered upon addition of acetone. However, very little product was obtained, as the benzyl proline proved difficult to precipitate. Attempts to follow the same method but using liquid–liquid extraction with dichloromethane (DCM) also failed to produce the desired product, and NMR suggested that benzylation was occurring at both the amine and carboxylic acid sites. Some product formation was achieved when the reaction was cooled below room

temperature in an ice bath, although yields remained around 20%. Using a different base, sodium methoxide, and concentrating the filtrate before recrystallising the product from methanol and diethyl ether significantly improved both yield and reproducibility, with the product obtained more consistently and at approximately 70% yield.

The second step involved activation of the carboxylic acid of BP through formation of the corresponding acyl chloride using thionyl chloride at low temperature, followed by condensation with 2-aminobenzophenone to form the ligand, (S)-2-[N-(N'-benzylprolyl)amino]benzophenone (BPB, compound **2**). Initially, the crude product was purified using normal-phase column chromatography (petroleum ether: ether), which resulted in a low yield, as some of the product remained in the column, due to strong interactions with the silica. To improve this, the purification step was changed to a recrystallisation from ethanol. This avoided losses associated with silica and solvent elution and allowed the product to crystallise directly from solution, giving a cleaner solid and substantially improving the isolated yield.

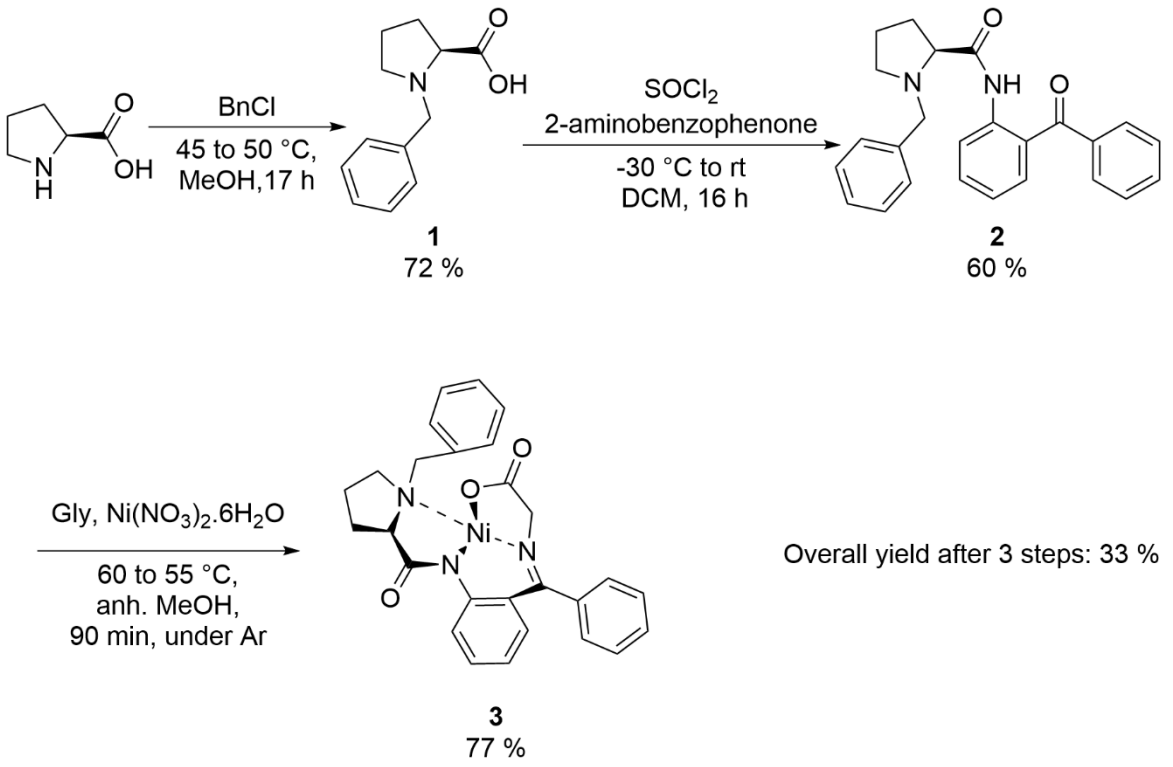
BPB was reacted with glycine and nickel(II) nitrate to form the Schiff base complex **3**. Initial complexation using 5 equivalents of glycine and 2 equivalents of $Ni(NO_3)_2 \cdot 6H_2O$ (Table 3, exp. A), following the method of Belokon, resulted in low conversion, with unreacted BPB still detectable by mass spectrometry and NMR. To optimise the reaction, several reagent ratios were trialled (Table 3), and the best result was obtained using 1.2 equivalents of nickel nitrate. For purification, a silica plug was employed: chloroform was used to elute residual BPB, followed by a chloroform: methanol (9:1) system to elute the nickel complex.

Table 3: *Reaction conditions and isolated yields for synthesis of compound **3**. Yields were calculated based on the mass of purified product relative to recovered starting material following silica plug purification.*

Experiment Number	Equivalents of Glycine	Equivalents of $Ni(NO_3)_2.H2O$	Equivalents of KOH	Yield,%
A	5	2	7	42

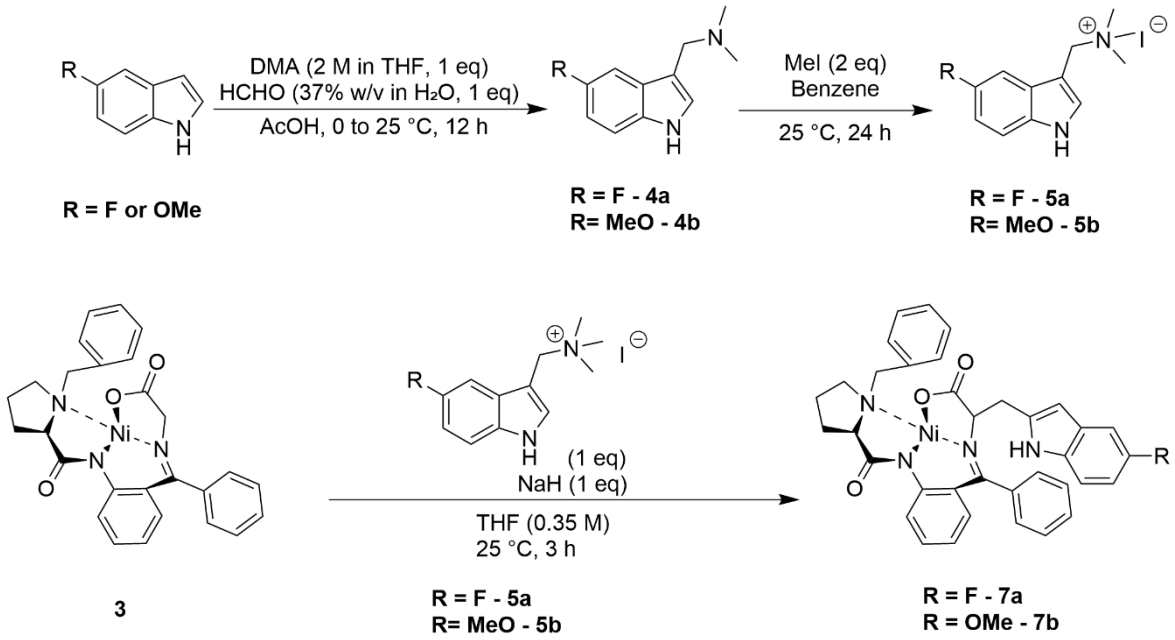
B	10	4	7	38
C	2	1.2	15 (of NaOMe)	77

These optimisations highlighted the need to adjust both reagent stoichiometry and purification methods to achieve a clean conversion to the Ni-Schiff base complex. Following these improvements, the synthesis proceeded reproducibly, with significantly higher yields than obtained following previously reported benzylation and BPB formation steps. A summary of the optimised synthetic route used in this work, including isolated yields for each step, is shown in Scheme 7.



Scheme 7: *Three-step synthesis of nickel(II) complex 3 from N-benzylproline. Overall yield: 33%.*

Once the synthesis of the nickel complex had been optimised, it was then used to make 5-fluoro and 5-methoxy tryptophan by the addition of a gramine side chain using a quaternary ammonium (Scheme 8).



Scheme 8: *Proposed synthetic route for functionalisation of nickel complex 3 with quaternary gramine derivatives to yield tryptophan analogues.*

Synthesis of the required gramines was adapted from the Pillaiyar *et al*. (2018) procedure for a Mannich reaction of indoles.[140] These were then used to synthesise 1-(5-Methoxy-1H-indol-3-yl)-trimethylammonium iodide (**5b**) and 1-(5-Fluoro-1H-indol-3-yl)-trimethylammonium iodide (**5a**), based on the method by Belokon for the synthesis of L-Trp. A new methodology was required for this synthesis due to the observation of two peaks in the mass spectra of the final product, as discussed above. The presence of this secondary peak suggested that the method produced not only the desired monomer but also a dimerised product (**6a, 6b**). NMR data confirmed this, and through analysis of the protons at the C10 position, it could be seen that the two products were formed in a 1:1 ratio (Fig. 14). The presence of this dimerised compound was not mentioned in the paper from which the method was obtained.[140] The integral values show a doubling of expected proton counts in most regions, indicating the presence of both monomer and dimer species. The aliphatic region (3.0 - 3.5 ppm) contains a peak with an integral of 9, corresponding to the three CH_3 groups of the monomer, alongside a peak with an integral of 3. Although the dimer contains two CH_3 groups, the lower integral reflects its approximately half concentration relative to the monomer. This pattern, mirrored across the aromatic

and other signals, demonstrates that the same amount of starting material contributed to both products, confirming the absence of selectivity in the reaction.

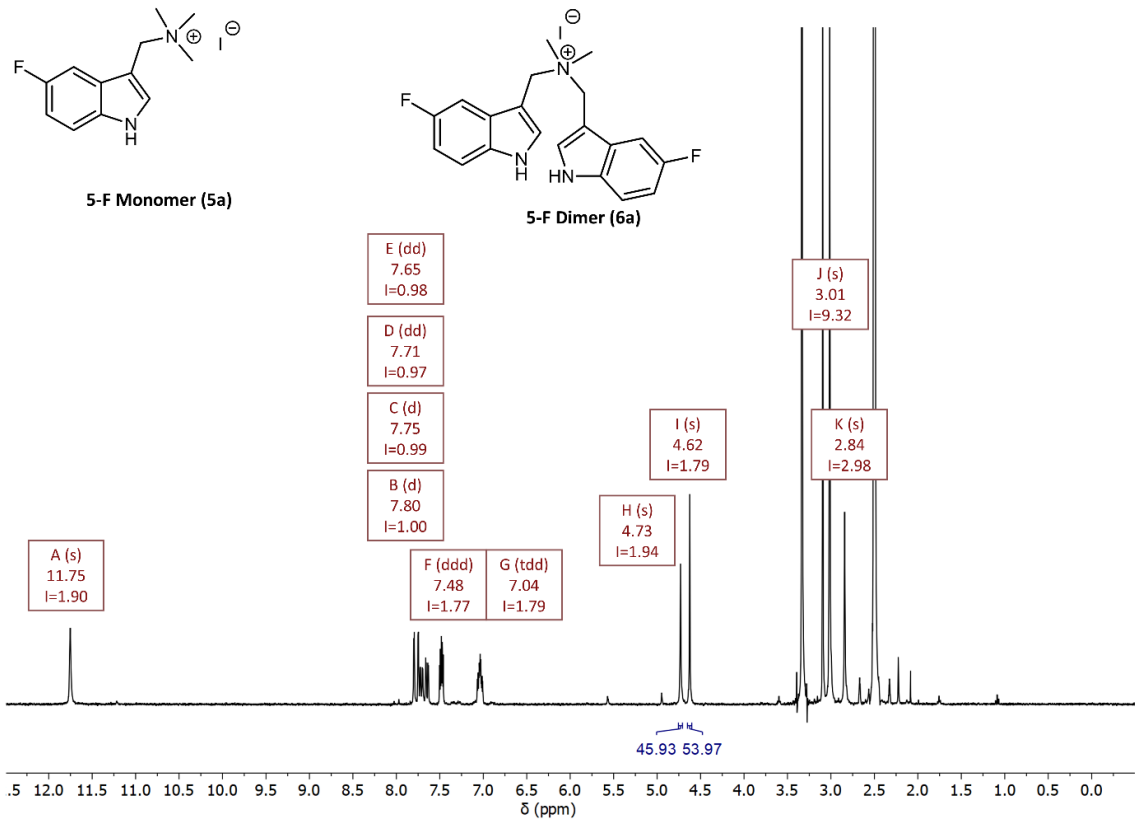


Figure 14: *Full ^{1}H NMR spectrum of the crude product from quaternisation of 5-fluorogramine, showing formation of both monomer and dimer species. Integration of the C10 proton signals indicates a 46:54 ratio, confirming equal contribution of starting material to each product despite differing molecular yields.*

Although the monomer contains two protons at the C10 position, and the dimer contains four protons in the equivalent environment, the integrals of these ^{1}H NMR signals can be directly compared. This is because the formation of each dimer molecule requires two units of the starting material, while only one unit is required to produce the monomer. Therefore, a 1:1 ratio of the integrals at the C10 position indicates that the same amount of starting material is contributing to both the monomer and the dimer even though the actual yield of dimer would be around half that of the monomer. The structures show the C10 position in the monomer and dimer products, highlighting the similarity in proton environments. Although the environments are closely related, they produce distinct NMR signals that appear near each other in the spectrum. Comparing the integrals at the C10 position in the

monomer and the dimer allows for the calculation of the percentage or ratio of the starting material that contributed to each product, directly measuring the reaction's selectivity towards monomer or dimer formation.

The presence of both monomer and dimer products (confirmed by NMR and mass spectrometry) demonstrated a lack of selectivity within this reaction, an observation which was not mentioned in the original literature. These findings highlighted the need to explore alternative synthetic approaches to achieve selective monomer or dimer formation. The following section outlines the methodology developed to control product selectivity by optimising reaction conditions and starting material modifications.

2.3.1.1 Methodology for quaternary ammonium monomer synthesis

After the identification of this issue with the methodology, an investigation into the optimisation of the reaction was conducted.

A short literature search for synthesis of both the monomer and dimer forms of the trimethyl ammonium methoxy indole showed five different conditions for the same reaction. In the literature, only one paper, from Lown and Weir (1978), discussed the production of the dimer when using the established method of quaternary amine synthesis and proposed better conditions for producing just the monomer.[141] However, no method for selectively making the dimer was found. The five methods, plus two more based on these methods, were trialled using 5-F gramine (synthesised from 5-F-indole), as 5-F-L-Trp was the next chosen target for synthesis (Table 4).

Table 4: *Reaction conditions trialled for the synthesis of monomeric quaternary ammonium salts from gramine derivatives, including crude yields and monomer-to-dimer ratios as determined by ^{1}H NMR integration.*

Exp. Number	Conditions	Yield (%)	Ratio monomer: dimer
1	100 eq methyl iodide, 2 h, rt	86%	87:13
2[140]	2 eq methyl iodide, benzene, 24h, rt	82%	50:50

3[142]	1.5 eq methyl iodide, THF, 1 h, 0 °C	59%	50:50
4	15 eq methyl iodide, EtOH,4 h, rt	70%	50:50
5[143]	1 eq methyl iodide, EtOH, 1 h, 0 °C	84%	13:87

Selectivity was determined using ^{1}H NMR comparison of the C10 proton environment (Fig. 15). Among the conditions trialled, the use of excess methyl iodide during quaternisation favoured monomer formation (Table 4, exp. 1), while reducing the equivalents of methyl iodide to 1 eq promoted dimer formation (Table 4, exp. 5).

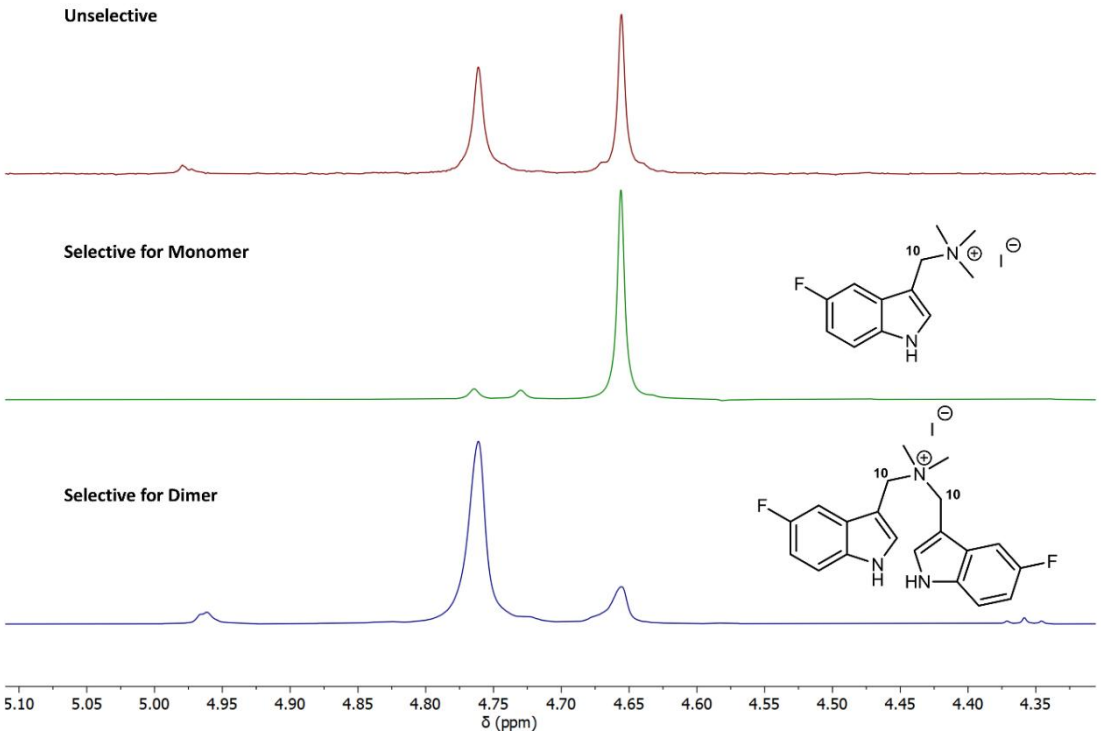
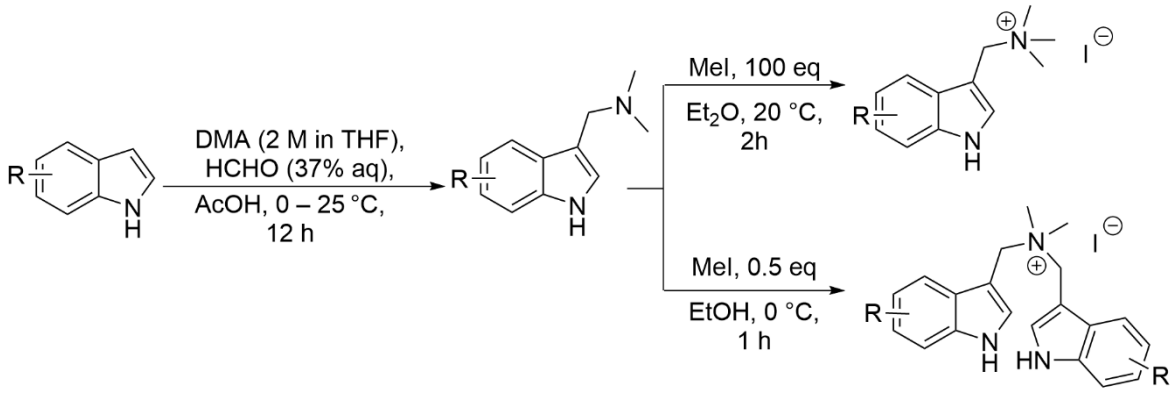


Figure 15: *Expanded view of the C10 region in the ^{1}H NMR spectra for reactions with 5-fluoro gramine (4a) under three conditions. Spectra show unselective (top), monomer-selective (middle), and dimer-selective (bottom) reactions. Differences in signal intensity reflect the influence of reaction conditions on product selectivity.*

The identified conditions for the selective synthesis of the monomer were optimised further by using excess diethyl ether as the monomer crashed out of solution in this solvent upon formation, reducing the formation of the dimer and for the conditions of

the dimer reaction, reducing the equivalents of methyl iodide further to 0.5 eq produced greater selectivity favouring the dimer.

These optimised methods provided a platform to further investigate the scope and potential applications of isolated monomer and dimer analogues (Scheme 9). The ability to selectively synthesise monomers bearing different ring substituents is particularly valuable, as these compounds can serve as precursors to diverse tryptophan analogues and enable the introduction of functional groups with potential roles as spectroscopic markers or interaction sites within proteins. In contrast, many of the corresponding dimers have not been previously reported, and earlier studies demonstrated that 4-substituted dimers in particular are challenging to obtain with high selectivity over the monomer and in synthetically useful yields.[144]



Scheme 9: *General synthesis for the preparation of quaternary ammonium salts, selectively yielding either monomeric or dimeric products, from substituted indoles via gramine intermediates. Substituents (R) at positions 4, 5, or 6 of the indole ring include H, halogens (F, Cl, Br, I), nitro (NO_2), methoxy (OMe), and methyl (Me).*

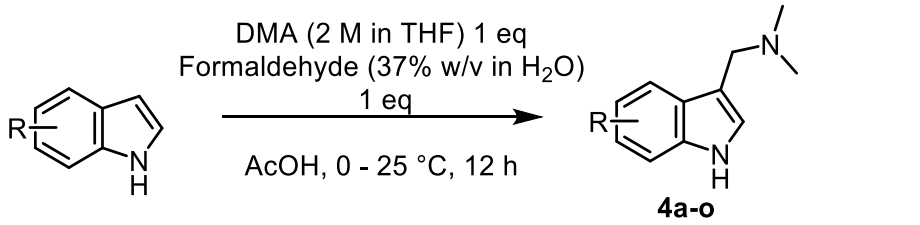
2.3.1.1.a Scope of application

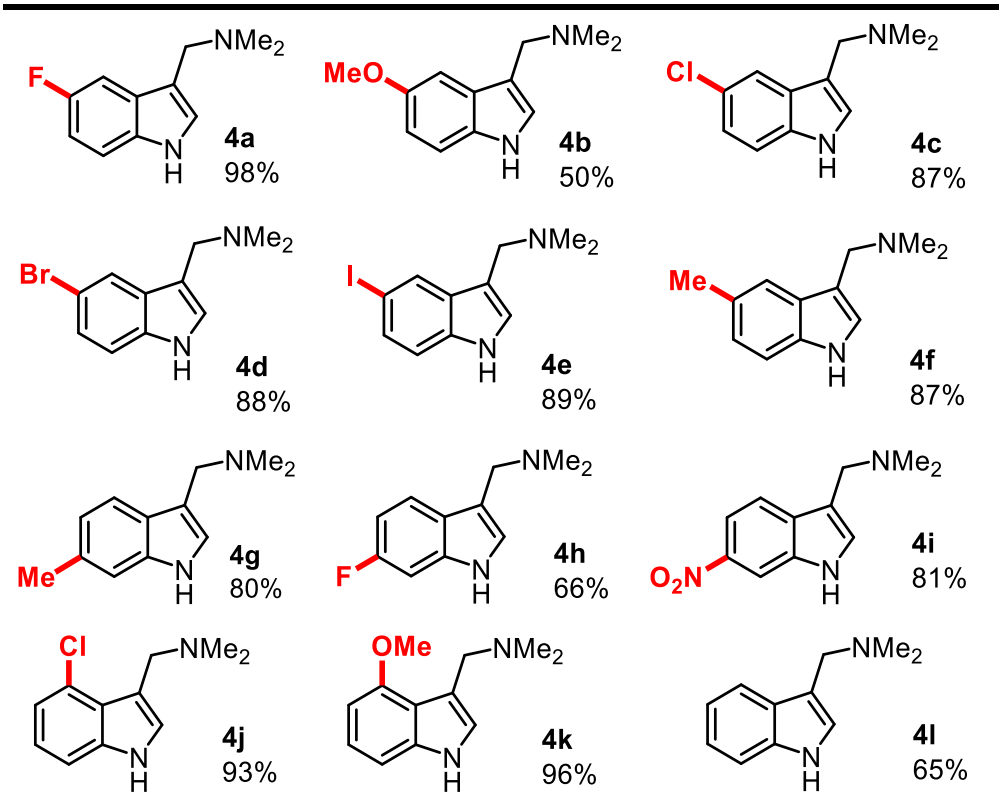
With the method optimised for 5-F gramine a broader scope of substrates was proposed using indoles with various functional groups at the 4-, 5-, and 6- positions on the benzyl ring. The aim was to determine whether the monomer (Table 6) and the dimer (Table 7) could be selectively synthesised under the optimised conditions. The substrates chosen included halogens (F, Cl, Br, I), methoxy, methyl and nitro groups as well as unsubstituted indole. These functional groups have diverse electronic effects, some acting as electron donors and others as electron-withdrawing groups. They have varied steric bulk ranging from proton (unsubstituted) to bulkier nitro

groups. Additionally, most of these functional groups have biological relevance, making them good potential candidates for future bioactivity studies.

To begin the set of 12 indoles were converted into their gramine derivatives (Table 5), which were the starting materials for the monomer and dimer synthesis.

Table 5: *Synthesis of gramine derivatives bearing functional groups at the 4-, 5-, or 6-position of the indole ring. Functional groups are highlighted in red. Compound identifiers are shown in bold; percentage yields are indicated below each structure. The general reaction scheme is shown above.*

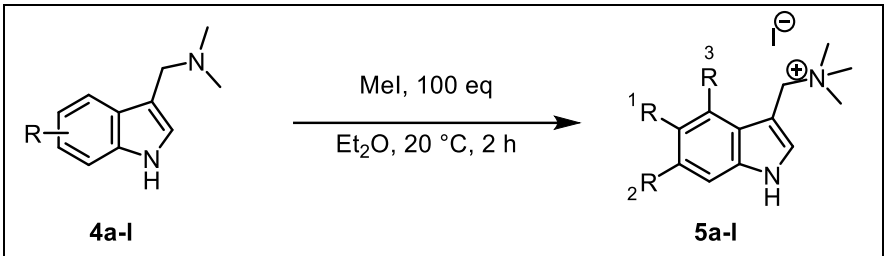




These gramine analogues had mostly high yields with an average of 82%. The substituted gramines were then used for the synthesis of monomers using the optimised conditions. Table 6 presents the overall yield and selectivity of the monomer formation from gramine derivatives under methylation conditions (MeI,

Et_2O, 20 °C, 2 h). Selectivity was determined by NMR, showing the percentage of monomer and minor dimer formation for each functionalised substrate. The synthesis of the monomers mostly had high yields, between 63 - 93%. Selectivity for the monomer over the dimer under these reaction conditions was also mostly very high, with many having less than 10% of the starting material going into the synthesis of the dimer.

Table 6: *Crude yields and monomer-to-dimer ratios for the synthesis of quaternary ammonium salts from substituted gramine derivatives using the monomer-selective methylation conditions. Substituents (R) vary at positions 4, 5, or 6 of the indole ring. Product ratios were determined by ^1H NMR integration.*



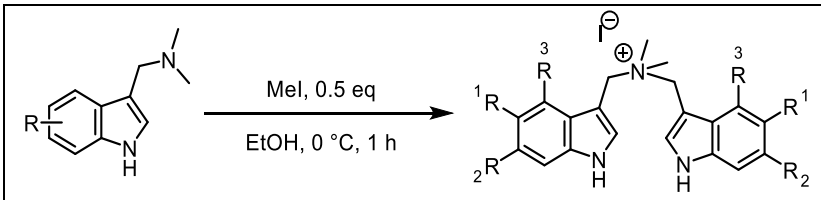
4a-l **5a-l**

	Overall Yield	Percentage monomer (by NMR)	Percentage dimer (By NMR)
(5a) R_1 = F	86%	92	8
(5b) R_1 = OMe	63%	99	1
(5c) R_1 = Cl	79%	89	11
(5d) R_1 = Br	80%	90	10
(5e) R_1 = I	86%	94	6
(5f) R_1 = Me	83%	99	1
(5g) R_2 = Me	83%	95	5
(5h) R_2 = F	86%	94	6

(5i) $R_2 = NO_2$	79%	81	19
(5j) $R_3 = Cl$	68%	99	1
(5k) $R_3 = OMe$	76%	80	20
(5l) $R_{1-3} = H$	93%	99	1

The same set of 12 gramines was then used with the optimised conditions for dimer synthesis. Table 7 presents the overall yield and selectivity of the dimer formation from gramine derivatives under controlled methylation conditions (MeI, EtOH, 0 °C, 1 h). Selectivity was again determined by NMR, showing the percentage of dimer as the main product and (mostly) minor monomer formation for each functionalised substrate.

Table 7: *Crude yields and dimer-to-monomer ratios for the synthesis of quaternary ammonium salts from substituted gramine derivatives using the dimer-selective methylation conditions. Substituents (R) vary at positions 4, 5, or 6 of the indole ring. Product ratios were determined by ^1H NMR integration.*



4a-l **6a-l**

	Overall Yield	Percentage dimer (by NMR)	Percentage Monomer (By NMR)
(6a) $R_1 = F$	63%	92	8
(6b) $R_1 = OMe$	52%	96	4
(6c) $R_1 = Cl$	76%	91	9

(6d) R_1 = Br	29%	77	23
(6e) R_1 = I	92%	56	44
(6f) R_1 = Me	62%	86	14
(6g) R_2 = Me	87%	82	18
(6h) R_2 = F	27%	89	11
(6i) R_2 = NO_2	15%	91	9
(6j) R_3 = Cl	34%	95	5
(6k) R_3 = OMe	62%	99	1
(6l) R_{1-3} = H	40%	93	7

The monomer and dimer forms of each gramine could be synthesised with overall improved selectivity for the synthesis of only the monomer or only the dimer respectively. The diversity of functional groups allowed for exploring how steric and electronic effects influence yield and selectivity under these conditions.

The synthesis of the dimer proved more challenging, with significantly lower yields, and more difficulty in achieving the desired selectivity (Table 7). Some compounds with higher yields had worse selectivity and some with good selectivity had lower yields. One notable trend observed in the dimer molecules was a decrease in selectivity from fluorine to iodine. The 5-Fluoro dimer reaction showed 93% selectivity for synthesis of the dimer, while the same reaction conditions using the 5-Iodo gramine exhibited no significant selectivity for the dimer, with about half the starting material going towards the monomer and half to the dimer. This trend suggests that steric and electronic effects could be contributing factors in dimer formation.

One potential explanation for this trend could be that at the 5-position on the indole ring, the fluorine group induces an electron-withdrawing effect within the molecule.[145, 146] This would slightly reduce the electron density at the tertiary amine, making it less nucleophilic and slowing its reaction with methyl iodide. This reduced reactivity could allow more time for the dimerisation to occur, therefore increasing selectivity. Moving down the halogen group, the elements exhibit less electronegativity, and so this effect would be reduced. However, electronic effects alone cannot fully explain the observed selectivity differences. This trend is not reflected in monomer formation, where all halogenated gramines exhibited similarly high selectivity. This suggests that additional factors, beyond simple steric and electronic effects are at play.

An explanation for the lack of this trend in the monomer reaction could be that in the monomer reaction, the product crashes out of solution upon formation, effectively removing it from the equilibrium and driving the reaction towards continued monomer synthesis. Due to this, steric and electronic effects play a less significant role in selectivity compared to the dimer reaction, where the product remains in solution until the workup.

Additionally, steric effects may play a role, as iodine is significantly larger than fluorine.[147] The increased bulk could disrupt the close π - π stacking interactions required for stable dimer formation, further reducing selectivity.

While steric and electronic effects provide one possible explanation for the observed trend, the underlying reason is likely to be more complex. Intermolecular interactions may significantly influence dimer selectivity, including π - π stacking, hydrogen bonding with solvent molecules, and dipole interactions. Additionally, differences in reaction kinetics could play a role, where variations in nucleophilicity and steric accessibility affect the rate at which monomer and dimer formation occur. To fully understand these trends, further studies, such as computational modelling of molecular conformations and studying of halogens at other ring positions, would be necessary to isolate the key driving forces behind dimer selectivity in halogen-substituted gramines.

Synthesis of four position dimers proved more difficult with lower yield and decreased selectivity with some functional groups found. These results align well with previously reported work by Melhado (1981), who found that the substituent position significantly influenced product distribution in the reaction of methyl iodide with nitro-substituted gramines.[144]

2.3.1.1.b Crystal structures

Crystal structures were successfully obtained providing proof of the proposed structures. A subset of these crystal structures, including four gramines (Fig. 16), one monomer (Fig. 17 A) and two dimers are presented below (Fig. 17 B and C).

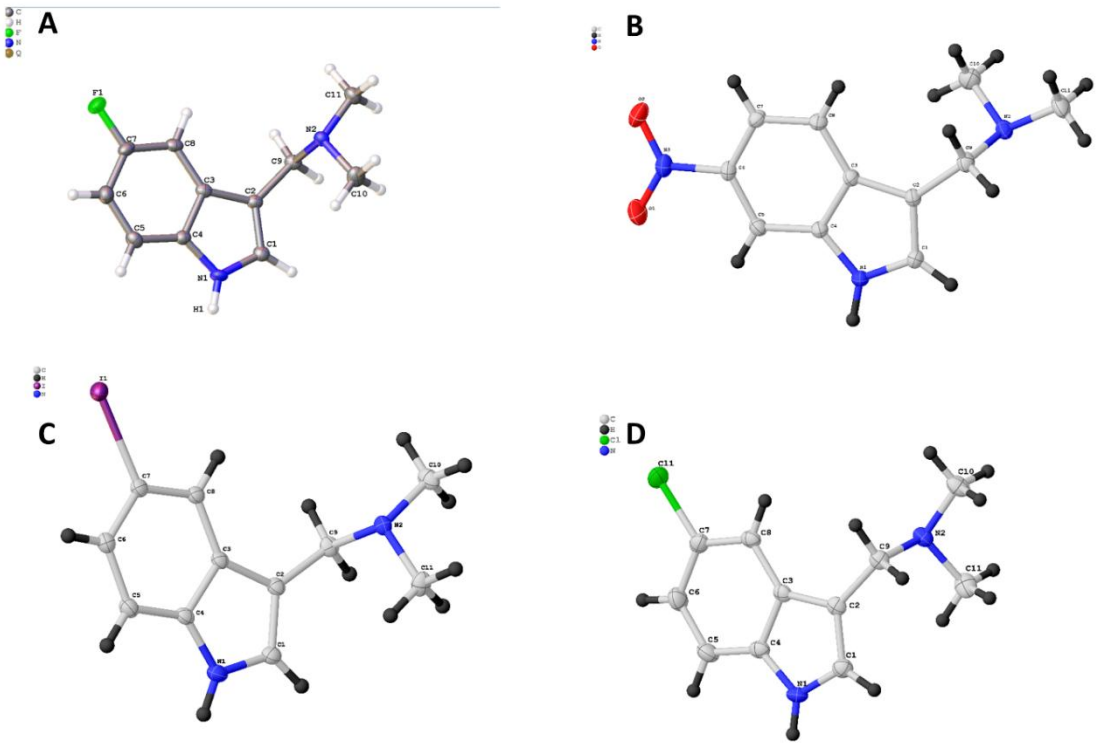


Figure 16: *X-ray crystal structures shown as ball-and-stick models for selected substituted gramine derivatives: (a) 5-fluoro gramine, (b) 6-nitro gramine, (c) 5-iodo gramine, and (d) 5-chloro gramine. Atom colours: carbon (grey), hydrogen (white), nitrogen (blue), oxygen (red), chlorine (dark green), fluorine (light green), iodine (purple).*

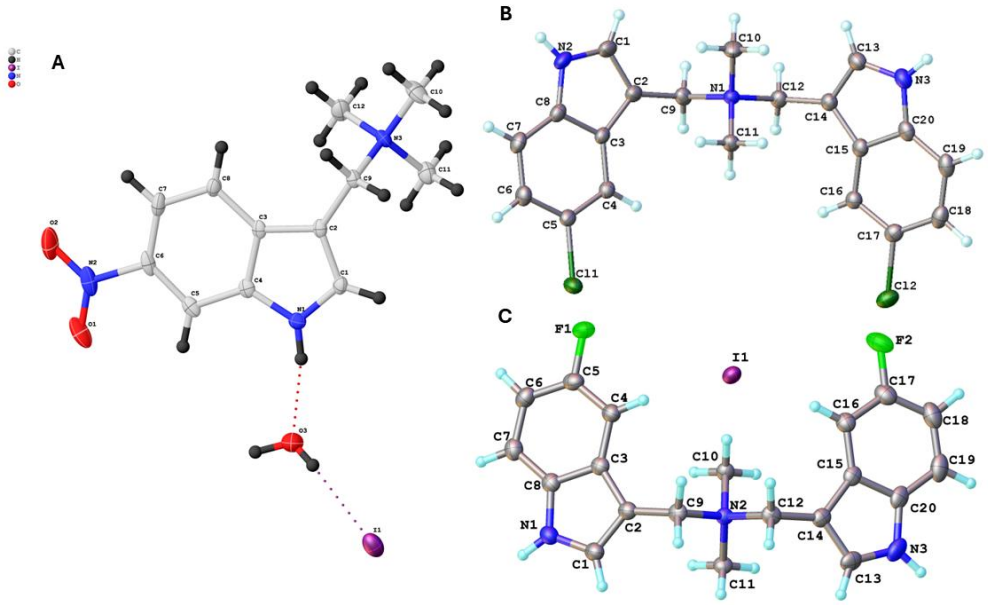


Figure 17: *Ball-and-stick representations of crystal structures for selected quaternary ammonium salts: (a) 6-nitro monomer, showing coordination with water; (b) 5-chloro dimer; and (c) 5-fluoro dimer. Atom colours: carbon (grey), hydrogen (white), nitrogen (blue), oxygen (red), chlorine (green), fluorine (light green).*

These crystal structures demonstrate that the proposed structures are seen with various functional groups attached at different positions on the ring. The packing of the crystals for dimer molecules within a unit cell (Fig. 18) shows how they assemble in the solid state, which may help explain the factors influencing their formation and selectivity. Packing diagrams show molecular arrangement within the unit cell for both dimers, highlighting key differences in intermolecular interactions.

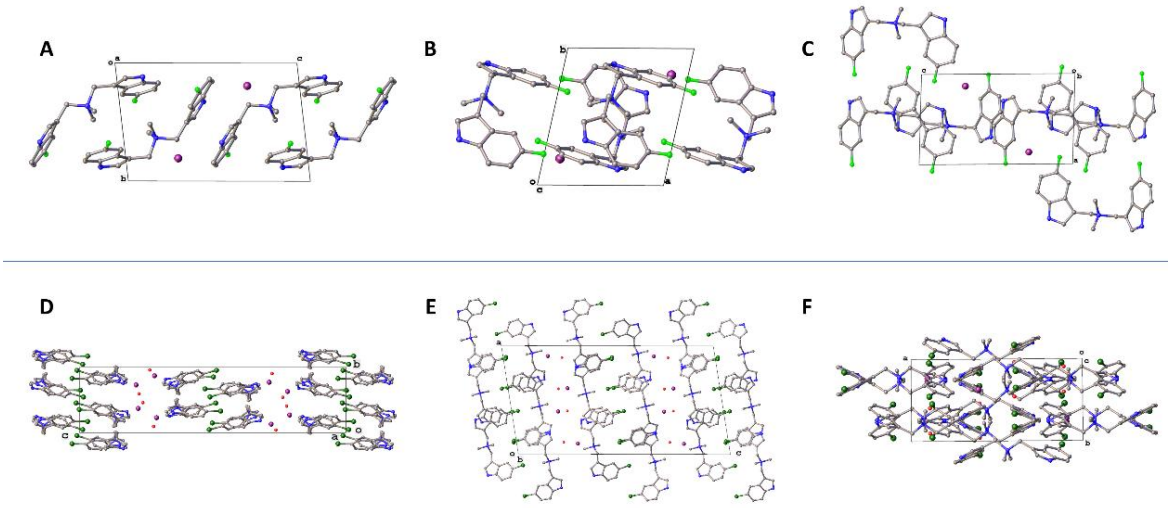
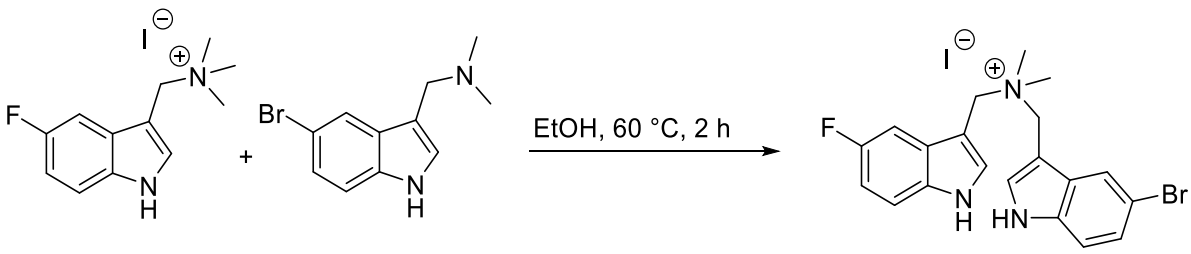


Figure 18: *Comparison of crystal packing in the 5-fluoro (top, panels A-C) and 5-chloro (bottom, panels D-F) quaternary ammonium dimers. Packing arrangements highlight differences in intermolecular interactions and aggregation motifs between the two derivatives.*The 5-fluoro dimer crystals (Fig. 18 A - C) are not as tightly packed as the 5-chloro dimer crystals (Fig. 18 D - F). The 5-chloro dimer is a monohydrate, coordinating with a single water molecule. The indole structure is key in both structures for intermolecular packing. Both indole NH groups hydrogen bond to the iodide but only one of them does this in the chlorinated compound. The other NH group favours hydrogen bonding with water instead. The differences in crystal packing between the 5-fluoro and 5-chloro dimers highlight potential structural factors that may influence their selectivity of formation. Further crystallographic research is needed to explore these packing trends and their relationship to dimerisation selectivity.

2.3.1.1.c Expansion for application - Heterodimers for Signalling

The concept for the heterodimer work originated from the suggestion that one ring could have a functional group capable of binding within a biological system while the other ring could carry a functional group designed to provide a detectable signal. This dual functionality offers the potential to develop a biomarker capable of both interacting with biological targets and facilitating their identification or monitoring. Searching the literature identified no methods for producing this, so method development was required. A method for synthesising a gramine dimer using one equivalent of monomer and one equivalent of gramine was found, suggesting that

stirring this starting material in ethanol would cause the product to form and crystallise out of the solution.[148] By using a monomer and gramine with different functional groups attached to the benzyl ring, it was proposed that this method could be used to make a heterosubstituted dimer. The original paper showed the formation of the dimer to be an equilibrium, so to force the reaction towards the dimer, stirring the starting material at either 0 °C or 60 °C was tried. Both methods formed dimer (as observed by mass spectrometry), but the heating method had greater yield and a stronger mass spectrometry signal for the desired product. The initial tests were done using 5-F monomer and 5-Br gramine to make a 5-F,5'-Br dimer (Scheme 10). However, while the desired product was observed in mass spectrometry, the homosubstituted 5-F dimer and 5-Br dimer were also observed. Attempts to isolate the desired product using high-performance liquid chromatography (HPLC) were unsuccessful.

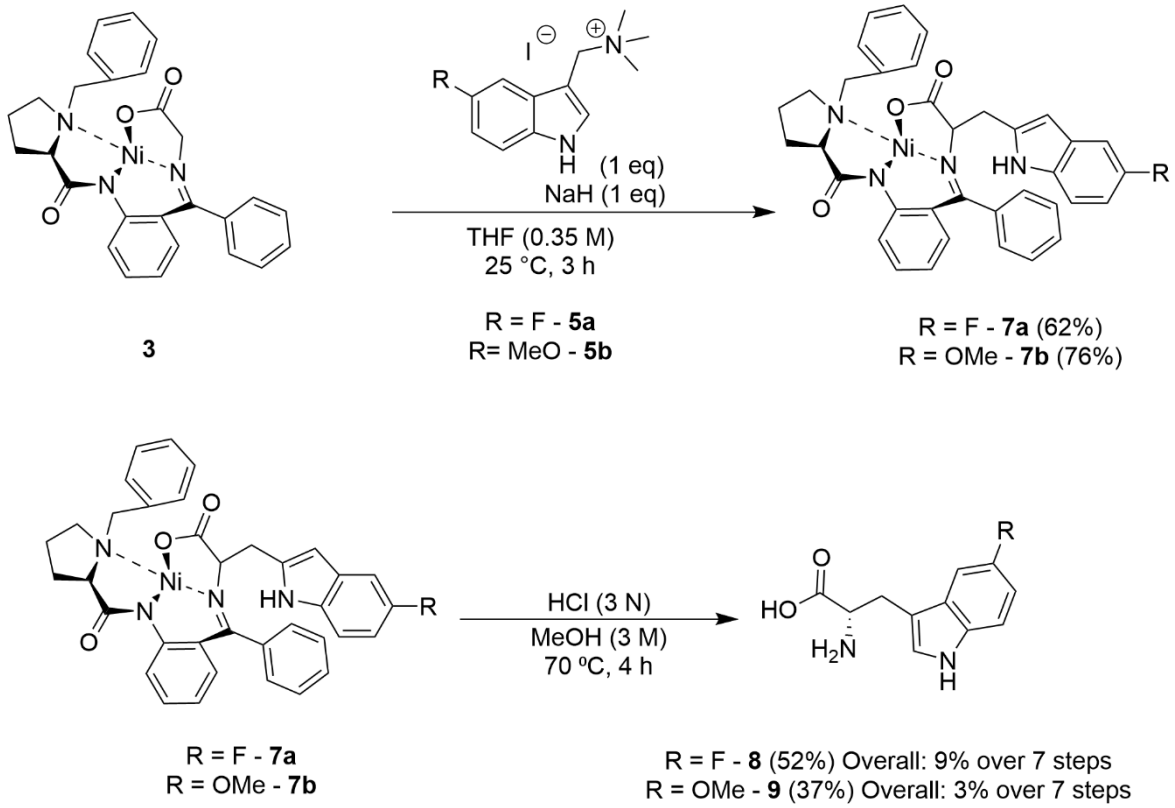


Scheme 10: *Synthesis of a heterosubstituted quaternary ammonium indole dimer using a novel method developed for selective dimer formation. The reaction combines 5-fluoro monomer and 5-bromo gramine to yield a mixed dimer bearing distinct substituents on each indole ring.*

Although initial attempts demonstrated heterodimer formation, the significant presence of undesired homodimers highlights the need for further optimisation of reaction conditions and purification methods.

2.3.1.2 Isolation of the amino acid

Following resolution of the side-chain substitution challenges, was applied to the 5-fluoro and 5-methoxy indole intermediates. This enabled successful installation of the amino acid backbone, yielding the corresponding tryptophan derivatives. The final products were obtained by acidic decomposition of the intermediate complexes (Scheme 11).



Scheme 11: *Addition of fluorine and methoxy-labelled indole side chain and isolation of 5-fluoro and 5-methoxy tryptophan derivatives from the nickel complex via acidic decomposition of intermediate complexes, using methods adapted from Belokon et al. (1988) and Xu et al. (2017).*[135, 139]

The resulting crude product was then purified using ion exchange resin (DOWEX® Resin 50X2) following the procedure of Georg Blaser (2007).[129] This proved a successful method, with yields consistently above 70% for 5-methoxytryptophan and 52% for 5-fluorotryptophan. Further purification was conducted via preparative HPLC.

2.3.1.2.a Confirming chirality

Chiral HPLC was used to confirm that the compounds synthesised using the Schiff base method had chiral purity, as the method hoped to afford. The initial method used a Chiralcel OJ column and a solvent system of 90% Hexane and 10% propan-2-ol. The retention time and peak profile of the synthesised compound (Fig. 19b) closely match those of the L-Trp standard (Fig. 19a), suggesting enantiomeric purity. However, small secondary peaks in both traces and three strong peaks in the racemic mixture (Fig. 19c) indicate poor separation. Therefore, while the available data

supports the identity of the synthesised compound as the L-isomer, further analysis with different chiral HPLC conditions would be required to confirm this assignment conclusively. A sample of 5-F-D-Trp was not available for use in these tests.

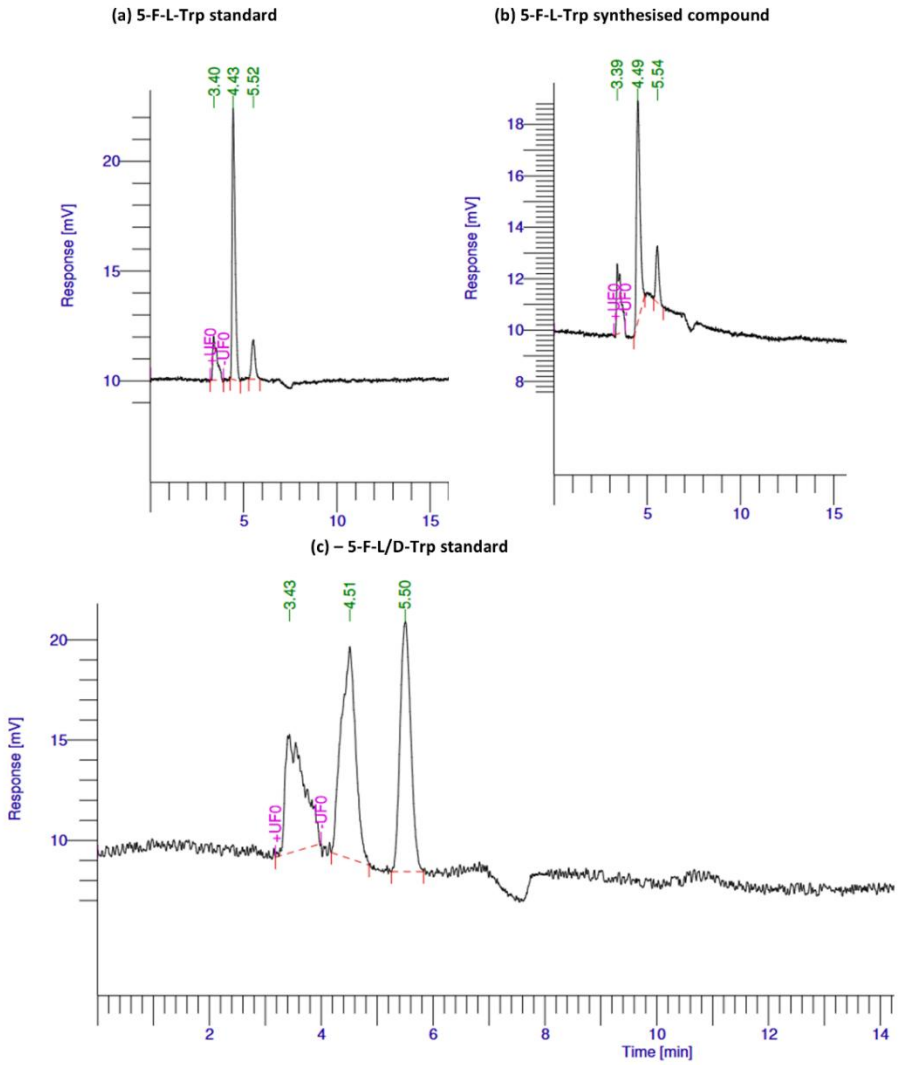


Figure 19: *Chiral HPLC traces for enantiomeric comparison of 5-fluorotryptophan (a) racemic 5-F-L/D-Trp standard, (b) 5-F-L-Trp standard, and (c) synthesised 5-F-L-Trp compound. Retention time alignment between (b) and (c) suggests the synthesised compound is the L-enantiomer. However, the presence of additional peaks and absence of a D-Trp standard preclude definitive assignment.*

The same conditions were used to confirm the chirality of synthesised 5-MeO-L-Trp, but this compound had much more severe separation issues with these conditions (Fig. 20). The separation achieved by the chiral HPLC method was insufficient to fully resolve the L- and D-isomers, as indicated by the overlapping peaks in the racemic standard (c). However, while the synthesised compound (d) closely resembles both

the racemic standard (c) and the 5-MeO-L-Trp standard (a), it displays a distinct chromatographic profile from the 5-MeO-D-Trp standard (b). While this suggests the synthesised compound is not the D-isomer, the lack of complete separation prevents a conclusive determination of enantiopurity.

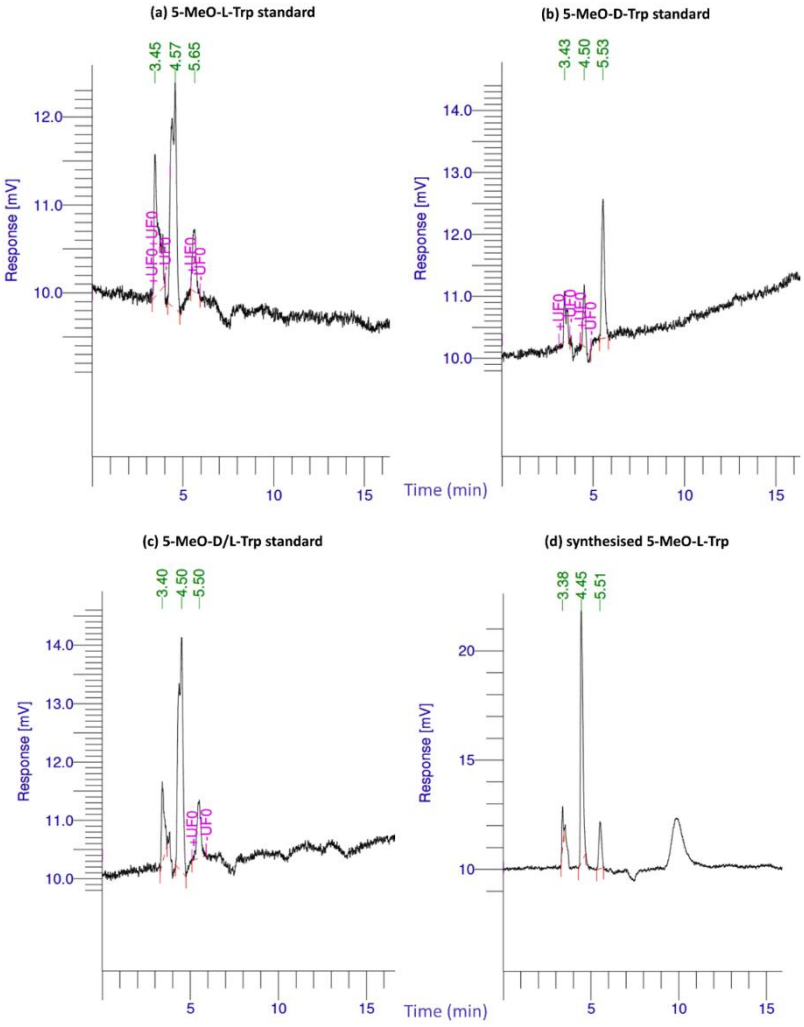


Figure 20: *Chiral HPLC traces for enantiomeric comparison of 5-methoxytryptophan. (a) 5-MeO-L-Trp standard, (b) 5-MeO-D-Trp standard, (c) racemic 5-MeO-D/L-Trp standard, and (d) synthesised 5-MeO-L-Trp compound. Overlapping peaks in the racemic standard indicate poor separation under these conditions. While the synthesised compound resembles the L-Trp standard and differs from the D-Trp trace, enantiopurity cannot be confirmed conclusively.*

A new method for chiral HPLC was proposed, based on the work by Hai *et al*. (2019), which described two sets of conditions for chiral separation. One method used for N-acetyl-D-tryptophan employed a CHIRALPAK® IA-3 column with a mobile phase of 80/20/0.1/0.1 hexanes/ethanol/TFA/DEA.[149] The second set of conditions, designed

for tryptophan derivatives including 5-MeO and 5-F analogues, employed a Crownpak® CR(+) column (150 mm x 4 mm x 3.5 µm, Daicel) with a mobile phase of 1% (w/v) aqueous $HClO_4$ supplemented with 15% (v/v) MeOH. However, the Crownpak® CR(+) column was not available within the department, leading to the adaptation of the CHIRALPAK® IA-3 method for the synthesised compounds. Unfortunately an instrument malfunction prevented this analysis from taking place.

2.3.2 Direct deuteration of tryptophan

Labelling Trp with at least two deuterium atoms shifts the mass of the molecule significantly enough to be distinguished from natural (unlabelled) Trp by mass spectrometry. The method of Winneca *et al*. (2009) was adopted to prepare both enantiomers of deuterated tryptophan.[150] This involved stirring commercially available L- and D- Trp in a 2:1 solution of deuterated Trifluoroacetic acid (TFA-*d*) and D_2O in darkness for several days with the solvent replaced at regular intervals. This process afforded deuterium incorporation at an average of three positions, as confirmed by mass spectrometry (Fig. 21). The major isotopic peak at *m/z* 208 corresponds to incorporation of three deuterium atoms in the dominant product, with some minor species containing more and some fewer deuterium atoms. However, it was not possible to determine which protons were deuterated using only mass spectrometry.

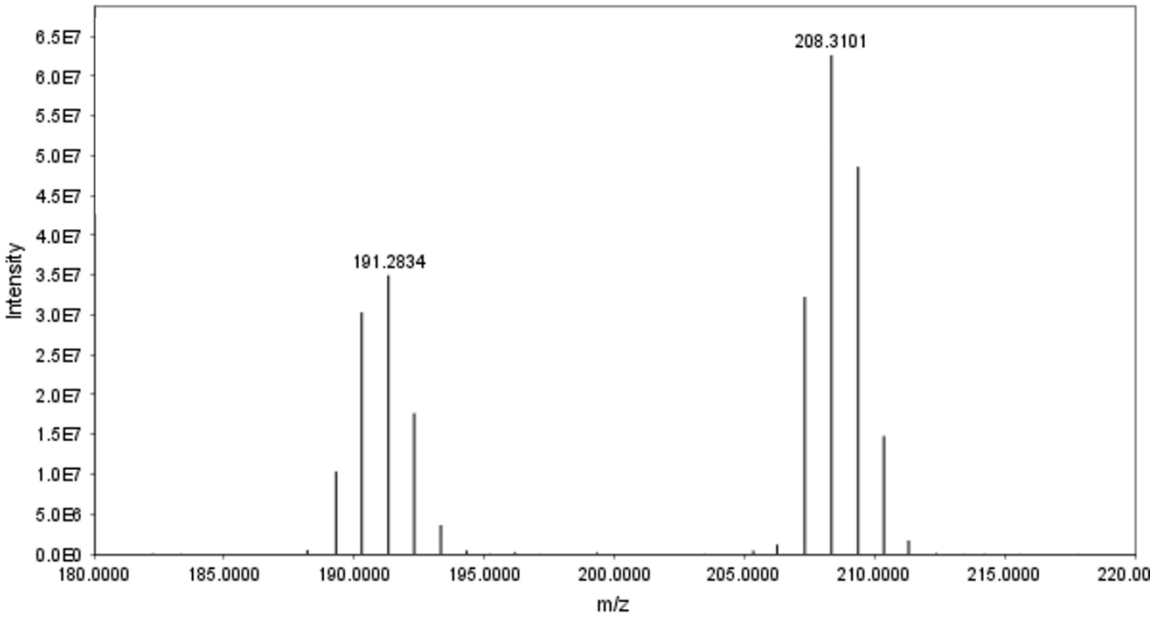


Figure 21: *Mass spectrum of partially deuterated L-tryptophan prepared using the method of Winneca et al. (2009).*[150] *The major isotopic peak corresponds to incorporation of three deuterium atoms.*

It is possible to say that the protons bonded to heteroatoms were not the isotopic labels seen in the mass spectrum. At these positions, proton exchange is rapid, and deuterium uptake would occur at these positions first. However, they would be washed out by the methanol used to make up the mass spectrometry sample and subsequently by the water used to prepare the spray for plant application.

2.3.2.1 Confirmation of deuterium positioning

^{1}H NMR (Fig. 22) showed three proton signals in the aromatic region, with integrations of 0.63, 0.54 and 0.39, respectively, when compared to a known undeuterated site, C10. This shows that two aromatic positions are fully deuterated and three are partially deuterated to the degree of 41%, 54% and 60%. This also confirms the mass spectrometry data showing three deuterium atoms as the most prevalent species.

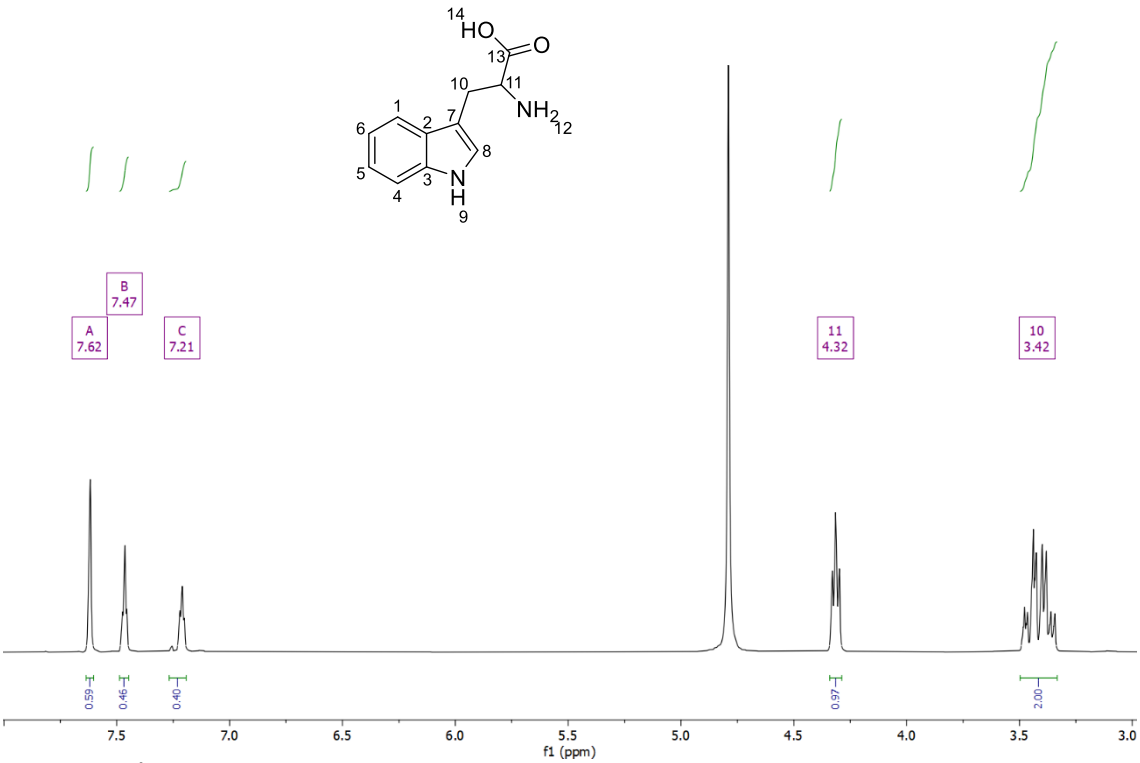


Figure 22: *^{1}H NMR spectrum of partially deuterated L-tryptophan (L-Trp-d_3), showing only three peaks in the aromatic region. Relative integrations of 0.59, 0.46, and 0.40 (normalised to undeuterated sites C10 and C11) indicate partial deuterium incorporation at three aromatic positions and full incorporation at two positions. Spectrum recorded in D_2O.*

One consideration with this method of calculating the amount of deuterium in the sample is that NMR can have a high integral error, with the concentration of the sample significantly affecting the accuracy of integration. At lower concentrations, the signal-to-noise ratio can be low, making for less accurate integral measurements. The signal-to-noise ratio for the triplet at δ 7.21 ppm, which was the smallest peak in the spectrum, was 295:1. This indicates good spectral quality and supports the accuracy of integral measurements used for calculating deuterium content.

To determine which proton relates to which position, the ^{13}C NMR of the deuterated sample was compared to the ^{13}C spectrum of a standard sample of L-Trp. Using literature precedents it was possible to assign the carbons of a standard sample of L-Trp and therefore of the deuterated sample (Fig. 23).[151] From this, it can be seen that C8 and C6 positions are fully deuterated while C1, C4 and C5 are partially deuterated.

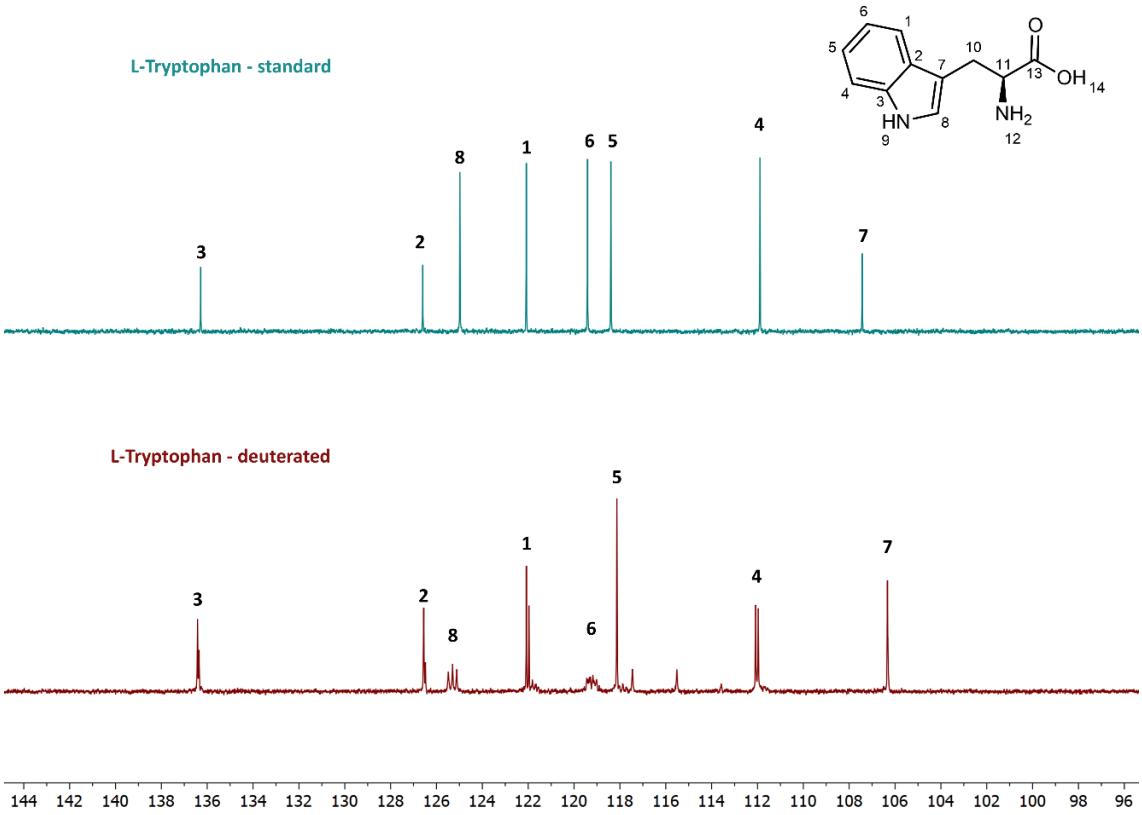


Figure 23: *^{13}C NMR spectra of L-tryptophan in the aromatic region: unlabelled standard (top) and deuterium-labelled sample (bottom). Comparison enables assignment of deuterated carbon positions. Triplet splitting patterns indicate full deuteration at C6 and C8, while singlet signals with minor splitting suggest partial deuteration at C1, C4, and C5.*

In ^{13}C NMR, the signal for a carbon bonded directly to a single deuterium atom is shown as a triplet. Only a triplet is visible for two of the aromatic carbon signals, whereas the other three have a singlet, followed by some splitting. This confirms that there are two fully deuterated positions (on C6 and C8) and partial deuteration on the rest of the aromatic carbons. The three partially deuterated positions were then assigned using Heteronuclear Multiple-Bond Correlation (HMBC) as C1 (54% deuteration), C4 (60% deuteration) and C5 (41% deuteration) (Fig. 24).

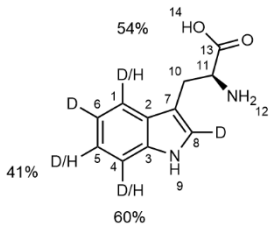


Figure 24: *Structure of partially deuterated L-tryptophan (L-Trp-d_3) showing assigned aromatic positions with D and D/H labels. Percentage deuteration at each site reflects combined data from mass spectrometry, ^{1}H NMR, and ^{13}C NMR analyses.*

2.3.2.2 D-Tryptophan

If L-Trp is metabolised to active intermediates that promote growth within the plant, rather than simply acting as a nitrogen source, a control experiment using D-Trp should produce no effect on growth. To further confirm whether the effect of exogenous amino acids on yield is related to the amino acids being metabolised by enzymes or some other factor, deuterated D-tryptophan (D-Trp-d_5) was synthesised using the same methods as for deuterated L-Trp. After the first 3 days of treatment, D-Trp-d_2 was observed as the prominent species, with the mass to charge ratio (*m/z)* for $[M+H]^+$ = 207.28. After another 3 days (6 days total), D-Trp-d_3 was the prominent species in mass spectrometry analysis, with *m/z* $[M+H]^+$ = 208.23. At the end of the 9th day of stirring, D-Trp-d_3 was still the prominent species present, having *m/z* $[M+H]^+$ = 208.27.

This experiment had a more substantial peak for D-Trp-d_4 $[M+H]^+$ = 209.30 than the deuterated L-Trp sample had (Fig. 25). The figure shows the base peak with an *m/z* of 208 and the second most prominent peak of 209. This shows that the most prominent compound was D-Trp-d_3 with D-Trp-d_4 being the second most prominent peak and is much stronger than the previous experiment with L-Trp. Although it would be assumed that using the same conditions over the same time frame should lead to the same level of deuteration, this was not seen to be the case. Possible reasons for this change could be the temperature of the laboratory while the experiments were being conducted, as the D-Trp reaction was done during the summer months and the L-Trp during the winter.

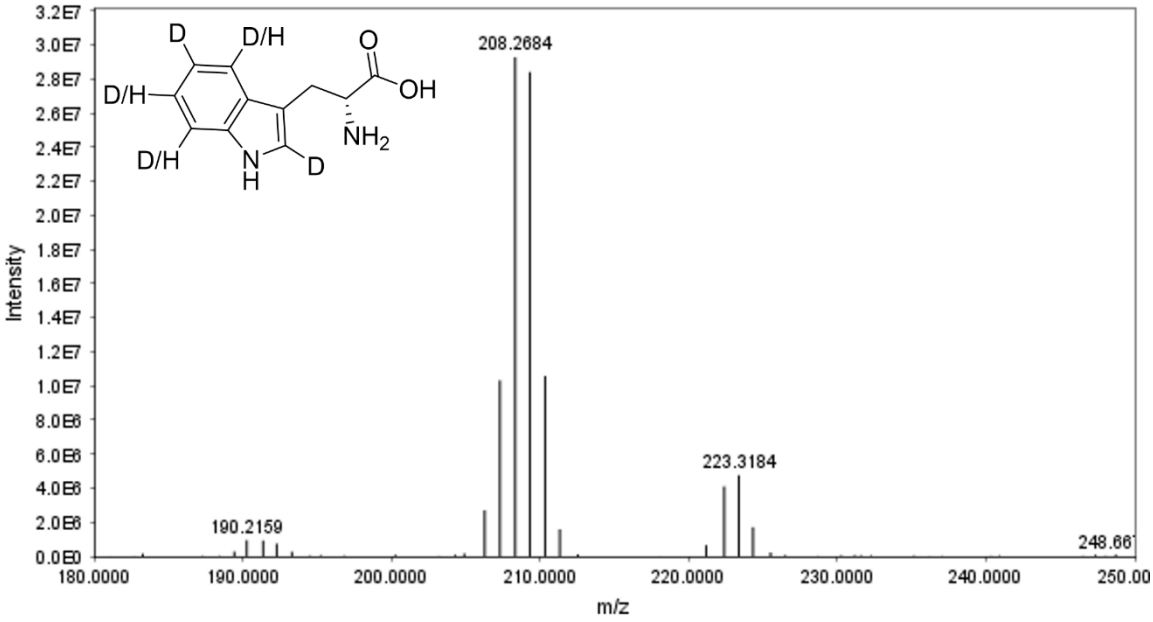


Figure 25: *Centroid extracted ion chromatogram (EIC) for deuterated D-tryptophan. The trace confirms the presence of partially deuterated species with retention behaviour consistent with D-Trp.*

2.3.2.3 Trifluoroacetic acid contamination

The method for synthesis of deuterated tryptophan given by Winnicka *et al.* (2009) removed the solvents through lyophilisation; however, due to the acid content, it was not possible to use the available freeze-dryer for the synthesis of compounds **10** and **11**.[150] Instead, a glass apparatus (Fig. 26) was made to act as a makeshift freeze dryer using the vacuum from a Schlenk line with an additional trap to remove the solvents. This meant that reaction mixture was frozen in a round bottom flask before being attached to the drying apparatus. Gentle heating was then conducted with a heat gun to encourage the evaporation of the solvents. A high-vacuum pump drew the solvents into a cold finger, condensing and dripping into the waste collection flask. This resulted in an off-white solid remaining in the initial round bottom flask.

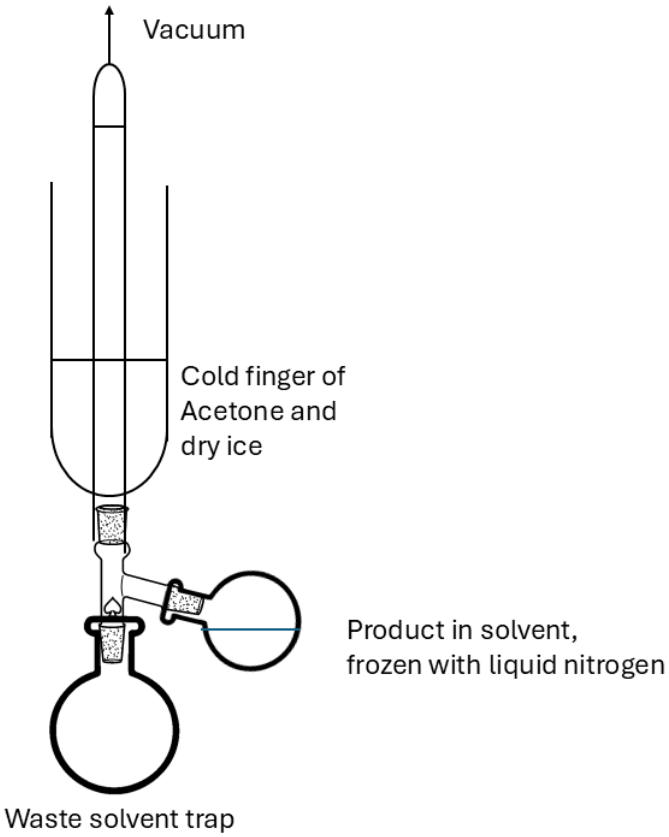


Figure 26: *Custom glass apparatus used to remove solvent from the reaction mixture during synthesis of deuterated tryptophan. Designed as a makeshift freeze-dryer, the system employed a Schlenk line vacuum, cold trap, and gentle heat to evaporate acidic solvents and isolate the deuterated product as an off-white solid.*

Once dried, the compounds were tested using ^{1}H and ^{13}C NMR and mass spectrometry to check for purity and were applied to plants. However, poor uptake was seen (Section 4.1.6), so further tests on the treatment liquid were needed. When a pH meter was used to check the pH values of the dissolved compounds which had been applied to the plants, those made using the deuteration method had pHs of around 1, while the 5-MeO tryptophan, made using the below described Schiff base method had a pH of 6.5. This suggested that some Trifluoroacetic Acid (TFA) remained stuck to the compound after drying. ^{19}F NMR confirmed this with a peak at δ -76, which correlates with literature data for TFA.[152]

To neutralise the increased acidity in the treatment, sodium bicarbonate was used to form a pH 7 buffer solution with each deuterated compound. The presence of TFA in the samples also meant that the intended concentration of the solutions applied to

plants was lower than initially planned. Therefore, it was necessary to calculate the amount of TFA present in each treatment to determine the final concentration of the amino acid applied.

To calculate this, 5-fluoro indole was used as a standard due to availability. It is a stable solid so it can be measured out accurately and contains a single fluorine group with a peak which is relatively close to the peak for TFA. A known amount of 5-fluoro indole (20 mg, 0.148 mM) was dissolved in 0.6 mL of deuterated dimethyl sulfoxide (DMSO-*d6)*, and ^{19}F NMR was performed using only a single spin. After allowing time for full relaxation between experiments, the same method was used on 20 mg of the impure deuterated L-Trp samples. This was also repeated with deuterated D-Trp samples. The absolute integral was then used to compare the amount of fluorine in the standard to that in the Trp-d_3 samples.

$$\frac{n_{TFA}}{n_{Indole}} = \frac{A_{TFA}/3}{A_{Indole}} \tag{2.1}$$

$$\frac{n_{TFA\ (L-Trp\ 1\ scan)}}{1.48 \times 10^{-4}} = \frac{\frac{35495.29}{3}}{21488.91} = 0.55$$

$$n_{TFA\ (L-Trp\ 1\ scan)} = 8.15 \times 10^{-5}\ mol$$

$$n_{TFA} \times M_{TFA} = mass\ of\ TFA$$

$$\left((8.15 \times 10^{-5}) \times 114.02\right) \times 1000 = 9.29\ mg$$

$$Mass\ \%\ TFA\ in\ sample = \frac{9.29\ mg}{20\ mg} \times 100 = 47\%$$

Where:
n_{TFA} is the number of moles of TFA in the sample
n_{Indole} is the number of moles of indole in the standard
A_{TFA} is the absolute integral for the TFA signal in ^{19}F NMR
A_{Indole} is the absolute integral for the 5-F indole signal in ^{19}F NMR
M_{TFA} is the molar mass of TFA

Accounting for TFA containing three times as much fluorine as the standard, the data showed that 47% of the L-Trp sample and 36% of the D-Trp sample were TFA. Similar numbers (47% and 35% respectively) were generated when the same calculations

were used on proton NMR data from 16 scans, demonstrating the precision of the method (Table 8). Integral error was estimated using the signal to noise ratio (calculated in MestreNova) for the signal region and a region of baseline using the following equations:

$$S/N\ Ratio = \frac{SNR\ (Signal)}{SNR\ (Noise)} \tag{2.2}$$

$$Integral\ Error = \frac{1}{S/N\ Ratio}$$

Where:
SNR (Signal) is the signal to noise ratio for an NMR peak
SNR (Noise) is the signal to noise ratio for part of the baseline
S/N Ratio is the signal to noise ratio for the NMR spectra

This approach ensures that the precision of the integral measurements is quantified, providing a reliable error estimate for the NMR data. The low integral error supports the reliability of the measurement, contributing to accuracy by reducing the impact of random noise.

Table 8: *Quantification of residual trifluoroacetic acid (TFA) in deuterated tryptophan samples using absolute integrals from ^{19}F NMR. Integrals from the deuterated samples were compared to those of a 5-Fluoro indole standard to estimate the amount of TFA remaining after solvent removal.*

	5-F Indole **standard**	L-Trp-d_3	D-Trp-d_4
Absolute integral from ^{19}F NMR - 1 Scan (Integral Error)	21482.5 (±0.28%)	37702.5 (±0.048%) 35863.9 (±0.59%)	27264.3 (±0.053%) 28076.8 (±0.074%)
Absolute integral from ^{19}F NMR - 16 Scans	334972.8 (±0.51%)	592568.6 (±0.53%)	444207.6 (±0.012%)

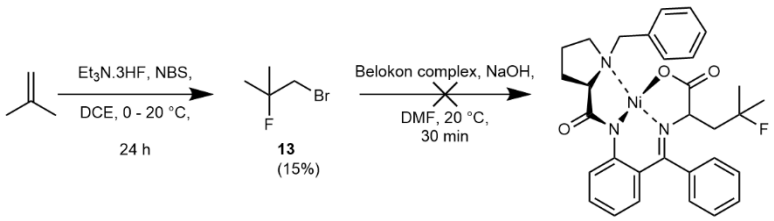
(Integral Error)			

2.4 4-Fluoroleucine

Following the successful synthesis of the labelled tryptophan derivatives, the focus moved to the synthesis of 4-Fluoroleucine. This compound was selected to explore whether the aliphatic nature of leucine, compared to the aromatic indole ring of tryptophan, would lead to differences in their modes of action, despite both being hydrophobic amino acids. The initial synthetic approach utilised the Schiff base methodology, mirroring the strategy applied to the fluorinated and methoxylated tryptophan derivatives to maintain methodological consistency and streamline reaction conditions. When this approach failed to produce the desired product, an alternative strategy involving photocatalysis was explored. Although 4-fluoroleucine was not fully isolated, these experiments provided valuable insights and are detailed below.

2.4.1 Schiff base method

The design for a side chain “clip” to be added to the Ni(II) complex was based on that outlined by Belokon (1988) for the synthesis of isoleucine, using bromide as the leaving group (Scheme 12). [135]



Scheme 12: *Proposed synthesis of 2-fluoro-L-leucine nickel complex using 2-methyl-1-propene as the starting reagent. The route was designed to introduce fluorine at the 2-position prior to chiral auxiliary coordination. However, the final step involving attachment of the fluorinated side chain to the Belokon nickel complex was unsuccessful under the tested conditions.*

Synthesis of 1-bromo-2-fluoro-2-methylpropane proved difficult due to the use of isobutylene, a gaseous reagent, which was introduced to the reaction vessel by

bubbling the gas through the solvent. Purification also proved problematic due to the volatile nature of the product and the delicate balance between getting a strong enough vacuum to distil the product without losing it to the vacuum. Failed reactions alongside the successive changes to the setup are given in Table 9.

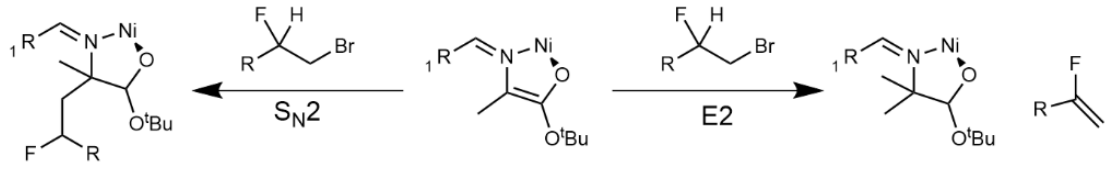
Table 9: *Summary of trialled methods for the synthesis of 1-bromo-2-fluoro-2-methylpropane. Experimental conditions, observed issues, and crude yields are listed for each attempt. Superscript a indicates product loss during purification.*

Attempt Number	Crude Yield	Reasoning for experiment Failure	Change needed for next attempt
3.2.1a	No product formed	Short needle used for gas line - gaseous reagent not able to saturate solution	Use a long needle to bubble the gaseous reagent through the solution. Change solvent from volatile DCM to a less volatile alternative (e.g. DCE)
3.2.1b	17%[a]	Product seen in crude NMR, lost during vacuum distillation	Small amount of product present, use Kugelrohr distillation for purification
3.2.1c	10%[a]	Product seen in crude NMR - lost during Kugelrohr distillation (degraded with heat)	Do not attempt purification. Take product on to next step as crude
3.2.1d	15%	No purification, Compound moved forward crude	

For both vacuum distillation and Kugelrohr distillation, no product was recovered in the receiving flask upon heating the distilling flask. In both cases, the solvent was removed upon prolonged heating of the flask, and a dark brown oil remained. Changing the solvent from dichloromethane (DCM) to dichloroethane (DCE), which is less volatile, reduced the risk of the reaction running dry while the gaseous reagent was bubbled through the solution with a long needle. This enabled a usable amount of the dihaloalkane to be isolated. The product was identified as the major component by NMR and gas chromatography mass spectrometry (GC-MS) and was used without further purification.

Further problems were encountered with the attachment of compound **13** to the Ni(II) complex. Increased time, temperature, basicity and equivalents of compound **13** were attempted to promote the reaction, but the complex was not recovered.

It was then thought that the positioning of the fluoride group might be interfering with the reactivity of the carbon to which the bromo group was attached and, therefore, hindering the ability of the bromide to act as a leaving group. Haufe *et al.* (1998) showed that for fluorobromoalkanes, under normal Schiff base alkylation conditions, an E2 elimination is favoured instead of synthesis of γ-fluoro-α-methyl amino acids (Scheme 13).[153]

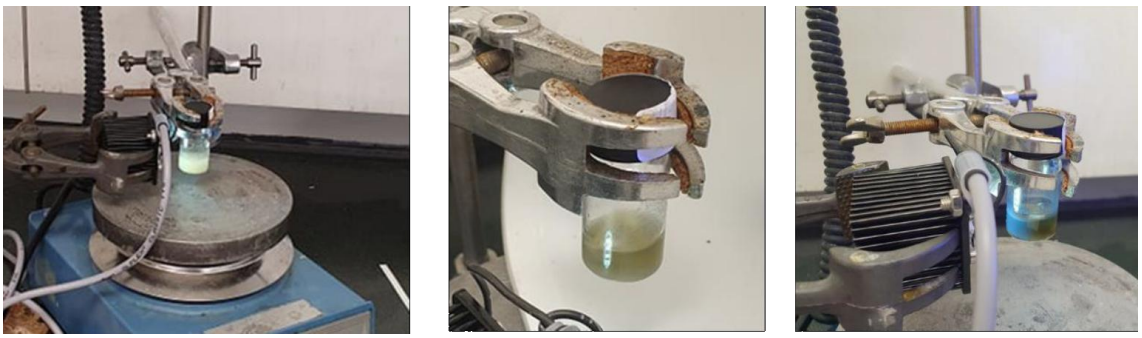


Scheme 13: *Products of S_N2 and E2 mechanisms arising from the reaction of a Schiff base with a fluoro-bromo-alkane. The scheme illustrates competing substitution and elimination pathways. (1998).*[153]Although E2 elimination is not possible with compound **12**, this suggests that the fluoride hinders the molecule from S_N2 alkylation. For the S_N2 alkylation reaction to occur, a strong base (KtBuO) with heat under inert conditions would be required. However, when this was tried using compound **12**, no product was recovered after 6 hours.

2.4.2 Photocatalyst method

Due to the challenges in producing 4-Fluoroleucine using the Schiff Base method, a different method was attempted using a photocatalyst, following the method of Nodwell *et al.* (2017).[154]

Direct fluorination of commercially available L-leucine using N-fluorobenzenesulfonimide (NFSI) as a source of fluorine and a photocatalyst (TBADT) to catalyse the reaction was attempted using a 365 nm ultraviolet light source. Colour change was seen over 20 hours (Fig. 27), and thin layer chromatography (TLC) monitoring showed the presence of a new compound.



10 minutes 6 hours 20 hours

Figure 27: *Photographs of the reaction vessel over time during photocatalytic fluorination of L-leucine using NFSI and TBADT under 365 nm UV light. A visible colour change from yellow to dark green was observed over 20 hours, indicating reaction progress.*

Though the desired product was identified by mass spectrometry of the crude material, it proved difficult to isolate. Therefore, the application of fluoroleucine to plant material was not taken beyond this point.

2.5 Isotopically labelled L-valine and L-leucine

The same Schiff Base method for synthesising methoxy- and 4-fluorotryptophan was proposed for the synthesis of isotopically labelled valine and leucine.

Incorporating ^{15}N-glycine into the Schiff base was used to label these amino acids. For the first ^{15}N labelled compound, leucine was chosen as a suitable target. Synthesis of the ^{15}N labelled Belokon complex was completed following the same

method as for the unlabelled complex in section 3.1, with ^{15}N labelled glycine used in place of normal glycine. The method for attaching the leucine side chain was adapted from the method for synthesis of isoleucine from Belokon (1988).[135] The method for attaching the leucine side chain was adapted from the method for synthesis of isoleucine from Belokon (1988).[135] This involved using 1-bromo-2-methylpropane for an S_N2 alkylation. However, NMR did not show the presence of the desired product.

Unlabelled material was used to test the synthesis of valine using the Schiff base method. This was successful, but due to time constraints, the synthesis of a labelled version using 2-bromo propane-*d7* was not made in time to be used in plant trials.

2.6 Summary

Overall, only tryptophan derivatives were fully synthesised for application to plants, though other amino acids were explored. The main synthetic routes were the deuteration of commercially available tryptophan and the synthesis of fluorine and methoxy-labelled L-tryptophan derivatives via Belokon's Schiff base complex.

Two major methodological modifications were required during the Schiff base reactions. Firstly, optimisation was required for the first steps in producing the ligand and complex. The main modifications came in synthesising a quaternary ammonium clip to add the indole side chain to the amino acids. This required optimisation of methods to selectively synthesise the monomer over the dimer. Further investigation into the promotion of dimer synthesis was also done for exploration of compounds which had not been previously reported in literature. 5-Fluoro-L-tryptophan, 5-methoxy-L-tryptophan, L-tryptophan-d_3 and D-Tryptophan-d_4 were all taken forward to plant trials.

Chapter 3: Development of plant and analytical systems for uptake studies

Before investigating the uptake of the synthesised amino acid derivatives, it was first necessary to establish appropriate biological and analytical systems to support the study. The choice of plant species was critical, as it could influence both the feasibility and outcome of the experiments. Once a suitable model was selected, methods for extracting and analysing free amino acids and metabolites from the plant material were developed to enable downstream detection and interpretation.

3.1 Evaluation of potential plant models

Multiple factors needed to be considered when choosing which species of plant to use for plant trials. The plant needed to be easy to grow within a growth chamber, have a relatively short growing cycle (ideally under 6 weeks), be low cost, have a comprehensive genetic profile, and have small trichomes and a non-waxy cuticle so that the treatment would have the best chance to enter the leaf. Additionally, selecting a plant with agricultural significance could enhance the industrial relevance of the research. Each plant model which was considered for the project was evaluated against these criteria for suitability. It would also be advantageous to use a plant which has easily isolatable plant parts for application of the treatment to specific areas.

3.1.1 Considered models

Several plant models were evaluated, including *Arabidopsis thaliana*, tobacco (*Nicotiana tabacum*), tomato (*Solanum lycopersicum*), lettuce (*Lactuca sativa*), and soybean (*Glycine max*). These species were selected as they represent commonly used research models as well as agriculturally relevant crops, and have been used in studies of amino acid metabolism and foliar uptake of applied compounds.[57, 95, 105, 155, 156] The suitability of each model was assessed against the criteria outlined above, with their advantages and limitations summarised below.

3.1.1.1 Arabidopsis

Arabidopsis thaliana is a well-studied and widely used plant model. This plant is most commonly used for genetic studies due to its fully sequenced genome and the availability of extensive genetic data.[157] It has a short growth cycle of around 6 weeks, which allows for rapid experimental turnover. The plant is small and so can be cultivated in a growth chamber.

However, despite its advantages, *Arabidopsis* presents certain limitations for this project. *Arabidopsis* is primarily a weed and so has little direct relevance to real-world agricultural practices. Moreover its leaf formation as small, tightly-packed rosettes further inhibits its use in this project. This is due to the difficulty of isolating the leaves for treatment application, which is necessary for distribution studies.

3.1.1.2 Tobacco

Tobacco (*Nicotiana tabacum*) is a widely grown industrial crop.[158] Many studies have used tobacco plants as models, particularly for studying uptake of applied compounds.[159, 160] This crop is also grown by the industry sponsor (Plant Impact Ltd) at their greenhouses at Rothamsted Research Ltd, Harpenden.

Tobacco plants take around 10 weeks to reach maturity, which is longer than other potential plant models. The leaves of the plant have a high density of large trichomes, which could potentially impact the ease of uptake of an applied treatment.[161] High trichome density can prevent the even spreading of liquid treatments across a leaf surface and cause the formation of droplets which roll off the surface. Trichomes can also act as a physical barrier, reducing the treatment's contact with the leaf surface where absorption occurs.

3.1.1.3 Tomato

The tomato plant (*Solanum lycopersicum*) is a widely cultivated crop with significant agricultural relevance. It has been used as a model plant for metabolomics research due to its small genome size which is fully sequenced.[162] Tomato's large leaves offer practical advantages for studies requiring substantial plant material, such as extraction and analytical chemistry workflows. The growth pattern of tomato leaves,

growing on distinct branches with significant gaps between groups of leaves, would allow for the isolation and targeted treatment of specific plant parts.

Tomato plants grow considerably larger and have more complex growth requirements than other potential plant models. They would require more space in a growth cabinet and demand more extensive care. They often have thick trichomes, which may interfere with the uptake of the treatment into the plant, potentially reducing the efficiency of the experimental procedure.

3.1.1.4 Lettuce

Lettuce (*Lactuca sativa*) plants are low cost and have a short growth period and mild growing conditions. They also have broad, non-waxy leaves which are good for efficient treatment application. The genetic characterisation of many lettuce varieties is also well documented.[163, 164] Extensive support and expertise in the lettuce varieties was available from the industrial sponsor. There is some agricultural relevance for lettuce, as it is a commonly used crop.[165]

Some lettuce varieties have a close leaf arrangement similar to that of Arabidopsis, which could make application of treatment to individual plant parts challenging.[166, 167] Lettuce can also be sensitive to environmental conditions such as water availability, temperature fluctuations and light conditions, which means great care is needed to maintain ideal conditions.[168, 169]

3.1.1.5 Soybean

Soybean (*Glycine max*) has significant agricultural relevance, with 394.97 million metric tons of soybean produced globally in 2023/24.[170] The genetic information of soybean is well characterised.[171] The leaves of the plant are large with a non-waxy cuticle. The industrial sponsor has significant experience with these plants. Soybean plants can be ready for treatment application around 4 weeks from germination, reaching full maturity within an additional 4 weeks. The trifoliate leaf structure, where three leaves grow at the end of each branch, allows for the precise isolation and targeted treatment of individual trifoliates.

Soybeans require warm conditions (20 - 30 °C) and need high light intensity with long photoperiods in order to achieve healthy growth. Maintaining these specific conditions in a shared growth cabinet could present challenges, mainly if other researchers are cultivating plants with differing environmental needs. These demands could lead to inconsistencies in growth and complicate experimental timelines.

3.1.2 Summary of plant options

All the proposed plant models had strengths and weaknesses regarding their suitability for this project. While each option offered specific advantages, two models (lettuce and soybean) demonstrated the greatest potential based on the key selection criteria (Table 10).

Table 10: *Comparison of potential plant models for treatment uptake studies Evaluated against key criteria including agricultural relevance, growth cycle length, ease of growth, treatment application potential, and cost. Tick (✔), cross (✘), and tilde (~) symbols indicate clear suitability, unsuitability, and partial suitability, respectively.*

Criteria	Arabidopsis	Tobacco	Tomato	Lettuce	Soybean
Agricultural Relevance	✘ Weed - limited industrial use	~ Common industrial crop, not food related	✔ Major crop species	~ relevance as a food crop but limited industrial application	✔ High relevance, major crop
Short Growth Cycle	✔ Around 6 weeks	✘ Around 10 weeks	✘ Around 8 - 10 weeks	✔ Around 6 - 8 weeks	✔ Treatment at around 4 weeks, harvest at 8 weeks

Ease of growth in lab conditions	✔ Small, Fast growing	✔ Common in research	✘ Large, complex care required	✔/(~) Mild conditions, though some special attention required	✘ Needs warm, high light
Ease of isolating plant parts	✘ Dense rosette leaves	✔ Large leaves, easy to isolate	✔ Distinct branches	~ Some varieties better than others	✔ Trifoliate leaves, easy to isolate
Genetic Resources	✔	✔	✔	✔	✔
Small trichomes/ non-waxy cuticle	✘ High trichome density	✘ High trichome density	✘ Thick trichomes	✔ Broad non-waxy leaves	✔ Non-waxy cuticle
Low cost	✔ Small, low maintenance	✘ High input costs	~ Moderate cost	✔ Low cost, widely available	~ Moderate cost

As the ability to conduct the growth within the department was of high importance for the start of the project, lettuce was chosen as the focus for initial trials as this was more suited for use in a growth cabinet.

3.1.3 Initial lettuce trials

Initial trials used lamb's lettuce (*Valerianella locusta*), and two lettuce cultivars 'Olenka' and 'Panisse' (*Lactuca sativa*) in a growth chamber. After seven weeks, lamb's lettuce performed best, but overall plant health and growth were poor, largely due to problems with the watering regime and growth cabinet environment. These issues made lettuce unsuitable as a model, and the approach was discontinued. A second trial with lamb's lettuce was planned but abandoned when COVID-19 restrictions halted laboratory access.

3.1.4 Final choice of soybean

Following the unsuccessful lettuce trials, it was decided to shift plant growth to the industrial partner (Plant Impact Ltd.) at Rothamsted, Harpenden. This change allowed the project to benefit from the controlled greenhouse facilities available at Rothamsted, providing a consistent and well-managed growth environment. Due to this decision, the plant choice for the trials was also changed to soybean as, with the improved resources and expertise for growing this plant was a better choice.

The switch to soybean as the plant model aligned the study with a major agricultural crop, enhancing the industrial relevance of the research. By focusing on soybeans rather than Arabidopsis, lettuce, or other commonly used plant models, the project ensured its findings could contribute directly to industrial agricultural practices.

The move to growing plants at an external site introduced new logistical considerations. After harvesting at the base of the stem, soybean samples were dried at 80 °C for 48 hours to prepare them for shipping from Harpenden to Durham for extraction and analysis.

3.2 Development of extraction methods

A range of extraction techniques were considered to isolate amino acids, plant metabolites, and hormones from plant tissues, ranging from traditional approaches such as maceration, hydrodistillation, and Soxhlet extraction, to more modern techniques including pressurised-liquid extraction, supercritical fluid extraction, microwave-assisted extraction, and enzyme-assisted extraction. These methods differ in extraction efficiency, throughput, solvent use, and suitability for temperature-sensitive compounds. The desired extraction method needed to be fast; extract amino acids, plant metabolites and hormones; be scalable for batch processing; and not cause the degradation of temperature sensitive compounds. Based on feasibility and compatibility with these analytical objectives, three methods: aqueous extraction, Soxhlet extraction, and ultrasonic-assisted extraction (UAE) were selected for trialling and method development.

3.2.1 Rationale for extraction technique selection

3.2.1.1 Traditional methods

Maceration and hydrodistillation are both traditional methods for extraction and isolation of plant-derived products. Both use relatively simple principles. These methods are particularly useful for small scale extraction of volatile compounds.

Maceration involves the soaking of ground plant material in a solvent within a sealed vessel for several days at room temperature with occasional agitation. This softens and breaks the cell walls, releasing chemicals from within the plant.[172] After the desired reaction time the mixture is filtered, leaving soluble compounds in the solvent. The method is simple and requires minimal equipment, making it accessible and practical with significant solvent flexibility and mild conditions. The time required for this type of extraction is a significant drawback and so it is impractical for studies using large numbers of samples. The time requirements can be reduced by introducing heat to the extraction process, but this can potentially cause degradation of compounds. Additionally, maceration involves high solvent use, potentially increasing costs and conflicting with green chemistry goals of minimising waste.[173]

The passive nature of maceration can also lead to incomplete extraction, limiting its efficiency for complex plant materials.

Hydrodistillation, by contrast, is used almost exclusively for extracting volatile compounds, especially essential oils, using steam and water.[172] There are three types of hydrodistillation: water, water and steam, and direct steam distillation.[174] In water distillation, the plant material is fully submerged in water, which is heated to produce steam. Volatile compounds are extracted into the steam and so are distilled off. The same process occurs with water and steam distillation, but rather than the plant material being submerged in the water, it is placed directly above the water. With steam distillation, the plant material is flushed with a continuous stream of steam, volatile organic compounds are extracted into the steam, cooled in the attached condenser and collected.[175] This last approach is gentler as the plant is not in direct contact with the boiling water, reducing thermal degradation.

Hydrodistillation can also enable fractional separation of compounds at different temperatures, with more volatile substances distilling at lower temperatures.[175] As it uses water rather than organic solvents, the method has environmental benefits and avoids the need for sample drying. However, its scope is limited to volatile compounds, and the elevated temperatures in all forms of hydrodistillation risk degrading heat-sensitive molecules.

Aqueous extraction is similar to both of these methods, in that it uses water but does not require heating. In this case the sample is sonicated to weaken cell walls and membranes to enhance solvent penetration and extraction of water-soluble molecules.[176] Compared to maceration, aqueous extraction is faster and requires less solvent, while also extracting a broader range of compounds than hydrodistillation. However, the resulting extracts often contain large amounts of chlorophyll, lipids, and other high-molecular-weight components, producing complex and messy spectra that obscure the target metabolites. As such, while simple and rapid, aqueous extraction did not meet the requirements for clean, reproducible analysis in this study.

Together these methods demonstrate the constraints of more traditional methods. Maceration, although mild, is too slow and solvent-intensive for large scale analysis. While hydrodistillation, though being a quicker procedure, is limited with the types of compounds which can be extracted and can cause the degradation of heat sensitive molecules. While better than the other two described methods, aqueous extraction still does not provide the scalability and clarity of spectra required for the project. As such, these methods did not meet the criteria for efficiency, scalability, and preservation of sensitive compounds required for this study, leading to further exploration of alternative approaches.

3.2.1.2 Solvent intensive methods

Solvent-intensive methods are widely employed for the extraction of plant-derived compounds, particularly when high yields are required. These approaches rely on large volumes of organic solvents, often under elevated temperature and pressure, to enhance solubility and mass transfer. Among the most established techniques are pressurised-liquid extraction (PLE), Supercritical Fluid Extraction (SFE) and Soxhlet extraction. Each of these methods provides distinct advantages in terms of efficiency and compound recovery but also presents challenges, including solvent consumption, potential degradation of thermolabile molecules, and environmental concerns.

Pressurised-liquid extraction (PLE) which is also known as accelerated solvent extraction (ASE), involves extracting compounds from solid samples using solvents at high temperatures (50 - 200 °C) and elevated pressure (10 - 15 MPa).[177, 178] The increased pressure allows the solvent to remain in a liquid state at temperatures higher than its normal boiling point. This speeds up the process of extraction compared to many traditional methods and improves the ability of the solvent to penetrate the plant material.

This technique is fast and can be easily automated for high throughput, making it particularly useful for studies with high sample turnover.[172] One significant advantage of PLE is its ability to reduce both extraction time and solvent consumption. By operating at high temperatures, PLE increases the solubility and mass transfer rates

of compounds into the solvent. Additionally, elevated temperatures decrease solvent viscosity, enhancing its ability to penetrate plant matrices effectively. Another advantage of PLE is its versatility in solvent choice, with it being possible to use either volatile or non-volatile solvents, depending on the target for extraction.

However, PLE also presents certain limitations. The high temperatures required for this technique can lead to the thermal degradation of heat-sensitive compounds, including some amino acids and plant metabolites such as tryptophan. The need for high pressure for this extraction technique requires specialised equipment which can be expensive. Working under high pressure also adds additional safety concerns.

Supercritical fluid extraction (SFE) shares some similarities in approach to PLE, but in this case a supercritical fluid which exhibits gaseous properties and has liquid solvating ability is used to extract from plant material. A commonly used supercritical fluid for this extraction method is CO2, which becomes a supercritical fluid above 31.1 °C and 7380 kPa.128 Once the analytes have been extracted into the CO2, the pressure can be reduced so that CO2 returns to its gaseous state leaving behind the extracted analytes in a concentrated form.

Carbon dioxide is non-toxic, non-flammable, cheap and easy to remove from the analyte. The method does not use organic solvents and the CO2 can be recovered and reused easily, so it is a more environmentally conscious method. Compared to PLE the conditions are much milder, which is beneficial for temperature-sensitive compounds. Modifiers can be used, such as ethanol, to improve the extraction of more polar compounds, as supercritical CO2 is more suited to the extraction of non-polar compounds.

The main drawbacks of this method is the high costs of the specialised high-pressure equipment and, as with PLE, the increased safety risks of working under high pressure. Though modifications can be used to improve polar extractions, SFE is best suited for non-polar compounds, such as essential oils, terpenes, and lipids, and less effective for polar plant metabolites and amino acids.

While both PLE and SFE offer more advanced and efficient options, their reliance on expensive, high-pressure instrumentation made them less feasible for this study. In contrast, Soxhlet extraction, provides a robust and accessible benchmark method that could realistically be implemented and compared alongside greener alternatives. Though Soxhlet extraction is an old technique, created by Franz von Soxhlet in 1879, it is still considered one of the best conventional extraction methods.[179] In Soxhlet extraction, a piece of glassware called Soxhlet apparatus is used (Fig. 28), with dried and ground plant material placed in a cellulose thimble within the extraction chamber.

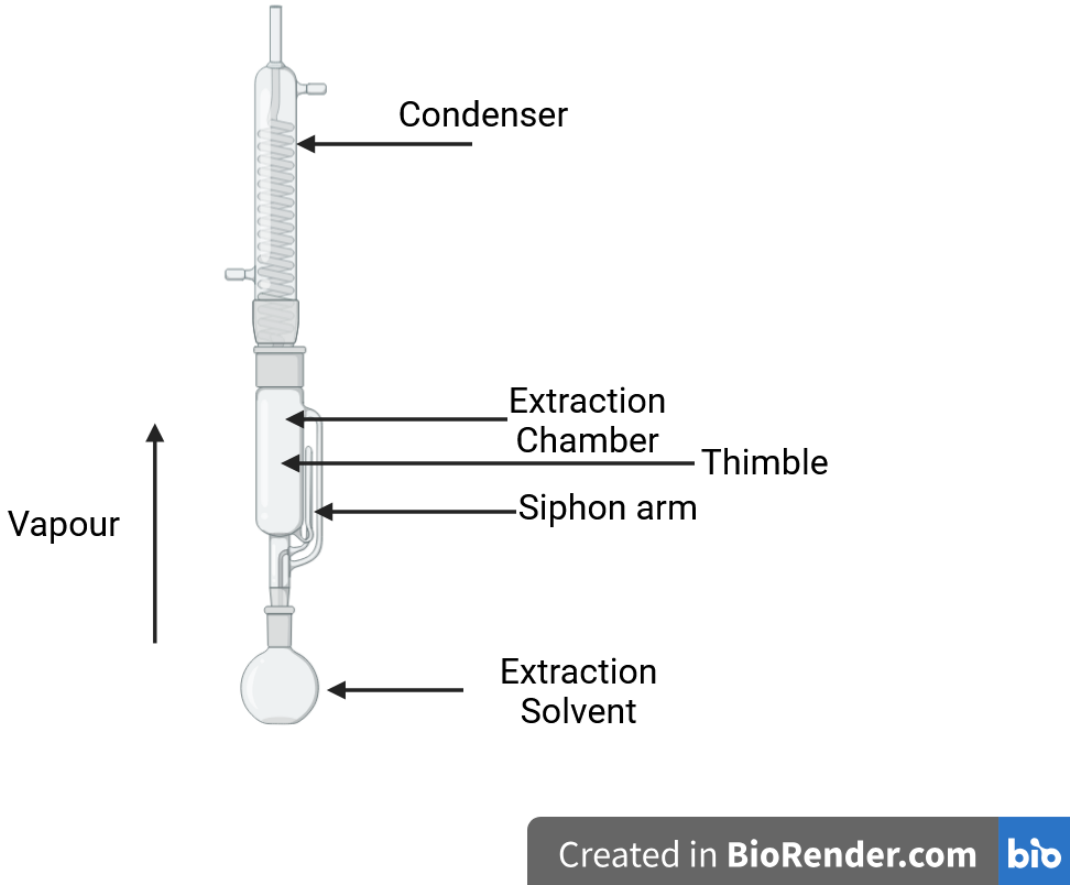


Figure 28: *Diagram of a Soxhlet apparatus used for solvent-based extraction of plant material. The system enables continuous cycling of heated solvent through a cellulose thimble containing dried plant matter, allowing efficient extraction of non-volatile and semi-polar compounds. Created with BioRender.com.*

The chosen solvent, held in the round-bottom flask, is heated to generate vapour. This vapour rises, condenses in the condenser, and accumulates in the extraction

chamber, allowing plant components to dissolve into the solvent. Once the chamber fills to the height of the siphon arm, the solvent automatically siphons back into the round-bottom flask, leaving the extracted material in the solvent. This cycle repeats until the extraction is complete, with the extracted material collected in the round bottom flask.

One significant benefit of Soxhlet extraction is the continuous cycling of solvent, ensuring that the solvent does not become oversaturated over time. The use of a cellulose thimble eliminates the need for filtration of the extraction liquid, simplifying sample preparation. Soxhlet extraction is highly versatile, suitable for extracting a wide range of compounds from plant material, including non-volatile and semi-polar compounds. The method is robust and well-established, offering high reproducibility and compatibility with many protocols available.

Soxhlet extraction requires less solvent than maceration however a significant amount of solvent is still required, particularly compared to other methods and when extraction of large numbers of samples is required. As the solvent used must be volatile enough to become vapour for the extraction to work, the scope of solvents is more limited than other methods. Soxhlet extraction can take many hours to complete, and when multiple successive solvents are needed, the process may extend over several days. This extended extraction time is impractical for studies with large sample numbers, and the continuous heating required raises environmental concerns due to the significant energy consumption involved.

Solvent-intensive methods represent a major step forward from traditional approaches, offering higher efficiency and adaptability. At the same time, their reliance on large volumes of solvents, specialised equipment, and energy-intensive conditions makes them less aligned with the environmental considerations of this study. These factors highlighted the need to look beyond solvent-intensive approaches and evaluate methods with a stronger focus on sustainability, while retaining the efficiency and reproducibility required for robust analysis.

3.2.1.3 Green Extraction Methods

The reliance of solvent-intense extraction techniques on large solvent volumes, energy input, and costly equipment highlighted the need for alternatives with a stronger environmental focus. This led to the consideration of so-called “green” extraction techniques, with the aim to reduce solvent use, lower energy demand, and improve the recovery of temperature-sensitive compounds while maintaining efficiency. Among these approaches, microwave-assisted extraction (MAE) and enzyme-assisted extraction (EAE) have attracted significant interest, while ultrasonic-assisted extraction (UAE) emerged as a particularly promising candidate for this study.

In EAE, enzymes are used to break down the cell walls of plant material to improve extraction.[173] There are two main approaches: enzyme-assisted aqueous extraction (EAAE) for the extraction of oils, and enzyme assisted cold pressing (EACP).[172]

For EAAE, plant material is mixed with water and the chosen enzymes. Enzymatic hydrolysis breaks down cell walls, releasing the intracellular compounds into the aqueous phase. This method is particularly effective for extracting oils, lipids, and water-soluble compounds. After extraction the aqueous phase is centrifuged and filtered to separate the oil or target compounds from the water. With EACP, the plant material is pretreated with the enzymes before being mechanically pressed at low temperatures to release the contents of the plant material. Commonly used enzymes include cellulases, pectinases, and proteases which target components of the cell wall.

EAAE is particularly effective for extracting water-soluble compounds and oils and can extract both polar and non-polar compounds. In contrast, EACP is beneficial when maintaining low temperatures, which is crucial when dealing with heat-sensitive compounds. EAE uses mild reaction conditions, with much lower temperatures required than other extraction techniques which is beneficial for the preservation of heat-sensitive compounds. This method can improve the amount and variety of material extracted from plants making it particularly useful for metabolomic investigations.

The main limitation of the EAE method is the cost of the enzymes, particularly for large scale projects where many plants will need to be processed. Potential enzyme inhibition by plant metabolites could also impact the extraction yield.

While EAE uses enzymes to break down cell walls, in MAE a microwave reactor is used to disrupt cell walls and enhance the release of intracellular compounds. Dried and ground plant material is mixed with a polar solvent such as ethanol, methanol or water. Microwave radiation is applied to the sample which interacts with the polar molecules in the solvent causing them to rotate rapidly. This rotation causes heat through molecular friction, heating the sample up rapidly. The heat and pressure caused by the microwave radiation increase the solubility of compounds, reducing extraction time from hours to minutes.[177] After extraction the mixture is cooled and filtered to leave the solvent containing extracted material.

Microwave extraction requires typically less than half the solvent required for Soxhlet extractions. It is fast, typically taking under an hour. Compared to traditional methods such as Soxhlet extraction, MAE significantly reduces both solvent use and extraction time, often achieving higher yields, particularly for polar compounds, while allowing the use of green solvents to support more environmentally friendly practices.

Limitations of this method include the risk of thermal degradation of heat-sensitive compounds and the high cost of obtaining specialised laboratory microwave reactors.

Ultrasonic assisted extraction (UAE) works in a similar way to MAE, but using ultrasonic waves instead of microwaves. These ultrasonic waves range from 20 kHz to 2000 kHz and are used to agitate plant material in a solvent, improving the release of compounds from within cells. The method is based on acoustic cavitation, where sound waves create microscopic bubbles in the solvent. When these bubbles collapse, they generate shear forces that break cell walls enhancing extraction efficiency. This effect also increases the surface contact between the solvent and the sample.[173]

UAE does not require volatile solvents so there is more flexibility in solvent choice than methods which rely on the vapourisation of the solvent. This also means that green solvents can be used such as water. The solvent requirement for this technique is very low. This method also allows for scaling and batch extraction, limited only by the number of extraction vials which will fit in the chosen ultrasonic bath. UAE is often more accessible than many modern extraction techniques because it only requires an ultrasonic bath or probe, which are common in most laboratories.

Prolonged running of the sonicator heats the water in the water bath surrounding the samples. This heating, if uncontrolled, could cause some damage to heat-sensitive compounds.

These greener extraction techniques demonstrate how modern approaches can balance efficiency, selectivity and environmental considerations. Use of enzymes

3.2.2 Trialled extraction techniques

Based on the comparative evaluation of extraction methods and the accessibility of available equipment, Soxhlet extraction was selected for initial trials while developing the mass spectrometry protocol. Its long-standing use in plant metabolite studies, straightforward implementation, and clear guidance in the literature made it a practical starting point. However, as the project advanced and the scale of sample processing became clear, a more efficient and sustainable approach was required. This led to the adoption of UAE as an alternative method.

To ensure consistent extraction across plant parts, each plant was systematically divided into different sections: a leaf from the trifoliate below application (T1), a leaf from the site of application (T2), a leaf from the trifoliate above the site of application (T3), a section of stem between T1 and T3 (STEM), the apical meristem (MS), and (where present) a seed pod (Pod) (Fig. 29). This sectioning allowed for targeted analysis of metabolite distribution and ensured reproducibility across replicates and time points.

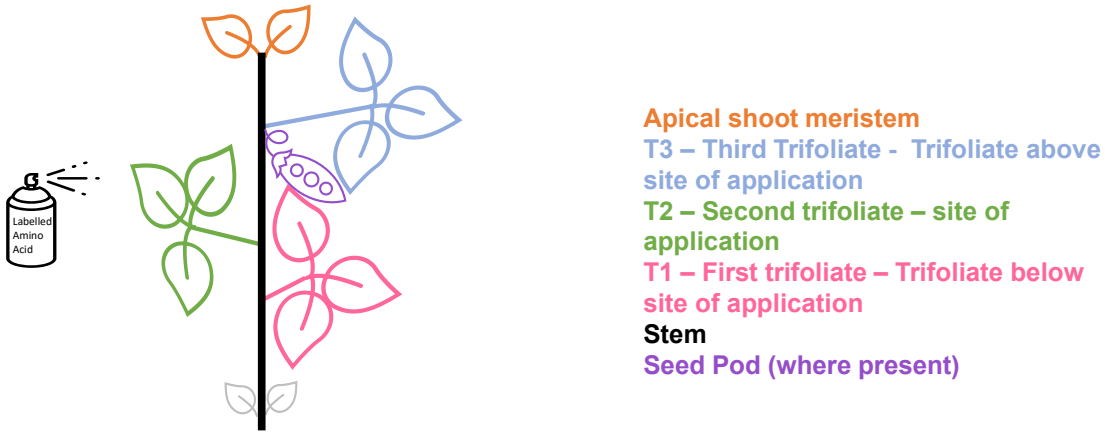


Figure 29: *Sectioning of whole plant for extraction from individual parts. Each plant was divided into six anatomical regions: T1 (trifoliate below application), T2 (site of application), T3 (trifoliate above application), stem (between T1 and T3), apical shoot meristem (MS), and seed pod (where present). This segmentation enabled targeted extraction from distinct tissues, allowing the tracing of deuterium labelled amino acid movement through the plant over time and across developmental zones.*

3.2.2.1 Soxhlet

A three-step Soxhlet extraction was undertaken to extract the relevant compounds following a three-step procedure modified by Vitalini *et al*. (2020).[180] Extraction using petroleum ether (40/60) followed by dichloromethane acted as defatting steps to remove chlorophyll as well as other large lipid components of the plant material. The crushed plant material was placed in a cellulose thimble within the extraction chamber. The chosen solvent, held in the round-bottom flask, was heated to generate vapour. The solvent vapour accumulated in the extraction chamber after condensing in the condenser, allowing plant components to dissolve into the solvent. Once the chamber filled to the height of the siphon arm, the solvent automatically siphoned back into the round-bottom flask, leaving the extracted material in the solvent. This cycle repeated until the extraction was complete, with the extracted material collected in the round bottom flask.

However, due to the number of samples to be processed, some drawbacks to this method were found. Processing capacity was restricted to five samples at a time due to equipment and space limitations. With each three-solvent batch requiring two days, the method proved too time-intensive for large-scale application. Additionally, solvent consumption exceeded 100 mL per sample, which was unsustainable given

the scope of the study, using over 100 plants, each segmented into five regions across multiple time points, resulting in nearly 1000 extractions. These logistical constraints rendered Soxhlet impractical for this project.

3.2.2.2 Ultrasonic assisted extraction (UAE)

To address the limitations of Soxhlet, ultrasonic-assisted extraction (UAE) was trialled as a low-cost and scalable alternative. Unlike many reported protocols employing probe-based ultrasonication, this study adapted the method for a 37 kHz laboratory ultrasonic bath.[181, 182] This approach enabled the simultaneous extraction of up to 24 samples with reduced solvent use and shorter extraction times, while remaining accessible within the constraints of available equipment.

Initial trials used the three previously used solvents for Soxhlet extraction but in glass vials placed into a sonic water bath. A sample of approximately 35 mg of leaf tissue was taken from the site of application at 4 hours - the time point at which deuterium labelled Trp was previously observed using Soxhlet extraction. The sample was ground following the general method for plant extraction (Section 6.3.3), then divided into vials containing methanol (Fig. 30, Method A), dichloromethane (DCM) (Fig. 30, Method B), and petroleum ether (Fig. 30, Method C), respectively.

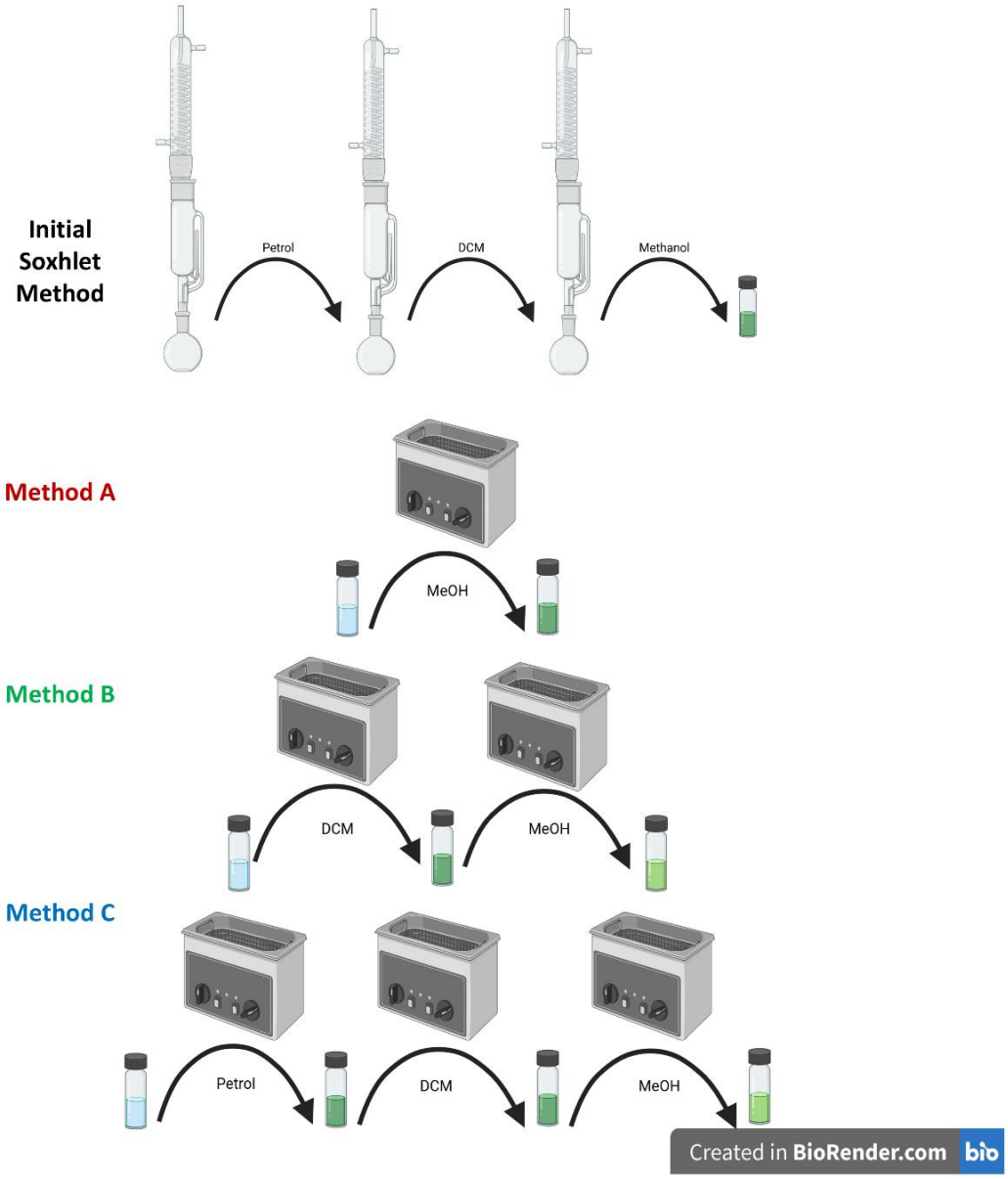


Figure 30: *Comparison of extraction methods trialled for ultrasonic-assisted extraction (UAE) of plant material containing the applied deuterium labelled L-Trp. The top panel shows the initial Soxhlet protocol using three solvents (petroleum ether, DCM, and methanol) in separate runs. The lower panels (Methods A-C) depict UAE adaptations using a 37 kHz ultrasonic bath: Method A trialled methanol-only extraction; Method B used sequential DCM then methanol; Method C replicated the Soxhlet solvent sequence. These trials aimed to identify a scalable, low-cost alternative to Soxhlet extraction with reduced solvent use and improved throughput. Created with BioRender.com.*

These vials were placed in an ultrasonic bath for 1 hour. After this time, the one step methanol extraction mixture (30a) was filtered to remove the plant material, the DCM was removed using a pipette from the two step extraction mixture (30b) and the

extraction mixture from the three-step extraction (30c) was centrifuged, and petroleum ether pipetted off. DCM was then added to the leaf material of the sample from Method B and methanol to Method C plant material. These were then sonicated for a further hour, and the Method C sample was treated with methanol following the same procedure.

^{1}H NMR of the final methanol samples under the three solvent conditions (Fig. 31) showed the most noise in the sample extracted with methanol alone (Fig. 31a). In contrast, little difference was observed in the aliphatic region between the two-step DCM-methanol extraction (Fig. 31b) and the three-step petroleum ether-DCM-methanol extraction (Fig. 31c).

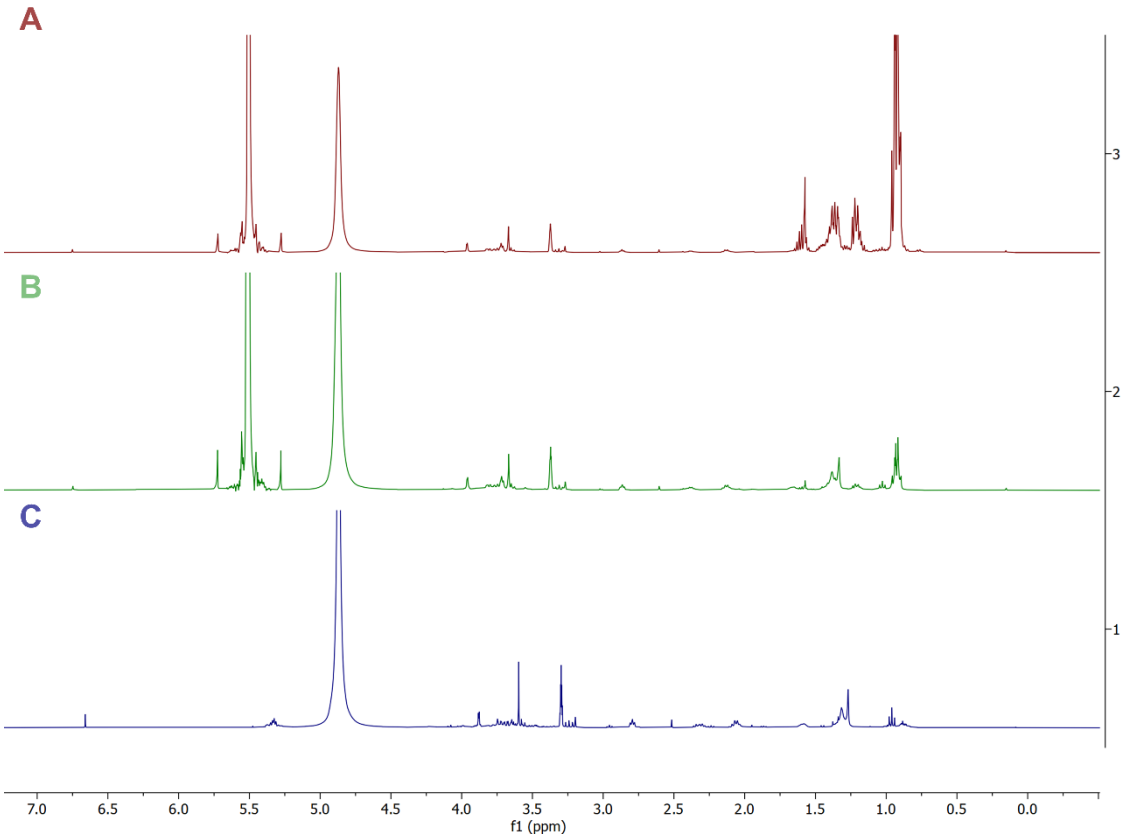


Figure 31: *Stacked ^{1}H NMR spectra of UAE-extracted plant material from the site of application at 4 hours, comparing three solvent conditions: methanol-only (A, red), DCM followed by methanol (B, green), and petroleum ether-DCM-methanol (C, blue). The two-step DCM-methanol method (B) shows reduced noise in the aliphatic region (0.5-2.5 ppm) and improved retention of aromatic signals (6.0-8.5 ppm), supporting its selection for subsequent analyses.*

The two-step DCM-methanol extraction (Fig. 31b) provides a strong balance between sample purity and retention of key signals. The aliphatic region (0.5-2.5 ppm) of this spectrum contains fewer impurities than the methanol-only extraction (Fig. 31a), indicating more effective removal of non-polar compounds. Although the petroleum ether-DCM-methanol method (Fig. 31c) also yields a clean aliphatic region, the additional petroleum ether step is not justified, as it increases solvent use and risks excessive loss of analytes. Furthermore, the two-step method retains more aromatic signals (6.0-8.5 ppm) than the three-step approach, which is particularly relevant for Trp analysis, as many tryptophan and indole-related resonances appear in this region. By reducing sample complexity while preserving key signals, the DCM-methanol method was selected for subsequent analyses, offering a practical balance between interpretability, efficiency, and solvent economy.

One problem found with this new method was removing the solvent without losing plant material. As extraction with multiple solvents was necessary, syringe filtration was not appropriate. Instead, using filter bags, commonly used for the infusion of teas, was examined as a way of keeping the plant material separate from the extraction solvent. As a trial of this method, a filter bag was exposed to the same conditions as those used for the plant extractions. Mass spectrometry of the extraction showed no difference between the sample with the filter present and a blank (Fig. 32).

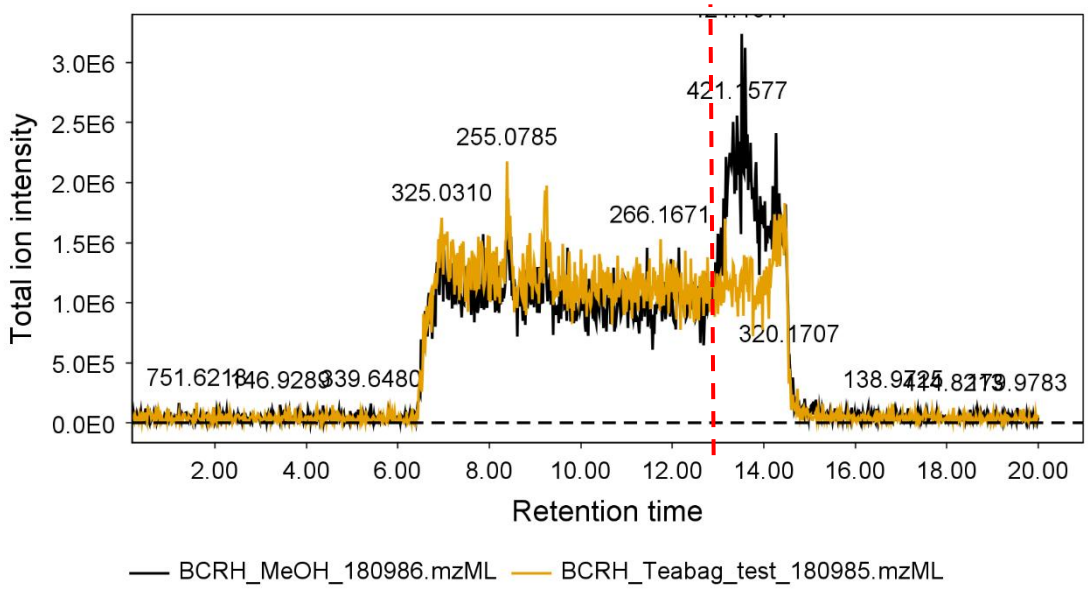


Figure 32: *A mass spectrum of a "blank" methanol sample (Black) overlayed with the mass spectrum from sonicating the filter bag material in DCM followed by methanol (yellow), following the protocol for UAE of plant materials. The vertical dashed red line marks the start of the column flush, during which all remaining material is removed from the system.*

In this test sample, as there are no additional peaks until after the 8-minute mark, it can be concluded that no contaminants leached out of the filter bags at detectable levels within the range of interest. This confirmed that the filter bags provided a clean and practical solution for solvent removal, ensuring reproducible extraction conditions in subsequent analyses.

3.2.2.3 Solvent trials with UAE

UAE allowed for a broader possible range of solvent systems to be trialled due to its lower solvent volume requirements and the ability to keep the extraction bag submerged throughout the process, unlike with Soxhlet. This flexibility allowed for the inclusion of aqueous mixtures, acids, and bases, which are typically incompatible with Soxhlet due to boiling point and apparatus constraints.

Seven solvent systems were tested (Table 11), selected to represent a range of polarities and chemical environments. These included pure organic solvents (methanol, ethanol), aqueous mixtures (aqueous dilutions of ethanol and methanol)

and buffered or pH modified systems (aqueous HCl, ammonium acetate and sodium bicarbonate).

Table 11: *Evaluation of final solvent choices for ultrasonic-assisted extraction (UAE) of tryptophan from plant material. Each solvent system was assessed for its ability to recover detectable levels of tryptophan, with outcomes based on mass spectrometric analysis of the extracted samples.*

Experiment Label	Solvents used	Presence of L-Trp-d_3 by mass spec?
A	100% methanol	No
B	50% methanol in water	No
C	100% ethanol	No
D	80% ethanol, 20% water	No
E	0.1 M hydrochloric acid	Yes
F	25 mM Ammonium acetate	Yes
G	10 mM sodium bicarbonate	No

Organic solvents alone (Experiments A-D) failed to recover detectable levels of L-Trp-d_3, suggesting that tryptophan was either poorly soluble under these conditions or retained within the plant matrix. The buffered base (sodium bicarbonate, Experiment G) also proved ineffective, likely due to poor solubility or degradation of the analyte under basic conditions.

In contrast, both 0.1 M hydrochloric acid (Experiment E) and 25 mM ammonium acetate (Experiment F) yielded detectable levels of L-Trp-d_3 (Fig. 33). While ammonium acetate offered a milder extraction environment and is commonly used in metabolomics workflows, it was ultimately not selected. This decision was made following consultation with Plant Impact Ltd., who advised against its use due to

concerns about downstream compatibility and potential interference with plant matrix components.

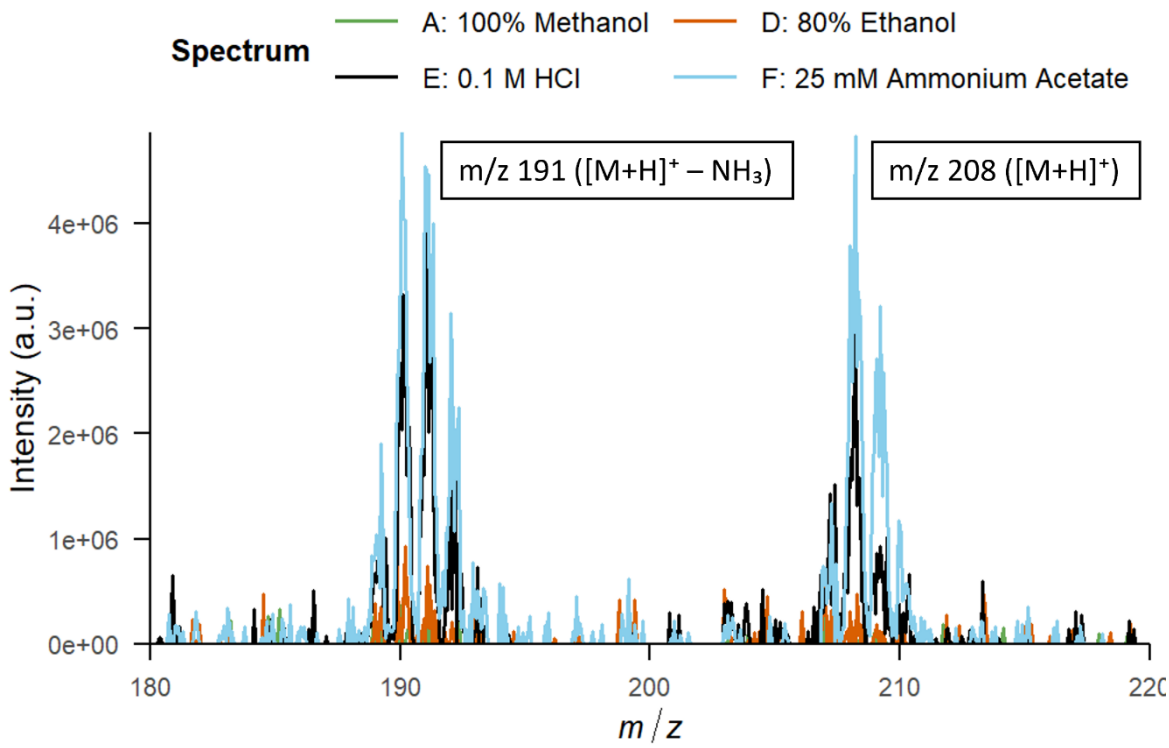


Figure 33: *Extracted ion chromatograms (EICs) for four of the trialled UAE solvent systems. Only 0.1 M hydrochloric acid (E) and 25 mM ammonium acetate (F) yielded detectable levels of L-Trp-d_3, as indicated by peaks near m/z 191 and 208. Organic solvents (A, D) showed no signal above baseline.*

Therefore, 0.1 M hydrochloric acid was chosen as the final extraction solvent. It provided consistent recovery of L-Trp-d_3, was compatible with mass spectrometric analysis, and aligned with industry guidance. This solvent was used as the second extraction step in all subsequent UAE protocols.

3.3 Detection techniques for monitoring uptake

The analytical experiments were designed with two primary aims: (i) to investigate changes in the chemical composition of plant extracts through the identification of metabolites and plant growth regulators, and (ii) to track the uptake and dispersal of the applied treatments within plant tissues.

Analysis of plant samples poses unique challenges due to their biochemical complexity. Even after the removal of large lipids and some chlorophyll through extraction with dichloromethane, the resulting extracts still contain many different compounds in addition to the applied treatment. To investigate the effect of the treatments on the plant, it is not just the treatment itself that must be identified from the samples but also other compounds that may only occur because of the applied treatment or in a different amount. Therefore, both broad, untargeted methods for comprehensive analysis of all compounds in the sample, as well as targeted methods for precise identification of the treatment compound, are required.

Several alternative methods of analysis were explored during this project. Each was selected for its potential to address specific research questions or improve analytical efficiency. However, most were discontinued after initial trials due to limited sensitivity, poor compatibility with plant extracts, or resource constraints that placed them beyond the scope of this study.

3.3.1 Discontinued techniques

3.3.1.1 Raman spectrometry

Raman spectrometry was explored for its potential to detect deuterated tryptophan through vibrational shifts outside of the fingerprint region. Although plant imaging with Raman has been demonstrated in previous studies, the low concentration of the deuterium labelled compound and strong chlorophyll fluorescence in leaf tissue posed significant challenges.[183] Chlorophyll absorbs in the visible range (400 - 700 nm) and typically fluoresces between 650-800 nm, interfering with Raman signals below 700 nm.[184] Mitigation would require advanced techniques such as NIR-FT-Raman, which were beyond the scope of this project.[183, 184] Given these limitations, the method was not pursued further.

3.3.1.2 Infrared spectrometry

Infrared (IR) spectrometry was trialled to distinguish vibrational differences between isotopes without liquid extraction. For initial tests, pure samples of both L-Trp and L-Trp-d_3 were analysed by IR. This showed a decrease in the C-H stretch in the

deuterated sample, though the intensity difference was not as significant as expected. This is most likely related to the deuteration of the aromatic region only being partial.

Ground leaf tissue from both the treated trifoliate and an adjacent untreated trifoliate was analysed by IR (Fig. 34). Minor spectral differences were observed in the fingerprint region, including a possible C-D stretch near 3000 cm^{-1} and an additional peak around 1500 cm^{-1} in the treated sample.

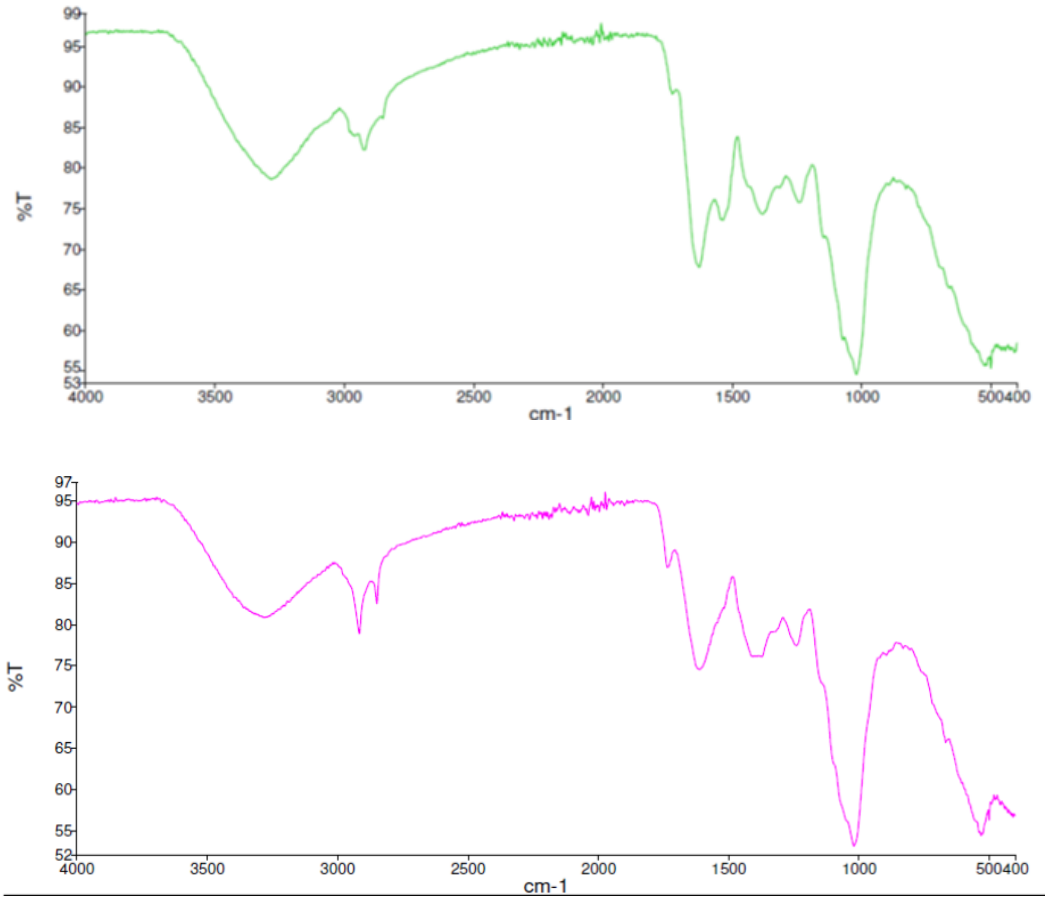


Figure 34: *Infrared (IR) spectra of leaf tissue from the trifoliate directly treated with L-Trp-d_5 (top) and an adjacent untreated trifoliate (bottom). Samples were ground and analysed without liquid extraction to assess vibrational differences between isotopically labelled and unlabelled tissue. Minor shifts in the fingerprint region and a potential C-D stretch near 3000 cm^{-1} were observed in the treated sample, though these did not fully align with pure standard spectra due to matrix complexity.*

However, these differences did not align with the spectra of the pure standards, likely due to the complexity of the plant material and overlapping signals from multiple compounds. Given the limited accessibility of the IR instrumentation and the challenges in achieving reliable quantification, this method was not pursued further.

3.3.1.3 Atmospheric solids analysis probe (ASAP) mass spectrometry

Atmospheric solids analysis probe (ASAP) mass spectrometry was also considered due to its ability to rapidly analyse solid samples with minimal sample preparation required, using direct ionisation to detect small molecules in complex matrices.[185] Initial tests on treated leaf tissue failed to detect clear deuterated Trp signals due to matrix complexity and lack of separation. The method lacked sufficient sensitivity and interpretability for this application and was not taken forward.

3.3.1.4 Solid state ^{19}F NMR

Solid-state NMR was initially explored for analysis of the deuterated Trp treatments; however, the low concentration of the deuterated compound was insufficient for detection using this method. Solid-state fluorine NMR was then considered for tracking the fluorinated Trp, as fluorine is not naturally present in the plant, so it would only be observed where the applied compound was present. However, logistical challenges prevented this analysis from being completed, and the method was not pursued further.

3.3.1.5 Fluorescence

Monitoring movement using fluorescence was initially considered as an analytical technique for the project due to the fluorescent properties of the methoxy-labelled tryptophan. However, early tests revealed that the compound was not as strongly fluorescent as expected, making it unsuitable for detecting the treatment at low concentrations. The addition of the methoxy group still allowed for detection via mass spectrometry, so this treatment could still be used in the analysis.

3.3.1.6 Solution NMR

The use of solution NMR was also considered, though the complexity of the plant extract would make analysis by NMR difficult. NMR also has detection limitations that are several orders of magnitude higher than mass spectrometry and so would not be suitable for the detection of treatment at lower concentrations.

3.3.2 Overview of selected detection methods

Mass spectrometry was ultimately selected as the primary analytical technique for this project due to its high sensitivity, minimal sample requirements, and versatility in performing both targeted and untargeted analyses. These attributes made it ideally suited for investigating treatment-induced changes within the biochemically complex plant extracts.

Full-scan mass spectrometry enabled broad detection of compounds, facilitating the identification of treatment-derived metabolites and potential plant responses. Alongside this, targeted multiple reaction monitoring (MRM) was employed for sensitive quantification of the applied amino acid treatments. Together, these complementary approaches provided both exploratory and focused insight into uptake, metabolism, and physiological impact.

Chapter 4: Monitoring uptake and physiological impact

Building on the analytical systems established in Chapter 3, this chapter applies mass spectrometry-based techniques to investigate the uptake, metabolism, and physiological effects of labelled amino acid treatments in soybean plants. Given the low concentrations of applied compounds (30-60 $mg \cdot mL^{-1}$) and their expected dispersal and transformation within plant tissues, sensitivity was a critical requirement.

Mass spectrometry data were converted into open-access mzML or mzXML formats for processing in MZmine, with downstream statistical analysis and data visualisation performed in R, allowing reproducible and efficient handling of large data sets.[186] These approaches enabled both targeted monitoring of labelled compounds and broader exploration of treatment-induced biochemical and physiological changes.

4.1 Targeted monitoring of labelled amino acids

To detect specific compounds at low concentrations, multiple reaction monitoring (MRM) can be used. MRM is a form of tandem mass spectrometry (MS/MS) used to monitor specific precursor-to-product ion transitions.[187] It works by filtering for selected ions over two mass measurements. First, the precursor ion is selected in the initial quadrupole (Q1), filtering out unrelated ions. The selected ions are then fragmented in the collision cell (Q2) through collision-induced dissociation (CID). Finally, the specific product ions are filtered and detected in the second quadrupole (Q3).[187] This approach allowed increased sensitivity over the full-scan method and reduces background interference in complex samples.

For this application, the chosen ions had *m/z* values of 205 (unlabelled Trp) and 208 (deuterium-labelled Trp) in positive ion mode. Endogenous Trp was used as an internal standard, providing a consistent reference for quantifying results across samples. This normalised the measurements and accounted for any variation in concentration and sample preparation.

Initial data analysis was based on the maximum signal value for peak intensity for each *m/z* value. This allowed quick results to compare the treated and untreated samples but did not account for the total signal across the peak. To address this, the area under the curve (AUC) of each *m/z* value was calculated using R (Appendix D.2).

4.1.1 Limit of detection for MRM quantification

Before using the targeted MRM system to analyse the plant samples, the limit of detection (LOD) needed to be determined. The LOD was determined by preparation of deuterated and undeuterated L-tryptophan standards ranging in concentration between 200 ng/mL and 25 ng/mL. Each standard was analysed using MRM on the Tandem Quadrupole Detector (TQD) mass spectrometer, and the response at each concentration was recorded. These results were plotted with response on the y-axis and concentration on the x-axis (Fig. 35). The LOD was defined as the lowest concentration at which the sample produced a measurable response, distinguishable from background noise.

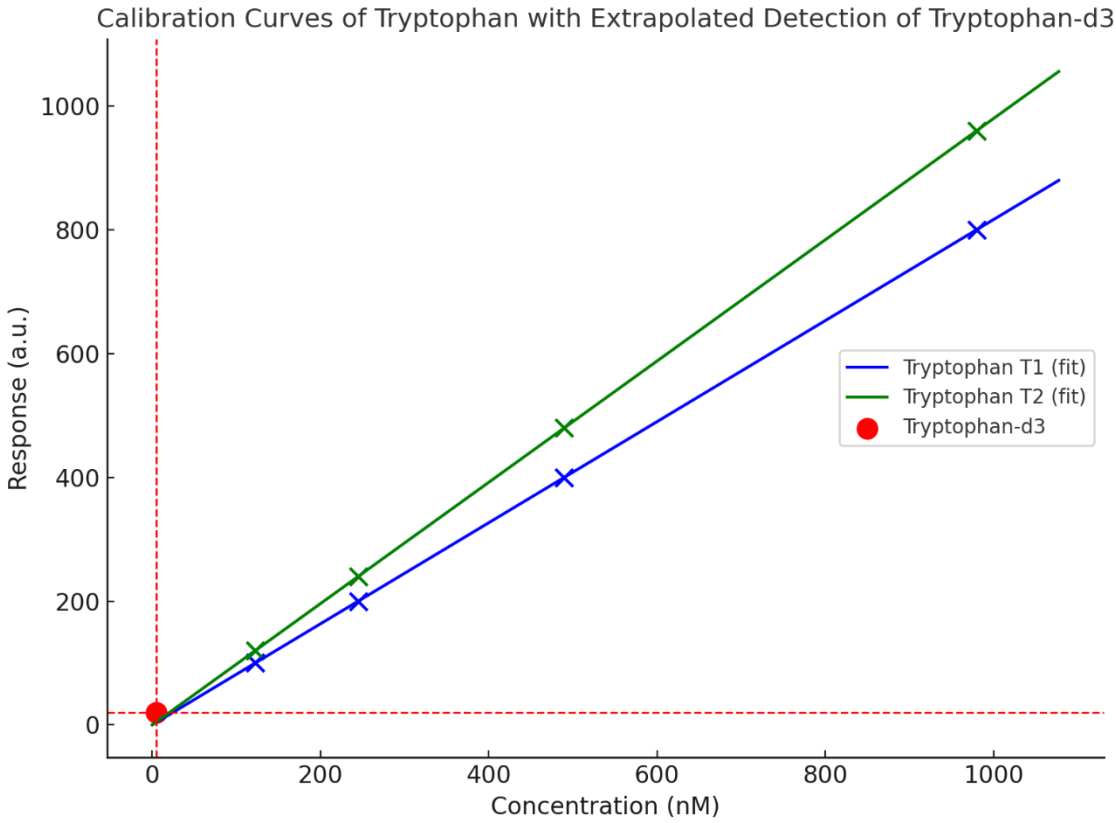


Figure 35: *Extrapolated calibration curves for two standard L-tryptophan samples (T1, blue; T2, green). Data points are shown as crosses, with fitted linear regressions extending below the lowest calibration point (25 ng/mL, equivalent to 122 nM). The red circle indicates the response of deuterated tryptophan (L-Trp-d_3) at 1 ng/mL (4.8 nM), with dashed red lines highlighting its position on the axes. The y-axis is reported in arbitrary units (a.u.), representing relative detector response from the mass spectrometer.*

The L-Trp-d_3 data point (Fig. 35) demonstrated that the treatment fits the trend of the untreated standards and so treatment is detectable even at low concentrations as it falls along the calibration curve. The dashed red lines indicate the position of the L-Trp-d_3 data point, highlighting its alignment with the predicted response. Figure 35 highlights the linear relationship between sample concentration and response across the tested range, demonstrating that the LOD lies below the lowest calibration standard of 122 nM.

4.1.2 Quantification by area under the curve

In chromatographic data, each compound elutes as a peak over time. The shape of this peak is shown by data points across the time frame. The size and spread of this peak reflect both the concentration of the compound and its ionisation efficiency.

Peak height (apex intensity) provides a quick estimate but does not account for peak shape, baseline noise, or partial elution. A more accurate measurement which considers the whole peak, not just the apex, is the area under the curve (AUC).
To calculate AUC accurately, peaks were modelled using Gaussian fitting, which assumes a normal distribution across the peak (Fig. 36). This approach reduced the influence of background noise and provided a more reliable estimation of peak area. Baseline correction was incorporated within the Gaussian fitting to further minimise noise, particularly in low-intensity regions.

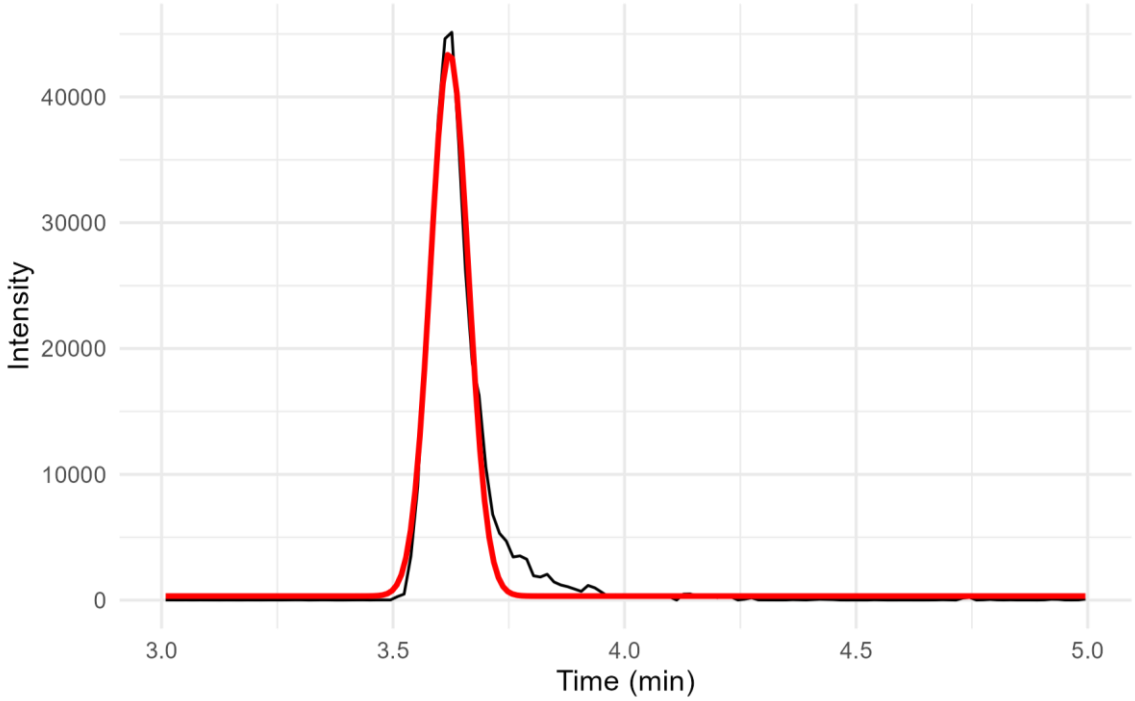


Figure 36: *Representative chromatogram showing a successful Gaussian fit (red) applied to a single elution peak. The fit models the peak shape using a normal distribution, enabling accurate calculation of area under the curve while accounting for baseline noise and partial elution. This approach improves quantification over simple apex intensity measurements, particularly in low-intensity regions.*

While Gaussian fitting provided the most accurate AUC estimates for well-resolved chromatograms, its applicability was limited in cases involving overlapping, poorly resolved, or asymmetrical peak shapes. To address these scenarios, alternative quantification methods were also assessed.

To reduce high-frequency noise while preserving peak integrity, spline smoothing was

applied to the raw mass spectrometry data. This method fits a continuous, piecewise polynomial curve through the data, allowing for flexible denoising that maintains the underlying signal structure. However, this method sometimes introduced artefacts in chromatograms where the compound was absent or present at very low levels. Trapezoidal integration was also considered, as it approximates the peak shape as a series of connected trapezoids, providing a simple way to capture the overall peak area.

Ultimately, a hybrid strategy was adopted where Gaussian fitting was used as the primary approach, but trapezoidal integration was employed in cases where peaks were poorly resolved or Gaussian fitting failed. This ensured that both well-defined and noisy chromatograms could be quantified consistently. With a reliable framework for quantifying peak areas in place, the next step was to evaluate which fragment ions provided the most consistent and representative signal for deuterium labelled and unlabelled tryptophan.

4.1.3 Interpretation of MRM fragmentation

To accurately interpret the MRM data, it was necessary to establish how the selected compounds fragmented under CID conditions, as this determined which chromatogram indices were suitable for quantification. Each .mzML file generated from the MRM method contained five chromatogram indices. Index 1 provided an overview of the full data set, while indices 2-5 each represented one precursor/product ion pair.

- **Index 2:** precursor ion *m/z* 205 (unlabelled Trp) → product ion *m/z* 188. This loss of 17 Da is consistent with the loss of the NH_2 group (Fig. 37B).
- **Index 3:** precursor ion *m/z* 205 → product ion *m/z* 146. This 59 Da loss is consistent with combined loss of NH_2 and CH_2CO, followed by rearrangement to form indole-methanol (Fig. 37C).
- **Index 4:** precursor ion *m/z* 208 (L-Trp-d_3) → product ion *m/z* 191.
- **Index 5:** precursor ion *m/z* 208 → product ion *m/z* 149.

These fragmentation pathways were consistent with assignments reported by Keller *et al.* (2013) for hypaphorine, a structurally similar compound to tryptophan.[188]

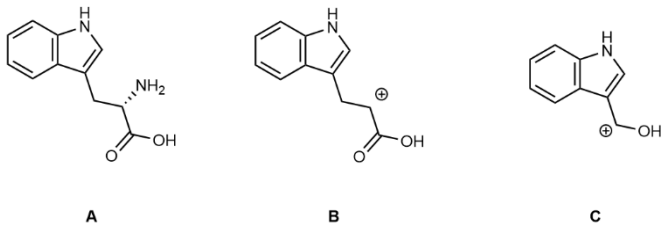


Figure 37: *Proposed fragmentation products of L-tryptophan (L-Trp) under collision-induced dissociation (CID) conditions, corresponding to precursor/product ion transitions observed in MRM analysis. A. Structure of L-Trp showing key functional groups; B. Product ion m/z 188, consistent with loss of the NH_2 group (17 Da) from the precursor ion m/z 205; C. Product ion m/z 146, consistent with combined loss of NH_2 and CH_2CO (59 Da), forming an indole-methanol structure.*

Fragment ion intensities derived from the same precursor ion appeared to behave proportionally (Fig. 38), suggesting consistent fragmentation across transitions. However, initial AUC calculations showed substantial variation between chromatogram indices, indicating that proportionality alone was insufficient for reliable quantification. This prompted further analysis of each index to assess baseline noise, peak shape, and fitting accuracy before selecting transitions for downstream quantification.

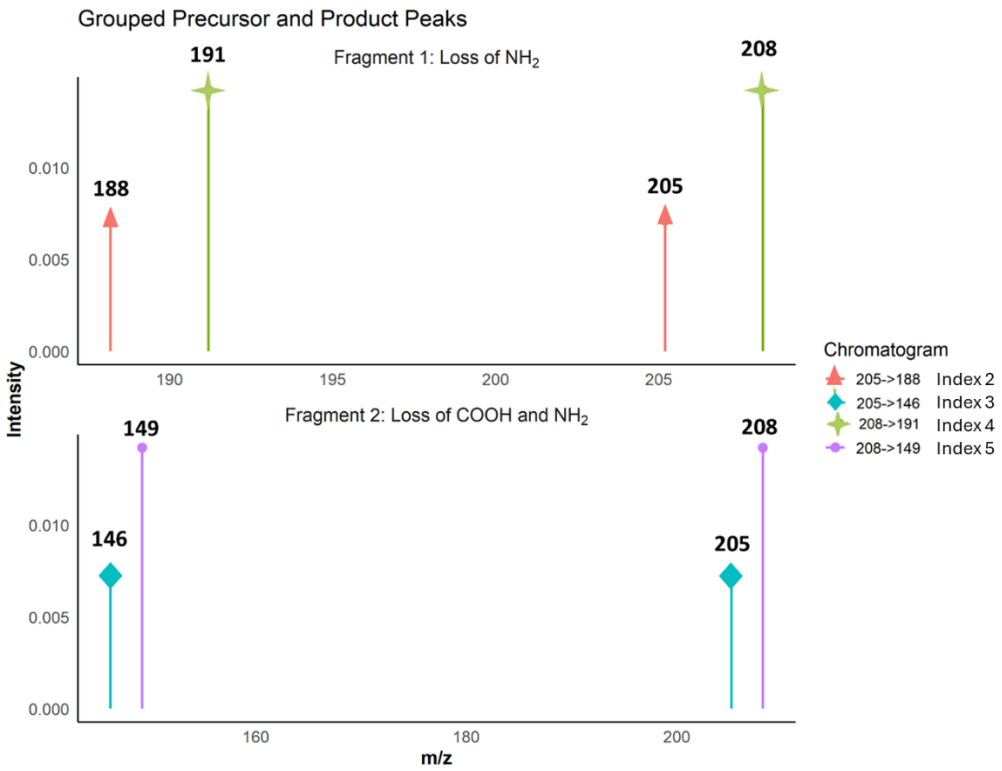


Figure 38: *Grouped MRM chromatograms showing precursor and product ion peaks from leaf tissue at the site of L-Trp-d_3 application, 4 hours post-treatment.*
A. Precursor ion m/z 208 (L-Trp-d_3) peak; B. Product ion m/z 191, corresponding to loss of NH_2; C. Product ion m/z 149, corresponding to combined loss of NH_2 and CH_2CO.

Scatterplots of AUC values for indices 2 vs 4 (Fig. 39A) and 3 vs 5 (Fig. 39B), representing fragment ion pairs from unlabelled and deuterated tryptophan respectively, showed strong linear correlation across 100 treated samples. Regression analysis confirmed that each pair of fragment ions behaved proportionally with respect to their shared precursor ion. However, this proportionality did not directly confirm consistency between the two fragment types (e.g., 188 vs 146 or 191 vs 149), which are used to infer tryptophan content. Further analysis was therefore required to assess the reliability of comparing different fragment types for quantification.

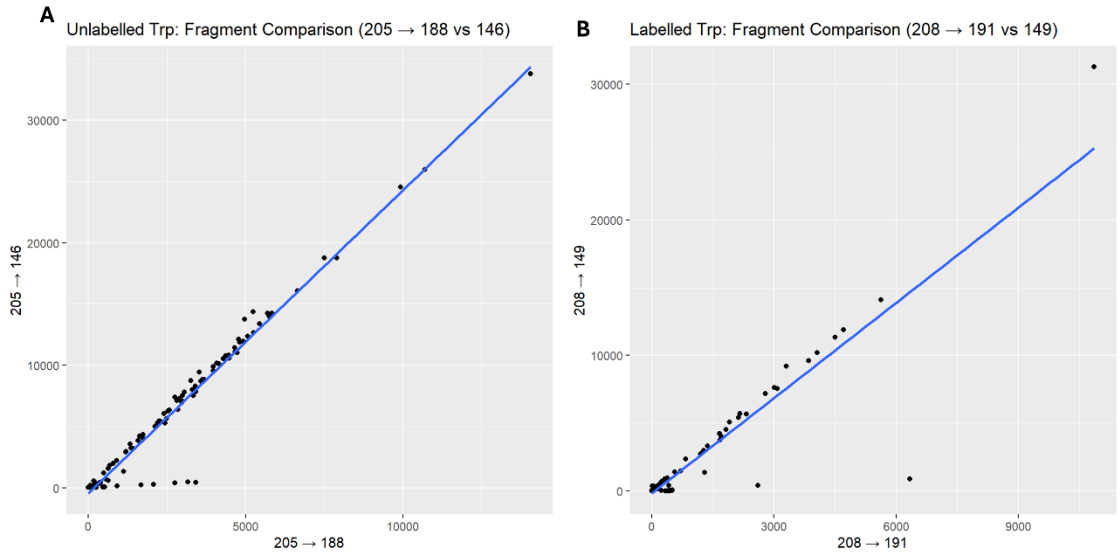


Figure 39: *Proportional relationship between fragment ion pairs derived from precursor ions 205 (A) and 208 (B). Scatterplots show linear correlation between AUC values for unlabelled tryptophan (205 → 188 vs 205 → 146) and deuterated tryptophan (208 → 191 vs 208 → 149) across 100 treated plant samples. Regression analysis confirmed proportional behaviour within each precursor, supporting the use of multiple transitions for quantification.*

To evaluate this, the ratios for index 2 to index 4, and index 3 to index 5 were calculated for each of the 100 treated samples. These were then compared to each other, and coefficient of variance (CV) values were calculated and plotted (Fig. 40).

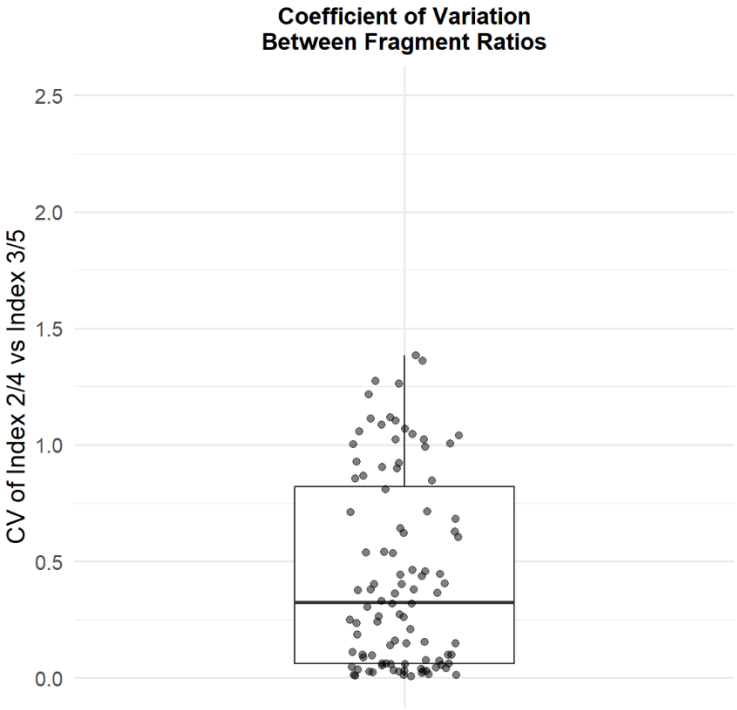


Figure 40: *Boxplot showing the coefficient of variation (CV) for fragment ion ratio comparisons between Index 2/4 and Index 3/5 across 100 treated samples. All CV values are plotted as a single group; the x-axis represents this aggregated comparison. The y-axis shows CV values, where lower values indicate greater consistency. Most values fall below 0.5, with a median of 0.33, suggesting strong agreement between fragment pairs used to quantify deuterium labelled and unlabelled tryptophan.*

With CV, the closer to 0 the value is the more consistent the data. This distribution confirms strong consistency, with most CV values below 0.5 and a median of 0.33. This suggested that both pairs of indices could be chosen for analysis. More consistent gaussian fitting was seen for indices 2 and 4 so this was the pair chosen for analysis.

4.1.4 Rate of dispersal and metabolism

Using the AUC calculated from each plant part at each time point, it was possible to monitor the movement of deuterium labelled tryptophan. The percentage of deuterated tryptophan relative to total tryptophan (Equation 4.1) could be calculated for each time point and plant part. This allowed the movement of the treatment throughout the plant to be tracked.

$$\%Trp = \frac{AUC_d}{(AUC_d + AUC_h)} \times 100 \quad (4.1)$$

Where:
%Trp is the percentage of deuterated Trp relative to the total Trp in a sample
AUC_d is the area under the curve of the deuterated Trp peak
AUC_h is the area under the curve of the unlabelled (natural) Trp peak

To standardise across variation in sample concentration and instrument response, the percentage of deuterium labelled tryptophan (L-Trp-d_3) relative to total tryptophan (labelled + unlabelled) was calculated for each sample. This approach ensured that observed differences reflected biological distribution rather than technical artefacts. Using these standardised values, it was possible to track the movement of the treatment over time: after initial uptake at the site of application, L-Trp-d_3 moved into the stem, and by 24 hours, further redistribution to the meristem and lower leaves was observed (Fig. 41).

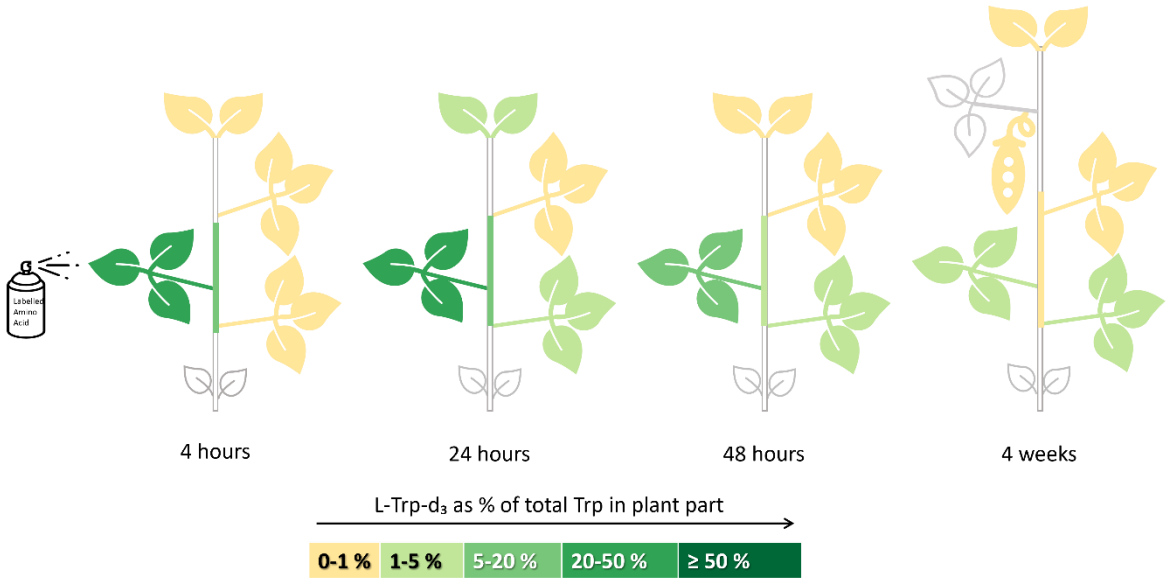


Figure 41: *Distribution of L-Trp-d_3 across plant tissues over time following foliar application. Colour-coded plant diagrams show the percentage of deuterium labelled Trp relative to total Trp in each tissue at 4 hours, 24 hours, 48 hours, and 4 weeks post-application. Mean values were calculated from n = 3 technical replicates for each plant part and time point. Colour legend: light yellow (0-1%), yellow-green (1-5%), light green (5-20%), green (20-50%), dark green (≥50%).*

The distribution of L-Trp-d_3 over time indicates an initial phase of rapid movement throughout the plant, followed by redistribution and retention in specific tissues. At 4

hours, most of the absorbed treatment remained at the site of application, with some movement into the stem already observed (Fig. 41, Table 12). At 24 hours, there had been significant movement both to higher and lower parts of the plant. This suggests phloem-based distribution through both acropetal and basipetal transport. At 48 hours the percentage of deuterated tryptophan across all samples had dropped. By 4 weeks, the proportion of L-Trp-d_3 in all tissues had decreased substantially. Only 1 - 5% remained in the application site and leaves below site of application, while all other plant parts, including the pods, contained less than 1%. This decrease in total Trp over time suggests that, beyond redistribution, the treatment was actively metabolised or incorporated into proteins rather than being retained in free form and that metabolism begins between 24 and 48 hours after application.

Table 12: *Average percentage of deuterium-labelled L-tryptophan (L-Trp-d_3) relative to total tryptophan in each plant tissue, calculated from three technical replicates per sample (n=3). Values are shown for each time point post-application, reflecting the distribution and decline of deuterated L-Trp across plant parts over time. Standard error of the mean (SEM) is given in brackets.*

	Internal L-Trp-d_3 (% of total Trp in each plant part)					
	T1	T2w	T3	Stem	Meristem	Pod
4 hours	0.1 (±0.09)	30 (±5.37)	0.3 (±0.05)	6 (±0.53)	0.5 (±0.20)	-
24 hours	2 (±0.63)	30 (±7.29)	0.4 (±0.17)	2 (±0.07)	2 (±0.29)	-
48 hours	1 (±0.18)	13 (±1.67)	0.2 (±0.01)	1 (±0.47)	0.3 (±0.08)	-
4 weeks	1 (±0.76)	3 (±0.26)	0.3 (±0.12)	0.2 (±0.08)	0.2 ($\pm 5.8\times10^{-4}$)	0.1 (±0.01)

Increased presence at the meristem at 24 hours may indicate that metabolism is occurring in actively growing tissues, where tryptophan is required for protein biosynthesis, auxin production, and secondary metabolite formation. The absence of significant accumulation in any one tissue at later time points further supports the

hypothesis that tryptophan is rapidly integrated into metabolic pathways rather than remaining in a free form.

To further explore the metabolic fate of the treatment, the proportion of deuterated L-Trp-d_3 relative to total Trp in each plant part was examined over time (Fig. 42). Calculating the proportion of L-Trp-d_3 within each plant part over time provides insight into where metabolic processing may occur and how rapidly the treatment is integrated into plant biochemistry.

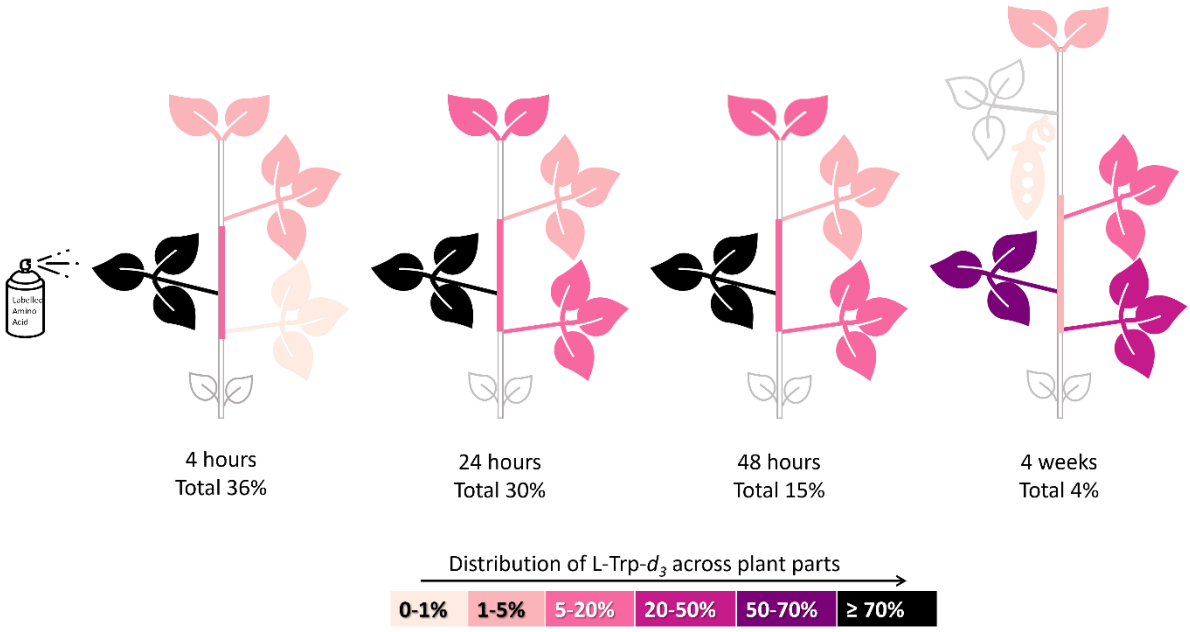


Figure 42: *Relative distribution of L-Trp-d_3 across soybean plant tissues over time, expressed as a percentage of total L-Trp-d_3 detected across all sampled tissues at each time point. Colour-coded diagrams show how the labelled fraction is partitioned between tissues, based on average deuterated Trp levels in each part. The percentage beneath each diagram indicates the proportion of total tryptophan (labelled + unlabelled) that was deuterated at that time point, contextualising overall recovery and metabolic fate. Mean values were calculated from n = 3 technical replicates for each plant part and time point. Colour legend: light pink (0–1%), pale rose (1–5%), soft mauve (5–20%), lavender (20–50%), plum (50–70%), deep purple (≥70%).*

At 4 hours, the highest proportion of deuterated L-Trp was detected at the site of application, where 80% of the deuterium labelled tryptophan was present (Fig. 42, Table 13). Early movement into the stem was also observed with 17% of the deuterated tryptophan being present in the stem. At 24 hours, 5% of the treatment was present in the meristem. The apical meristem is a significant site of auxin

biosynthesis and since Trp is a precursor to the auxin indole-3-acetic acid (IAA), it is possible that the transported L-Trp-d_3 was converted into auxin at this site. The corresponding decrease in transport activity over time further supports the idea that initial movement was driven by rapid phloem translocation, while later retention reflects metabolic integration.

Table 13: *Proportion of deuterium-labelled L-tryptophan (L-Trp-d_3) detected in individual plant tissues at 4 hours, 24 hours, 48 hours, and 4 weeks post-application. Values represent the percentage of total labelled Trp measured across all sampled tissues at each time point. T1: trifoliate below the site of application; T2w: washed leaf from the application site; T3: trifoliate above the site of application. "Stem" refers to the internodal segment between T1 and T3; "meristem" to the apical shoot meristem; "pod" to a seed pod where present. Mean values were calculated from n = 3 technical replicates for each plant part and time point. Standard error of the mean (SEM) is given in brackets.*

	Proportion of L-Trp-d_3 in each plant part (% of whole plant total)					
	T1	T2w	T3	Stem	Meristem	Pod
4 hours	0.4 (±0.25)	80 (±14.68)	1 (±0.16)	17 (±1.43)	1 (±0.40)	
24 hours	5 (±1.82)	85 (±20.76)	1 (±0.51)	5 (±0.20)	5 (±3.45)	
48 hours	5 (±1.24)	85 (±10.73)	1 (±0.02)	8 (±3.24)	2 (±0.58)	
4 weeks	23 (±17.49)	62 (±5.87)	6 (±2.60)	4 (±2.06)	4 (±0.01)	1 (±0.26)

To confirm this metabolic fate, further analysis to attempt to identify deuterated metabolites within plant samples was required.

4.1.5 Rate of uptake

Uptake efficiency is a critical factor in determining the viability of biostimulant treatments for agricultural use. Fast uptake, ideally within 24 hours, is essential for industrial applications as after this time there is greater potential for the treatment to be washed away by rain.[189] If the treatment cannot penetrate plant tissue rapidly enough, it can be lost to environmental factors, reducing the potential benefits of the

application. This could also lead to the need for increased levels of treatment, which is an economically unsustainable option for end users.

MRM data from the site of application was used to calculate how much of the applied treatment had been absorbed into the leaf at each time point by comparing the washed and unwashed leaf at the application site. The unwashed leaf gave the total content of deuterated tryptophan, present both inside and on the surface of the leaf, while the washed leaf represented the internalised treatment (Fig. 43).

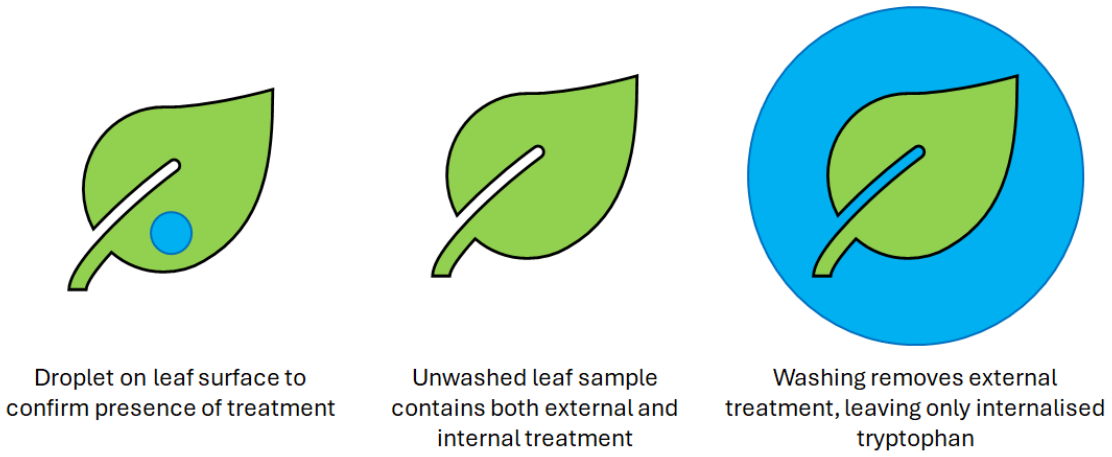


***Figure 43**: Sample preparation used to assess internalisation of L-Trp-d_3 at the application site. Droplet applied to leaf surface to confirm treatment presence (left). Unwashed leaf samples provided the total Trp-d_3 content, including both surface residue and internalised compound (centre). Washed leaf samples, rinsed prior to extraction, represented only the internalised fraction (right). Comparing these two sample types enabled estimation of uptake efficiency and absorption dynamics over time.*

The internal treatment fraction was calculated using the following equation (Equation 4.2):

$$Internal\ treatment\ fraction = \frac{\%Trp_w}{\%Trp_{uw}} \qquad (4.2)$$

Where:
$\%Trp_w$ is the percentage of deuterated L-Trp relative to the total Trp in the washed sample
$\%Trp_{uw}$ is the percentage of deuterated L-Trp relative to the total Trp in the unwashed sample

Using the calculated percentages for washed and unwashed samples rather than the raw intensity data allows you to compare overall sample intensities which could arise due to variations in extraction efficiency, sample concentration, or measurement conditions.

Plotting the internal treatment fraction over time can provide a graph to estimate the uptake rate. This can illustrate whether the treatment was absorbed into the plant within the timeframe identified as necessary for industrial applications.

To distinguish between surface residue and internalised treatment, three sample types were analysed: T2spot (surface solution), T2uw (total leaf content), and T2w (internalised content only). Comparing these values over time provides insight into how much of the applied L-Trp-d_3 remained on the leaf surface, how much was absorbed, and how this balance shifted during the treatment period. This approach enables estimation of uptake efficiency and helps identify when redistribution or metabolism begins (Table 14).

Table 14: *Percentage of total tryptophan (Trp) in leaf samples represented by deuterium-labelled L-Trp (L-Trp-d_3), used to assess uptake and internalisation over time. Samples include a surface spot of dilute hydrochloric acid solution (T2spot), an unwashed leaf from the application site (T2uw), and a leaf washed for 10 minutes with dilute hydrochloric acid (T2w). T2spot reflects the composition of the applied solution; T2uw represents total L-Trp-d_3 content (surface + internal); T2w indicates internalised L-Trp-d_3 only. Values are averaged across three technical replicates (n=3) per time point, with standard error of the mean shown in brackets.*

Plant part	Time point	Average% of deuterated Trp (± SEM)
T2spot	4 hours	99 (±0.47)
T2spot	24 hours	99 (±0.86)
T2spot	48 hours	100 (±0.05)
T2spot	4 weeks	7 (±6.69)
T2uw	4 hours	58 (±3.35)

T2uw	24 hours	46 (±2.00)
T2uw	48 hours	47 (±3.01)
T2uw	4 weeks	8 (±6.19)
T2w	4 hours	29 (±5.36)
T2w	24 hours	30 (±7.29)
T2w	48 hours	13 (±1.67)
T2w	4 weeks	0.39 (±0.26)

To confirm the presence of externally applied L-Trp-d_3, a droplet of 0.1 M HCl was placed on the surface of the leaf for 10 minutes and analysed (referred to in Table 14 as T2spot). This confirmed that treatment remained on the leaf surface for up to 48 hours after application. By 4 weeks, the percentage of deuterium labelled Trp recovered from the surface had dropped substantially, suggesting that most of the treatment had either been absorbed or metabolised. This reduction was not attributed to environmental wash-off, as plants were watered at soil level to minimise surface disturbance.

While Table 14 presents the proportion of deuterium labelled Trp in each sample type, these values alone do not quantify uptake. To estimate how much of the treatment was absorbed into the leaf, the percentage of internalised Trp was calculated by comparing the washed and unwashed leaf samples (equation 4.2). This derived value reflects the fraction of total L-Trp-d_3 that had entered the leaf tissue, excluded surface residue, and provided a clearer picture of uptake dynamics over time (Table 15).

Table 15: *Calculated internalisation percentages of deuterium-labelled L-tryptophan (L-Trp-d_3) in soybean leaves over time. Values represent the proportion of total L-Trp-d_3 absorbed into leaf tissue (washed sample) relative to the total L-Trp-d_3 present (unwashed sample) at each time point, as calculated using Equation 4.2. Data are averaged from three technical replicates (n=3) per time point, with standard error of the mean given in brackets.*

Time point	deuterated L-Trp in Unwashed Leaf (%)	deuterated L-Trp in Washed Leaf (%)	Internalised (%)
4 hours	58 (±3.35)	29 (±5.36)	50 (±9.67)
24 hours	46 (±2.00)	30 (±7.29)	65 (±16.10)
48 hours	47 (±3.01)	13 (±1.67)	28 (±3.98)
4 weeks	8 (±6.19)	0.39 (±0.26)	5 (±4.98)

To better visualise the uptake dynamics, the internalised percentage values from Table 15 were plotted over time (Fig. 44). This graph highlights the rapid absorption of L-Trp-d_3 within the first 24 hours, followed by a decline at 48 hours and minimal retention by 4 weeks. The trend supports the interpretation that redistribution and metabolic assimilation begin shortly after initial uptake.

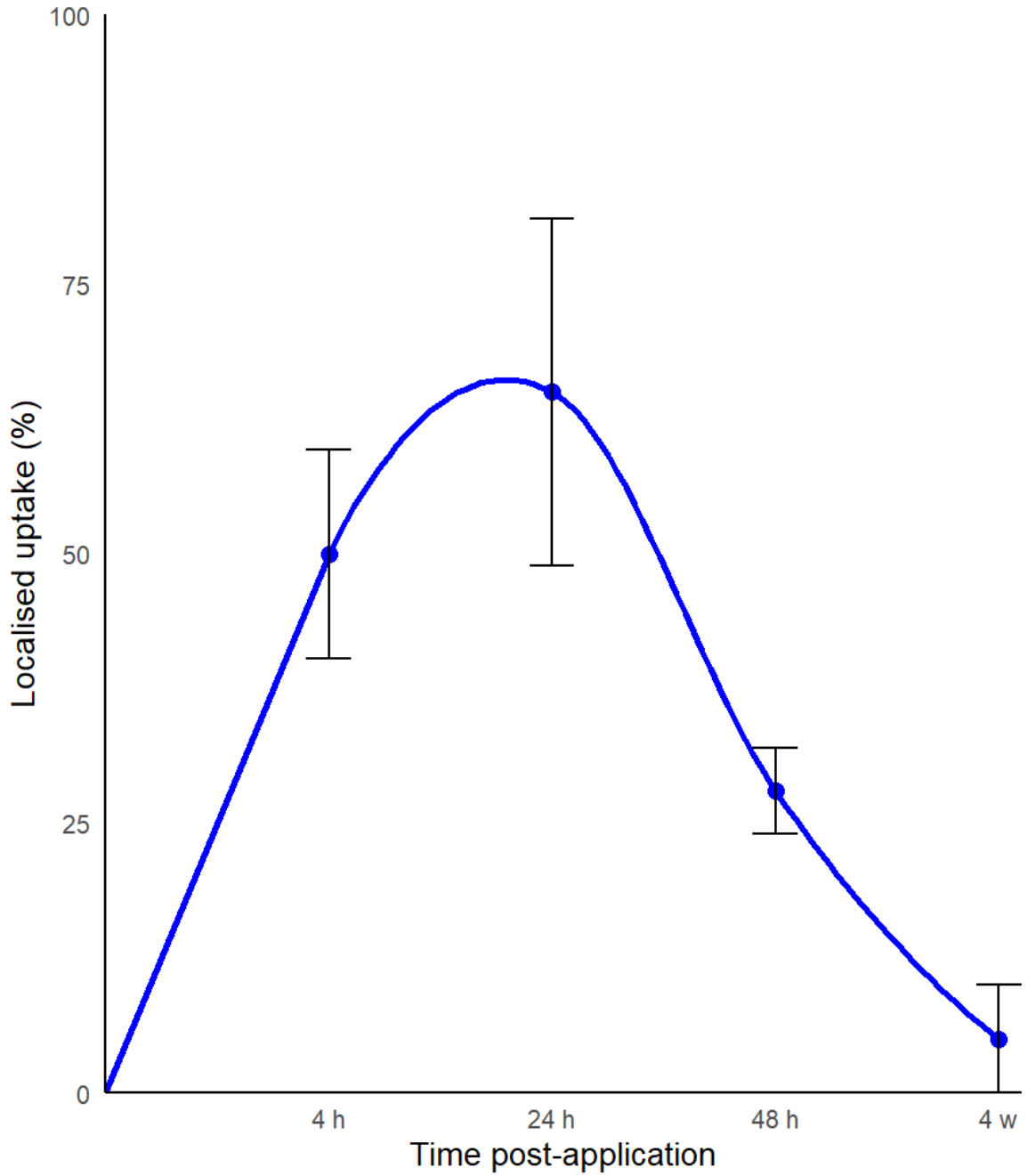


Figure 44: *Internalised deuterium-labelled L-tryptophan (L-Trp-d_3) in soybean leaves following foliar application. Values represent the percentage of total Trp-d_3 absorbed into the leaf (washed sample) relative to total L-Trp-d_3 present (unwashed sample) at each time point. Data are averaged from three technical replicates per time point. Error bars were calculated from SEM.*

The results (Table 15) show that initial uptake was rapid, with 50% of the treatment on the leaf having been absorbed into the leaf after 4 hours. This value had increased to 65% by 24 hours, showing strong uptake within the first 24 hours. At 48 hours, the percentage decreased to an average of 28%. This suggests a redistribution of the

absorbed L-Trp-d_3 within the plant and that this redistribution was occurring faster than continued uptake. After 4 weeks, the percentage was 5%. The overall levels of deuterated L-Trp in both washed and unwashed samples were very low compared to previous samples. This suggests that external L-Trp-d_3 had mostly been absorbed into the plant or removed by environmental factors and internalised treatment had mostly been metabolised or redistributed to other parts of the plant.

These findings suggest that the uptake of tryptophan occurs rapidly within the first 24 hours, with a significant portion absorbed early in the treatment period. The observed slowing of uptake after 24 hours suggests redistribution within the plant potentially through phloem transport or incorporation into metabolic pathways.

While this study provides insights into the uptake dynamics of tryptophan in soybean plants under controlled greenhouse conditions, several factors remain to be explored in future research. Variations to uptake due to environmental conditions such as temperature, humidity, light exposure and time of year were not assessed and could affect the use of this in general agricultural applications. As shown in the literature review (Section 1.3) the plant species used plays a significant role in determining how treatments interact with physiological processes. Therefore, broader research across multiple species is necessary to evaluate the general applicability of tryptophan as a biostimulant in diverse crop systems.

4.1.6 Effect of pH and enantiomers

The application of a highly acidic L-Trp solution resulted in a slightly reduced uptake compared to the neutral pH solution (46% vs 50% at 4 hours). However, the same calculation at 4 weeks showed very different results with 61% of deuterium labelled Trp at the site of application being internalised. This suggests that while initial uptake remains relatively unchanged, transportation and metabolism within the plant may be affected by strongly acidic conditions (Fig. 45).

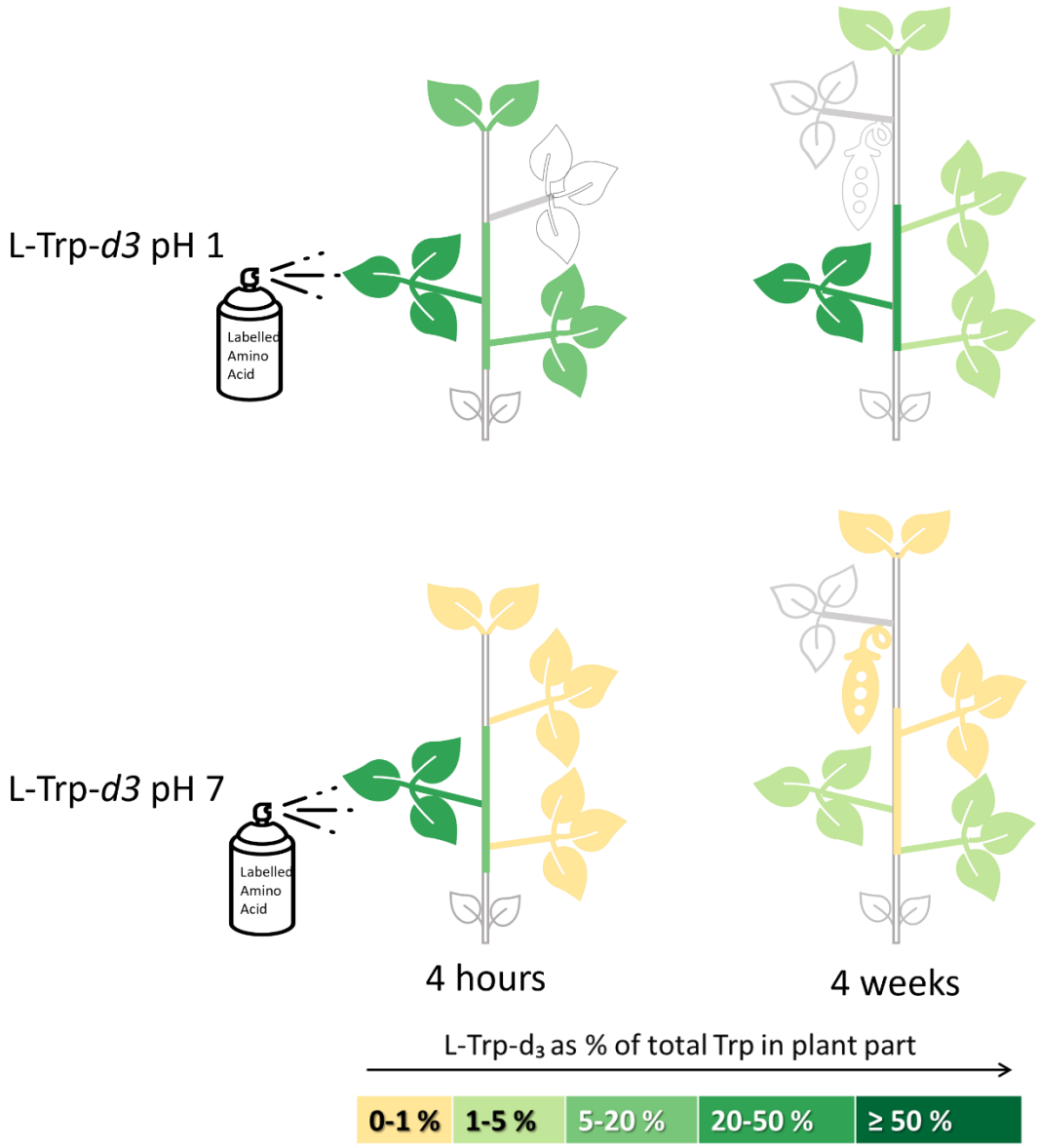


Figure 45: *Distribution of deuterium-labelled L-tryptophan (L-Trp-d_3) across plant tissues following foliar application at two pH levels: acidic (pH 1) and neutral (pH 7). Plant diagrams show the percentage of labelled Trp relative to total tryptophan in each tissue, enabling comparison of uptake and distribution under differing pH conditions. For all 4h L-Trp-d_3 pH 1 samples n=1, while for 4 weeks and for all L-Trp-d_3 pH 7 samples n=3, where n is a technical replicate. Colour legend: yellow (0-1%), yellow-green (1-5%), light green (5-20%), green (20-50%), dark green (≥50%).*

The large percentage of labelled tryptophan present at all parts after 4 weeks for the acid-treated plants suggests significant slowing in metabolism caused by the pH of the treatment. This may be as a result of the acidity causing stress to enzymes responsible for tryptophan metabolism, causing them to become less efficient or denatured.[190] It may also cause cellular damage such as increased oxidative stress,

disruption of cell wall integrity or membrane destabilisation.[191, 192] This would lead to the plant activating stress response pathways to mitigate the damage. This response uses energy which would otherwise be allocated to growth or metabolism.

The high percentage of deuterium labelled L-Trp remaining in the T2spot sample at 4 weeks (87%, Table 16) suggests poor internalisation under acidic conditions. At pH 1, tryptophan exists predominantly in a positively charged form (Trp^+), which may bind to the negatively charged surfaces of aqueous cuticular pores. This electrostatic interaction, combined with increased hydration and reduced lipophilicity, likely hinders translocation across the cuticle.[193]

This difference in uptake at varying pH levels (Table 16) indicates that pH had a significant influence on the absorption of exogenous tryptophan. In the initial experiments, residual trifluoracetic acid (TFA) was found to have unintentionally lowered the pH of the solution, which may have affected absorption and translocation. Although this issue was later resolved, comparing the outcomes from the acidic and neutral solutions sheds light on how pH impacts the movement of treatments within the plant for both initial uptake and distribution.

Table 16: *Average percentage of deuterium-labelled L-tryptophan (L-Trp-d_3) relative to total tryptophan in each sample from plants treated with an acidic (pH 1) foliar solution. Values represent the proportion of deuterated L-Trp detected in each tissue, enabling assessment of uptake and distribution following low-pH treatment. Data are averaged across replicates (where available), with SEM shown in brackets.*

Plant Part	Time Point	No. of replicates	Average% of deuterated Trp in total Trp
Meristem	4 hours	1	8
Stem	4 hours	1	11
T1	4 hours	1	13
T2spot	4 hours	1	95

T2uw	4 hours	1	77
T2w	4 hours	1	36
Meristem	4 weeks	3	5 (±1.14)
Stem	4 weeks	1	26
T1	4 weeks	3	2 (±0.13)
T2spot	4 weeks	3	87 (±1.23)
T2uw	4 weeks	3	18 (±0.70)
T2w	4 weeks	3	11 (±2.06)
T3	4 weeks	3	2 (±0.08)

D-Trp-d_4 at pH 7 showed impressive uptake at 4 hours with the comparison of the washed and unwashed samples suggesting an uptake of 81% at 4 hours (Table 17). This shows that uptake into the leaf is non-selective for enantiomers. However subsequent dispersal did not follow the same trends as L-Trp, with no deuterated material seen in the meristem or leaves above or below the site of application at 1 week. The absence of deuterated D-Trp in the meristem and distal tissues contrasts with the redistribution seen in L-Trp-d_3 treated plants. The lack of dispersal and amount of intact D-Trp-d_4 may reflect stereoselective transport and metabolism. Where enzymes and selective proteins are involved in the movement and metabolism of amino acids, the mechanisms are preferential for L-enantiomers, which are more common in nature and serve as precursors for key signalling molecules such as auxin.[190, 194]

Table 17: *Average percentage of deuterium-labelled D-tryptophan (D-Trp-d_4) relative to total tryptophan in each sample from plants treated with a neutral (pH 7) foliar solution. Values represent the proportion of labelled Trp detected in each tissue, enabling assessment of uptake and distribution following D-Trp-d_4 application. Data are averaged across three technical replicates, with SEM given in brackets.*

Plant Part	**Time Point**	**Average% of D-Trp-d_4 in total Trp**
Meristem	4 hours	2 (±1.12)
Stem	4 hours	6 (±2.65)
T1	4 hours	11 (±5.68)
T2uw	4 hours	66 (±4.21)
T2w	4 hours	54 (±2.99)
T3	4 hours	4 (±1.82)
Meristem	1 week	0
Pod	1 week	15 (±11.21)
Stem	1 week	29.35 (±8.66)
T1	1 week	0
T2uw	1 week	20 (±12.76)
T2w	1 week	20 (±13.63)
T3	1 week	0

One problem with the data from the D-Trp-d_4 treated plants was high standard deviation across all plant parts. This is a result of the samples from D-Trp-d_4 at pH 7 generally having weaker signals and so suffering more from increased noise compared to the L-Trp pH 7 samples which mostly had higher peak intensities (Fig. 46). These weaker signals were due to the poorer uptake and dispersal of the D-Trp-d_4 treatment by the plant resulting in less deuterated tryptophan in the sample to be detected.

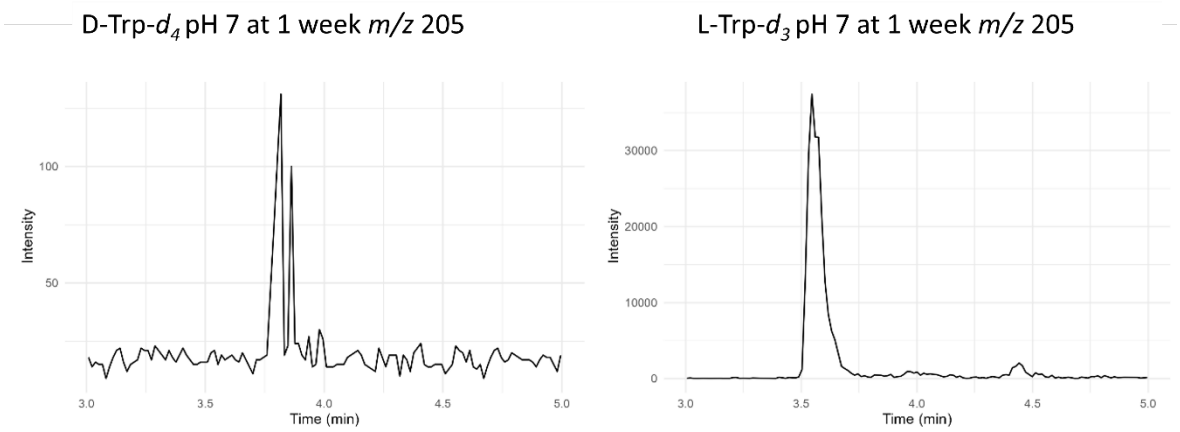


Figure 46: *MRM chromatograms of unlabelled tryptophan detected in the meristem of soybean plants one week after foliar application of deuterium-labelled L-tryptophan (L-Trp-d_3) and D-tryptophan (D-Trp-d_4), respectively. The chromatogram from the D-Trp-treated plant shows markedly lower signal intensity and increased baseline noise compared to the L-Trp-treated sample.*

Increasing sample concentration may reduce saturation effects and improve peak definition. This may also improve the comparability of samples.

4.2 Full-scan metabolite analysis

To investigate the broader metabolic effects of applied tryptophan treatments, a full-scan mass spectrometry approach was used to detect a wide range of compounds across plant tissues and time points. Unlike targeted MRM analysis, which focused on predefined ion transitions, full-scan analysis enabled the detection of both known and unexpected metabolites, including those formed through downstream metabolic processes. Deuterium labelling provided a distinct isotopic signature, allowing treatment-derived compounds to be distinguished from endogenous metabolites.

Samples were analysed using an Acquity UPLC (Waters Ltd, UK) coupled to a TQD with an electrospray ionisation (ESI) source, operating in both positive and negative ionisation modes. The use of both positive and negative ionisation modes maximised the detection of a wide range of plant metabolites with varying chemical properties. Chromatographic separation was performed on a Zorbax Eclipse XDB-C18 column, using a gradient elution of water with 0.1% (v/v) formic acid and methanol. The method was optimised for detecting a broad range of plant metabolites within a scan range of 100 - 700 *m/z* over a 16-minute run time.

Initial screening was performed using R scripts to identify m/z values unique to treated samples (Appendix D.3), cross-referenced against known tryptophan metabolites and plant hormones (Appendices C.1 and D.4). Extracted ion chromatograms (EICs) were generated for candidate compounds (Appendix D.5), but challenges such as inconsistent retention times and low signal intensity prompted a shift to manual validation using MZmine 4.

To support interpretation, total ion chromatograms (TICs) were overlaid using R (Appendix D.6) to visualise treatment-induced changes in chemical profiles. A separate script was used to explore auxin-related metabolites (Appendix D.7), given tryptophan's role as a precursor to indole-3-acetic acid. Blank samples and control comparisons were used to identify background peaks and minimise false positives (Table 18). Peaks occurring after this point were not included in the search as the column impurities distorted them. Similarly, comparing control samples produced a list of 13 common peaks across all plant parts and time points (Table 18). The presence of these peaks in the treatment samples could then be expected, and these peaks could be ruled out as belonging or being in response to the treatment.

Table 18: *Base peaks consistently identified in control samples across all plant parts and time points, identified as background signals. Peaks occurring after 5 minutes were excluded due to column-derived impurities observed in blank (methanol-only) samples.*

Base peak *m/z*	**RT (minutes)**	**Potential compound identity**
133	3.96	Aromatic fragment[195]
205	3.82	Unlabelled Trp

188	3.82	Possible Trp derived metabolite
115	3.1	Trp derived metabolite - possibly serotonin or tryptamine[196]
147	2.78	Possible kynurenine fragment[197]
166	2.62	
119	2.18	Possible kynurenic acid fragment
132	1.89 - 2.03	
175	1.59	Possible kynurenine fragment[197]
118	1.47	
104	1.29	
133	1.13 - 1.17	Aromatic fragment
132	0.73	

The combined approach of using automated screening and manual validation provided a framework for identifying treatment derived responses and assessing their potential role.

4.2.1 Initial metabolomics screening

Metabolomics provides a powerful tool for understanding biochemical changes within plants.[198, 199] Identifying specific metabolites within the samples, linked to the applied compound, such as sharing the isotope pattern, allows the identification of which chemical pathways are activated or influenced by the compound. The techniques used for identifying metabolites can also be applied to the search for plant growth regulators that may be activated in response to the presence of the treatment.

To identify treatment-derived metabolites, total ion chromatograms (TICs) were generated for both control and treatment samples across all plant parts and time points using an R script (Appendix D.6). Figure 47 provides a representative example of this analytical method, demonstrating how chromatograms were overlaid using R to visualise differences in chemical profiles. The R script allowed systematic comparison of chromatograms by highlighting differences in peak intensities between treatment and control samples. These initial observations guided further analysis in MZmine, where specific *m/z* values unique to treatment samples were identified, enabling a focused search for potential treatment-derived metabolites.[200]

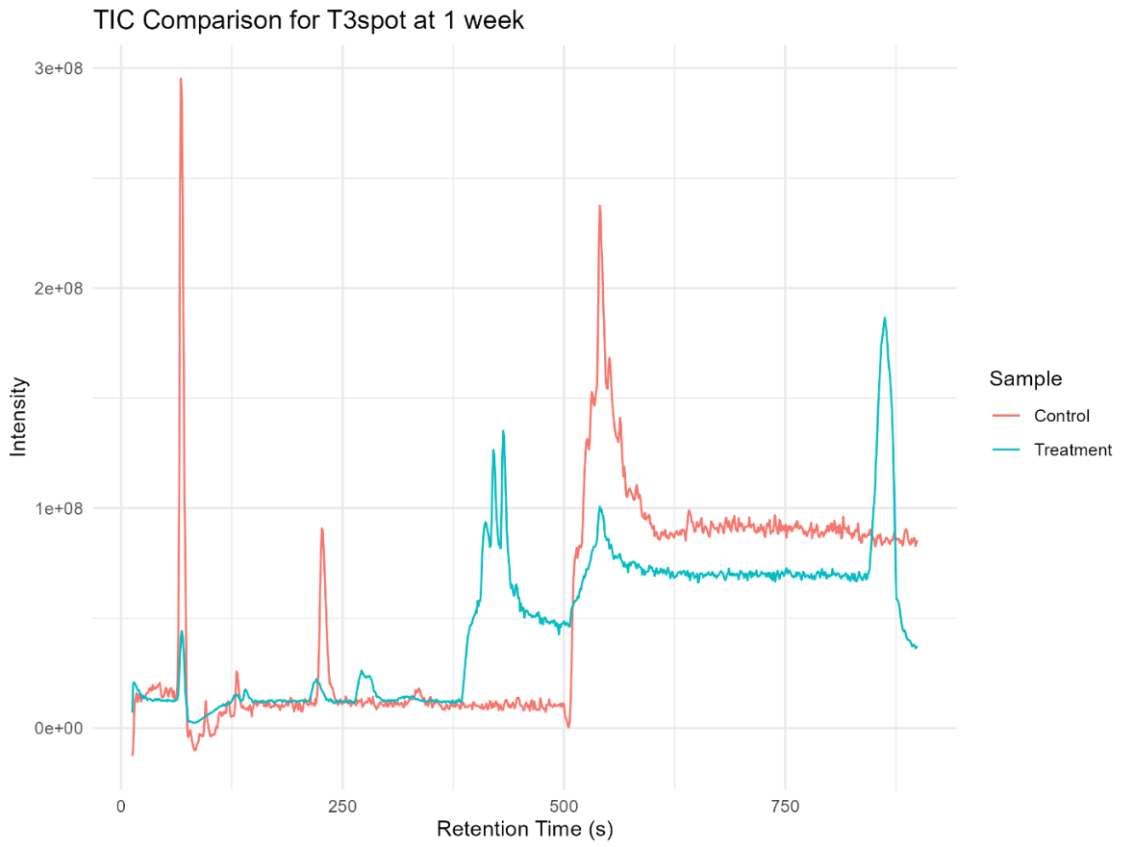
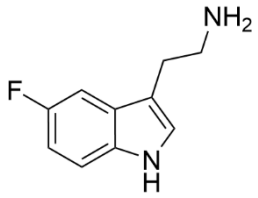
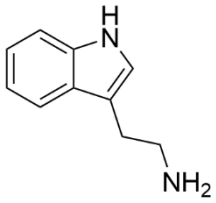
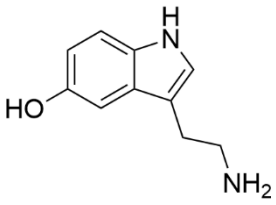
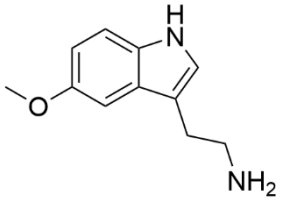
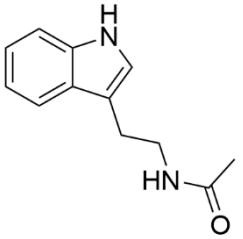
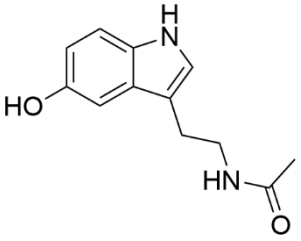


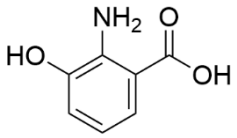
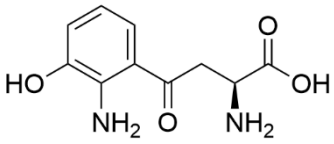
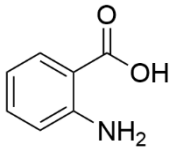
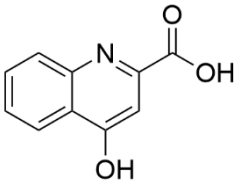
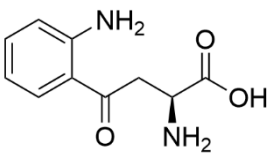
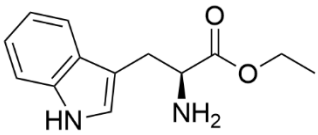
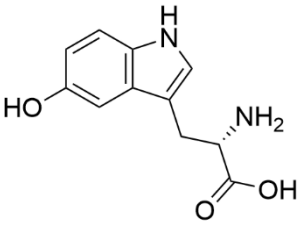
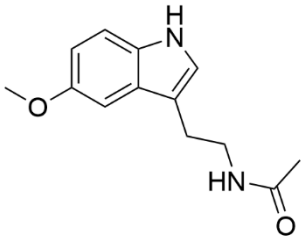
Figure 47: Overlay of liquid chromatography - mass spectrometry (LC-MS) chromatograms from leaf extract from site of application after one week of growth comparing L-Trp-d_3 (pH 1) and control treatments. Chromatograms were overlaid using R to visualise differences in chemical profiles, enabling systematic comparison of peak intensities.

An R script (Appendix D.3) was used to identify any *m/z* values in the treatment file but not in the control file from these overlays. Base peaks were then extracted for both positive and negative scans of these unique peaks.

The *m/z* values for these base peaks could then be compared to known *m/z* values for metabolites of tryptophan identified by Vitalini *et al.* (2020) as well as plant growth hormones (Appendix C.1 and D.4).[180] Potential tryptophan metabolites produced from the deuterium labelled treatment would have higher *m/z* values than those naturally occurring in plants, so a list of *m/z* values with which to compare unique peaks from the plant data was produced (Table 19). The centroid data was checked for the distinct isotopic pattern of deuterium-containing compounds.

Table 19: *Known tryptophan metabolites and plant growth hormones with their corresponding m/z values, associated biosynthetic enzymes, and expected m/z values for deuterium-labelled analogues.*

Metabolite	Structure	*m/z* $[M+H]^+$	Enzyme responsible	Anticipated *m/z* for deuterated compounds
5-F Tryptamine		179.1	Tryptophan decarboxylase (with substrate modification)	181.1
Tryptamine		161.1	Tryptophan decarboxylase	164.1
5-OH Tryptamine		177.1	Tryptophan hydroxylase	179.1
5-OCH3 Tryptamine		191.1	Hydroxyindole-O-methyltransferase	193.1
N-Ac Tryptamine		203.1	N-acetyltransferase	206.1
N-Ac-5-OH Tryptamine		219.1	N-acetyltransferase	221.1

3-OH Anthranilic acid		154.1	Kynureninase	155.1
3-OH Kynurenine		225.1	Kynurenine monooxygenase	226.1
Anthranilic acid		138.1	Kynureninase	140.1
Kynurenic acid		190.1	Kynurenine aminotransferase	192.1
Kynurenine		209.2	Tryptophan-2,3-dioxygenase	211.1
Tryptophan ethyl ester		233.1	Enzymatic modification (specific esterase or chemical modification)	236.1
5-OH Tryptophan		221.1	Tryptophan hydroxylase	223.1
Melatonin		233.1	Hydroxyindole-O-methyltransferase + N-acetyltransferase	235.1

To verify the authenticity of identified unique *m/z* values, the methodology involved comparing retention times and isotopic patterns with expected values for deuterated metabolites and known plant compounds. Blanks were incorporated into the analysis to minimise potential artefacts arising from noise or sample preparation. This reduced the risk of cross-contamination from the instrument. Replicate analyses were planned to assess the reproducibility of compound identification across biological replicates to ensure robust and reliable findings.

Several challenges arose from this approach. Many EICs showed inconsistent retention times across replicates and lacked clear isotope patterns. This suggests that these peaks were unlikely to represent the proposed compounds. Some peaks displayed low intensity (below 100,000 counts) making isotopic confirmation unreliable. Some detected compounds had unexpected fragmentation patterns which further complicated identification, particularly with compounds with similar m/z values. Due to these limitations in automated screening, a manual inspection using MZmine was undertaken to improve confidence in compound identification.

4.2.2 Compound identification

Following the initial screening and manual validation steps, several compounds were identified as potential treatment-derived metabolites.

Deuterated L-Trp was consistently identified in the full scan mass spectra of most of the samples where it had been identified from the targeted, more sensitive MRM approach. Its characteristic isotope pattern included a base peak at *m/z* 209 ($[M+H]^+$ for Trp-d_4), with 208 ($[M+H]^+$ for L-Trp-d_3) just slightly lower in intensity, and 210 and 207 also visible (Fig. 48). Fragment ions at m/z 191 and 165 likely correspond to $[M-H_2O]^+$ and $[M-COOH]^+$, respectively, consistent with known Trp fragmentation pathways.

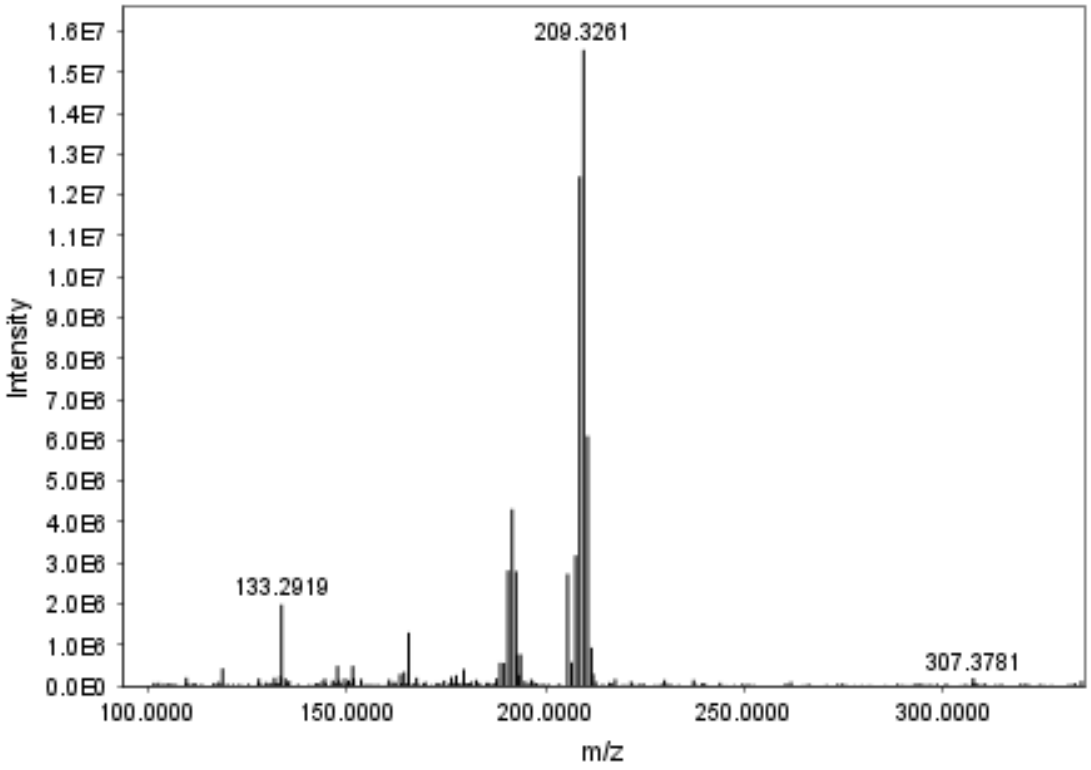


Figure 48: *Full scan mass spectrum of deuterium-labelled tryptophan (L-Trp-d_3/d_4) detected in soybean leaf tissue. The base peak at m/z 209 corresponds to $[M+H]^+$ for L-Trp-d_4, with adjacent isotope peaks at m/z 208 (Trp-d_3), 210, and 207 also visible. Fragment ions include m/z 191, likely representing $[M-H_2O]^+$ from loss of water at the carboxylic acid group.*

In addition to the applied compound, several other peaks were observed that may represent metabolites derived from Trp or compounds influenced by its presence. These were assessed based on their *m/z* values, isotope patterns, retention times, and absence from control samples.

4.2.2.1 Potential metabolite (*m/z* 220.33)

A potential metabolite was found on the outside of the leaf at the site of application in one sample at 72 hours. This compound has the highest intensity peak at *m/z* of 220.33. With the isotopic pattern seen in the chromatogram (Fig. 49) the presence of deuterium in the compound is very likely. The presence of the co-eluting peak with *m/z* 133.29 also suggests the compound is a tryptophan derivative. While these peaks could be fragments of the main peak, it is also possible that these are just compounds or parts of compounds which elute at the same time as the deuterated sample. The lack of deuterium isotope patterns for the two co-eluting compounds

suggests these may not be related to the largest intensity peak in the chromatogram, or alternatively, that during fragmentation, deuterium was lost from the resulting fragment ions. The loss of 55 (220->165) in the fragmentation could indicate the loss of a fragment containing an N-methyl group.

From a search of tryptophan derived metabolites with *m/z* values around 218 (for an undeuterated version), a few potential compounds were found. The same *m/z* value has been identified as a major fragment of Indole-3-carboxylic acid (3ICA) in mass spectrometry.[201] This is a precursor for synthesis of IAA. Another potential candidate is the auxin 2,4-Dichlorophenoxyacetic acid which again has been shown to appear in mass spectra as *m/z* 218.[202] N-Ac-5-OH Tryptamine has an undeuterated *m/z* value of 219 and is a known metabolite of tryptophan, making it another possibility for the compound this spectrum came from.

Two of these compounds are related to or involved in the synthesis of auxins. However, as this was found on the external surface of the plant only, it is possible that this was formed due to prolonged exposure to environmental factors while not entering the leaf. IAA synthesis has been found in a number of different bacteria as well as plants, so this could also be caused as a response to the plant's microbiome, rather than internal effects.[203]

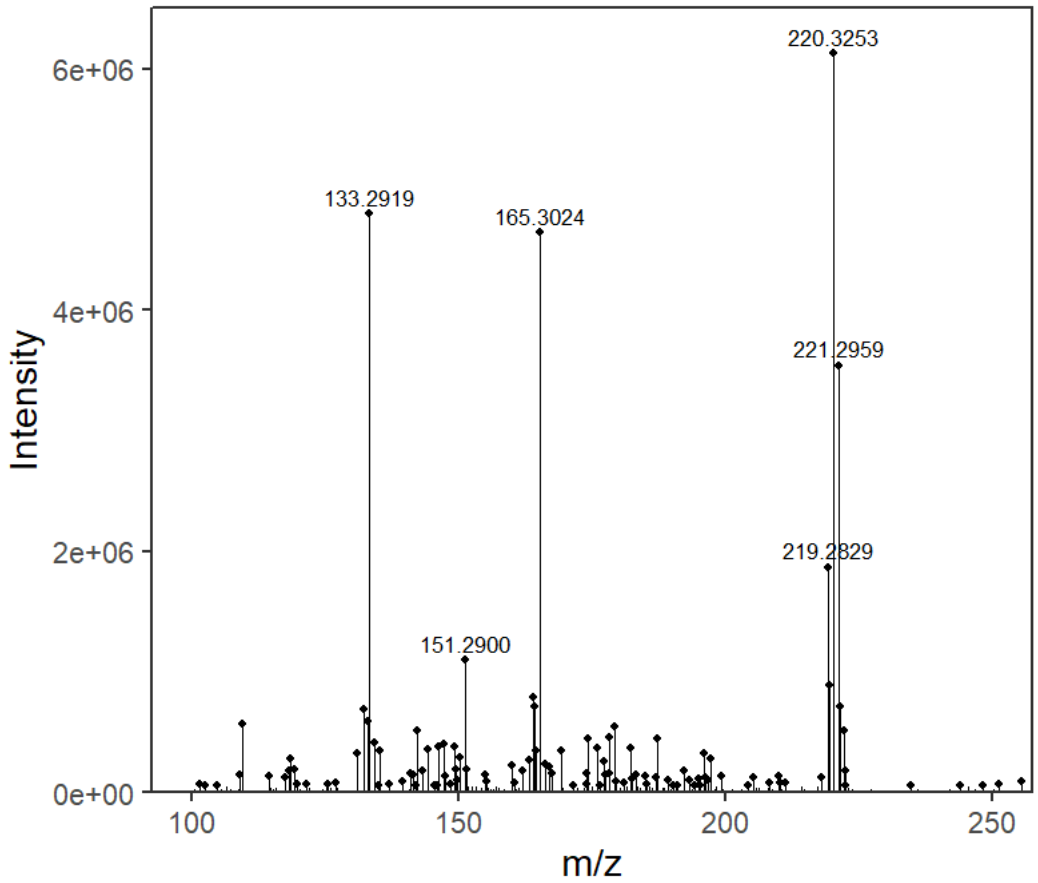


Figure 49: *EIC of a compound with a base peak at m/z 220, detected in soybean leaf tissue following foliar application of deuterated tryptophan (L-Trp-d_3/d_4). The isotope pattern surrounding this peak is consistent with the presence of deuterium. This signal was absent in control samples, supporting its identification as a treatment-derived metabolite.*

Although this metabolite is interesting, it was only found in one sample, so it does not provide robust evidence of the effect of the deuterium labelled L-Trp. Further studies specifically looking for this compound would be needed to draw any significant conclusions from its presence. It may also be possible that this compound could be found in other samples, but the concentration is too low to see it using the current analytical technique. Notably, the sample where this metabolite was identified had a different preparation method than the other plant parts, as it was explicitly intended to demonstrate the efficacy of the washing technique in removing the treatment from the external leaf surface. This sample consisted of a droplet of aqueous hydrochloric acid solution, which was applied to the leaf surface for 10 minutes and then transferred to a mass spectrometry vial with an insert to accommodate smaller volumes. As a result, this sample was more concentrated than those from other plant

parts. Consequently, it is possible that if more concentrated extracts of other samples were analysed, small amounts of the deuterated compound might also be detected. In order to accurately identify this compound, the potential candidates would need to be run under the same experimental conditions to observe if they have the same retention time and fragmentation pattern.

4.2.2.2 Reduction in unlabelled tryptophan

In all three stem sample replicates for 72 hours post-application, there was an absence of peaks (Fig. 50) with *m/z* values of 205 (Trp), 115 and 119 despite the presence of these peaks in two of the three control samples, with the third anomalously showing the same peaks as the treatment samples. This suggests a potentially localised metabolic effect driven by the treatment.

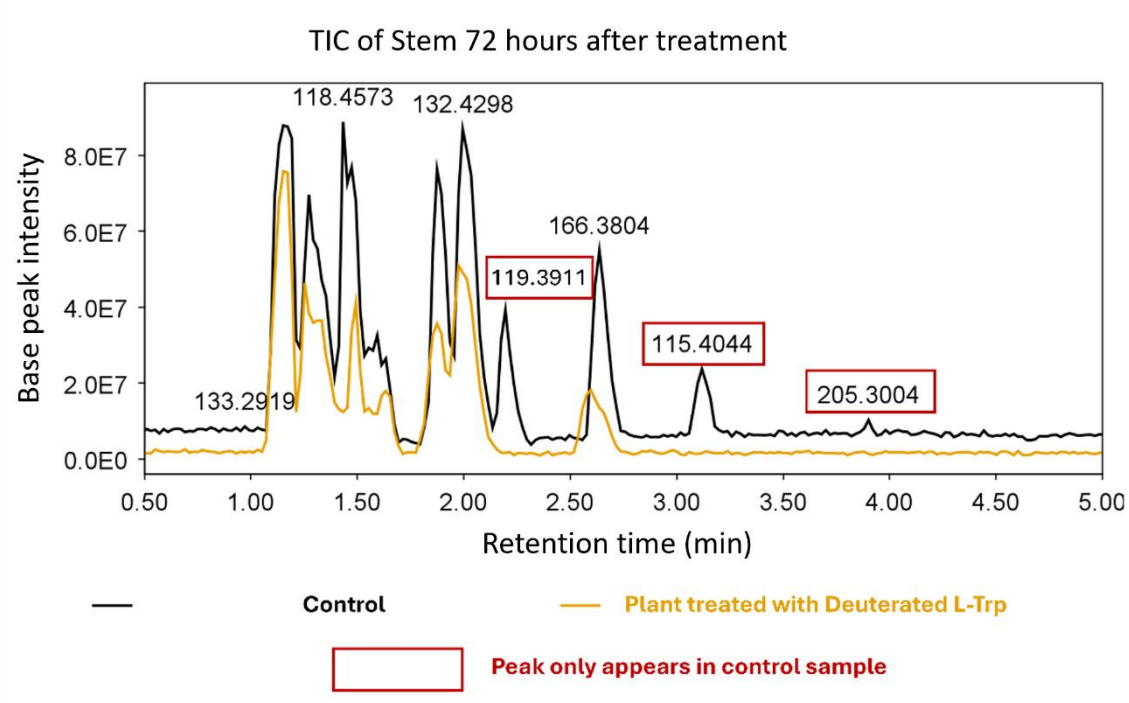


Figure 50: *Overlayed total ion chromatogram (TIC) of stem extracts from soybean plants 72 hours after foliar application of deuterium-labelled L-tryptophan (L-Trp-d_3) generated using R. The black trace represents the control sample, and the yellow trace represents the treated sample. Several peaks present only in the control (highlighted in red) were absent from the treated sample, suggesting treatment-induced suppression or metabolic alteration.*

The absence of these peaks suggests that the presence of the exogenous L-Trp has some suppressing effect in the stem. The missing peak with the highest intensity *m/z* value of 115 is likely a fragment from a Trp-derived secondary metabolite such as tryptamine or serotonin.

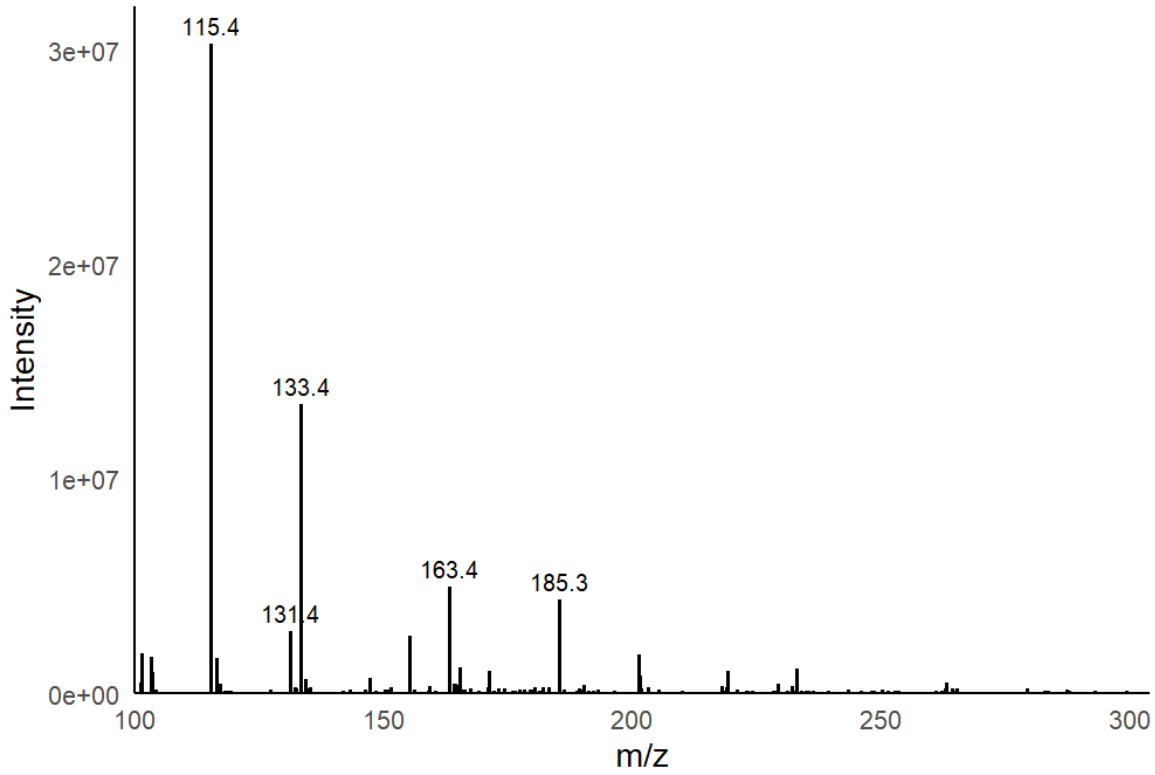


Figure 51: *EIC from control soybean stem samples showing a potential tryptophan-derived metabolite eluting at 3.1 minutes, with a base peak at m/z 115. This signal was consistently observed in control samples and absent or suppressed in treated samples, suggesting it may represent an endogenous compound downregulated or displaced by deuterated Trp application.*

The observed fragmentation pattern (Fig. 51) suggests that this compound contains an indole structure which is likely derived from Trp. In positive electron spray ionisation (ESI), tryptamine ($[M+H]^+$ = 161) and indole-containing structures are known to form indole-based fragment ions around *m/z* 117-119 due to cleavage at the C_α-N bond to give an indole cation. A peak with *m/z* of around 133 also suggests an indole fragment.

The chromatogram for the peak with a 2.18 min retention time showed the strongest intensity for *m/z* 119 (Fig. 52). This is also likely a Trp-derived metabolite. The most

likely candidate is kynurenic acid (KYNA), as it has fragment ions of *m/z* 161 and 189, which are consistent with known KYNA fragmentation patterns. While *m/z* 161 could also originate from tryptamine, the absence of other characteristic tryptamine fragments makes KYNA the stronger candidate for this peak. The fragment at *m/z* 101 suggests a loss of NH_3 or CO from the 119 base peak.

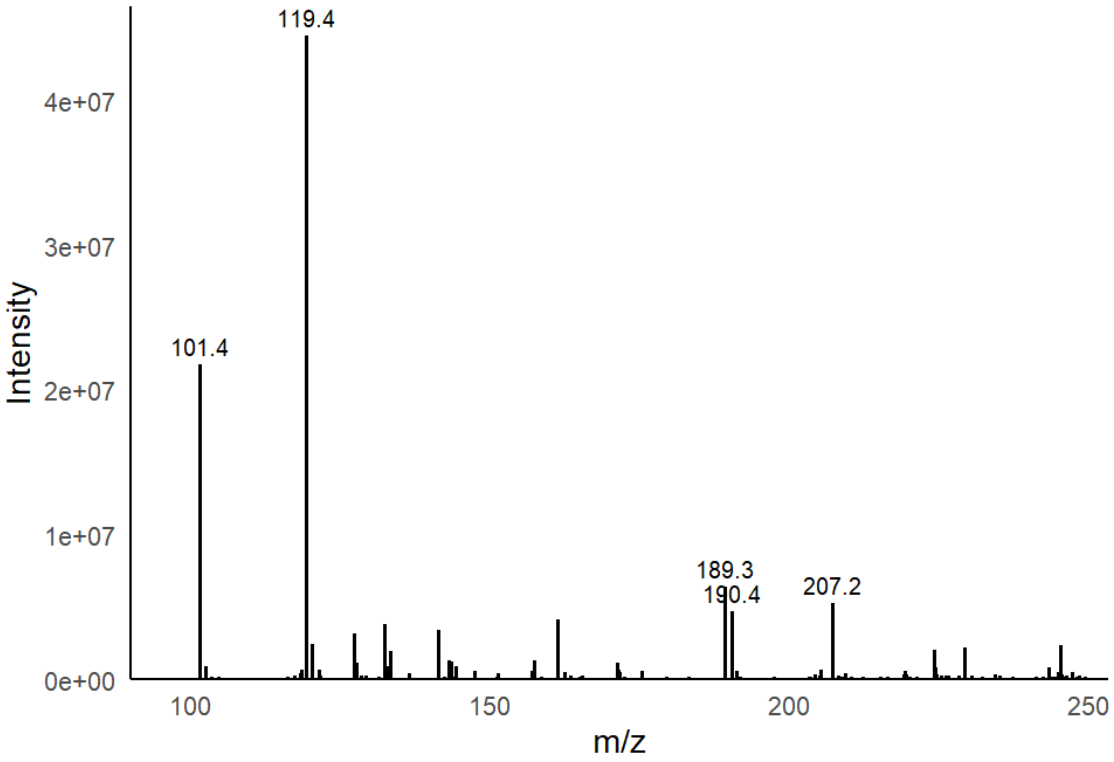


Figure 52: *EIC from control soybean stem samples at 72 hours post-treatment, showing a compound eluting at 2.18 minutes.*
This signal was observed consistently in control samples but not in treated samples.

The observed reductions in Trp and its metabolites in these stem samples may result from metabolic or transport-related effects. As both deuterium labelled and unlabelled Trp are missing from the samples, it appears that the treatment is not affecting endogenous Trp synthesis and uptake through competitive inhibition or regulatory feedback but may instead be altering its metabolism, distribution or transport within the plant. One possibility is metabolic reallocation, where the treatment causes the promotion of some metabolic pathways over others. In this case, protein synthesis or auxin biosynthesis may be favoured over the kynurenine

and indoleamine pathways. Protein synthesis is the more likely action as these large molecules would not be extracted with the current extraction method, due to the focus on free amino acids and small molecules.

Alternatively, the absence of these compounds from the stem could be related to a disruption in phloem transport. If the treatment modified phloem loading it could prevent the movement of Trp derivatives from source tissues. Changes in auxin biosynthesis could affect phloem loading as auxins regulate phloem function.[204] Therefore, if the treatment was up-regulating auxin production, there could be a reduction in general distribution of Trp as a nitrogen source through the plant, with distribution focused on younger plant tissues.

Interestingly, both deuterium labelled and unlabelled Trp were still detected in MRM scans, indicating that these amino acids were not completely absent but likely present at concentrations below the detection threshold. Further quantitative analysis would be required to confirm whether exogenous L-Trp application affects endogenous amino acid metabolism.

4.2.2.3 Auxin

In an initial experiment in which deuterated L-Trp was applied to leaves at 30 mM concentration at around pH 1, a new compound was observed at the end of the spectrum for the site of application, 24 hours after application. This was identified as potentially being 4-chloro-indole-3-acetic acid due to the *m/z* value (*m/z* 209 = $[M]^+$), fragmentation pattern (*m/z* 164 = $[M\text{-}COOH]^+$) and an isotopic pattern showing the presence of chlorine (Fig. 53). The lack of deuterium shifting in the isotope pattern shows that this is not produced as a metabolite of the deuterium labelled L-Trp treatment and was formed from endogenous tryptophan.

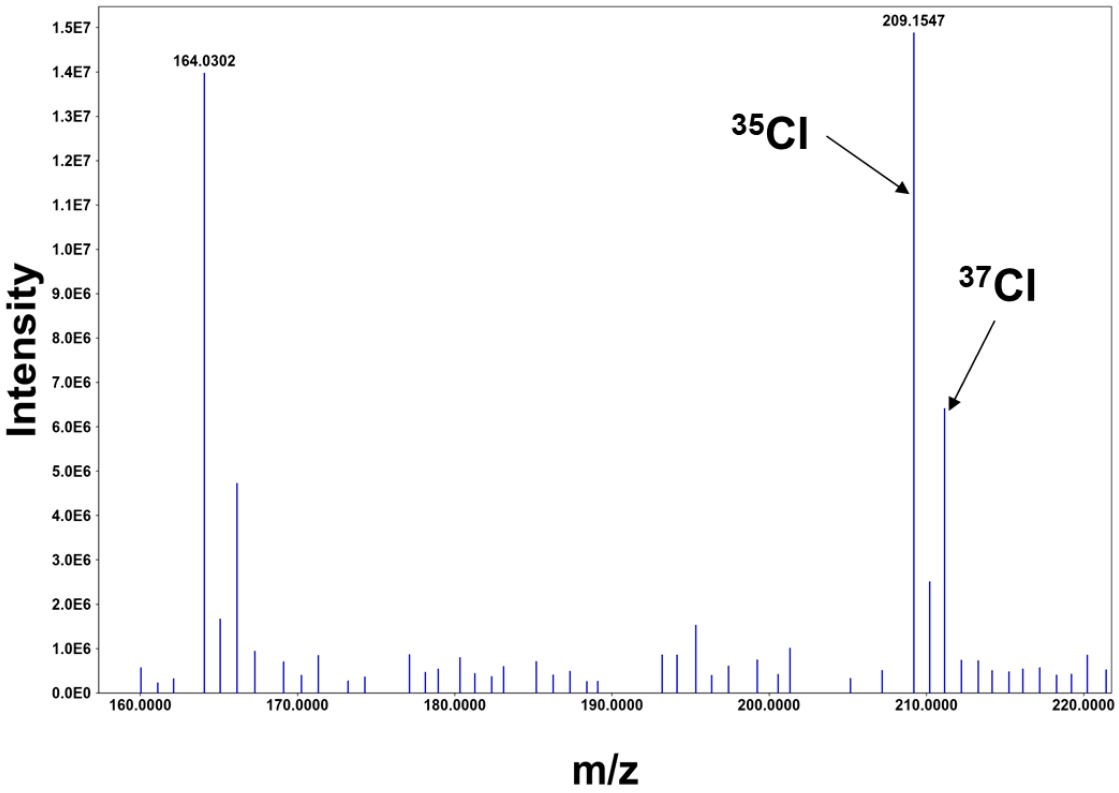


Figure 53: *Mass spectrum of a compound detected in soybean leaf extract from the unwashed site of application 24 hours after foliar treatment with deuterium-labelled tryptophan (L-Trp-d_3/d_4). The spectrum shows a distinct isotopic pattern consistent with deuterium incorporation, supporting identification of the compound as treatment-derived.*

Not much is known about the biosynthesis of 4-chloro-indole-3-acetic acid in plants but, as an auxin, it is known to promote growth.[205, 206] Auxins are produced at the apical meristems of plants as well as at enlarging plant organs.[207] Attempts to find this auxin in other samples have been unsuccessful. This suggests that rather than being in response to the treatment itself, this auxin may be present due to the use of a highly acidic solution when applying the treatment.

4.3 Physiological and growth responses to treatment

Assessing physiological and growth differences between treatment and control is essential for understanding the treatment's impact on plant health and development. In addition to monitoring biochemical changes, physical parameters such as dry mass, stem length, pod production, and chlorophyll content provide tangible indicators of plant vitality and yield. Growth metrics reflect overall health, while

chlorophyll content offers insight into photosynthetic capacity and physiological status. While the initial hypothesis was that neutral pH L-Trp treatment would enhance growth and pod production, the literature indicates considerable variability in amino acid effects across plant species and treatment conditions.

4.3.1 Chlorophyll soil plant analysis development (SPAD) data

One common theme within the research literature was the observation of increased chlorophyll after foliar application of amino acids.

These measurements were taken using a Soil Plant Analysis Development (SPAD) meter, which compares how much red and infrared light a leaf absorbs to measure chlorophyll content within the leaf.[208] This non-destructive method allows rapid in situ assessment of the chlorophyll concentration within a leaf. If the application of Trp improved growth through increased photosynthesis, this would be observed through increased levels of chlorophyll compared to untreated plants.

Initial measurements of chlorophyll levels, obtained using a SPAD meter, found a significant rise in chlorophyll levels in plants treated with L-Trp-d_3 after 1 week. However, after 4 weeks, a difference was not observed (Fig. 54). This pattern suggested that the treatment may have initially stimulated chlorophyll production, potentially improving plant health and yield, but that this effect diminished as the treatment was metabolised.

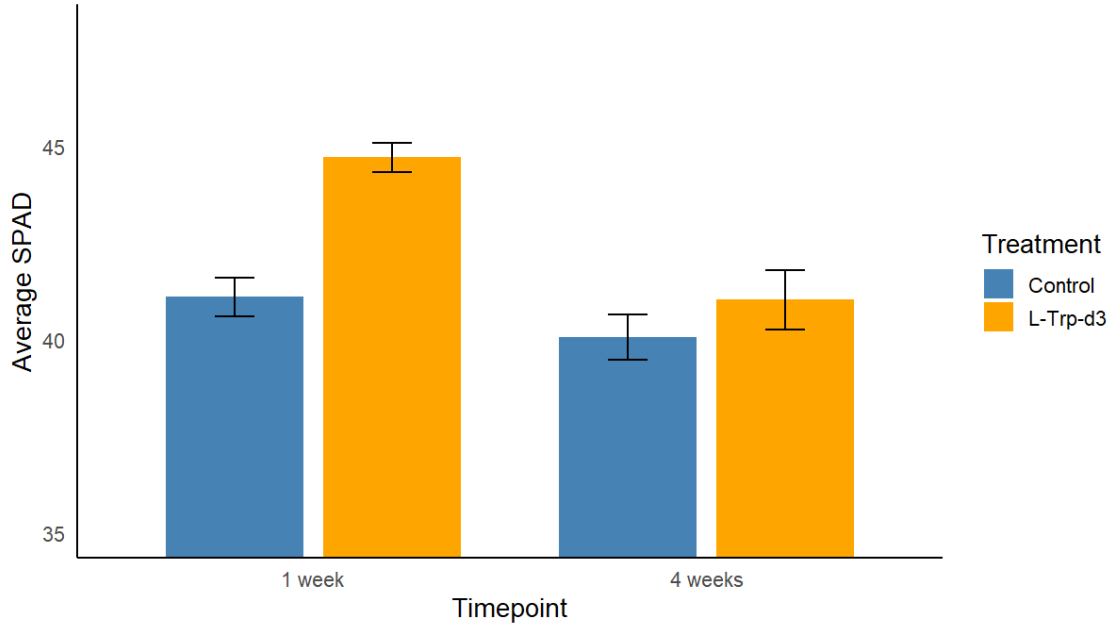


Figure 54: *Average SPAD values for soybean plants treated with deuterium-labelled L-tryptophan (L-Trp-d_3) compared to untreated control plants. SPAD readings were used as a measure of chlorophyll content and leaf greenness. Means were calculated for four technical replicates per timepoint and treatment (n=4). Error bars represent the standard error of the mean.*

Due to these findings, a follow-up study was conducted to replicate and further investigate these results, with more time points and differently labelled Trp derivatives included (Fig. 55). Data for 24 hours and 1 week after application was not included in the analysis. This was due to incomplete data provided by the industrial partner. Although SPAD data for all treatments at all timepoints was requested, for 24 hours only control was provided, and for 1 week control and L-Trp-d_3 were provided. As full data could not be shown for the other treatments, only timepoints with complete data for all treatments were analysed. In contrast to the initial study, no significant difference was seen between the control and the treated plants.

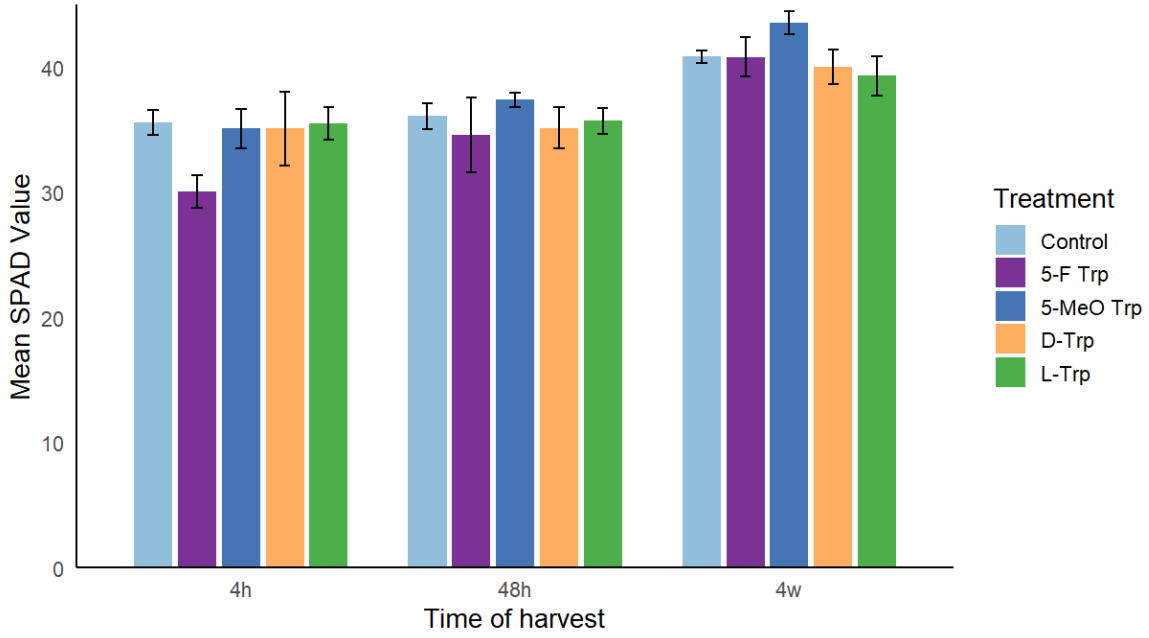


Figure 55: *Average SPAD values for soybean plants treated with tryptophan derivatives at 4 hours, 48 hours, and 4 weeks after foliar application. No significant differences were observed between treated and control plants at any time point. Error bars show the standard error of the mean, where three technical replicates were used for each mean (n=3).*

Statistical analyses were performed using Python to assess differences within the data. These tests included the Shapiro-Wilk test to confirm normality, Levene's test to assess homogeneity of variances, independent t-tests to compare mean SPAD values between treatments and controls, and a one-way analysis of variance (ANOVA) to validate the t-test results and analyse overall differences across treatments. All tests indicated no statistically significant differences between control and treated plants.

The absence of the 1-week measurements in the follow-up study raises further questions about the observed increase in chlorophyll content after 1 week in the initial study. It remains possible that chlorophyll levels transiently increase within the period between 48 hours and 1 week, before returning to baseline levels by 4 weeks. Future work should prioritise more frequent sampling during this period to better characterise any potential transient chlorophyll increase.

4.3.2 Effect on plant growth

In addition to chlorophyll content, plant biomass and pod production were measured to assess the impact of treatment on vegetative growth and yield.

4.3.2.1 Mass comparisons

Throughout the experiment, no significant differences in mass were observed at most time points between control plants and those treated with the neutral L-Trp spray. At the 4-hour, 24-hour, 72-hour, and 1-week time points, t-tests revealed no statistically significant differences in average mass ($p > 0.05$), indicating that the treatment did not have a measurable impact on growth during these stages (Table 20).

Table 20: *Summary of plant mass measurements across treatment groups and timepoints. Values represent the mean ± SEM, with n values stated for each average. n refers to the number of technical replicates; when n > 3, this indicates that measurements from a second biological replicate were included.*

Group	**4 hours**	**24 hours**	**48 hours**	**72 hours**	**1 week**	**4 Weeks**
Control	1.74 (±0.31) (n=4)	2.33 (±0.29) (n=7)	2.86 (±0.14) (n=3)	3.75 (±0.10) (n=3)	5.44 (±0.10) (n=3)	15.58 (±0.73) (n=3)
L-Trp	2.21 (±0.087) (n=4)	2.45 (±0.2) (n=5)	3.35 (±0.1) (n=3)	3.93 (±0.1) (n=3)	5.52 (±0.95) (n=2)	11.91 (±1.05) (n=3)
D-Trp	1.93 (±0.18) (n=3)	-	-	-	4.01 (±0.16) (n=2)	6.70 (±0.22) (n=2)
L-Trp pH 1	2.00 (±0.18) (n=3)	-	-	-	-	7.02 (±0.29) (n=3)

5-F Trp	2.12 (±0.04) (n=3)	-	-	-	3.91 (±0.31) (n=3)	12.89 (±0.51) (n=3)

At the 48-hour timepoint, a two-sample t-test indicated a borderline significant increase in mass for the L-Trp treated plants compared to the control ($p = 0.051$, two-tailed), suggesting a potential short-term growth response to the treatment that warrants further investigation (see Appendix F for full statistical data).

At the 4-week timepoint, a two-sample t-test assuming equal variances revealed a statistically significant difference in plant mass between the control and plants treated with the neutral L-Trp spray ($p = 0.046$, two-tailed). Control plants exhibited a higher average mass (15.58 ± 0.73 g) than treated plants (11.91 ± 1.05 g), suggesting that the treatment may have reduced biomass accumulation over this period.

Following this unexpected mass difference at 4 weeks, additional analyses were performed to compare mass differences across a broader range of Trp derivatives. A one-way ANOVA (Appendix F.6) revealed significant differences between the mass of plants treated with L-Trp pH 7, D-Trp pH 7, L-Trp pH 1, and 5-F Trp, alongside the untreated control group ($F(4,9) = 30.05$, $p = 3.22e^{-5}$). Pairwise comparisons indicated substantial mass reductions for D-Trp (6.70 g) and L-Trp pH 1 (7.02 g) relative to the control (15.58 g). In contrast, 5-F Trp (12.89 g) exhibited only a moderate reduction compared to the control (Fig. 56). This suggests that structural or pH modifications can influence the plant's growth response to Trp derivatives. These findings highlight the potential role of stereochemistry and environmental pH in modulating physiological responses and warrant further investigation.

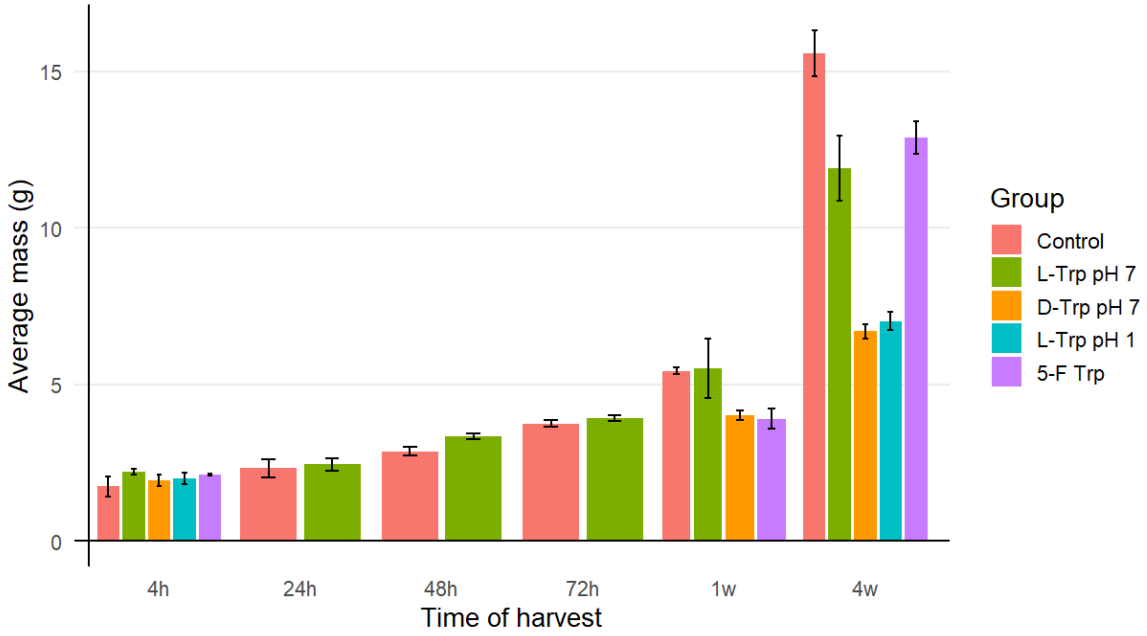


Figure 56: *Average plant mass (± SEM) at selected timepoints across treatment groups. Means were calculated from technical replicates as listed in Table 20.*

4.3.2.2 Pod production

Pod production was assessed at the 1-week and 4-week time points to evaluate potential treatment effects on yield (Table 21). At 1 week, plants treated with L-Trp exhibited a slightly higher average pod number than the control group (4 vs 2.5, respectively); however, a t-test assuming unequal variances found no statistically significant difference between the groups (p = 0.408, two-tailed).

Table 21: *Mean number of pods observed in soybean plants treated with deuterium-labelled L-tryptophan (L-Trp-d_3) and in untreated control plants at 1 and 4 weeks after foliar application. The mean was calculated from technical replicates (n), with SEM and n value provided in brackets for each timepoint and treatment.*

Treatment	Time point	Mean number of pods
Control	1 week	2.5 (±0.4) (n=3)
L-Trp pH 7	1 week	4 (± 1.0) (n=2)
Control	4 weeks	13 (±3) (n=2)

L-Trp pH 7	4 weeks	12 (±2.1) (n=3)

At the 4-week timepoint, initial results showed a slightly higher pod number in L-Trp-treated plants compared to the control (12 vs 9, respectively). Still, the difference was again not statistically significant ($p = 0.579$, two-tailed). Upon removal of an anomalous data point, the analysis was repeated, resulting in revised means of 13 for control plants at 4 weeks. A t-test assuming equal variances confirmed that the difference between treatments remained non-significant ($p = 0.793$, two-tailed).

These results suggest that L-Trp treatment did not substantially affect pod production over the experimental period. Observed variations may reflect natural variability rather than a treatment effect. Future work could explore larger sample sizes to provide greater statistical power to detect subtle differences in reproductive output.

Although the results from both mass and pod production comparisons did not align with the initial hypothesis of increased growth following L-Trp treatment, it remains possible that the treatment may have conferred physiological benefits not reflected in mass or pod production, such as improved chlorophyll levels. This possibility was explored in the following section.

4.4 5-Methoxy and 5-Fluoro-tryptophan

MRM measurements were collected for plants treated with 5-fluorotryptophan (5-F-Trp). However, the intensity of naturally occurring Trp was not recorded alongside these measurements. As a result of this, the presence of the treatments could be confirmed, but the percentage relative to total Trp could not be calculated due to the lack of an internal standard for these samples.

Despite these limitations, the data (Appendix E.1.f) indicates that at 1 week, 5-F Trp was present at the site of application (both washed and unwashed) in stem samples and in one meristem sample. At 4 weeks, peaks above the noise were observed again in the washed and unwashed samples at the site of application and in one replicate, it was seen in the leaf below the site of application.

4.5 Summary of results

This chapter presented the findings on uptake, distribution, metabolism and physiological effects of applying a foliar spray of labelled tryptophan to soybean plants. Initial uptake was rapid with 50% of treatment being internalised within 4 hours of application. Redistribution was detected in most plant parts, with levels decreasing after 48 hours. This showed that metabolism of the free amino acid treatment began between 24 and 48 hours after application.
Metabolomics analysis revealed a potential treatment-derived metabolite with isotopic labelling, though identification of the compound was not successful. Reductions in some naturally occurring tryptophan derived compounds shows that the treatment causes some effect on transport or metabolic processes within the plant. Potential auxin presence, inferred from metabolite screening and isotopic patterns, could indicate a hormone response to the treatment.

The use of D-Trp and acidic conditions reduced movement of the treatment within the plant. Physiological measurements showed little difference between control samples and L-Trp pH 7 samples, but some changes at 1 week were seen when chlorophyll levels were measured. Together, these results demonstrate clear uptake and metabolic processing of applied Trp, with preliminary evidence of downstream physiological and hormonal effects.

Chapter 5: Conclusions and future work

This thesis set out to investigate the fate of amino acids following foliar application to plants, with a particular focus on tryptophan. While the original aim included understanding mechanisms that might influence growth and plant health, the treatments applied in this study did not produce measurable changes in either parameter. As such, no conclusions can be drawn about the agricultural efficacy of foliar-applied Trp at the doses used, and recommendations regarding application frequency or formulation would be premature. However, the study yielded important insights into the uptake, retention, and movement of Trp within the plant, and highlighted key limitations in how such data can be interpreted. These findings contribute to a broader understanding of amino acid behaviour in plants and offer a foundation for future work exploring physiological relevance under different conditions.

Both L- and D-Trp were rapidly taken up after foliar application, with over 60% of the applied treatment absorbed into the leaf within the first 24 hours. Rapid uptake is important for foliar treatments, as sprays are vulnerable to wash-off by rainfall within the first 24-48 hours.[189]

Key insights came from tracking the movement of labelled amino acids using MRM mass spectrometry. Within 4 hours of application, Trp had already translocated into the stem and was subsequently distributed to all major plant tissues. Metabolisation appeared to begin around 48 hours post-application. While no physiological effects were observed, this rapid redistribution suggests that sustained presence may require repeated application, if future studies confirm bioactivity at higher doses.

The detection of a deuterated potential auxin precursor alongside an unlabelled auxin in different samples suggests that Trp may influence auxin-related pathways in soybean. Additionally, the reduction in certain peaks observed in the stem at 72 hours after treatment may indicate altered compound movement or a suppressive effect on endogenous metabolite production. Identification of these compounds would be necessary to further understand the underlying cellular responses.

Physiological assessment revealed no consistent differences between treated and control plants in terms of biomass, pod number, or chlorophyll content after 4 weeks. One experiment showed a significant increase in chlorophyll at 1 week after treatment. However, the 1-week result was missing from the repeat of this experiment so further conclusions cannot be drawn in this project regarding physiological impact.

The original scope of this project encompassed a broader survey of amino acid uptake, metabolism, and physiological impact. However, due to disruptions caused by the COVID-19 pandemic and subsequent health-related challenges, the study was narrowed to focus on tryptophan as the primary target.

Overall, this thesis has provided insight into the uptake, distribution and potential metabolic pathways of foliar-applied tryptophan in soybeans. By demonstrating rapid absorption and movement within the plant as well as identifying potential links to auxin production, this work has laid solid groundwork for the development of novel bioactive plant treatments.

Through the synthesis of a range of labelled tryptophan analogues and the novel synthesis of monomer and dimer indole derivatives, this research not only advanced the primary aim of understanding amino acid uptake in plants but also demonstrated a new method for selective synthesis addressing a significant gap in the existing literature. The dimer in particular holds promise for future bioactivity structures due to its structural similarity to the backbone of rebeccamycin, a compound known for its therapeutic potential. Future work is required to refine the methodology to see improved yields for this application.

By integrating the principles of green chemistry, this research aligns with the broader goal of developing environmentally responsible agricultural treatments. The combined use of full-scan and targeted MRM mass spectrometry provided a comprehensive view of tryptophan bioavailability and potential modes of action.

5.1 Summary of aims and objectives

The overarching aim of this thesis was to investigate the uptake, metabolism, and physiological effects of foliar-applied tryptophan in soybean plants. This was supported by six specific objectives, each of which is addressed below with reference to the relevant chapters and findings. Addressing these objectives required a combination of synthetic chemistry (Chapter 2), analytical method development (Chapter 4), and plant-based experimentation (Chapter 3 and 4), supported by custom data analysis tools (Appendix D).

1. Understand the metabolic pathways of foliar applied tryptophan within the plant species

Labelled tryptophan analogues were synthesised (Section 2.3) to explore metabolic pathways following treatment. Metabolomics data (Section 4.2) suggested a possible treatment derived metabolite. Although full identification was not possible, isotopic labelling confirmed that the compound came from the applied treatment.

2. Investigate the mechanisms of foliar uptake of tryptophan including factors affecting absorption efficiency

Uptake rates were investigated in Section 4.1.5, with analysis of the effects of pH and enantiomers conducted in Section 4.1.6. These experiments demonstrated that the treatment was rapidly absorbed following application.

3. Develop and validate analytical methods for tracking labelled tryptophan within plant parts using mass spectrometry.

Method development is detailed across Chapters 2, 3, and 4, including isotope labelling strategies, washing protocols, and mass spectrometric detection.

4. Determine the uptake, transport, and conversion of tryptophan and its derivatives in different plant tissues

Uptake, transport and metabolism of tryptophan were explored in Sections 4.1 and 4.2 using deuterium-labelled tryptophan analogues (Section 2.3.2)

5. Assess the impact of tryptophan application on plant growth, yield, and stress response.

The physiological effects of the tryptophan application were discussed in Section 4.3, with no consistent differences observed between treated and control plants.

6. Explore the hypothesis that amino acid interactions with plant hormones drive specific physiological effects, with different amino acid influencing different hormonal pathways.

Although full hormonal mapping was outside the project scope, some evidence supporting hormone-related effects was presented in Section 4.2, including discussion of auxin-like responses and altered metabolite profiles.

From compound synthesis and method development to uptake analysis, metabolite tracking, and physiological assessment, this work presents a multifaceted investigation into the fate and function of foliar-applied tryptophan in plants. These findings contribute both practical techniques and new insights to the study of amino acid biostimulants and their mechanisms within plants and lay the groundwork for future exploration into hormone-amino acid interactions and bioactive compound development.

5.2 Future work

5.2.1 Synthetic expansion and bioactivity of indole dimers

Building on the synthetic strategies developed in Chapter 2, future work should explore both the biological relevance and synthetic optimisation of the dimeric compounds. Bioactivity screening for antimicrobial, anticancer, and antibacterial is recommended, given the structural similarity of these dimers to natural bioactive compounds such as rebeccamycin, which exhibit significant biological activity. Screening against a range of microbial strains or cancer cell lines could help identify potential applications in medicine or agriculture.

Further work is also needed to refine the synthesis and optimisation of the heterodimer, with particular emphasis on minimising homodimer formation to maximise heterodimer yield. This optimisation is essential to fully assess the potential of these molecules as viable bio-tracers.

Expanding the original set of substituted indoles to include a complete series of halogens (F, Cl, Br, I) at each position on the indole ring would offer valuable insights into the electronic and steric effects of these substituents on dimer formation. Understanding these trends will clarify substituent effects on selectivity and guide future optimisation of the synthesis conditions.

As well as this, future work could explore the design and synthesis of novel amino acid analogues with enhanced bioactivity, building on insights from tryptophan metabolism and other amino acids yet to be studied in detail. These analogues may offer more pronounced effects on plant growth, health, and yield, and could support the development of targeted agricultural treatments with improved efficacy and minimal side effects. The design of such compounds could also focus on improving leaf permeability through strategic structural modifications. Studying their movement within the plant could ensure consistent distribution to all tissues, mirroring the transport characteristics observed for tryptophan.

5.2.2 Amino acid analogues and plant bioactivity

While this study advances knowledge on amino acid uptake and distribution, further work is needed to clarify the precise metabolic pathways involved. One limitation of the project was the sensitivity of the instrumentation, which reduced the ability to detect deuterated fragments following metabolism. To overcome this in future studies, employing alternative instrumentation such as Quadrupole Time-of-Flight (Q-TOF) could afford higher resolution and improved detection at lower concentrations.

In addition, it is possible that a sufficient portion of the deuterated L-Trp was incorporated into larger molecules, such as peptides and proteins. The method utilised in this study focused on free amino acids, and so larger molecules could not be identified. Using different extraction techniques and mass spectrometry methods, such as Matrix-Assisted Laser Desorption Ionisation Time-of-Flight (MALDI-TOF), would allow for the investigation of both larger molecules and the small free amino acids already studied.

Future work should also expand the scope of amino acids investigated to better understand the relationship between their modes of action in plant systems. Comparative studies across a broader selection of amino acids could help clarify uptake dynamics, metabolic fate, and physiological impact.
Although the focus of this research was on uptake and metabolisation, growth parameters were included to ensure that any observed changes in plant biochemistry aligned with measurable improvements in plant growth. Future studies could build on these findings by incorporating additional metrics such as leaf area, root-to-shoot ratio, or nutrient content, providing a more detailed understanding of treatment impact on plant physiology. These metrics could not be included within the constraints of this project.

In addition to expanding analytical scope, future work should adopt a “benign by design” approach. This aligns with the principles of green chemistry and considers not only the effectiveness of a treatment but also its environmental impact. Such an approach involves evaluating the full lifecycle of a compound from synthesis to plant

metabolism and eventual removal. These considerations guide the development of treatments that balance agricultural productivity with environmental responsibility, helping to avoid unintended consequences such as those seen in the aftermath of the Green Revolution. By integrating principles of green chemistry and ethical design, future research could contribute to agricultural solutions that support both food security and environmental health, aligning scientific advancement with societal needs.

Chapter 6: Experimental

This chapter outlines the experimental procedures, including the synthesis of amino acid derivatives (8.2), the growth conditions for plant trials (8.3), and the analytical methods used (8.4).

6.1 Materials and general methods

All reagents and solvents are commercially available (Sigma Aldrich, Fischer, Fluorochem, and TCI) and were used without further purification unless stated otherwise. Anhydrous methanol was supplied through the departmental Solvent Purification System (SPS).

Reactions were carried out without drying, without an inert atmosphere unless stated otherwise. Room temperature (r.t.) reactions were carried out between 15 - 25 °C. Column chromatography was typically carried out using 60 Å (70 - 230 mesh) silica gel from VWR.[209] TLC was conducted using 2 cm x 5 cm aluminium-backed plates coated with silica matrix (0.2 mm) and fluorescent indicator (254 nm). Visualisation of TLC was carried out using a UV lamp and/or appropriate staining such as Ninhydrin or phosphomolybdic acid (PMA).[210]

6.1.1 Nuclear magnetic resonance (NMR)

NMR spectra were recorded on either a Bruker Avance III-HD-400 spectrometer with operating frequencies of 400.07 MHz for ^{1}H, 100.60 MHz for ^{13}C, 376.45 MHz for ^{19}F or a Varian VNMRS-600 spectrometer with operating frequencies of 599.42 MHz for ^{1}H, 150.72 MHz for ^{13}C, 564.02 MHz for ^{19}F at 298 K. Spectra were processed using MestReNova (V 15.1) software. ^{1}H NMR Chemical shifts were referenced to residual non-deuterated solvent peaks within the NMR solvent; $CHCl_3$ (δH = 7.26 ppm), CH_2Cl_2 (δH = 5.32 ppm), CH_3CN (δH = 1.94 ppm), CH_3OH (δH = 3.31 ppm) DMSO (δH = 2.50 ppm). The multiplicity of ^{1}H signals are indicated as: s = singlet; d = doublet; t = triplet; q = quartet; quint. = quintet; sex. = sextet; sept. = septet; m = multiplet; br = broad; and combinations thereof. Coupling constants (J) are quoted in Hz and reported to the nearest 0.1 Hz. Chemical shifts for ^{13}C NMR spectra were referenced

to deuterated solvent peaks in the NMR solvent; $CDCl_3$ (δC = 77.16 ppm), CD_3CN (δC = 1.32 ppm), CD_3OD (δC = 49.00 ppm) DMSO-d_6 (δC = 39.52 ppm). All ^{13}C resonances are reported to the nearest 0.1 ppm in general or to 0.01 ppm to aid in the differentiation of closely resolved signals. Product identification and peak assignments were completed using 2D experiments (COSY, NOSEY, HSQC, HMBC) and pulse experiments (DEPT-135), where appropriate. Where full assignment has not been possible, signals have been labelled as Ar (or Ar-H for proton NMR) to indicate that a carbon or proton is part of an aromatic system and Al (or Al-H for proton NMR) to indicate that a carbon or proton is part of an aliphatic system.

6.1.2 Mass spectrometry

All mass spectrometry for characterisation was carried out using a tandem Acquity UPLC (Waters Ltd, UK) and a Single Quadrupole Detector (SQD) with an ESI mass spectrometer (set to EI+ mode and EI- mode where appropriate). The UPLC was equipped with an Acquity UPLC BEH C_{18} 1.7 μm (2.1 mm x 50 mm) column, and mobile phase composition of H_2O containing formic acid (0.1% v/v): Acetonitrile mobile phase (gradient elution; t = 0 min, 95% : 5%, t = 4 min, 5% : 95%), set at a flow rate of 0.6 $mL.min^{-1}$.

Mass Spectrometry for plant analysis was carried out using a tandem Acquity UPLC (Waters Ltd, UK) and a Triple Quadrupole Detector (TQD) with an ESI mass spectrometer (set to EI+ mode and EI- mode where appropriate). The UPLC was equipped with a Zorbax Eclipse XDB-C18 HPLC column (4.5 mm × 50 mm, 1.8 μm particle size) column, and mobile phase composition of H_2O containing formic acid (0.1% v/v) : methanol mobile phase (gradient elution; t=0 min, 80% : 20%, t=5.1 min, 1% : 99%, t= 13 min, 1% : 99%, t=13.1 min, 80% : 20%), set at a flow rate of 0.4 $mL.min^{1}$.

6.1.3 Melting point data

Melting points of solid and crystalline products were measured using a Sanyo Gallenkamp variable heater equipped with a 300 °C mercury thermometer. Melting

points are uncorrected and solvents of crystallisation are listed along with the observed melting point range where appropriate.

6.1.4 Infrared spectroscopy

Infra-red spectra were acquired using a Perkin Elmer Spectrum 100 FT-IR spectrometer equipped with a UATR attachment and CsI window. Spectra were recorded from a range of 4,000 - 380 cm^{-1}. Absorbance shape and intensity are described as w = weak; m = medium; s = strong; sh = sharp; br = broad.

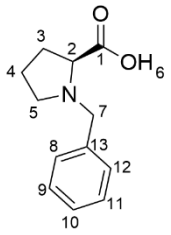
6.1.5 X-ray crystallography (XRD)

X-ray single crystal data was collected using λMoKα radiation (λ = 0.71073 Å) on an Agilent XCalibur (Sapphire-3 CCD detector, fine-focus sealed tube, graphite monochromator; compounds (rac)-1b and -2c) and Bruker D8 Venture (Photon100 CMOS detector, IμS-microsource, focusing mirrors; all other compounds) diffractometers equipped with the Cryostream (Oxford Cryosystems) open-flow nitrogen cryostats at a temperature of 120.0 K. All structures were solved by direct methods and refined by full-matrix least squares on F2 for all data using Olex2 and SHELXTLsoftware.[211] All non-disordered non-hydrogen atoms were refined anisotropically, the hydrogen atoms were placed in the calculated positions and refined in riding mode. All X-ray single crystal data was solved by Dr Toby Blundell, Dr Natalie Pridmore, and Dr Matthew Kitching. For compounds with crystal structures matching known data in the Cambridge Structural Database (CSD), database identifier codes are provided. Full crystallographic data is given when available, where no crystal data is available in the CSD.

6.2 Synthetic experimental

6.2.1 General synthesis of Belokon complex Gly-Ni-(S)-2-[N-(N-benzylpropyl)amino]benzophenone

6.2.1.1 1-Benzyl-L-proline (**1**)



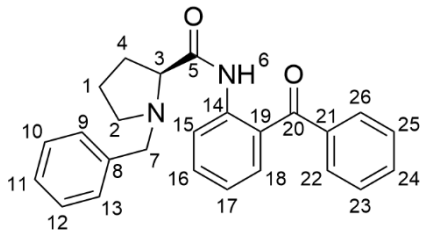
1

Commercially available L-proline (8.10 g, 70.4 mmol) was dissolved in methanol (40 mL). To this, a solution of sodium methoxide (7.61 g, 140.8 mmol, 2 eq.) in methanol (25.4 mL) was added, and the reaction mixture was heated to 45 °C. After 10 minutes, benzyl chloride (8.4 mL, 73.2 mmol, 1.05 eq.) was added dropwise over 1 hour. The mixture was stirred at 50 °C for a further 16 hours, over which time a white precipitate formed.

The reaction mixture was then cooled to room temperature, chloroform (40 mL) was added, and the reaction mixture was stirred for a further 90 minutes. The precipitate formed was removed by filtration and washed with cold chloroform, followed by acetone. The filtrate was concentrated to give an orange cloudy oil. This was then crystalised from methanol and diethyl ether, and a white solid was removed by filtration. The mother liquor from this crystallisation was then concentrated to give an off-white solid, which was dried over phosphorus pentoxide overnight to give 1-benzyl-L-proline (**1**) (11.13 g, 50.76 mmol, 72% yield).

^{1}H NMR (400 MHz, DMSO-d_6) δ 7.40 - 7.28 (m, 5 H, $H_{8\text{-}12}$), 4.08 (d, J = 13.0 Hz, 1 H, $H_{7'}$), 3.76 (d, J = 13.0 Hz, 1 H, $H_{7''}$), 3.34 (dd, J = 9.0, 5.7 Hz, 1 H, H_2), 3.07 (ddd, J = 9.8, 7.5, 3.7 Hz, 1 H, $H_{5'}$), 2.60 (ddd, J = 9.8, 8.7, 7.4 Hz, 1 H, $H_{5''}$), 2.16 - 2.04 (m, 1 H, $H_{4'}$), 1.92 - 1.63 (m, 3 H, $H_{4''}$, H_3). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 172.3 (C_1), 136.1 (C_{13}), 129.3 (C_9, C_{11}), 128.3 (C_8, C_{12}), 127.7 (C_{10}), 65.9 (C_2), 57.3 (C_7), 52.8 (C_5), 28.6 (C_3), 22.8 (C_4). **LC-MS** (ESI+) *m/z* 206.20 $[M+H]^+$. **mp**: 164 - 166 °C (Lit.: 165 - 166 °C).[212] **FT-IR** (ν_{max} / cm^{-1}): 3401w, 2981w, 2796w, 1577s, 1405s, 1300s, 1183w, 1077w, 701s. NMR and Mass Spectrometry data closely match literature values.[6, 139, 213]

6.2.1.2 (S)-1-Benzylpyrrolidine-2-carboxylic acid (2-benzoylphenyl) amide (**2**)

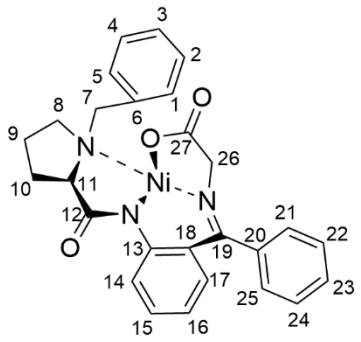


2

To a stirred solution of benzyl proline (**1**) (3.42 g, 16.6 mmol) in dichloromethane (8.5 mL), thionyl chloride (1.5 mL, 20 mmol, 1.2 eq.) was added dropwise over 10 minutes at -20 °C. Stirring was continued at -10 °C until the reaction mixture became almost transparent (5 minutes), at which point a solution of 2-aminobenzophenone (3.27 g, 16.6 mmol, 1 eq.) in dichloromethane (8.5 mL) was added to the reaction mixture at -30 °C. After 5 minutes, triethylamine (5.5 mL, 39.3 mmol, 2.4 eq.) was added. Stirring was continued overnight at room temperature (around 21 °C). The reaction mixture was then quenched with 10% aqueous sodium carbonate (20 mL) at 0 °C. The reaction mixture was then extracted with dichloromethane (3 x 50 mL). The organic layers were combined and dried over magnesium sulphate before being filtered. The filtrate was concentrated *in vacuo* to give a brown residue (5.42 g crude). The compound was recrystalised from ethanol to give a yellow solid (3.84 g, 9.96 mmol, 60%).

^{1}H NMR (599 MHz, DMSO-d_6) δ 11.03 (s, 1H, H_6), 8.29 (dd, J = 8.4, 1.1 Hz, 1H, H_{18}), 7.76 - 7.71 (m, 2H, $H_{9,13}$), 7.71 - 7.65 (m, 1H, H_{17}), 7.61 - 7.52 (m, 3H, $H_{23\text{-}25}$), 7.46 (dd, *J* = 7.8, 1.6 Hz, 1H, H_{15}), 7.33 - 7.28 (m, 2H, $H_{22,26}$), 7.19 (td, *J* = 7.5, 1.2 Hz, 1H, H_{16}), 7.15 - 7.07 (m, 3H, $H_{10\text{-}12}$), 3.74 (d, *J* = 12.8 Hz, 1H, $H_{7'}$), 3.52 (d, *J* = 12.9 Hz, 1H, $H_{7''}$), 3.17 (dd, *J* = 10.0, 4.5 Hz, 1H, H_3), 2.97 (ddd, *J* = 9.1, 6.4, 2.2 Hz, 1H, $H_{2'}$), 2.34 (td, *J* = 9.5, 6.3 Hz, 1H, $H_{2''}$), 2.11 (dtd, *J* = 12.3, 10.4, 9.9, 8.6 Hz, 1H, $H_{4'}$), 1.74 - 1.62 (m, 2H, H_1), 1.60 (dddd, *J* = 9.7, 6.5, 4.9, 2.6 Hz, 1H, $H_{4''}$).**^{13}C NMR** (151 MHz, DMSO-d_6) δ 196.9 (C_{20}), 173.2 (C_5), 138.2 (C_8), 137.9 (C_{19}), 137.6 (C_{21}), 132.9 (C_{24}), 132.9 (C_{17}), 131.6 (C_{15}), 129.8 ($C_{9,13}$), 128.9 ($C_{22,26}$), 128.5 ($C_{23,25}$), 128.0 ($C_{10,12}$), 126.9 (C_{11}), 126.2 (C_{14}), 122.8 (C_{16}), 121.1 (C_{18}), 67.3 (C_3), 59.0 (C_7), 53.4 (C_2), 30.1 (C_4), 23.6 (C_1). **LC-MS** (ESI +) *m/z:* 385.9 $[M+H]^+$ **mp:** 88 - 90 °C (Lit: 96 - 98 °C).[139] **FT-IR** (ν_{max} / cm^{-1}): 3249w, 2970w, 2813w, 1644m, 1511m, 1443m, 1264m, 1160m, 923m, 768m, 698s. **XRD:** Sample crystalised in acetonitrile to produce colourless blocks. Structure matched CSD-EXIKAY in the Cambridge Structural Database.[214] NMR data closely resembles literature values.[215, 216]

6.2.1.3 Gly-Ni-(S)-2-[N-(N-benzylpropyl)amino]benzophenone (**3**)



3

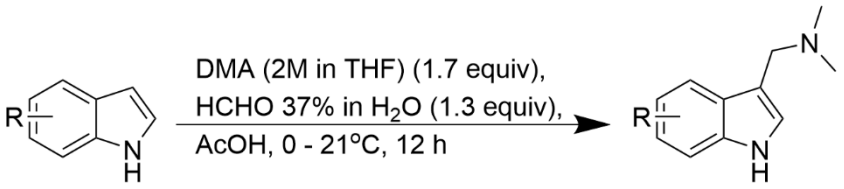
A mixture of **2** (0.50 g, 1.3 mmol), nickel(II) nitrate hexahydrate (0.45 g, 1.56 mmol, 1.2 eq.), and glycine (0.195 g, 2.6 mmol, 2 eq.) in anhydrous methanol (7 mL) was stirred under argon at 60 °C. To this a solution of sodium methoxide (1.08 g, 20 mmol) in anhydrous methanol (8 mL) was added.

The reaction mixture was stirred for 90 minutes at 55 °C and then poured onto water (100 ml) with acetic acid (5 mL). The red precipitate was recovered by Bückner filtration and purified by silica plug using chloroform followed by chloroform: methanol (9:1 v/v) to give the title compound (0.50g, 1.00 mmol, 77%)

^{1}H NMR (599 MHz, DMSO-d_6) δ 8.45 - 8.36 (m, 2H, $H_{1,5}$), 8.15 (dd, J = 8.7, 1.2 Hz, 1H, H_{23}), 7.65 - 7.57 (m, 1H, H_{16}), 7.57 - 7.52 (m, 2H, $H_{15,17}$), 7.46 (t, J = 7.7 Hz, 2H, $H_{2,4}$), 7.32 - 7.26 (m, 2H, $H_{3,14}$), 7.15 (ddd, J = 8.7, 6.7, 1.9 Hz, 2H, $H_{21,25}$), 6.73 - 6.64 (m, 2H, $H_{22,24}$), 4.13 (d, J = 12.4 Hz, 1H, $H_{7'}$), 3.71 (d, J = 19.7 Hz, 1H, $H_{26'}$), 3.59 (d, J = 12.4 Hz, 1H, $H_{7''}$), 3.56 - 3.49 (m, 3H, $H_{8',11,26''}$), 3.16 - 3.02 (m, 1H, $H_{9'}$), 2.44 - 2.34 (m, 1H, $H_{10'}$), 2.27 (td, J = 10.6, 5.9 Hz, 2H, $H_{8'',10''}$), 2.06 (ddq, J = 14.3, 5.7, 3.0 Hz, 1H, $H_{9''}$). **^{13}C NMR** (151 MHz, DMSO-d_6) δ 181.2 (C_{12}), 175.7 (C_{27}), 170.1 (C_{19}), 142.5 (C_{13}), 134.8 (C_6), 134.5 (C_{18}), 132.5 (C_{25}), 131.7 ($C_{1,5}$), 131.2 (C_{24}), 129.3 ($C_{15,17}$) 129.1 (C_{16}), 128.6 ($C_{2,3,4}$), 126.6 (C_{14}), 126.1 (C_{21}), 124.7 (C_{20}), 123.8 (C_{23}), 120.2 (C_{22}), 69.2 (C_{11}), 62.6 (C_7), 61.0 (C_{26}), 57.9 (C_8), 30.4 (C_{10}), 23.3 (C_9). **LC-MS** (ESI +) *m/z*: 499.2 [M+H$^+$]. **mp:** 189 - 191 °C (Lit: 206 - 210 °C).[217] **FT-IR** (ν_{max} / cm^{-1}): 3451w (br), 2982w, 2871w, 1671m, 1442m, 1336m, 1333m, 1174m, 1065m, 959w, 757m, 701m. **XRD:** Sample crystallised in methanol to produce orange planks: Crystal data for $C_{27}H_{25}N_3NiO_3$ (m = 498.21 g/mol): orthorhombic, space group $P2_12_12_1$. NMR data closely resembles literature values.[213]

6.2.2 Synthesis of gramine monomers and dimers

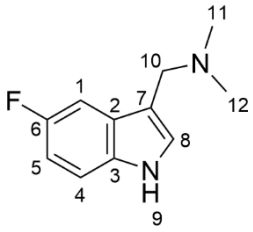
6.2.2.1 General synthesis for gramine analogues (**4a-l**)



4a-l

To a solution of dimethylamine (2 M in tetrahydrofuran, 8.5 mL, 17.05 mmol, 1 eq.) at 0 °C, formaldehyde (37% w/v in H_2O, 1.3 mL, 17.05 mmol, 1 eq.) and glacial acetic acid (1.5 mL, 39.25 mmol, 2.3 eq.) were added and the solution stirred for 10 minutes. To this mixture, the chosen indole analogue (17.05 mmol, 1 eq.) dissolved in glacial acetic acid (2 mL) was added dropwise. The reaction was then brought to room temperature and stirred overnight. Once gone to completion, as observed by thin layer chromatography (TLC, Hexane: Ethyl acetate, 6:4), the reaction mixture was poured into water, and aqueous sodium hydroxide (2 M) was added until a pH of 10 was reached. The formed precipitate was filtered off, washed with water, and dried under vacuum.

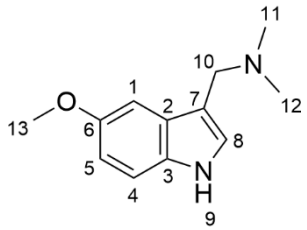
6.2.2.1.a 5-Fluorogramine (**4a**)



4a

Yield (98%) **^{1}H NMR** (400 MHz, DMSO-d_6) δ 10.99 (s, 1H, H_9), 7.36 - 7.23 (m, 3H, H_1,H_4,H_8), 6.89 (td, J = 9.2, 2.6 Hz, 1H, H_5), 3.48 (s, 2H, H_{10}), 2.13 (s, 6H, H_{11},H_{12}). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 156.6 (d, J = 230.8 Hz, C_6), 133.0 (C_7), 127.8 (d, J = 9.9 Hz, C_3), 126.4 (C_8), 112.2 (d, J = 9.8 Hz, C_1), 112.0 (d, J = 4.8 Hz, C_5), 109.0 (d, J = 26.2 Hz, C_2), 103.6 (d, J = 23.1 Hz, C_4), 54.3 (C_{10}), 44.9 ($C_{11,12}$). **^{19}F NMR** (376 MHz, $CDCl_3$) δ -124.34 to -125.52 (m). **LC-MS** (ESI +) *m/z*: 193.3 $[M+H]^+$. **FT-IR** (ν_{max} / cm^{-1}): 3043m (br), 2954m, 2929m, 2861m (sh), 2823m (sh), 2775m (sh), 1468m (sh), 1365m, 1170m, 1065m, 794s (sh). **XRD:** Sample crystallised in methanol to produce clear colourless plates: Crystal data for $C_{11}H_{13}FN_2$ (m = 192.23 g/mol): orthorhombic, space group $P2_12_12_1$. NMR data closely resembles literature values.[140]

6.2.2.1.b 5-Methoxygramine (**4b**)

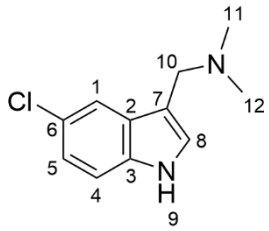


4b

Yield (50%) **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.29 (s, 1H, H_9), 7.22 (dd, *J* = 8.8, 0.6 Hz, 1H, H_5), 7.14 (d, *J* = 2.5 Hz, 1H, H_4), 7.11 - 7.05 (m, 1H, H_8), 6.85 (dd, *J* = 8.7, 2.5 Hz, 1H, H_1), 3.85 (s, 3H, H_{13}), 3.60 (s, 2H, H_{10}), 2.30 (s, 6H, H_{11}, H_{12}). **^{13}C NMR** (101 MHz, $CDCl_3$) d 154.2 (C_6), 131.5 (C_3), 128.4 (C_2), 124.8 (C_8), 112.8 (C_7), 112.2 (C_5), 111.9 (C_4), 101.2 (C_1), 56.0 (C_{13}), 54.6 (C_{10}), 45.3 (C_{11}, C_{12}). **mp**: 120 - 122 °C (Lit: 120 - 122 °C).[218] **LC-MS** (ESI +) *m/z*: 205.2 [M+H$^+$]. **FT-IR** (v_{max} / cm^{-1}): 3115w (br), 2937m, 2857m 2815m, 2775w, 1585w (sh), 1447m, 1243m, 1213s, 993m, 826m, 788s (sh). NMR data closely resembles literature values.[219]

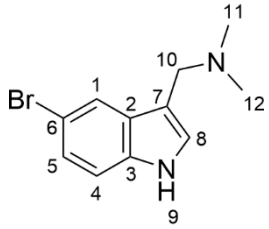
6.2.2.1.c 5-Chlorogramine (**4c**)

Yield (87%) **^{1}H NMR** (400 MHz, DMSO-d_6) δ 11.09 (s, 1H, H_9), 7.60 (d, *J* = 2.1 Hz, 1H, H_1), 7.35 (dd, *J* = 8.6, 0.6 Hz, 1H, H_4), 7.31 - 7.26 (m, 1H, H_8), 7.10 - 7.02 (m, 1H, H_5), 3.51 - 3.46 (m, 2H, H_{10}), 2.13 (s, 6H, H_{11}, H_{12}). **^{13}C NMR** (151 MHz, DMSO-d_6) δ 134.9 (C_6), 128.6 (C_7), 126.1 (C_8), 123.0 (C_2), 120.9 (C_5), 118.4 (C_1), 112.9 (C_4), 111.8 (C_3), 54.3 (C_{10}), 44.9 ($C_{11,12}$). **LC-MS** (ESI +) *m/z*: 209.2 [M+H]$^+$. **FT-IR** (vmax / cm^{-1}): 2824w, 1454m, 1344m (sh), 1233m (sh), 1168m, 995m, 858m, 788s (sh), 418m. **XRD:** Sample crystallised in methanol to produce clear colourless plates: Crystal data for $C_{11}H_{13}ClN_2$ (m = 208.68 g/mol): monoclinic, space group $P2_1/n$. NMR data closely resembles literature values.[219]



4c

6.2.2.1.d 5-Bromogramine (**4d**)

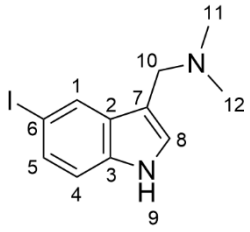


4d

Yield (88%) **^{1}H NMR** (400 MHz, DMSO-d_6) δ 11.10 (s, 1H, H_9), 7.75 (d, *J* = 2.0 Hz, 1H, H_1), 7.31 (d, *J* = 8.6 Hz, 1H, H_4), 7.27 (d, *J* = 2.0 Hz, 1H, H_8), 7.17 (dd, *J* = 8.6, 2.0 Hz, 1H, H_5), 3.50 - 3.46 (m, 2H, H_{10}), 2.12 (s, 6H, H_{11}, H_{12}). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 135.5 (C_3), 129.8 (C_2), 126.4 (C_5), 123.8 (C_8), 121.9 (C_1), 113.8 (C_4), 112.2 (C_6), 111.4 (C_7), 54.8 (C_{10}), 45.4 ($C_{11,}$ C_{12}). **LC-MS** (ESI

+) *m/z*: 253.2 [M+H^{79}Br]$^+$, 255.1 [M+H^{81}Br]. **FT-IR** (ν_{max} / cm^{-1}): 2818s (sh), 1451s (sh), 1437s (sh), 1233m, 1169m (sh), 1113m, 996s (sh), 860m, 818m, 763s (sh), 515m, 482m, 781s. **XRD:** Sample crystalised in acetonitrile to produce colourless plates. Structure matched CSD-EVIDIY in the Cambridge Structural Database.[220] NMR data closely resembles literature values.[221]

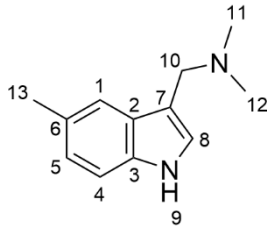
6.2.2.1.e 5-Iodogramine (**4e**)



4e

Yield (89%) **^{1}H NMR** (400 MHz, DMSO-d_6) δ 11.10 (s, 1H, H_9), 7.94 (d, J = 1.8 Hz, 1H, H_1), 7.32 (dd, J = 8.5, 1.7 Hz, 1H, H_5), 7.27 - 7.18 (m, 2H, $H_{4,8}$), 3.47 (s, 2H, H_{10}), 2.12 (s, 6H, $H_{11,12}$). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 135.5 (C_3), 130.2 (C_2), 128.8 (C_1), 127.6 (C_4), 125.5 (C_8), 113.9 (C_5), 111.4 (C_7), 82.1 (C_6), 54.4 (C_{10}), 44.9 (C_{11}, C_{12}). **LC-MS** (ESI +) *m/z*: 301.08 [M+H]$^+$ **FT-IR** (ν_{max} / cm^{-1}): 3102m, 2991m, 2855m, 2810s, 1450m, 1436m, 1362m, 1335m, 1250m, 1168m, 997s, 865m, 792s (sh), 419m**. XRD**: Sample crystallised in methanol to produce clear colourless plates: Crystal data for $C_{11}H_{13}IN_2$ (m = 300.13 g/mol): monoclinic, space group $P2_1$.

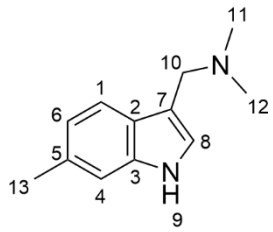
6.2.2.1.f 5-Methylgramine (**4f**)



4f

Yield (87%) **^{1}H NMR** (400 MHz, DMSO-d_6) δ 10.81 (s, 1H, H_9), 7.38 - 7.33 (m, 1H, H_1), 7.22 (d, J = 8.3 Hz, 1H, H_4), 7.13 (d, J = 2.3 Hz, 1H, H_8), 6.88 (dd, J = 8.3, 1.7 Hz, 1H, H_5), 3.49 - 3.45 (m, 2H, H_{10}), 2.36 (s, 3H, H_{13}), 2.12 (s, 6H, $H_{11,12}$). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 134.8 (C_6), 127.8 (C_3), 126.6 (C_2), 124.4 (Ar), 122.5 (Ar), 118.7 (Ar), 111.3 (C_7), 111.0 (Ar), 54.6 (C_{10}), 45.0 ($C_{11,12}$), 21.3 (C_{13}). **LC-MS** (ESI +) *m/z*: 189.2 [M+H]$^+$, 144.1 [M-NMe$_2$]$^+$. **FT-IR** (ν_{max} / cm^{-1}): 3104s (br), 3016m, 2988m, 2949m, 2855m, 2816m, 1482m, 1442m, 1368m, 1322m, 1235m, 1170w, 1106w, 1036w, 828w, 788w, 636w, 421w. **XRD:** Sample crystallised in methanol to produce clear colourless plates: Crystal data for $C_{12}H_{16}N_2$ (m = 188.27 g/mol): monoclinic, space group $P2_1/n$. NMR data closely resembles literature values.[222]

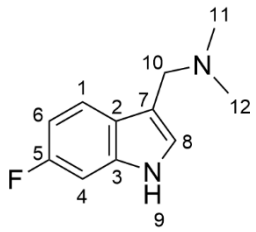
6.2.2.1.g 6-Methylgramine (**4g**)



4g

Yield (80%) **^{1}H NMR** (400 MHz, DMSO-d_6) δ 10.74 (s, 1H, H_9), 7.47 (d, J = 8.0 Hz, 1H, H_1), 7.13 (s, 1H, H_8), 7.10 (d, J = 2.3 Hz, 1H, H_4), 6.80 (dd, J = 8.1, 1.5 Hz, 1H, H_6), 3.49 (s, 2H, H_{10}), 2.37 (d, J = 8.9 Hz, 3H, H_{13}), 2.13 (s, 6H, $H_{11,12}$). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 136.8 (C_5), 129.8 (C_3), 125.5 (C_2), 123.6 (C_1), 120.1 (C_6), 118.8 (C_4), 111.5 (C_7), 111.1 (C_8), 54.5 (C_{10}), 44.9 ($C_{11,12}$), 21.4 (C_{13}). **LC-MS** (ESI +) *m/z*: 189.2 $[M+H]^+$, 144.1 $[M-NMe_2]^+$. **FT-IR** (ν_{max} / cm^{-1}): 2145w, 2855m, 2812m, 2767m, 1626m, 1542m, 1502m, 1448m, 1420m, 1339m, 1239m, 1169m, 1074m, 852s, 796s (sh), 592m, 422m. NMR data closely resembles literature values.[223]

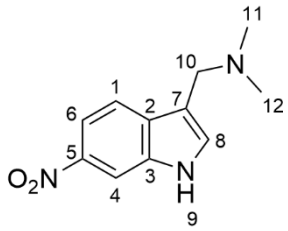
6.2.2.1.h 6-Fluorogramine (**4h**)



4h

Yield (66%) **^{1}H NMR** (400 MHz, DMSO-d_6) δ 10.95 (s, 1H, H_9), 7.57 (dd, J = 8.7, 5.6 Hz, 1H, H_1), 7.19 (d, J = 2.3 Hz, 1H, H_8), 7.11 (dd, J = 10.2, 2.3 Hz, 1H, H_4), 6.82 (ddd, J = 9.9, 8.7, 2.4 Hz, 1H, H_6), 3.49 (s, 2H, H_{10}), 2.13 (s, 6H, H_{11},H_{12}). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 158.8 (d, J = 233.7 Hz, C_5), 136.1 (d, J = 12.7 Hz, C_2), 124.8 (d, J = 3.4 Hz, C_1), 124.3 (C_3), 120.1 (d, J = 10.2 Hz, C_8), 112.0 (C_7), 106.7 (d, J = 24.3 Hz, C_4), 97.2 (d, J = 25.3 Hz, C_6), 54.4 (C_{10}), 44.9 ($C_{11,}C_{12}$). **^{19}F NMR** (376 MHz, DMSO-d_6) δ -122.46. **LC-MS** (ESI +) *m/z*: 193.2 $[M+H^+]$. **FT-IR** (ν_{max} / cm^{-1}): 3125m, 3082m (br), 2816m, 2778m, 1625m, 1550m, 1452m, 1438m, 1308m, 1150m, 805m, 794s (sh), 609m, 481m. NMR data closely resembles literature values.[140]

6.2.2.1.i 6-Nitrogramine (**4i**)

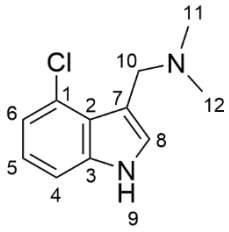


4i

Yield (81%) **^{1}H NMR** (400 MHz, DMSO) δ 11.75 (s, 1H, H_9), 8.38 - 8.27 (m, 2H, H_8), 7.88 (ddd, J = 8.9, 4.0, 2.2 Hz, 2H, Ar-H), 7.75 - 7.62 (m, 2H, Ar-H), 3.55 (s, 2H, H_{10}), 2.13 (s, 6H, $H_{11,12}$). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 142.3 (C_5), 135.2 (C_3), 132.7 (C_2), 131.9 (Ar), 119.8 (Ar), 114.1 (Ar), 113.7 (C_7), 108.7 (Ar), 54.3 (C_{10}), 45.3 ($C_{11,12}$). **LC-MS** (ESI +) *m/z*: 220.2 $[M+H^+]$.

mp: 158 - 160 °C (Lit. 176 - 178 °C)[224] **FT-IR** (v_{max} / cm^{-1}): 3266m (br), 2638w, 1588s, 1372s, 1246s, 1141m, 1054m, 812w, 540w. **XRD**: Sample crystallised in methanol to produce clear colourless plates: Crystal data for $C_{11}H_{13}N_3O_2$ (m = 219.24 g/mol): monoclinic, space group $P2_1/c$. ^1H NMR data matches literature values, no carbon NMR values available.[225]

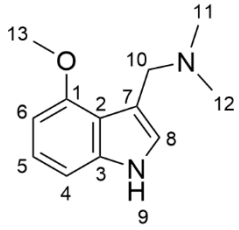
6.2.2.1.j 4-Chlorogramine (**4j**)



4j

Yield (93%) **^{1}H NMR** (400 MHz, DMSO-d_6) δ 11.27 (s, 1H, H_9), 7.32 (dd, J = 7.9, 1.2 Hz, 1H, H_6), 7.28 (s, 1H, H_8), 7.07 - 6.95 (m, 2H, $H_{4,5}$), 3.67 (s, 2H, H_{10}), 2.17 (s, 6H, $H_{11,12}$).**^{13}C NMR** (101 MHz, DMSO-d_6) δ 138.0 (C_1), 126.5 (C_6), 124.9 (C_2), 123.8 (C_3), 121.6 (C_4), 119.5 (C_5), 112.0 (C_7), 110.7 (C_8), 54.5 (C_{10}), 44.8 ($C_{11,12}$). **LC-MS** (ESI +) *m/z*: 209.2 $[M+H]^+$. **FT-IR** (v_{max} / cm^{-1}): 2973m, 2855m, 2819m, 1450m, 1427m, 1339m, 1240m, 1171m, 1047m, 853m, 776s, 751s, 731s, 597m, 580m. NMR data closely resembles literature values.[226, 227]

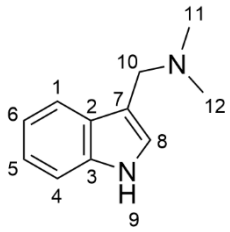
6.2.2.1.k 4-Methoxygramine (**4k**)



4k

Yield (96%) **^{1}H NMR** (400 MHz, DMSO-d_6) δ 10.89 (s, 1H, H_9), 7.02 (d, J = 2.0 Hz, 1H, H_6), 6.97 - 6.90 (m, 2H, $H_{5,8}$), 6.43 (dd, J = 6.3, 2.2 Hz, 1H, H_4), 3.82 (s, 3H, H_{13}), 3.68 (s, 2H, H_{10}), 2.17 (s, 6H, $H_{11,12}$). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 154.2 (C_1), 137.7 (C_7), 122.9 (C_6), 121.6 (C_5), 117.2 (C_2), 111.6 (C_3), 104.8 (C_8), 99.0 (C_4), 54.9 (C_{13}), 54.8 (C_{10}), 44.7 ($C_{11,12}$). **LC-MS** (ESI +) *m/z*: 205.2 $[M+H]^+$. **FT-IR** (v_{max} / cm^{-1}): 2941m, 2853m, 2806m, 2773m, 1541m, 1514m, 1466m, 1351m, 1278m, 1243m, 1084w, 994m, 850m, 782s, 755s, 617m. NMR data closely resembles literature values.[228]

6.2.2.1.l Gramine (**4l**)

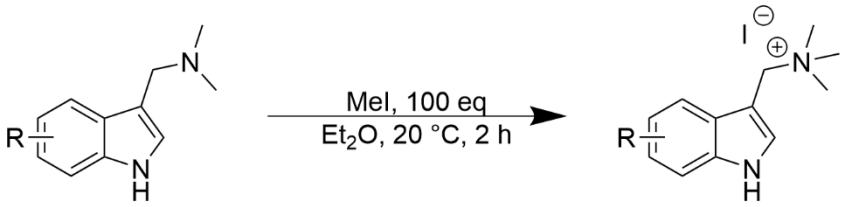


4l

Yield (65%) **^{1}H NMR** (400 MHz, DMSO-d_6) δ 10.93 (s, 1H, H_9), 7.59 (d, J = 7.8 Hz, 1H, H_1), 7.34 (d, J = 8.2 Hz, 1H, H_4), 7.19 (s, 1H, H_8), 7.06 (t, J = 8.2 Hz, 1H, H_5), 6.96 (t, J = 7.4 Hz, 1H, H_6), 3.52 (s, 2H, H_{10}), 2.14 (s, 6H, $H_{11,12}$). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 136.3 (C_3), 127.5 (C_2), 124.3 (C_8), 120.8 (C_5), 119.0 (C_6), 118.3 (C_1), 111.7 (C_7), 111.3 (C_4), 54.5 (C_{10}), 44.9 ($C_{11,12}$). **LC-MS** (ESI +) *m/z*: 175.2 $[M+H]^+$.

XRD: Sample crystallised in methanol to produce clear colourless plates: Crystal data for $C_{11}H_{14}N_2$ (m = 174.24 g/mol): orthorhombic, space group Pbca. NMR data closely resembles literature values.[229]

6.2.2.2 General synthesis for 1-(1H-indol-3-yl) trimethyl ammonium iodide analogues (monomers)

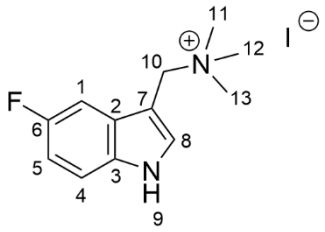


4a-l **5a-l**

Gramine **4a-l** (0.5 mmol, 1 eq) was added slowly to a solution of methyl iodide (50 mmol, 100 eq) in diethyl ether (100 mL). This solution was stirred at room temperature (20 ^{0}C) for 2 hours. After this time, a solid precipitate had formed, which was removed by Bückner filtration.

As these experiments were to assess the selectivity of monomer over dimer formation by spectroscopic or chromatographic means, isolation and purification of the product from the crude mixture was not conducted.

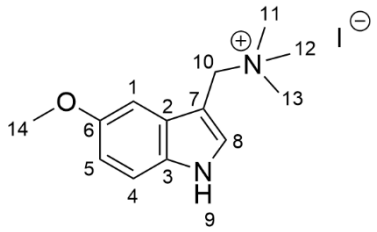
6.2.2.2.a **(5-Fluoro-3-indolylmethyl)trimethylammonium iodide (5a)**



5a

Yield (86%) Ratio of monomer to dimer by NMR (92:8). **^{1}H NMR** (400 MHz, DMSO-d_6) δ 11.74 (s, 1H, H_9), 7.76 (s, 1H, H_8), 7.66 (dd, J = 10.2, 2.5 Hz, 1H, H_1), 7.48 (dd, J = 8.8, 4.6 Hz, 1H, H_4), 7.03 (td, J = 9.1, 2.5 Hz, 1H, H_5), 4.66 (s, 2H, H_{10}), 3.03 (s, 9H, H_{11}, H_{12}, H_{13}). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 157.9 (d, J = 232.7 Hz, C_6), 132.9 (C_7), 132.2 (C_8), 128.5 (d, J = 10.2 Hz, C_3), 113.5 (d, J = 9.9 Hz, C_1), 110.4 (d, J = 26.2 Hz, C_5), 103.7 (d, J = 24.0 Hz, C_4), 102.3 (d, J = 4.6 Hz, C_2), 60.7 - 60.4 (m, C_{10}), 51.6 - 51.2 (m, $C_{11\text{-}13}$). **^{19}F NMR** (376 MHz, DMSO-d_6) δ -123.38. **LC-MS** (ESI +) *m/z*: 207.3 $[M]^+$. **FT-IR** (ν_{max} / cm^{-1}): 3229m, 1486m, 1462m, 1245m, 1182m, 803s. **mp:** 158 - 160 °C. (No literature value available for the iodide salt). NMR data closely resembles literature values (for sulphate salt).[230]

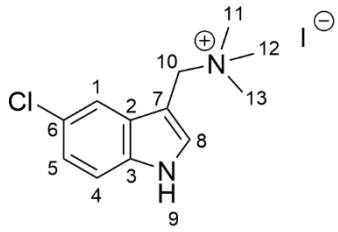
6.2.2.2.b (5-OMe-3-indolylmethyl)trimethylammonium iodide (**5b**)



5b

Yield (63%) Ratio of monomer to dimer by NMR (99:1) **^{1}H NMR** (400 MHz, DMSO-d_6) δ 11.50 (s, 1H, H_9), 7.64 (s, 1H, H_8), 7.45 - 7.33 (m, 2H, H_1, H_4), 6.81 (dd, J = 8.9, 2.4 Hz, 1H, H_5), 4.71 (s, 2H, H_{10}), 3.79 (s, 3H, H_{14}), 3.07 (s, 9H, H_{11}, H_{12}, H_{13}). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 154.8 (C_6), 131.5 (C_3), 130.9 (C_8), 128.78 (C_2), 113.3 (C_5), 112.6 (C_4), 102.1 (C_7), 101.0 (C_1), 61.1 (C_{10}), 56.2 (C_{14}), 51.6 (C_{11}), 51.6 (C_{12}), 51.6 (C_{13}). **LC-MS (ESI+):** *m/z* 219.3 $[M]^+$. **FT-IR** (ν_{max} / cm^{-1}): 3238m (br), 2956m, 1623m, 1583m, 1338m, 1254m, 1187m, 988s, 866s. NMR data closely resembles literature values (for sulphate salt).[230]

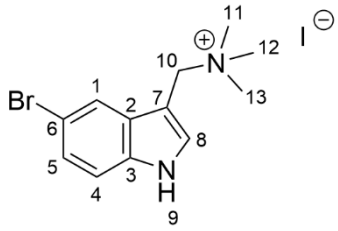
6.2.2.2.c (5-Chloro-3-indolylmethyl)trimethylammonium iodide (**5c**)



5c

Yield (79%) Ratio of monomer to dimer by NMR (89:11) **^{1}H NMR** (400 MHz, DMSO-d_6) δ 11.83 (d, J = 2.8 Hz, 1H, H_9), 7.97 (d, J = 2.0 Hz, 1H, H_1), 7.77 (d, J = 2.6 Hz, 1H, H_4), 7.50 (d, J = 8.7 Hz, 1H, H_8), 7.18 (dd, J = 8.7, 2.0 Hz, 1H, H_5), 4.68 (s, 2H, H_{10}), 3.04 (s, 9H, H_{11},H_{12},H_{13}). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 134.5 (C_3), 131.8 (C_8), 129.0 (C_6), 125.0 (C_2), 121.9 (C_1), 117.9 (C_5), 113.7 (C_4), 101.7 (C_7), 60.0 (C_{10}), 51.2 - 51.1 (m, $C_{11,12,13}$). **LC-MS** (ESI +) *m/z*: 223.2 $[M]^+$. **FT-IR** (v_{max} / cm^{-1}): 3216m, 1538m, 1459m, 1333m, 1242m, 1108m, 870m, 800m.

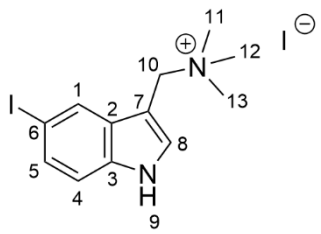
6.2.2.2.d (5-Bromo-3-indolylmethyl)trimethylammonium iodide (**5d**)



5d

Yield (80%) Ratio of monomer to dimer by NMR (90:10). **^{1}H NMR** (400 MHz, DMSO-d_6) δ 11.83 (s, 1H, H_9), 8.10 (d, J = 1.9 Hz, 1H, H_1), 7.75 (s, 1H, H_8), 7.45 (d, J = 8.6 Hz, 1H, H_4), 7.29 (dd, J = 8.6, 1.9 Hz, 1H, H_5), 4.67 (s, 2H, H_{10}), 3.03 (s, 9H, H_{11}, H_{12}, H_{13}). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 134.8 (C_3), 131.6 (C_2), 129.7 (C_5), 124.5 (C_8), 120.8 (C_1), 114.1 (C_4), 113.0 (C_6), 101.6 (C_7), 60.0 (C_{10}), 51.2 - 51.1 ($C_{11,12,13}$). **LC-MS** (ESI+) *m/z*: 267.1 $[M+^{79}Br]^+$, 269.1 $[M+^{81}Br]^+$. **FT-IR** (v_{max} / cm^{-1}): 3229m, 2181w, 1537m, 1453m, 1288m, 1248m, 803s.

6.2.2.2.e (5-Iodo-3-indolylmethyl)trimethylammonium iodide (**5e**)



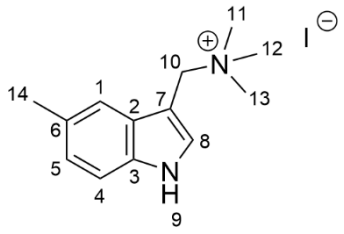
5e

Yield (86%) Ratio of monomer to dimer by NMR (94:6) **^{1}H NMR** (400 MHz, DMSO-d_6) δ 11.80 (s, 1H, H_9), 8.26 (d, J = 1.6 Hz, 1H, H_1), 7.70 (d, J = 1.8 Hz, 1H, H_8), 7.44 (dd, J = 8.5, 1.6 Hz, 1H, H_5), 7.34 (d, J = 8.5 Hz, 1H, H_4), 4.66 (s, 2H, H_{10}), 3.03 (s, 9H, H_{11}, H_{12}, H_{13}). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 135.6 (C_3), 131.6 (C_1), 130.9 (C_2), 130.4 (C_5), 127.3 (C_8),

115.0 (C_4), 101.8 (C_7), 84.9 (C_6), 60.5 (C_{10}), 51.7 - 51.6 ($C_{11\text{-}13}$). **LC-MS** (ESI+) *m/z*: 315.1 $[M]^+$. **FT-IR** (ν_{max} / cm^{-1}): 3264w, 2200w, 1988w, 1483m, 1451m, 1250m, 1108m, 880m (sh).

6.2.2.2.f (5-Me-3-indolylmethyl)trimethylammonium iodide (**5f**)

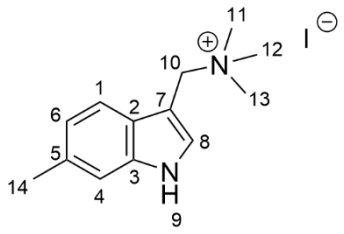
Yield (83%) Ratio of monomer to dimer by NMR (99:1)



5f

^{1}H NMR (400 MHz, DMSO) δ 11.50 (s, 1H, H_9), 7.64 (t, *J* = 2.6 Hz, 2H, $H_{1,8}$), 7.36 (d, *J* = 8.3 Hz, 1H, H_4), 7.00 (dd, *J* = 8.3, 1.6 Hz, 1H, H_5), 4.68 (s, 2H, H_{10}), 3.05 (s, 9H, $H_{11\text{-}13}$), 2.40 (s, 3H, H_{14}). **^{13}C NMR** (101 MHz, DMSO) δ 134.4 (C_6), 130.1 (C_3), 128.8 (Ar), 127.9 (C_2), 123.4 (Ar), 118.0 (Ar), 111.8 (Ar), 101.2 (C_7), 60.6 (C_{10}), 51.2 - 51.1 ($C_{11\text{-}13}$), 21.3 (C_{14}). **LC-MS (ESI+):** *m/z* 203.3 $[M]^+$. **FT-IR** (ν_{max} / cm^{-1}): 3231m (br), 3004m, 2967m, 1542s, 1482s, 1214m, 1107m, 872w, 796w, 617w.

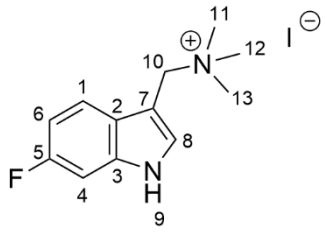
6.2.2.2.g (6-Me-3-indolylmethyl)trimethylammonium iodide (**5g**)



5g

Yield (83%) Ratio of monomer to dimer by NMR (95:5)
^{1}H NMR (599 MHz, DMSO-d_6) δ 11.48 (s, 1H, H_9), 7.70 (d, *J* = 8.1 Hz, 1H, H_1), 7.60 (d, *J* = 2.6 Hz, 1H, H_8), 7.26 (s, 1H, H_4), 6.96 (dd, *J* = 8.2, 1.5 Hz, 1H, H_6), 4.66 (s, 2H, H_{10}), 3.03 (s, 9H, $H_{11\text{-}13}$), 2.40 (s, 3H, H_{14}). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 136.5 (C_3), 131.1 (C_5), 129.5 (C_8), 125.6 (C_2), 121.9 (C_6), 118.2 (C_1), 111.8 (C_4), 101.6 (C_7), 60.7 (C_{10}), 51.2 - 51.1 ($C_{11\text{-}13}$), 21.2 (C_{14}). **LC-MS (ESI+):** *m/z* 203.3 $[M]^+$. **FT-IR** (ν_{max} / cm^{-1}): 3198m (br), 1505m, 1454m, 1249m, 1060m, 988s, 870s.

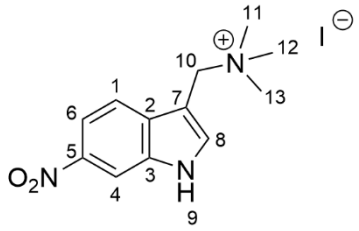
6.2.2.2.h (6-Fluoro-3-indolylmethyl)trimethylammonium iodide (**5h**)



5h

Yield (86%) Ratio of monomer to dimer by NMR (94:6) **^{1}H NMR** (400 MHz, DMSO-d_6) δ 11.67 (s, 1H, H_9), 7.88 (dd, J = 8.8, 5.4 Hz, 1H, H_1), 7.72 (d, J = 2.8 Hz, 1H, H_8), 7.26 (dd, J = 9.9, 2.4 Hz, 1H, H_4), 7.00 (td, J = 9.2, 2.4 Hz, 1H, H_6), 4.74 (s, 2H, H_{10}), 3.07 (s, 9H, $H_{11\text{-}13}$). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 159.0 (d, J = 235.5 Hz, C_5), 135.9 (d, J = 12.8 Hz, C_2), 130.9 (d, J = 3.3 Hz, C_1), 124.5 (C_3), 119.8 (d, J = 10.3 Hz, C_8), 108.6 (d, J = 24.6 Hz, C_4), 102.0 (C_7), 98.0 (d, J = 25.7 Hz, C_6), 60.6 - 59.9 (m, C_{10}), 51.4 - 51.0 (m, $C_{11\text{-}13}$). **^{19}F NMR** (376 MHz, DMSO-d_6) δ -121.10. **LC-MS (ESI+):** *m/z* 207.2 $[M]^+$, 148.1 $[M\text{-}NMe_2]^+$. **FT-IR** (ν_{max} / cm^{-1}): 3266m, 1628m, 1502m, 1457m, 1420m, 1357m, 1208m, 833s, 800s.

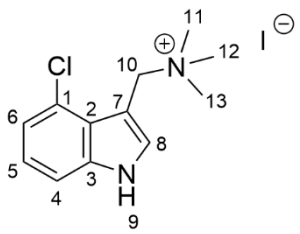
6.2.2.2.i (6-Nitro-3-indolylmethyl)trimethylammonium iodide (**5i**)



5i

Yield (79%) Ratio of monomer to dimer by NMR (81:19). **^{1}H NMR** (400 MHz, DMSO-d_6) δ 12.33 (s, 1H, H_9), 8.43 (dd, J = 4.7, 2.0 Hz, 1H, Ar-H), 8.11 (s, 1H, Ar-H), 8.07 - 8.02 (m, 2H, H, Ar-H), 4.77 (s, 2H, H_{10}), 3.06 (s, 9H, $H_{11\text{-}13}$). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 142.5 (C_5), 136.5 (C_3), 134.6 (C_2), 132.7 (C_7), 119.0 (Ar), 115.2 (Ar), 108.9 (Ar), 103.2 (Ar), 59.7 (C_{10}), 51.3 - 51.3 ($C_{11\text{-}13}$). **LC-MS** (ESI+): *m/z* 234.2 $[M]^+$. **FT-IR** (ν_{max} / cm^{-1}): 3196m (br), 3011m, 2956m, 1589s, 1621s, 1483m, 1247m, 1079w, 821w, 733w. ^{1}H NMR data matches literature values (no carbon available).[144]

6.2.2.2.j (4-Cl-3-indolylmethyl)trimethylammonium iodide (**5j**)

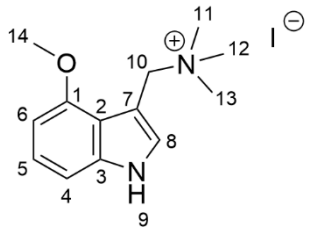


5j

Yield (68%) Ratio of monomer to dimer by NMR (99:1) **^{1}H NMR** (400 MHz, DMSO-d_6) δ 11.75 (s, 1H, H_9), 7.84 (d, J = 8.5 Hz, 1H, H_4), 7.73 (d, J = 2.4 Hz, 1H, H_8), 7.53 (d, J = 2.0 Hz, 1H, H_6), 7.17 (dd, J = 8.5, 1.9 Hz, 1H, H_5), 4.67 (s, 2H, H_{10}), 3.02 (s, 9H, $H_{11\text{-}13}$). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 136.5 (C_1), 131.3 (C_6), 126.6 (C_2), 126.6 (C_3), 120.5 (C_4), 120.0 (C_5), 111.7

(C_8), 102.2 (C_7), 60.2 (d, C_{10}), 51.3 - 51.2 (m, $C_{11\text{-}13}$). **LC-MS** (ESI +) *m/z*: 223.2 $[M]^+$. **FT-IR** (ν_{max} / cm^{-1}): 3187m (br), 1483m, 1430m, 1339m, 943s, 738s, 686s.

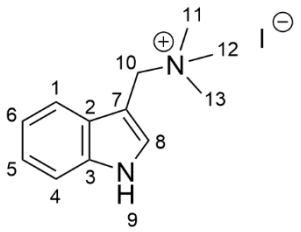
6.2.2.2.k (4-MeO-3-indolylmethyl)trimethylammonium iodide (**5k**)



5k

Yield (76%) Ratio of monomer to dimer by NMR (80:20) **^{1}H NMR** (400 MHz, DMSO-d_6) δ 11.72 (d, *J* = 10.4 Hz, 1H), 7.67 (d, J = 2.8 Hz, 1H, H_8), 7.35 - 7.3231 (m, 1H, H_4), 7.08 - 7.03 (m, 1H, H_5), 6.87 (dt, J = 87.1, 0.9 Hz, 1H), 7.06 (dd, *J* = 8.1, 7.2 Hz, 1H), 6.87 (dt, *J* = 7.1, 1.0 Hz, 1H), 1.0 Hz, 1H, H_6), 4.75 (s, 2H), 3.12 - 3.07 (m, 3H), 3.02 (s, 9H, $H_{11\text{-}14}$). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 136.6 (C_1), 130.9 (C_5), 128.7 (C_3), 125.5 (C_2), 122.4 (C_4), 121.9 (C_6), 110.2 (C_8), 101.6 (C_7), 61.9 (C_{10}), 51.3 - 51.2 ($C_{11\text{-}13}$), 20.8 (C_{14}). **LC-MS** (ESI +): *m/z* 203.1 $[M\text{-}CH_3I+H]^+$ **FT-IR** (ν_{max} / cm^{-1}): 3197m (br), 2997m, 2187w, 1612m, 1530m, 1478m, 1340m, 919s, 783s, 753s.

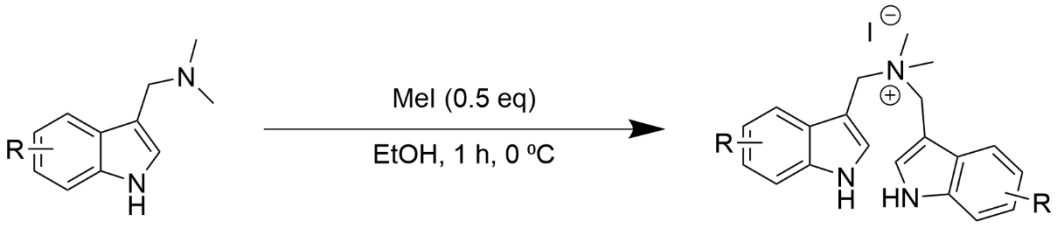
6.2.2.2.l **(3-indolylmethyl)trimethylammonium iodide (5l)**



5l

Yield (83%) Ratio of monomer to dimer by NMR (99:1) **^{1}H NMR** (400 MHz, DMSO-d_6) δ 11.67 (s, 1H, H_9), 7.84 (d, *J* = 7.8 Hz, 1H, H_1), 7.70 (s, 1H, H_8), 7.48 (d, *J* = 7.9 Hz, 1H, H_4), 7.24 - 7.09 (m, 2H, $H_{5,6}$), 4.71 (s, 2H, H_{10}), 3.05 (s, 9H, $H_{11\text{-}13}$). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 136.0 (C_3), 130.2 (C_2), 127.6 (C_8), 121.8 (C_5), 120.1 (C_6), 118.5 (C_1), 112.1 (C_7), 101.7 (C_4), 60.6 - 60.5 (C_{10}), 51.2 - 51.1 ($C_{11\text{-}13}$). **LC-MS** (ESI+) *m/z* 189.2 $[M]^+$. **FT-IR** (ν_{max} / cm^{-1}): 3187w, 1459m, 1482m, 1341m, 1254w, 851w, 748m. **mp:** 156-158 °C (Lit.: 157.6-157.9 °C).[231] ^{1}H NMR data matches literature values (no carbon available).[231]

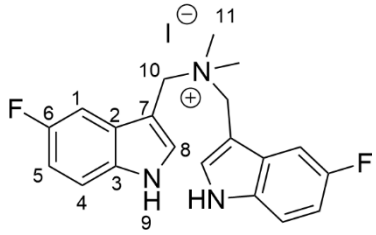
6.2.2.3 General synthesis of bis(indol-3-ylmethyl)dimethylammonium iodide



4a-l analogues **6a-l**

To a solution of Gramine **4a-l** (0.05 mmol, 1 eq) in ethanol (1.25 mL), iodomethane (0.019 mL, 0.5 eq) was added, and the solution was stirred for 1 hour at 0 ^{o}C. After this time, an excess of ethyl acetate was added, and a precipitate formed. The precipitate was isolated and dried by Bückner filtration.

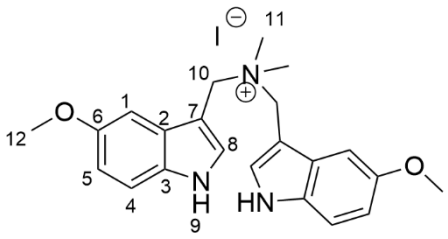
6.2.2.3.a 1-(5-Fluoro-1H-indol-3-yl)-N-((5-fluoro-1H-indol-3-yl)methyl)-N,N-dimethylmethanaminium iodide (**6a**)



6a

Yield (63%) Ratio of monomer to dimer by NMR (92:8). **^{1}H NMR** (400 MHz, DMSO-d_6) δ 11.75 (d, J = 2.7 Hz, 2H, H_7), 7.80 (d, J = 2.6 Hz, 2H, H_8), 7.71 (dd, J = 10.1, 2.5 Hz, 2H, H_1), 7.49 (dd, J = 8.8, 4.5 Hz, 2H, H_4), 7.04 (td, J = 9.1, 2.6 Hz, 2H, H_5), 4.74 (s, 4H, H_{10}), 2.85 (s, 6H, $H_{11,12}$). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 158.2 (d, J = 232.5 Hz, C_6), 133.3 (C_7), 132.7 (C_8), 129.0 (d, J = 10.0 Hz, C_3), 113.8 (d, J = 9.7 Hz, C_4), 110.6 (d, J = 25.9 Hz, C_5), 104.0 (d, J = 23.5 Hz, C_1), 102.3 (d, J = 4.7 Hz, C_2), 59.3 (C_{10}), 47.7 (C_{11}). **^{19}F NMR** (376 MHz, DMSO-d_6) δ -123.42. **LC-MS** (ESI+) m/z 340.3 $[M]^+$. **FT-IR** (ν_{max} / cm^{-1}): 3206w, 1477m, 1487m, 1431m, 941s, 801m, 613m. **XRD:** Sample crystallised in methanol to produce clear colourless plates: Crystal data for $C_{20}H_{20}F_2IN_3$ (m = 467.29 g/mol): triclinic, space group P-1.

6.2.2.3.b 1-(5-methoxy-1H-indol-3-yl)-N-((5-methoxy-1H-indol-3-yl)methyl)-N,N-dimethylmethanaminium iodide (**6b**)

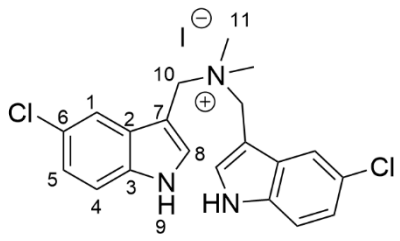


6b

Yield (52%) Ratio of dimer to monomer by NMR (96:4).

^{1}H NMR (400 MHz, DMSO-d_6) δ 11.50 (s, 1H, H_9), 7.68 (d, J = 2.6 Hz, 2H, H_8), 7.45 - 7.32 (m, 4H, $H_{1,4}$), 6.84 (dd, J = 8.8, 2.4 Hz, 2H, H_5), 4.77 (s, 4H, H_{10}), 3.81 (s, 6H, H_{12}), 2.86 (s, 6H, H_{11}). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 154.3 (C_6), 131.1 (C_2), 130.6 (C_5), 128.5 (C_3), 112.8 (C_4), 111.9 (C_1), 101.4 (C_8), 100.6 (C_7), 58.9 (C_{10}), 55.7 (C_{12}), 54.4 ($C_{11'}$), 47.1 ($C_{11''}$). **LC-MS** (ESI +) *m/z*: 364.4 $[M]^+$. **FT-IR** (ν_{max} / cm^{-1}): 3261m (br), 1584m, 1534m, 1438m, 1253m, 944s, 794m, 434w. **XRD:** Sample crystallised in methanol to produce clear colourless plates: Crystal data for $C_{22.5}H_{26.75}IN_{3.25}O_2$ (m = 501.62 g/mol): monoclinic, space group C2/c.

6.2.2.3.c 1-(5-Chloro-1H-indol-3-yl)-N-((5-chloro-1H-indol-3-yl)methyl)-N,N-dimethylmethanaminium iodide (**6c**)

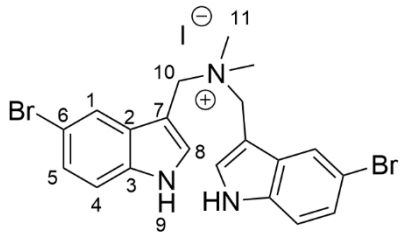


6c

Yield (76%) Ratio of dimer to monomer by NMR (76:24).

^{1}H NMR (400 MHz, DMSO-d_6) δ 11.76 (s, 2H, H_9), 7.80 (s, 2H, H_8), 7.72 (dd, J = 10.1, 2.5 Hz, 2H, H_5), 7.49 (dd, J = 8.9, 4.6 Hz, 2H, H_4), 7.04 (td, J = 9.2, 2.5 Hz, 2H, H_1), 4.74 (s, 4H, H_{10}), 2.84 (s, 6H, H_{11}). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 136.3 (C_6), 127.5 (C_2), 124.3 (C_5), 120.9 (C_3), 119.1 (C_4), 118.3 (C_1), 111.7 (C_8), 111.3 (C_7), 54.4 (C_{10}), 44.9 (C_{11}). **LC-MS** (ESI +) *m/z*: 372.2 $[M]^+$. **FT-IR** (ν_{max} / cm^{-1}): 3206m (br), 1704m, 1538m, 1452m, 1342m, 1235m, 1131m, 792m, 691m, 612m.

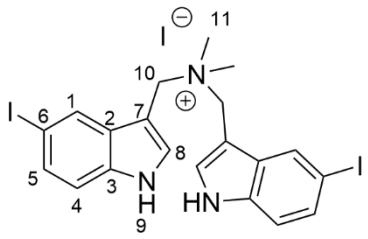
6.2.2.3.d 1-(5-Bromo-1H-indol-3-yl)-N-((5-bromo-1H-indol-3-yl)methyl)-N,N-dimethylmethanaminium iodide (**6d**)



6d

Yield (29%) Ratio of dimer to monomer by NMR (77:23). **^{1}H NMR** (400 MHz, DMSO-d_6) δ 11.84 (d, J = 2.8 Hz, 3H, H_9 (plus H_9 from monomer)), 8.17 (d, J = 2.0 Hz, 2H, H_1), 7.79 (d, J = 2.4 Hz, 2H, H_8), 7.46 (dd, J = 8.6, 4.3 Hz, 3H, H_4 (plus 1H from monomer), 7.30 (ddd, J = 8.7, 5.0, 1.9 Hz, 3H, H_5 (plus 1H from monomer)), 4.77 (s, 4H, H_{10}), 3.03 (s, 6H, H_{11}). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 134.9 (C_6), 131.9, 129.9 (C_3), 124.5 (Ar), 120.9 (Ar), 114.2 (Ar), 113.0 (C_2), 101.4 (C_7), 58.5 (C_{10}), 51.2 (C_{11}). **LC-MS** (ESI +) *m/z*: 462.1 $[M+2H]^+$. **FT-IR** (vmax/cm^{-1}): 3155m (br), 3034m, 2949m, 1535m, 1453m, 1322m, 1288m, 1213m, 1107m, 1059m, 1007m, 944m, 880w, 801m, 763m, 606m, 524w.

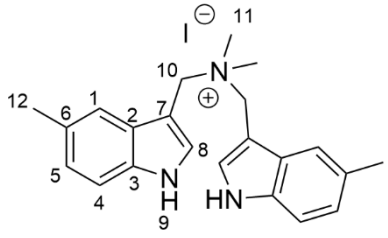
6.2.2.3.e 1-(5-Iodo-1H-indol-3-yl)-N-((5-iodo-1H-indol-3-yl)methyl)-N,N-dimethylmethanaminium (**6e**)



6e

Yield (92%) Ratio of dimer to monomer by NMR (53:47). **^{1}H NMR** (400 MHz, DMSO) δ 11.80 (t, J = 3.2 Hz, 3H, H_9 (plus H9 from monomer)), 8.36 - 8.31 (m, 2H, H_1), 7.74 (d, J = 2.6 Hz, 2H, H_8), 7.44 (ddd, J = 8.5, 5.1, 1.6 Hz, 3H, H_5 (plus H from monomer)), 7.34 (dd, J = 8.5, 4.5 Hz, 3H, H_4 (plus H from monomer)), 4.77 (s, 4H, H_{10}), 2.85 (s, 6H, H_{11}). **^{13}C NMR*** (101 MHz, DMSO) δ 135.2 (d, J = 7.4 Hz, C_6), 131.2 (d, J = 21.4 Hz, Ar), 130.5 (d, J = 21.8 Hz, Ar), 129.9 (C_3), 126.9 (d, J = 3.5 Hz, Ar), 114.5 (d, J = 3.5 Hz, Ar), 101.1 (d, J = 24.5 Hz, C_2), 84.5 (C_7), 60.0 (d, J = 2.9 Hz, C_{10}), 54.39 (t, J = 4.0 Hz, C_{11}). **LC-MS** (ESI +) *m/z*: 556.0 $[M+H]^+$. **FT-IR** (ν_{max} / cm^{-1}): 1449m, 1108m, 803m, 606m, 444m, 421m. *As it is not possible to determine which is related to the dimer and which to the monomer; here they have been reported as doublets. The triplets observed are due to the proximity of the carbons to the quaternary nitrogen.

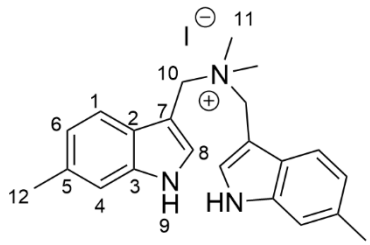
6.2.2.3.f 1-(5-Methyl-1H-indol-3-yl)-N-((5-Methyl-1H-indol-3-yl)methyl)-N,N-dimethylmethanaminium iodide (**6f**)



6f

Yield (62%) Ratio of monomer to dimer by NMR (86:14). **^{1}H NMR** (400 MHz, DMSO-d_6) δ 11.52 (d, J = 2.8 Hz, 2H, H_9), 7.64 (dt, J = 20.6, 2.7 Hz, 4H), 7.36 (dd, J = 8.3, 3.2 Hz, 2H), 7.01 (dt, J = 9.0, 2.3 Hz, 2H), 4.76 (s, 4H, H_{10}), 2.85 (s, 6H, H_{11}), 2.41 (d, J = 4.7 Hz, 6H, H_{12}). **^{13}C NMR** (101 MHz, DMSO) δ 134.5 (C_3), 130.3 (Ar), 128.8 (C_6), 128.1 (C_2), 123.5 (Ar), 118.0 (Ar), 111.8 (Ar), 101.0 (C_7), 59.0 (C_{10}), 47.2 (C_{11}), 21.3 (C_{12}). **LC-MS** (ESI +) *m/z*: 332.4 $[M]^+$. **FT-IR** (ν_{max} / cm^{-1}): 3238m (br), 2969m, 2359w, 1582s, 1540s, 1432m, 1245m, 1107w, 867w, 797w, 428w.

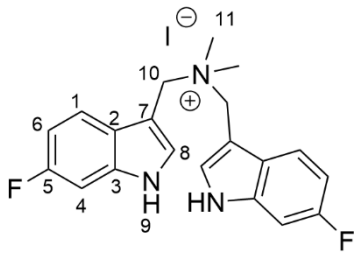
6.2.2.3.g 1-(6-Methyl-1H-indol-3-yl)-N-((6-Methyl-1H-indol-3-yl)methyl)-N,N-dimethylmethanaminium iodide (**6g**)



6g

Yield (87%) Ratio of dimer to monomer by NMR (81:19) **^{1}H NMR** (400 MHz, DMSO-d_6) δ 11.49 (d, *J* = 2.6 Hz, 2H, H_9), 7.73 (d, *J* = 8.1 Hz, 2H, H_1), 7.64 (d, *J* = 2.4 Hz, 2H, H_8), 7.26 (s, 2H, H_4), 6.97 (ddd, *J* = 8.3, 3.9, 1.4 Hz, 2H, H_6), 4.76 (s, 4H, H_{10}), 2.83 (s, 6H, H_{11}), 2.41 (s, 6H, H_{12}). ^{13}C NMR (101 MHz, DMSO) δ 136.6 (C_3), 131.0 (C_5), 129.7 (Ar), 125.8 (C_2), 121.9 (Ar), 118.3 (Ar), 111.8 (Ar), 101.4 (C_7), 59.0 (C_{10}), 47.2 (C_{11}), 21.3 (C_{12}). **LC-MS** (ESI +) *m/z*: 332.4[M]$^+$. **FT-IR** (ν_{max} / cm^{-1}): 3262m (br), 3008m, 2916w, 2162w, 2027w, 2054w, 1625m, 1541m, 1450m, 1411w, 1302w, 1196w, 944m, 867w, 808w.

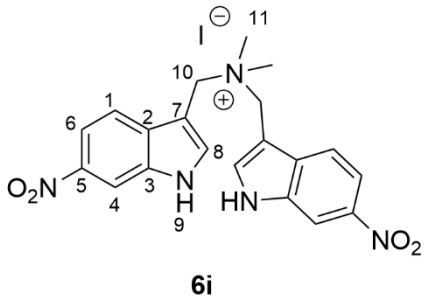
6.2.2.3.h 1-(6-fluoro-1H-indol-3-yl)-N-((6-fluoro-1H-indol-3-yl)methyl)-N,N-dimethylmethanaminium iodide (**6h**)



6h

Yield (27%) Ratio of dimer to monomer by NMR (89:11). **^{1}H NMR** (400 MHz, DMSO-d_6) δ 11.69 (d, J = 2.6 Hz, 2H, H_9), 7.87 (dd, J = 8.8, 5.3 Hz, 2H, H_1), 7.73 (d, J = 2.5 Hz, 2H, H_8), 7.27 (dd, J = 9.8, 2.4 Hz, 2H, H_4), 7.09 - 6.97 (m, 2H, H_6), 4.77 (s, 4H, H_{10}), 2.85 (s, 6H, H_{11}). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 159.0 (d, J = 235.5 Hz, C_5), 136.1 (d, J = 12.8 Hz, C_3), 131.1 (C_8), 124.7 (C_2), 119.8 (d, J = 10.2 Hz, C_4), 108.7 (d, J = 24.6 Hz, C_6), 101.8 (C_7), 98.1 (d, J = 25.5 Hz, C_1), 58.7 (C_{10}), 47.3 (C_{11}). **^{19}F NMR** (376 MHz, DMSO-d_6) δ -121.14 (d, J = 7.9 Hz, 1F). **LC-MS** (ESI +) *m/z*: 340.3 $[M]^+$. **FT-IR** (ν_{max} / cm^{-1}): 3248m (br), 2201w, 2058w, 2022w, 1638m, 1541m, 1458m, 1413m, 1345m, 1098m, 823m, 800m, 667m (sh).

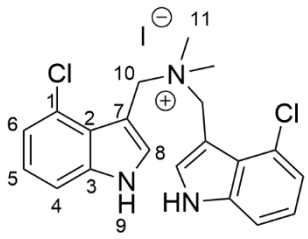
6.2.2.3.i 1-(6-Nitro-1H-indol-3-yl)-N-((6-Nitro-1H-indol-3-yl)methyl)-N,N-dimethylmethanaminium iodide (**6i**)



6i

Yield (15%) Ratio of dimer to monomer by NMR (91:9).
^{1}H NMR (400 MHz, DMSO-d_6) δ 12.35 (s, 2H, H_9), 8.44 (d, J = 2.1 Hz, 2H, H_8), 8.16 - 8.02 (m, 6H, $H_{1,4,6}$), 4.87 (s, 4H, H_{10}), 2.90 (s, 6H, H_{11}). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 142.5 (C_5), 136.8 (Ar), 134.7 (C_3), 132.9 (C_2), 119.1 (Ar), 115.2 (Ar), 108.9 (Ar), 102.9 (C_7), 58.3 (C_{10}), 47.4 (C_{11}). **LC-MS** (ESI +) *m/z*: 394.2 $[M]^+$. **FT-IR** (ν_{max} / cm^{-1}): 3194m (br), 1483s, 1458m, 1367m, 1246m, 1079w, 865w, 733w, 670w.

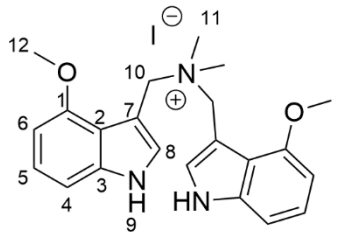
6.2.2.3.j 1-(4-Chloro-1H-indol-3-yl)-N-((4-chloro-1H-indol-3-yl)methyl)-N,N-dimethylmethanaminium iodide (**6j**)



6j

Yield (34%) Ratio of dimer to monomer by NMR (95:5). **^{1}H NMR** (599 MHz, DMSO-d_6) δ 12.10 (d, J = 2.8 Hz, 2H, H_9), 7.95 (d, J = 2.7 Hz, 2H, H_8), 7.50 (dd, J = 8.0, 1.0 Hz, 2H, H_6), 7.19 - 7.11 (m, 4H, H_4, H_5), 5.02 (s, 4H, H_{10}), 2.91 (s, 6H, H_{11}). **^{13}C NMR** (151 MHz, DMSO-d_6) δ 137.8 (C_1), 132.7 (C_8), 123.5 (C_2), 123.4 (C_3), 122.8 (C_6), 121.4 (C_4), 111.8 (C_5), 100.6 (C_7), 58.6 (C_{10}), 47.4 - 47.3 (C_{11}). **LC-MS** (ESI +) *m/z*: 374.0 $[M]^+$. **FT-IR** (ν_{max} / cm^{-1}): 3207m (br), 1475m, 1420m, 1341w, 1190w, 842w, 916w, 787w, 744w, 733w, 574w, 505w, 804w, 733w.

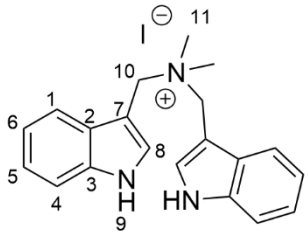
6.2.2.3.k 1-(4-methoxy-1H-indol-3-yl)-N-((4-methoxy-1H-indol-3-yl)methyl)-N,N-dimethylmethanaminium (**6k**)



6k

Yield (62%) Ratio of dimer to monomer by NMR (99:1). **^{1}H NMR** (400 MHz, DMSO) δ 11.67 (s, 2H, H_9), 7.65 (s, 2H, H_8), 7.13 - 7.03 (m, 4H, $H_{4,6}$), 6.57 (p, J = 4.8 Hz, 2H, H_5), 4.79 (s, 4H, H_{10}), 3.73 (s, 6H, H_{12}), 2.82 (s, 6H, H_{11}). **^{13}C NMR** (101 MHz, DMSO) δ 153.1 (C_1), 137.6 (C_2), 129.2 (Ar), 122.9 (Ar), 117.2 (C_3), 105.5 (Ar), 101.0 (C_7), 100.5 (C_8), 59.5 (C_{10}), 54.7 (C_{12}), 47.0 (C_{11}). **LC-MS** (ESI +) *m/z*: 364.4 $[M]^+$. **FT-IR** (ν_{max} / cm^{-1}): 3195m (br), 1550m, 1510m, 1473m, 1453m, 1383w, 1331w, 1092s, 822w, 783w, 730w.

6.2.2.3.l bis(indol-3-ylmethyl)dimethylammonium iodide (**6l**)



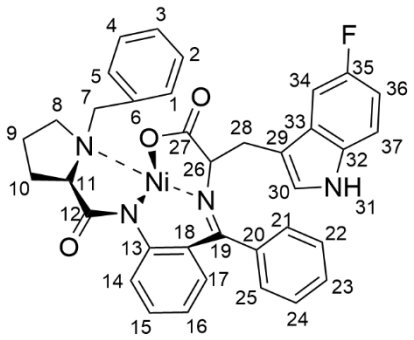
6l

Yield (40%) Ratio of dimer to monomer by NMR (94:6) **^{1}H NMR** (400 MHz, DMSO-d_6) δ 11.69 (s, 2H, H_9), 7.90 - 7.85 (m, 2H, H_4), 7.74 (s, 2H, H_8), 7.51 - 7.47 (m, 2H, H_1), 7.24 - 7.09 (m, 4H, $H_{5,6}$), 4.81 (s, 4H, H_{10}), 2.86 (s, 6H, H_{11}). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 136.2 (C_3), 130.4 (C_8), 127.9 (C_2), 121.9, 120.1, 118.6, 112.2, 101.5 (C_7), 59.0 (C_{10}), 47.3 (C_{11}).

LC-MS (ESI +) *m/z*: 304.4 $[M]^+$. **FT-IR** (ν_{max} / cm^{-1}): 3244m (br), 1543m, 1469m, 1458m, 1424w, 1250w, 1098w, 979w, 787w, 764w, 614w, 556w. NMR data matches literature values.[144]

6.2.3 Synthesis of 5-Fluorotryptophan

6.2.3.1 5-Fluoro-Tryptophan-Ni-(S)-2-[N-(N-benzylpropyl) amino]benzophenone (**7a**)



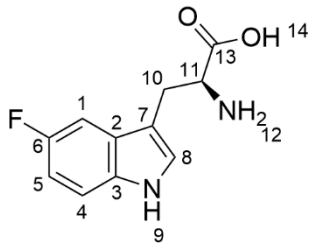
7a

To a stirred solution of compound **3** (0.25 g, 0.50 mmol) and compound **5a** (0.16 g, 0.75 mmol, 1.5 eq.) in acetonitrile (10 mL), finely ground sodium hydroxide (0.05 g, 1.25 mmol, 2.5 eq.) was added. The reaction mixture was stirred under argon for 3 hours at room temperature.

After this time, the reaction mixture was quenched with a 3 M HCl aqueous solution (50 mL) and extracted with chloroform (3 x 50 mL). Organic layers were combined and washed with sodium bicarbonate (2 x 50 mL), followed by water (2 x 50 mL), and brine (2 x 50 mL). This was then dried with sodium sulphate and solvent removed *in vacuo* to give a red solid (0.20 g, 0.31 mmol, 62%). The compound was taken forward without further purification.

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.26 (d, J = 8.1 Hz, 1H, H_{31}), 8.14 (d, J = 7.5 Hz, 1H, Ar-H), 8.04 - 7.89 (m, 1H, Ar-H), 7.80 (d, J = 7.5 Hz, 1H, Ar-H), 7.66 (d, J = 21.5 Hz, 3H, Ar-H), 7.52 (s, 4H, Ar-H), 7.49 - 7.29 (m, 5H, Ar-H), 7.17 (t, J = 7.1 Hz, 1H), 7.13 (s, 8H, Ar-H + $H_{11,26}$), 6.92 (dd, J = 18.3, 9.4 Hz, 1H, Al-H), 6.69 (dd, J = 6.0, 3.2 Hz, 1H, Al-H), 3.75 - 3.64 (m, 1H, Al-H), 3.59 - 3.39 (m, 1H, Al-H), 1.28 (s, 2H, Al-H), 0.09 (s, 4H, Al-H). **LC-MS** (ESI +) *m/z*: 646.3 $[M+H]^+$.

6.2.3.2 5-Fluoro-L-tryptophan (**8**)



8

A solution of compound **7a** (0.58 g, 0.93 mmol) in methanol (3 mL) was added to a stirred solution of aqueous 3 M HCl and methanol (1:1 v/v, 0.6 mL) at 70 °C. Upon colour change of red to purple (4 h) the reaction mixture was cooled to room temperature and solvent removed *in vacuo*. Water was then added to the residue followed by excess concentrated ammonia solution and this was extracted with chloroform (3 x 20 mL).

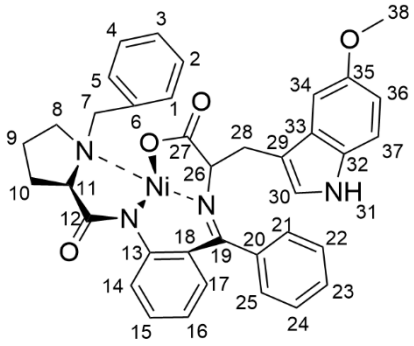
The organic layers were combined and dried with magnesium sulphate, solvent was removed *in vacuo* and the resulting residue purified by silica plug, to return recovered (S)-1-Benzylpyrrolidine-2-carboxylic acid (2-benzoylphenyl)amide (**2**).

The aqueous layer was concentrated *in vacuo* with toluene azeotrope and then purified by ion-exchange chromatography using DOWEX® Resin 50X2. The resin was prepared by washing with water and methanol before running a 0.1 M HCl solution to protonate acceptor sites. The product, as part of the aqueous layer, was eluted with water (100 mL) followed by 10% ammonia solution in methanol (300 mL). Extractions with UV activity and a purple stain in ninhydrin were collected and solvent removed *in vacuo*. Upon further purification by HPLC, a pale yellow solid was obtained (0.108 g, 0.49 mmol, 52%)

^{1}H NMR (400 MHz, DMSO) δ 11.25 (s, 1H, H_9), 7.67 - 7.61 (br s, H_{12}, exchangeable), 7.36 - 7.31 (m, 3H, $H_{1,4,5}$), 6.88 (td, *J* = 9.1, 2.4 Hz, 1H, H_8), 3.54 (s, 1H, H_{11}), 3.30 - 3.00 (m, 2H, H_{10}). **^{13}C NMR** (151 MHz, DMSO-d_6) δ 170.9 (C_{13}), 156.7 (d, *J* = 230.4 Hz, C_6), 133.0 (C_3), 127.7 (d, *J* = 10.2 Hz, C_4), 126.6 (C_2), 112.2 (d, *J* = 9.7 Hz, C_5), 109.3 (C_8), 109.0 (d, *J* = 26.2 Hz, C_5), 103.3 (d, *J* = 23.2 Hz, C_1), 54.5 (C_{11}), 26.7 (C_{10}). **^{19}F NMR** (376 MHz, DMSO) δ -125.54. **LC-MS** (ESI +) *m/z*: 223.4 $[M+H]^+$. NMR data aligns with literature values.[129]

6.2.4 Synthesis of 5-Methoxytryptophan

6.2.4.1 5-Methoxy-Tryptophan-Ni-(S)-2-[N-(N-benzylpropyl) amino]benzophenone (**7b**)



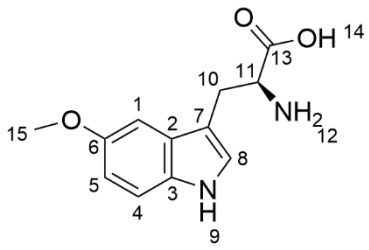
7b

To a stirred solution of compound **6** (0.25 g, 0.50 mmol) and compound **5b** (0.26 g, 0.75 mmol, 1.5 eq.) in acetonitrile (10 mL), finely ground sodium hydroxide (0.05 g, 1.25 mmol, 2.5 eq.) was added. The reaction mixture was stirred under argon for 3 hours at room temperature.

After this time the reaction mixture was quenched with 3 M HCl aqueous solution (50 mL) and extracted with chloroform (3 x 50 mL). Organic layers were combined and washed with sodium bicarbonate (2 x 50 mL), followed by water (2 x 50 mL), and brine (2 x 50 mL). This was then dried with sodium sulphate and solvent removed *in vacuo* to give a red solid (0.25 g, 0.38 mmol, 76%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.61 - 8.56 (m, 1H, H_{31}), 8.27 (d, *J* = 8.7 Hz, 1H, Ar-H), 7.99 - 7.92 (m, 2H, Ar-H), 7.53 - 7.42 (m, 2H, Ar-H), 7.28 (d, *J* = 8.2 Hz, 4H, benzyl Ar-H), 7.18 - 7.09 (m, 2H, Ar-H), 6.96 (d, *J* = 2.4 Hz, 1H, H_{30}), 6.85 - 6.73 (m, 3H, Ar-H), 6.72 - 6.64 (m, 2H, Ar-H), 4.29 (t, *J* = 4.9 Hz, 1H, H_{26}), 4.18 (d, *J* = 12.6 Hz, 1H), 3.87 - 3.68 (m, 1H), 3.47 (s, 3H, H_{38}), 3.42 - 3.31 (m, 2H), 3.17 (dd, *J* = 10.1, 7.4 Hz, 1H), 3.08 - 2.95 (m, 1H), 2.81 (ddd, *J* = 11.2, 6.8, 4.1 Hz, 2H), 2.47 (d, *J* = 15.0 Hz, 1H), 2.12 (ddt, *J* = 12.7, 9.9, 7.9 Hz, 1H), 1.98 (d, *J* = 15.1 Hz, 1H), 1.86 - 1.70 (m, 2H). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 180.3 (C_{12}), 179.7 (C_{27}), 170.9 (C_{19}), 154.5 (C_{35}), 143.0 (C_{13}), 134.2 ($C_{6,18}$), 133.6 ($C_{21,25}$), 133.3 (C_{31}), 132.4 ($C_{22,24}$), 131.6 ($C_{1,5}$), 129.8 (C_{15}), 129.1 (C_{17}), 128.8 ($C_{2,4}$), 128.1 (C_{16}), 127.4 (C_3), 126.2 (C_{20}), 125.1 (C_{14}), 123.6 (C_{33}), 120.6 (C_{23}), 113.5 (C_{37}), 112.2 (C_{36}), 109.8 (C_{29}), 100.6 (C_{34}), 77.4 (C_{30}), 71.6 (C_{26}), 70.5 (C_{11}), 63.2 (C_7), 57.0 (C_8), 55.6 (C_{38}), 30.8 (C_{28}), 30.6 (C_{10}), 22.8 (C_9). **LC-MS (ESI +)** *m/z*: 658.4 $[M+H]^+$.

6.2.4.2 5-Methoxytryptophan (**9**)



9

A solution of compound **7b** (0.58 g, 0.93 mmol) in methanol (3 mL) was added to a stirred solution of aqueous 3 M HCl and methanol (1:1 v/v, 0.6 mL) at 70 °C. Upon colour change of red to purple (4 h) the reaction mixture was cooled to room temperature and solvent removed *in vacuo*. Water was then added to the residue followed by excess concentrated ammonia solution and this was extracted with chloroform (3 x 20 mL). The organic layers were combined and dried with magnesium sulphate, solvent was removed *in vacuo* and the resulting residue purified by silica plug, to return recovered (S)-1-Benzylpyrrolidine-2-carboxylic acid (2-benzoylphenyl)amide (**5**).

The aqueous layer was concentrated *in vacuo* with toluene azeotrope and then purified by ion-exchange chromatography using DOWEX® Resin 50X2. The resin was prepared by washing with water and methanol before running a 0.1 M HCl solution to protonate acceptor sites. The product, as part of the aqueous layer, was eluted with water (100 mL) followed by 10% ammonia solution in methanol (300 mL). Extractions with UV activity and a purple stain in ninhydrin were collected and solvent removed *in vacuo* to give a brown solid (**8**) (0.08 g, 0.34 mmol, 37%).

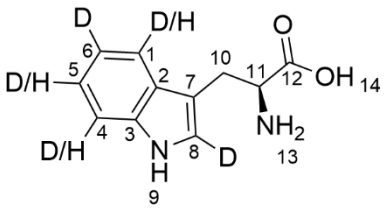
Further purification was undertaken by HPLC. This purification comprised three preparative runs on an InterChim Puriflash with Waters C18 XBridge column (100 x 19mm, 5um). The mobile phase was water with 0.1% formic acid and acetonitrile.

^{1}H NMR (599 MHz, DMSO-d_6) δ 10.85 (s, 1H, H_9), 7.26 - 7.19 (m, 2H, $H_{4,8}$), 7.09 (d, J = 2.4 Hz, 1H, H_1), 6.70 (dd, J = 8.8, 2.2 Hz, 1H, H_5), 3.75 (s, 3H, H_{15}), 3.49 (s, 1H, H_{11}), 3.27 (s, 1H, $H_{10'}$), 2.98 (s, 1H, $H_{10''}$). **^{13}C NMR** (151 MHz, DMSO-d_6) δ 170.5 (C_{13}), 153.0 (C_6), 131.5 (C_2), 127.7 (C_8), 124.9 (C_3), 112.0 (C_4), 111.1 (C_5), 109.3 (C_7), 100.3 (C_1), 55.4 (C_{15}), 54.8 (C_{11}), 27.1 (C_{10}). **LC-MS** (ESI +) *m/z*: 235.3 $[M+H]^+$. **FT-IR** (ν_{max} / cm^{-1}): 3186m (br), 2921m, 2226w, 2162w, 2032w, 1578m, 1509m, 1449m, 1188w, 1031w, 943w, 787w, 733w, 593w. NMR data closely aligns with literature values.[129]

6.2.5 General synthesis for deuterated tryptophan

Tryptophan (1 g, 0.49 mmol) was added to a round bottom flask containing deuterium oxide and deuterated trifluoroacetic acid (1.2 mL, 1:2 v/v). The flask was then sealed and stirred at room temperature in darkness for 3 days. After this time, the deuterated trifluoroacetic acid was removed under vacuum and the remaining solvent was removed by lyophilisation. This procedure was repeated for a total duration of 9 days.

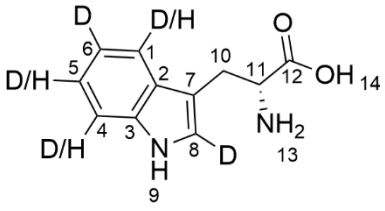
6.2.5.1 L-Trp-d_3 (**10**)



10

^{1}H NMR (400 MHz, D_2O) δ 7.62 (d, *J* = 2.2 Hz, 0.59H, H_1), 7.47 (q, *J* = 3.2 Hz, 0.46H, H_4), 7.26 (s, 0.02H, H_8), 7.21 (t, *J* = 4.2 Hz, 0.38H, H_5), 4.32 (ddd, *J* = 7.2, 5.3, 1.8 Hz, 1H, H_{11}), 3.51 - 3.32 (m, 2H, H_{10}). **^{13}C NMR** (151 MHz, D_2O) δ 172.4 (C_{12}), 136.4 (d, *J* = 9.7 Hz, C_3), 126.5 (d, *J* = 10.4 Hz, C_2), 125.6 - 124.9 (m, C_8), 122.3 - 121.4 (m, C_5), 119.6 - 118.8 (m, C_6), 118.1 (d, *J* = 1.8 Hz, C_1), 112.0 (d, *J* = 16.4 Hz, C_4), 106.3 (C_7), 53.7 (C_{11}), 25.9 (d, *J* = 2.5 Hz, C_{10}). **LC-MS** (ESI +) *m/z*: 208. $[M+H]^+$. **FT-IR** (ν_{max} / cm^{-1}): 3399m, 2981m (br), 1724s, 1650s, 1409s, 1148s, 1137s. NMR data for the undeuterated parts of the compound closely align with literature values.[232]

6.2.5.2 D-Trp-d_3 (**11**)



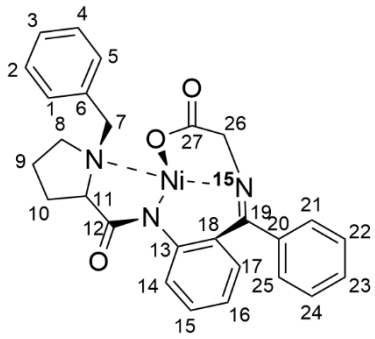
11

^{1}H NMR (400 MHz, D_2O) δ 7.63 (s, 0.47H, H_1), 7.46 (s, 0.39H, H_4), 7.21 (s, 0.28H, H_5), 4.30 (s, 1H, H_{11}), 3.44 (dd, *J* = 25.5, 11.3 Hz, 2H, H_{10}).**^{13}C NMR** (101 MHz, D_2O) δ 171.7 (C_{12}), 136.4 - 136.1 (m, C_3), 126.5 - 126.4 (m, C_2), 125.3 (C_8), 122.0 (t, *J* = 9.9 Hz, C_1), 119.7 - 119.1 (m, C_6), 118.1 (d, *J* = 11.5 Hz, C_5), 111.9 (d, *J* = 10.9 Hz, C_4), 106.1 (d, *J* = 17.2 Hz, C_7), 53.1 (C_{10}), 25.6 (C_{11}). **LC-MS** (ESI +) *m/z*: 209.3 $[M+H]^+$. NMR data for the undeuterated parts of the compound closely align with literature values.[232]

6.2.6 Successfully synthesised compounds excluded from further study

6.2.6.1 ^{15}N-Gly-Ni-(S)-2-[N-(N-benzylpropyl)amino]benzophenone (**12**)

A mixture of **2** (0.252 g, 0.657 mmol), nickel(II) nitrate hexahydrate (0.229 g, 0.789 mmol, 1.2 eq.), and ^{15}N-glycine (0.100 g, 1.315 mmol, 2 eq.) in anhydrous methanol (3 mL) was stirred under argon, at 60 °C. To this a solution of sodium methoxide (0.547 g, 10.124 mmol) in anhydrous methanol (8 mL) was added. This mixture was stirred for 90 minutes at 55 °C and then poured onto water (100 ml) with acetic acid (5 mL). A red precipitate was recovered by Bückner filtration and purified by silica plug using chloroform followed by chloroform: methanol (9:1 v/v) to give (0.19 g, 0.381 mmol, 58%).

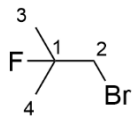


12

^{1}H NMR (400 MHz, $CDCl_3$) d 8.31 (d, *J* = 8.8 Hz, 1H, H_{23}), 8.13 - 8.06 (m, 2H, $H_{1,5}$), 7.53 (s, 3H, $H_{15,16,17}$), 7.45 (t, *J* = 7.8 Hz, 2H, $H_{2,4}$), 7.38 - 7.19 (m, 5H, benzyl),7.13 (d, *J* = 7.2 Hz, 1H, Ar), 4.52 (d, *J* = 12.7 Hz, 1H, H_{11}), 3.81 (d, *J* = 20.1 Hz, 1H, H_8), 3.70 (d, *J* = 15.5 Hz, 2H, H_{10}), 3.38 (d, *J* = 11.0 Hz, 1H, H_9), 2.60 (s, 1H, $H_{7'}$), 2.45 (s, 1H, $H_{7''}$), 2.15 - 2.10 (d, *J* = 11.3 Hz, 1H, $H_{26'}$), 2.10 (s, 1H, $H_{26''}$). **LC-MS** (ESI+) *m/z* = 500.2 $[M+2H]^+$. NMR data for the unlabelled compound closely align with literature values.[213]

6.2.6.2 Attempted synthesis of F-Leucine (**14**)

6.2.6.2.a Synthesis of 1-bromo-2-fluoro-2-methylpropane (**13**)



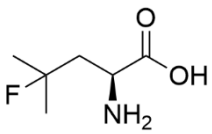
13

An excess of isobutene gas was bubbled through a stirred mixture of triethyl amine trihydrofluoride (14.7 mL, 90 mmol, 1.4 eq) and dichloroethane (60 mL). N-bromosuccinimide (NBS) (11.8 g, 66 mmol, 1 eq) was added at 0 °C. After 15 minutes of stirring the reaction mixture was brought up to room temperature and stirred for 16 hours. After this time, the reaction was quenched in slightly basic ice water (900 mL ice water: 25 mL aqueous ammonia [48%]) and extracted in dichloromethane. The combined organic layers

were washed with hydrochloric acid (2 x 150 mL, 0.1N) then a 5% aqueous solution of sodium carbonate and dried with magnesium sulphate. The solvent was removed *in vacuo* to give the product as a yellow oil (1.542 g, 9.95 mmol, 15%).

^{1}H NMR (400 MHz, CDCl3) δ 3.50 - 3.41 (m, 2H, H_2), 1.51 (d, *J* = 21.0 Hz, 6H, $H_{3,4}$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 43.6 (C_2), 39.9 (d, *J* = 28.7 Hz, C_1), 25.5 (d, *J* = 24.1 Hz, $C_{3,4}$). **^{19}F NMR** (376 MHz, $CDCl_3$) δ -138.6. **LC-MS** (ESI +) *m/z*: 154.5 $[M+H]^+$. **FT-IR** (ν_{max} / cm^{-1}): 2979m, 1708s, 1454m, 1428m, 1384m, 1270m, 1178m, 1100w, 1143w, 1036w, 668w.

6.2.6.2.b Synthesis of F-Leucine (**14**)

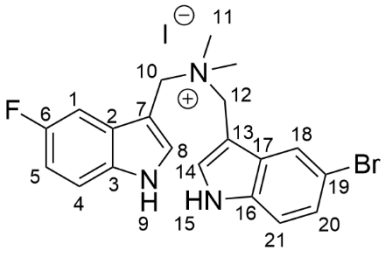


14

L-Leucine (0.05 g, 0.4 mmol), tetrabutylammonium decatungstate (TBADT) (0.066 g, 0.02 mmol, 0.05 eq) and N-fluorobenzenesulfonimide (NFSI) (0.378 g, 1.2 mmol, 3 eq) were added to a solution of acetonitrile and water (20:1 v/v, 1 mL) under UV irradiation (365 nm). After 20 hours, a colour change from yellow to blue was observed. Product observed in the crude mass spectrum. Isolation of the product was unsuccessful.

LC-MS (ESI +) *m/z*: 150.3 $[M+H]^+$.

6.2.6.3 1-(5-bromo-1H-indol-3-yl)-N-((5-fluoro-1H-indol-3-yl)methyl)-N,N-dimethylmethanaminium iodide (**15**)



15

5-Bromo gramine (4d) (0.076 g, 0.23 mmol, 1 eq) and **(5-**fluoro-3-indolylmethyl)trimethylammonium iodide (5a) (0.1 g, 0.23 mmol, 1 eq) were stirred with ethanol (10 mL) at 60 °C for 2 hours. After this time, an excess amount of diethyl ether was added, forming a precipitate (0.098 g, 0.22 mmol, 96%). The desired product was identified within the crude products. However, preparatory HPLC failed to isolate the product from the Fluoro-dimer and the Bromo dimer, which were also present. Proton NMR of the crude product is provided.

^{1}H NMR (400 MHz, DMSO) δ 11.75 (s, 1H, NH), 11.50 (s, 1H, NH), 7.80 (t, *J* = 3.1 Hz, 1H, Ar-H), 7.74 - 7.64 (m, 2H, Ar-H), 7.49 (dd, *J* = 8.9, 4.5 Hz, 1H, Ar-H), 7.41 - 7.33 (m, 2H, Ar-H), 7.04 (td, *J* = 9.2, 2.5 Hz, 1H, Ar-H), 6.84 (dd, *J* = 8.9, 2.3 Hz, 1H, Ar-H), 4.73 (d, *J* = 3.6 Hz, 4H, $H_{10,12}$), 2.84 (s, 6H, H_{11}). **LC-MS** (ESI +) *m/z*: 402.3 $[M+H]^+$.

6.3 Biological experimental

Plant trials were completed by Jahmel Neblett and Dr Clara Piccinini (Croda Europe Ltd.) at Rothamsted Research, Harpenden. The experimental section for the growth of the plants and application of treatment is based on information provided in the summary report of these trials, produced by Jahmel Neblett.

6.3.1 Preparation of treatment

Compounds for use as treatments were shipped to Rothamsted as pre-weighed pellets (Fig. 57a) or in an aqueous buffer solution of sodium bicarbonate. Deionised water was added to the samples to achieve a final volume of 80 mL, resulting in a 60 mM amino acid solution. The solutions were then sonicated in an ultrasonic bath for 2 minutes or until the concentrate was fully dissolved in the solution (Fig. 57b).

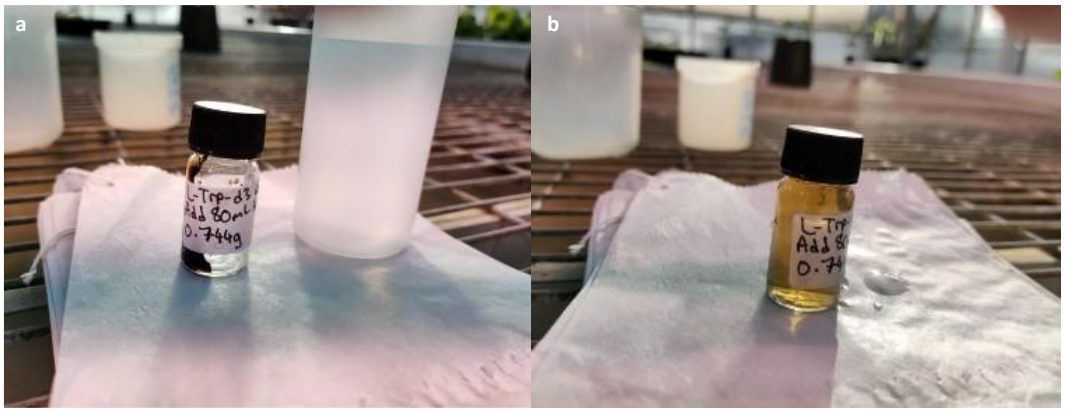


Figure 57: *Trp treatment as a pellet (a) and after the addition of water and sonication (b)Photographs provided by Jahmel Neblett*Plant preparation and treatment application

The glasshouse settings were 22 °C day/ 19 °C night with supplemental LED lighting for 14 hours of daylight. The dimensions of the pots selected were 9x9x14cm with 300g of Levington's M3 compost. Thirty soybean (*Glycine max*, Abelina variety) seeds

were sown per treatment to be run with two seeds sown per pot at a depth of 3cm. Plants were thinned to one plant per pot at the VC growing stage. More seeds were sown than the number of plants needed (24), so the most uniform plants could be selected for the experiment.

The biological controls used were Bioline (*Amblyseius cucumeris*) to control western flower thrips. Yellow sticky traps were used to control Scarid flies.

The desired growth stage for the application of the treatment was at V3 (where the third trifoliate of the plant is present (Fig. 58).

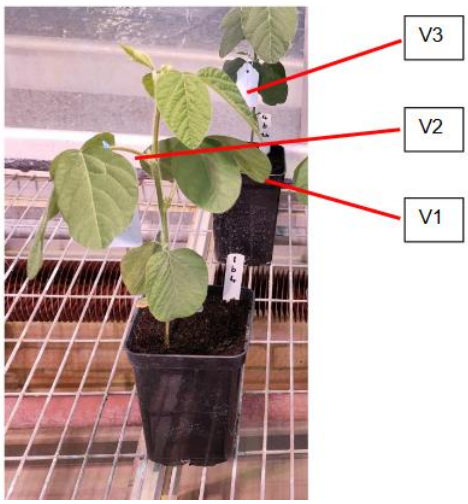


Figure 58: *Soybean plant at V3 growth stage, prior to application of amino acid treatment. Photograph provided by Jahmel Neblett*

The second trifoliate (T2) was isolated by using a plastic bag with a hole cut into the side to allow only T2 to be exposed (Fig. 59). This ensured treatment exposure to only the target leaf. The bags were not removed for approximately 1-hour following application to reduce contamination from any residual treatment solution.

Figure 59: *Soybean plant shortly after treatment application, with a protective bag to isolate the 2nd trifoliate.*

Treatments were applied using a custom designed microsprayer system to produce a fine mist, calibrated to spray an output of 200 $L.ha^{-1}$. There were 24 plants per treatment level, with 12 sprayed simultaneously. Once spray applications were complete and the plants dry, the bags were removed, and the plants were arranged into blocks to account for variation in the glasshouse.

Before application of treatment, plants were watered from above following treatment each plant was watered individually at the base to reduce risk of run-off.

There were six harvest time points where the trifoliate that received direct treatment were removed from the plant and placed in paper bags. The remaining part of the plant was placed in larger bags, and all parts dried in a drying oven set to 60 °C for 48 hours.

The six harvest time points were: 4 hours, 24 hours, 48 hours, 72 hours, 1 week, and 4 weeks post-application. After drying times these plants were shipped to Durham and stored in dry, cool dark conditions prior to extraction.

6.3.3 Preparation of plant extracts

From each plant, six or seven sections were taken (Fig. 60); an unwashed leaf from the site of treatment application (T2uw), a washed leaf from the site of treatment

application (T2w), a leaf from the trifoliate below the site of application (T1), a section of stem spanning from slightly below to slightly above the T2 trifoliate (Stem), a leaf from the trifoliate above the site of application (T3), the apical shoot meristem (MS), a Pod from plants old enough to have produced them (Pod).

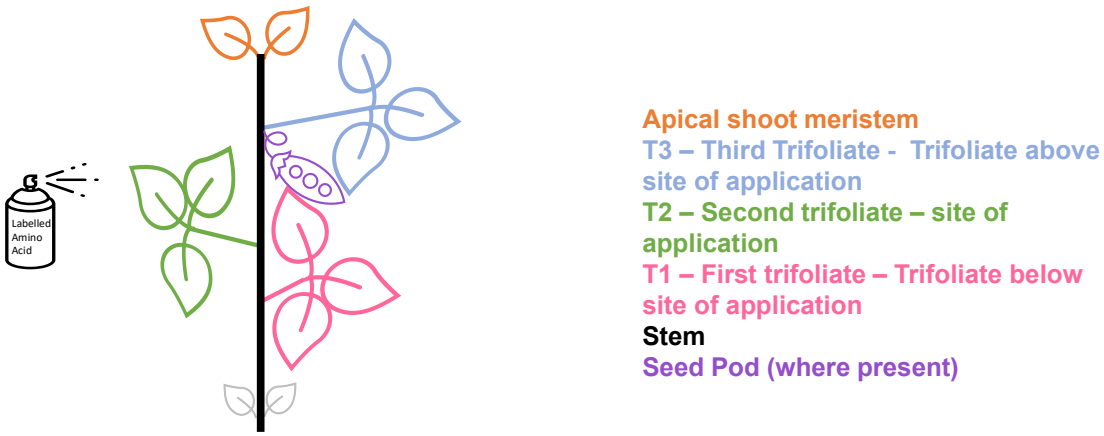


Figure 60: *Schematic illustration of plant sectioning for extraction. Sample regions include treated and untreated trifoliates, stem, meristem, and pod (where present).*

For one leaf at the site of application, a wash with 0.1 M HCl was done for 10 min to remove treatment from the outside of the leaf.

Samples were frozen using liquid nitrogen and ground using a pestle and mortar to a fine powder. The stem and pod samples were too hard and fibrous for pestle and mortar grinding, so a manual herb grinder (PPpanda Grinder for Dry Herb Portable Zinc Alloy Metal Grinder, 5x4 cm, Amazon UK) and an electric grinder (ELMWAY Electric Burr Coffee Grinder, Amazon UK) were purchased to produce a finer powder consistency.

Each plant material sample was placed in a paper filter bag (repurposed teabag filters, Whittard of Chelsea, UK) and a glass vial with DCM (6 mL). These vials were sonicated at 25 °C for 1 hour. After this time, the bags were removed from the DCM, dried, and placed in fresh vials with a hydrochloric acid solution (0.1 M, 4 mL). The vials were then sonicated again for 1 hour at 50 °C. Sample bags were removed, the solution was filtered through cotton wool to remove any remaining particulates, and then the extraction liquid was analysed using mass spectrometry.

6.4 Analytical experimental

6.4.1 Mass spectrometry protocol

The mass spectrometry protocol was developed in collaboration with and conducted by Mr Peter Stokes, Durham University. This setup was used for both full-scan and multiple-reaction monitoring (MRM) mass spectrometry experiments (Section 6.1.2).

6.4.2 Initial analytical methods

Initial analysis of the MRM data collected the intensity at the apex of peaks for *m/z* values of 205 (unlabelled Trp) and 208 (deuterium labelled Trp). These values were collated into an excel spreadsheet (Appendix D.1). The percentage of total Trp in each sample which was deuterated was calculated using Equation 4.1 (Section 4.1.4) but substituting the integrated peak areas with peak intensities instead.

6.4.3 R scripts

All mass spectrometry data files were converted from raw data files into mzML files using MSConvert (ProteoWizard, version 3.0.25015-b8edf2a, 64-bit) on a Windows platform.[233] Files were converted to mzML format with 64-bit binary encoding precision, zlib compression, and indexed output enabled. TPP compatibility mode was also selected. For most samples, only mass spectrometry level 1- data were retained, while for some targeted analyses, the ‘SRM as spectra’ option was enabled. Peak picking using the vendor algorithm was applied selectively where required. These files were then used within R to analyse the mass spectra. statistical analyses and data processing were conducted in R (version 4.4.2, "Pile of Leaves", R Foundation for Statistical Computing) on a Windows (x86_64-w64-mingw32) platform.

The R scripts were adapted from those initially written by John Sanderson, with assistance from AI models for debugging and developing script sections.[234] Full scripts are available in the appendices (Appendix D).

6.4.3.1 Rate of uptake

MRM mass spectrometry data from the site of application was processed in R using the mzR (Version 2.40.0) package. The script (Appendix D.1) assumes a folder structure of BaseDir/PlantPart/TimePoint/*.mzML and so can categorise the output based on both plant part and time point. It processes each file that has both indices 2 and 4, and calculates the AUC for each using a hybrid method (Gaussian fit with robust baseline; fallback to trapezoidal). It then computes the percentage of deuterated Trp as: $(auc_{index4} / (auc_{index2} + auc_{index4})) \times 100$ and produces a summary CSV file using dplyr (Version 1.1.4) with averages by plant part and timepoint. In addition, overlay plots of the smoothed chromatograms are generated using ggplot2 (Version 3.5.1).

6.4.3.2 Metabolomics

Initial methods for metabolomics search used an R script (Appendix D.3) to compare control and treatment mzml files and identify unique peaks in the treatment files. The package fuzzy joins (Version 0.1.6) was used for approximate matching of peaks, and the results are saved as CSV files for further analysis. The resultant CSV files were then used with another R script (Appendix D.4) to search for specific metabolites and plant growth hormones *m/z* values which matched those in the unique peaks files. EICs were then plotted for each matched *m/z* value using ggplot2 (Version 3.5.1).

6.4.3.2.a MZmine

Metabolite identification was also performed using MZmine (Version 4.5.0).[235] LC-MS data in mzML format were imported into MZmine and subjected to mass detection with a set noise threshold of 1.0E4 to remove background signals. Next the Chromatogram Builder module was used to create a feature list of chromatograms for different retention times. The resulting lists were then searched for specific *m/z* values corresponding to target metabolites identified in section 4.3.2. As well as this, total ion chromatograms (TICs) from treatment and control samples were overlayed to visually identify unique peaks. Finally extracted ion chromatograms (EICs) were generated for these candidate peaks to evaluate their identity.

References

1. T. A. Wise, *A scoping paper to assess the evidence,* Global Development and Environment Institute Working Paper 13-02, Tufts University, Medford MA, 2013.
2. M. C. Hunter, R. G. Smith, M. E. Schipanski, L. W. Atwood and D. A. Mortensen, *Bioscience*, 2017, **67**, 386-391.
3. A. Steensland, *Global Agricultural Productivity Report: Productivity Growth for Sustainable Diets, and More*, Virginia Tech College of Agriculture and Life Sciences, Report No. GAP Report 2019, 2019.
4. J. Agnew and S. Hendery, *Every Farmer Every Tool: 2023 Gap Report Executive Summary*, Virginia Tech College of Agriculture and Life Sciences, 2023.
5. Plant Impact Limited, *What We Do*, [online] Available at: https://www.plantimpact.com/about-us/what-we-do [Accessed 13 March 2020].
6. Y. Yang *et al.*, *Science*, 2024, **385**, eadn3747.
7. S. Perveen, N. Iqbal, M. Saeed, S. Zafar and Z. Arshad, *Brazilian Journal of Botany*, 2018, **41**, 805-815.
8. M. Habib-ur-Rahman, A. Ahmad, A. Raza, M. U. Hasnain, H. F. Alharby, Y. M. Alzahrani, A. A. Bamagoos, K. R. Hakeem, S. Ahmad and W. Nasim, *Front. Plant Sci.*, 2022, **13**, 925548.
9. Y. Rouphael and G. Colla, *Front. Plant Sci.*, 2018, **9**, 1655.
10. H. K. Jain, *Green Revolution: History, Impact and Future*, Studium Press LLC, Houston, TX, 2010.
11. M. V. Joshi, *Green-revolution and its impacts*, APH Publishing, New Delhi 1999.
12. L. F. Hesser, The Man Who Fed the World: Nobel Peace Prize Laureate Norman Borlaug and His Battle to End World Hunger: An Authorized Biography, Leon Hesser, Dallas, TX, 2006.
13. V. Shiva, *The violence of the green revolution: Third world agriculture, ecology, and politics*, University Press of Kentucky, 1991.
14. B. Glaeser, *The Green Revolution revisited*, Allen & Unwin, London, 1 edn., 1987.
15. T. M. Addiscott, A. P. Whitmore and D. S. Powlson, *Farming, fertilizers and the nitrate problem*, CAB International (CABI), 1991.
16. E. P. Rao and K. Puttanna, *Curr. Sci.*, 2000, **79**, 1163-1168.
17. G. Breitenbeck, A. Blackmer and J. Bremner, *Geophys. Res. Lett.*, 1980, **7**, 85-88.
18. P. Du Jardin, *Sci. Hortic.*, 2015, **196**, 3-14.
19. S. Corsi, G. Ruggeri, A. Zamboni, P. Bhakti, L. Espen, A. Ferrante, M. Noseda, Z. Varanini and A. Scarafoni, *Agronomy*, 2022, **12**, 1257.
20. O. Ali, A. Ramsubhag and J. Jayaraman, *Plants*, 2021, **10**, 531.
21. C. Xu and D. I. Leskovar, *Sci. Hortic.*, 2015, **183**, 39-47.
22. R. Haworth, *Soil Science*, 1971, **111**, 71-79.
23. H.-S. Shin, H. V.-M. Nguyen and J. Hur, in *Humus and Humic Substances - Recent Advances*, ed. A. Makan, IntechOpen, Rijeka, 2022, DOI: 10.5772/intechopen.105518.
24. J. S. Gaffney, N. A. Marley and S. B. Clark, ACS Publications, 1996.
25. A. Zavarzina, N. Danchenko, V. Demin, Z. Artemyeva and B. Kogut, *Eurasian Soil Sci.*, 2021, **54**, 1826-1854.
26. L. P. Canellas, F. L. Olivares, N. O. Aguiar, D. L. Jones, A. Nebbioso, P. Mazzei and A. Piccolo, *Sci. Hortic.*, 2015, **196**, 15-27.
27. L. P. Canellas and F. L. Olivares, *Chem. Biol. Technol. Agric.*, 2014, **1**, 1-11.
28. P. Rathor, P. Upadhyay, A. Ullah, L. Y. Gorim and M. S. Thilakarathna, *AoB Plants*, 2024, **16**, plae018.

29. J. Gerke, *J. Plant Nutr. Soil Sci.*, 2021, **184**, 329-338.
30. N. J. Ayon, *AIMS Mol. Sci.*, 2020, **7**, 229-268.
31. J. Hastings, G. Owen, A. Dekker, M. Ennis, N. Kale, V. Muthukrishnan, S. Turner, N. Swainston, P. Mendes and C. Steinbeck, *Nucleic acids Res.*, 2016, **44**, D1214-D1219.
32. D. C. Stateras and N. K. Moustakas, *J. Plant Nutr.*, 2018, **41**, 186-196.
33. P. Tsouvaltzis, A. Koukounaras and A. S. Siomos, *Int. J. Agric. Biol.*, 2014, **16**.
34. J. Cheng, J. Cheng, R. Sun, S. Cao, X. Wang and H. Yang, *Cogent Food Agric.*, 2024, **10**, 2321680.
35. S. Jahanbani, H. Mumivand, B. Zahedi and S. Argento, *Agronomy*, 2024, **14**, 2950.
36. A. Fersht, *Structure and mechanism in protein science: a guide to enzyme catalysis and protein folding*, Macmillan, New York, 1999.
37. M. SH Sadak, M. T. Abdelhamid and U. Schmidhalter, *Acta Biol. Colomb.*, 2015, **20**, 141-152.
38. A. Heidarzadeh, *Discover Plants*, 2025, **2**, 237.
39. L. Zhang, Q. Tan, R. Lee, A. Trethewy, Y.-H. Lee and M. Tegeder, *The Plant Cell*, 2010, **22**, 3603-3620.
40. T. Ohyama, N. Ohtake, K. Sueyoshi, Y. Ono, K. Tsutsumi, M. Ueno, S. Tanabata, T. Sato and Y. Takahashi, Amino Acid:New Insights Roles Plant Anim., 2017, 171-196.
41. K. Arand, D. Stock, M. Burghardt and M. Riederer, *J. Exp. Bot.*, 2010, **61**.
42. V. Fernández, H. A. Bahamonde, J. Javier Peguero-Pina, E. Gil-Pelegrín, D. Sancho-Knapik, L. Gil, H. E. Goldbach and T. Eichert, *J. Exp. Bot.*, 2017, **68**, 5293-5306.
43. J. C. Stiegler, M. D. Richardson, D. E. Karcher, T. L. Roberts and R. J. Norman, *Crop Sci.*, 2013, **53**, 1148-1152.
44. M. T. Tyree, T. D. Scherbatskoy and C. A. Tabor, *Plant Physiol.*, 1990, **92**, 103-109.
45. M. Tegeder and D. Rentsch, *Mol. Plant*, 2010, **3**, 997-1011.
46. A. Ortiz-Lopez, H. C. Chang and D. R. Bush, *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 2000, **1465**, 275-280.
47. X. Yao, J. Nie, R. Bai and X. Sui, *Plants*, 2020, **9**, 972.
48. I. T. Mertz, N. E. Christians and A. W. Thoms, *HortTechnology*, 2019, **29**, 833-837.
49. S. Khan, H. Yu, Q. Li, Y. Gao, B. N. Sallam, H. Wang, P. Liu and W. Jiang, *Agronomy*, 2019, **9**, 266.
50. V. Fernández and T. Eichert, *Crit. Rev. Plant Sci.*, 2009, **28**, 36-68.
51. J. P. Santos, P. Torres, A. M. Amorim, B. N. T. da Silva, D. Y. A. C. dos Santos and F. Chow, *Brazilian Journal of Botany*, 2023, 1-8.
52. M. F. Mondal, M. Asaduzzaman, H. Tanaka and T. Asao, *Sci. Hortic.*, 2015, **192**, 453-459.
53. M. A. M. El Hadi, F.-J. Zhang, F.-F. Wu, C.-H. Zhou and J. Tao, *Molecules*, 2013, **18**, 8200-8229.
54. W. Gou, L. Zhang, F. Chen, Z. Cui, Y. Zhao, P. Zheng, L. Tian, C. Zhang and L. Zhang, *Pak. J. Bot*, 2015, **47**, 2257-2262.
55. K. Matysiak, R. Kierzek, I. Siatkowski, J. Kowalska and R. Krawczyk, *Agronomy*, 2020, **10**, 769.
56. Y. A. Noroozlo, M. K. Souri and M. Delshad, *Adv. Hortic. Sci.*, 2019, **33**.
57. Y. A. Noroozlo, M. K. Souri and M. Delshad, *Open Agric.*, 2019, **4**, 164-172.
58. G. A. F. Hendry and A. K. Stobart, *Phytochemistry*, 1977, **16**, 1567-1570.
59. A. S. Khan, B. Ahmad, M. J. Jaskani, R. Ahmad and A. U. Malik, Int. J. Agric. Biol., 2012, **14**, 383-388.
60. P. Harinasut, K. Tsutsui, T. Takabe, M. Nomura, T. Takabe and S. Kishitani, *Biosci., Biotechnol., Biochem.*, 1996, **60**, 366-368.
61. W. Xing and C. Rajashekar, *Environ. Exp. Bot.*, 2001, **46**, 21-28.
62. C. Meek, D. Oosterhuis and J. Gorham, *Crop management*, 2003, **2**, 1-10.

63. F. Cruz, G. Castro, D. S. Júnior, R. Festucci-Buselli and H. Pinheiro, *Photosynthetica*, 2013, **51**, 102-108.
64. S. Lutts, V. Majerus and J. M. Kinet, *Physiol. Plant.*, 1999, **105**, 450-458.
65. C. Ben Ahmed, B. Ben Rouina, S. Sensoy, M. Boukhriss and F. Ben Abdullah, *J. Agric. Food Chem.*, 2010, **58**, 4216-4222.
66. R. Tonhati, S. C. Mello, P. Momesso and R. M. Pedroso, *Sci. Hortic.*, 2020, **268**, 109370.
67. M. Ozden, U. Demirel and A. Kahraman, *Sci. Hortic.*, 2009, **119**, 163-168.
68. W. Claussen, *Plant Science*, 2005, **168**, 241-248.
69. M. Ashraf and M. R. Foolad, *Environ. Exp. Bot.*, 2007, **59**, 206-216.
70. M. N. Alyemeni, Q. Hayat, S. Hayat, M. Faizan and A. Faraz, *Legume Res.*, 2016, **39**.
71. Q. Ali, M. Ashraf and H.-U.-R. Athar, *Pak. J. Bot.*, 2007, **39**, 1133-1144.
72. H. Yaqoob, N. A. Akram, S. Iftikhar, M. Ashraf, N. Khalid, M. Sadiq, M. N. Alyemeni, L. Wijaya and P. Ahmad, *Plants*, 2019, **8**, 588.
73. X. Liang, L. Zhang, S. K. Natarajan and D. F. Becker, *Antioxid. Redox Signaling*, 2013, **19**, 998-1011.
74. M. Takeuchi, C. Arakawa, Y. Kuwahara and H. Gemma, *Acta Hortic.* 2007, **800**, 549-554.
75. Ž. Tarasevičienė, A. Velička and A. Paulauskienė, *Plants*, 2021, **10**, 599.
76. G. Gutiérrez-Gamboa, T. Garde-Cerdán, J. Portu, Y. Moreno-Simunovic and A. M. Martínez-Gil, *Food Res. Int.*, 2017, **96**, 46-53.
77. T. Garde-Cerdán, R. López, J. Portu, L. González-Arenzana, I. López-Alfaro and P. Santamaría, *Food Chem.*, 2014, **163**, 136-141.
78. T. Garde-Cerdán, P. Santamaría, P. Rubio-Bretón, L. González-Arenzana, I. López-Alfaro and R. López, *LWT-Food Science and Technology*, 2015, **60**, 684-689.
79. S. Marín-San Román, T. Garde-Cerdán, E. Baroja, P. Rubio-Bretón and E. P. Pérez-Álvarez, *Sci. Hortic.*, 2020, **272**, 109515.
80. J. Portu, I. López-Alfaro, S. Gómez-Alonso, R. López and T. Garde-Cerdán, *Food Chem.*, 2015, **180**, 171-180.
81. X. Cheng, X. Wang, A. Zhang, P. Wang, Q. Chen, T. Ma, W. Li, Y. Liang, X. Sun and Y. Fang, *J. Agric. Food Chem.*, 2020, **68**, 15390-15402.
82. H. Wang, G. Cao and R. L. Prior, *J. Agric. Food Chem.*, 1997, **45**, 304-309.
83. P. Waffo Teguo, B. Fauconneau, G. Deffieux, F. Huguet, J. Vercauteren and J.-M. Mérillon, *J. Nat. Prod.*, 1998, **61**, 655-657.
84. Y. Zhao, *Molecular plant*, 2012, **5**, 334-338.
85. A. Mustafa, A. Hussain, M. Naveed, A. Ditta, Z.-e.-H. Nazli and A. Sattar, *Soil Environ.*, 2016, **35**, 76-84.
86. A. Mustafa, M. Imran, M. Ashraf and K. Mahmood, *Pedosphere*, 2018, **28**, 16-34.
87. Y. A. Khalifa, A. El-Naem, F. Gamal and M. A. Mahmoud, *Egypt. J. Agron.*, 2020, **42**, 47-61.
88. S. Rao, A. Qayyum, A. Razzaq, M. Ahmad, I. Mahmood and A. Sher, *J. Anim. Plant Sci*, 2012, **22**, 768-772.
89. A. Sanada and S. Agehara, *Plants*, 2023, **12**, 186.
90. W. F. Mosa, H. M. Ali and N. R. Abdelsalam, *Environ. Sci. Pollut. Res.*, 2021, **28**, 1983-1991.
91. S. El-sharabasy, F. Issa, G. Hammad and M. El-Dawayaty, *Egypt. J. Genet. Cytol.,* 2016, **44**.
92. S. S. Sharma and K.-J. Dietz, *J. Exp. Bot.*, 2006, **57**, 711-726.
93. M. Rizwan, S. Ali, M. Z. Akbar, M. B. Shakoor, A. Mahmood, W. Ishaque and A. Hussain, *Environ. Sci. Pollut. Res.*, 2017, **24**, 21938-21947.

94. D. Lv, C. Yu, L. Yang, S. Qin, H. Ma, G. Du, G. Liu and S. Khanizadeh, *Eur. J. Hortic. Sci.*, 2009, **74**, 204.
95. S. A. Akladious and S. M. Abbas, *Bangladesh Journal of Botany*, 2013, **42**, 31-44.
96. E. A. A. Abou-Zaid and M. A. Eissa, *J. Soil Sci. Plant Nutr.*, 2019, **19**, 725-733.
97. Y. Cao, Z. Gao, J. Li, G. Xu and M. Wang, *Acta Hortic.*, 2006, **856**, 91–98.
98. R. M. McCoy, G. W. Meyer, D. Rhodes, G. C. Murray, T. G. Sors and J. R. Widhalm, *Agronomy*, 2020, **10**, 358.
99. M. A. Gagne, R. Minocha, S. Long and S. C. Minocha, Can. J. For. Res., 2019, **49**, 1548-1559.
100. G. Hoza, M. M. Jasim, G. Neata, M. Dinu, A. Becherescu, I. A. Apahidean and H. H. Shalal, *Sci. Pap. Ser. B Hortic.*, 2019, **63**, 417-423.
101. S. Ghasemi, A. H. Khoshgoftarmanesh, M. Afyuni and H. Hadadzadeh, *Eur. J. Agron.*, 2013, **45**, 68-74.
102. S. Ghasemi, A. H. Khoshgoftarmanesh, M. Afyuni and H. Hadadzadeh, *Sci. Hortic.*, 2014, **165**, 91-98.
103. J. Singh and B. Dhankhar, *Veg. Sci.*, **1991**, *18*, 110–113.
104. M. Rafie, A. Khoshgoftarmanesh, H. Shariatmadari, A. Darabi and N. Dalir, *Sci. Hortic.*, 2017, **216**, 160-168.
105. W. F. Teixeira, E. B. Fagan, L. H. Soares, R. C. Umburanas, K. Reichardt and D. D. Neto, *Front. Plant Sci.*, 2017, **8**.
106. W. F. Teixeira, L. H. Soares, E. B. Fagan, S. da Costa Mello, K. Reichardt and D. Dourado-Neto, J. Plant Growth Regul., 2019, 1-15.
107. M. Alfosea-Simón, E. A. Zavala-Gonzalez, J. M. Camara-Zapata, J. J. Martínez-Nicolás, I. Simón, S. Simón-Grao and F. García-Sánchez, *Sci. Hortic.*, 2020, **272**, 109509.
108. G.-f. FU, S. Jian, J. Xiong, Y.-r. LI, H.-z. CHEN, M.-k. LE and L.-x. TAO, *Agric. Sci. China*, 2011, **10**, 1016-1025.
109. M. H. Aminifard and A. Jorkesh, *Z. Arznei-Gewurzpflanzen*, 2019, **24**, 189-192.
110. C. Kaya, S. Aydemir, O. Sonmez, M. Ashraf and M. Dikilitas, *Acta Bot. Croat.*, 2013, **72**, 157-168.
111. J. A. Roberts, *Plant growth regulators*, Springer Science & Business Media, 2012.
112. S. Abel and A. Theologis, Cold Spring Harb. Perspect. Biol., 2010, **2**, a004572.
113. M. Sarwar and R. J. Kremer, *Plant Soil*, 1995, **172**, 261-269.
114. A. Bajguz and A. Piotrowska-Niczyporuk, *Metabolites*, 2023, **13**, 884.
115. T.-p. Sun, in *Plant hormones: biosynthesis, signal transduction, action!*, ed. P. J. Davies, Springer Science & Business Media, 2004, pp. 308-328.
116. F. N. Ritonga, D. Zhou, Y. Zhang, R. Song, C. Li, J. Li and J. Gao, *Plants*, 2023, **12**, 1243.
117. S. M. Southwick and K. Glozer, *HortTechnology*, 2000, **10**, 744-751.
118. B. Vishal and P. P. Kumar, *Front. Plant Sci.*, 2018, **9**, 838.
119. A. Osugi and H. Sakakibara, *BMC Bio.*, 2015, **13**, 102.
120. S. Gajdošová, L. Spíchal, M. Kamínek, K. Hoyerová, O. Novák, P. I. Dobrev, P. Galuszka, P. Klíma, A. Gaudinová and E. Žižková, *J. Exp. Bot.*, 2011, **62**, 2827-2840.
121. R. Mertens, B. Deus-Neumann and E. Weiler, *FEBS letters*, 1983, **160**, 269-272.
122. P. F. Wareing, *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 1978, **284**, 483-498.
123. E. Nambara and A. Marion-Poll, *Annu. Rev. Plant Biol.*, 2005, **56**, 165-185.
124. K. Bürstenbinder, G. Rzewuski, M. Wirtz, R. Hell and M. Sauter, *Plant J*, 2007, **49**, 238-249.
125. N. Iqbal, N. A. Khan, A. Ferrante, A. Trivellini, A. Francini and M. Khan, *Front. Plant Sci.*, 2017, **8**, 475.
126. F. Fiorani, G. M. Bogemann, E. J. Visser, H. Lambers and L. A. Voesenek, *Plant Physiol.*, 2002, **129**, 1382-1390.

127. T. Eichert and V. Fernández, in *Marschner's Mineral Nutrition of Plants (Fourth Edition)*, eds. Z. Rengel, I. Cakmak and P. J. White, Academic Press, San Diego, 2023, DOI: https://doi.org/10.1016/B978-0-12-819773-8.00014-9, pp. 105-129.
128. X. Ma, H. Wu, L. Liu, Q. Yao, S. Wang, R. Zhan, S. Xing and Y. Zhou, *Sci. Hortic.*, 2011, **129**, 102-107.
129. G. Blaser, J. M. Sanderson, A. S. Batsanov and J. A. Howard, *Tetrahedron Lett.*, 2008, **49**, 2795-2798.
130. A. Strecker, *Justus Liebigs Ann. Chem.*, 1850, **75**, 27-45.
131. L. Yet, *Angew. Chem., Int. Ed*, 2001, **40**, 875-877.
132. U. Schöllkopf, U. Groth and C. Deng, *Angew. Chem., Int. Ed. Engl.*, 1981, **20**, 798-799.
133. J. Ji, C. Chen, J. Cai, X. Wang, K. Zhang, L. Shi, H. Lv and X. Zhang, *Org. Biomol. Chem.*, 2015, **13**, 7624-7627.
134. R. Noyori and S. Hashiguchi, *Acc. Chem. Res.*, 1997, **30**, 97-102.
135. Y. N. Belokon, V. I. Bakhmutov, N. I. Chernoglazova, K. A. Kochetkov, S. V. Vitt, N. S. Garbalinskaya and V. M. Belikov, *J. Chem. Soc., Perkin Trans. 1*, 1988, 305-312.
136. Y. N. Belokon, S. Harutyunyan, E. V. Vorontsov, A. S. Peregudov, V. N. Chrustalev, K. A. Kochetkov, D. Pripadchev, A. S. Sagyan, A. K. Beck and D. Seebach, *Arkivoc*, 2004, **2004**, 132-150.
137. Y. N. Belokon, V. I. Tararov, V. I. Maleev, T. F. Savel'eva and M. G. Ryzhov, *Tetrahedron: Asymmetry*, 1998, **9**, 4249-4252.
138. Y. Belokon, *Janssen Chim. Acta*, 1992, **10**, 4-4.
139. Y. Xu, I. Correia, T. Ha-Duong, N. Kihal, J.-L. Soulier, J. Kaffy, B. Crousse, O. Lequin and S. Ongeri, *Beilstein J. Org. Chem.*, 2017, **13**, 2842-2853.
140. T. Pillaiyar, E. Gorska, G. Schnakenburg and C. E. Müller, *J. Org. Chem.*, 2018, **83**, 9902-9913.
141. J. W. Lown and G. L. Weir, *Can. J. Chem.*, 1978, **56**, 249-257.
142. J. J. Füller, R. Röpke, J. Krausze, K. E. Rennhack, N. P. Daniel, W. Blankenfeldt, S. Schulz, D. Jahn and J. Moser, *J. Bio. Chem.,* 2016, **291**, 20068-20084.
143. T. Geissman and A. Armen, *J. Am. Chem. Soc.*, 1952, **74**, 3916-3919.
144. L. L. Melhado, *The Journal of Organic Chemistry*, 1981, **46**, 1920-1923.
145. A. Bauzá, D. Quiñonero, A. Frontera and P. M. Deyà, *Physical Chemistry Chemical Physics*, 2011, **13**, 20371-20379.
146. H. Szatylowicz, A. Jezuita and T. M. Krygowski, *Structural Chemistry*, 2019, **30**, 1529-1548.
147. B. Cordero, V. Gómez, A. E. Platero-Prats, M. Revés, J. Echeverría, E. Cremades, F. Barragán and S. Alvarez, *Dalton Transactions*, 2008, 2832-2838.
148. C. Schopf and J. Thesing, *Angew. Chem*, 1951, **63**, 377-378. (Translated from German by J. Backhaus)
149. Y. Hai, M. Jenner and Y. Tang, *J. Am. Chem. Soc.*, 2019, **141**, 16222-16226.
150. E. Winnicka and M. Kańska, *J. Radioanal. Nucl. Chem.*, 2009, **279**, 675-678.
151. D. S. Wishart, C. Knox, A. C. Guo, R. Eisner, N. Young, B. Gautam, D. D. Hau, N. Psychogios, E. Dong and S. Bouatra, *Nucleic Acids Res.*, 2009, **37**, D603-D610.
152. K.-C. Liu, P.-S. Lee, S.-Y. Wang, Y.-S. E. Cheng, J.-M. Fang and C.-H. Wong, *Bioorg. Med. Chem.*, 2011, **19**, 4796-4802.
153. G. Haufe, K. W. Laue, M. U. Triller, Y. Takeuchi and N. Shibata, *Tetrahedron*, 1998, **54**, 5929-5938.
154. M. B. Nodwell, H. Yang, M. Čolović, Z. Yuan, H. Merkens, R. E. Martin, F. Bénard, P. Schaffer and R. Britton, *J. Am. Chem. Soc.*, 2017, **139**, 3595-3598.
155. W. F. Teixeira, E. B. Fagan, L. H. Soares, J. N. Soares, K. Reichardt and D. D. Neto, *Front. Plant Sci.*, 2018, **9**, 396.

156. J. Sun, Y. Zhao, X. Qin, Z. Hu, J. Li, Y. Guo, G. Sun, Z. Chen, H. Huang and C. Hu, *Front. Agronomy*, 2024, **6**, 1427571.
157. M. Koornneef and D. Meinke, *The Plant Journal*, 2010, **61**, 909-921.
158. H. Sun, X. Sun, H. Wang and X. Ma, *Hereditas*, 2020, **157**, 5.
159. R. Bozhinova, *Bulg. J. Agric. Sci.*, 2023, **29**.
160. R. Bozhinova, Y. Dyulgerski, V. Nikolova and N. Nikolov, *Bulg. J. Soil Sci. Agrochem. Ecol.*, 2024, **58**, 16-26.
161. H. Cui, S.-T. Zhang, H.-J. Yang, H. Ji and X.-J. Wang, *BMC Plant Biology*, 2011, **11**, 76.
162. S. P. Kashyap, N. Kumari, P. Mishra, D. P. Moharana and M. Aamir, *Genet. Resour. Crop Evol.*, 2021, **68**, 2207-2233.
163. R. V. Kesseli, I. Paran and R. W. Michelmore, *Genetics*, 1994, **136**, 1435-1446.
164. B. S. Landry, R. V. Kesseli, B. Farrara and R. W. Michelmore, *Genetics*, 1987, **116**, 331-337.
165. M. V. Shatilov, A. F. Razin and M. I. Ivanova, *IOP Conf. Ser.: Earth Environ. Sci.*, 2019, 395, 012001.
166. M. Davey, P. Anthony, P. Van Hooff, J. Power and K. Lowe, in *Transgenic Crops IV*, Springer, 2007, pp. 221-249.
167. Á. Kovácsné Madar and M. Takácsné Hájos, *Int. J. Hortic. Sci.*, 2022, **28**, 50–56, DOI: 10.31421/ijhs/28/2022/10314.
168. C. G. Abdel, *Mesopotamia J. Agric.*, 2005, **33**, 2.0-7.0.
169. M. Gallardo, L. Jackson, K. Schulbach, R. Snyder, R. Thompson and L. Wyland, *Irrig. Sci.*, 1996, **16**, 125-137.
170. F. A. S. U.S Department of Agriculture, Production - Soybeans (commodity code 2222000), https://www.fas.usda.gov/data/production/commodity/2222000, (accessed 12th September, 2025).
171. R. M. Stupar and J. E. Specht, *Adv. Agron.*, 2013, **118**, 177-204.
172. J. Azmir, I. S. M. Zaidul, M. M. Rahman, K. M. Sharif, A. Mohamed, F. Sahena, M. H. A. Jahurul, K. Ghafoor, N. A. Norulaini and A. K. Omar, *J. Food Eng.*, 2013, **117**, 426-436.
173. N. Azwanida, *Journal*, 2015.
174. P. S. Vankar, *Resonance*, 2004, **9**, 30-41.
175. D. A. Starmans and H. H. Nijhuis, *Trends Food Sci. Technol.*, 1996, **7**, 191-197.
176. P. Safaiee, A. Taghipour, F. Vahdatkhoram and K. Movagharnejad, *Chem. Pap.*, 2019, **73**, 3067-3073.
177. C. W. Huie, Anal. Bioanal. Chem., 2002, **373**, 23-30.
178. G. Alvarez-Rivera, M. Bueno, D. Ballesteros-Vivas, J. A. Mendiola and E. Ibañez, in *Liquid-phase extraction*, Elsevier, 2020, pp. 375-398.
179. M. L. De Castro and F. Priego-Capote, *J. Chromatogr. A*, 2010, **1217**, 2383-2389.
180. S. Vitalini, M. Dei Cas, F. M. Rubino, I. Vigentini, R. Foschino, M. Iriti and R. Paroni, *Molecules*, 2020, **25**, 311.
181. K. Kumar, S. Srivastav and V. S. Sharanagat, *Ultrason Sonochem*, 2021, **70**, 105325.
182. M. I. Prawira-Atmaja and S. Puangpraphant, *J. Saudi Soc. Agric. Sci.*, 2025, **24**, 10.
183. N. Gierlinger and M. Schwanninger, *J. Spectrosc.*, 2007, **21**, 69-89.
184. A. Saletnik, B. Saletnik and C. Puchalski, *Molecules*, 2022, **27**, 4454.
185. Waters Corporation, Atmospheric Solids Analysis Probe (ASAP), https://www.waters.com/nextgen/gb/en/services/instrument-services/instrument-upgrades/mass-spectrometry-ms-upgrades/atmospheric-solids-analysis-probe.html, (accessed 01 March, 2025).
186. A. Yerlekar and M. Kshirsagar, *J. Eng. Res. Appl*, 2014, **4**, 17-23.
187. K. Oetjen, Mastering LC-MS/MS: Essential Fundamentals and Theory with SCIEX (LC-MS/MS 101), https://www.youtube.com/watch?v=uvn6fg9id3Q, (accessed 08 January, 2025).

188. B. O. Keller, B. T. Wu, S. S. Li, V. Monga and S. M. Innis, *J. Agric. Food Chem.*, 2013, **61**, 7654-7660.
189. E. Sarhat, *Foliar Nutrition*, [Publisher Unknown], 2002.
190. M. Trovato, D. Funck, G. Forlani, S. Okumoto and R. Amir, *Front. Plant Sci.*, 2021, **12**, 772810.
191. A. Khan, M. Numan, A. L. Khan, I.-J. Lee, M. Imran, S. Asaf and A. Al-Harrasi, *Plants*, 2020, **9**, 407.
192. N. Khan, S. Ali, P. Zandi, A. Mehmood, S. Ullah, M. Ikram, I. Ismail and M. Babar, *Pak. J. Bot*, 2020, **52**, 355-363.
193. E. C. Tredenick, T. W. Farrell, W. A. Forster and S. T. Psaltis, *Front. Plant Sci.*, 2017, **8**, 746.
194. E. R. Radwanski and R. L. Last, *The Plant Cell*, 1995, **7**, 921.
195. F. Westphal and T. Junge, *Forensic Sci. Int.*, 2012, **223**, 97-105.
196. D. Asakawa, H. Mizuno, E. Sugiyama and K. Todoroki, *Analyst*, 2021, **146**, 2292-2300.
197. I. Sadok, A. Gamian and M. M. Staniszewska, *J. Sep. Sci.*, 2017, **40**, 3020-3045.
198. F. P. Carrera, C. Noceda, M. G. Maridueña-Zavala and J. M. Cevallos-Cevallos, *Agronomy*, 2021, **11**, 824.
199. R. Kumar, A. Bohra, A. K. Pandey, M. K. Pandey and A. Kumar, *Front. Plant Sci.*, 2017, **8**, 1302.
200. S. Heuckeroth, T. Damiani, A. Smirnov, O. Mokshyna, C. Brungs, A. Korf, J. D. Smith, P. Stincone, N. Dreolin and L.-F. Nothias, *Nat. Protoc.*, 2024, **19**, 2597-2641.
201. A. Pautova, Z. Khesina, M. Getsina, P. Sobolev, A. Revelsky and N. Beloborodova, *Molecules*, 2020, **25**, 3258.
202. H. Zhao, M. Jiang, Q. Liang, C. Xie, S. Song, J. Wang, G. Bai and G. Luo, *Int. J. Mass Spectrom.*, 2013, **335**, 7-15.
203. M. M. A. Khan, A. Khatun and M. T. Islam, *Microbes in Action; Nova Science Publishers: New York, NY, USA*, 2016, 1-43.
204. Y. Ren, Z. Zhang, D. Zhanakhmetova, W. Li, S. Chen, T. Werner and J. Liesche, *The Plant Journal*, 2024, **120**, 2305-2318.
205. A. Walter, L. Caputi, S. O'Connor, K.-H. van Pée and J. Ludwig-Müller, *Int. J. Mol. Sci.*, 2020, **21**, 2567.
206. Y. Zhao, *Annu. Rev. Plant Biol.*, 2010, **61**, 49-64.
207. A. C. Leopold, *Auxins and plant growth*, Univ of California Press, 1955.
208. J. Markwell, *The American Biology Teacher*, 1999, **61**, 672-676.
209. W. C. Still, M. Kahn and A. Mitra, *The Journal of Organic Chemistry*, 1978, **43**, 2923-2925.
210. M. C. Pirrung, *The synthetic organic chemist's companion*, John Wiley & Sons, 2007.
211. O. V. Dolomanov, L. J. Bourhis, R. J. Gildea, J. A. Howard and H. Puschmann, *J. Appl. Crystallogr.*, 2009, **42**, 339-341.
212. S. A. Yakukhnov and V. P. Ananikov, *Adv. Synth. Catal.*, 2019, **361**, 4781-4789.
213. W. Vuong, F. Mosquera-Guagua, R. Sanichar, T. R. McDonald, O. P. Ernst, L. Wang and J. C. Vederas, *Org. Lett.*, 2019, **21**, 10149-10153.
214. S. Nayab, H.-I. Lee and J. H. Jeong, *Structure Reports*, 2011, **67**, o2478-o2478.
215. Y. N. Belokon, A. G. Bulychev, S. V. Vitt, Y. T. Struchkov, A. S. Batsanov, T. V. Timofeeva, V. A. Tsyryapkin, M. G. Ryzhov and L. A. Lysova, *J. Am. Chem. Soc.*, 1985, **107**, 4252-4259.
216. L. Li, B. C. Chenna, K. S. Yang, T. R. Cole, Z. T. Goodall, M. Giardini, Z. Moghadamchargari, E. A. Hernandez, J. Gomez and C. M. Calvet, *J. Med. Chem.*, 2021, **64**, 11267-11287.
217. H. Kaur, P. W. R. Harris, P. J. Little and M. A. Brimble, *Org. Lett.*, 2015, **17**, 492-495.
218. A. Gouyette, *Eur. J. Med. Chem.*, 1980, **15**, 503.

219. H. Fujita, R. Nishikawa, O. Sasamoto, M. Kitamura and M. Kunishima, *J. Org. Chem.*, 2019, **84**, 8380-8391.
220. R. Lajarin-Cuesta, C. Nanclares, J.-A. Arranz-Tagarro, L. Gonzalez-Lafuente, R. L. Arribas, M. Araujo de Brito, L. Gandia and C. de Los Ríos, *J. Med. Chem.*, 2016, **59**, 6265-6280.
221. S.-G. Li and S. Z. Zard, *Org. Lett.*, 2013, **15**, 5898-5901.
222. M. Bandini and A. Eichholzer, *Angew. Chem.*, 2009, **121**, 9697-9701.
223. Bergman, *Acta Chem. Scand.*., 1971, **25**, 3296.
224. J. B. Hester Jr, *J. Org. Chem.*, 1964, **29**, 1158-1160.
225. I. Hernández, G. Domínguez, V. A. Soloshonok, A. Landa and M. Oiarbide, *J. Org. Chem.*, 2024, **89**, 17291-17309.
226. J. Brauer, R. Wiechert, A. Hahn and T. Opatz, *Org. Lett.*, 2024, **26**, 4314-4317.
227. M. Somei and M. Tsuchiya, *Chem. Pharm. Bull.*, 1981, **29**, 3145-3157.
228. S. V. Ley and A. Priour, *Eur. J. Org. Chem.*, 2002, **2002**, 3995-4004.
229. V. D. Nimbarte, J. Wirmer-Bartoschek, S. L. Gande, I. Alshamleh, M. Seibert, H. R. Nasiri, F. Schnütgen, H. Serve and H. Schwalbe, *ChemMedChem*, 2021, **16**, 1667-1679.
230. M. Reinfelds, K. Kalinins, D. Katkevica, R. Zemribo and M. Katkevics, *Tetrahedron Lett.*, 2015, **56**, 5882-5885.
231. N. Netz and T. Opatz, *J. Org. Chem.*, 2016, **81**, 1723-1730.
232. S. Liu, C. Sun, Y. Ha, M. Ma, N. Wang, Y. Zhou and Z. Zhang, *Tetrahedron Lett.*, 2024, **137**, 154935.
233. M. C. Chambers, B. Maclean, R. Burke, D. Amodei, D. L. Ruderman, S. Neumann, L. Gatto, B. Fischer, B. Pratt and J. Egertson, *Nat. Biotechnol.*, 2012, **30**, 918-920.
234. OpenAi, 2025.
235. R. Schmid, S. Heuckeroth, A. Korf, A. Smirnov, O. Myers, T. S. Dyrlund, R. Bushuiev, K. J. Murray, N. Hoffmann and M. Lu, *Nat. Biotechnol.*, 2023, **41**, 447-449.

Appendices

Appendix A: Control conditions in amino acid literature review

Title of Paper	**Reference**	**Amino Acid Discussed**	**Notes on Controls**
Effect of foliar application of amino acids on plant yield and some physiological parameters in bean plants irrigated with seawater.	SH Sadak, M., Abdelhamid, M. T. & Schmidhalter, U. *Acta Biol. Colomb.* **20**, 141-152 (2015).	mixtures	Control plants irrigated with tap water. Daytime temperatures ranged from 14.5 to 30.2 ◦C with an average of 23.2 ± 3.8 ◦C whereas temperatures at night were 12.4 ± 1.8 ◦C, with minimum and maximum of 8.0 and 17.6 ◦C, respectively. Daily relative humidity averaged 57.7± 9.6% in a range between from 38.1 to 78.7%.
Foliar absorption of various inorganic and organic nitrogen sources by creeping bentgrass.	Stiegler, J. C., Richardson, M. D., Karcher, D. E., Roberts, T. L. & Norman, R. J. *Crop Sci.* **53**, 1148-1152 (2013).	Glycine, Glutamic acid, proline	Control plants irrigated with tap water (compared to saline). Soil moisture controlled by pot weight and water loss supplemented twice daily
Branched-chain amino acids for use as a nitrogen source on creeping bentgrass.	Mertz, I. T., Christians, N. E. & Thoms, A. W. *HortTechnology* **29**, 833-837 (2019).	Leucine, Isoleucine, vaine	Control plants not treated with nitrogen fertilizer

Exogenous Application of Amino Acids Improves the Growth and Yield of Lettuce by Enhancing Photosynthetic Assimilation and Nutrient Availability.	Khan, S. *et al*. *Agronomy* **9**, 266 (2019).	Methionine, glycine, Trp	The concentration of L-methionine applied was 2200 mg/L, 220 mg/L, and 22 mg/L for the treatment and 2.2 mg/L, 0.2 mg/L, and 0.02 mg/L for the control. Average minimum and maximum monthly temperatures were set to 24 °C and 34 °C. Plants were provided with natural sunlight with a light intensity of approximately 900-1000 µmoles m−2/s. At pre-emergence stages, the nutrient solution was applied once a week. Plants with at least 2 fully expanded leaves 30 days after sowing were transferred to a closed-loop hydroponic system.
Application of amino acids improves lettuce crop uniformity and inhibits nitrate accumulation induced by the supplemental inorganic nitrogen fertilization. International	Tsouvaltzis, P., Koukounaras, A. & Siomos, A. S. *Journal of Agriculture & Biology* **16** (2014).	mixtures	Average temperature min and max maintained (24-34 C). Light intensity measured. pH and electrical conduction of nutrient solution maintained. All plants were transferred to nutrient solution 8 days before amino acid application to prevent nutrient shock. Control plant used for both experimental sections.
Effects of amino acids on the growth and flowering of Eustoma grandiflorum under autotoxicity in closed hydroponic culture.	Mondal, M. F., Asaduzzaman, M., Tanaka, H. & Asao, T.Sci. Hortic. 192, 453-459, doi:https://doi.org/10.1016/j.scienta.2015.05.024 (2015).	Alanine, Arginine, Asparagine, Aspartic acid, Cysteine, Glutamatic acid, Glutamine, Glycine, Lysine, Proline, Serine, Threonine, Trp, Methionine , Leucine, Isoleucine, Histidine, Phenylalanine, Valine	Distilled water as control treatment. Growth chamber used to ensure controlled light conditions, humidity and temperature,

Foliar application of amino acids modulates aroma components of'fuji'apple (Malus Domestica L.)	Gou, W. *et al.* . *Pak. J. Bot* **47**, 2257-2262 (2015).	Leucine, isoleucine, valine, alanine	Control mentioned but not adequately explained
Effect of Exogenous Application of Amino Acids L-Arginine and Glycine on Maize under Temperature Stress.	Matysiak, K., Kierzek, R., Siatkowski, I., Kowalska, J. & Krawczyk, R. *Agronomy* **10**, 769 (2020).	Arginine, Glycine	Control plants had seeds soaked in pure distilled water rather than amino acid treatments. Foliar treatment with water instead of amino acids
Effects of foliar application of glycine and glutamine amino acids on growth and quality of sweet basil.	Noroozlo, Y. A., Souri, M. K. & Delshad, M. *Advances in Horticultural Science* **33** (2019).	Glycine, Glutamine	Distilled water as control treatment. Growth chamber used to ensure controlled light conditions, humidity and temperature,
Stimulation Effects of Foliar Applied Glycine and Glutamine Amino Acids on Lettuce Growth.	Noroozlo, Y. A., Souri, M. K. & Delshad, M. *Open Agric.* **4**, 164-172, doi:10.1515/opag-2019-0016 (2019).	Glycine, Glutamine	Distilled water as control treatment.
Glycine metabolism and chlorophyll synthesis in barley leaves.	Hendry, G. A. F. & Stobart, A. K. Phytochemistry 16, 1567-1570, doi:https://doi.org/10.1016/0031-9422(77)84024-7 (1977).	Glycine	Feeding was continued in the light for 22 h. Controls fed glycine alone rather than glycine and glyoxylate

Foliar Application of Mixture of Amino Acids and Seaweed (Ascophylum nodosum) Extract Improve Growth and Physicochemical Properties of Grapes.	Khan, A. S., Ahmad, B., Jaskani, M. J., Ahmad, R. & Malik, A. U. *Int. J. Agric. Biol.* **14**, 383-388 (2012).	mixtures	Control vines were unsprayed.
Glycine betaine involvement in freezing tolerance and water stress in Arabidopsis thaliana.	Xing, W. & Rajashekar, C. *Environ. Exp. Bot.* **46**, 21-28 (2001).	Glycine Betaine	Control plants were not cold treated. freezing conditions controlled by programable fridge to decrease by 2 C per hour
Does Foliar-applied Glycine Betaine Affect Endogenous Betaine Levels and Yield In Cotton?	Meek, C., Oosterhuis, D. & Gorham, J. *Crop management* **2**, 1-10 (2003).	Glycine Betaine	No controls given
Exogenous glycine betaine modulates ascorbate peroxidase and catalase activities and prevent lipid peroxidation in mild water-stressed Carapa guianensis plants.	Cruz, F., Castro, G., Júnior, D. S., Festucci-Buselli, R. & Pinheiro, H. *Photosynthetica* **51**, 102-108 (2013).	Glycine Betaine	Control plants mentioned but not explained
NaCl effects on proline metabolism in rice (Oryza sativa) seedlings.	Lutts, S., Majerus, V. & Kinet, J. M. *Physiol. Plant.* **105**, 450-458 (1999).	Proline	controls were performed by omitting individual substrates

Effects of proline on antioxidant system in leaves of grapevine (Vitis vinifera L.) exposed to oxidative stress by H2O2.	Ozden, M., Demirel, U. & Kahraman, A. Sci. Hortic. 119, 163-168, doi:https://doi.org/10.1016/j.scienta.2008.07.031 (2009).	Proline	T_1 = incubation in sterile dH_2O for 24 h (control); 24 C 60% humidity 16 h photoperiod
Proline as a measure of stress in tomato plants.	Claussen, W. Plant Science 168, 241-248, doi:https://doi.org/10.1016/j.plantsci.2004.07.039 (2005).	Proline	Large changes in temperature and light: Day/night air temperatures were 20-22 °C/18-20 °C, with midday temperatures occasionally rising to 23-32 °C in summer. Relative humidity was between 45 and 85%, and midday photosynthetic photon flux density (PAR) ranged from 350 to 1200 µmol m−2 s−1 (April-October).
Exogenously applied proline at different growth stages enhances growth of two maize cultivars grown under water deficit conditions.	Ali, Q., Ashraf, M. & Athar, H.-U.-R. *Pak. J. Bot.* **39**, 1133-1144 (2007).	Proline	Control samples were grown without water stress conditions
Exogenous proline application enhances the efficiency of nitrogen fixation and assimilation in chickpea plants exposed to cadmium.	Alyemeni, M. N., Hayat, Q., Hayat, S., Faizan, M. & Faraz, A. *Legume Research: An International Journal* **39** (2016).	Proline	Control plants received double distilled water. Five sets of earthen pots (10 inch diameter) filled with sandy loam soil and farmyard manure (6:1) arranged under a simple randomized block design. At the start of the experiment, Out of these five sets of prepared pots, four sets were supplemented with different doses (0, 25, 50 or 100 mg per kg of soil) of Cd, respectively and one set of pots was left untreated serving as control. No temperature controls given

Seed pretreatment and foliar application of proline regulate morphological, physio-biochemical processes and activity of antioxidant enzymes in plants of two cultivars of quinoa (Chenopodium quinoa Willd.)	Yaqoob, H. *et al.*. *Plants* **8**, 588 (2019).	Proline	During the course of the experiment, the average temperature of 28 °C (day) and 20 °C (night), relative humidity of 61% and day-length of 7.5 h were recorded. Three concentrations (0, 25 and 50 mM) of proline were prepared in distilled water containing 0.1% Tween-20 as a surfactant
Exogenous Proline Effects on Photosynthetic Performance and Antioxidant Defence System of Young Olive Tree.	Ben Ahmed, C., Ben Rouina, B., Sensoy, S., Boukhriss, M. & Ben Abdullah, F. *J. Agric. Food. Chem.* **58**, 4216-4222, doi:10.1021/jf9041479 (2010).	Proline	Control plants (nonstressed plants irrigated with fresh water) plants received nutrient solution instead of saline. humidity not given precise temp and light controls not given (ambient environmental conditions with natural sunlight and temperature)
L-proline alleviates heat stress of tomato plants grown under protected environment.	Tonhati, R., Mello, S. C., Momesso, P. & Pedroso, R. M. Sci. Hortic. 268, 109370, doi:https://doi.org/10.1016/j.scienta.2020.109370 (2020).	Proline	Control plants were sprayed with deionized water only. temperature, relative humidity and photosynthetic active radiation, PAR inside the greenhouse) were recorded every 30 min using a Watchdog 2475 plant growth station (Spectrum Technologies, 3600 Thayer Court, Aurora, IL 60504), placed at a height of 1.5 m.

Effects of L-proline foliar application on the quality of 'Kosui' Japanese pear.	Takeuchi, M., Arakawa, C., Kuwahara, Y. & Gemma, H. *Acta Hortic.* **800**, 549-554 (2007).	Proline	(1) 30 plants having high levels of chlorophyll content in their leaves were fed 3 g N as a supplementary top dressing and were treated with L-proline at 200 mg L-1 concentration with 0.1% surfactant during the fruit growing period (hereinafter, N+P+); (2) 30 plants having high levels of chlorophyll content in the leaves were fed 3 g N the same as in (1) but without L-proline application (N+P-); (3) 30 plants with low chlorophyll content in the leaves were treated by L-proline at 200 mg L-1 concentration with 0.1% surfactant (N-P+); (4) 30 plants with low chlorophyll content were treated the same as in (3) but without L-proline application (N-P).
Foliar nitrogen application in Cabernet Sauvignon vines: Effects on wine flavonoid and amino acid content.	Gutiérrez-Gamboa, G., Garde-Cerdán, T., Portu, J., Moreno-Simunovic, Y. & Martínez-Gil, A. M.*Food Res. Int.* **96**, 46-53 (2017).	Arginine	The vineyard was equipped with a drip irrigation system using 4 L/h drippers, to assure plant water needs. Vines were irrigated when the midday leaf water potential (ψl) reached 1.0 to 1.2 MPa. The vineyard plot was homogeneous on its vegetative expression and fruit load. The site's annual average temperature is 14.5 °C with a minimum of −2.5 °C (July) and a maximum of 36.7 °C (January), and an average annual rainfall of 583.8 mm. The vineyard soil is clay loam classified as Cunculén series Vertic Haploxeralfs. 18 untreated plants left in the same row and two rows between replicates to avoid contamination.

Foliar application of proline, phenylalanine, and urea to Tempranillo vines: Effect on grape volatile composition and comparison with the use of commercial nitrogen fertilizers.	Garde-Cerdán, T. *et al*. LWT-Food Science and Technology **60**, 684-689 (2015).	Proline, Phenylalanine	Control plants were sprayed with water solution of Tween 80 alone (without the amino acid treatments). Grapes were harvested at their optimum maturity, i.e. when the weight of 100 berries remained constant and the probable alcohol reached 13 mL of ethanol/L for control sample. Extraction temp controlled. treatments applied at same interval. Growth conditions not provided
Study of the effects of proline, phenylalanine, and urea foliar application to Tempranillo vineyards on grape amino acid content. Comparison with commercial nitrogen fertilisers.	Garde-Cerdán, T. *et al.Food Chem*. **163**, 136-141 (2014).	Proline, Phenylalanine	Control plants were sprayed with a water solution of Tween 80 alone
Changes on grape phenolic composition induced by grapevine foliar applications of phenylalanine and urea.	Portu, J., López-Alfaro, I., Gómez-Alonso, S., López, R. & Garde-Cerdán, T. *Food Chem*. **180**, 171-180 (2015).	Phenylalanine	Control plants were sprayed with water solutions of Tween 80 alone. treatments applied at same interval. Growth conditions not provided

Foliar application of phenylalanine plus methyl jasmonate as a tool to improve Grenache grape aromatic composition.	Marín-San Román, S., Garde-Cerdán, T., Baroja, E., Rubio-Bretón, P. & Pérez-Álvarez, E. P. *Sci. Hortic.* **272**, 109515 (2020).	Phenylalanine	Control plants were sprayed with aqueous solution of Tween 80 alone. All treatments were performed twice, at veraison and one week later. The meteorogical data were obtained from the Agroclimatic Information Service of La Rioja (SIAR), selecting the station located about 5 km from the place where the vineyard was located. The collected data were the rain accumulated from the beginning of April until 1st of September, being 114.1 mm; the duration of leaf wetness, 848 h; and the average maximum, mean and minimum temperatures, being 27.5 °C, 19.9 °C and 12.7 °C, respectively. The plots were managed according to the viticultural practices of the region.
Foliar phenylalanine application promoted antioxidant activities in Cabernet Sauvignon by regulating phenolic biosynthesis.	Cheng, X. *et al.* *J. Agric. Food. Chem.* **68**, 15390-15402 (2020).	Phenylalanine	The control check treatment with an aqueous solution (CK). From June to harvest date, the temperature was adjusted using electric fans to keep the day temperature in the range of 20-35 °C and the night temperature above 15 °C, and the relative humidity in greenhouse was in 40-60%, mainly according to the weather and time of day.
Effect of tryptophan and ascorbic acid on yield and some chemical constituents of lupine (Lupines termis L.) plants.	Khalifa, Y. A., El-Naem, A., Gamal, F. & Mahmoud, M. A. *Egypt. J. Agron.* **42**, 47-61 (2020).	Trp, ascorbic acid	Control left untreated. Temp/humidity/light controls not given

Impact of Foliar Application of Amino Acids on Total Phenols, Phenolic Acids Content of Different Mints Varieties under the Field Condition.	Tarasevičienė, Ž., Velička, A. & Paulauskienė, A. *Plants* **10**, 599 (2021).	phenylalanine, Trp, tyrosine	Foliar spray with water was used as a positive control. The effects of all treatments were compared with a negative control (without spraying). Soil composition given including pH of soil. graph of precipitation and temperature variation included.
Application of amino acids improves lettuce crop uniformity and inhibits nitrate accumulation induced by the supplemental inorganic nitrogen fertilization.	Tsouvaltzis, P. Koukounaras, A. & Siomos, A. S. *International Journal of Agriculture and Biology* **16** (2014).	mixtures	Plants without any supplemental fertilization were used as control.
Effect of different amino acids at different con-centrations on multiplication and rooting stage of in vitro propagation of strawberries (Fragaria X Ananassa Duch cv. Chandler).	El-sharabasy, S., Issa, F., Hammad, G. & El-Dawayaty, M. *Egypt J Genet Cytol* **44** (2015).	tyrosine, arginine, glutamine	After inoculation, the culture jars were maintained at a temperature of 25±2C with a 16 hours/day photoperiod. Lighting was supplied using fluorescent lamps with 1000 lux for the multiplication stages and 3000 lux for the rooting stages. Controls treated with MS nutrient medium supplemented with (1 $mg.L^{-1}$) Naphthalene acetic acid (NAA) and (2.5$g.L^{-1}$) activated charcoal powder with 0 mg/L amino acid conc.
Alleviation of sea water stress on tomato plants by foliar application of aspartic acid and glutathione.	Akladious, S. A. & Abbas, S. M. *Bangladesh J. Bot.* **42**, 31-44 (2013).	aspartic acid, glutathione	The 1st group of pots was irrigated with tap water to serve as control. No data on growing conditions controls in place.
Foliar application of aspartic acid lowers cadmium uptake and Cd-induced oxidative stress in rice under Cd stress.	Rizwan, M. *et al. Environ Sci Pollut R* **24**, 21938-21947 (2017).	aspartic acid	Control plants not treated with amino acid. No data on growing conditions controls in place.

The significance of amino acids and amino acid-derived molecules in plant responses and adaptation to heavy metal stress.	Sharma, S. S. & Dietz, K.-J. *J. Exp. Bot.* **57**, 711-726 (2006).	proline, histidine	Untreated control (in the absence of excess heavy metal)
Thompson seedless grapevines growth and quality as affected by glutamic acid, vitamin B, and algae.	Abou-Zaid, E. A. A. & Eissa, M. A. *J. Soil Sci. Plant Nutr.* **19**, 725-733 (2019).	glutamic acid	(C) Control (spraying with water). plant growing conditions not given
Effects of Foliar-Applied L-Glutamic Acid on the Diurnal Variations of Leaf Gas Exchange and Chlorophyll Fluorescence Parameters in Hawthorn (Crataegus pinnatifida Bge.).	Lv, D. *et al.* Eur. J. Hortic. Sci. **74**, 204 (2009).	glutamic acid	Distilled water as a control. Variation in temp given in graph
Effects of extraneous glutamic acid on nitrate contents and quality of Chinese chive	Cao, Y., Gao, Z., Li, J., Xu, G. & Wang, M. in International Symposium on Vegetable Safety and Human Health 856. 91-98.	glutamic acid	water control

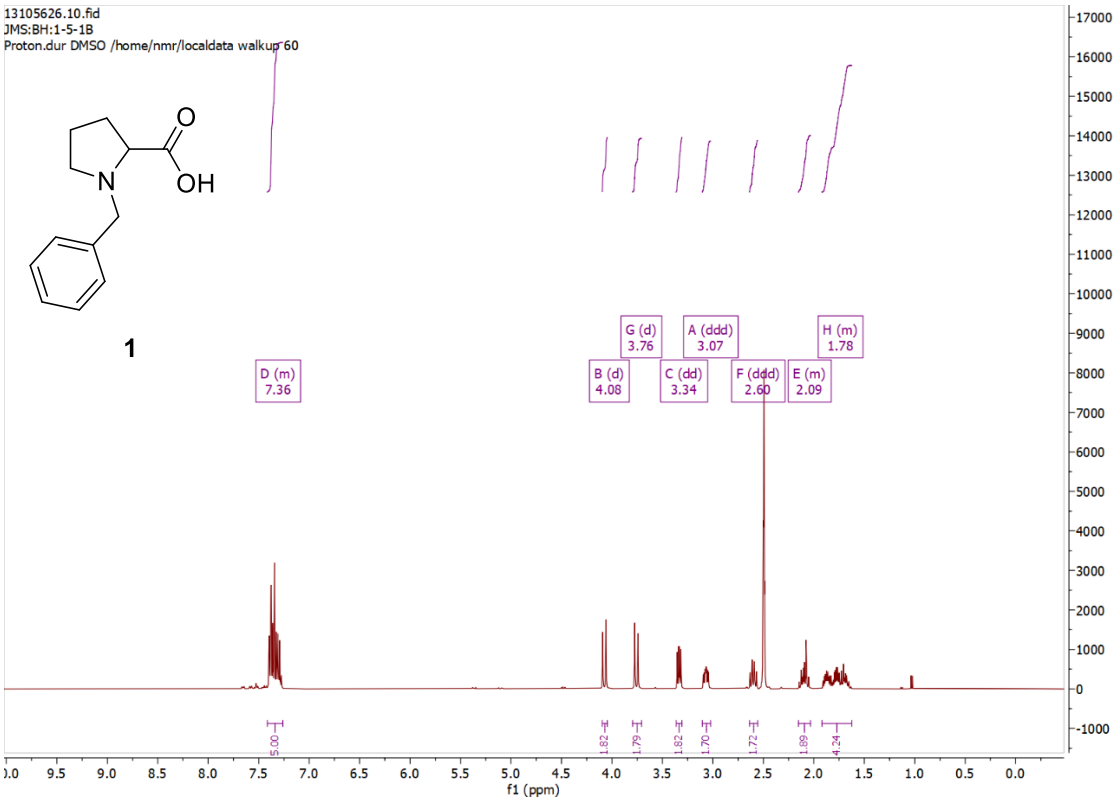
Exploratory Study on the Foliar Incorporation and Stability of Isotopically Labeled Amino Acids Applied to Turfgrass.	McCoy, R. M. *et al*. *Agronomy* **10**, 358 (2020).	glutamate	23 - 24 °C with an average humidity of 45% under daylight spectrum fluorescent lighting with 12-h days. lants were watered weekly and fertilized once at germination with a fertilizer containing 12% nitrogen, 6% phosphorus (P2O5), 6% potassium (K2O), and micronutrients boron, copper, iron, manganese, and zinc. No control plants mentioned
Effects of different foliar nitrogen fertilizers on cellular nitrogen metabolism and biomass of two shrub willow cultivars.	Gagne, M. A., Minocha, R., Long, S. & Minocha, S. C. *Can. J. For. Res.* **49**, 1548-1559 (2019).	arginine	No fertilizer spray. ambient light and temperature conditions no precise values given
Effect of foliar spraying with arginine and cysteine and the number of stems on the growth and yield of cherry tomatoes grown in protected culture.	Hoza, G. *et al*. Scientific Papers-Series B, Horticulture **63**, 417-423 (2019).	arginine, cysteine	Control was also conducted on 1 and 2 stems, without application of amino acids. No growing conditions given
Iron (II)-amino acid chelates alleviate salt-stress induced oxidative damages on tomato grown in nutrient solution culture.	Ghasemi, S., Khoshgoftarmanesh, A. H., Afyuni, M. & Hadadzadeh, H. *Sci. Hortic.* **165**, 91-98 (2014).	arginine, histidine, glycine	8h light period at intensity of 390 mol m−2 s−1, 25/20 ◦C day/night temperature, and 65-75% relative humidity. Nutrient solution composition given. Not compared to untreated control plants

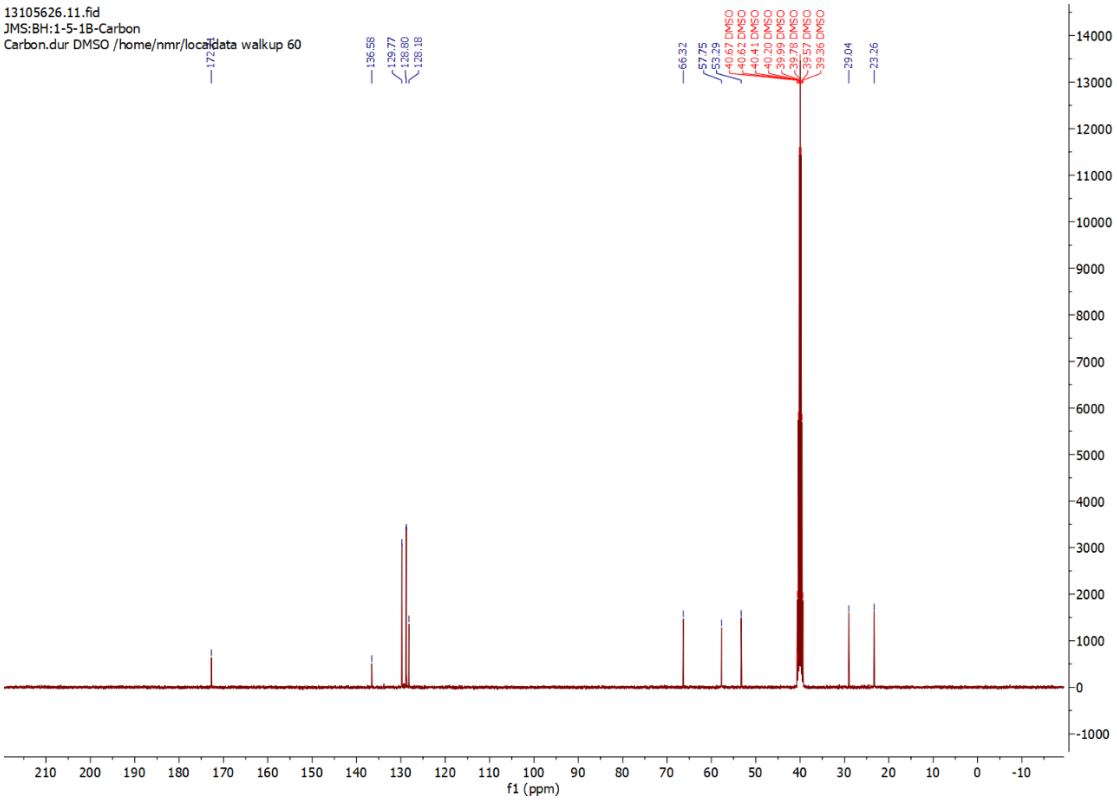
The effectiveness of foliar applications of synthesized zinc-amino acid chelates in comparison with zinc sulfate to increase yield and grain nutritional quality of wheat.	Ghasemi, S., Khoshgoftarmanesh, A. H., Afyuni, M. & Hadadzadeh, H. *Eur. J. Agron.* **45**, 68-74, doi:10.1016/j.eja.2012.10.012 (2013).	arginine, histidine, glycine	temperature and precipitation values for growth period given. Control plant not given Zn fertilizer.
Influence of foliar-applied zinc in the form of mineral and complexed with amino acids on yield and nutritional quality of onion under field conditions.	Rafie, M., Khoshgoftarmanesh, A., Shariatmadari, H., Darabi, A. & Dalir, N.*Sci. Hortic.* **216**, 160-168 (2017).	lysine, methionine, threonine	A control treatment free of Zn and amino acids was used. The mean monthly maximum and minimum temperatures at the site were 43.16 ◦C (on June) and 4.33 ◦C (on December), respectively. The total annual precipitation during the growing season was 356.6 mm, all of which falling between November and May. Soil cores (0-30 cm depth) were taken from the experimental location. The samples were air-dried, ground, sieved to less than 2-mm and used for chemical analysis.
Role of foliar application of sulfur-containing compounds on maize (Zea mays L. var. Malka and hybrid DTC) under salt stress.	Perveen, S., Iqbal, N., Saeed, M., Zafar, S. & Arshad, Z. *Brazilian Journal of Botany* **41**, 805-815 (2018).	cysteine	control plants not sprayed. Control conditions used alongside salt stress conditions. No growing conditions given
Foliar and Seed Application of Amino Acids Affects the Antioxidant Metabolism of the Soybean Crop.	Teixeira, W. F. *et al. Front. Plant Sci.* **8**, doi:10.3389/fpls.2017.00327 (2017).	glutamate, phenylalanine, cysteine, glycine	control plants not treated with amino acid. No data on growing conditions controls in place (temp/humidity/soil pH ect.)

Amino Acids as Stress Reducers in Soybean Plant Growth Under Different Water-Deficit Conditions.	Teixeira, W. F. *et al*. *J. Plant Growth Regul.*, 1-15 (2019).	proline, glutamate	In the control treatments, distilled water was applied to the seeds or leaves. No temperature/humidity/soil pH conditions given
Effect of foliar application of amino acids on the salinity tolerance of tomato plants cultivated under hydroponic system.	Alfosea-Simón, M. *et al*. *Sci. Hortic*. **272**, 109509 (2020).	Arganine, proline, glutamine, Trp, methionine and arganine, methionine and Trp, glutamine and proline	The photoperiod was set for 16 hours of light with an intensity of 500 µmol $m^{-2}s^{-1}$, the temperature was set at 25 °C day/night, with a relative humidity of 60%. Control was without added salt or amino acid treatment.
Foliar spray of asparagine amino acid on biochemical and morphological traits of garden cress (Lepidium sativum L.) plants under greenhouse conditions.	Aminifard, M. H. & Jorkesh, A. *Zeitschrift fur arznei-& gewurzpflanzen* **24**, 189-192 (2019).	asparagine	randomised complete design with three replications. A control group received no Asn. The greenhouse maintained a temperature of 25 °C during the day and 15 °C at night, with a CO2 concentration of 350 ppm, 40% relative humidity, and a 16-hour light/8-hour dark photoperiod.
. Regulation of growth and some key physiological processes in salt-stressed maize (Zea mays L.) plants by exogenous application of asparagine and glycerol.	Kaya, C., Aydemir, S., Sonmez, O., Ashraf, M. & Dikilitas, M *Acta Bot. Croat.* **72**, 157-168 (2013).	aparagine	Control treatment (nutrient solution alone) Five sterilized caryopses were planted in 8 kg peat contained in each plastic pot and all pots placed in a growth room at 27±2 °C with light intensity 350 mmol photons m^{-2} s^{-1} and RH ranging from 60 to 70%.
Physiological response of onion plants to foliar application of putrescine and glutamine.	Amin, A., Gharib, F. A., El-Awadi, M. & Rashad, E.-S. M. *Sci. Hortic*. **129**, 353-360 (2011).	glutamine	untreated plants (control), were sprayed only with tap water. Same time of year used

Appendix B: Spectra for synthetic materials

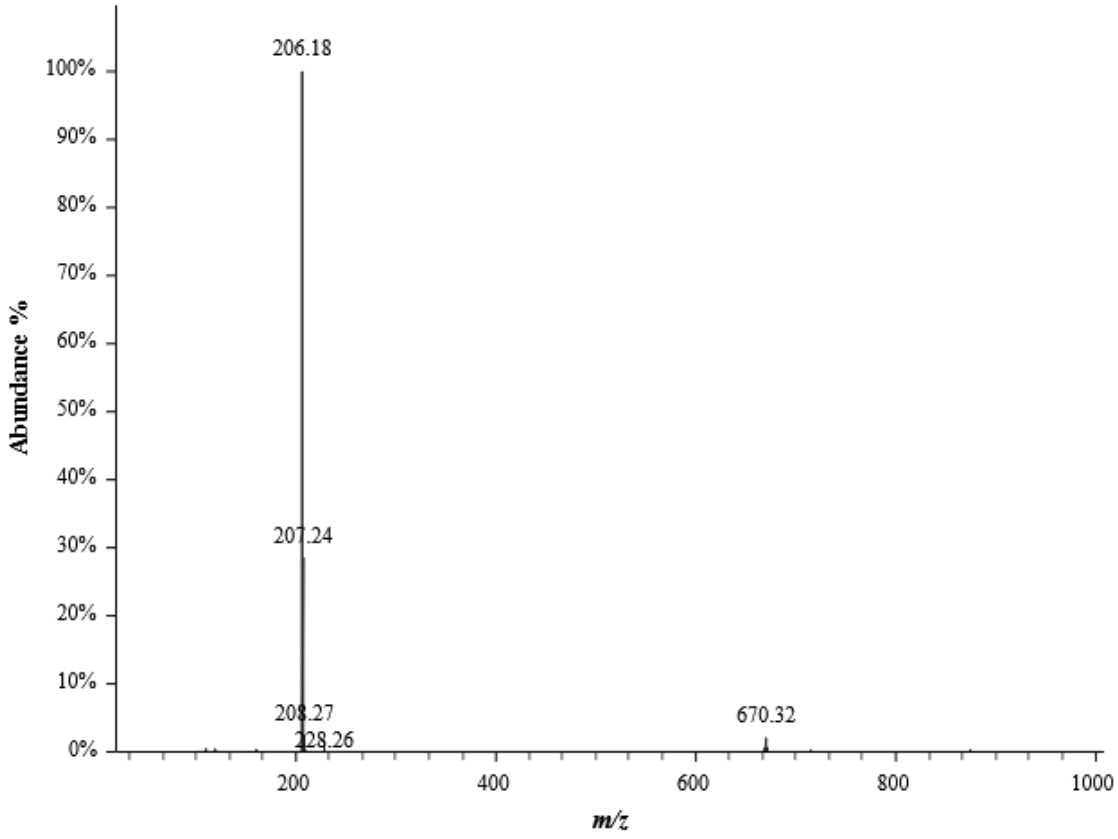
B.1. Benzyl proline (**1**)

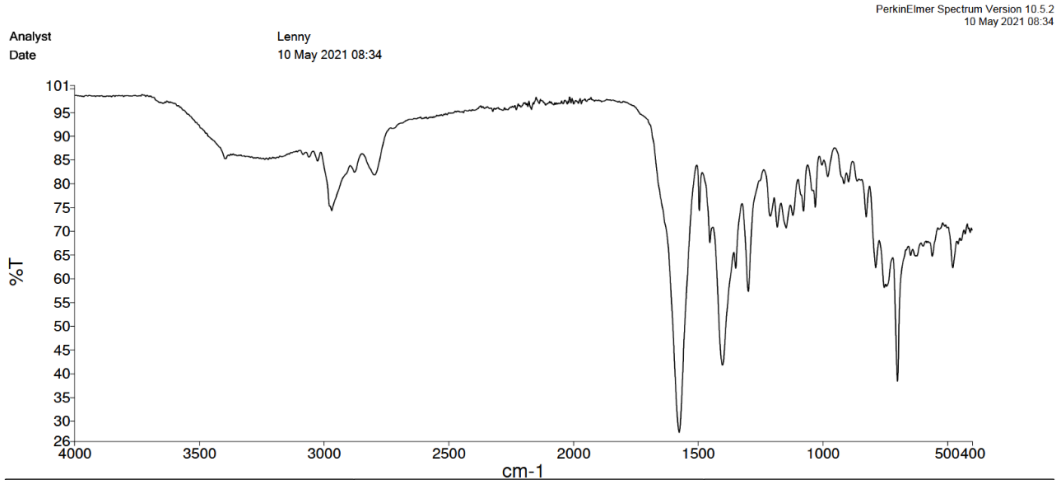




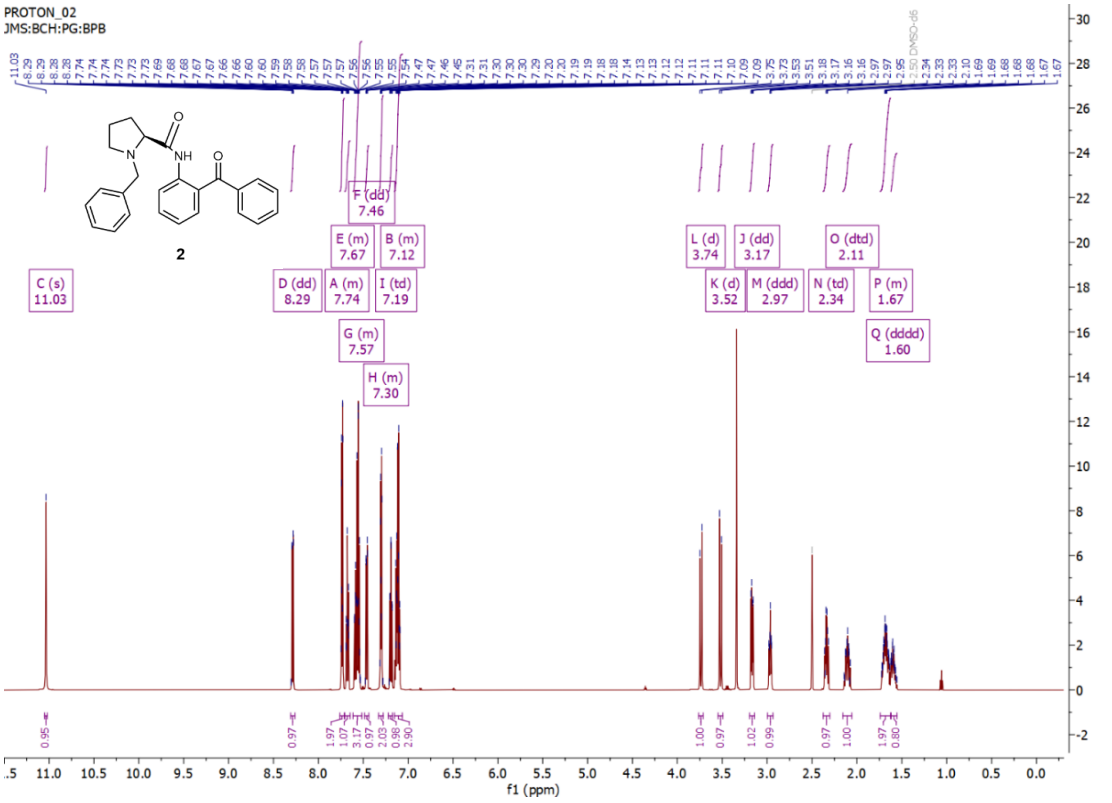
BCRH_BnPro_212300

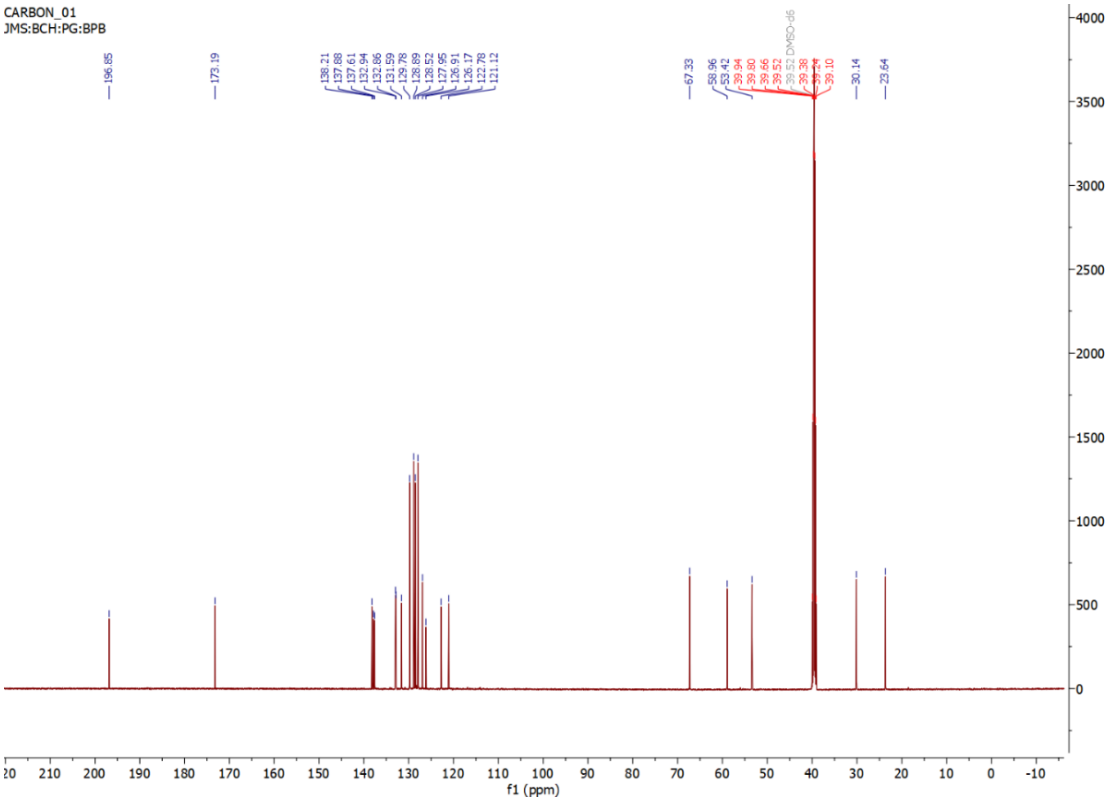
RT: 0.6658 minutes, Scan 71, 1: MS ES+ c (100.0-2000.0), NL 1.21e+8





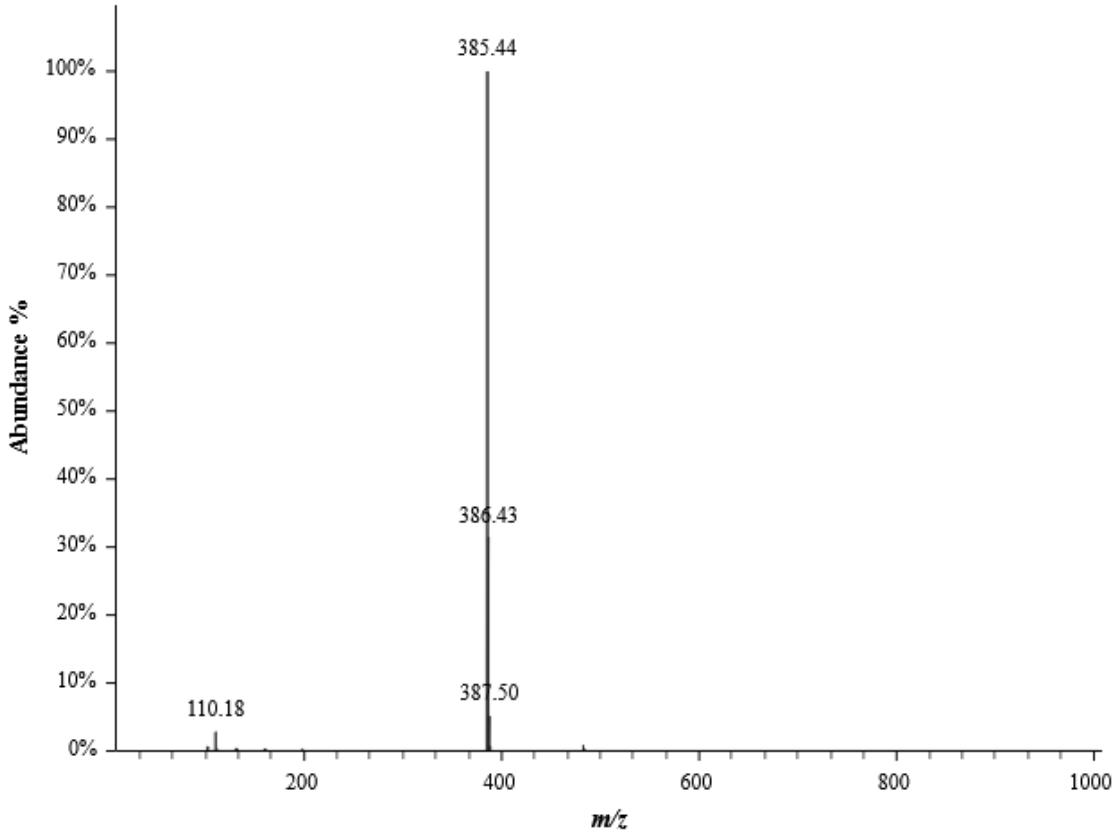
B.2. (S)-1-Benzylpyrrolidine-2-carboxylic acid (2-benzoylphenyl) amide (**2**)

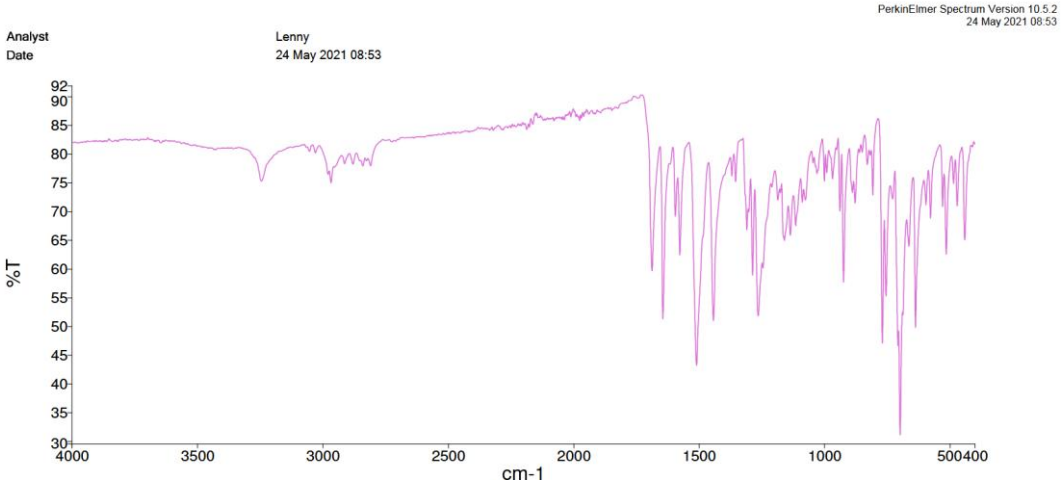




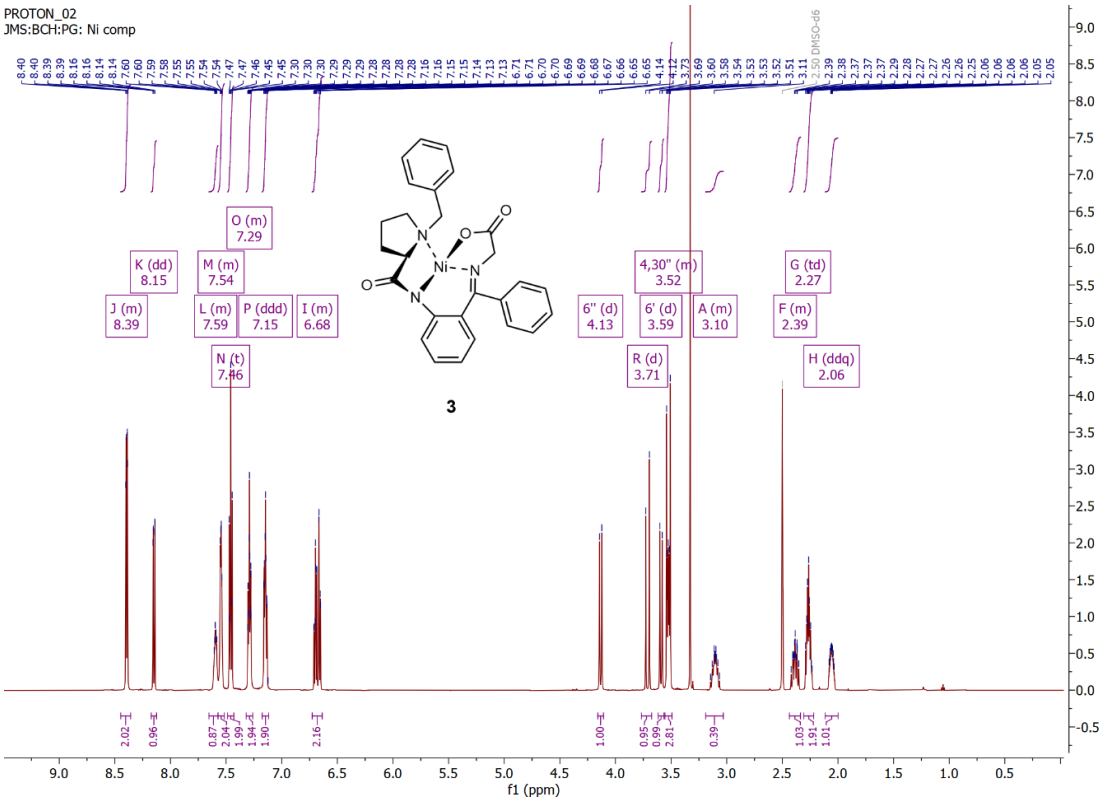
BCRH_BPB_5_197771

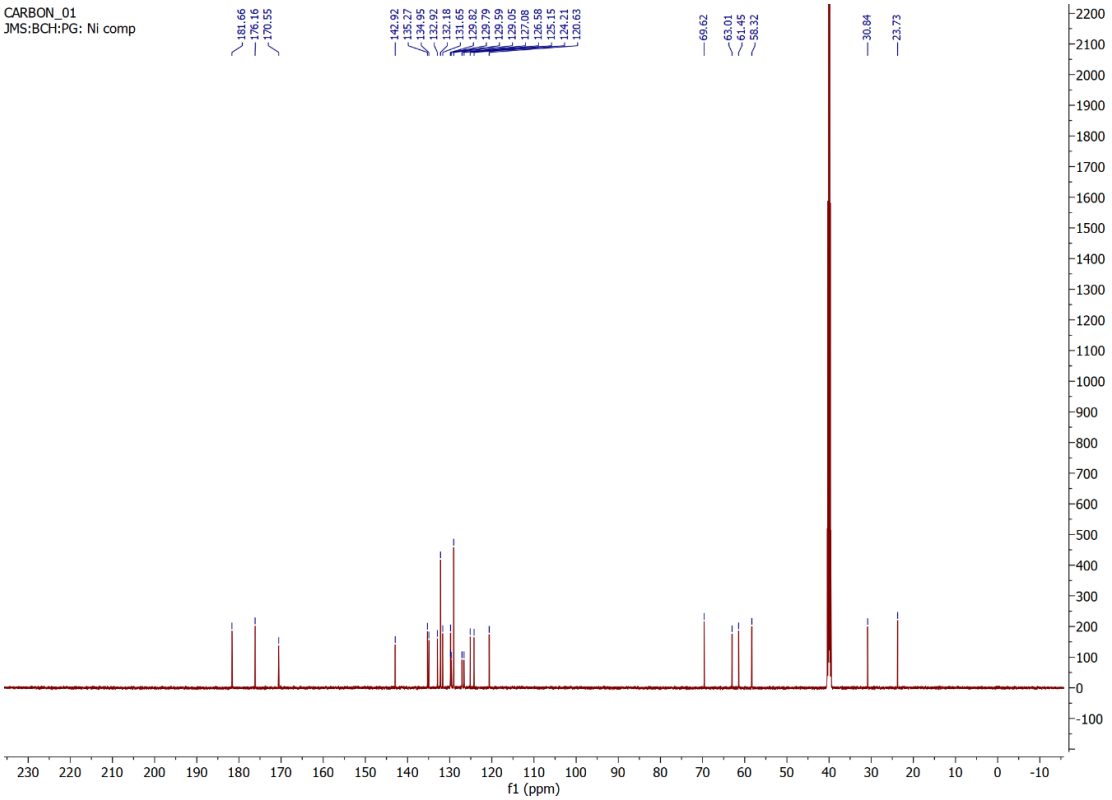
RT: 1.8287 minutes, Scan 205, 1: MS ES+ c (100.0-2000.0), NL 8.95e+7





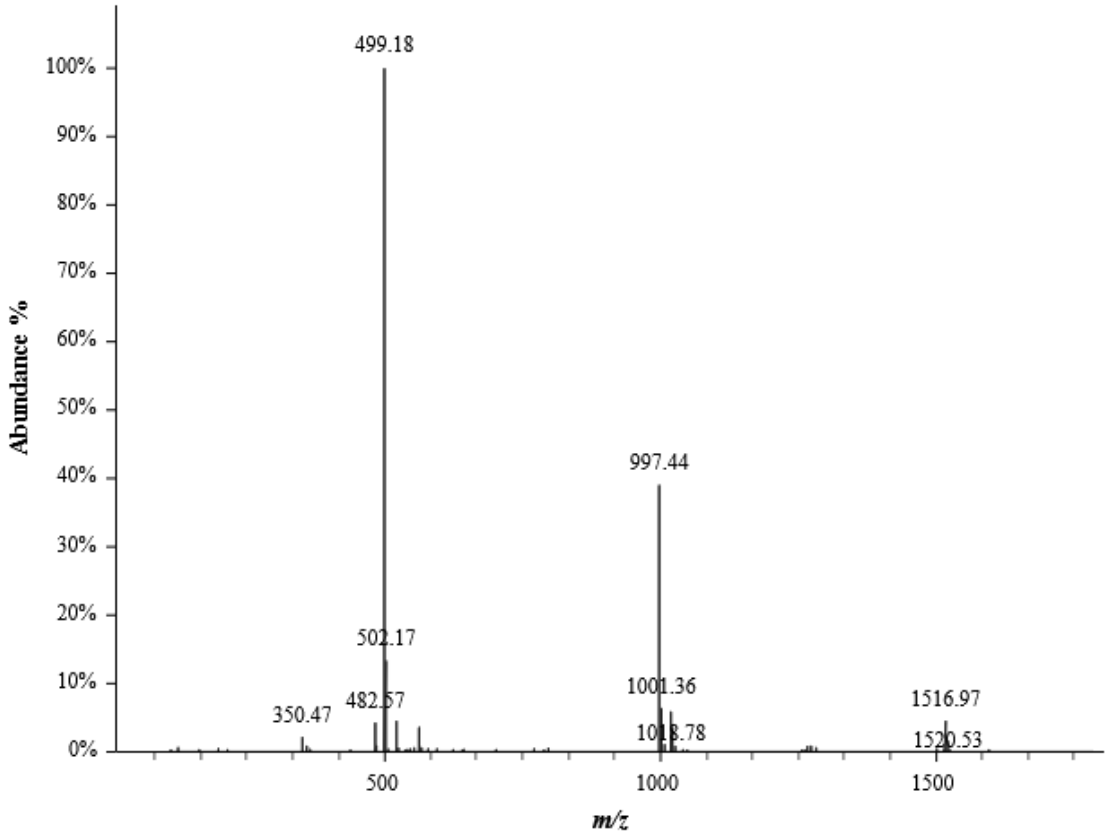
B.3. Gly-Ni-(S)-2-[N-(N-benzylpropyl)amino]benzophenone (**3**)

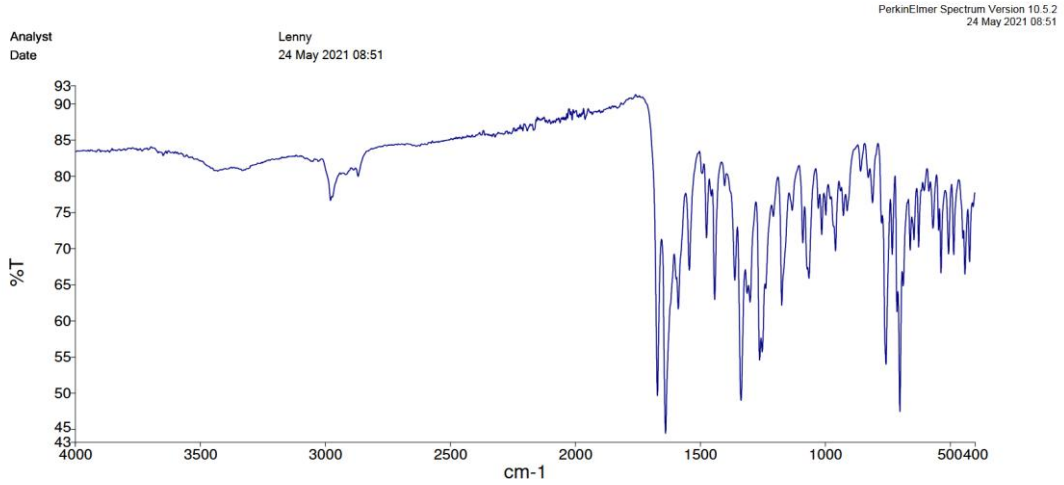




BCRH_BH015001_156870

RT: 2.7868 minutes, Scan 213, 1: MS (100-2000) ES+, NL 5.12e+7

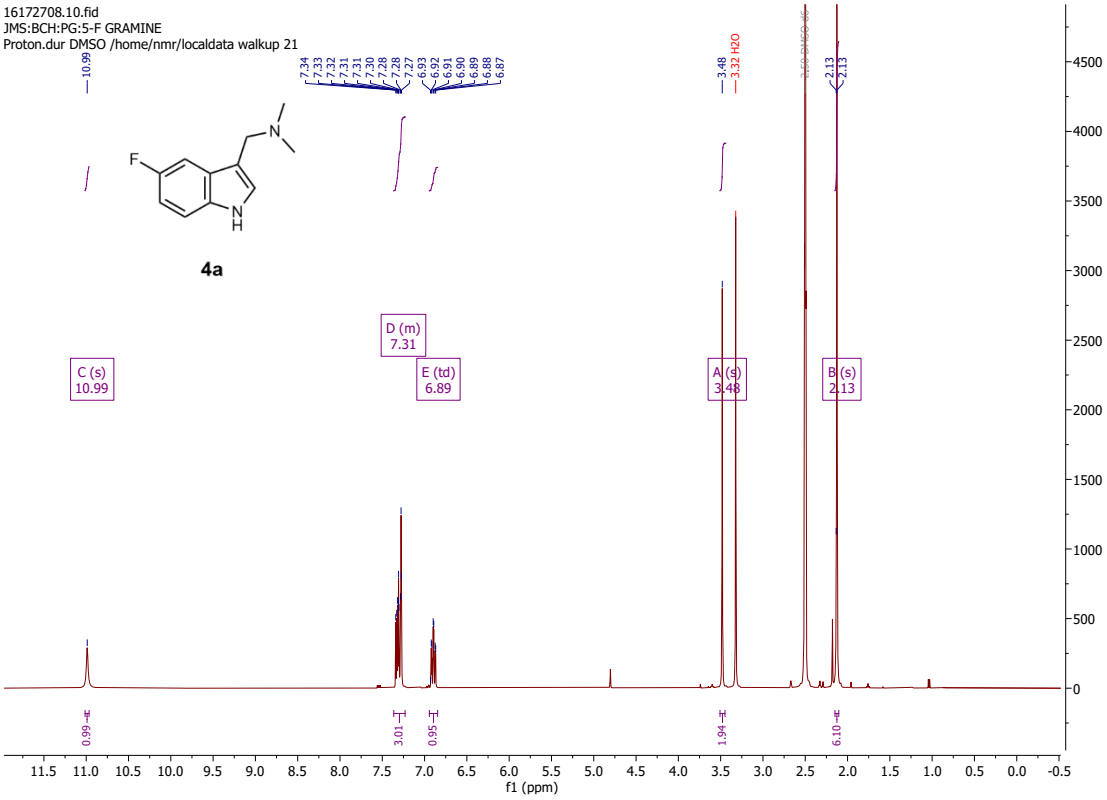


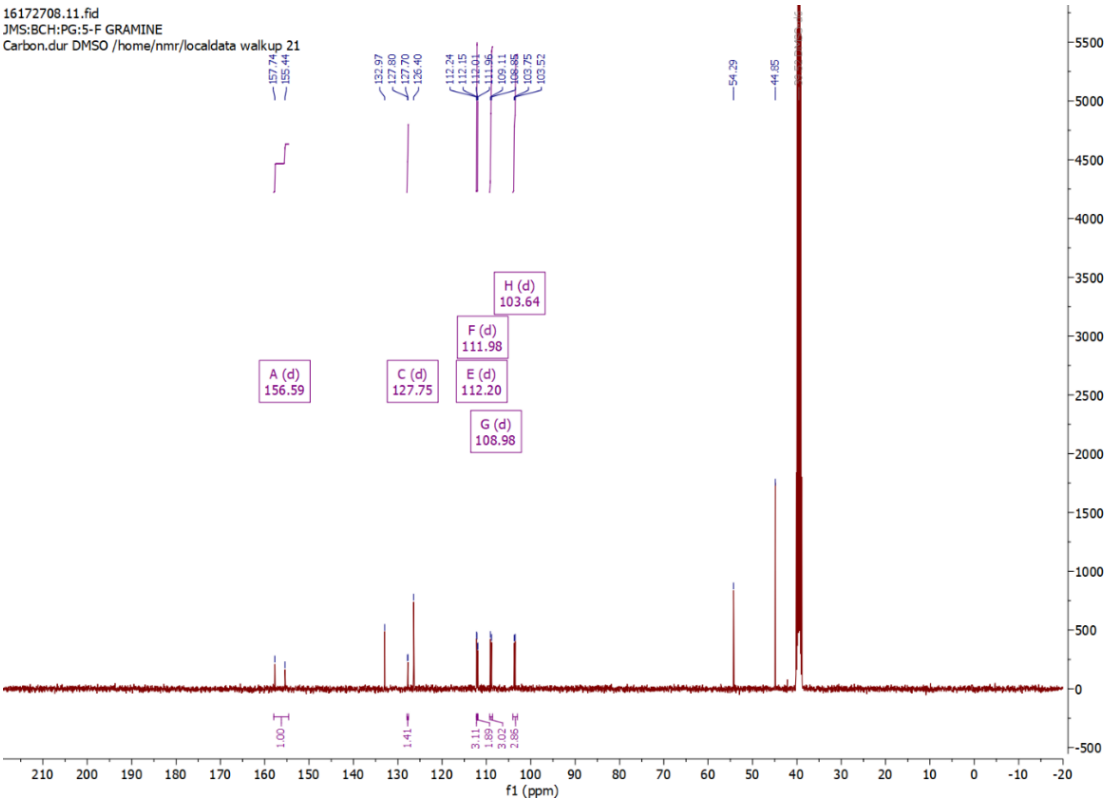


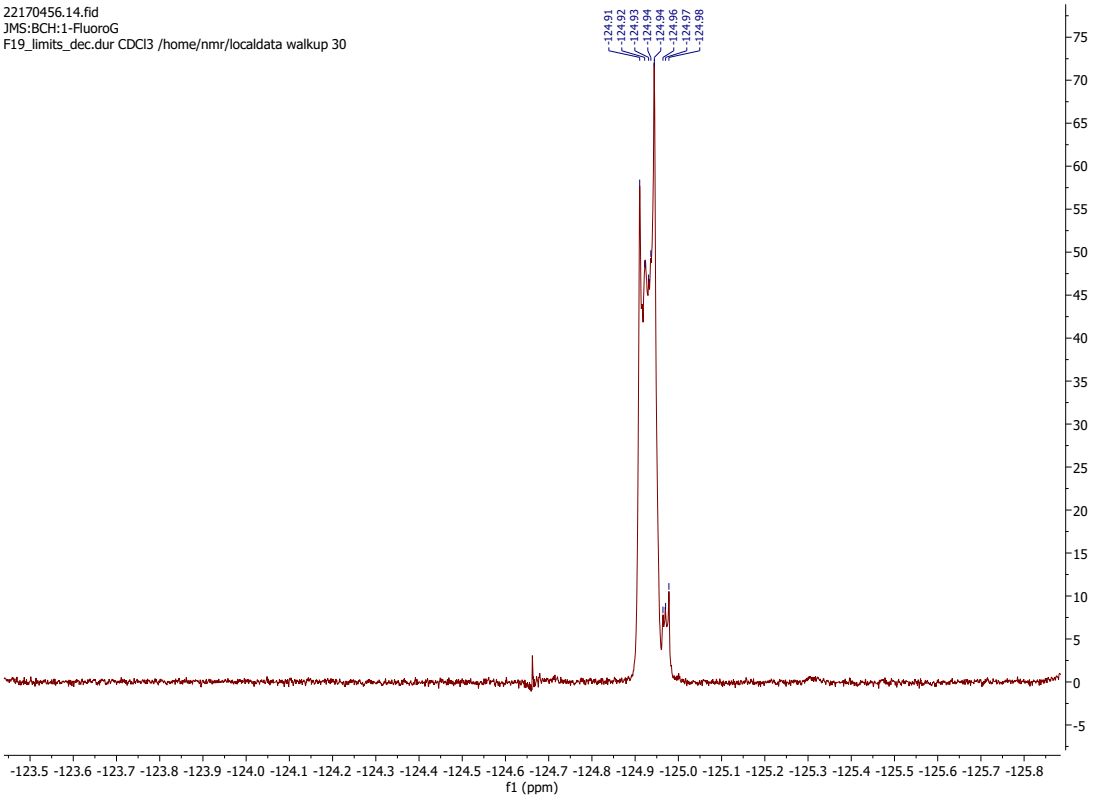
Crystal data and structure refinement for Compound 3	
Empirical formula	$C_{27}H_{25}N_3NiO_3$
Formula weight	498.21
Temperature/K	120
Crystal system	Orthorhombic
Space group	$P2_12_12_1$
a/Å	8.9825(3)
b/Å	9.6496(3)
c/Å	26.2557(8)
α/°	90
β/°	90
γ/°	90
Volume/Å^3	2275.78(12)
Z	4
ρ_{calc}g/cm^3	1.454
μ/mm^{-1}	0.888
F(000)	1040.0
Crystal size/mm^3	0.34 × 0.15 × 0.038
Radiation	Mo Kα (λ = 0.71073)
2Θ range for data collection/°	4.498 to 55.752
Index ranges	-11 ≤ h ≤ 11, -12 ≤ k ≤ 12, -34 ≤ l ≤ 34
Reflections collected	54575
Independent reflections	5429 [R_{int} = 0.0590, R_{sigma} = 0.0294]
Data/restraints/parameters	5429/0/307
Goodness-of-fit on F^2	1.078
Final R indexes [I>=2σ (I)]	R_1 = 0.0303, wR_2 = 0.0554
Final R indexes [all data]	R_1 = 0.0339, wR_2 = 0.0564
Largest diff. peak/hole / e Å^{-3}	0.27/-0.32
Flack parameter	0.008(5)

B.4. Gramine analogues

B.4.a. 5-Fluorogramine (**4a**)

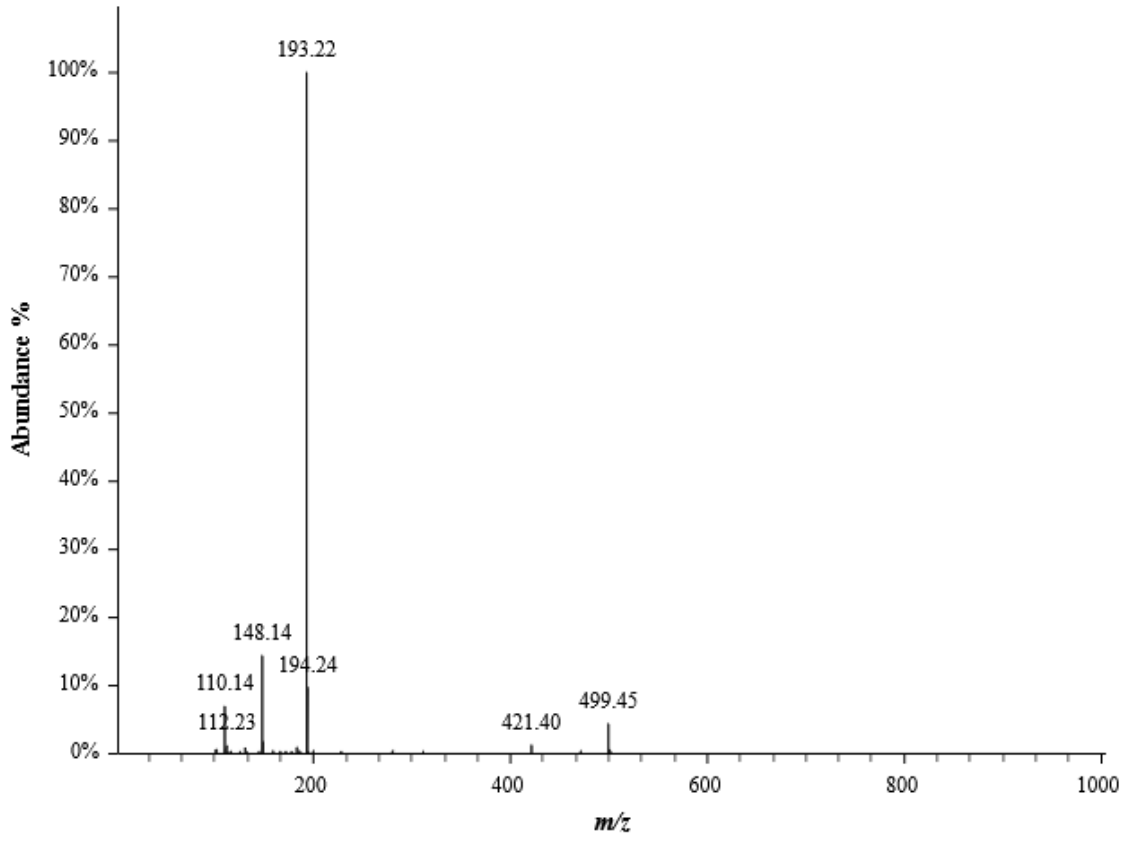


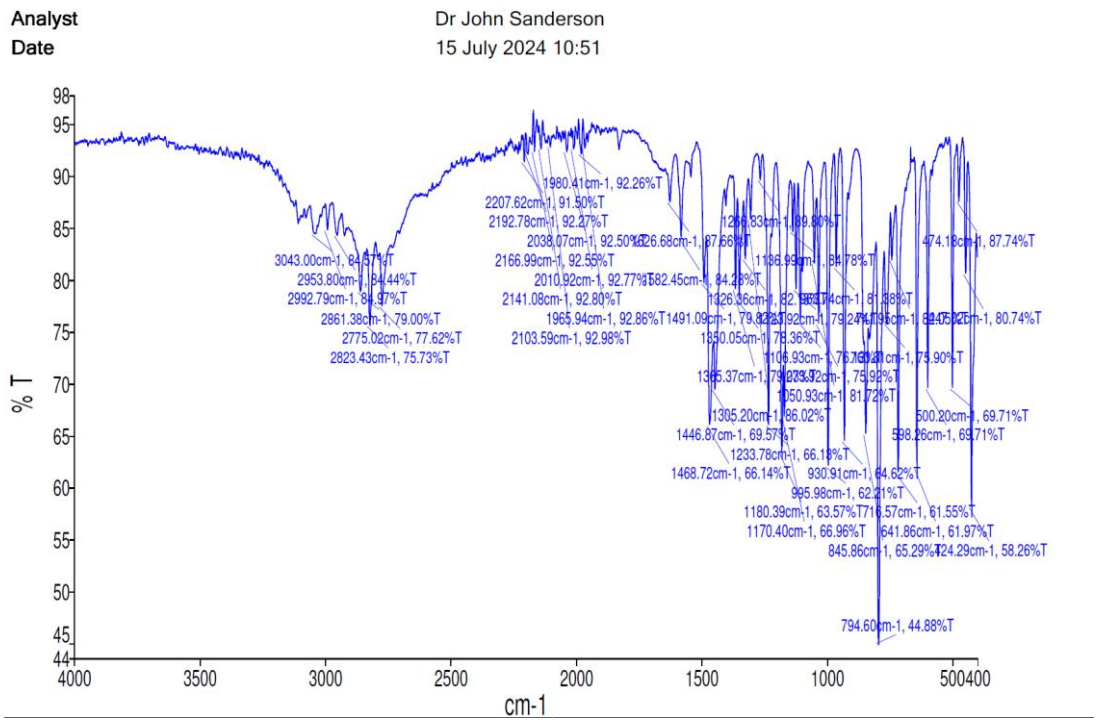




BCRH_5FGramine_206771

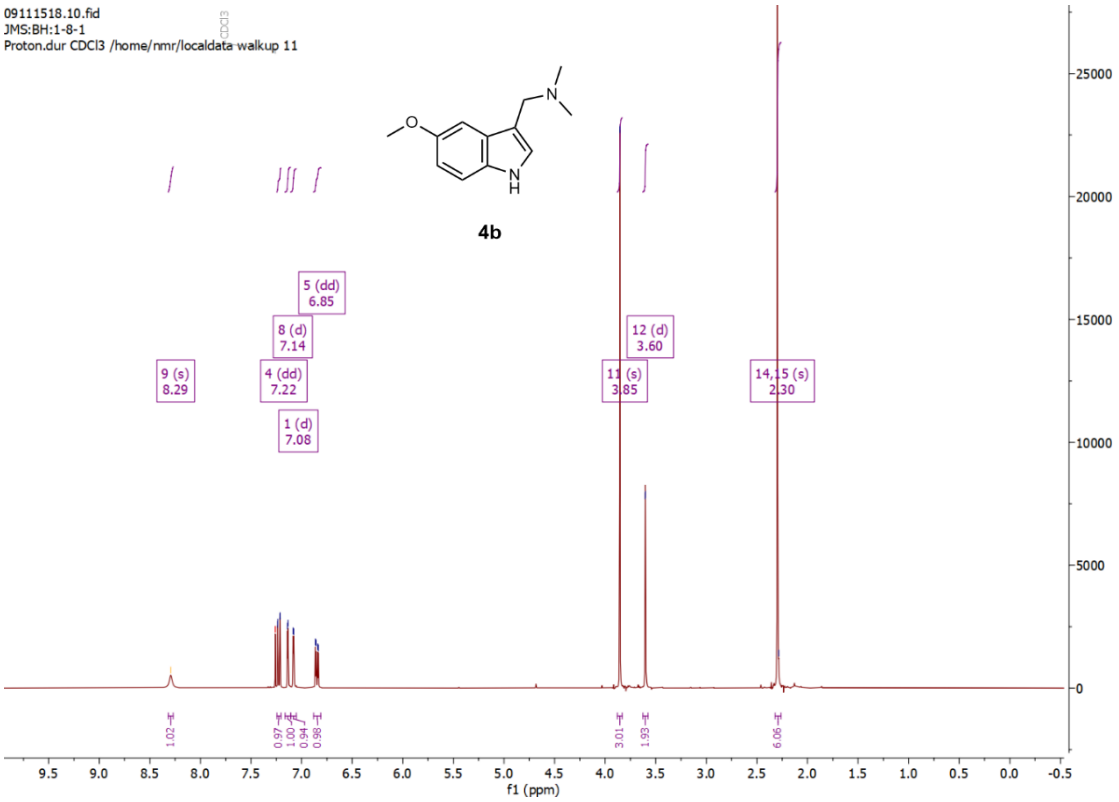
RT: 0.9262 minutes, Scan 101, 1: MS ES+ c (100.0-2000.0), NL 2.85e+7

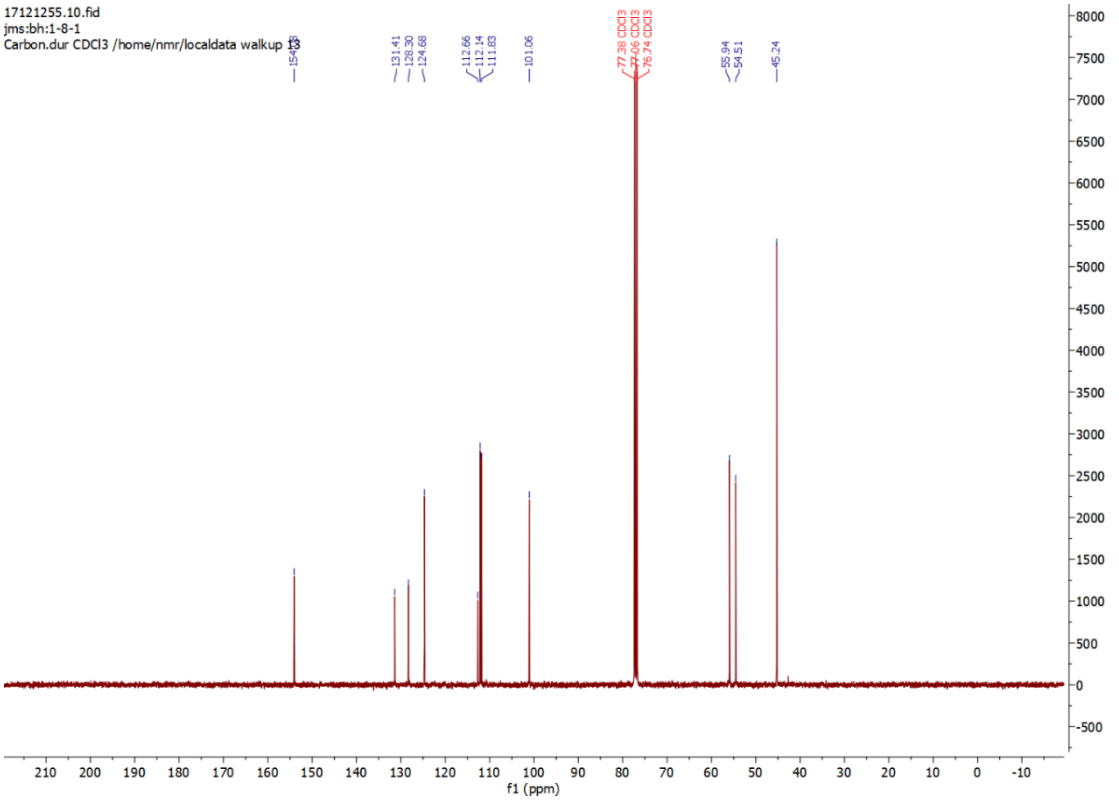



Analyst
Dr John Sanderson
Date
15 July 2024 10:51
% T
cm-1
3043.00cm-1, 84.57%T
2953.80cm-1, 84.44%T
2992.79cm-1, 84.37%T
2861.38cm-1, 79.00%T
2775.02cm-1, 77.62%T
2823.43cm-1, 75.73%T
1980.41cm-1, 92.26%T
2207.62cm-1, 91.50%T
2192.78cm-1, 92.27%T
2038.07cm-1, 92.50%T
2166.99cm-1, 92.55%T
2010.92cm-1, 92.77%T
2141.08cm-1, 92.80%T
1965.94cm-1, 92.86%T
2103.59cm-1, 92.98%T
1626.68cm-1, 87.66%T
1582.45cm-1, 84.28%T
1491.09cm-1, 79.82%T
1850.05cm-1, 78.36%T
1365.37cm-1, 79.23%T
1305.20cm-1, 86.02%T
1446.87cm-1, 69.57%T
1233.78cm-1, 66.13%T
1468.72cm-1, 66.14%T
930.91cm-1, 64.62%T
995.98cm-1, 62.21%T
1180.39cm-1, 63.57%T
1170.40cm-1, 66.96%T
845.86cm-1, 65.29%T
794.60cm-1, 44.88%T
474.18cm-1, 87.74%T
1050.93cm-1, 81.72%T
500.20cm-1, 69.71%T
598.26cm-1, 69.71%T
641.86cm-1, 61.97%T
724.29cm-1, 58.26%T

Crystal data and structure refinement for Compound 4a.	
Empirical formula	$C_{11}H_{13}FN_2$
Formula weight	192.23
Temperature/K	120.00
Crystal system	Orthorhombic
Space group	$P2_12_12_1$
a/Å	5.5669(3)
b/Å	9.5282(4)
c/Å	18.5946(8)
α/°	90
β/°	90
γ/°	90
Volume/Å^3	986.30(8)
Z	4
ρ_{calc}g/cm^3	1.295
μ/mm^{-1}	0.091
F(000)	408.0
Crystal size/mm^3	0.44 × 0.318 × 0.04
Radiation	Mo Kα (λ = 0.71073)
2Θ range for data collection/°	4.382 to 54.992
Index ranges	-7 ≤ h ≤ 7, -12 ≤ k ≤ 12, -24 ≤ l ≤ 24
Reflections collected	24225
Independent reflections	2261 [R_{int} = 0.0648, R_{sigma} = 0.0289]
Data/restraints/parameters	2261/0/133
Goodness-of-fit on F^2	1.115
Final R indexes [I>=2σ (I)]	R_1 = 0.0476, wR_2 = 0.0949
Final R indexes [all data]	R_1 = 0.0516, wR_2 = 0.0965
Largest diff. peak/hole / e Å^{-3}	0.21/-0.20

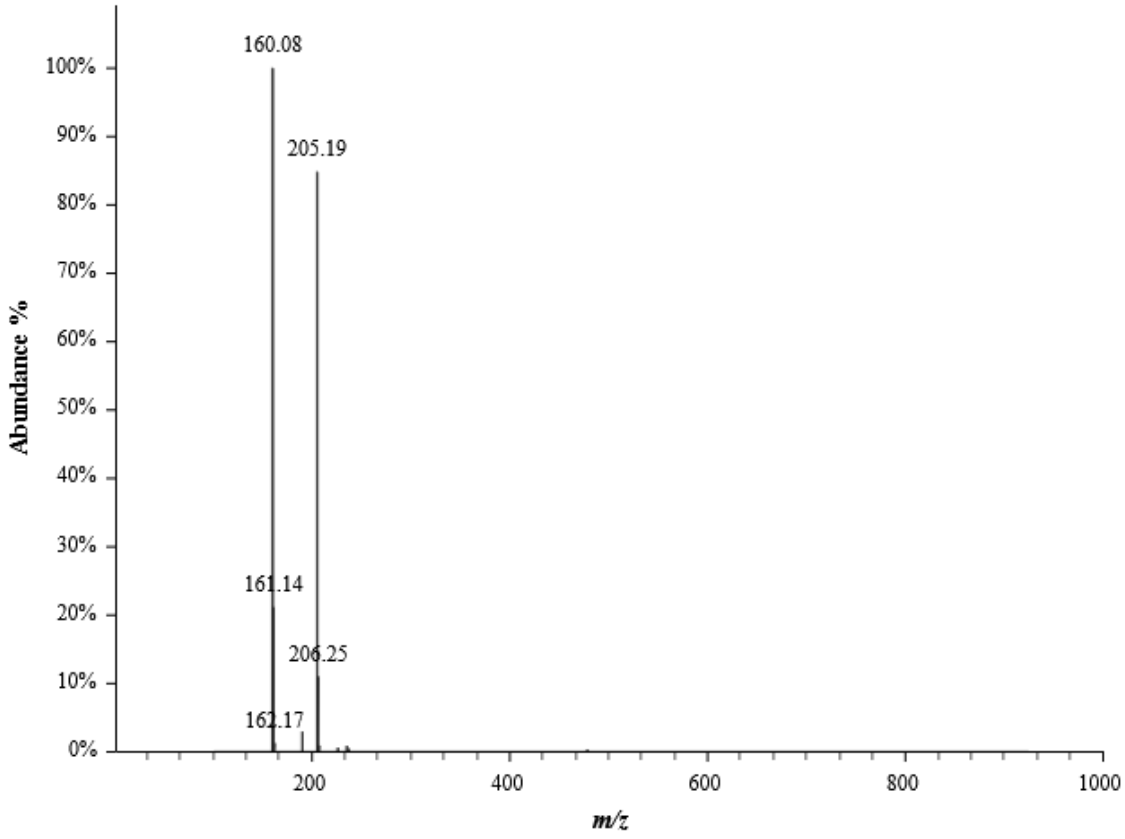
B.4.b. 5-Methoxygramine (**4b**)

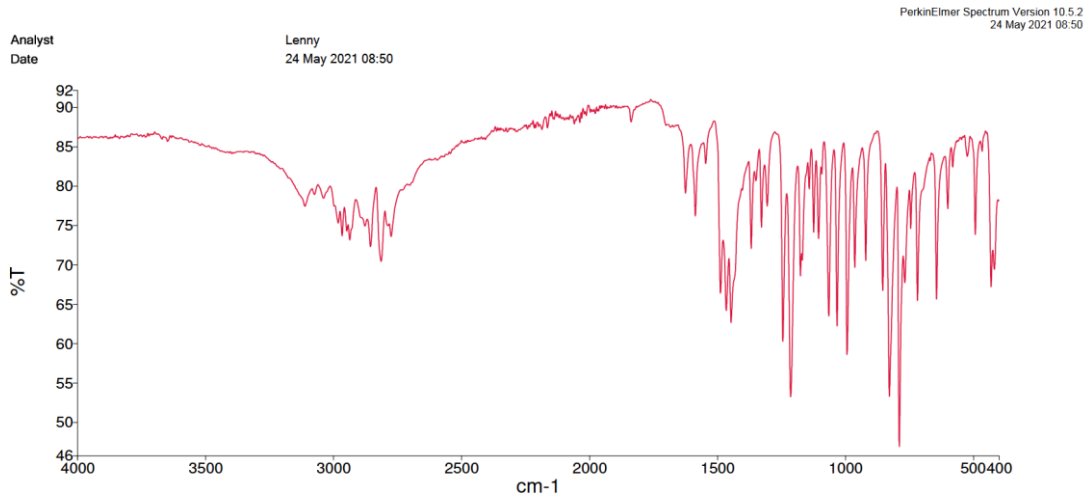




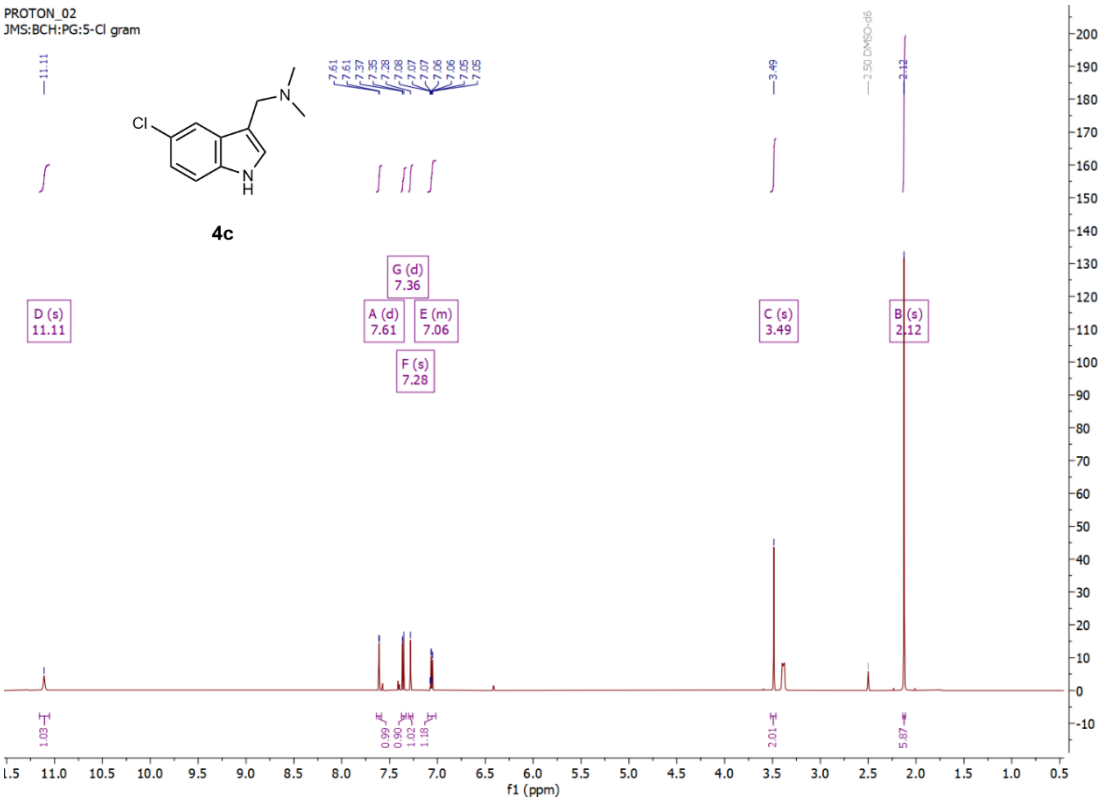
BCRH_5MeO_Gram_212079

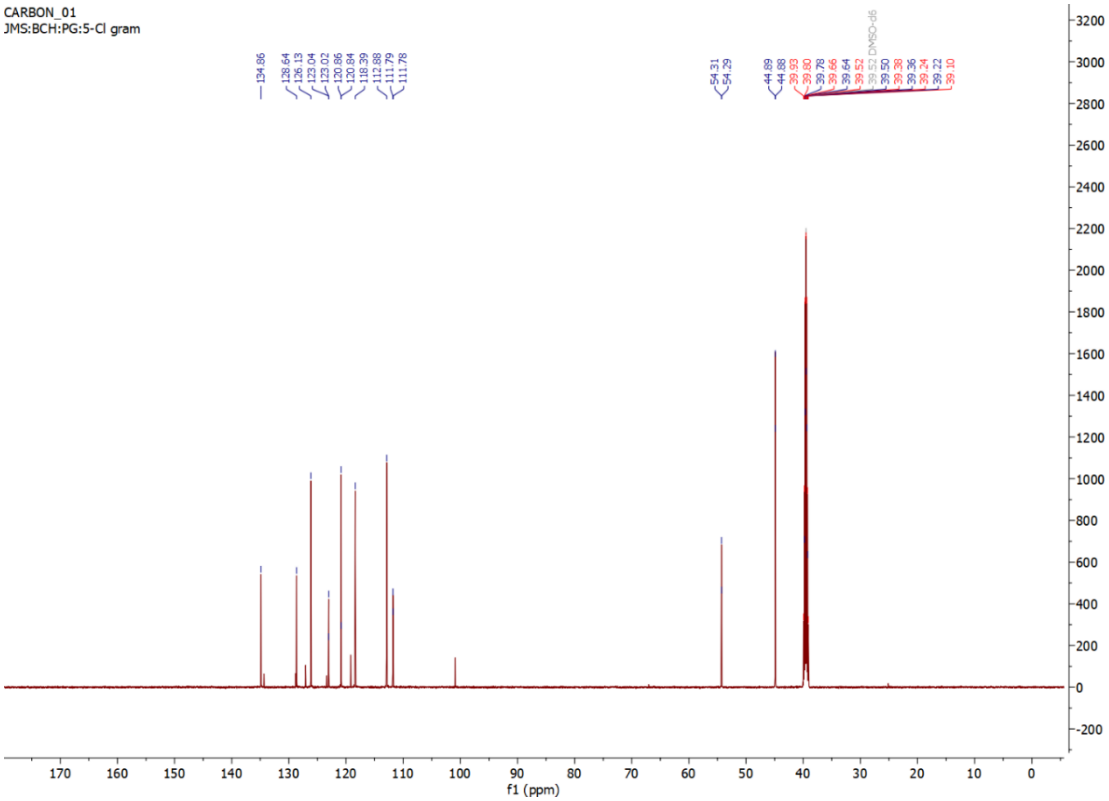
RT: 0.8568 minutes, Scan 93, 1: MS ES+ c (100.0-2000.0), NL 1.12e+8





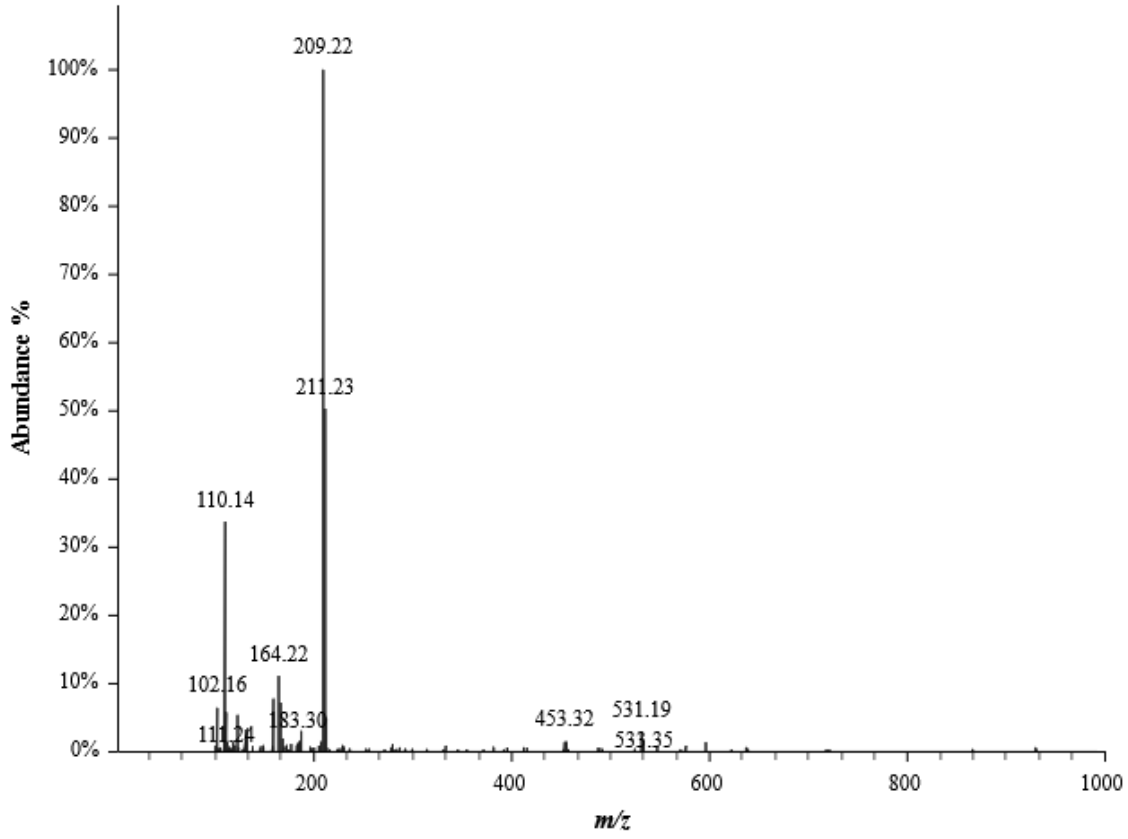
B.4.c. 5-Chlorogramine (**4c**)

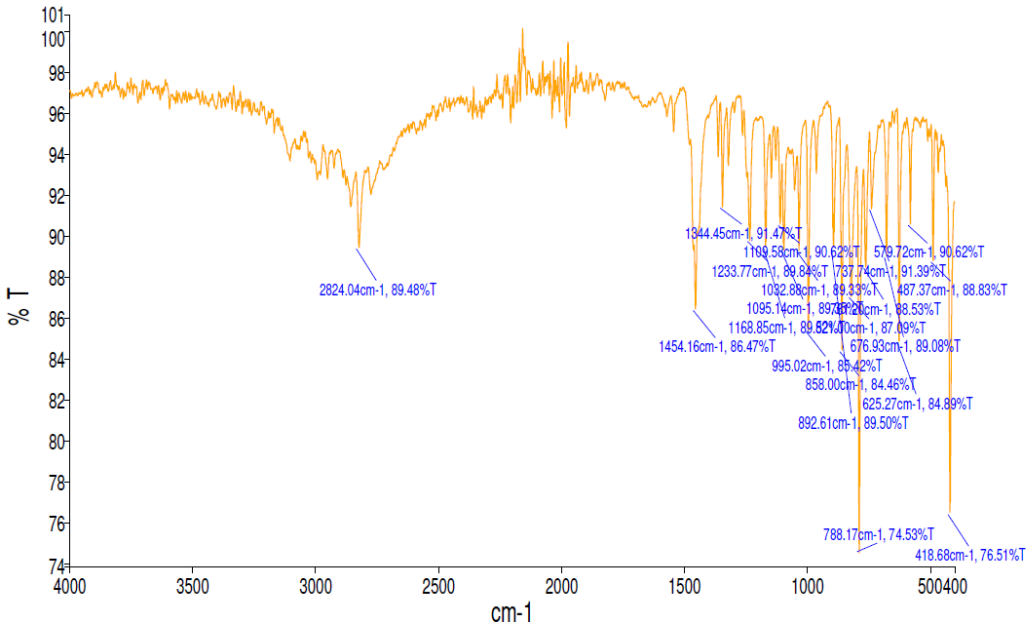




BCRH_5Clgramine_206790

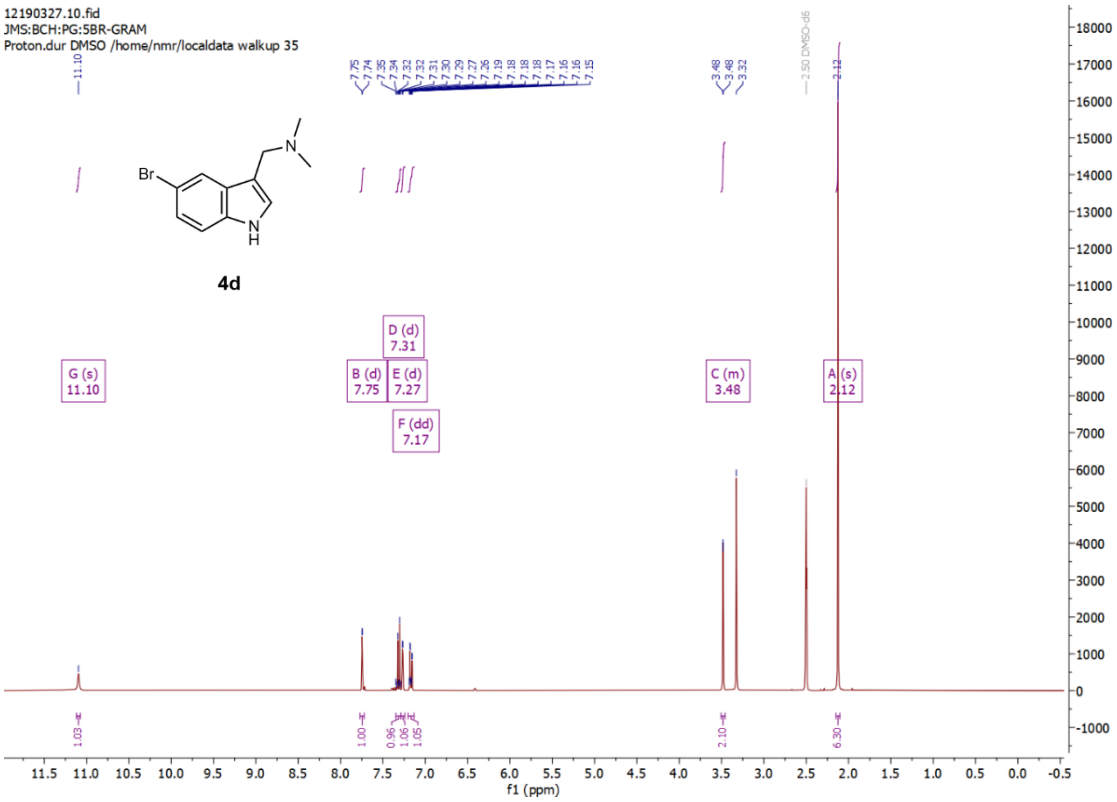
RT: 1.1691 minutes, Scan 129, 1: MS ES+ c (100.0-2000.0), NL 7.94e+6

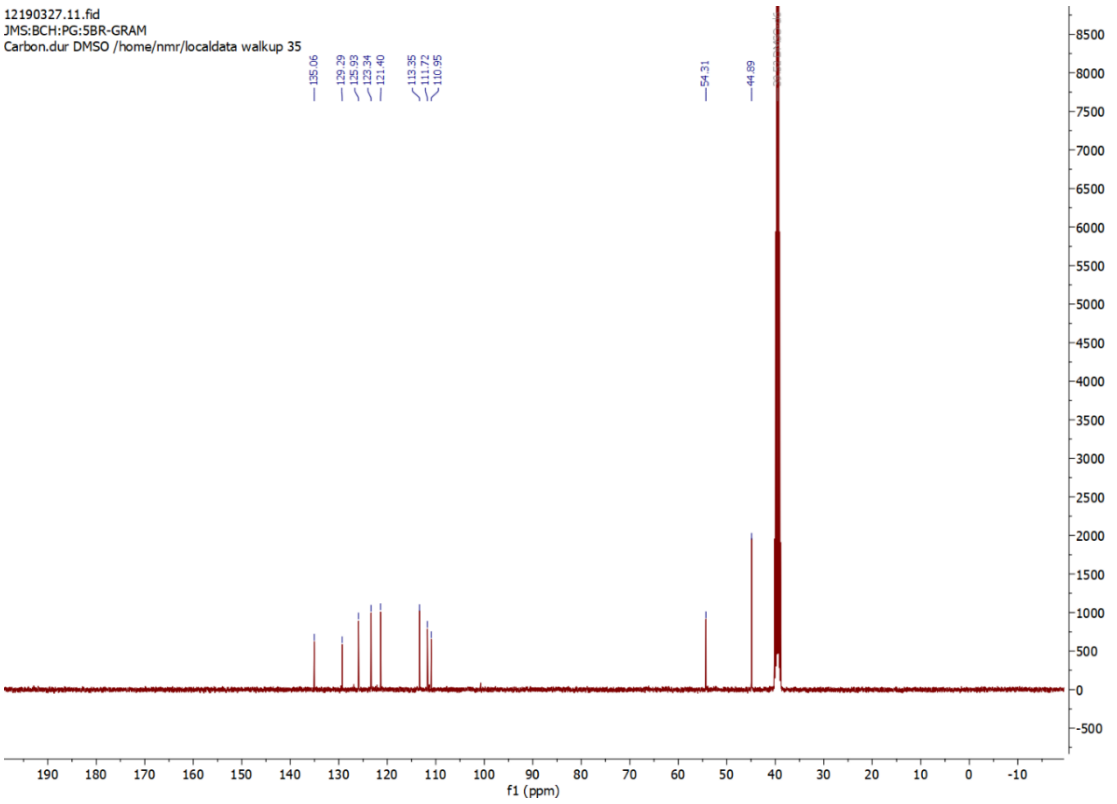




Crystal data and structure refinement for Compound 4c.	
Empirical formula	$C_{11}H_{13}ClN_2$
Formula weight	208.68
Temperature/K	120.00
Crystal system	monoclinic
Space group	$P2_1/n$
a/Å	5.5846(2)
b/Å	9.6618(4)
c/Å	19.4177(8)
α/°	90
β/°	94.981(2)
γ/°	90
Volume/Å^3	1043.77(7)
Z	4
ρ_{calc}g/cm^3	1.328
μ/mm^{-1}	0.326
F(000)	440.0
Crystal size/mm^3	0.162 × 0.114 × 0.036
Radiation	Mo Kα (λ = 0.71073)
2Θ range for data collection/°	4.212 to 52.94
Index ranges	-6 ≤ h ≤ 7, -12 ≤ k ≤ 12, -24 ≤ l ≤ 24
Reflections collected	27909
Independent reflections	2141 [R_{int} = 0.0556, R_{sigma} = 0.0247]
Data/restraints/parameters	2141/0/133
Goodness-of-fit on F^2	1.081
Final R indexes [I>=2σ (I)]	R_1 = 0.0376, wR_2 = 0.0777
Final R indexes [all data]	R_1 = 0.0489, wR_2 = 0.0818
Largest diff. peak/hole / e Å^{-3}	0.25/-0.21

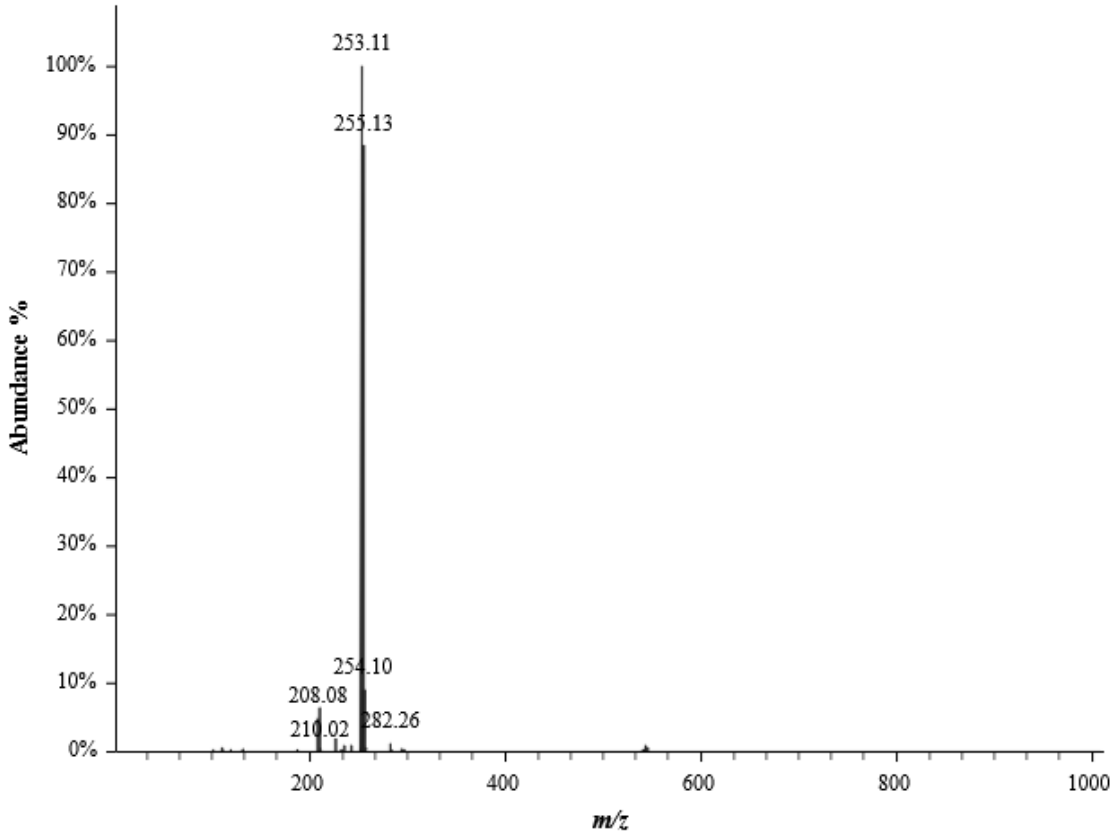
B.4.d. 5-Bromogramine (**4d**)

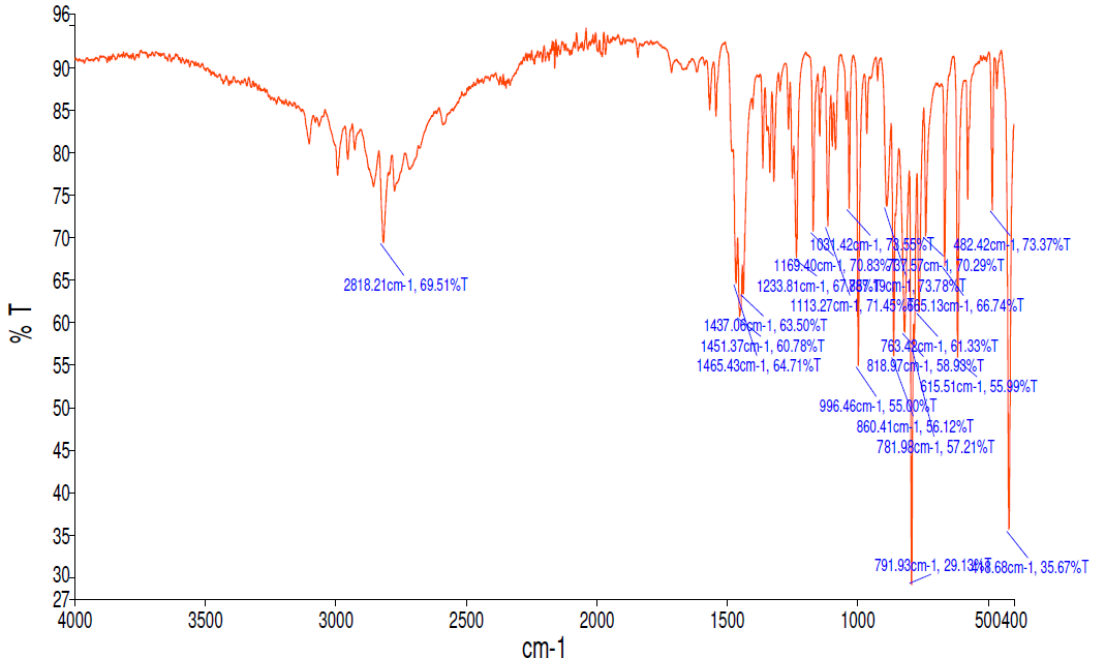




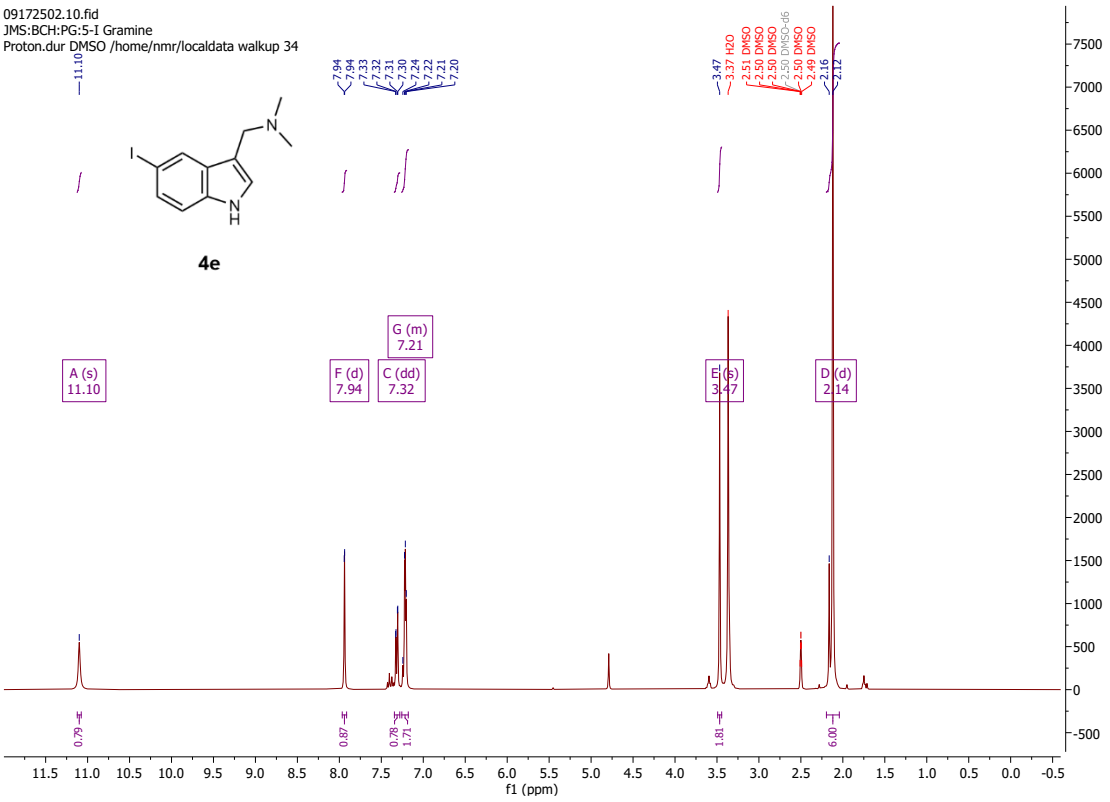
BCRH_5Br_Gramine_212066

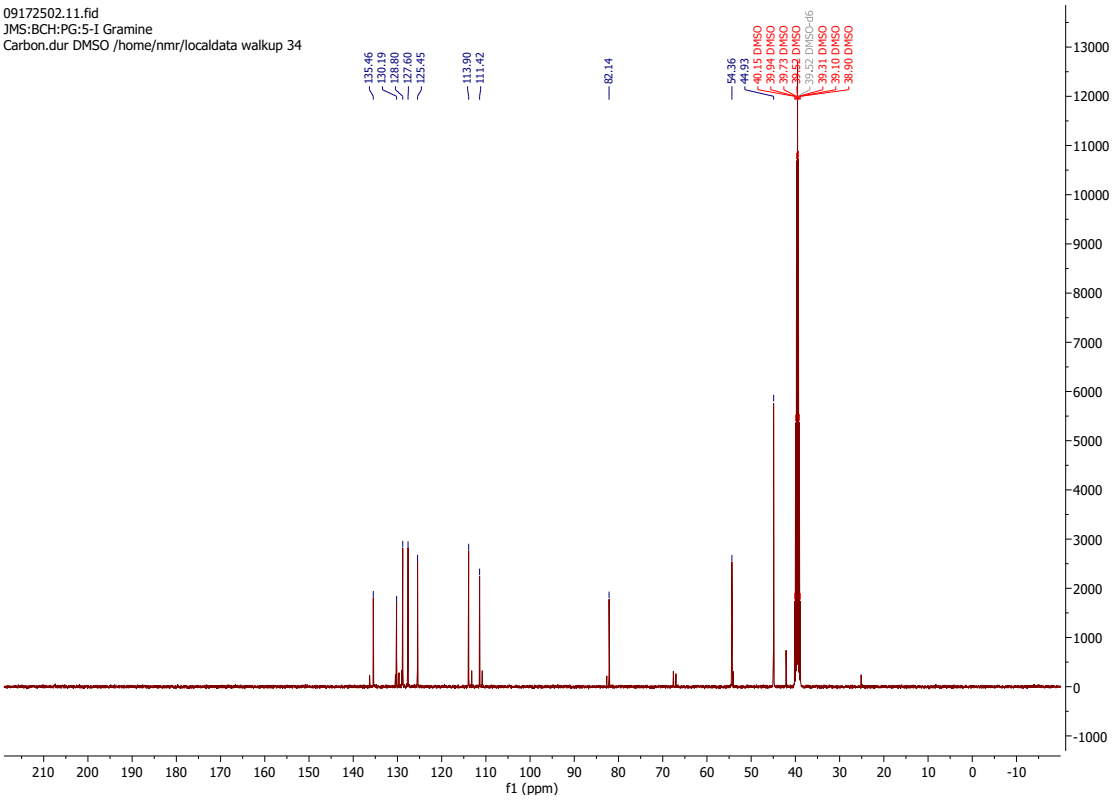
RT: 1.2559 minutes, Scan 139, 1: MS ES+ c (100.0-2000.0), NL 6.59e+7





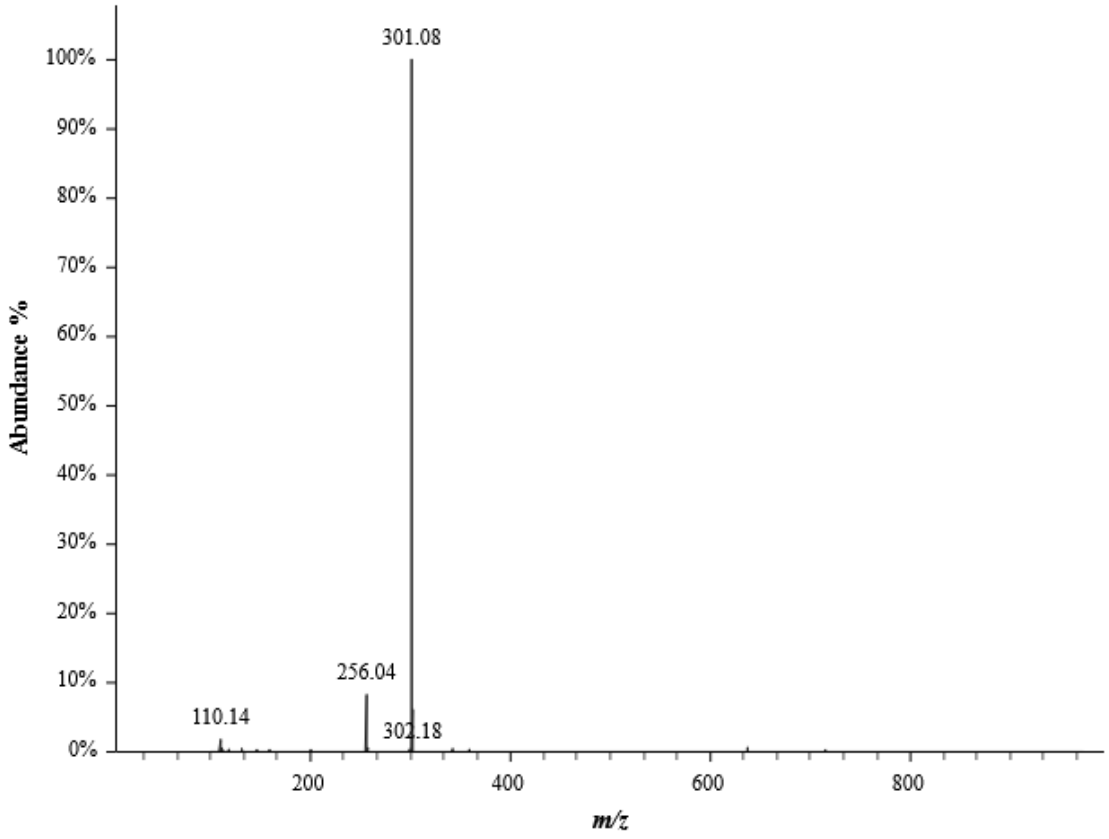
B.4.e. 5-Iodogramine (**4e**)

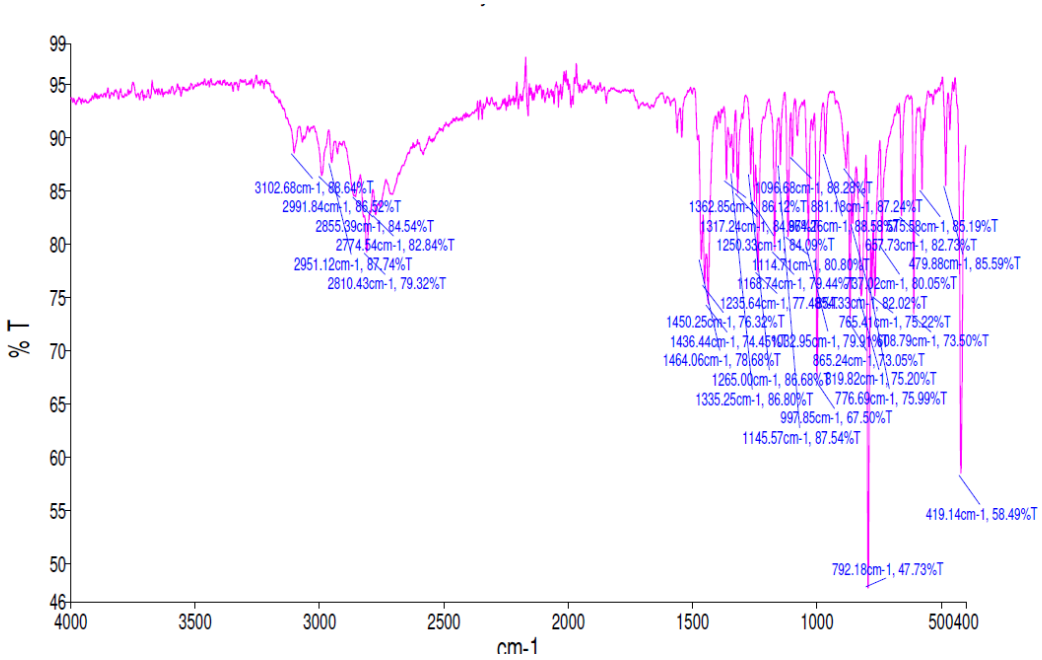




BCRH_5IGramine_210391

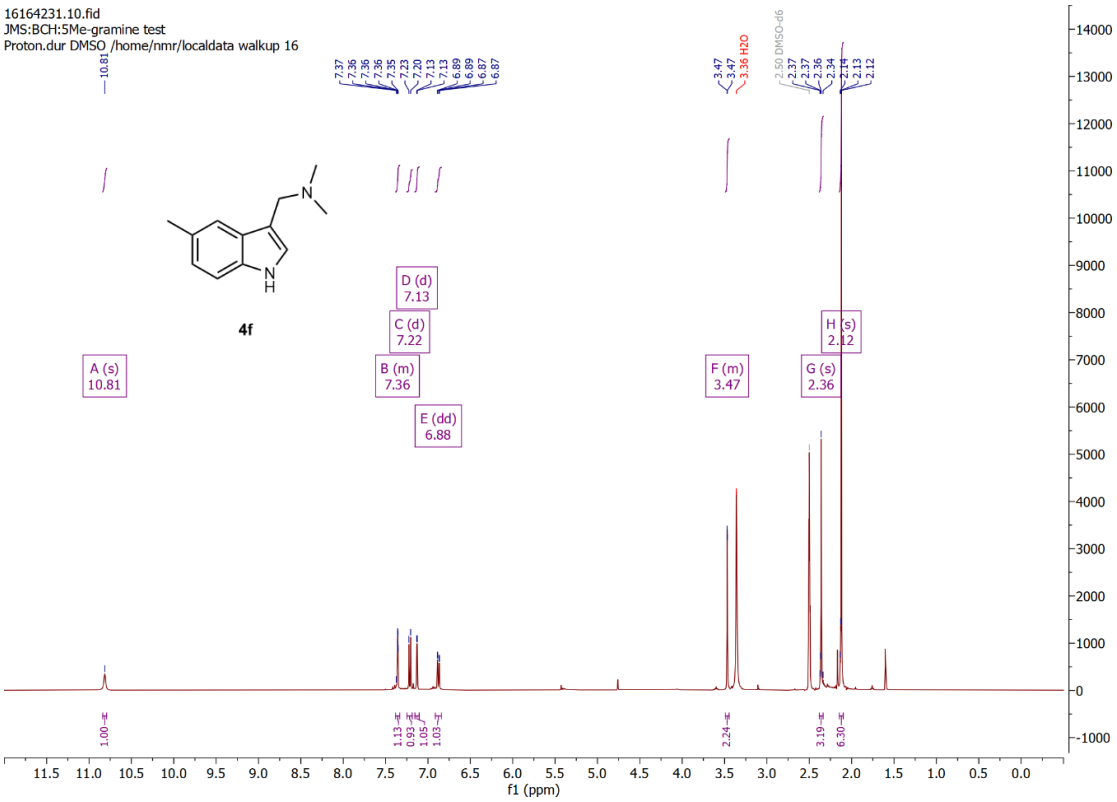
RT: 1.3427 minutes, Scan 149, 1: MS ES+ c (100.0-2000.0), NL 9.85e+7

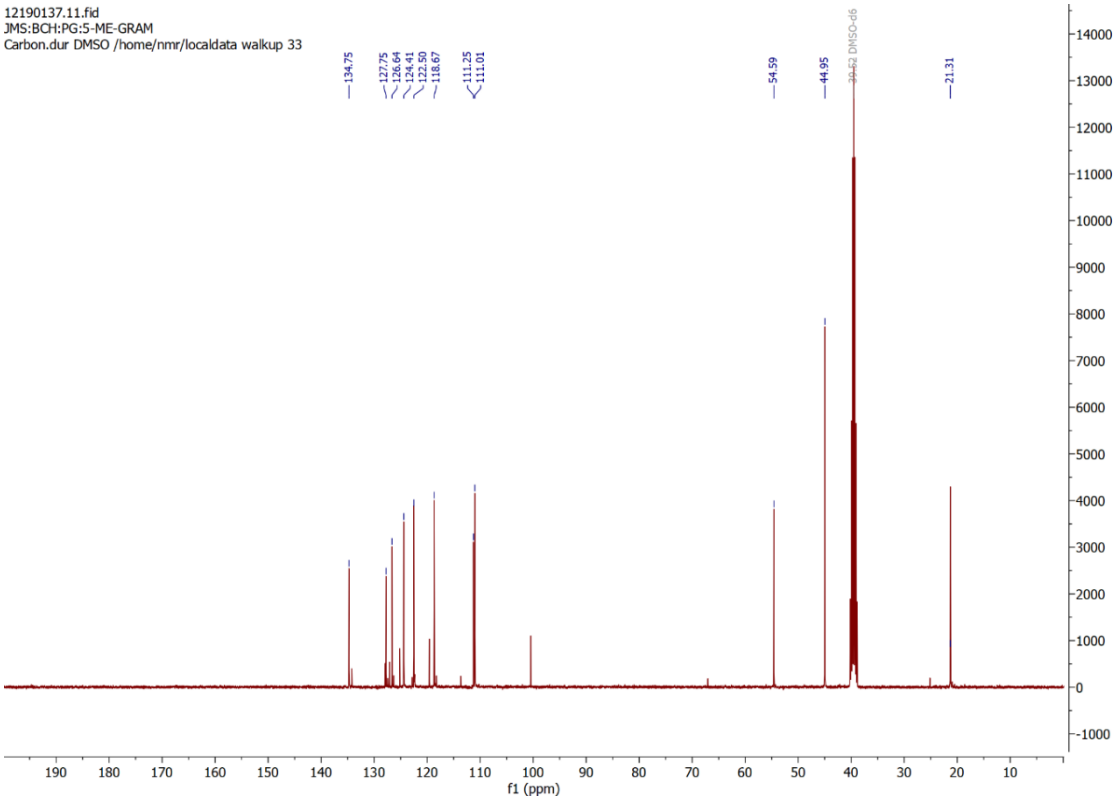




Crystal data and structure refinement for Compound 4e.	
Empirical formula	$C_{11}H_{13}IN_2$
Formula weight	300.13
Temperature/K	120.00
Crystal system	monoclinic
Space group	$P2_1$
a/Å	5.6866(2)
b/Å	9.8188(4)
c/Å	10.3151(4)
α/°	90
β/°	102.4980(10)
γ/°	90
Volume/Å^3	562.30(4)
Z	2
ρ_{calc}g/cm^3	1.773
μ/mm^{-1}	2.811
F(000)	292.0
Crystal size/mm^3	0.584 × 0.378 × 0.282
Radiation	Mo Kα (λ = 0.71073)
2Θ range for data collection/°	4.044 to 72.718
Index ranges	-9 ≤ h ≤ 9, -16 ≤ k ≤ 16, -17 ≤ l ≤ 17
Reflections collected	25910
Independent reflections	5242 [R_{int} = 0.0360, R_{sigma} = 0.0280]
Data/restraints/parameters	5242/1/132
Goodness-of-fit on F^2	1.093
Final R indexes [I>=2σ (I)]	R_1 = 0.0242, wR_2 = 0.0517
Final R indexes [all data]	R_1 = 0.0263, wR_2 = 0.0524
Largest diff. peak/hole / e Å^{-3}	0.52/-1.44
Flack parameter	-0.013(12)

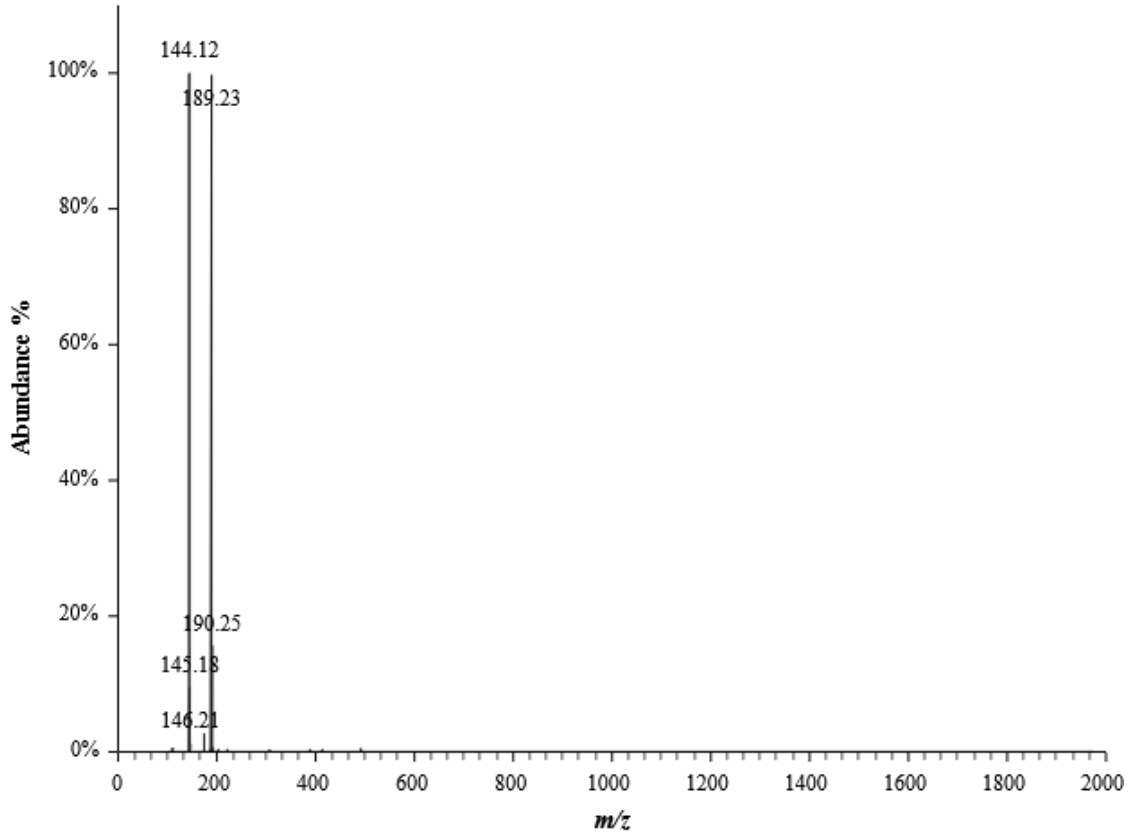
B.4.f. 5-Methyl gramine (**4f**)

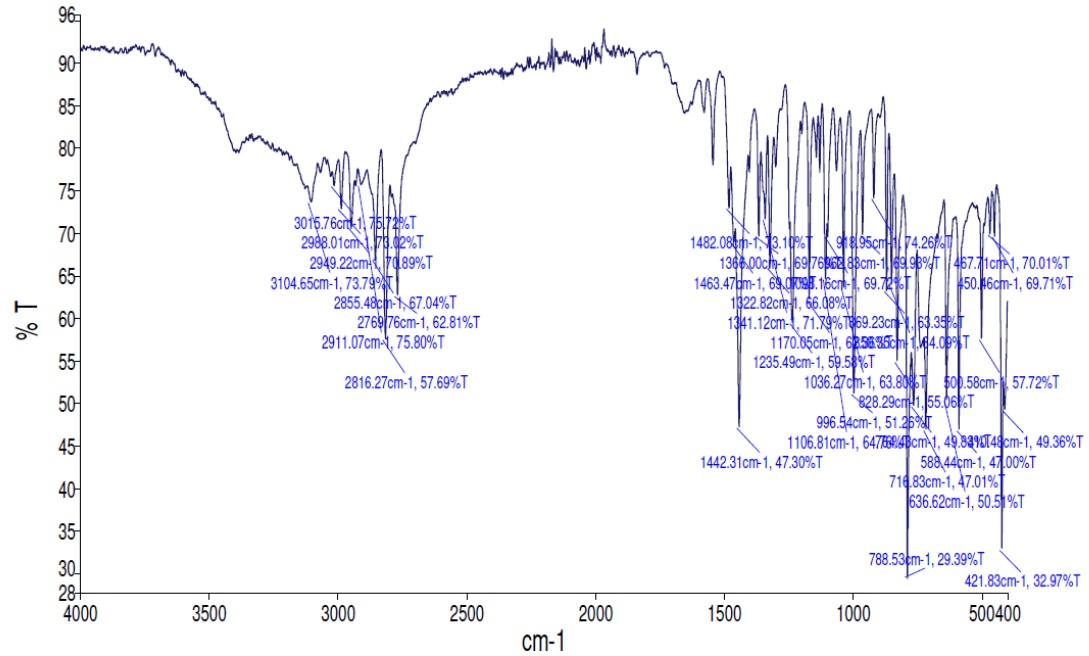




BCRH_5MeGramine_210477

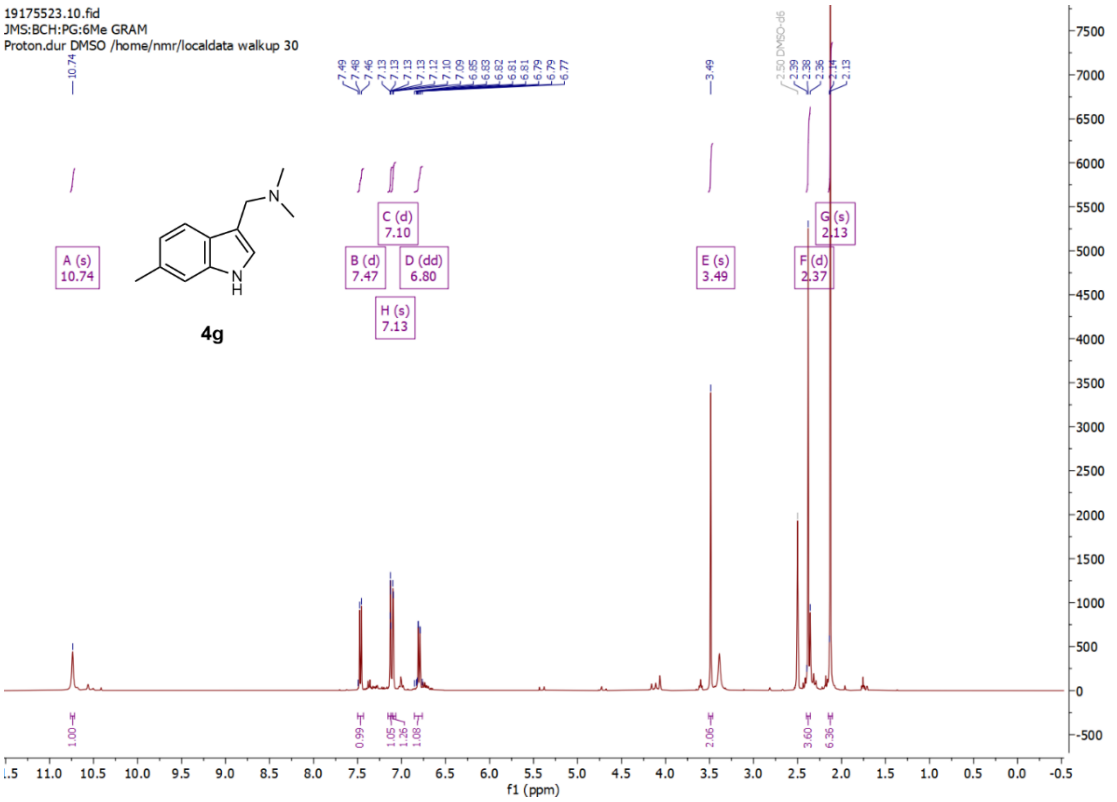
RT: 1.0477 minutes, Scan 115, 1: MS ES+ c (100.0-2000.0), NL 8.91e+7

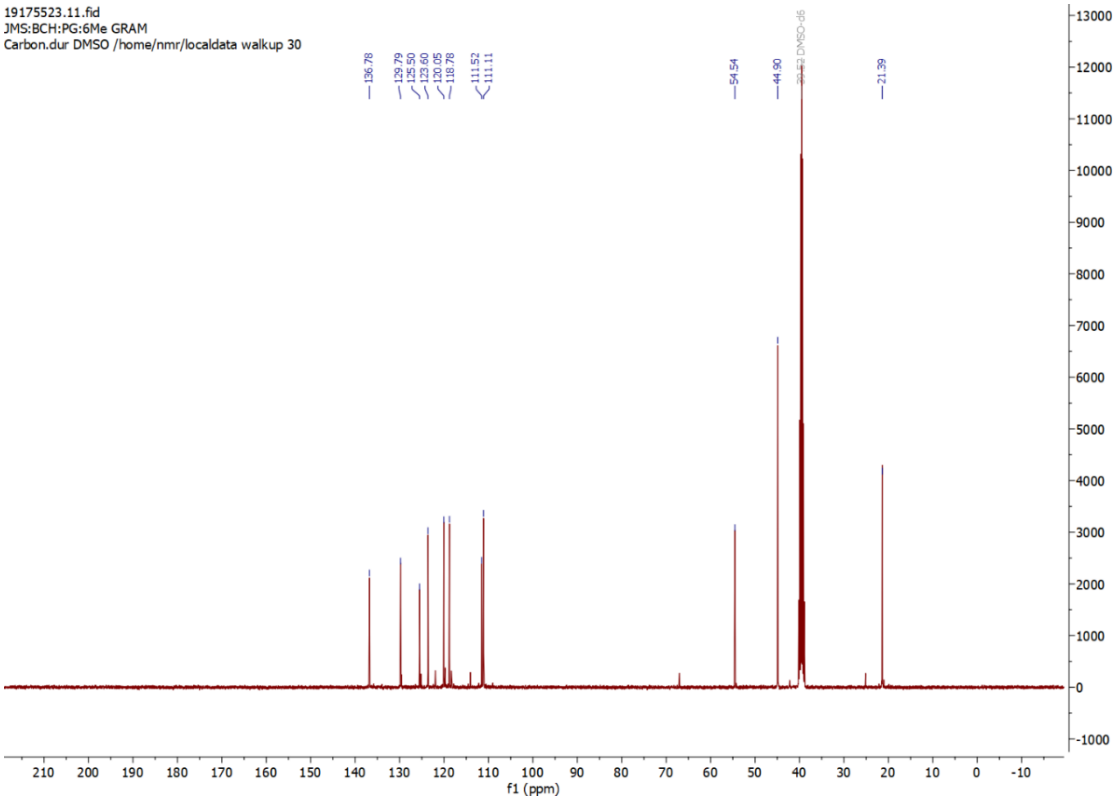




Crystal data and structure refinement for Compound 4f.	
Identification code	24srv187
Empirical formula	$C_{12}H_{16}N_2$
Formula weight	188.27
Temperature/K	120.00
Crystal system	monoclinic
Space group	$P2_1/n$
a/Å	5.7816(3)
b/Å	9.5129(5)
c/Å	19.3858(10)
α/°	90
β/°	93.981(2)
γ/°	90
Volume/Å^3	1063.64(10)
Z	4
ρ_{calc}g/cm^3	1.176
μ/mm^{-1}	0.070
F(000)	408.0
Crystal size/mm^3	0.318 × 0.193 × 0.016
Radiation	Mo Kα (λ = 0.71073)
2Θ range for data collection/°	4.212 to 57.302
Index ranges	-7 ≤ h ≤ 7, -12 ≤ k ≤ 12, -26 ≤ l ≤ 26
Reflections collected	28010
Independent reflections	2728 [R_{int} = 0.0648, R_{sigma} = 0.0341]
Data/restraints/parameters	2728/0/134
Goodness-of-fit on F^2	1.185
Final R indexes [I>=2σ (I)]	R_1 = 0.0638, wR_2 = 0.1305
Final R indexes [all data]	R_1 = 0.0770, wR_2 = 0.1359
Largest diff. peak/hole / e Å^{-3}	0.31/-0.28

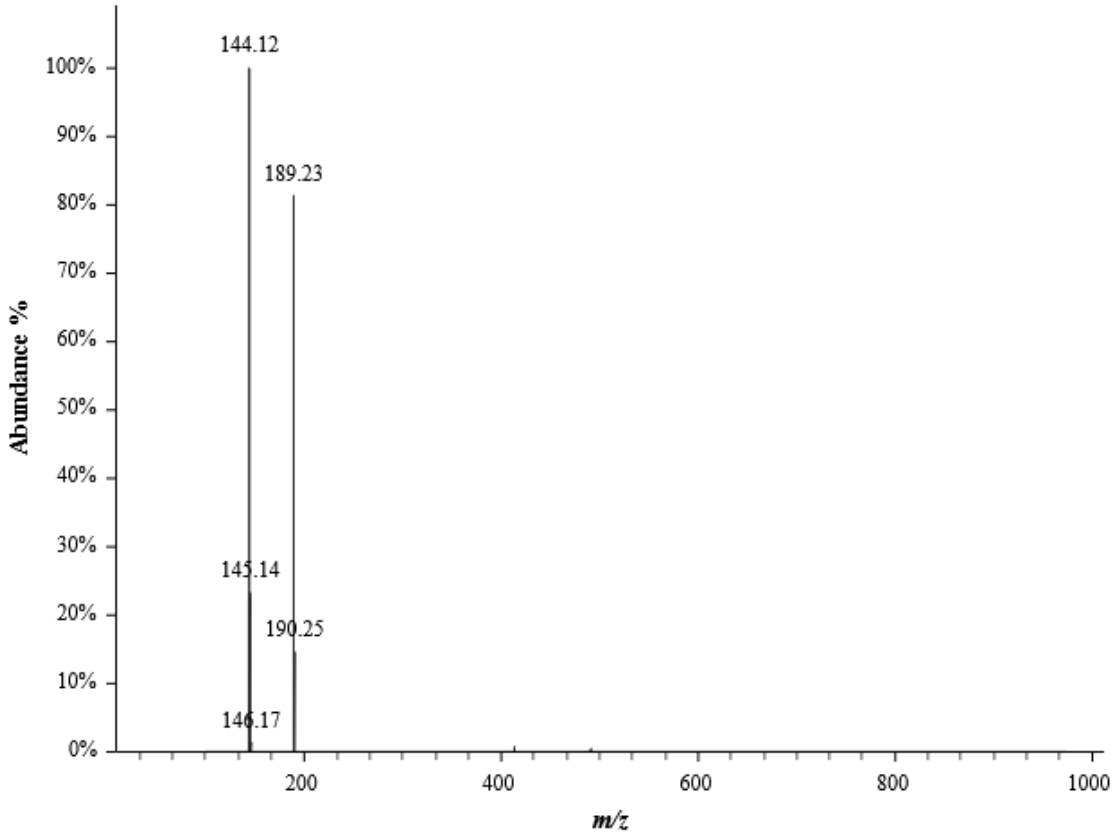
B.4.g. 6-Methylgramine (**4g**)

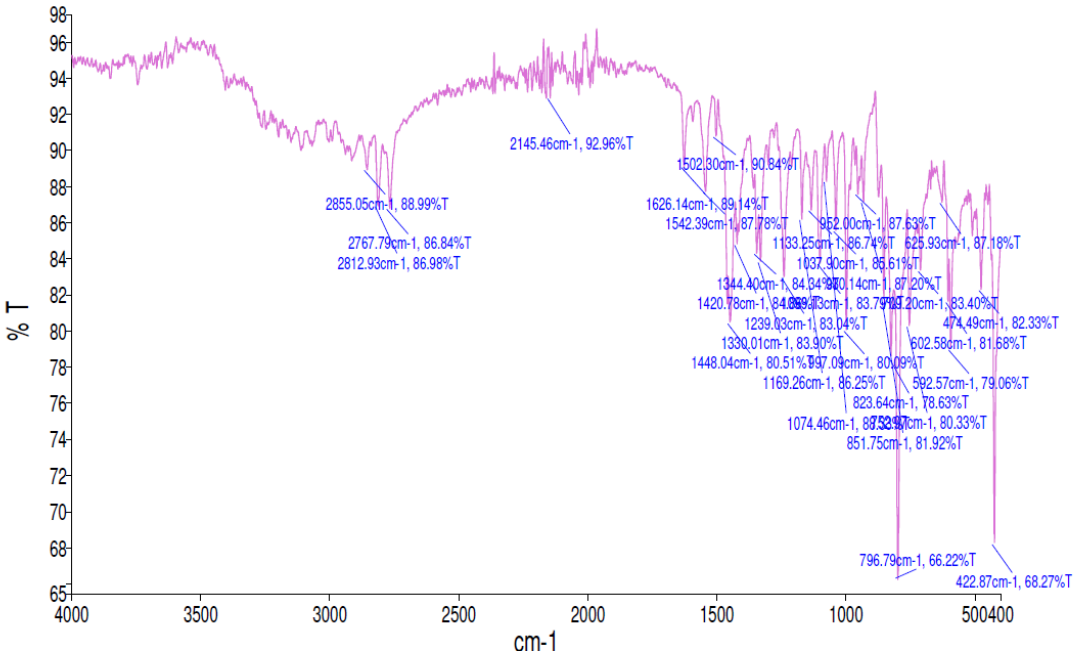




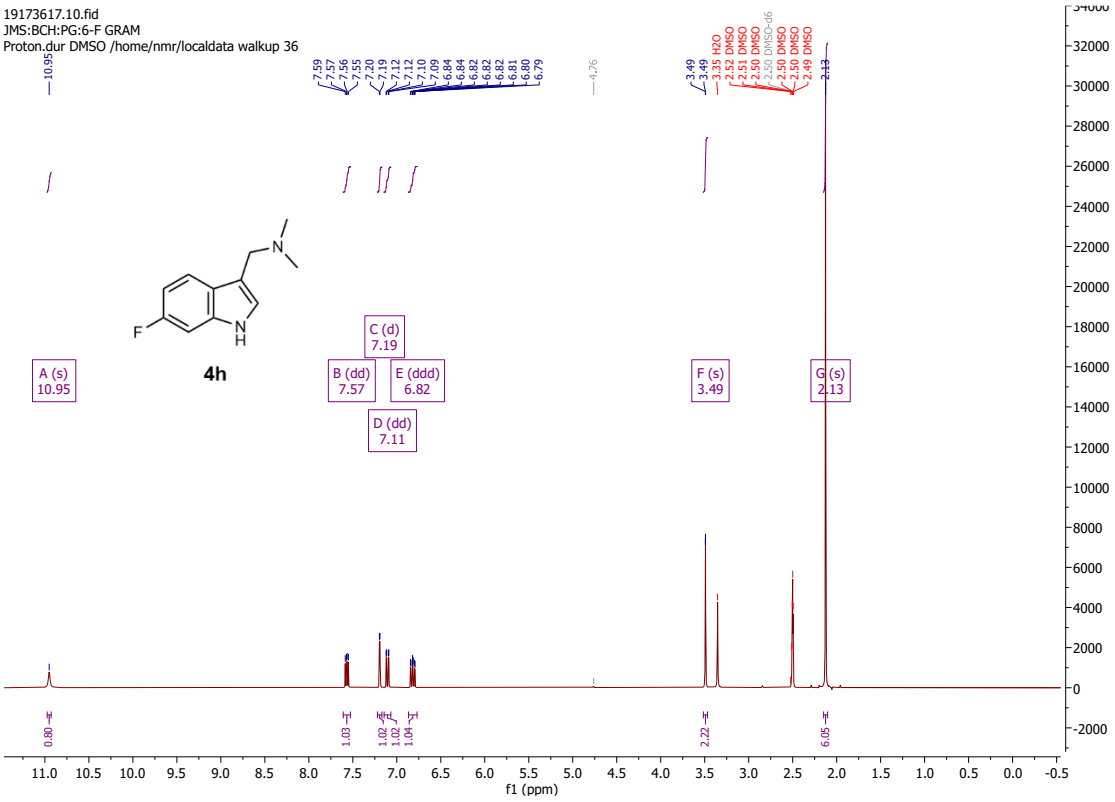
BCRH_6MeGramine_210509

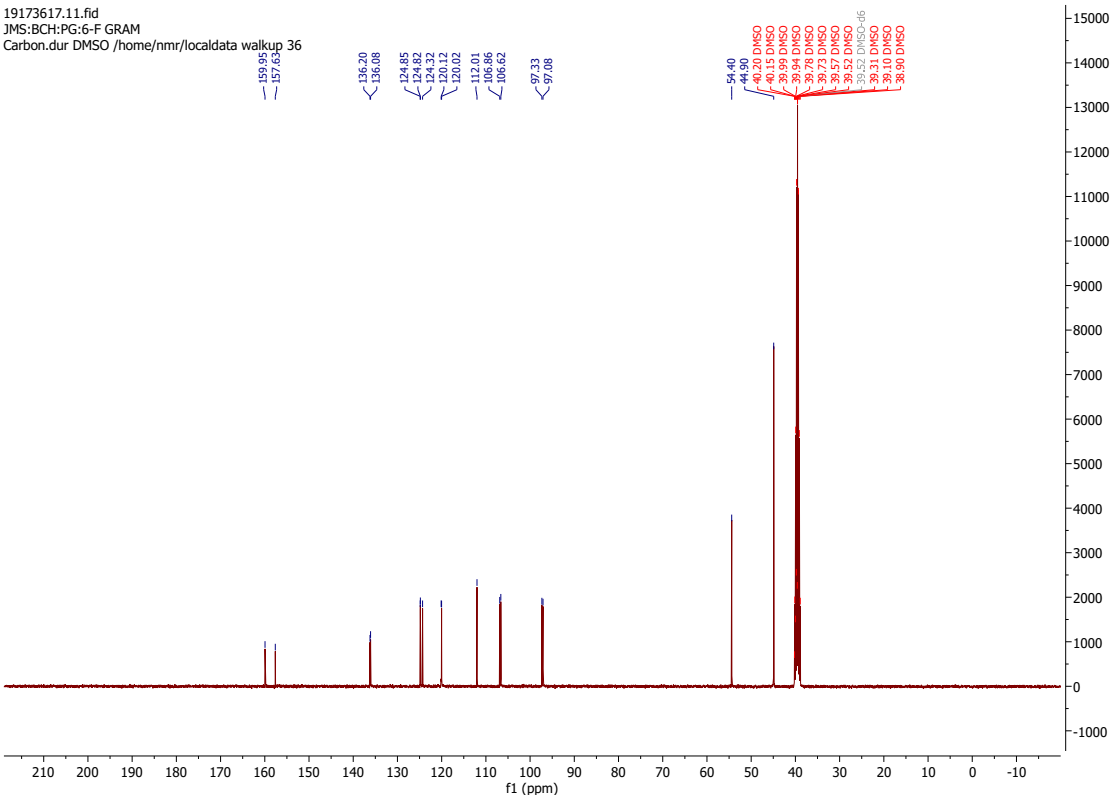
RT: 1.0303 minutes, Scan 113, 1: MS ES+ c (100.0-2000.0), NL 1.17e+8

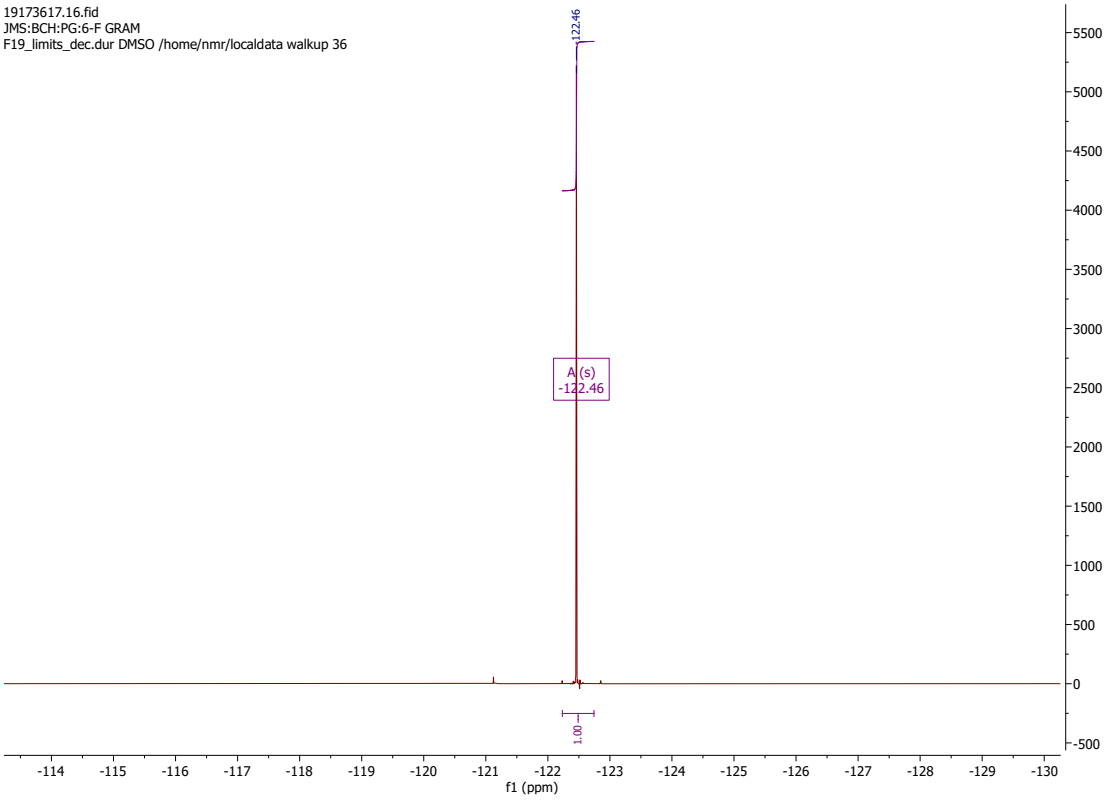




B.4.h. 6-Fluorogramine (**4h**)

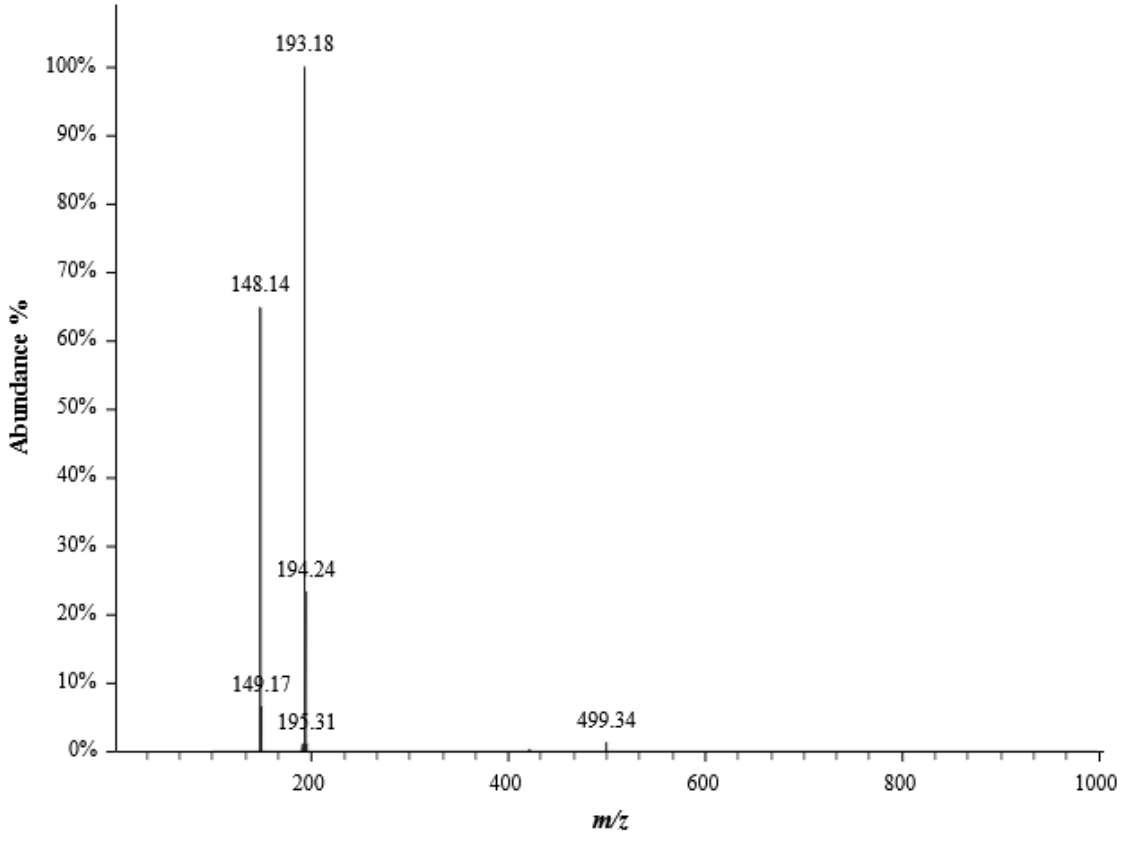


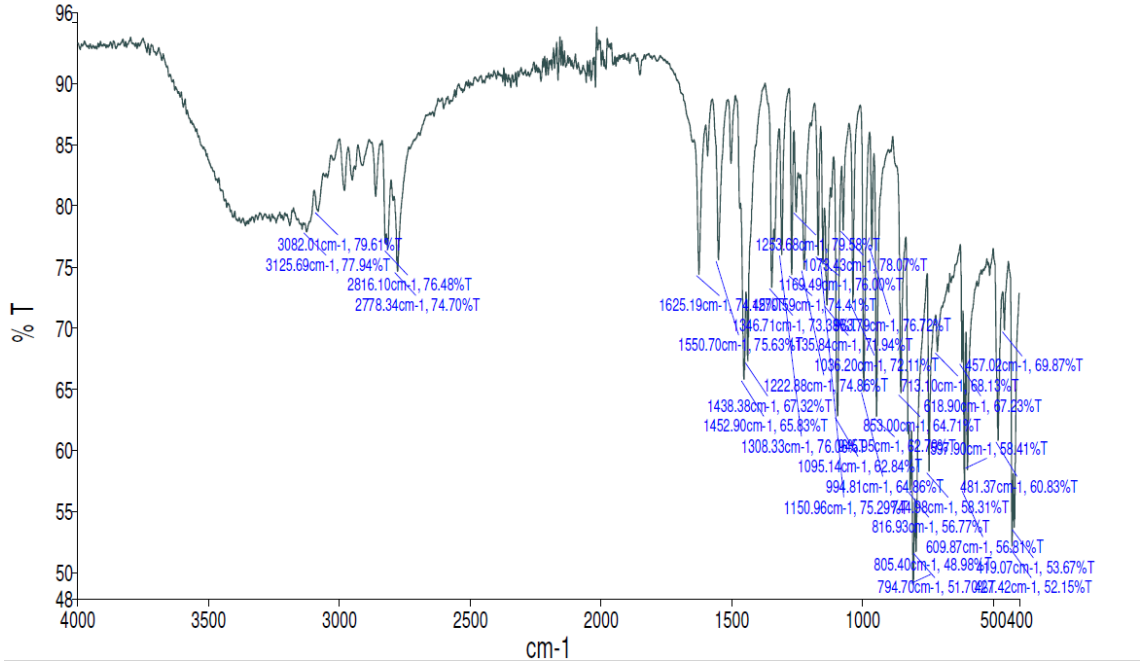




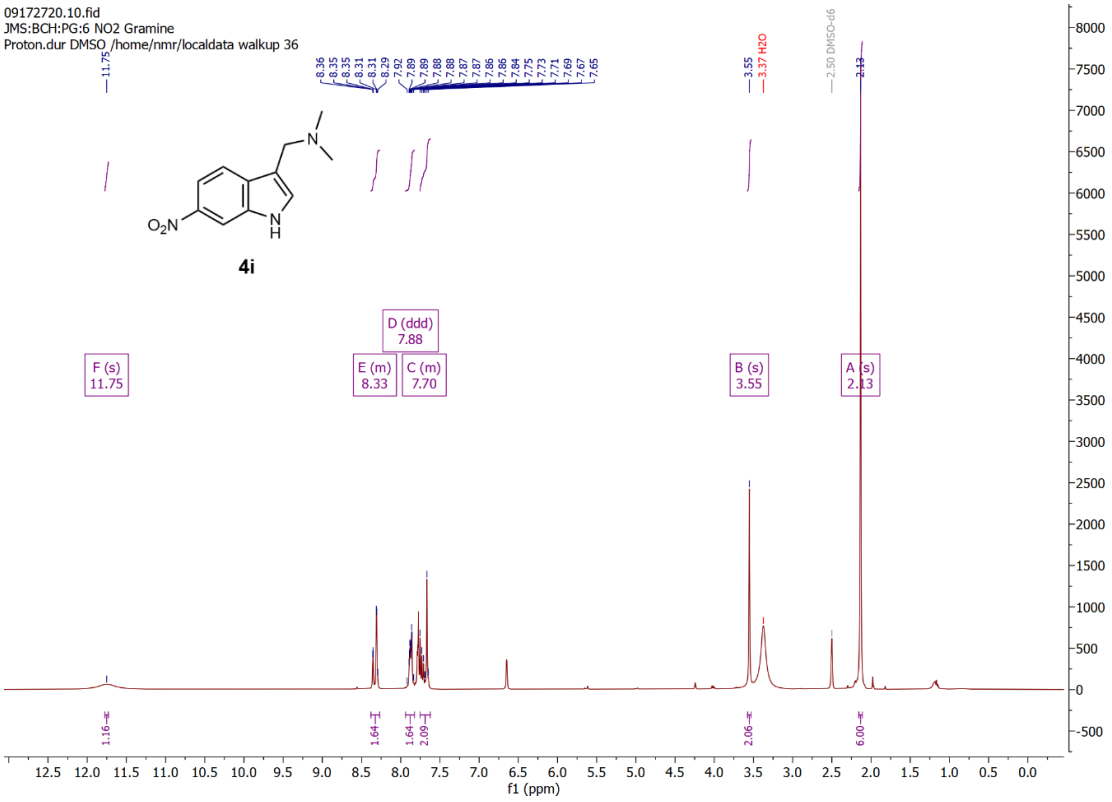
BCRH_6FGramine_210472

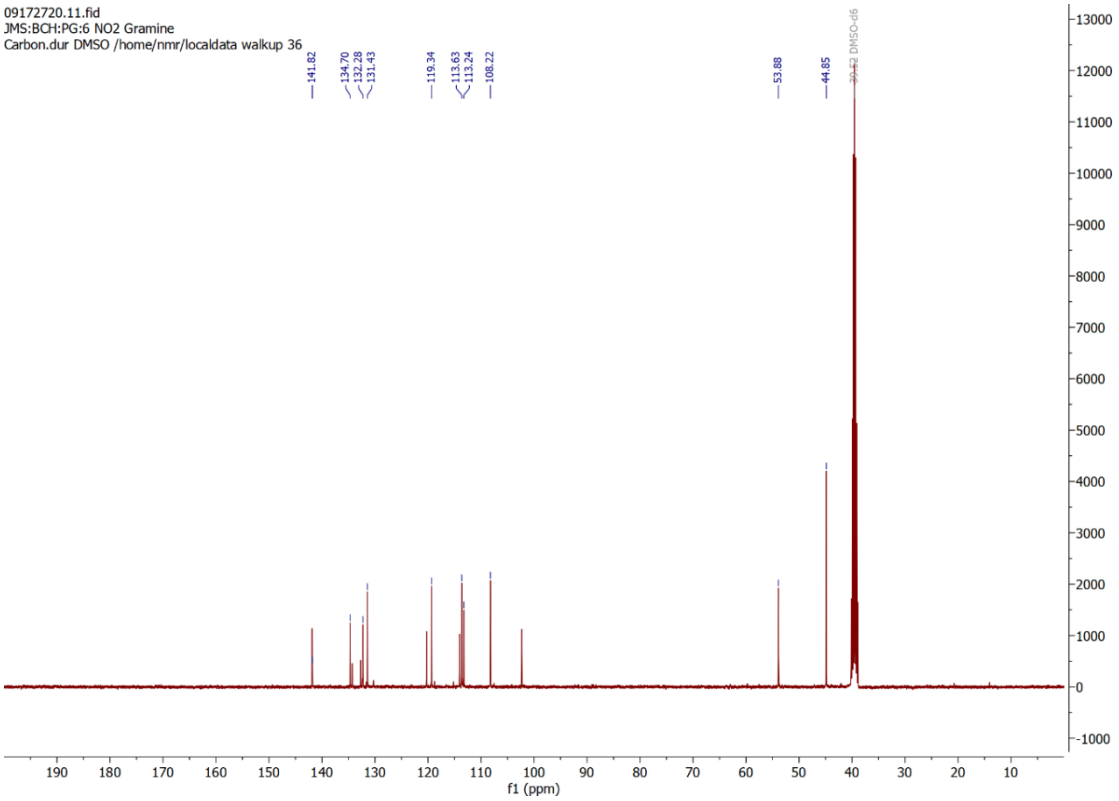
RT: 0.7873 minutes, Scan 85, 1: MS ES+ c (100.0-2000.0), NL 1.18e+8





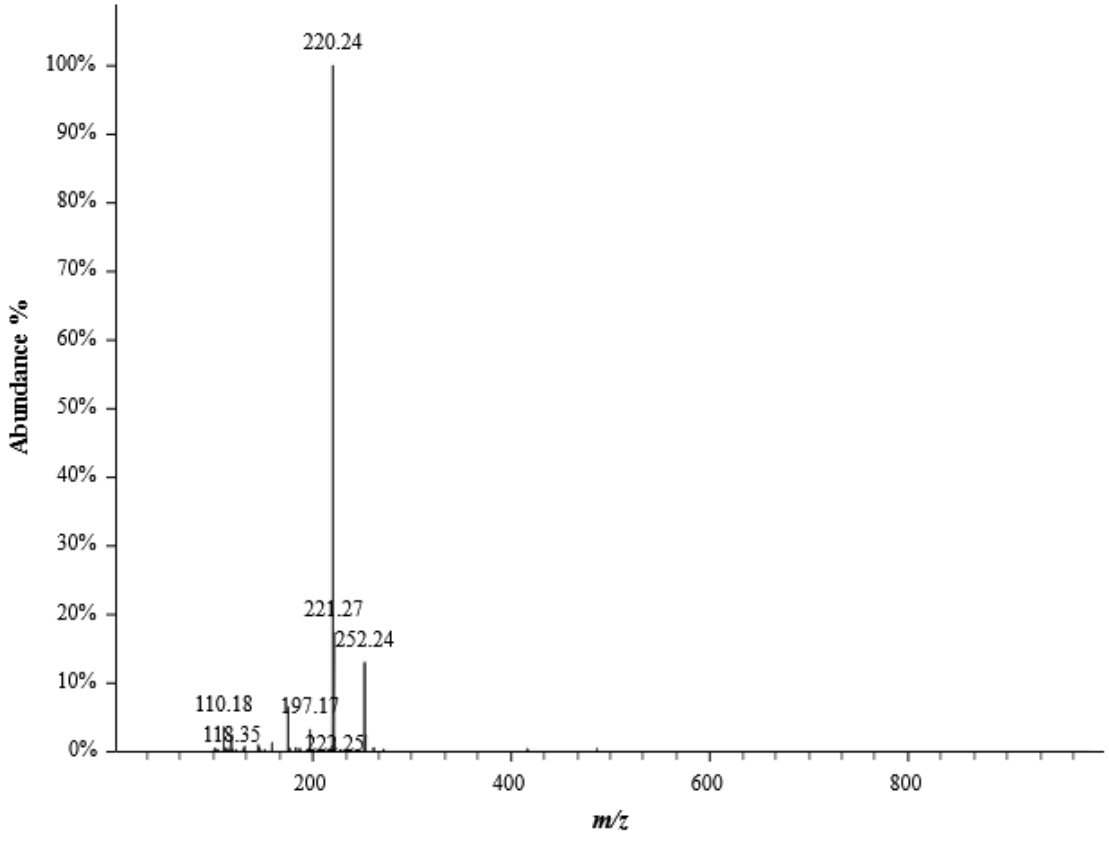
B.4.i. 6-Nitrogramine (**4i**)

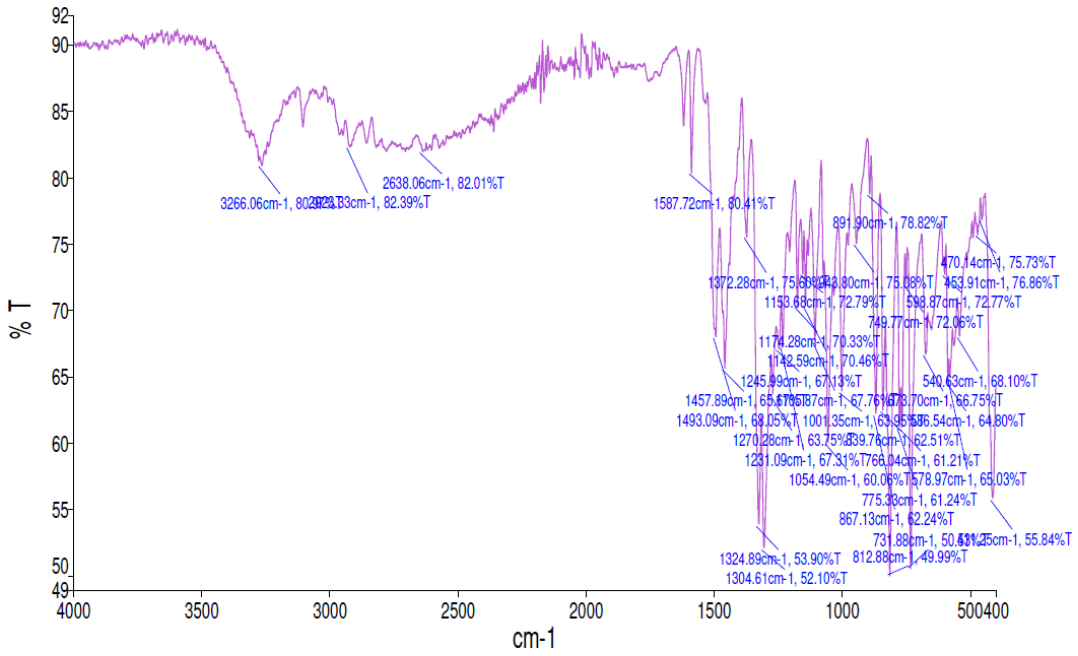




BCRH_6NO2_gramine_211806

RT: 0.9435 minutes, Scan 103, 1: MS ES+ c (100.0-2000.0), NL 2.68e+7

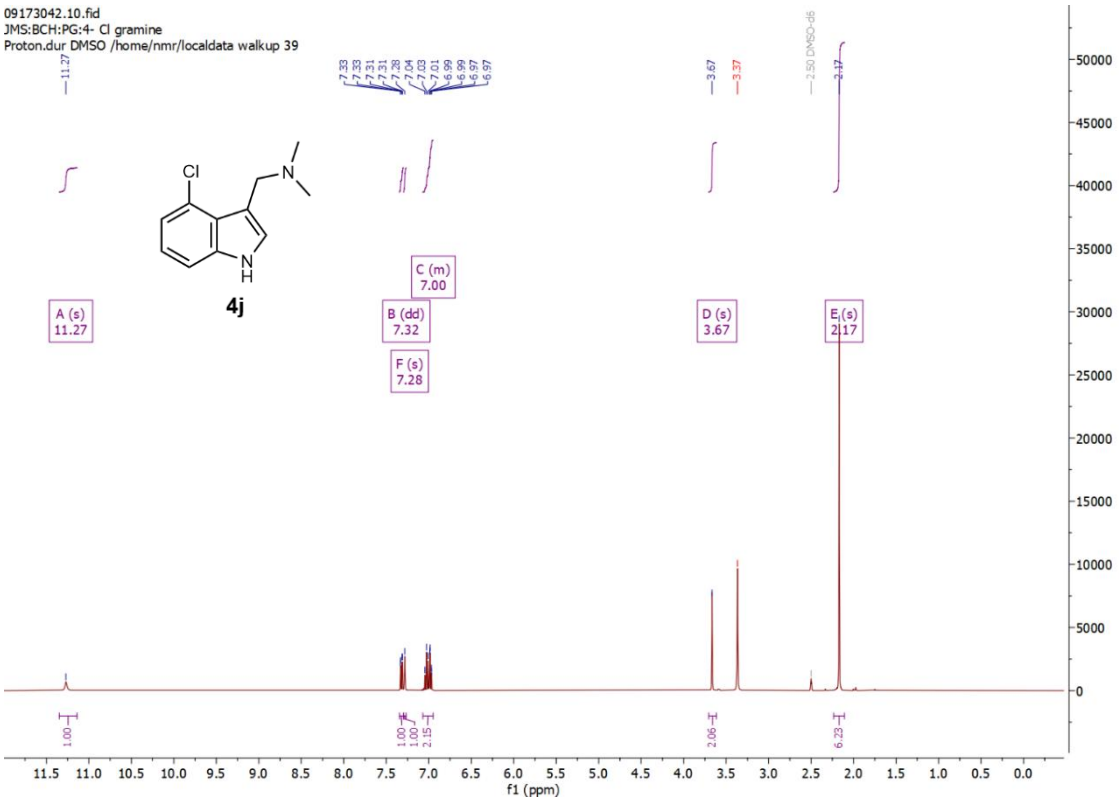


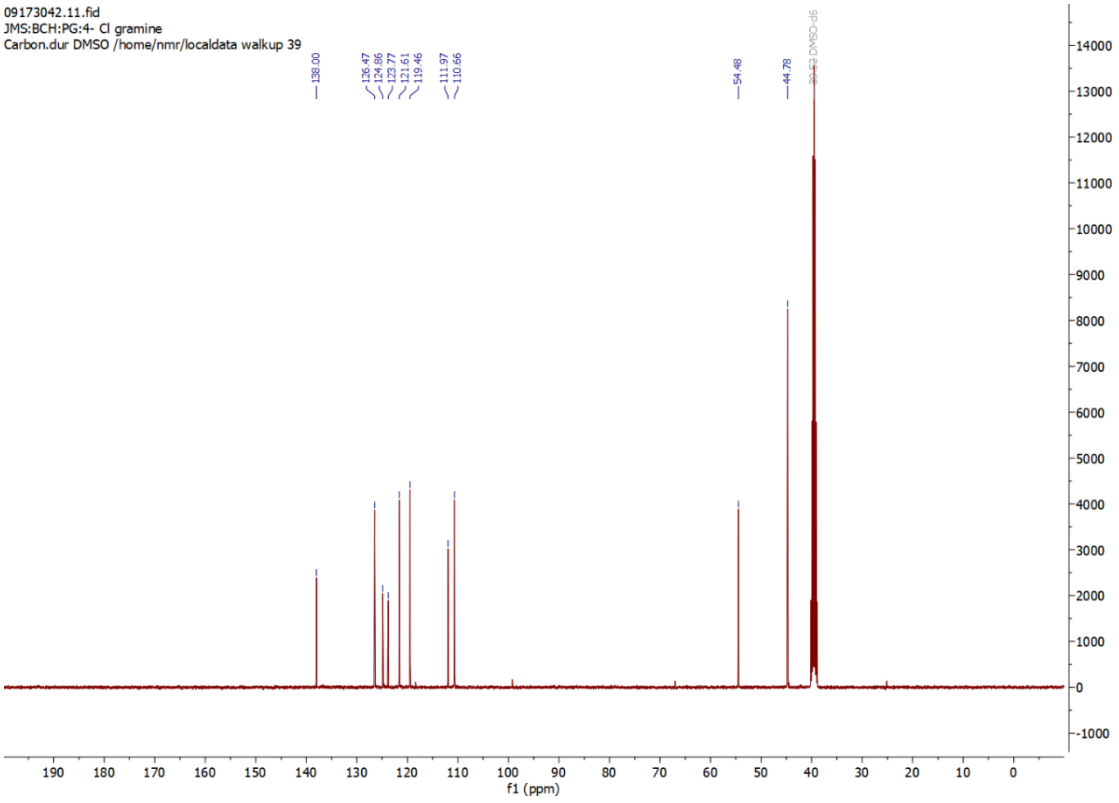


Crystal data and structure refinement for Compound 4i.	
Identification code	24srv205
Empirical formula	$C_{12}H_{18}IN_3O_3$
Formula weight	379.19
Temperature/K	120.00
Crystal system	monoclinic
Space group	$P2_1/c$
a/Å	6.8617(5)
b/Å	22.0043(16)
c/Å	10.3976(8)
α/°	90
β/°	102.477(2)
γ/°	90
Volume/Å^3	1532.8(2)
Z	4
ρ_{calc}g/cm^3	1.643
μ/mm^{-1}	2.097
F(000)	752.0
Crystal size/mm^3	0.347 × 0.139 × 0.037
Radiation	Mo Kα (λ = 0.71073)
2Θ range for data collection/°	3.702 to 54.93
Index ranges	-8 ≤ h ≤ 8, -28 ≤ k ≤ 28, -13 ≤ l ≤ 13
Reflections collected	36910
Independent reflections	3508 [R_{int} = 0.0333, R_{sigma} = 0.0162]
Data/restraints/parameters	3508/0/178
Goodness-of-fit on F^2	1.300
Final R indexes [I>=2σ (I)]	R_1 = 0.0402, wR_2 = 0.0867
Final R indexes [all data]	R_1 = 0.0415, wR_2 = 0.0873

Largest diff. peak/hole / e Å^{-3}	0.92/-1.43

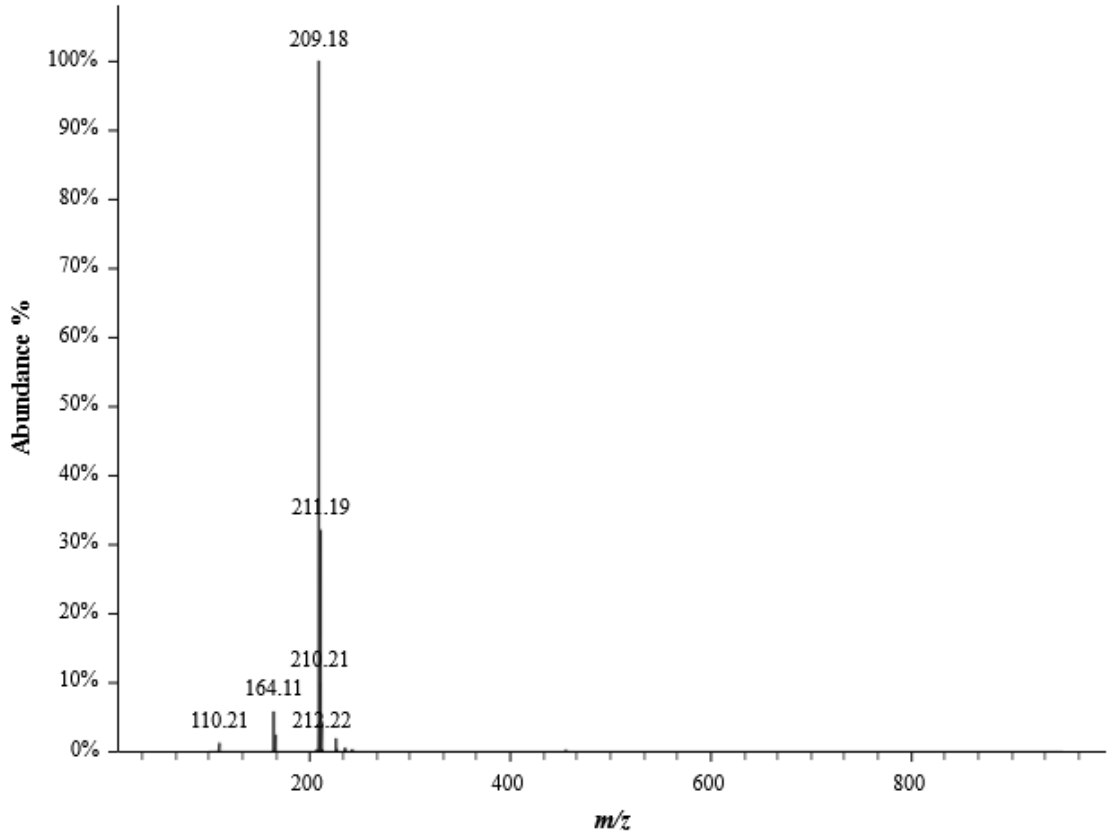
B.4.j. 4-Chlorogramine (**4j**)

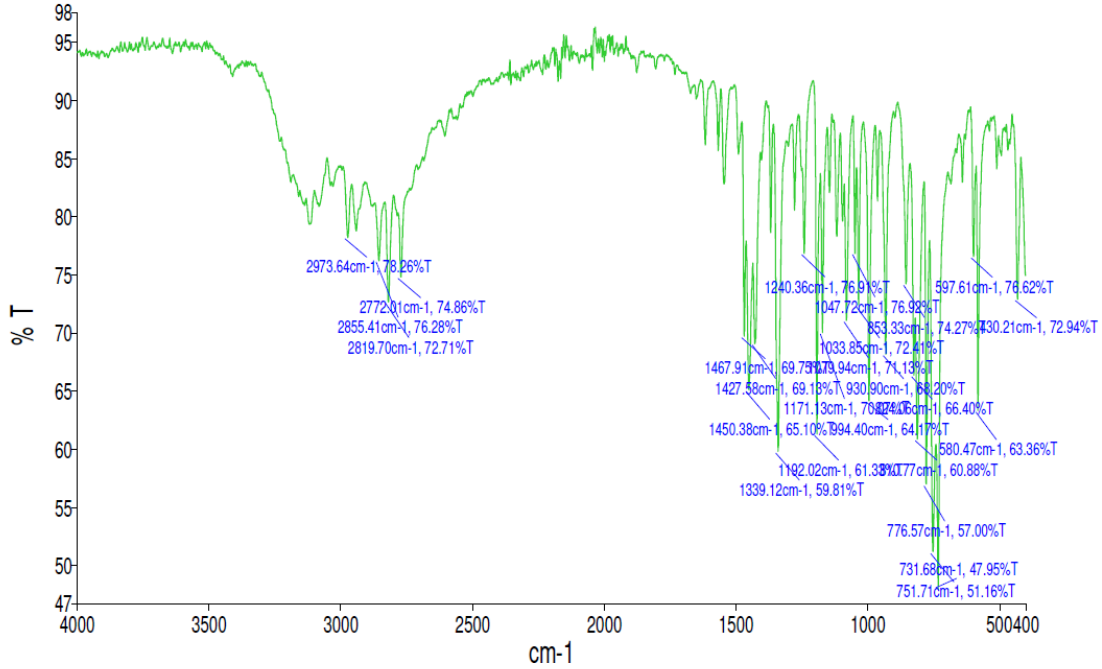




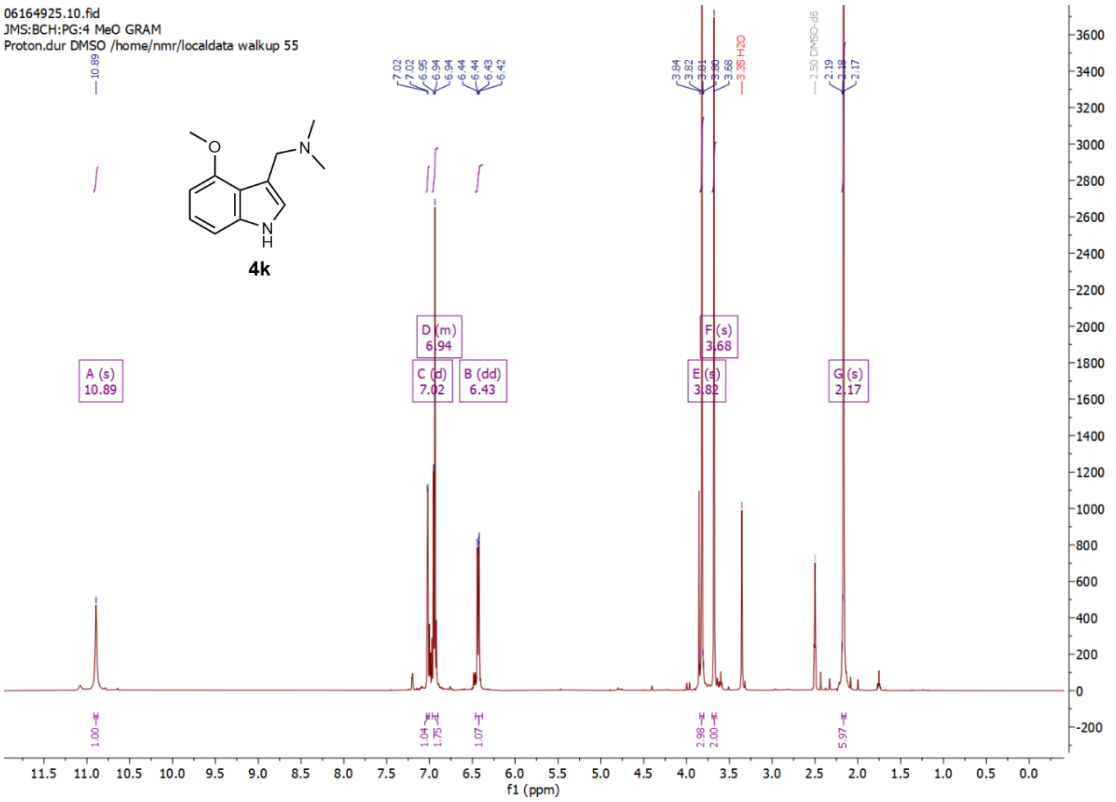
BCRH_4Cl_gram_212064

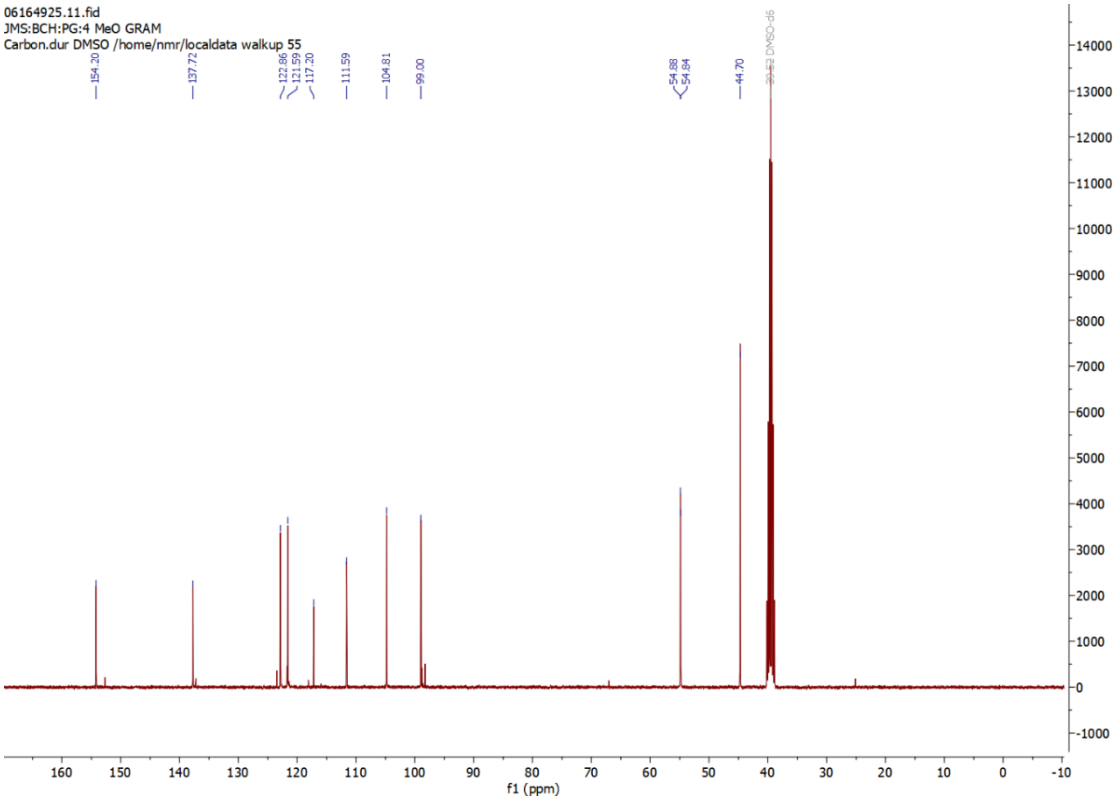
RT: 1.1171 minutes, Scan 123, 1: MS ES+ c (100.0-2000.0), NL 7.57e+7





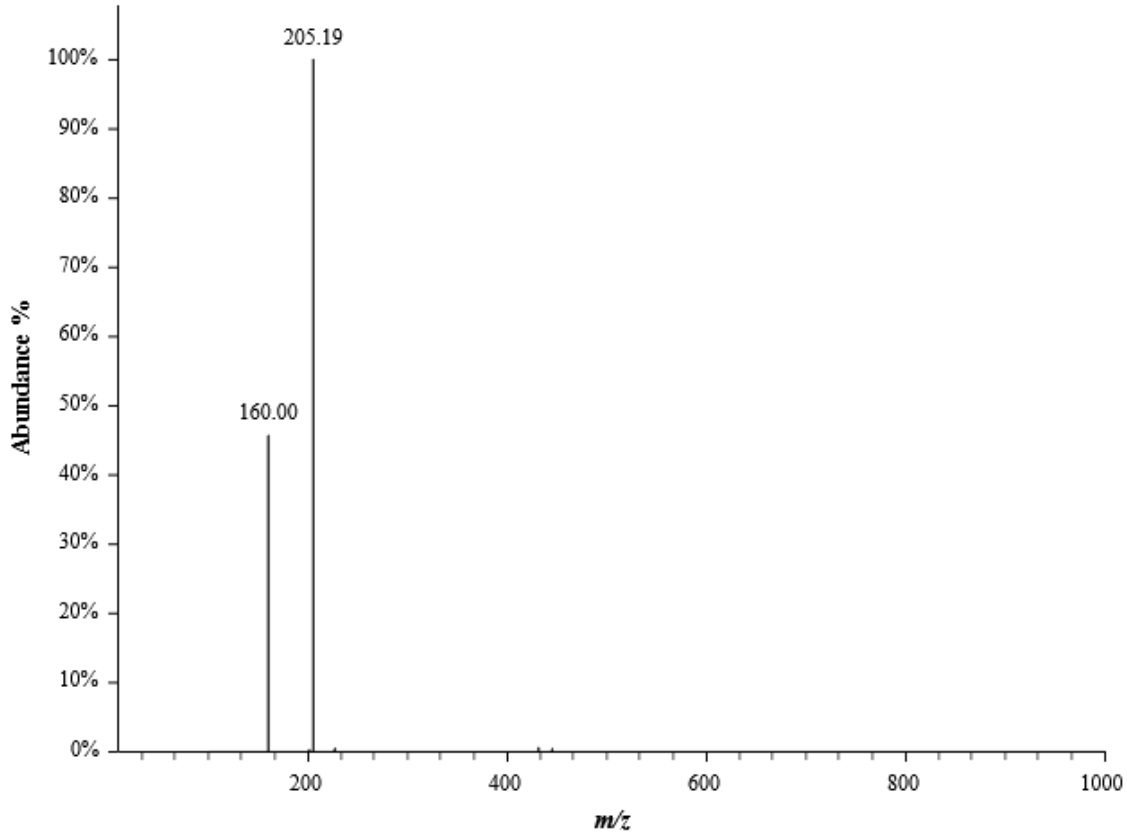
B.4.k. 4-Methoxygramine (**4k**)

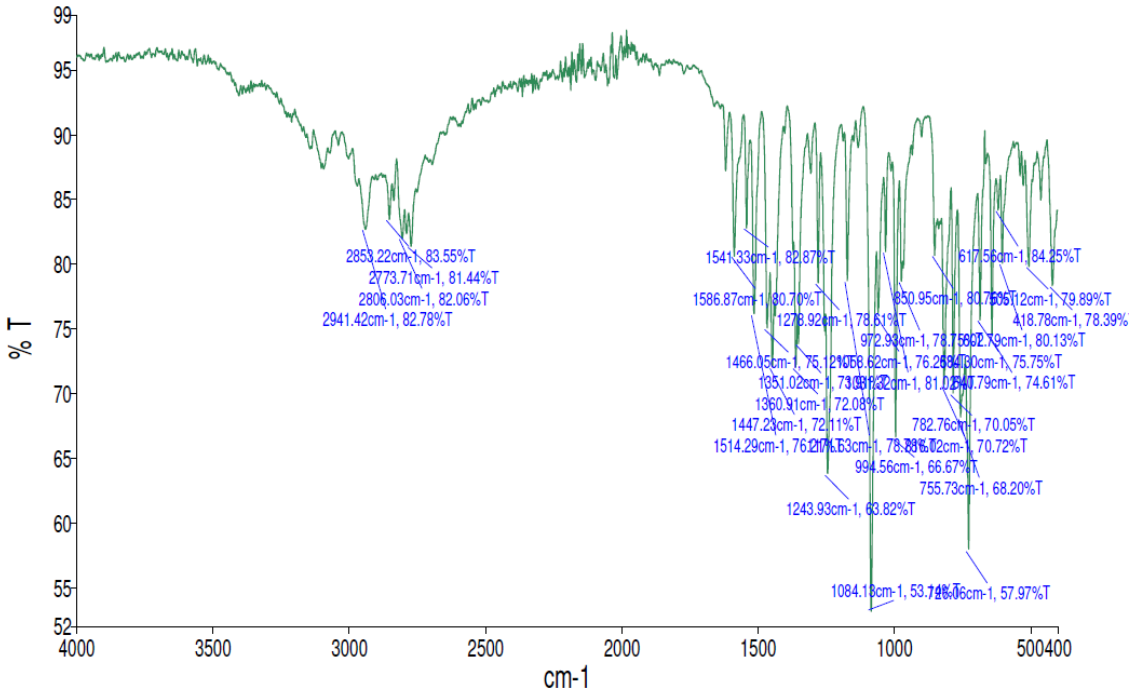




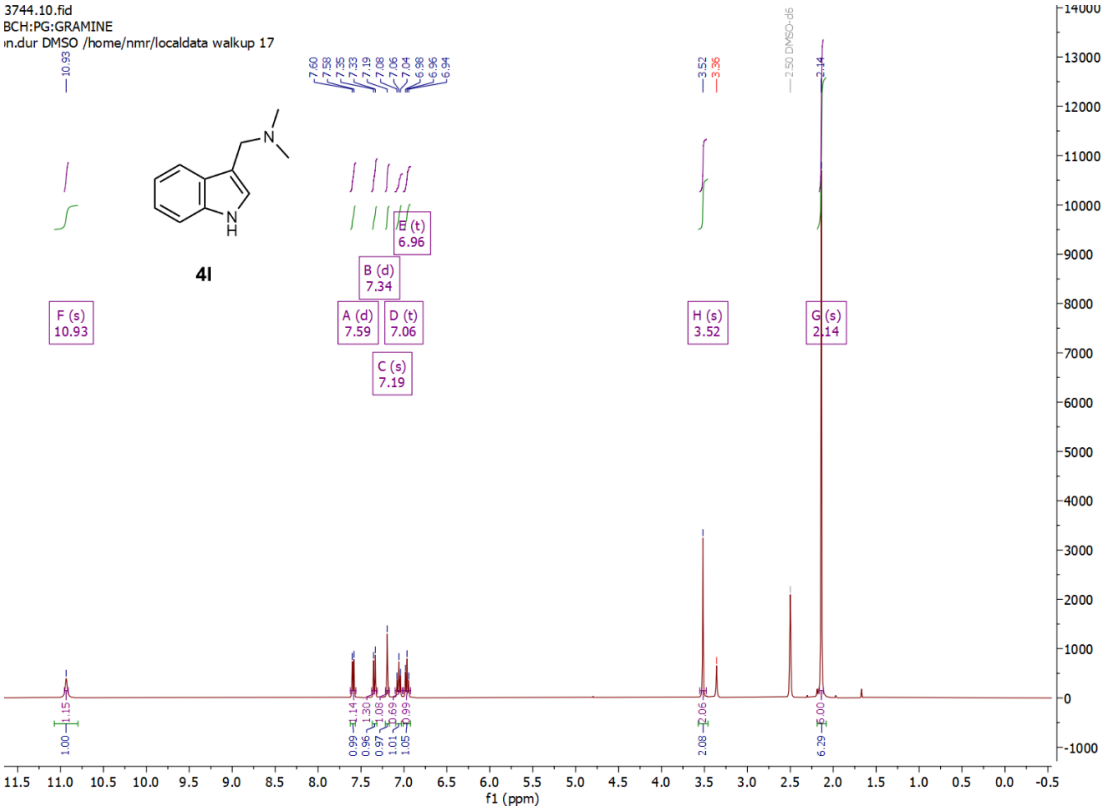
BCRH_4_MeO_gramine_222797

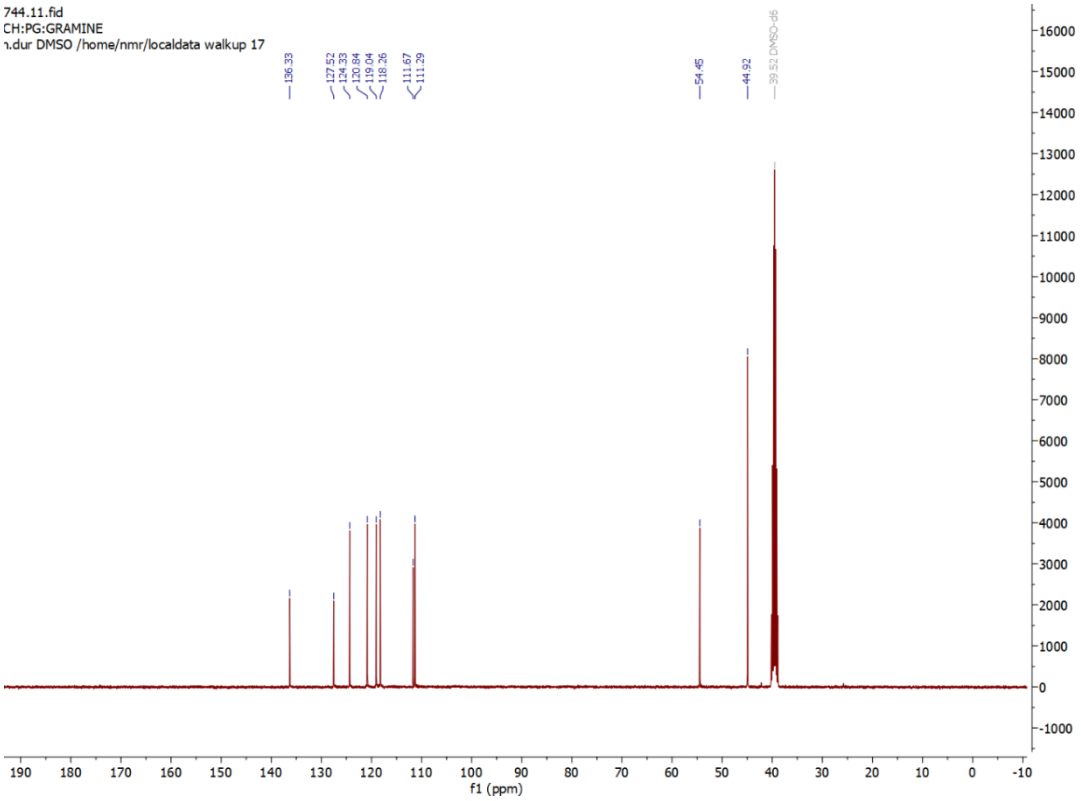
RT: 0.9782 minutes, Scan 107, 1: MS ES+ c (100.0-2000.0), NL 2.85e+8





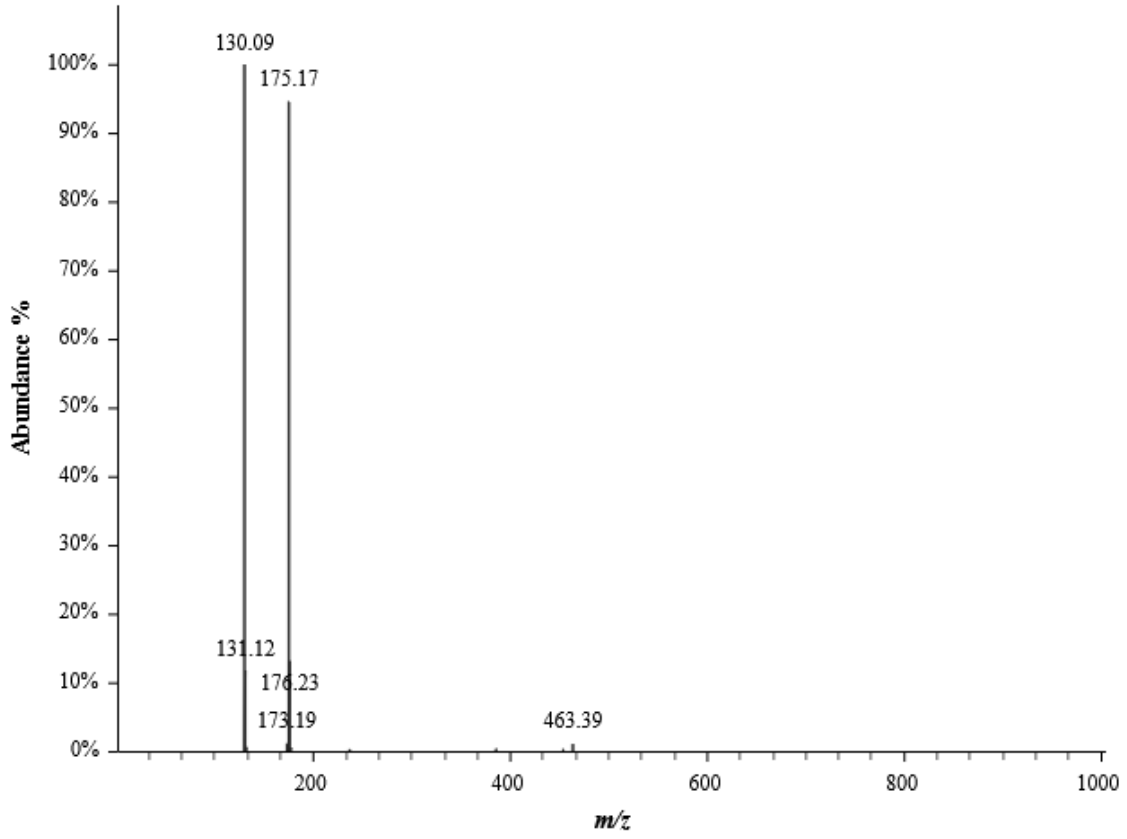
B.4.l. Gramine (**4l**)





BCRH_Gramine_211805

RT: 0.7699 minutes, Scan 83, 1: MS ES+ c (100.0-2000.0), NL 9.85e+7

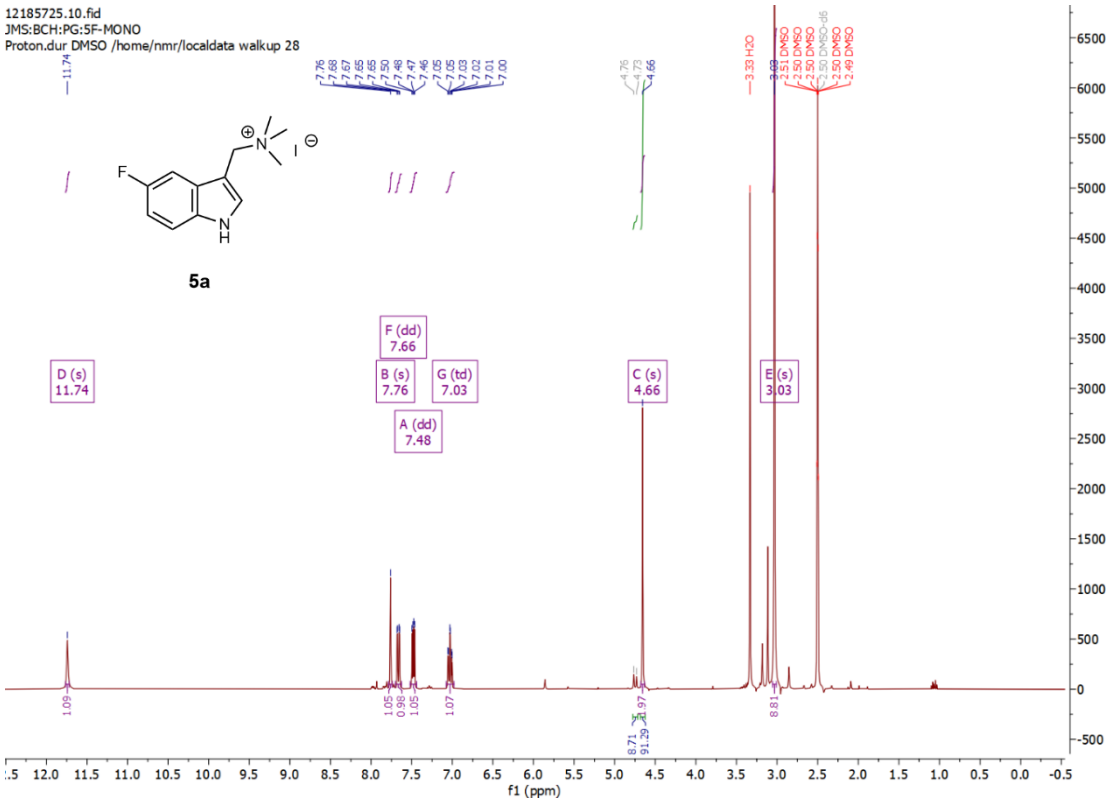


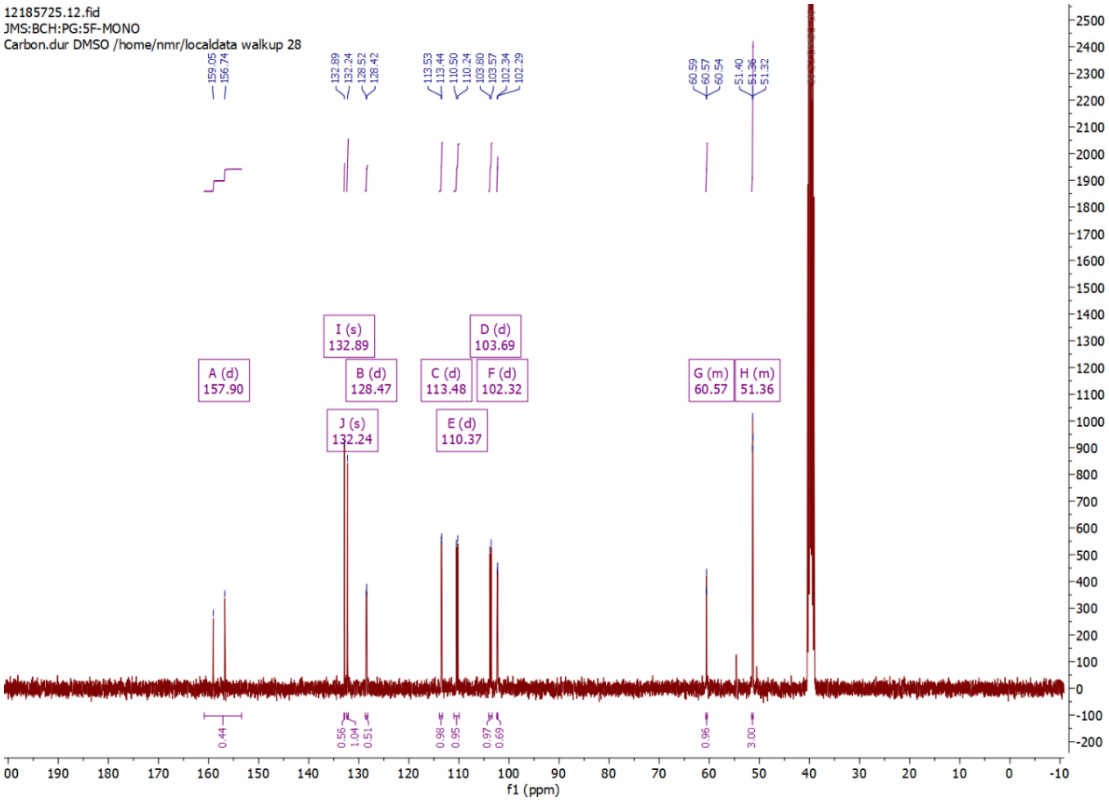
Crystal data and structure refinement for Compound 4l.	
Empirical formula	$C_{11}H_{14}N_2$
Formula weight	174.24
Temperature/K	120.00
Crystal system	orthorhombic
Space group	Pbca
a/Å	10.740(2)
b/Å	9.161(2)
c/Å	19.784(4)
α/°	90
β/°	90
γ/°	90
Volume/Å^3	1946.4(7)
Z	8
ρ_{calc}g/cm^3	1.189
μ/mm^{-1}	0.072
F(000)	752.0
Crystal size/mm^3	0.521 × 0.1 × 0.054
Radiation	MoKα (λ = 0.71073)
2Θ range for data collection/°	5.598 to 49.976

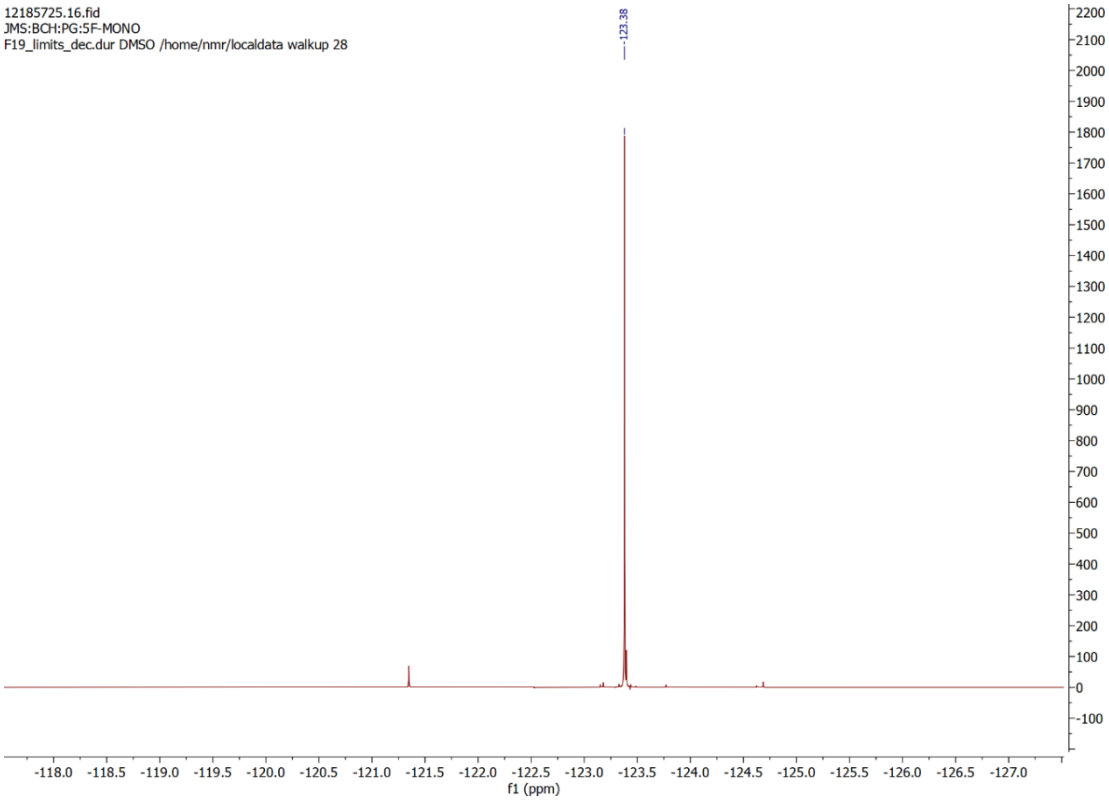
Index ranges	-11 ≤ h ≤ 12, -10 ≤ k ≤ 10, -21 ≤ l ≤ 23
Reflections collected	14738
Independent reflections	1713 [R_{int} = 0.1342, R_{sigma} = 0.0763]
Data/restraints/parameters	1713/0/124
Goodness-of-fit on F^2	1.103
Final R indexes [I>=2σ (I)]	R_1 = 0.1033, wR_2 = 0.2055
Final R indexes [all data]	R_1 = 0.1383, wR_2 = 0.2236
Largest diff. peak/hole / e $Å^{-3}$	0.26/-0.37

B.5. 1-(1H-indol-3-yl) trimethyl ammonium iodide analogues (Monomers)

B.5.a. (5-Fluoro-3-indolylmethyl)trimethylammonium iodide **(5a)**

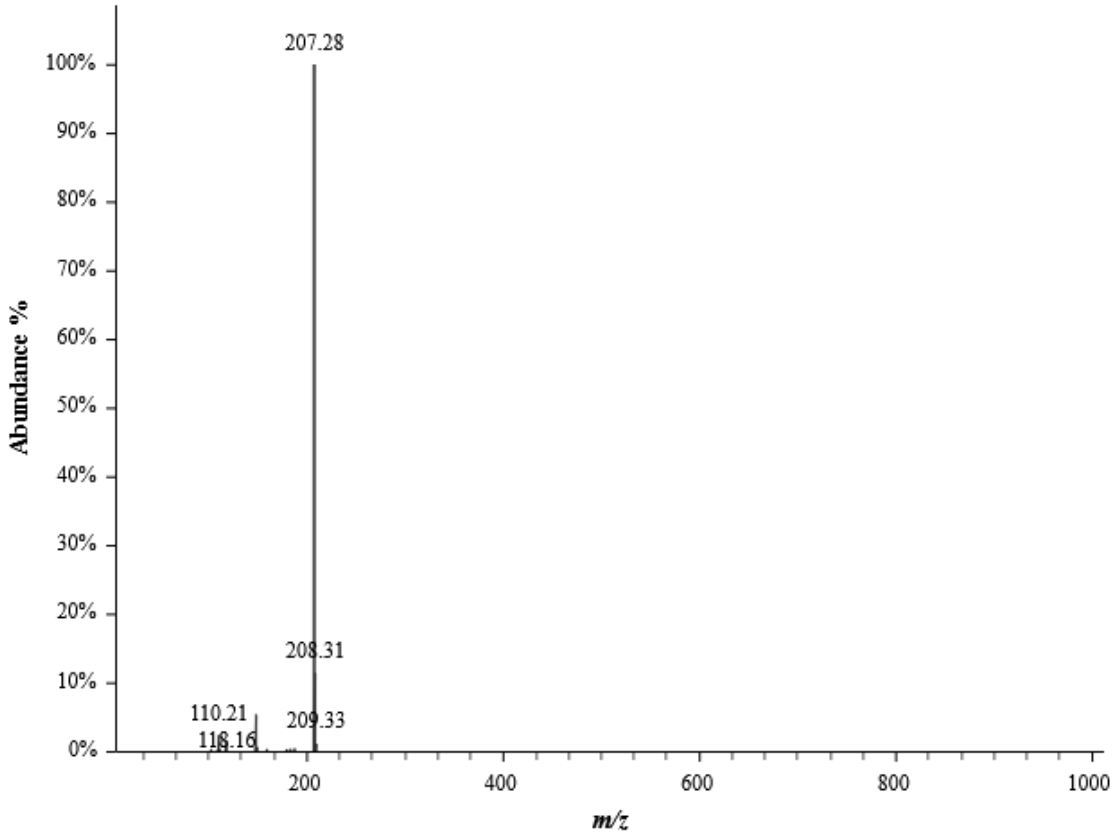


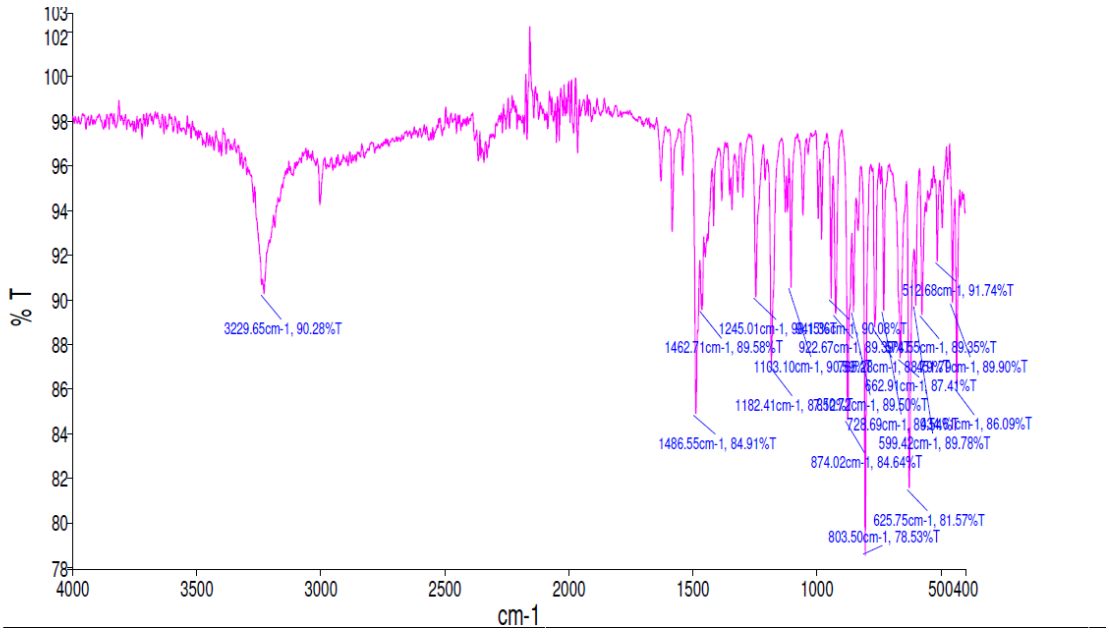

12185725.12.fid
JMS:BCH:PG:5F-MONO
Carbon.dur DMSO /home/nmr/localdata walkup 28
A (d) 157.90
B (d) 128.47
C (d) 113.48
D (d) 103.69
E (d) 110.37
F (d) 102.32
G (m) 60.57
H (m) 51.36
I (s) 132.89
J (s) 132.24
f1 (ppm)


12185725.16.fid
JMS:BCH:PG:5F-MONO
F19_limits_dec.dur DMSO /home/nmr/localdata walkup 28
-123.38
f1 (ppm)

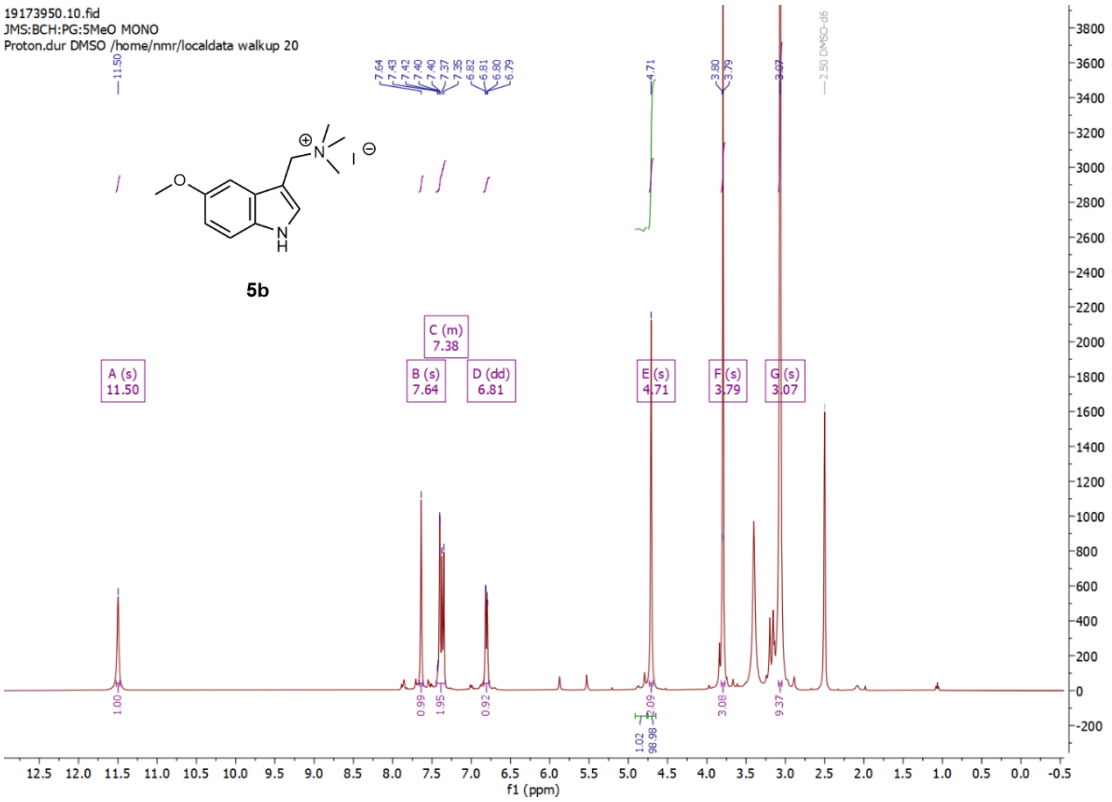
BCRH_5FMono_205609

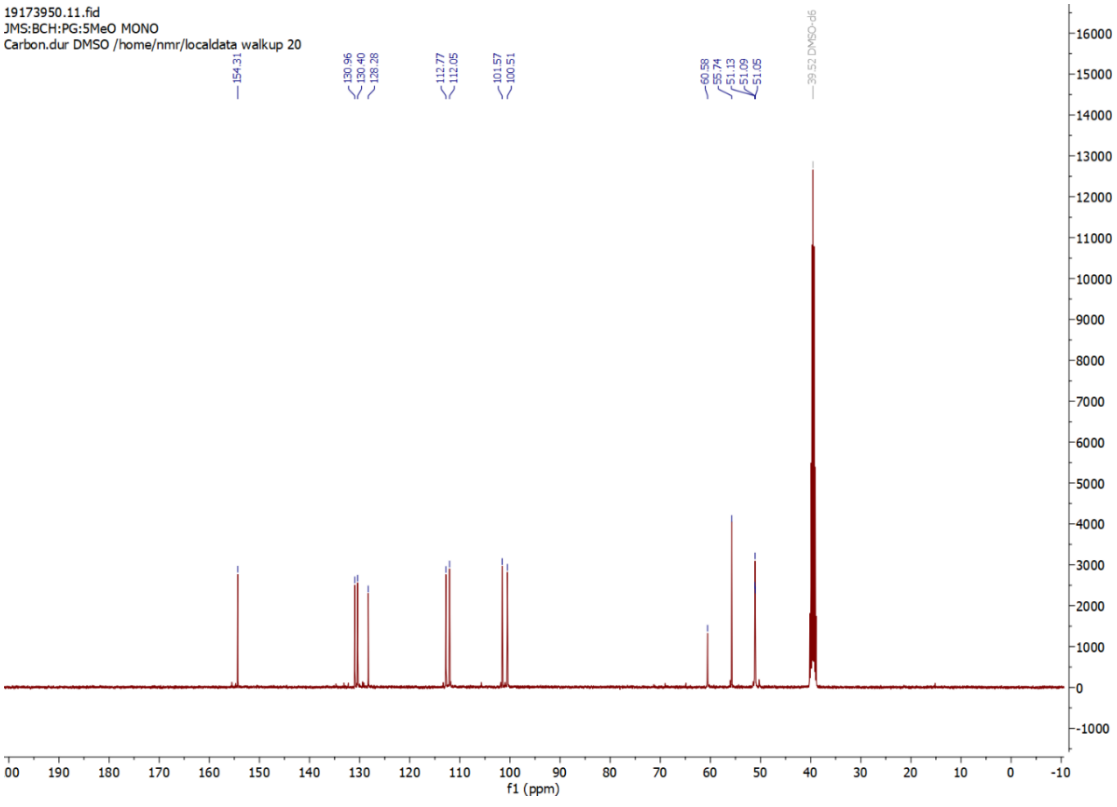
RT: 0.9262 minutes, Scan 101, 1: MS ES+ c (100.0-2000.0), NL 4.37e+7





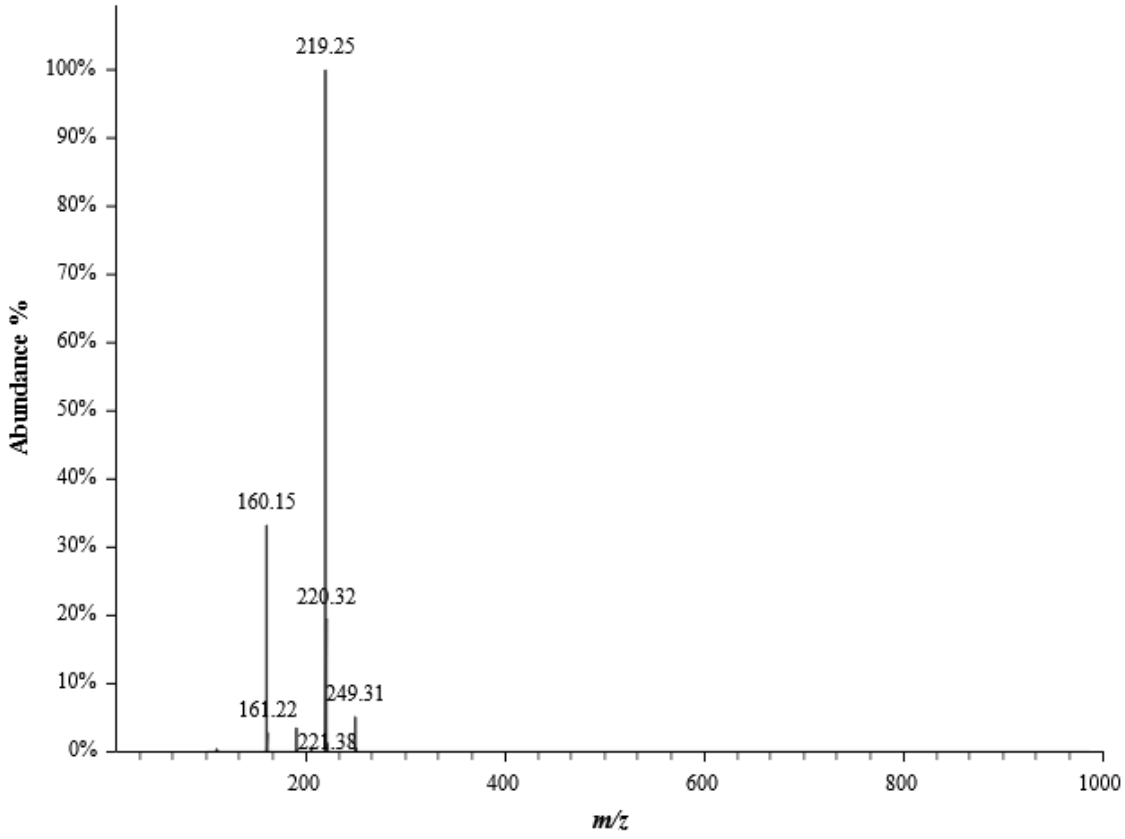
B.5.b. (5-Methoxy-3-indolylmethyl)trimethylammonium iodide (**5b**)

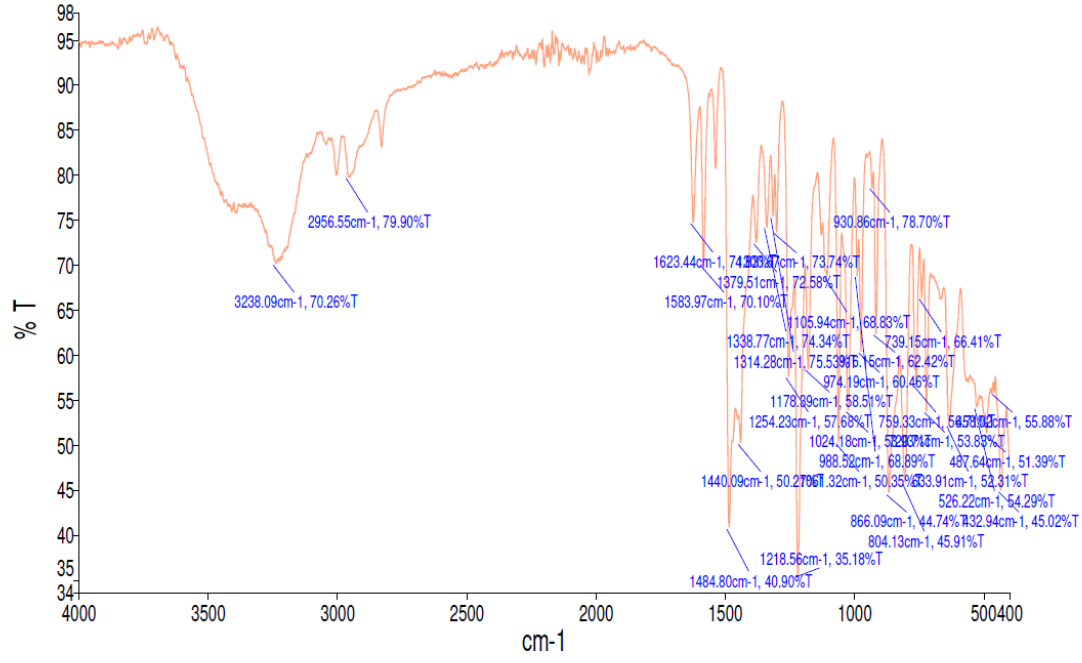




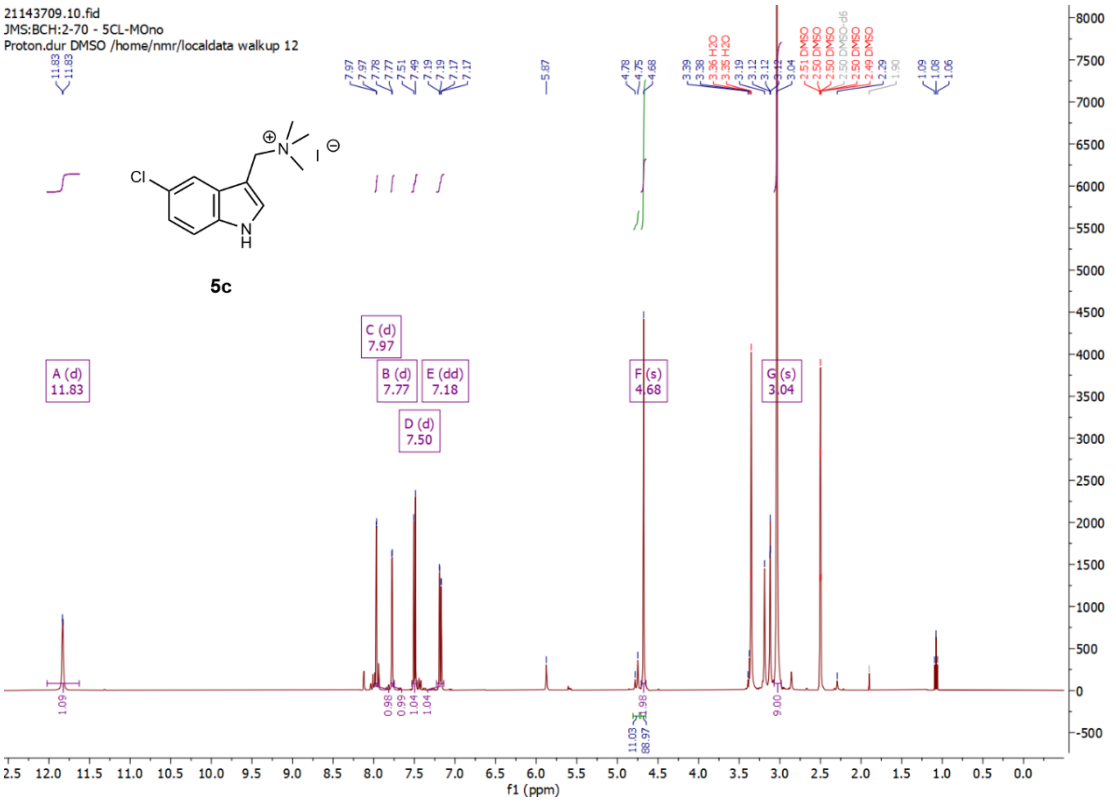
BCRH_5MeOMonomer_210475

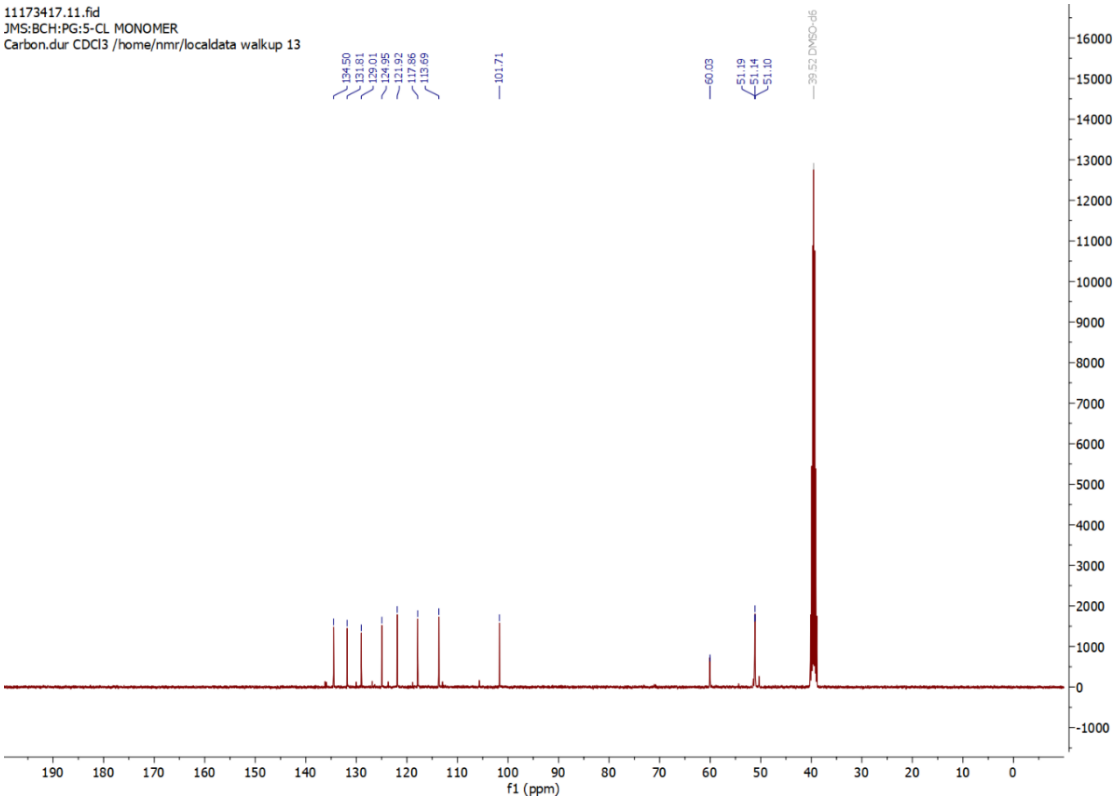
RT: 0.8741 minutes, Scan 95, 1: MS ES+ c (100.0-2000.0), NL 9.31e+7





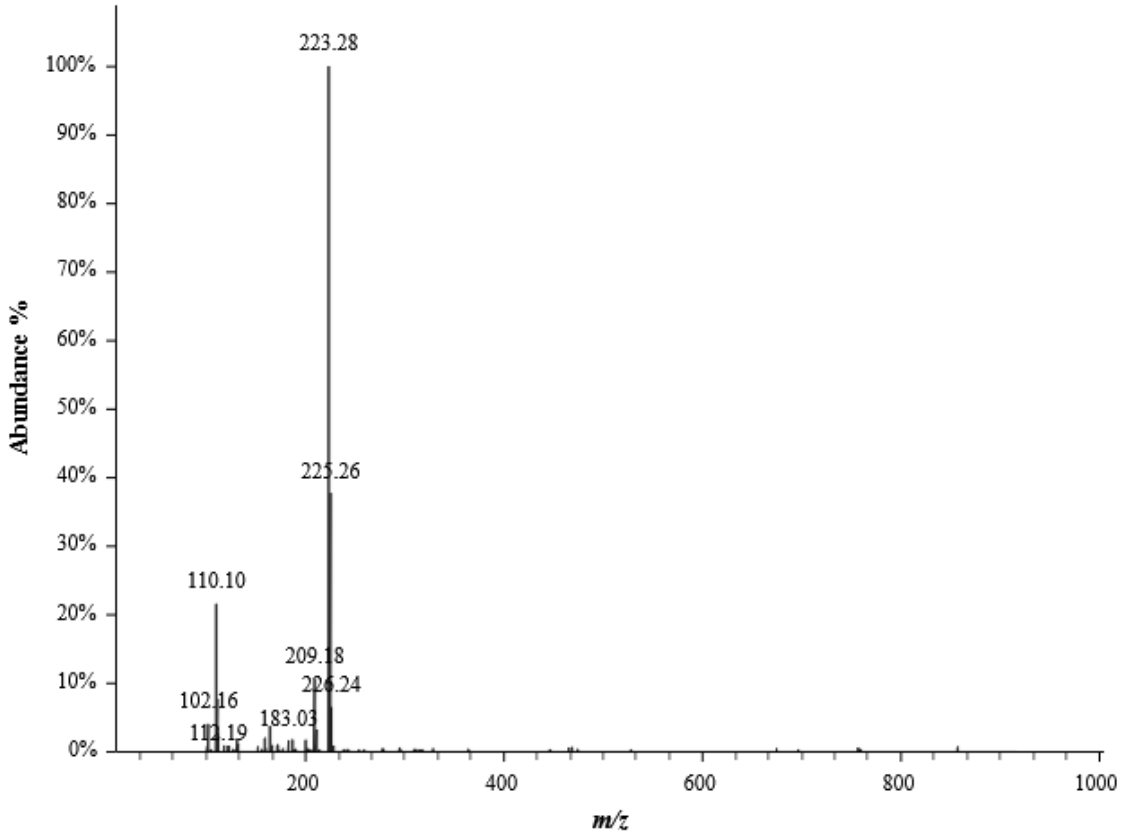
B.5.c. (5-chloro-3-indolylmethyl)trimethylammonium iodide (**5c**)

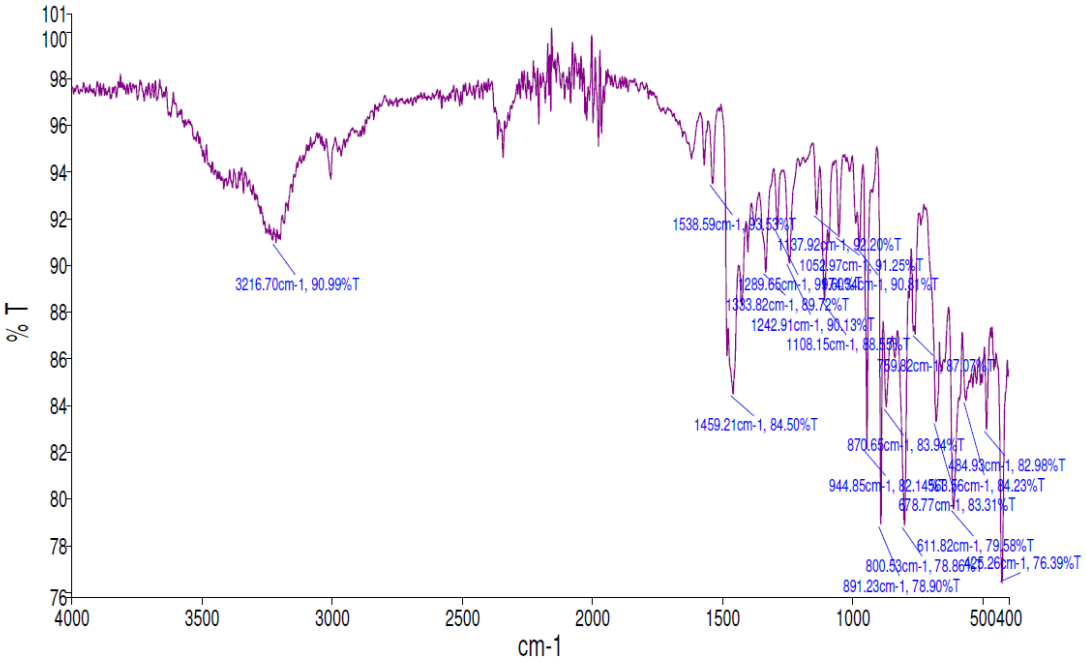




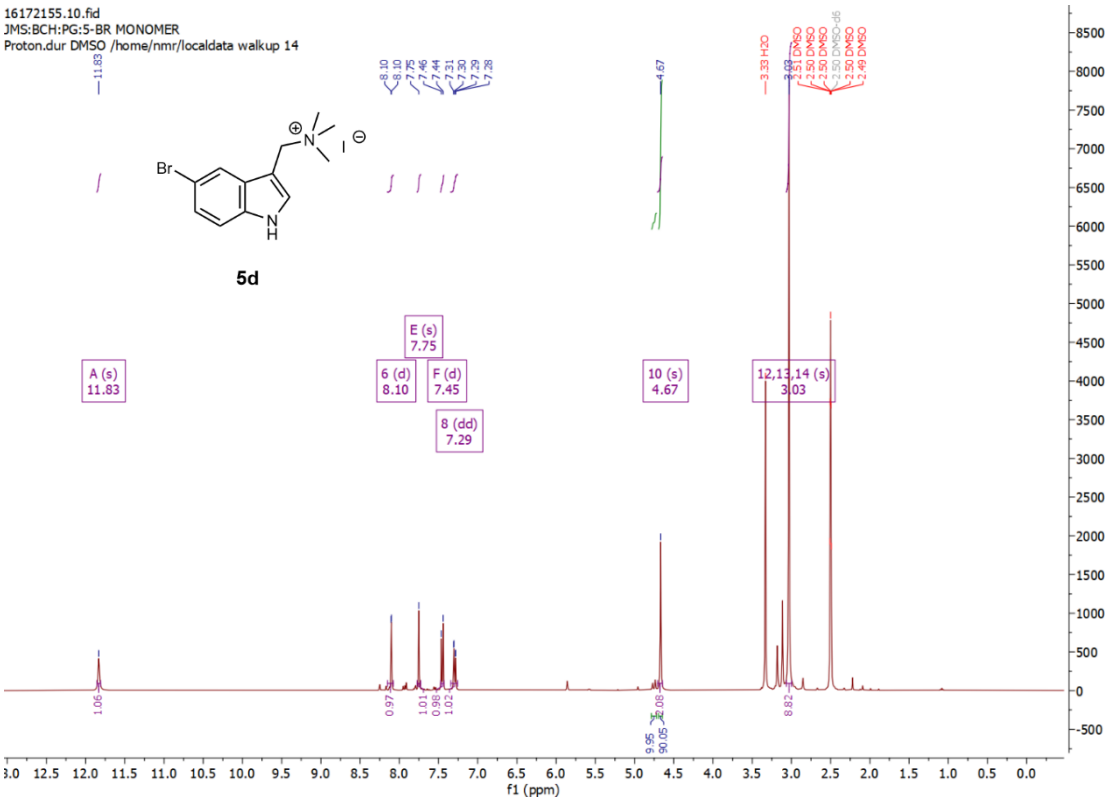
BCRH_5Clmono_206792

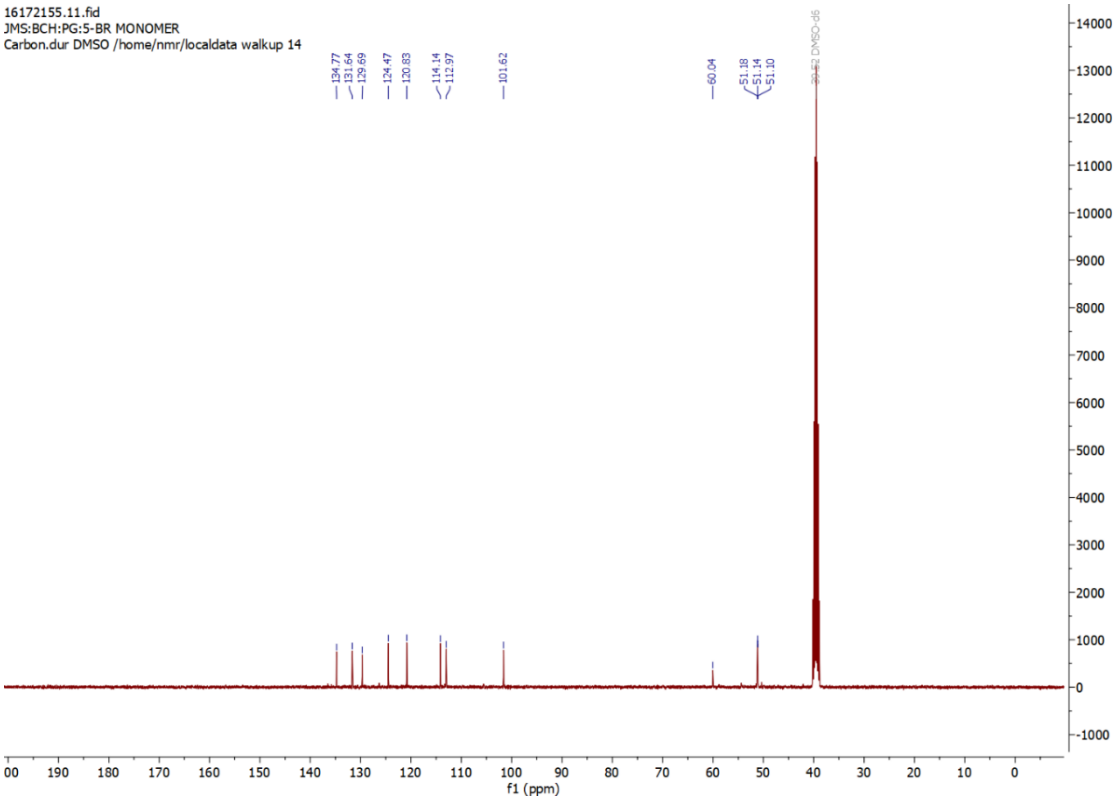
RT: 1.1691 minutes, Scan 129, 1: MS ES+ c (100.0-2000.0), NL 1.21e+7





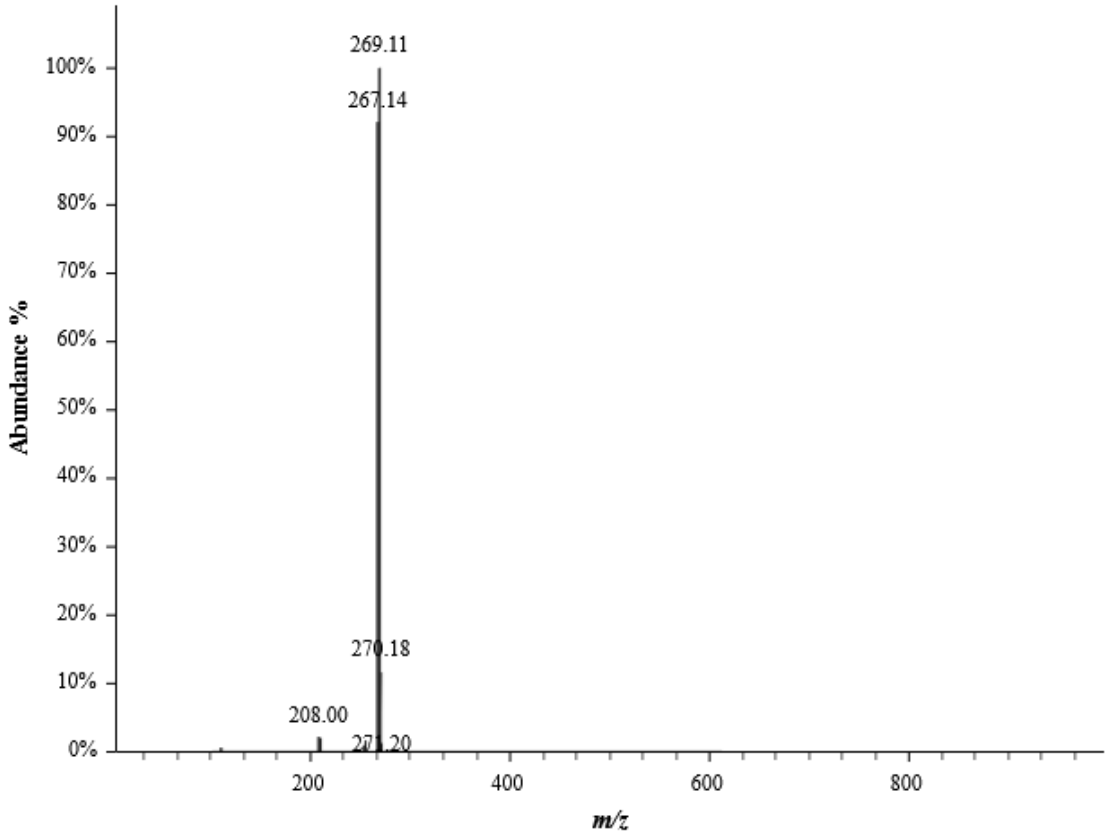
B.5.d. (5-Bromo-3-indolylmethyl)trimethylammonium iodide (**5d**)

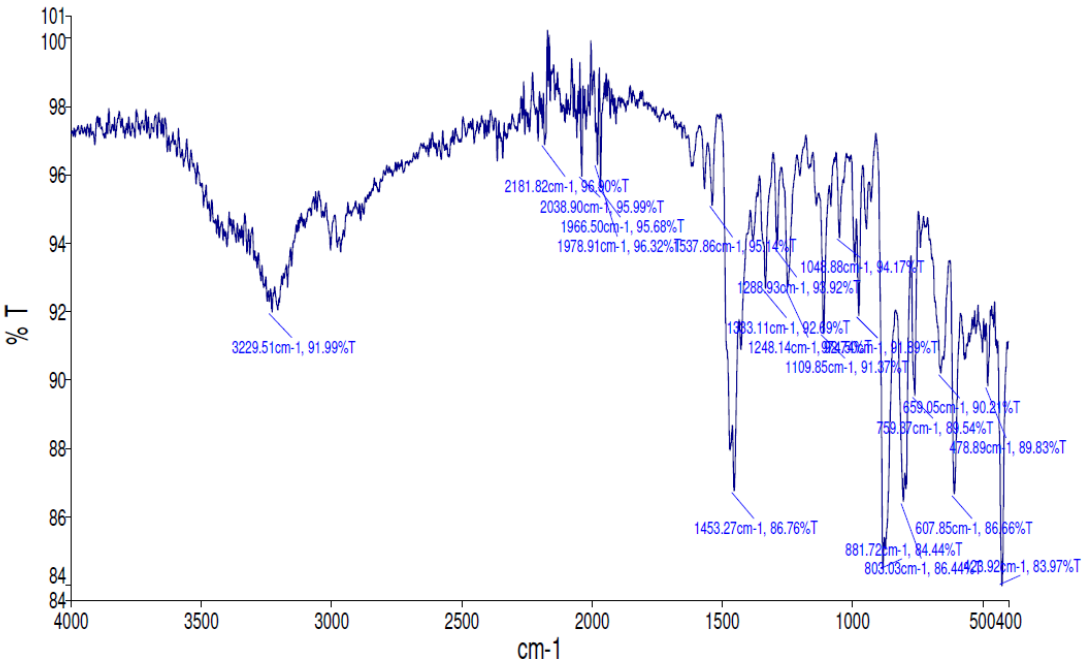




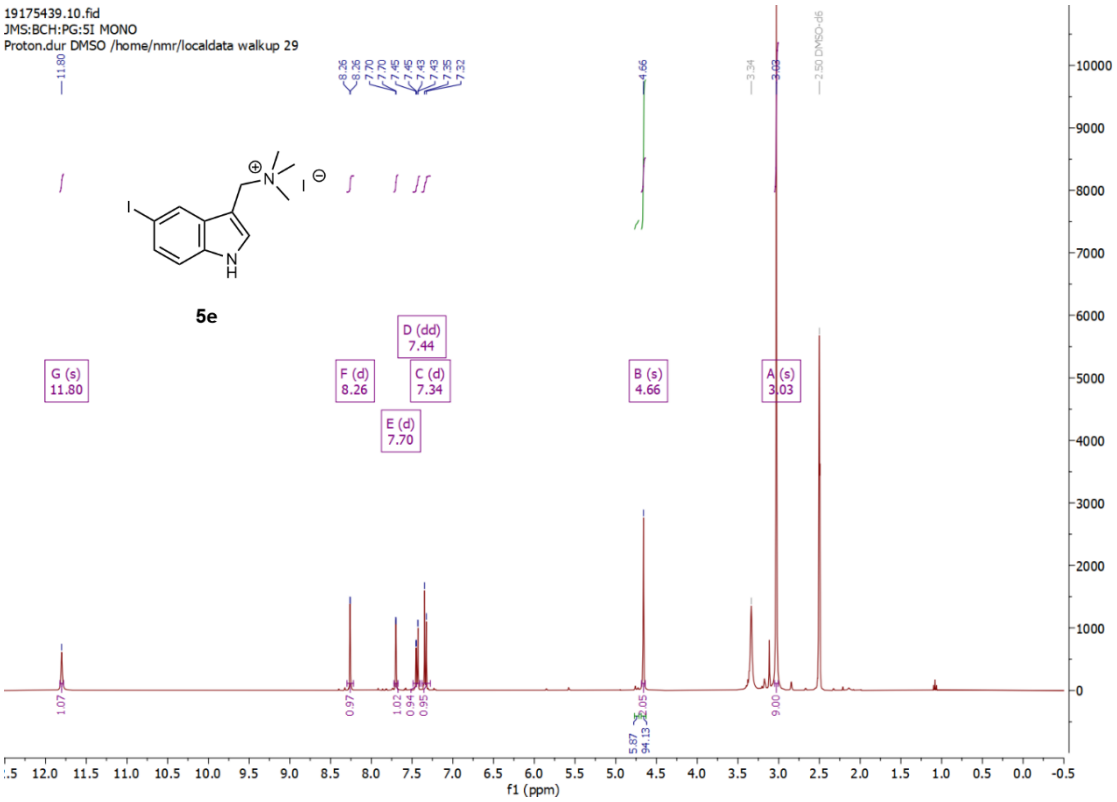
BCRH_5Br_mono_211786

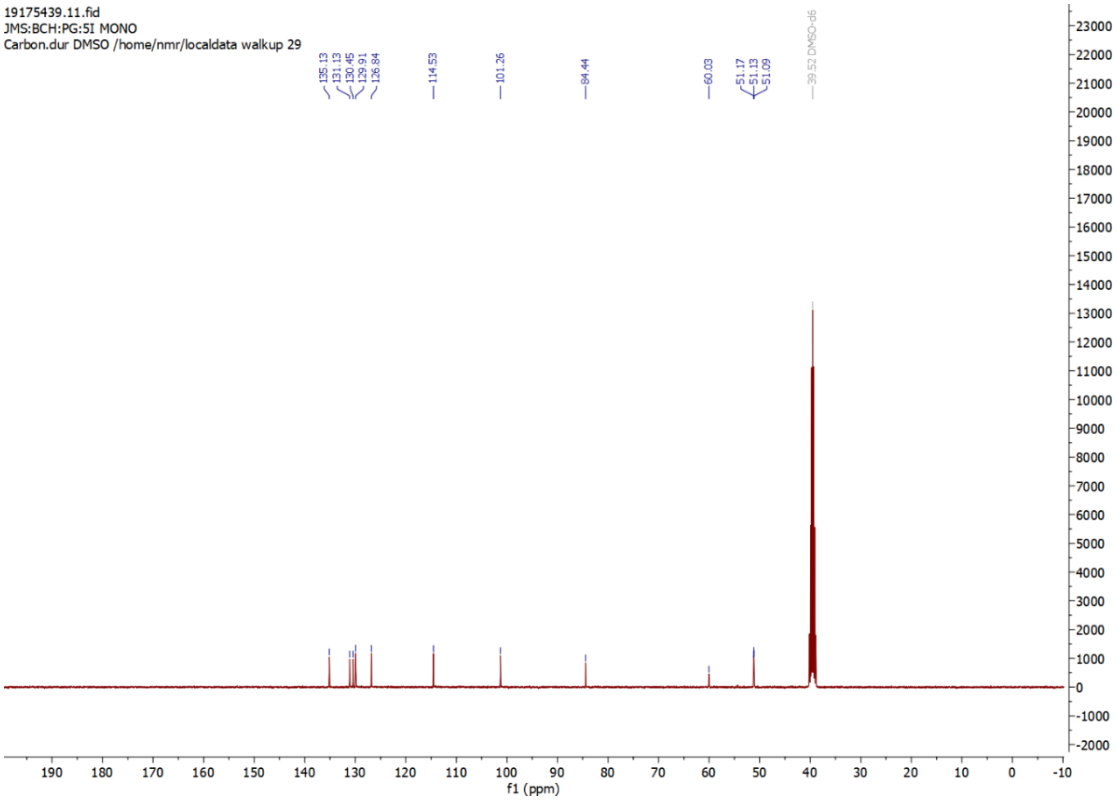
RT: 1.2039 minutes, Scan 133, 1: MS ES+ c (100.0-2000.0), NL 1.05e+8





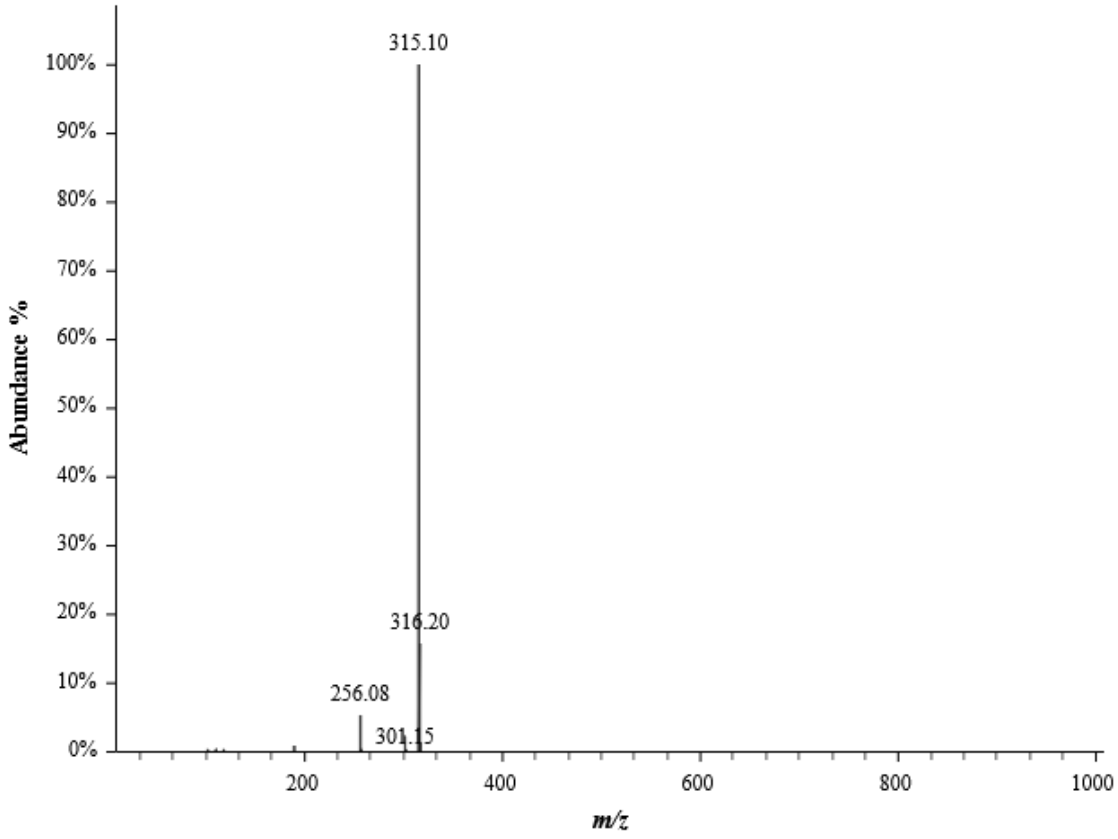
B.5.e. (5-Iodo-3-indolylmethyl)trimethylammonium iodide (**5e**)

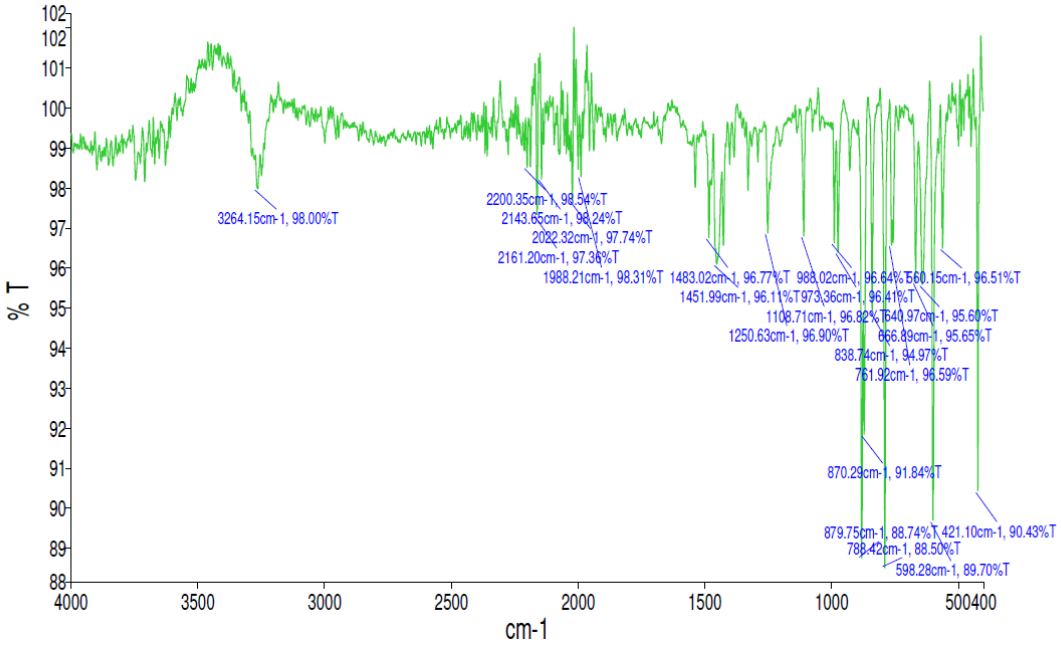




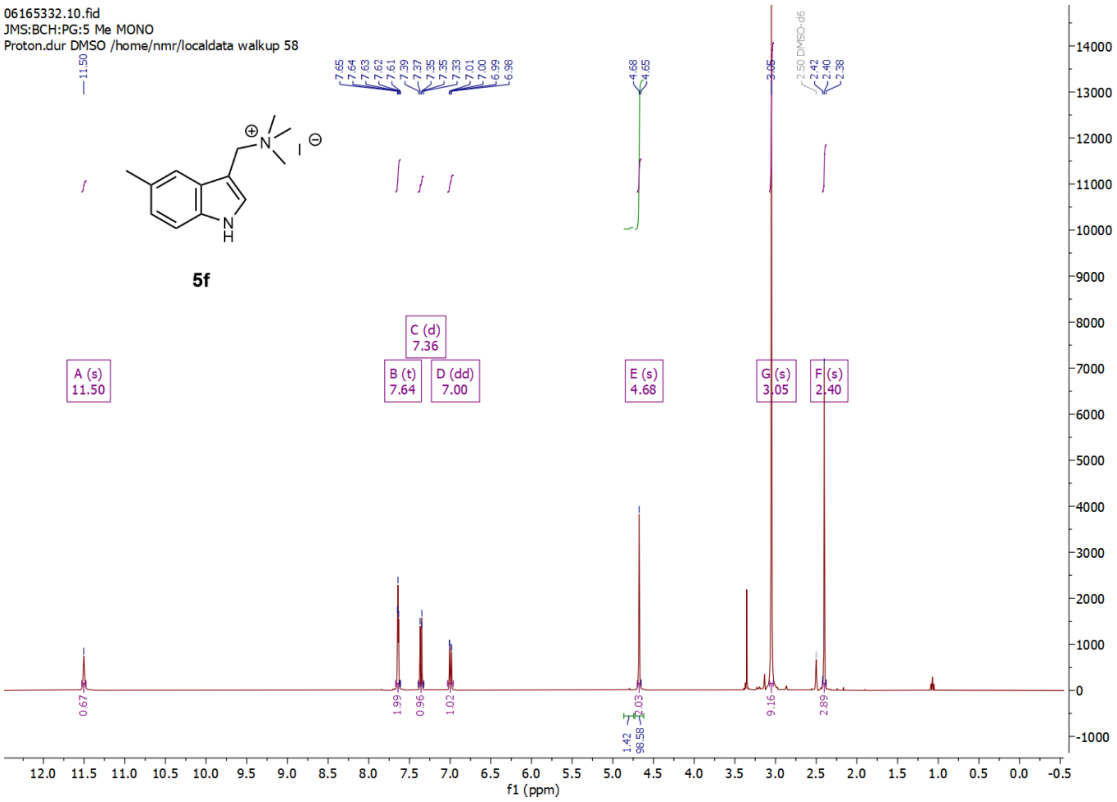
BCRH_5Imonomer_210473

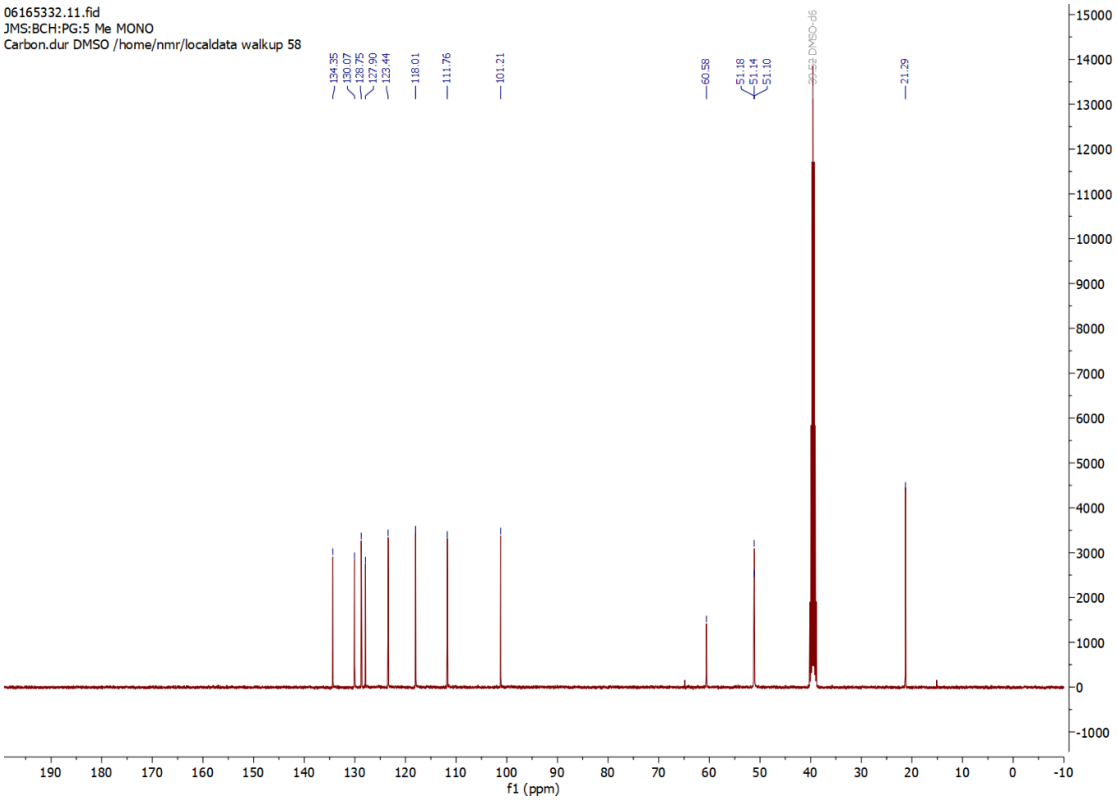
RT: 1.3080 minutes, Scan 145, 1: MS ES+ c (100.0-2000.0), NL 1.19e+8





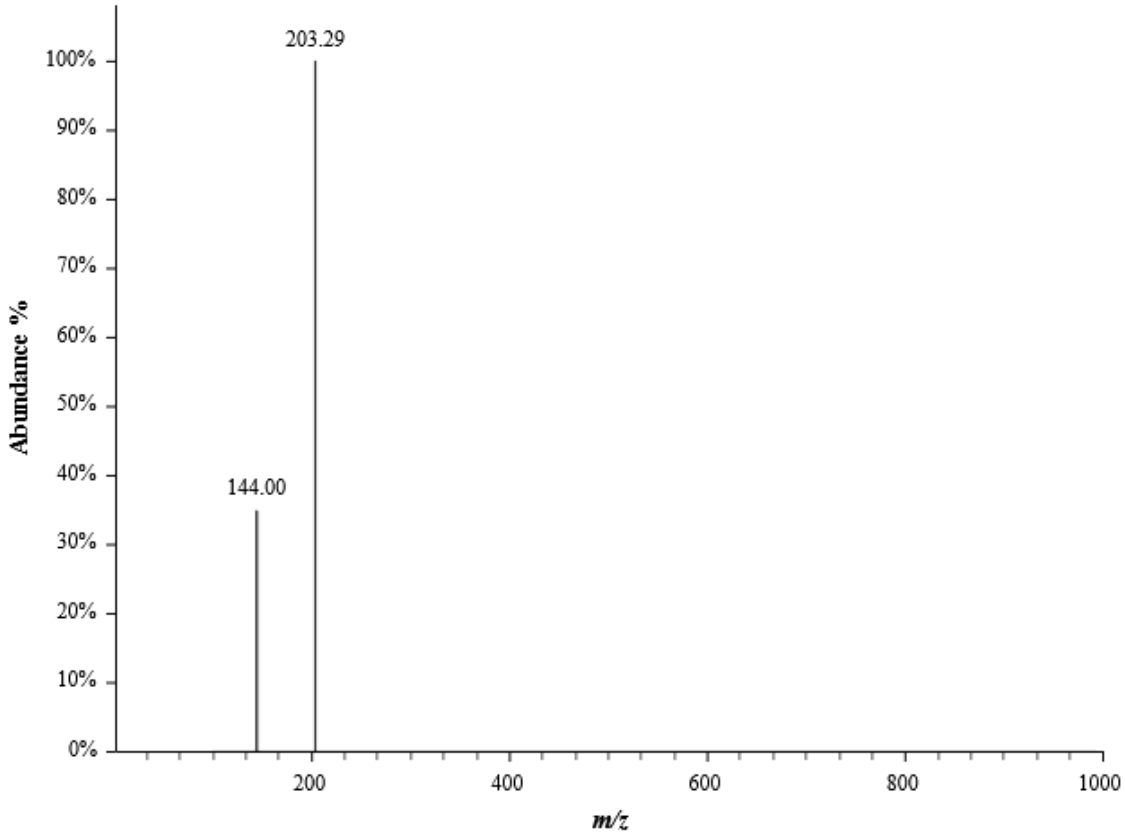
B.5.f. (5-Methyl-3-indolylmethyl)trimethylammonium iodide (**5f**)

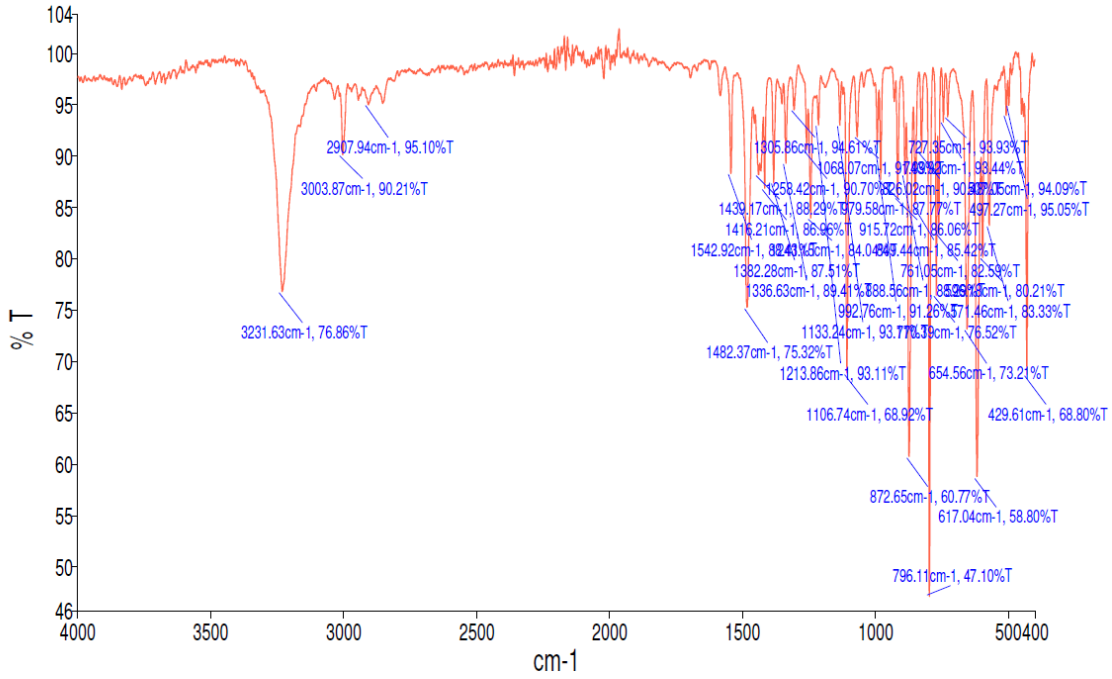




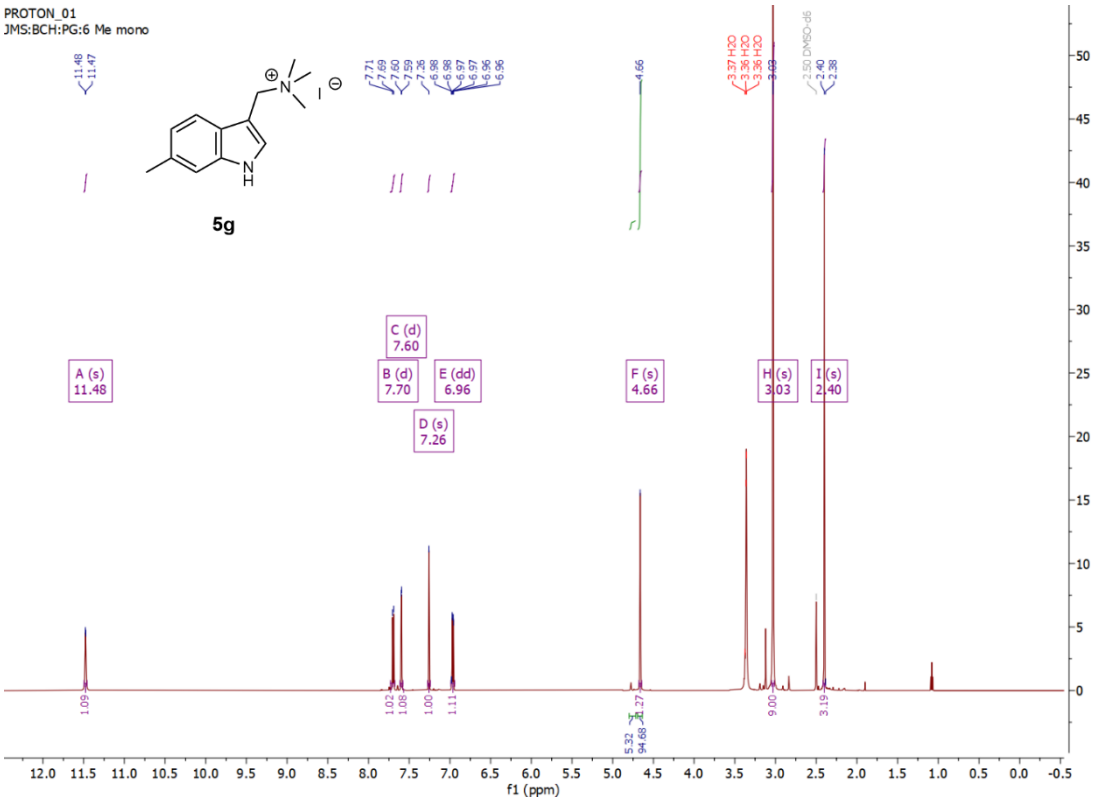
BCRH_5_Me_mono_222793

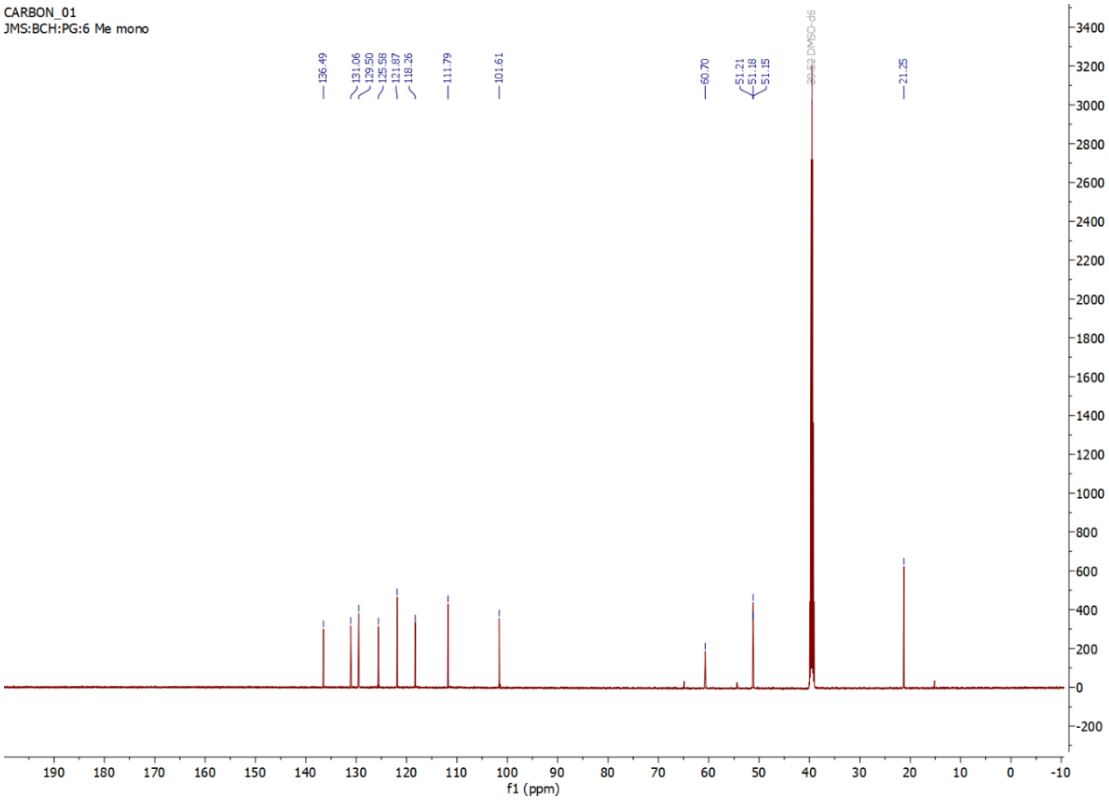
RT: 0.9956 minutes, Scan 109, 1: MS ES+ c (100.0-2000.0), NL 2.76e+8





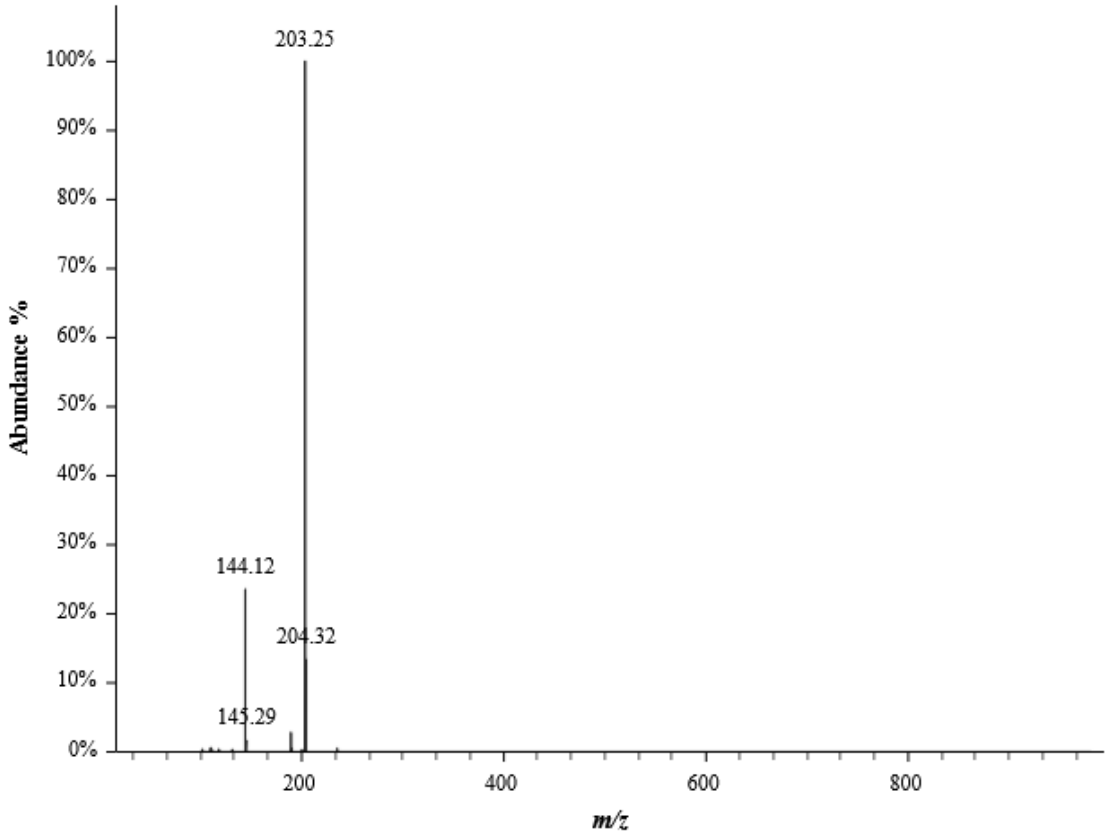
B.5.g. (6-Methyl-3-indolylmethyl)trimethylammonium iodide (**5g**)

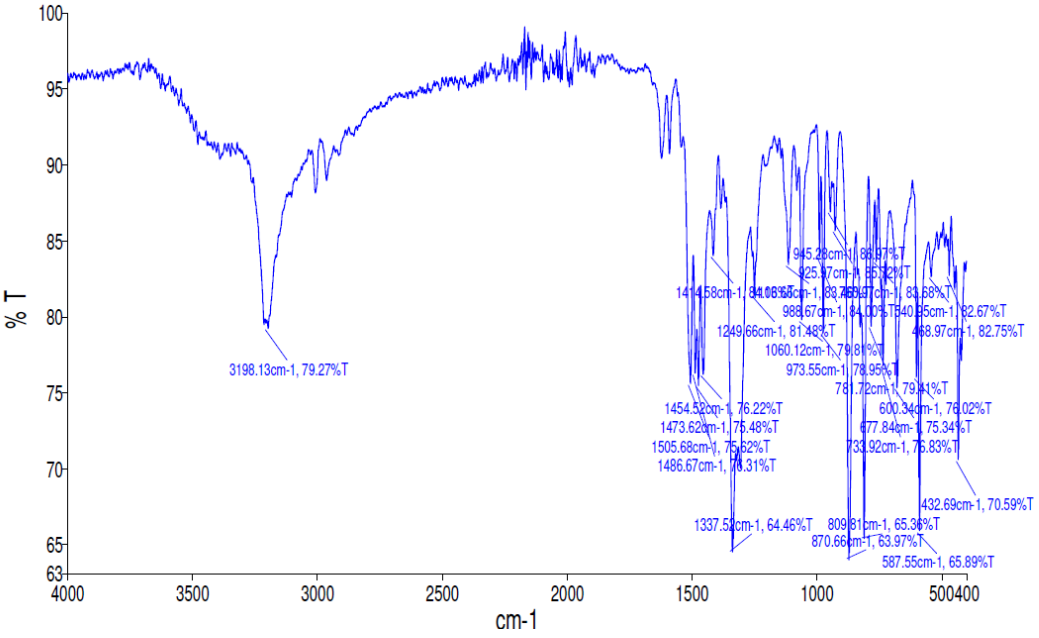




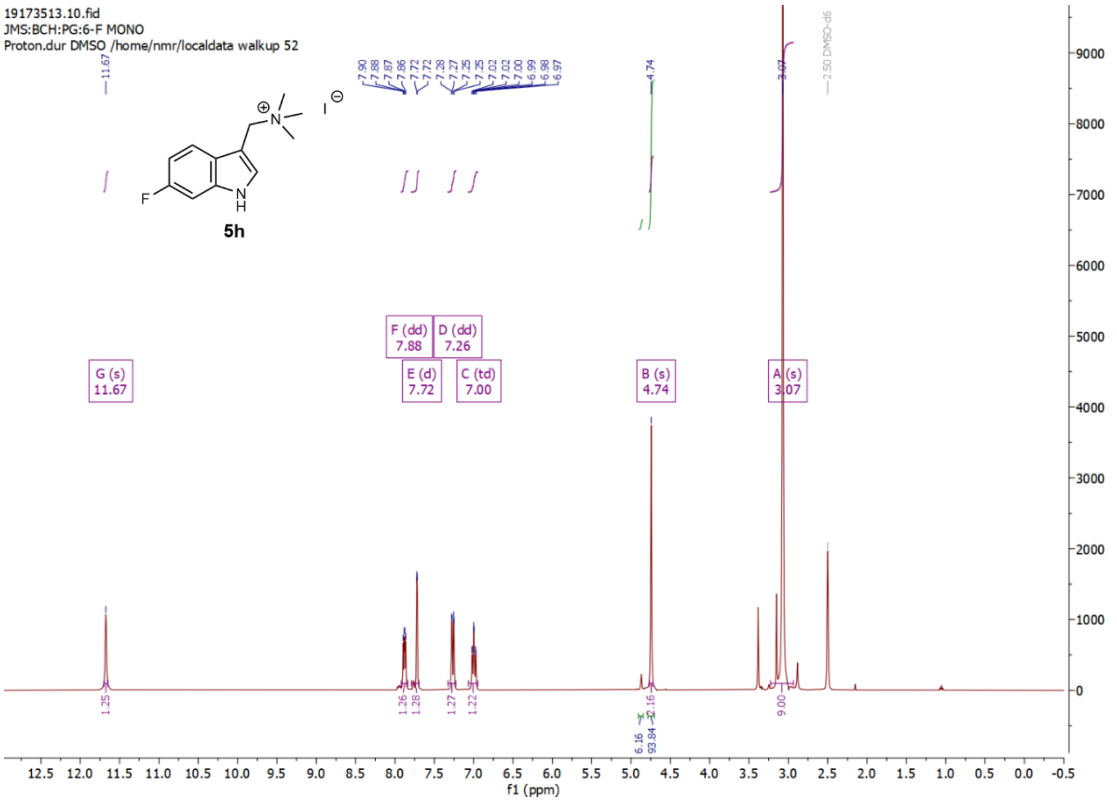
BCRH_6MeMonomer_210505

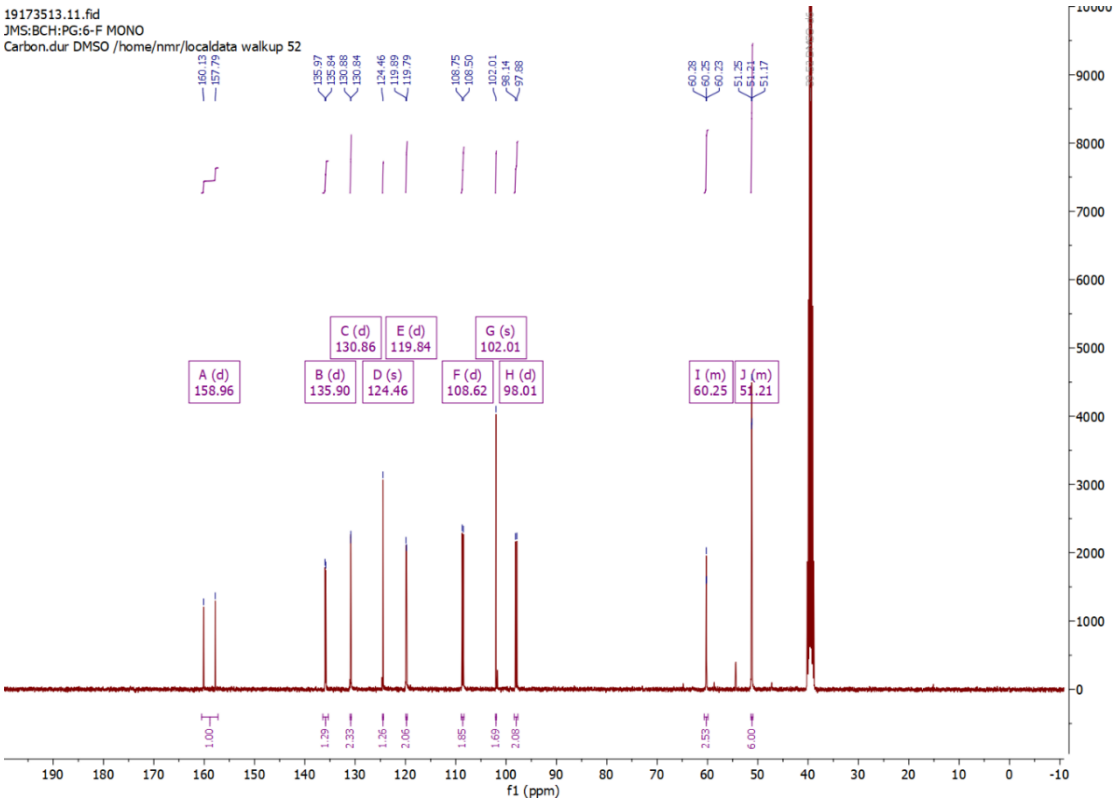
RT: 1.0824 minutes, Scan 119, 1: MS ES+ c (100.0-2000.0), NL 7.18e+7

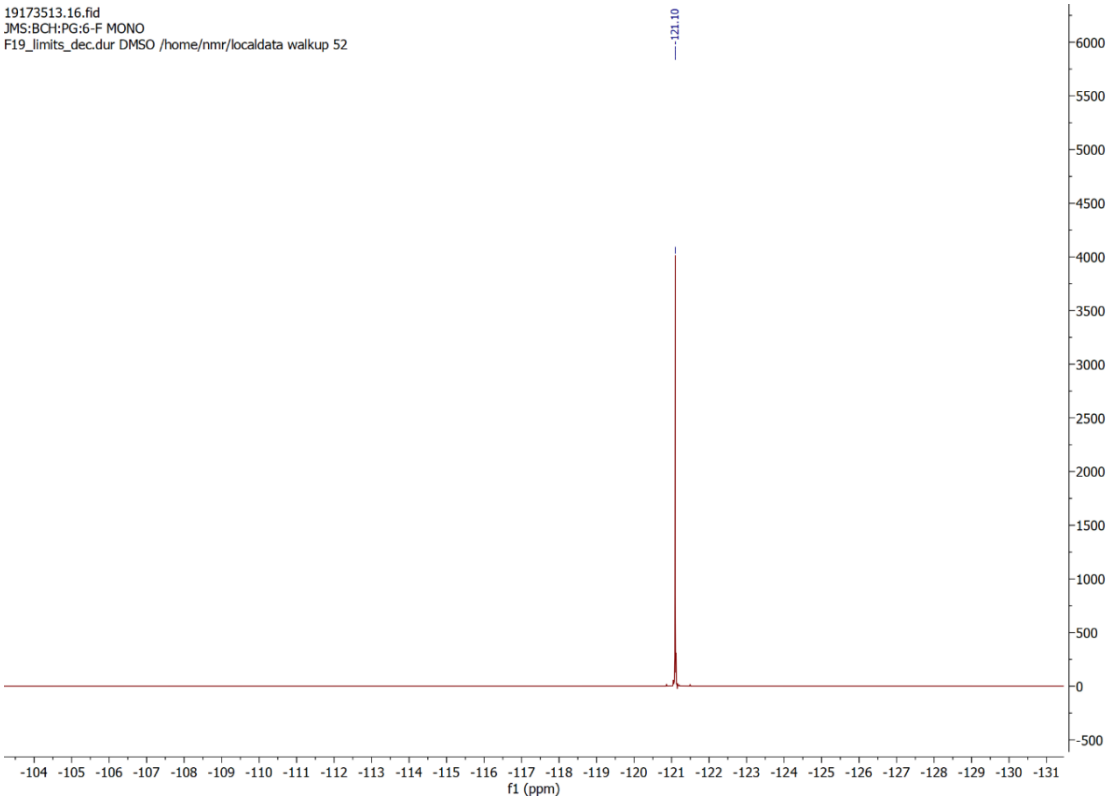




B.5.h. (6-Fluoro-3-indolylmethyl)trimethylammonium iodide (**5h**)

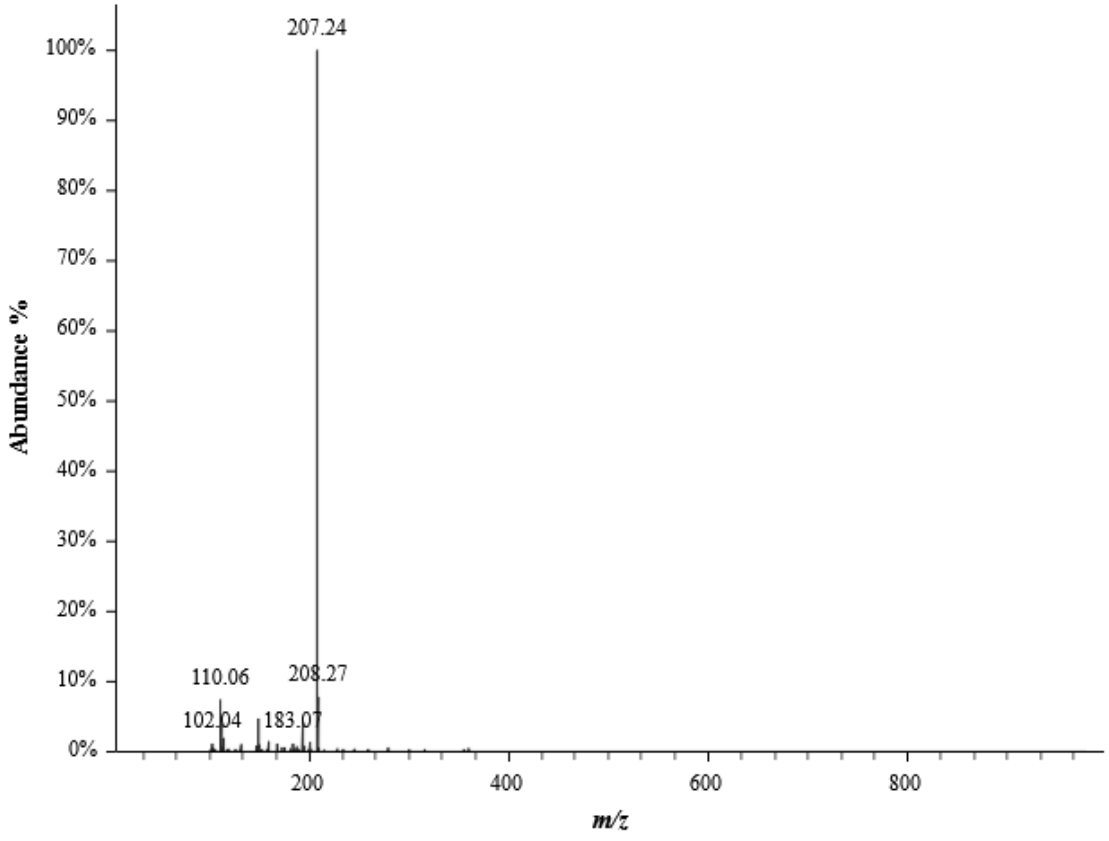


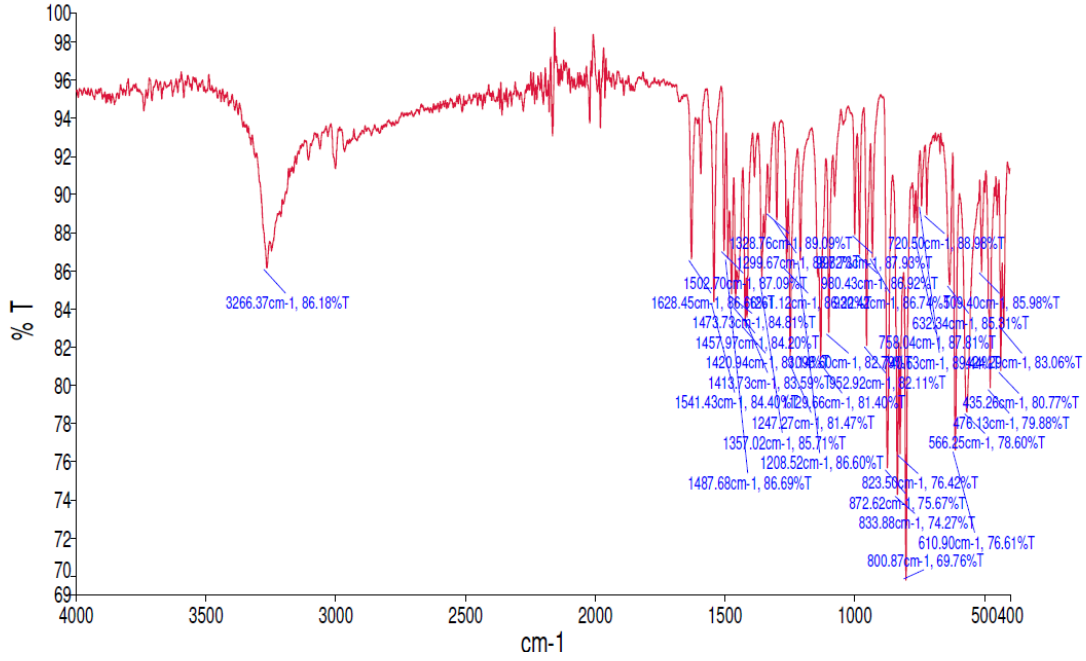




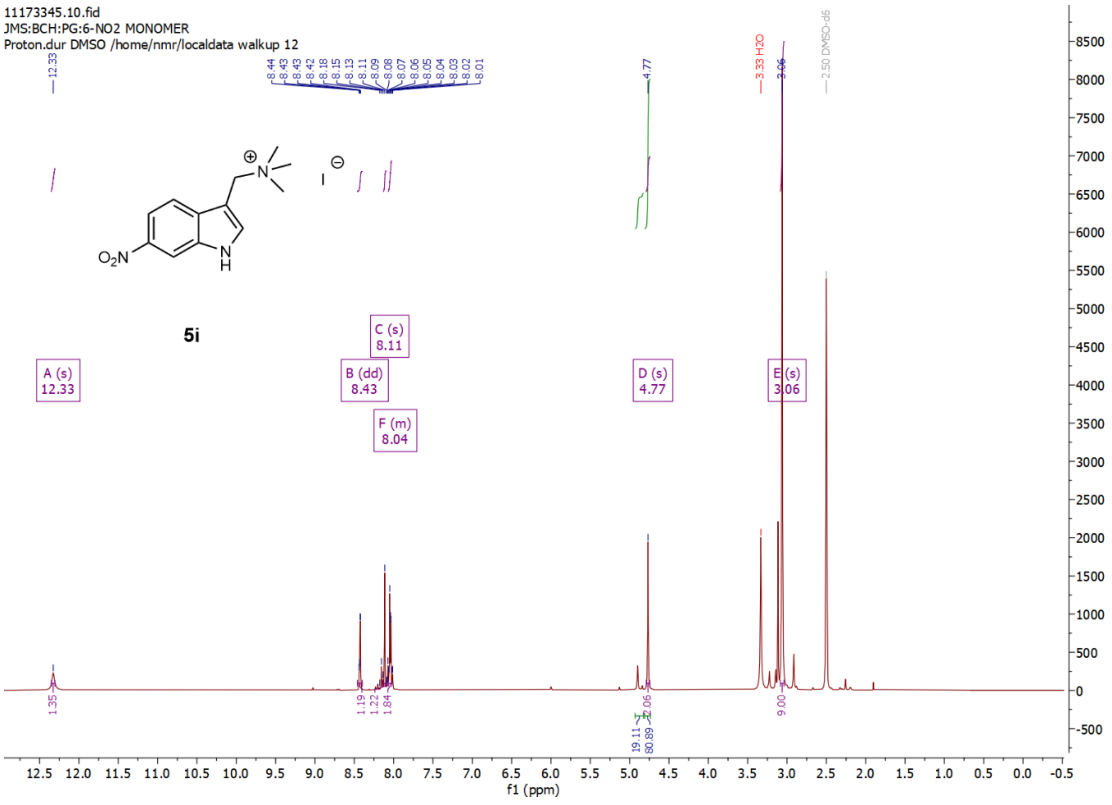
BCRH_6FMonomer_210504

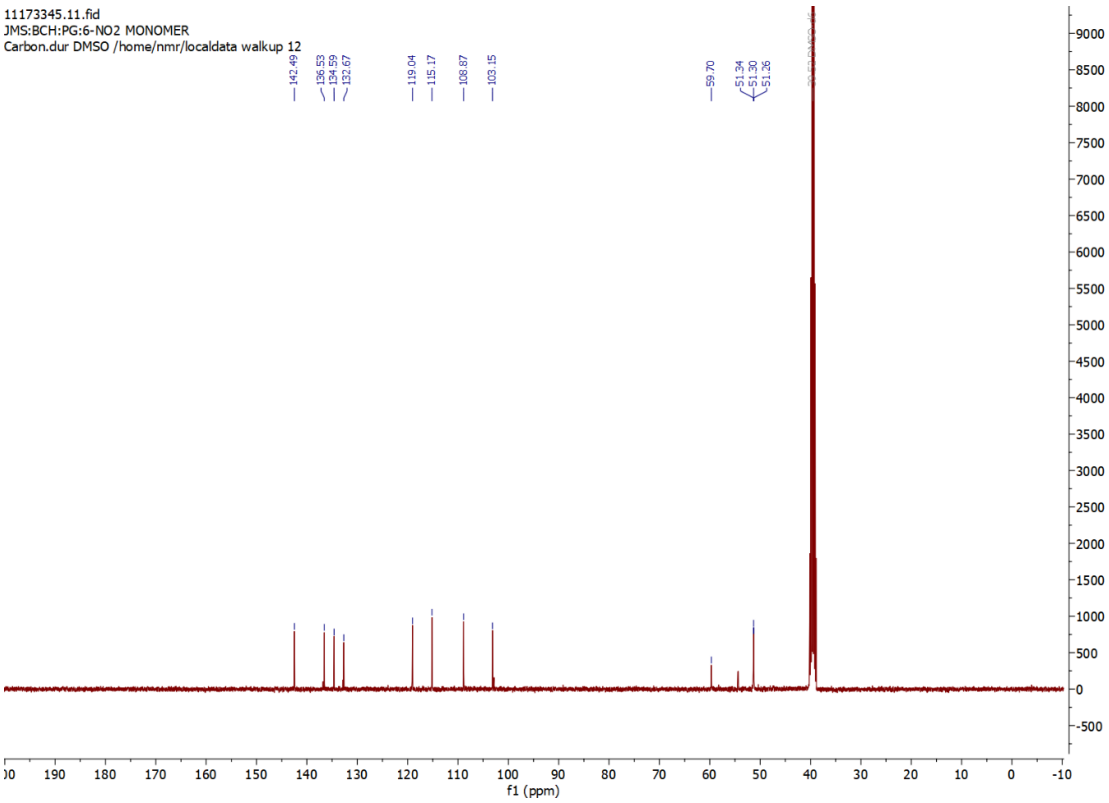
RT: 0.9262 minutes, Scan 101, 1: MS ES+ c (100.0-2000.0), NL 2.43e+7





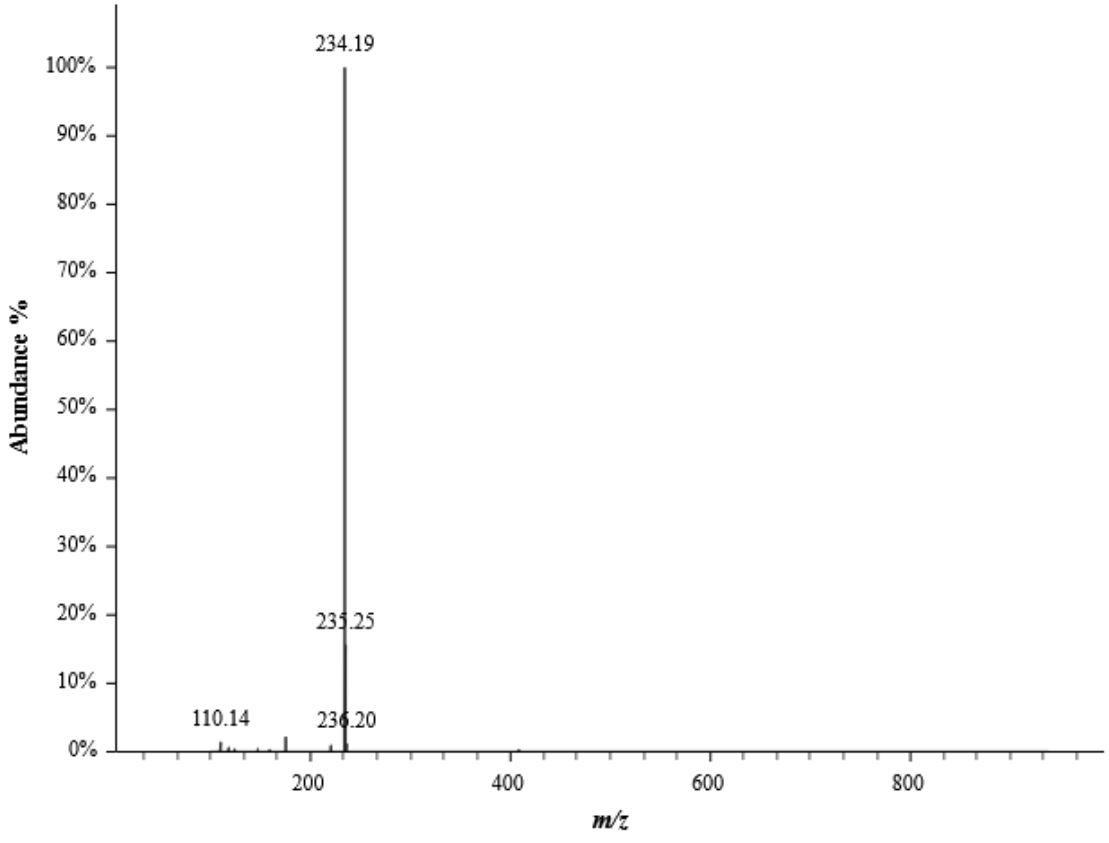
B.5.i. (6-Nitro-3-indolylmethyl)trimethylammonium iodide (**5i**)

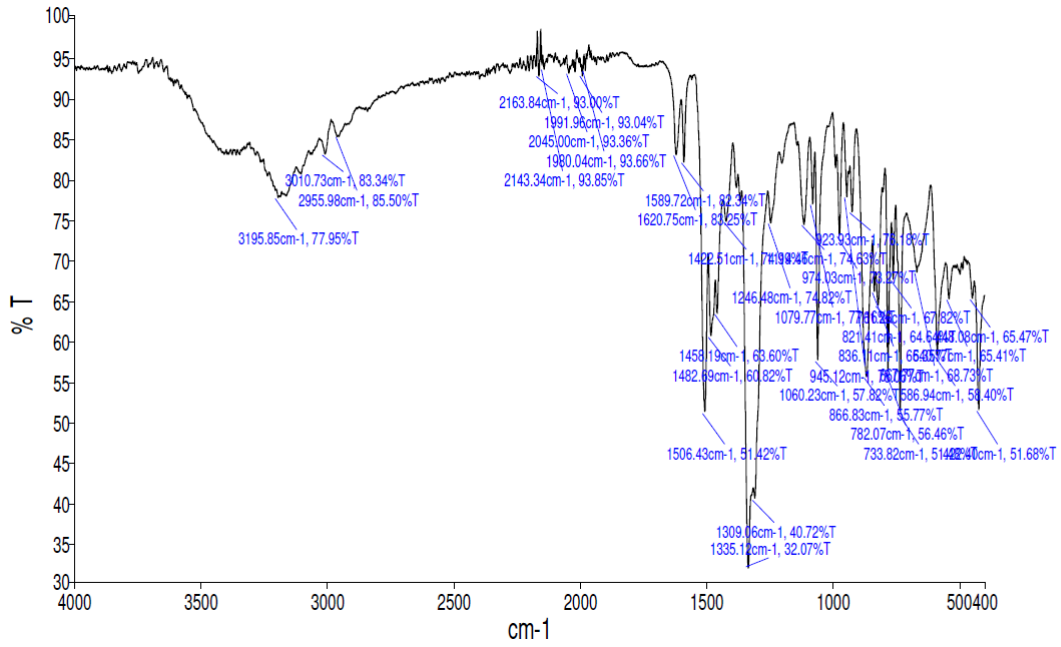




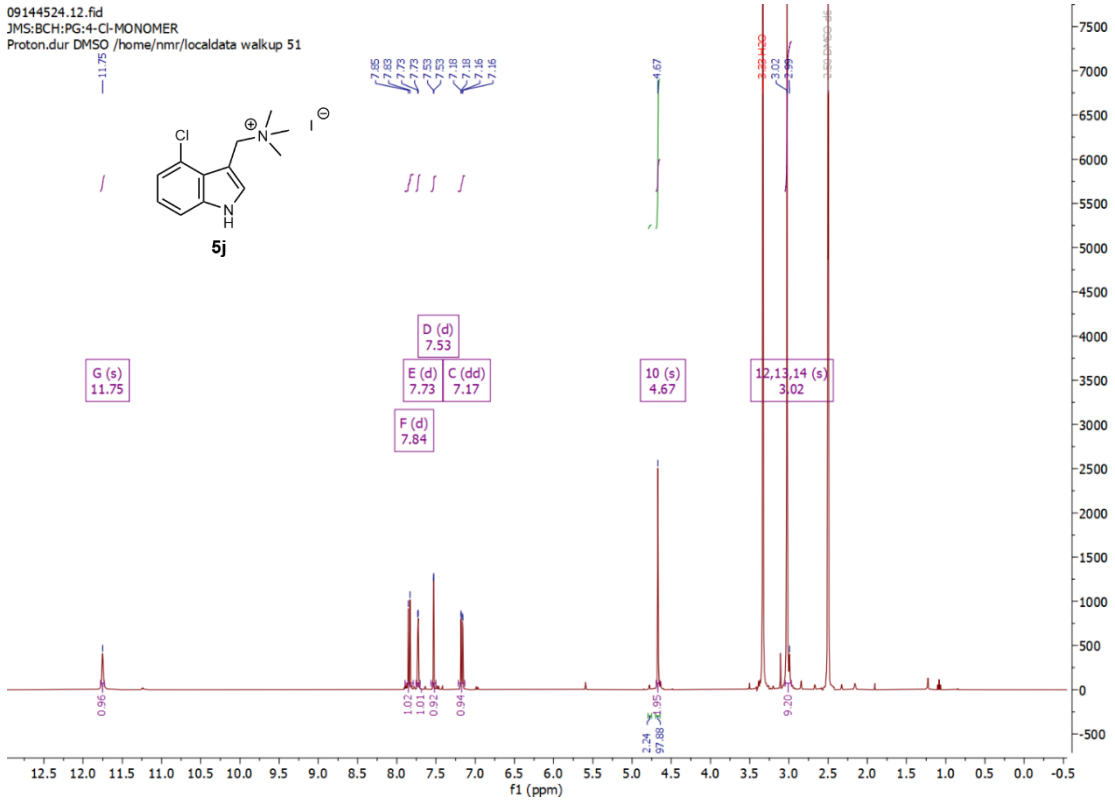
BCRH_6NO2_mono_211738

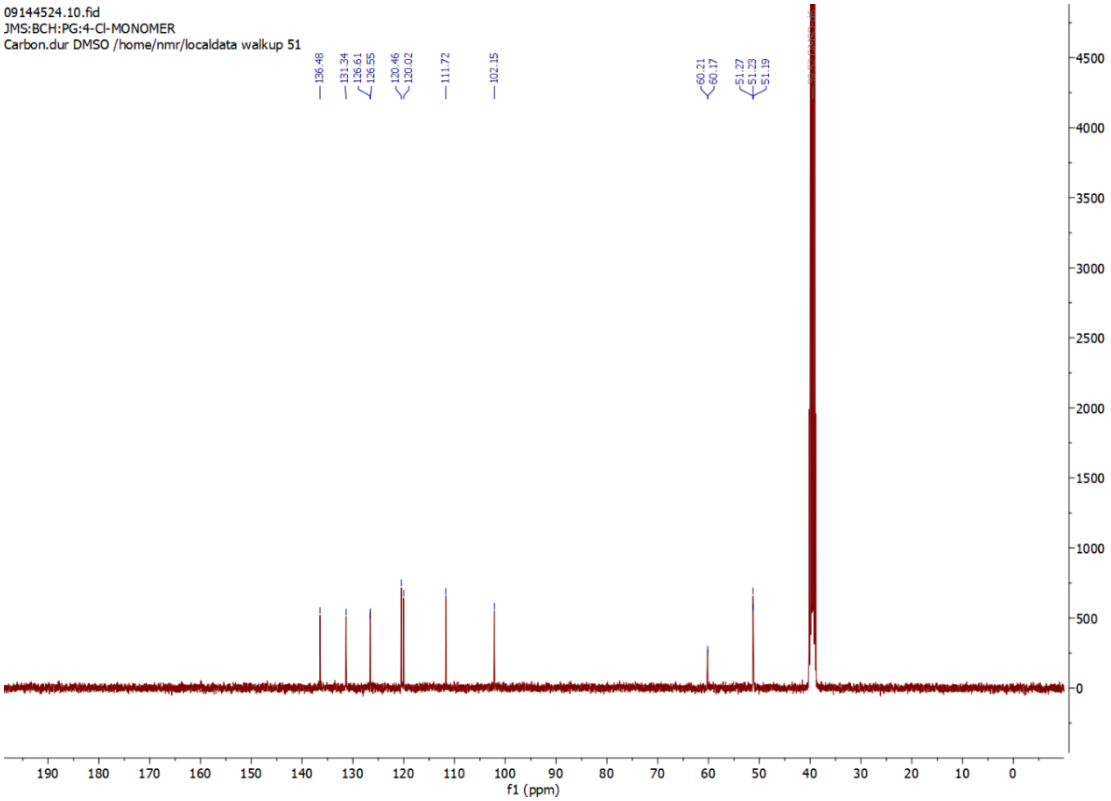
RT: 0.8915 minutes, Scan 97, 1: MS ES+ c (100.0-2000.0), NL 8.86e+7





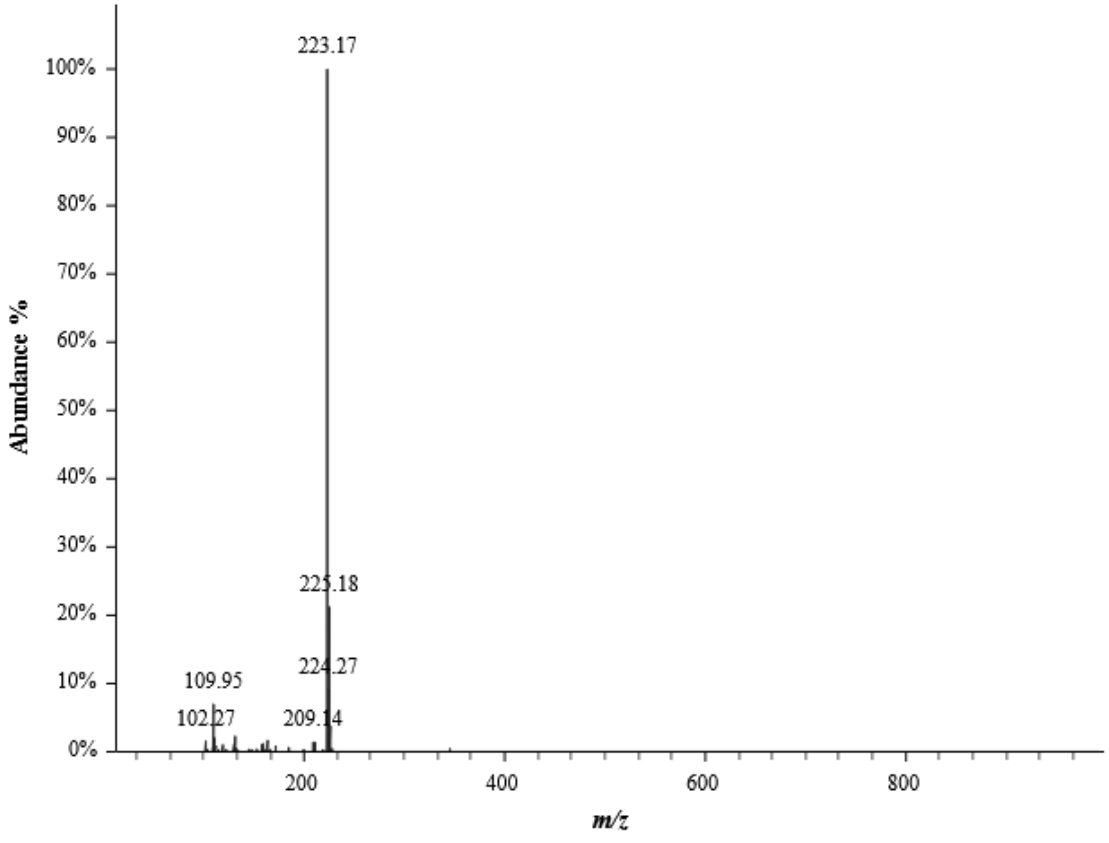
B.5.j. (4-Cl-3-indolylmethyl)trimethylammonium iodide (**5j**)

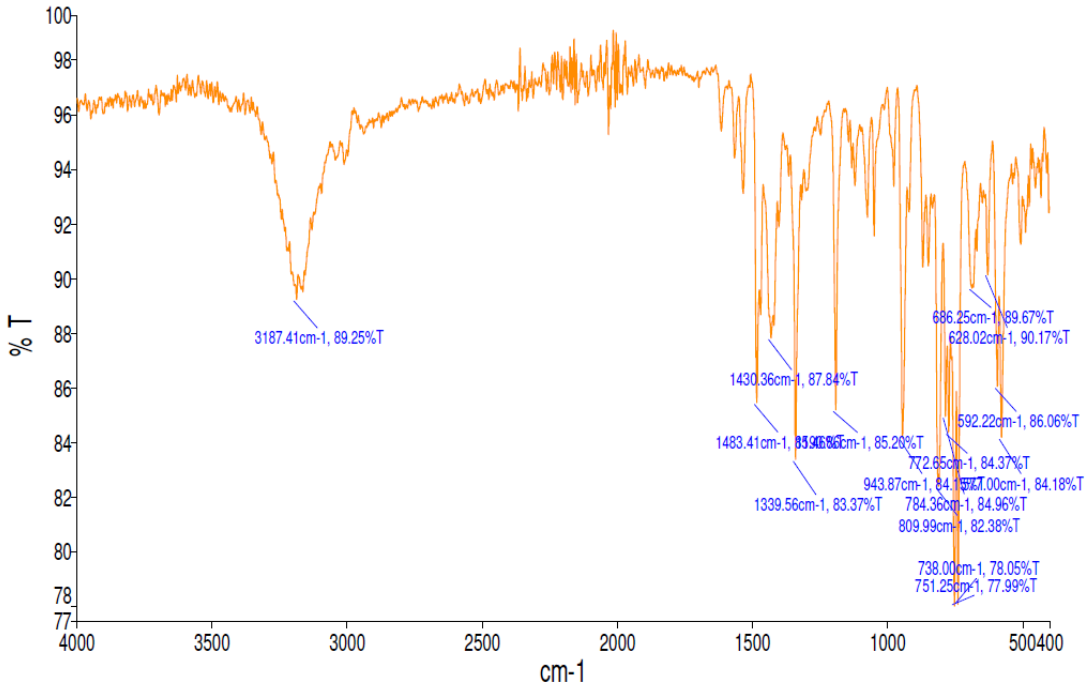




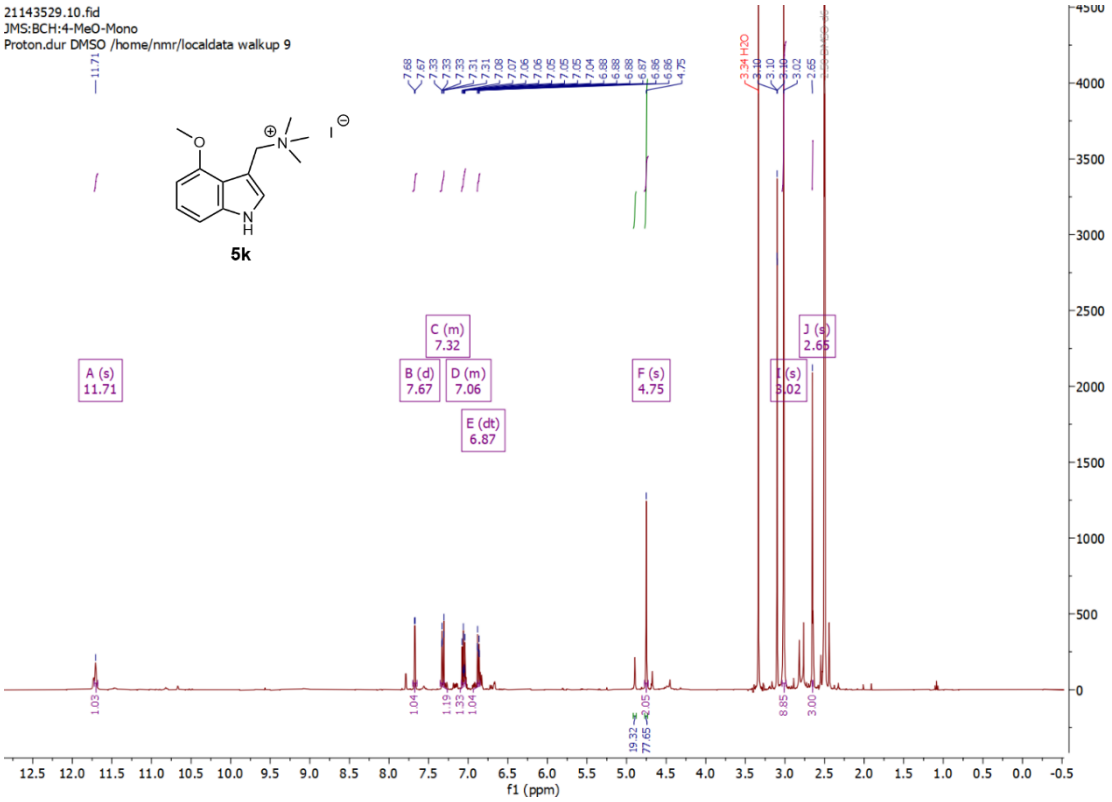
BCRH_4Cl_mono_211740

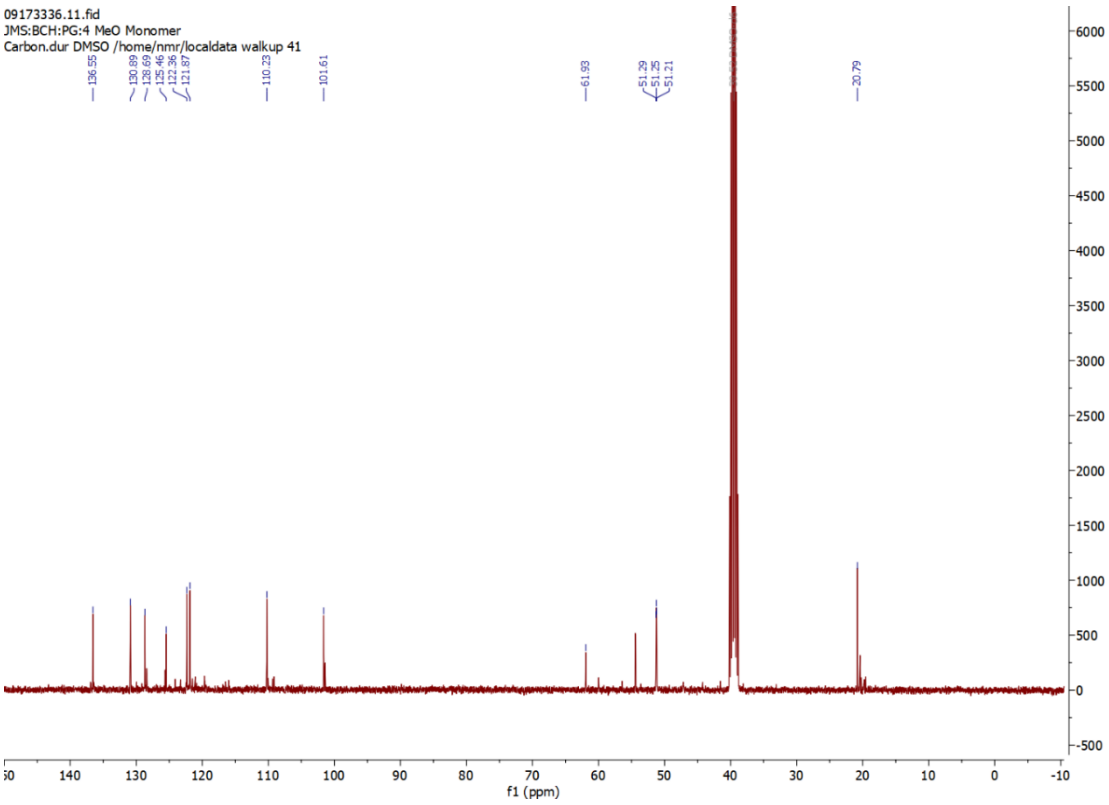
RT: 1.1865 minutes, Scan 131, 1: MS ES+ c (100.0-2000.0), NL 2.37e+7





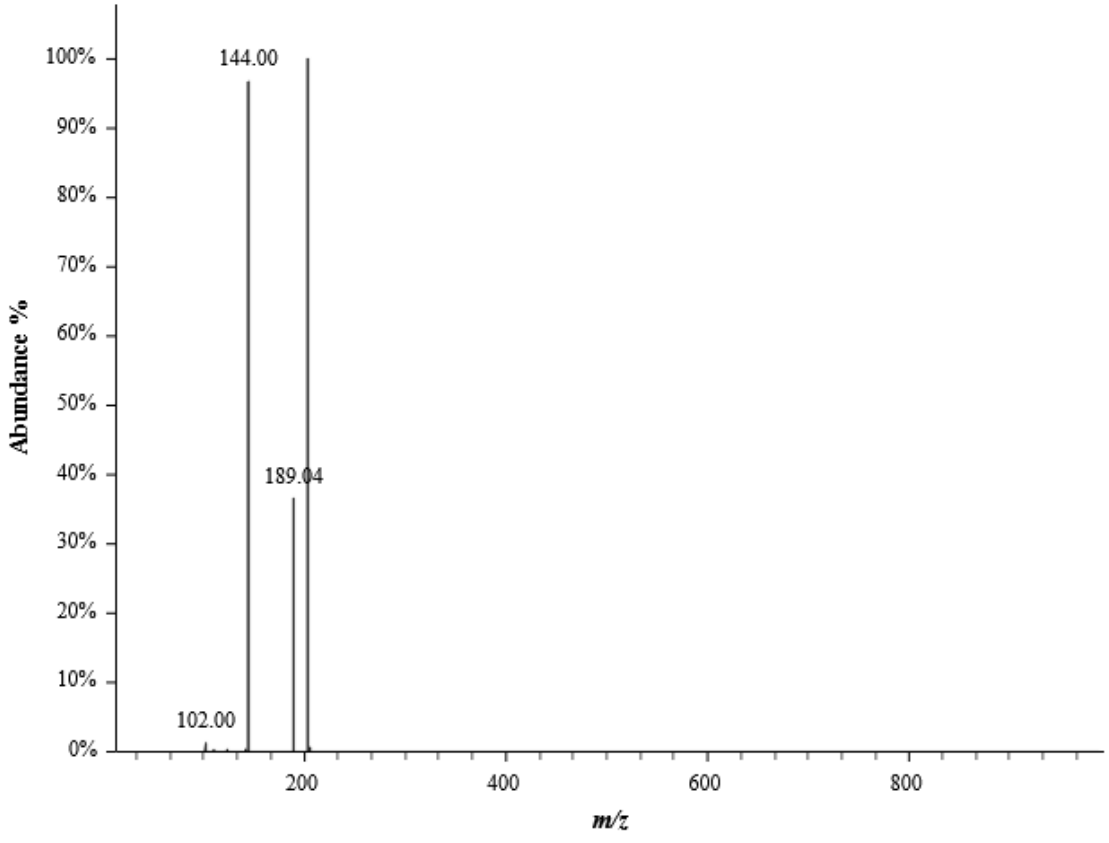
B.5.k. (4-MeO-3-indolylmethyl)trimethylammonium iodide (**5k**)

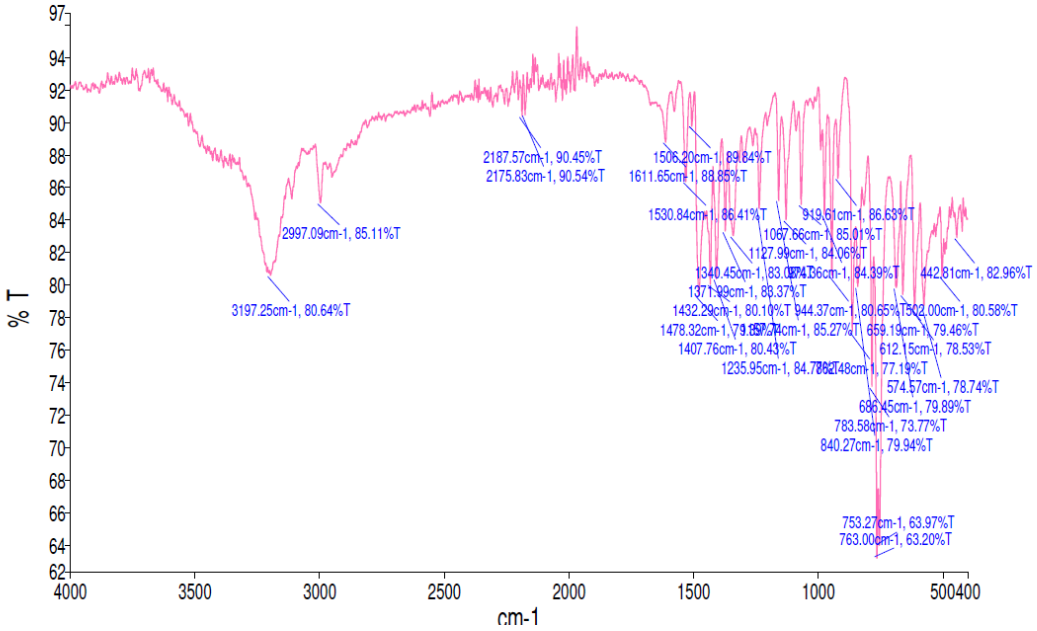




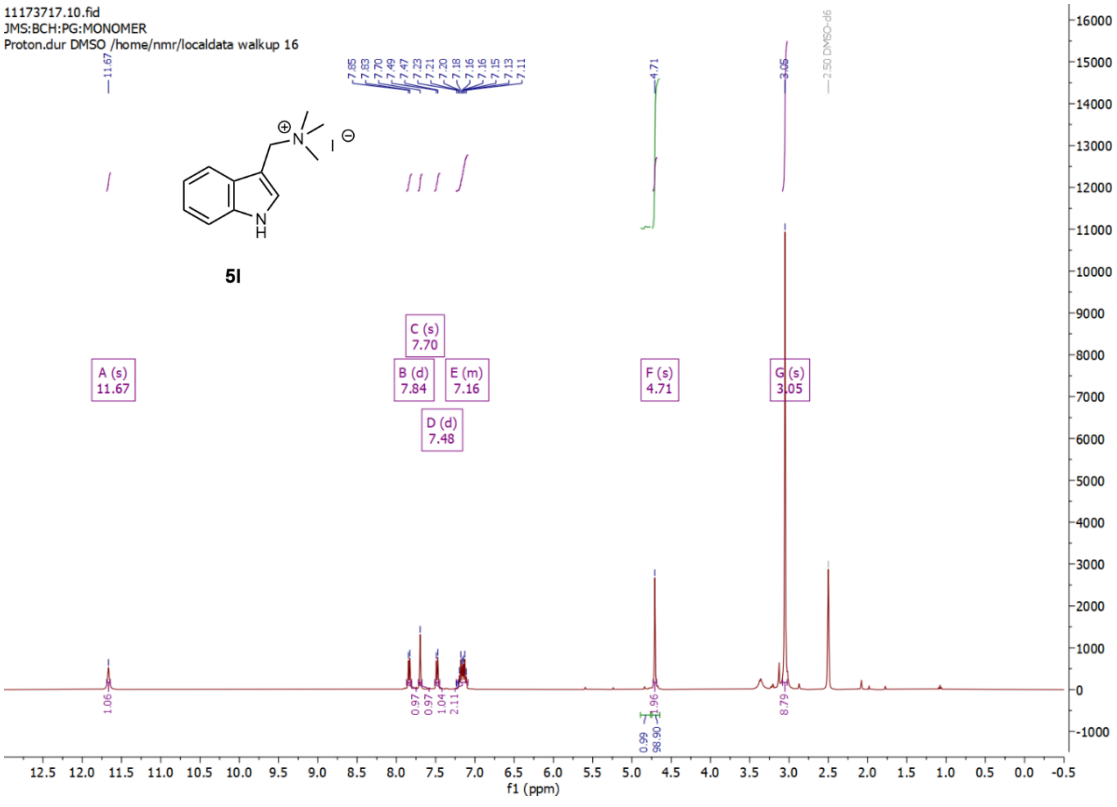
BCRH_4_MeO_mono_222786

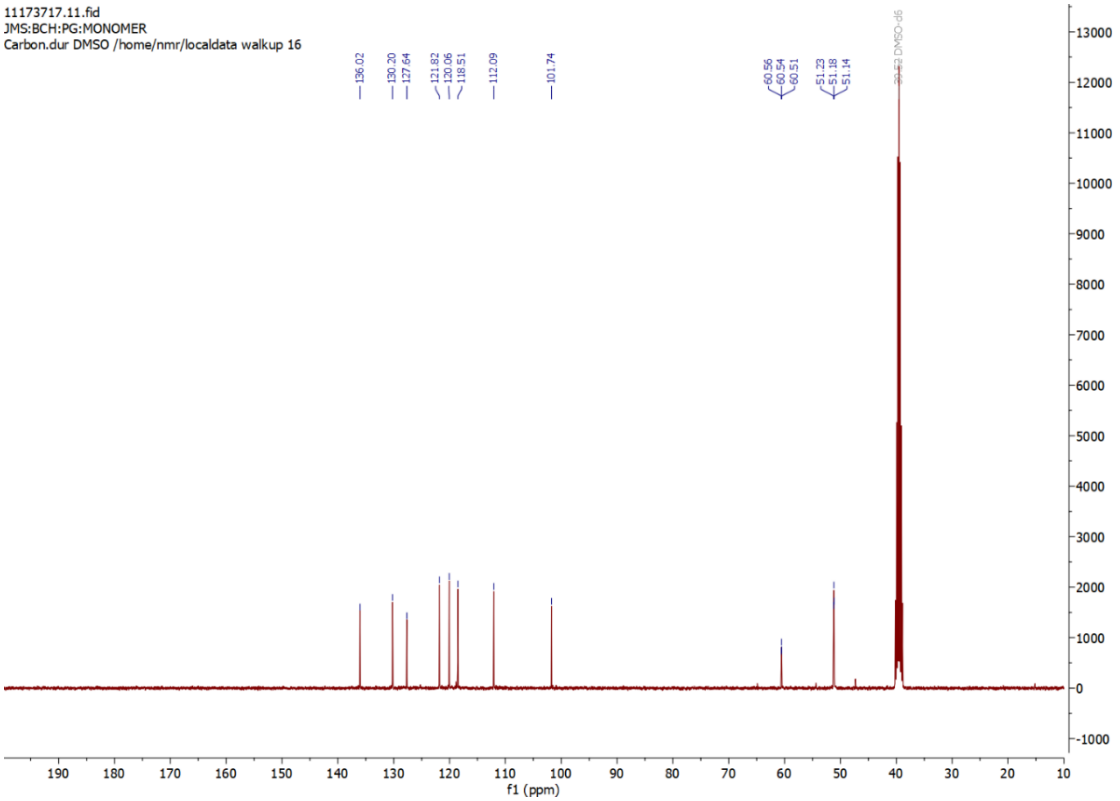
RT: 1.0303 minutes, Scan 113, 1: MS ES+ c (100.0-2000.0), NL 1.35e+8





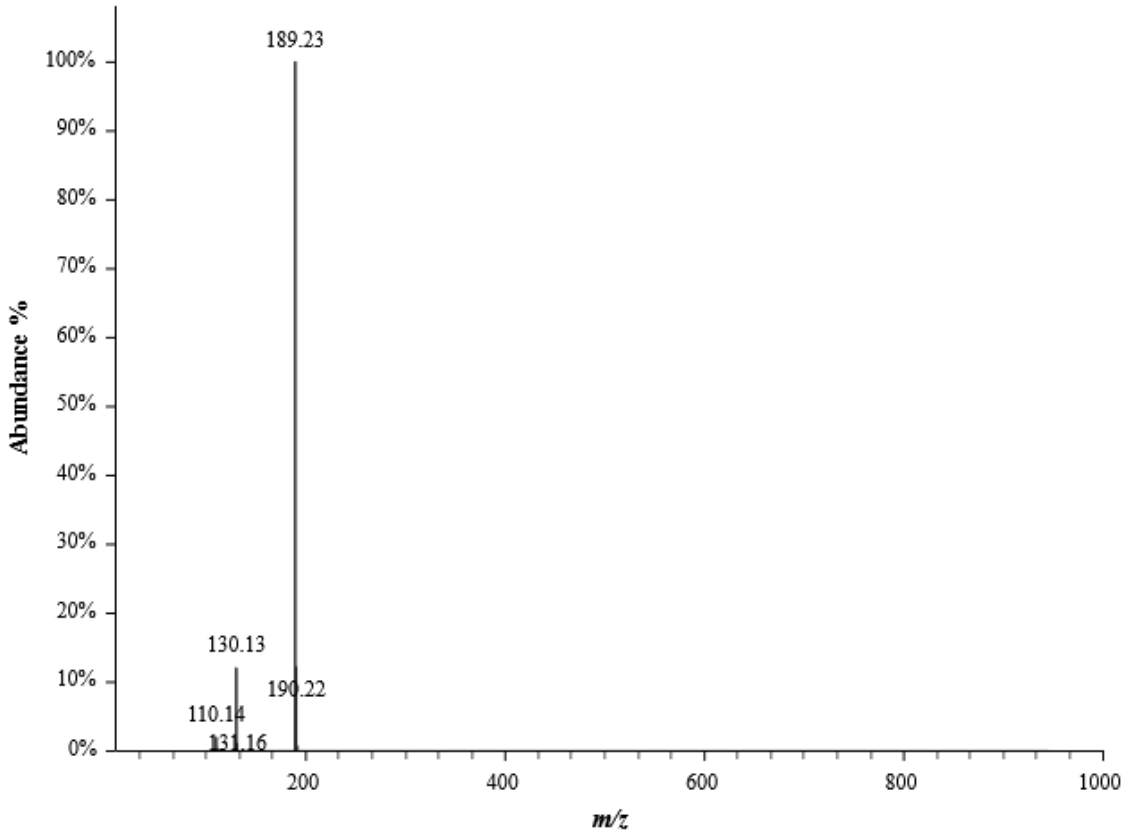
B.5.l. Monomer (3-indolylmethyl)trimethylammonium iodide (**5l**)

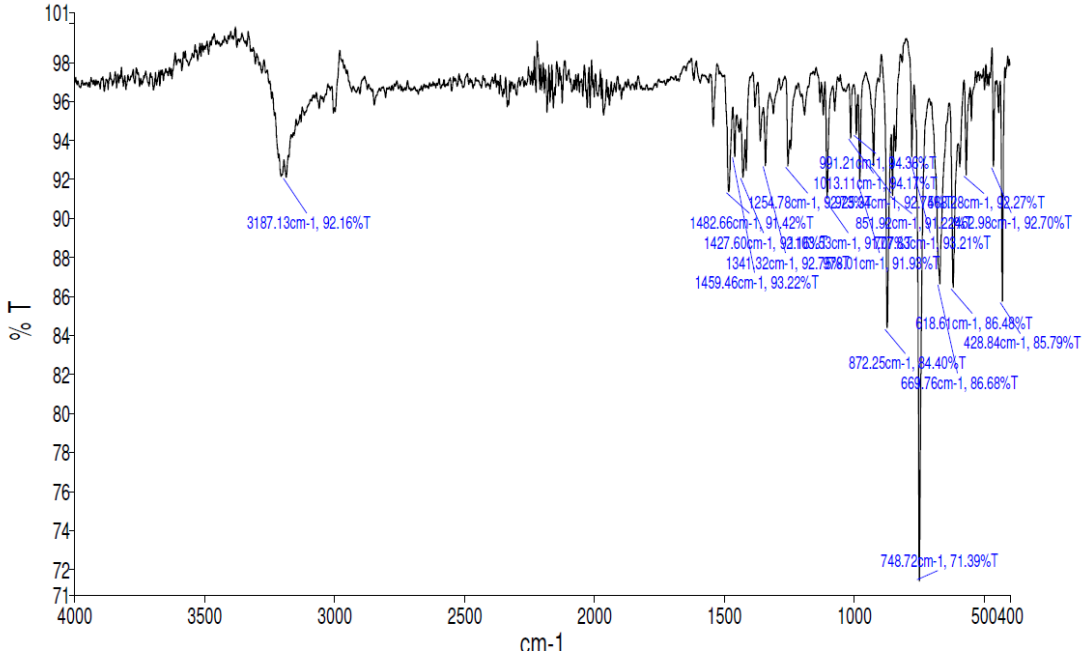




BCRH_monomer_211785

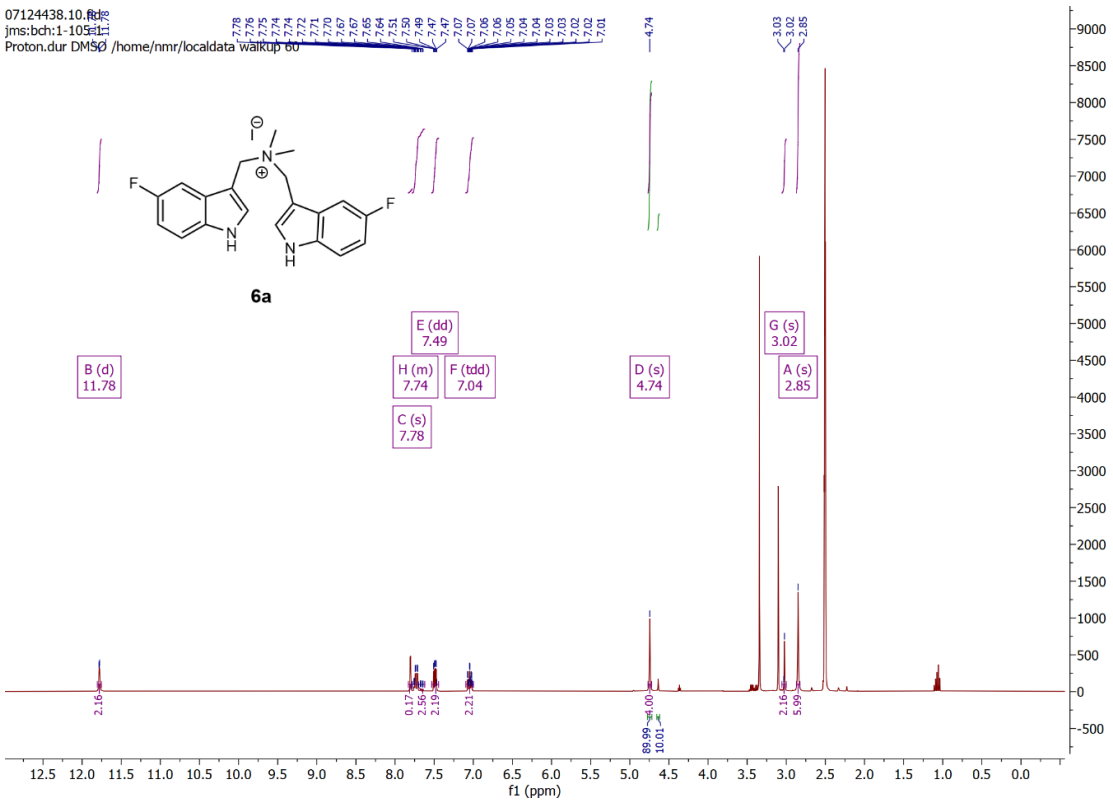
RT: 0.8394 minutes, Scan 91, 1: MS ES+ c (100.0-2000.0), NL 6.25e+7

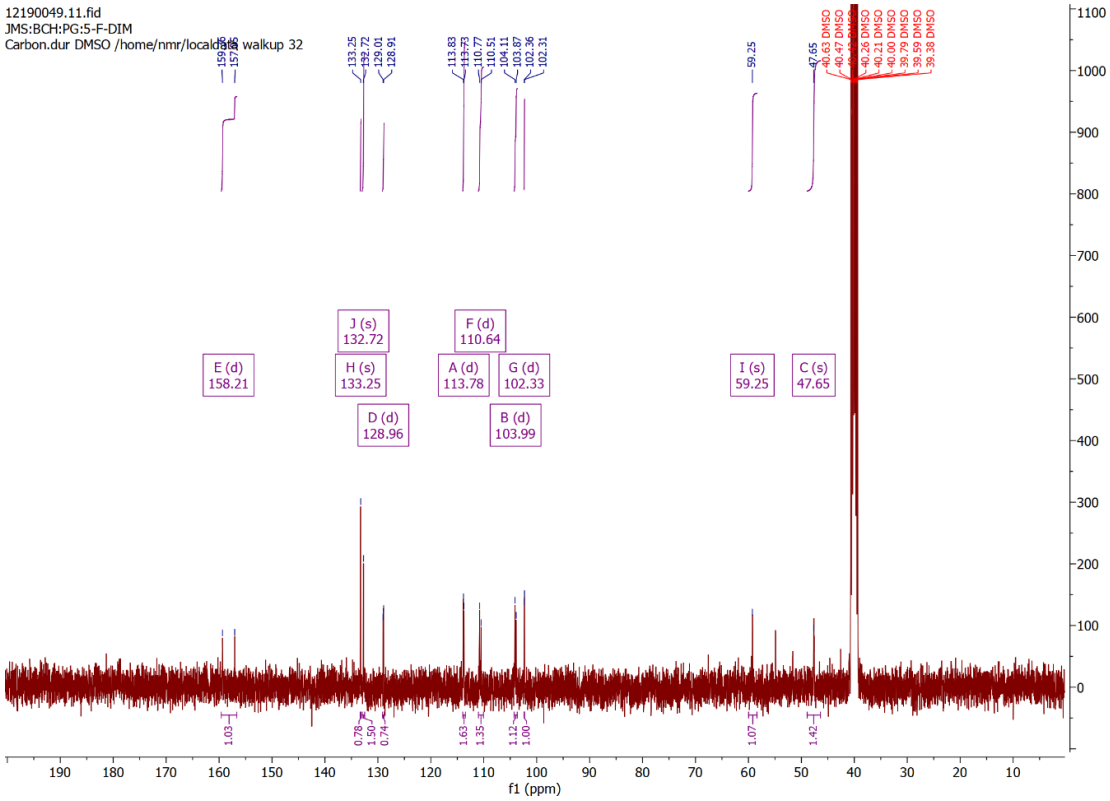


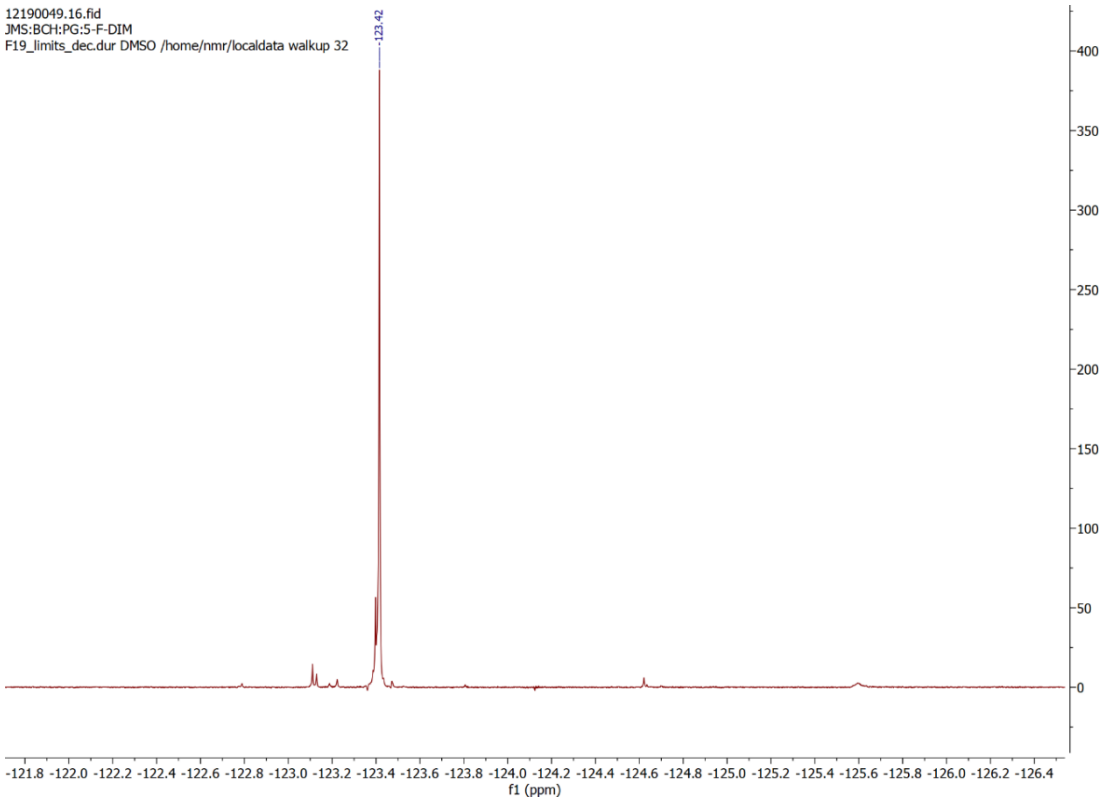


B.6. Indole Dimers

B.6.a. 1-(5-fluoro-1H-indol-3-yl)-N-((5-fluoro-1H-indol-3-yl)methyl)-N,N-dimethylmethanaminium iodide (**6a**)

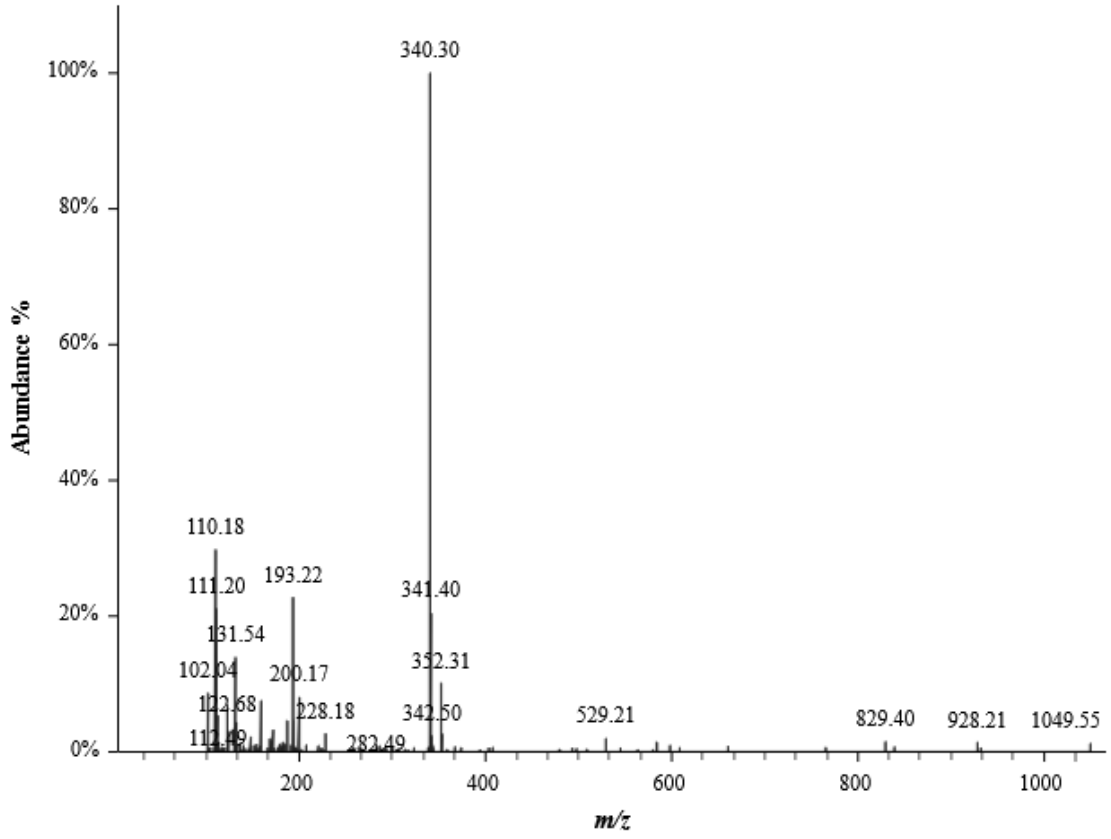


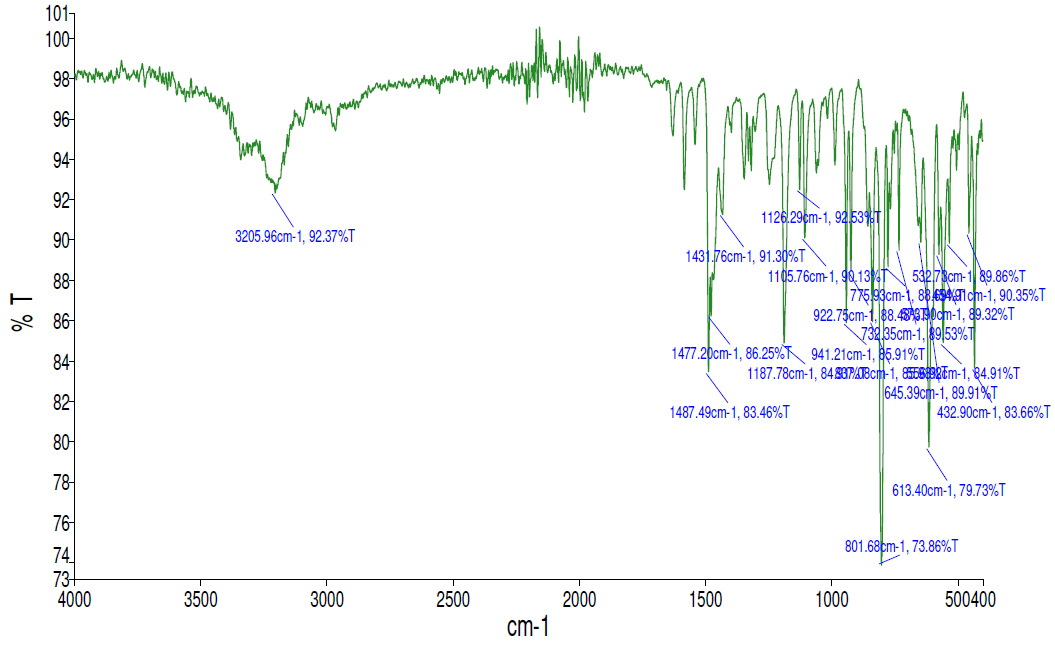

12190049.11.fid
JMS:BCH:PG:5-F-DIM
Carbon.dur DMSO /home/nmr/localdata walkup 32
J (s) 132.72
F (d) 110.64
E (d) 158.21
H (s) 133.25
A (d) 113.78
G (d) 102.33
I (s) 59.25
C (s) 47.65
D (d) 128.96
B (d) 103.99
f1 (ppm)


12190049.16.fid
JMS:BCH:PG:5-F-DIM
F19_limits_dec.dur DMSO /home/nmr/localdata walkup 32
f1 (ppm)

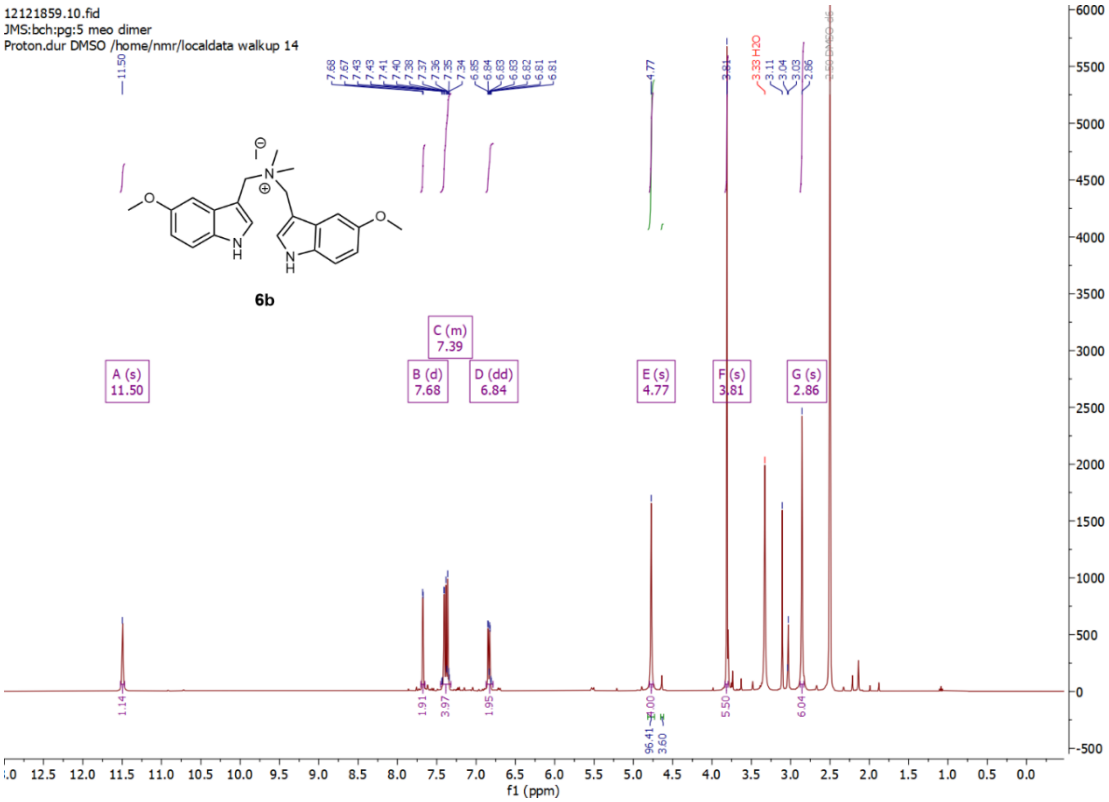
BCRH_5FDimer_206772

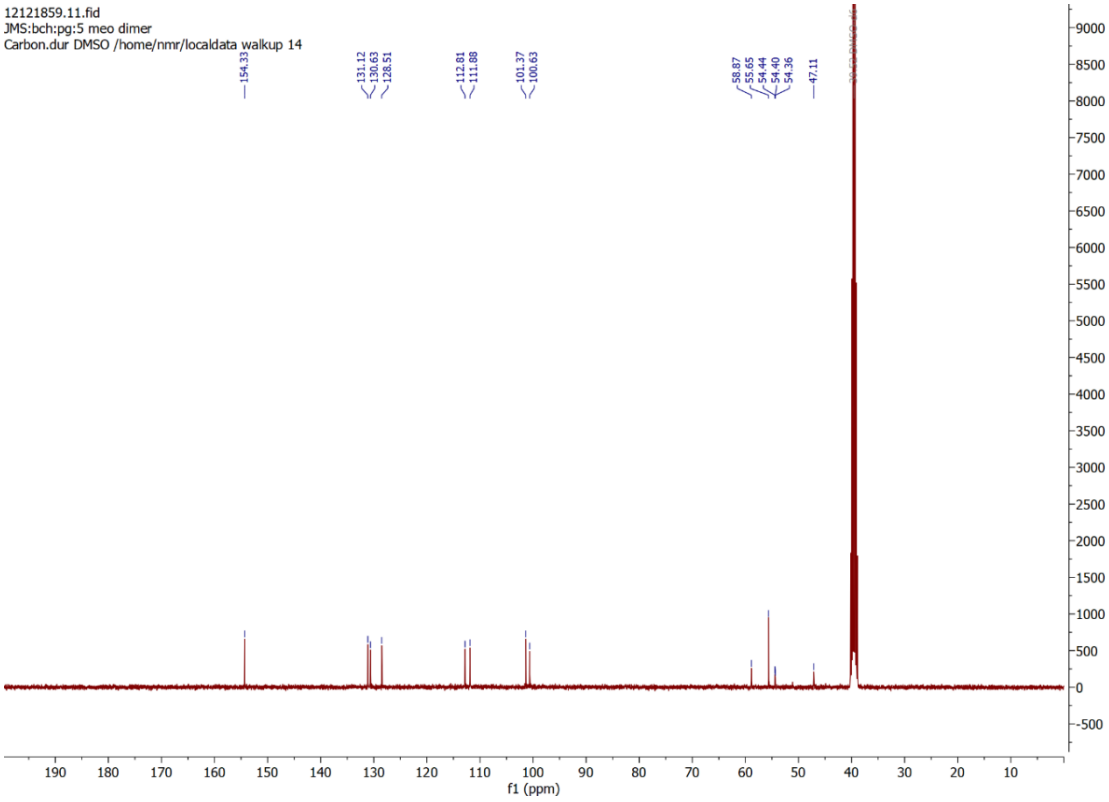
RT: 1.6551 minutes, Scan 185, 1: MS ES+ c (100.0-2000.0), NL 7.50e+6





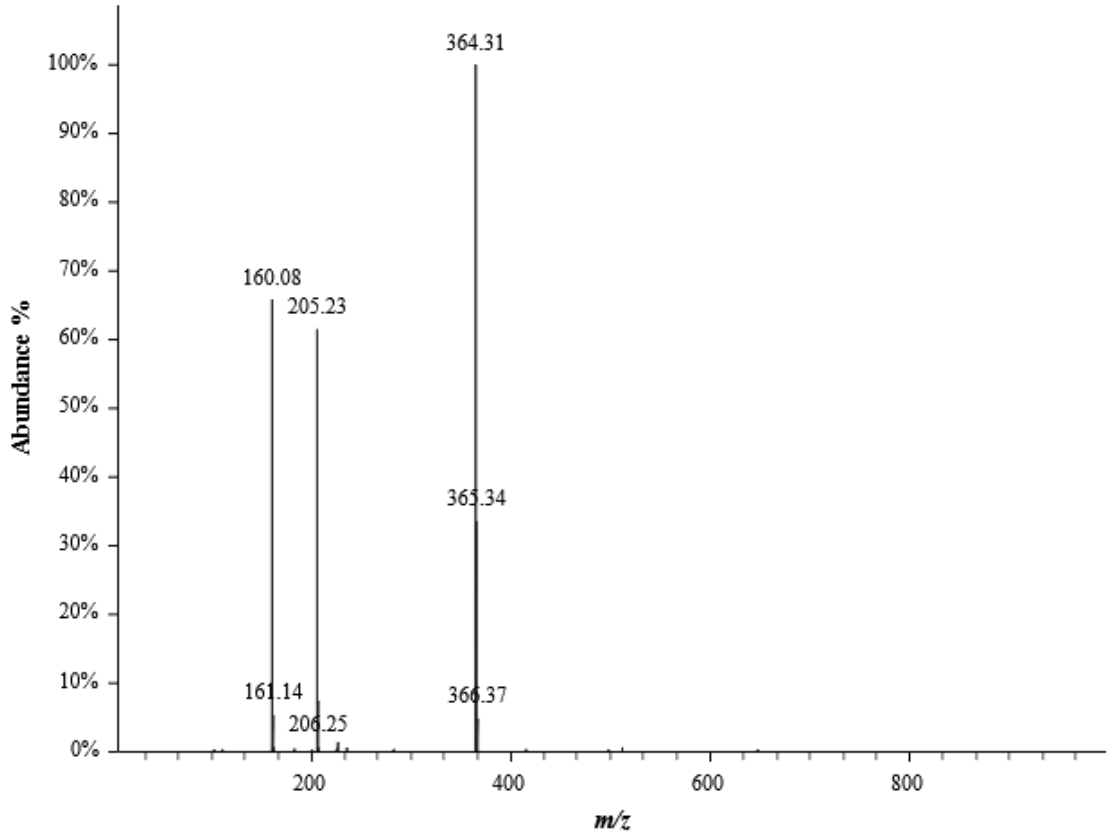
B.6.b. 1-(5-methoxy-1H-indol-3-yl)-N-((5-methoxy-1H-indol-3-yl)methyl)-N,N-dimethylmethanaminium iodide (**6b**)

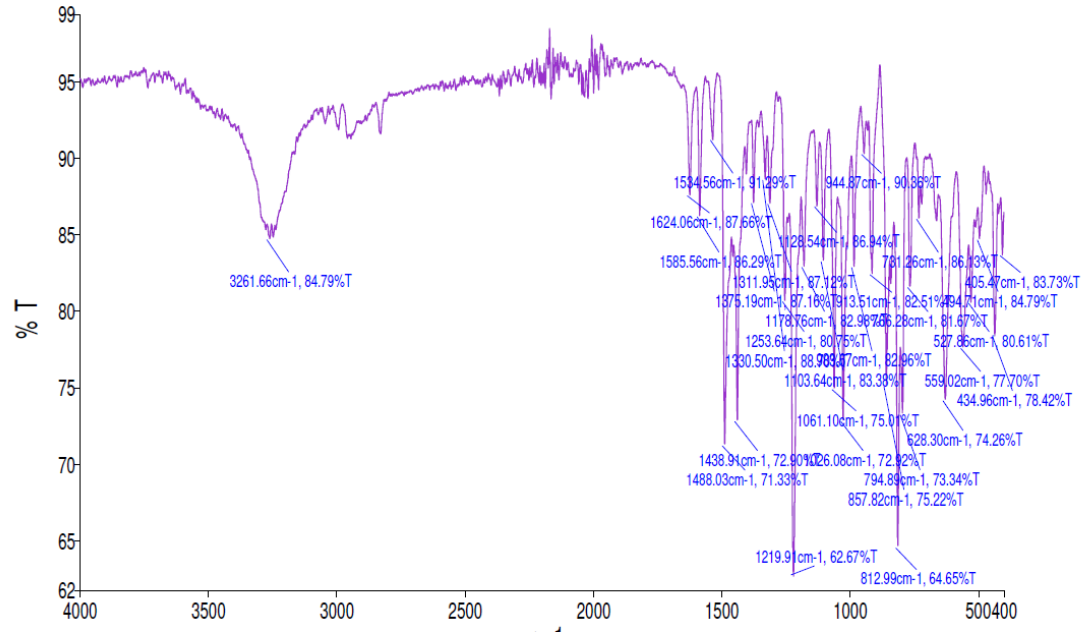




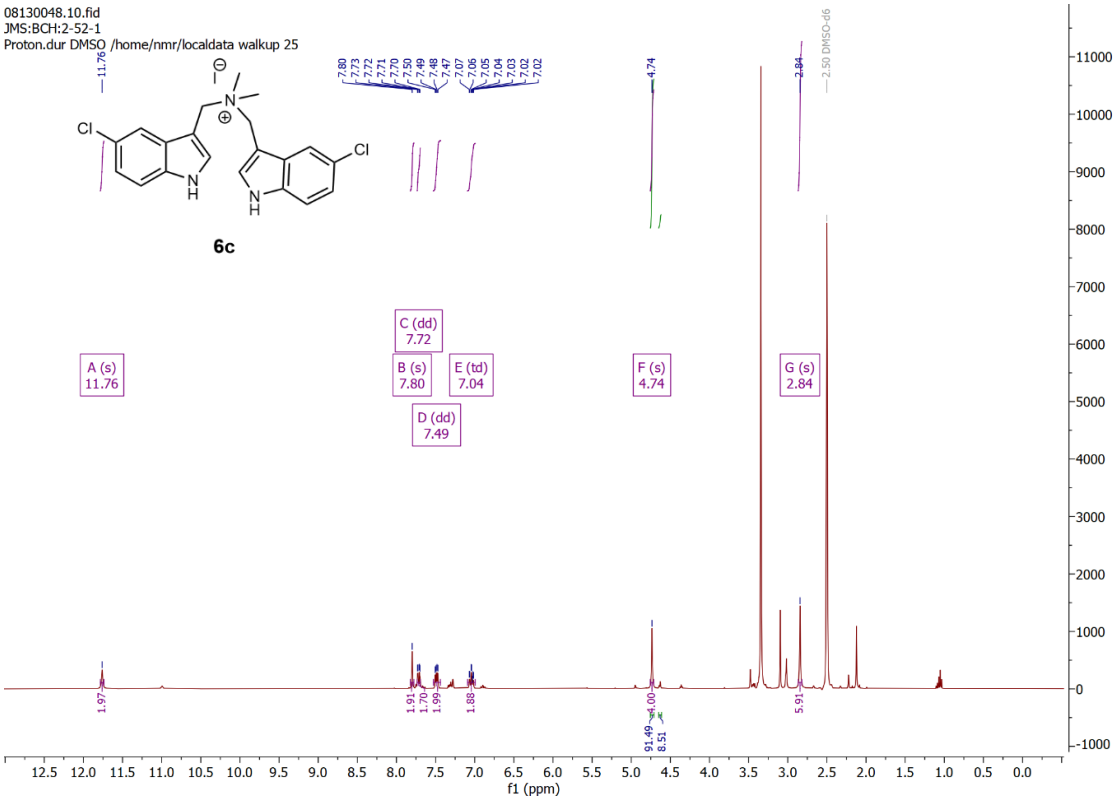
BCRH_5MeO_Dimer_212081

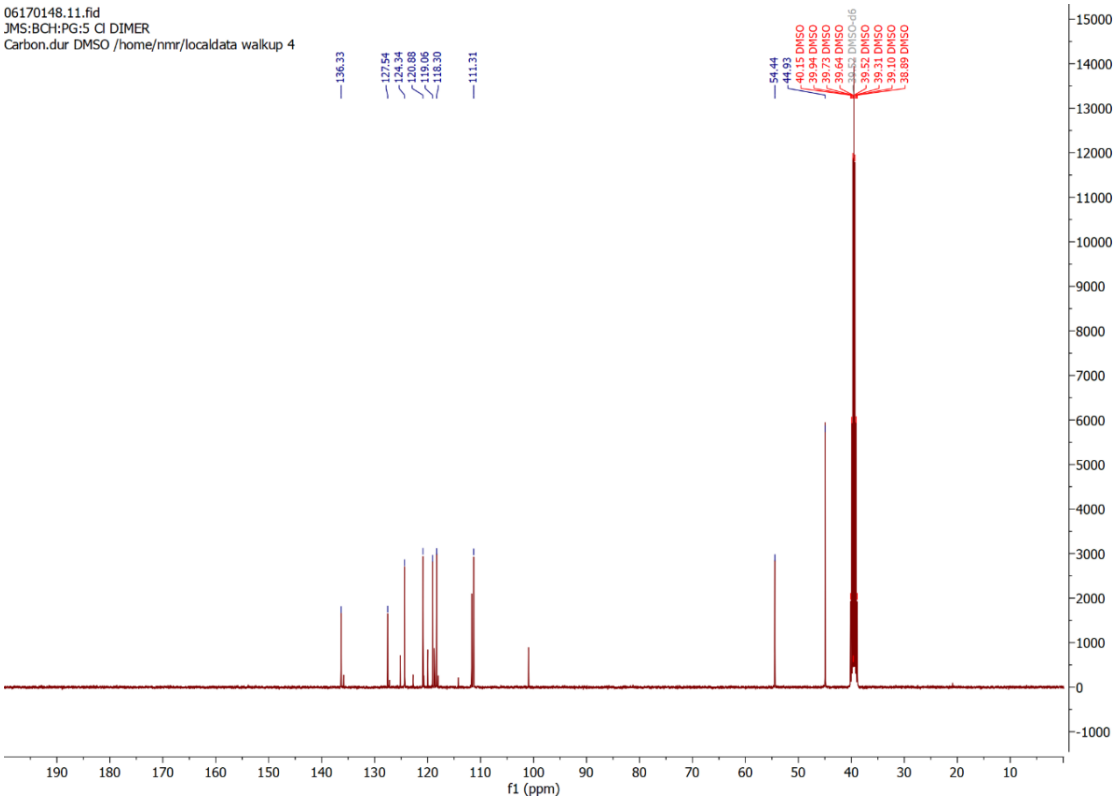
RT: 1.6031 minutes, Scan 179, 1: MS ES+ c (100.0-2000.0), NL 1.10e+8





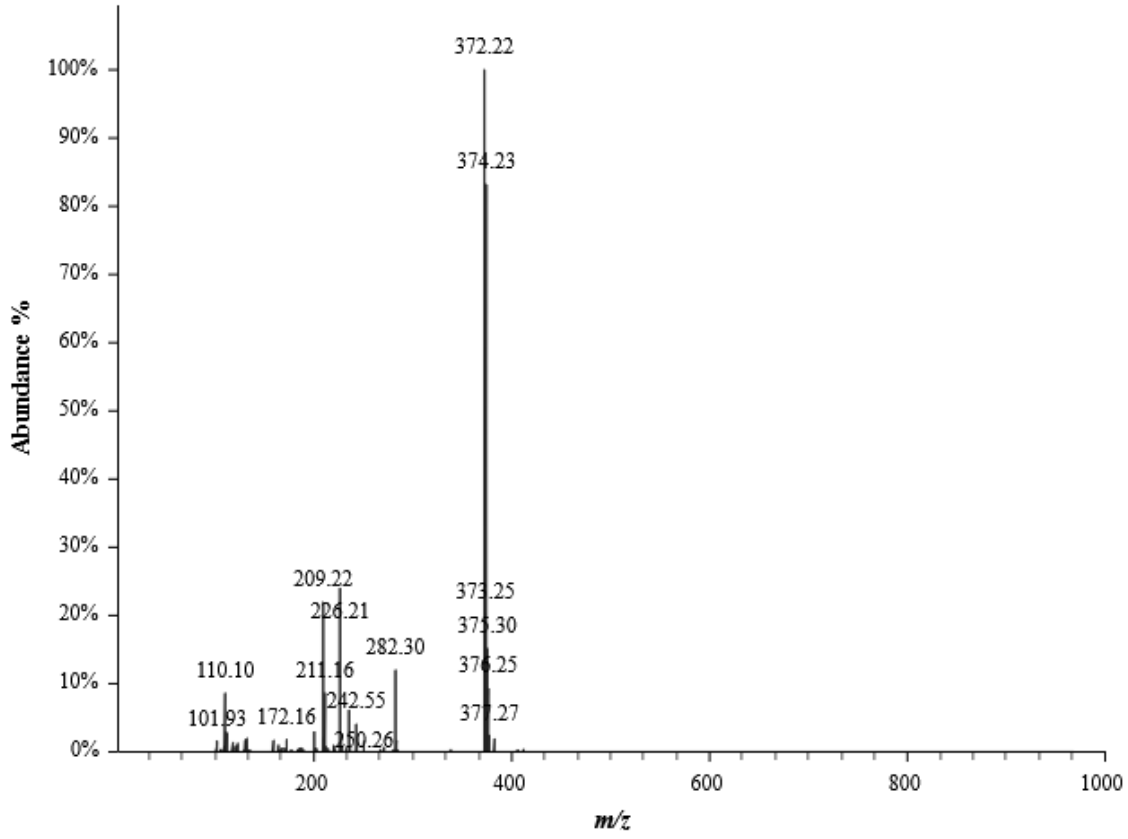
B.6.c. 1-(5-Chloro-1H-indol-3-yl)-N-((5-chloro-1H-indol-3-yl)methyl)-N,N-dimethylmethanaminium iodide (**6c**)

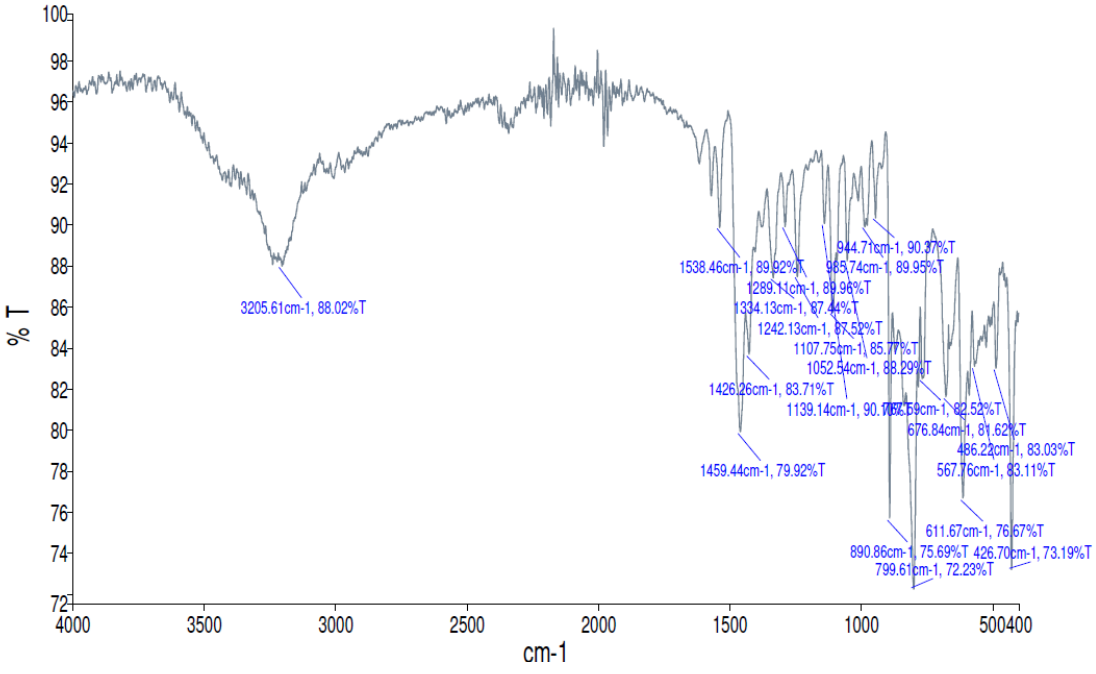




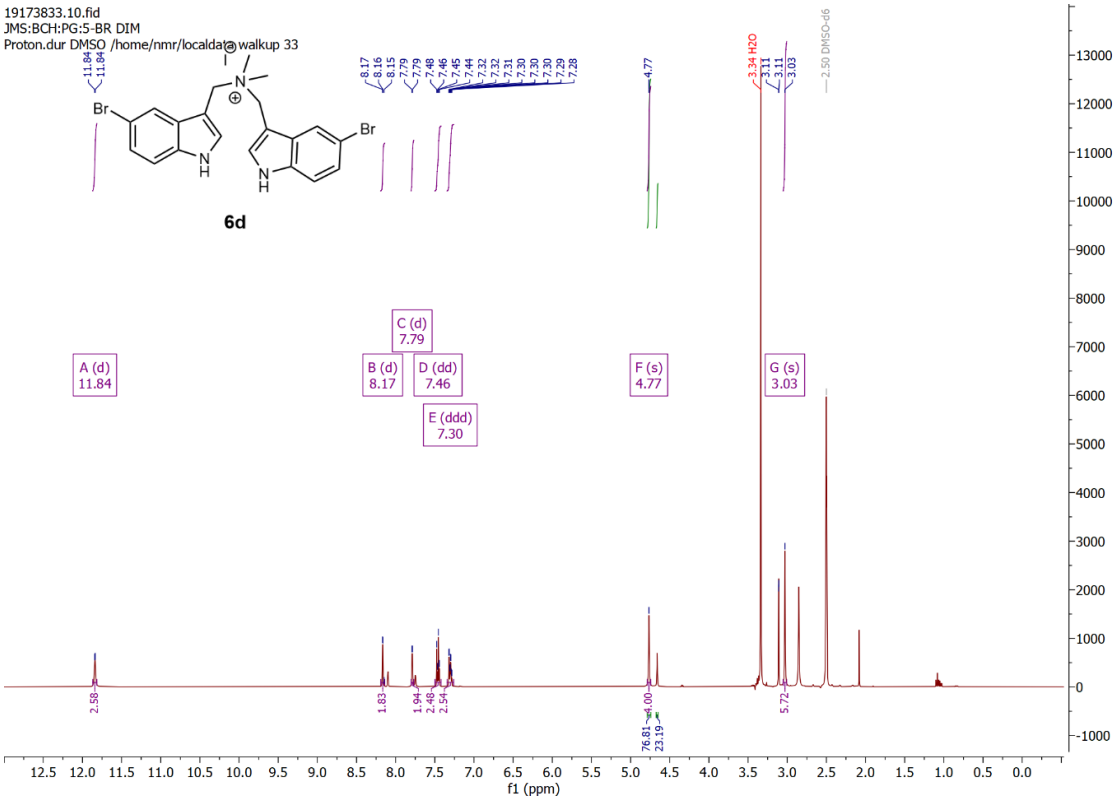
BCRH_5Cl_dimer2_212090

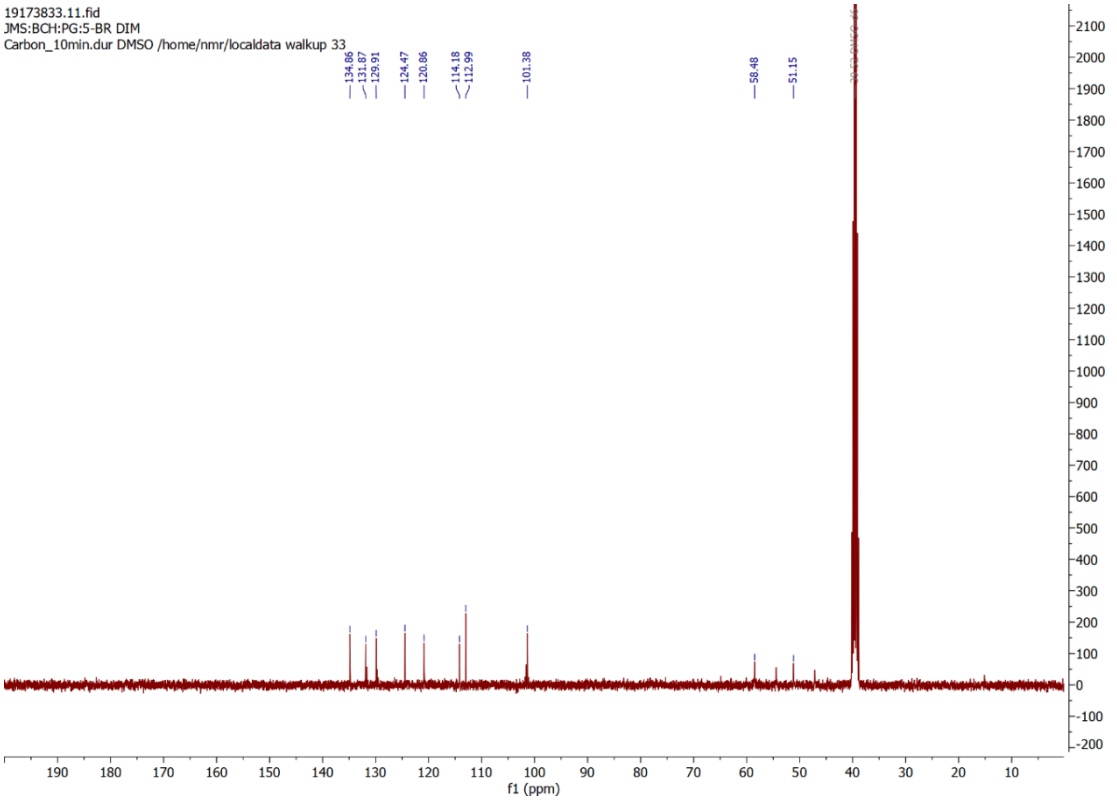
RT: 2.0022 minutes, Scan 225, 1: MS ES+ c (100.0-2000.0), NL 2.19e+7





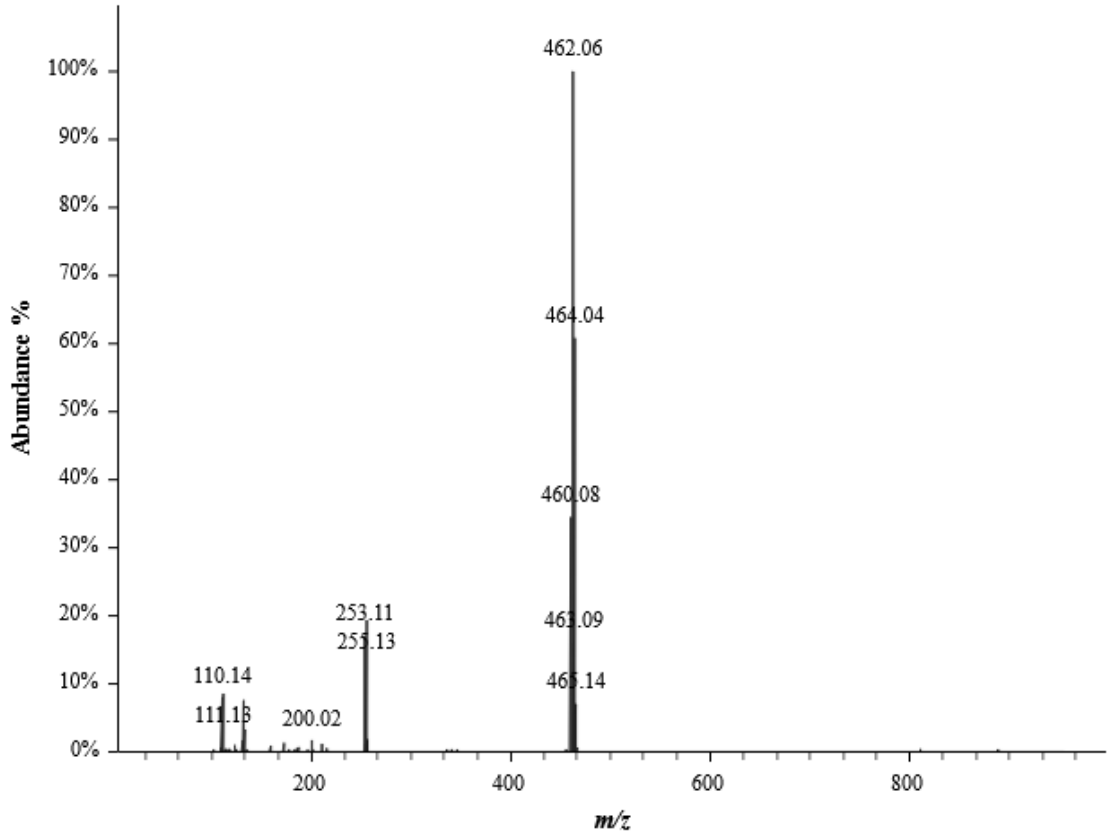
B.6.d. 1-(5-Bromo-1H-indol-3-yl)-N-((5-bromo-1H-indol-3-yl)methyl)-N,N-dimethylmethanaminium iodide (**6d**)

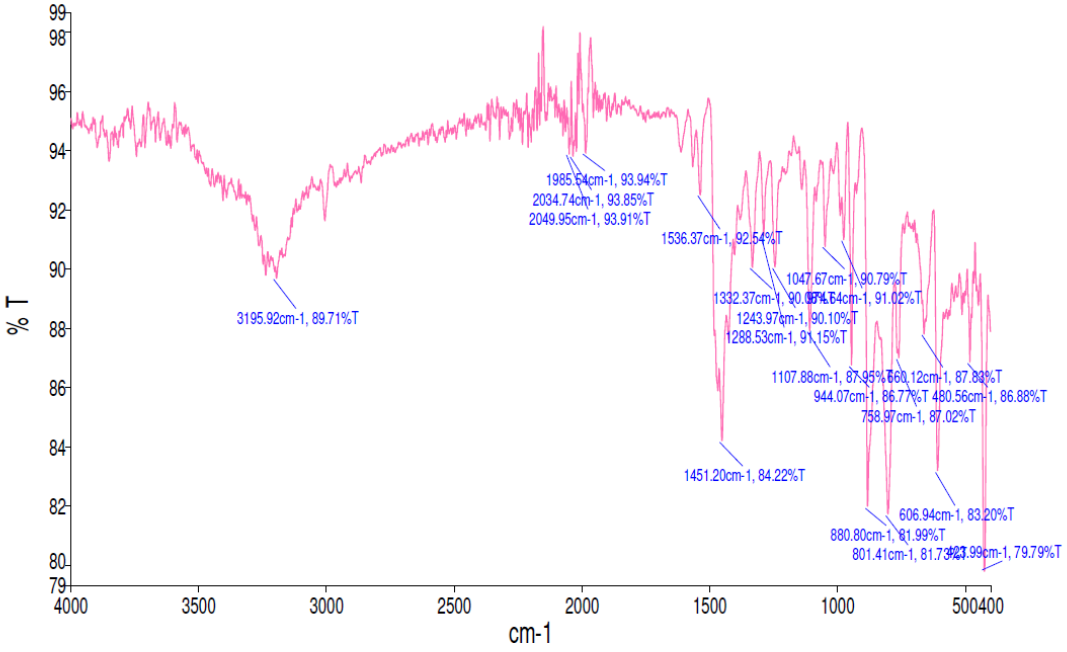




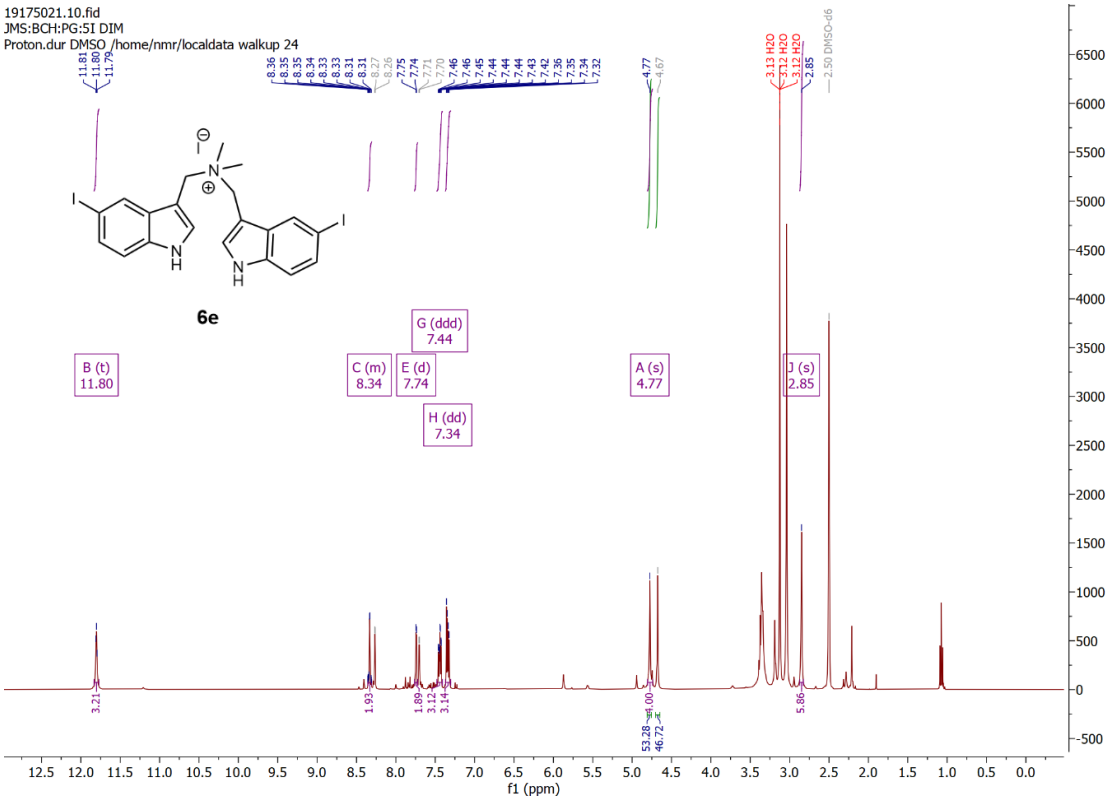
BCRH_5BrDimer_210392

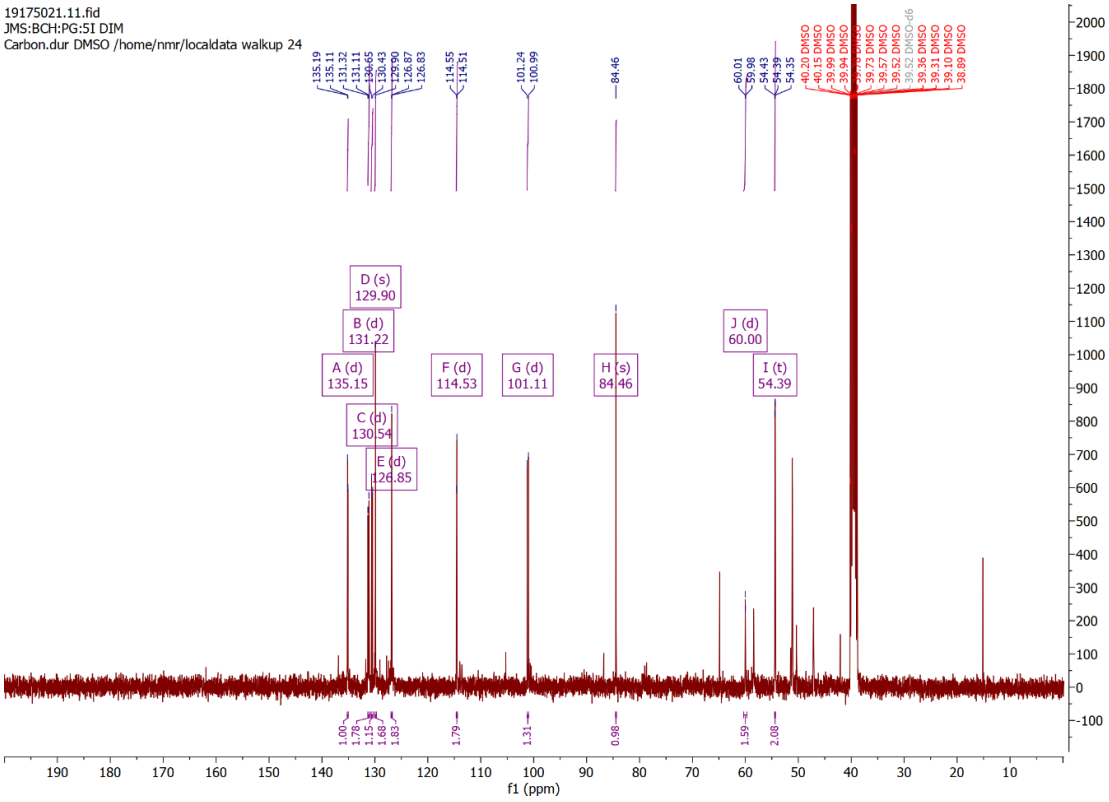
RT: 2.0543 minutes, Scan 231, 1: MS ES+ c (100.0-2000.0), NL 3.14e+7





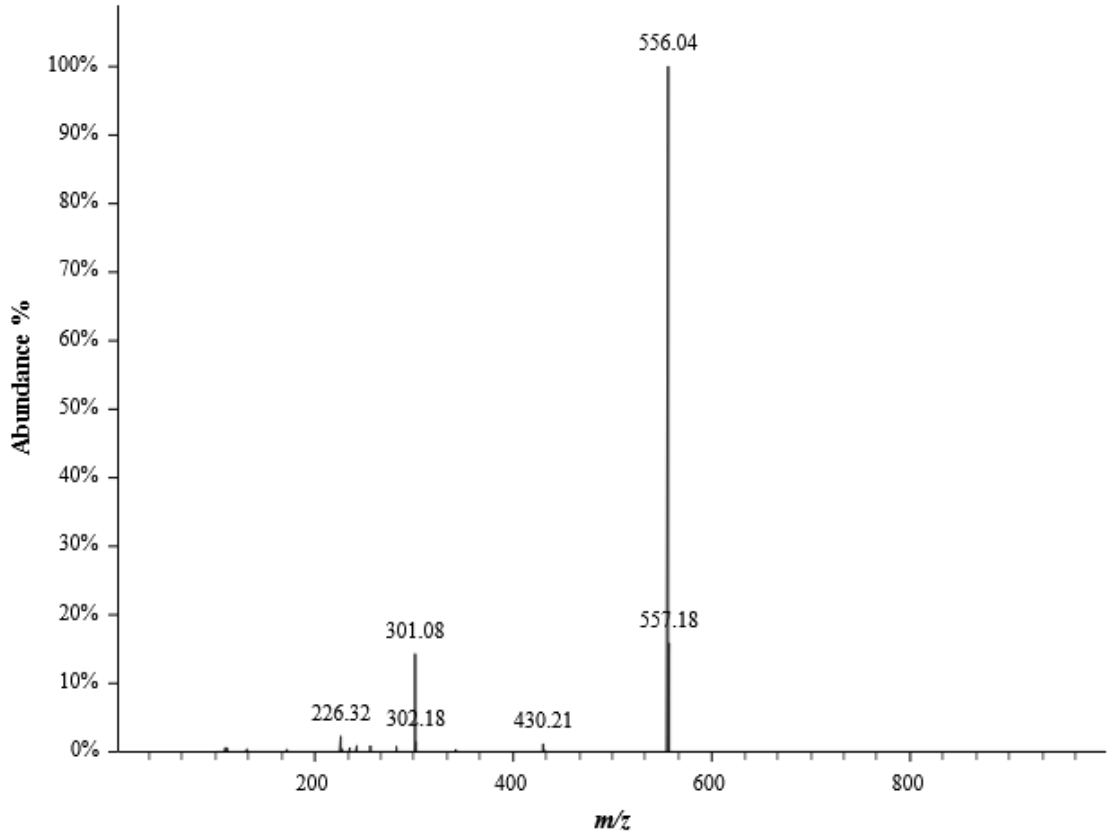
B.6.e. 1-(5-Iodo-1H-indol-3-yl)-N-((5-iodo-1H-indol-3-yl)methyl)-N,N-dimethylmethanaminium (**6e**)

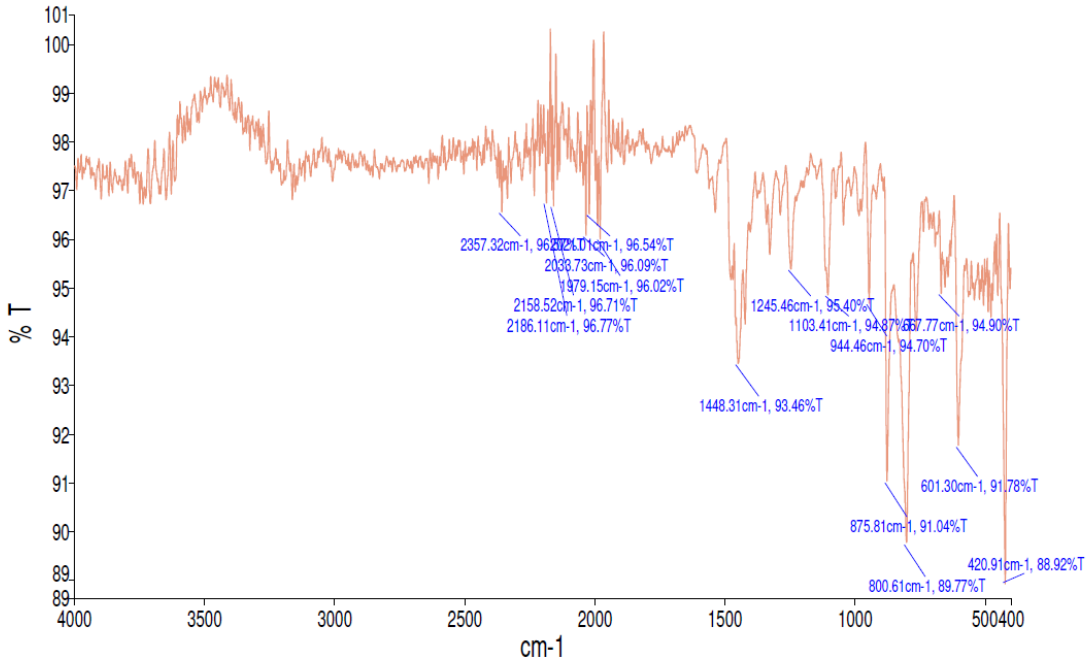




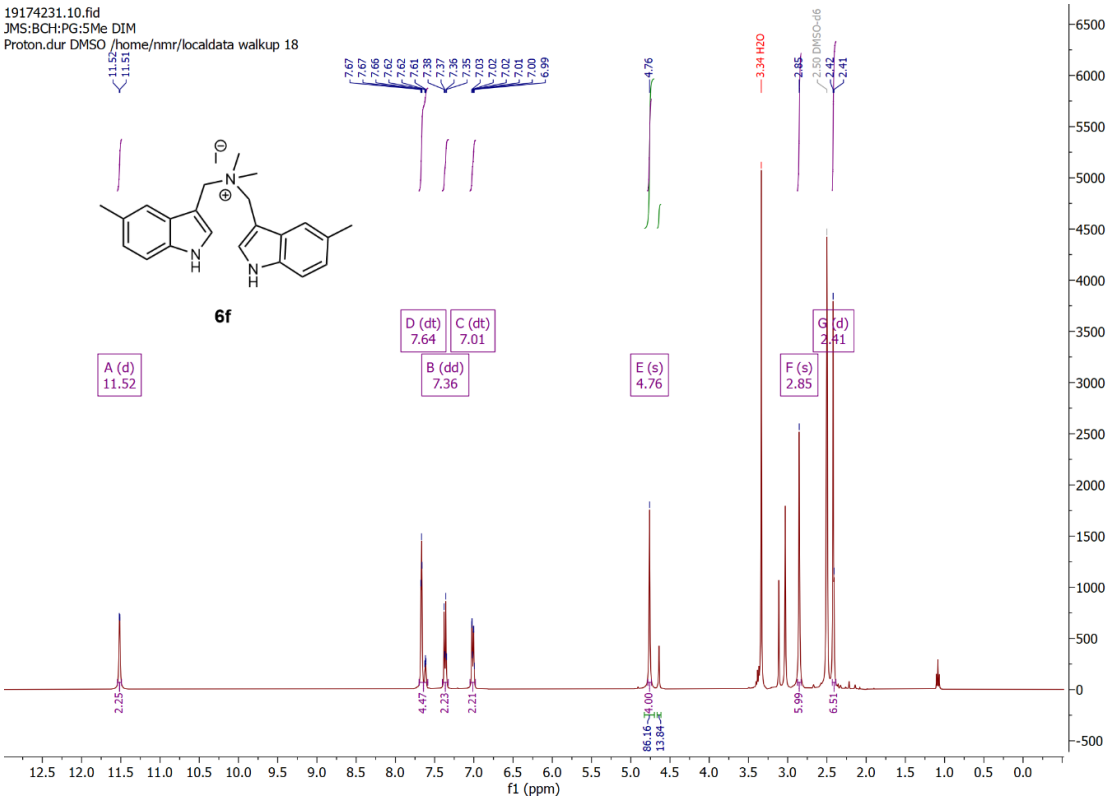
BCRH_5I_dimer_212068

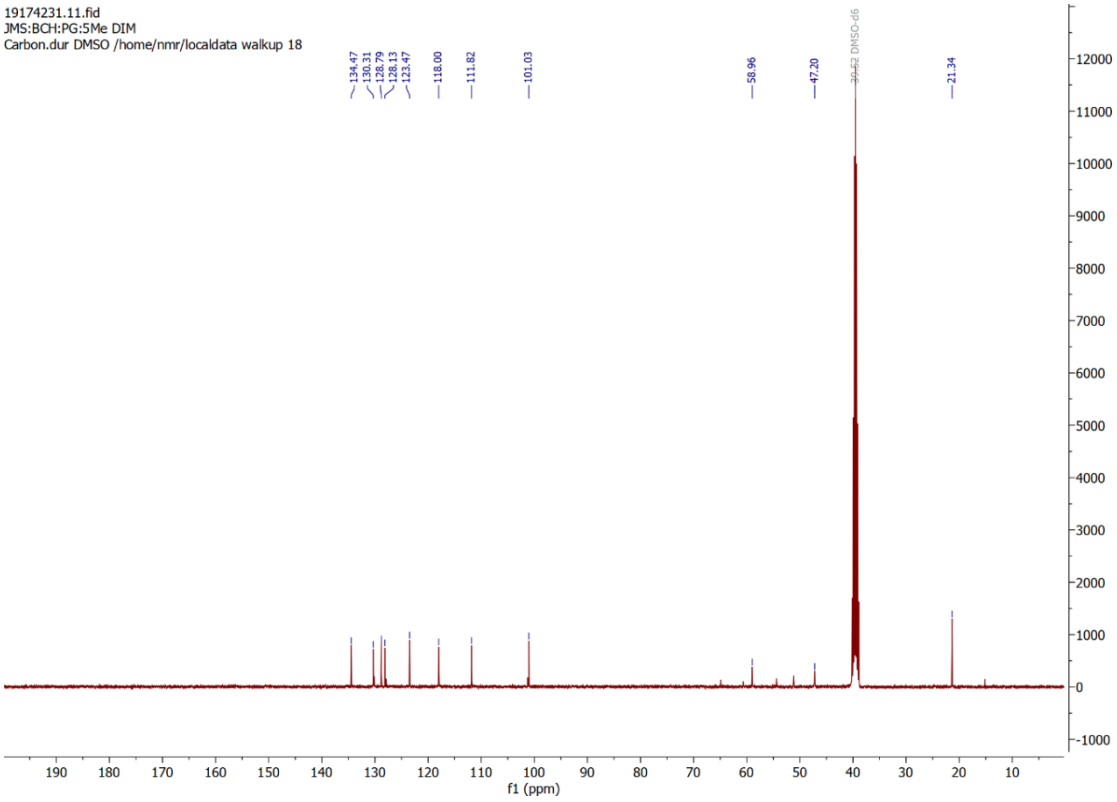
RT: 2.1932 minutes, Scan 247, 1: MS ES+ c (100.0-2000.0), NL 1.10e+8





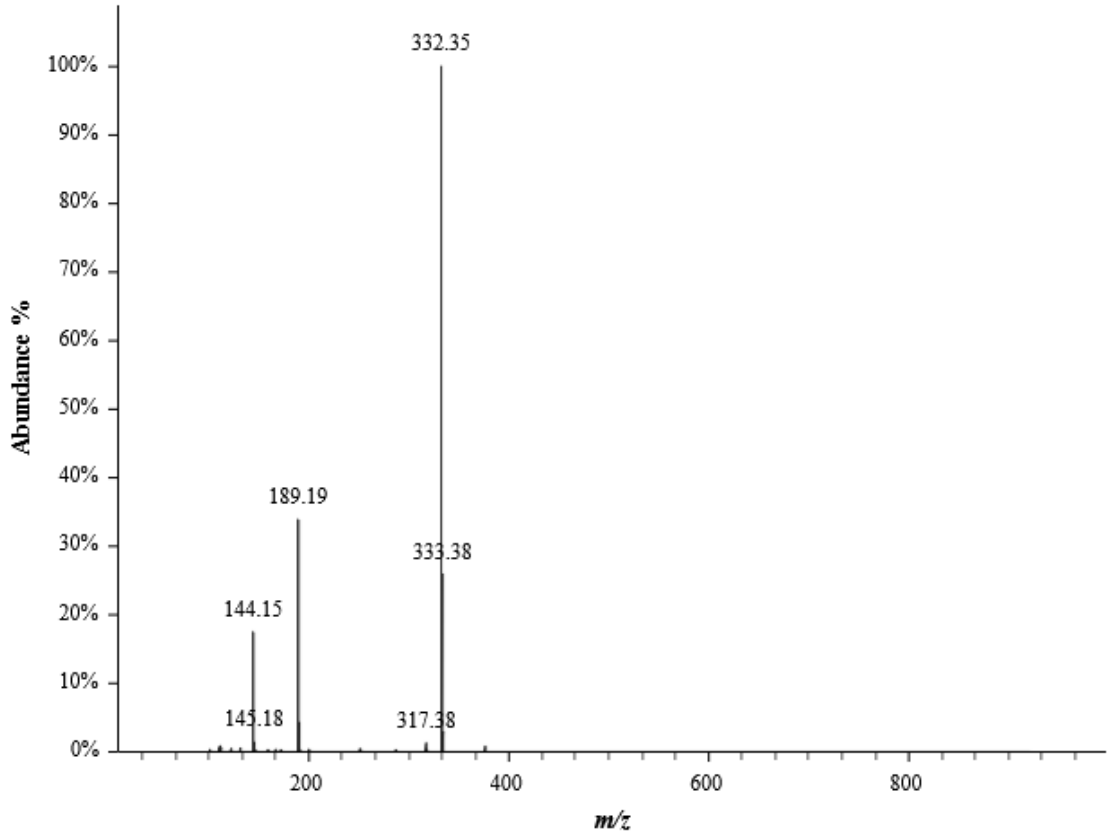
B.6.f. 1-(5-Methyl-1H-indol-3-yl)-N-((5-methyl-1H-indol-3-yl)methyl)-N,N-dimethylmethanaminium iodide (**6f**)

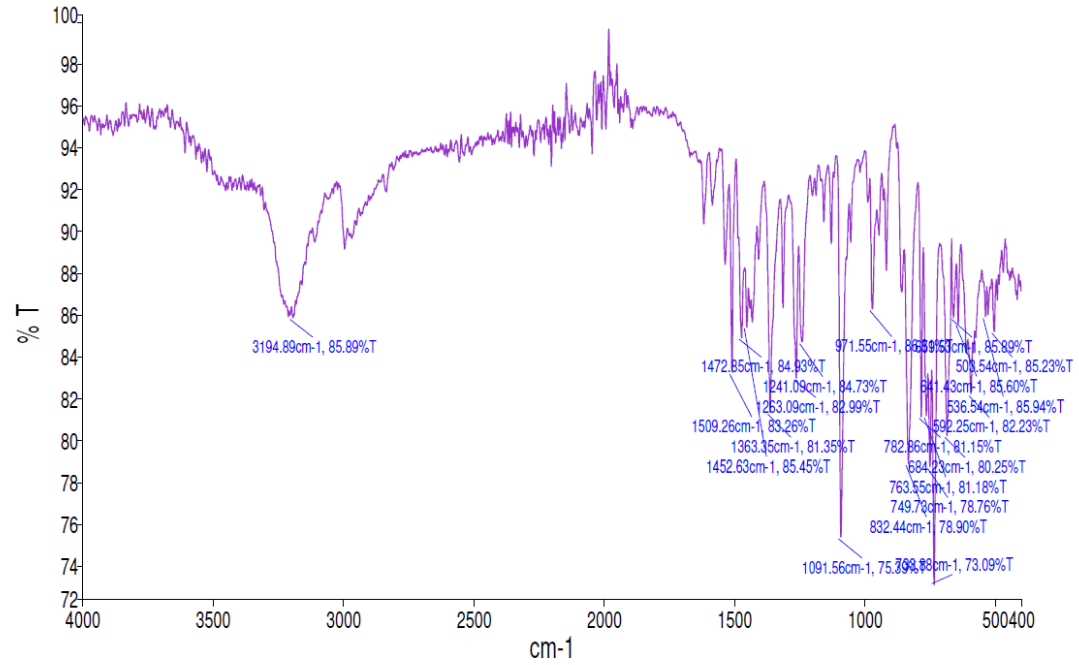




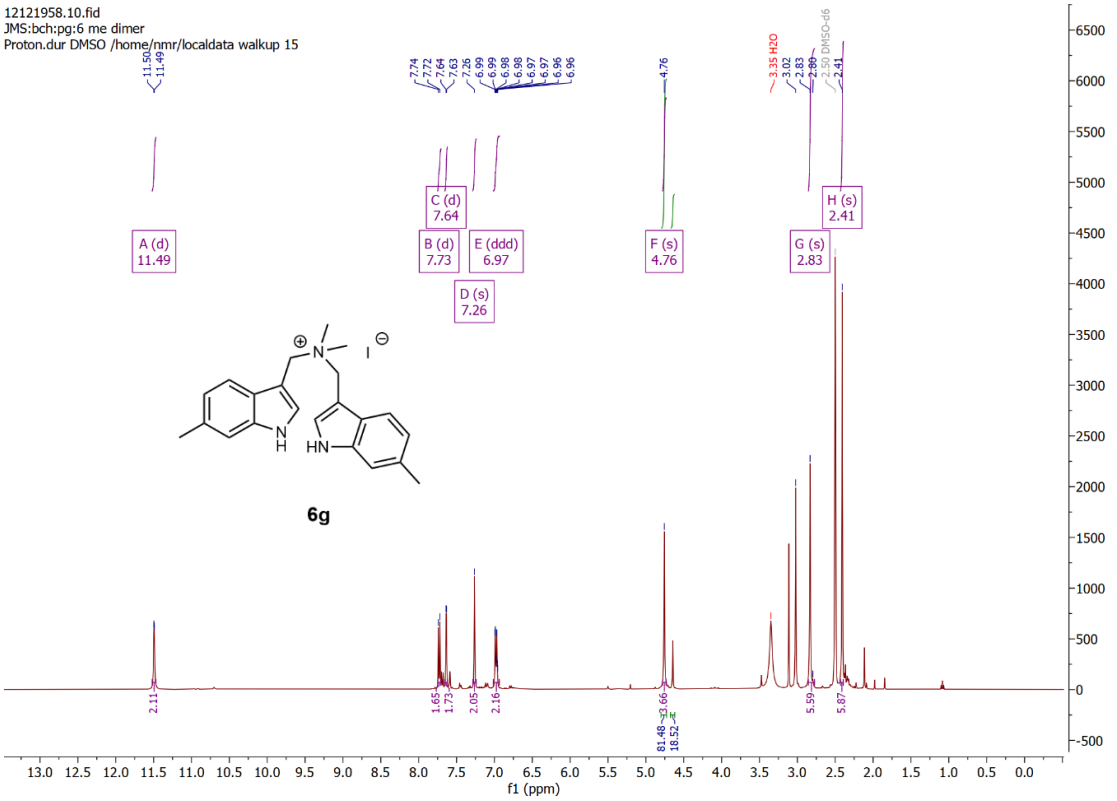
BCRH_5MeDimer_210507

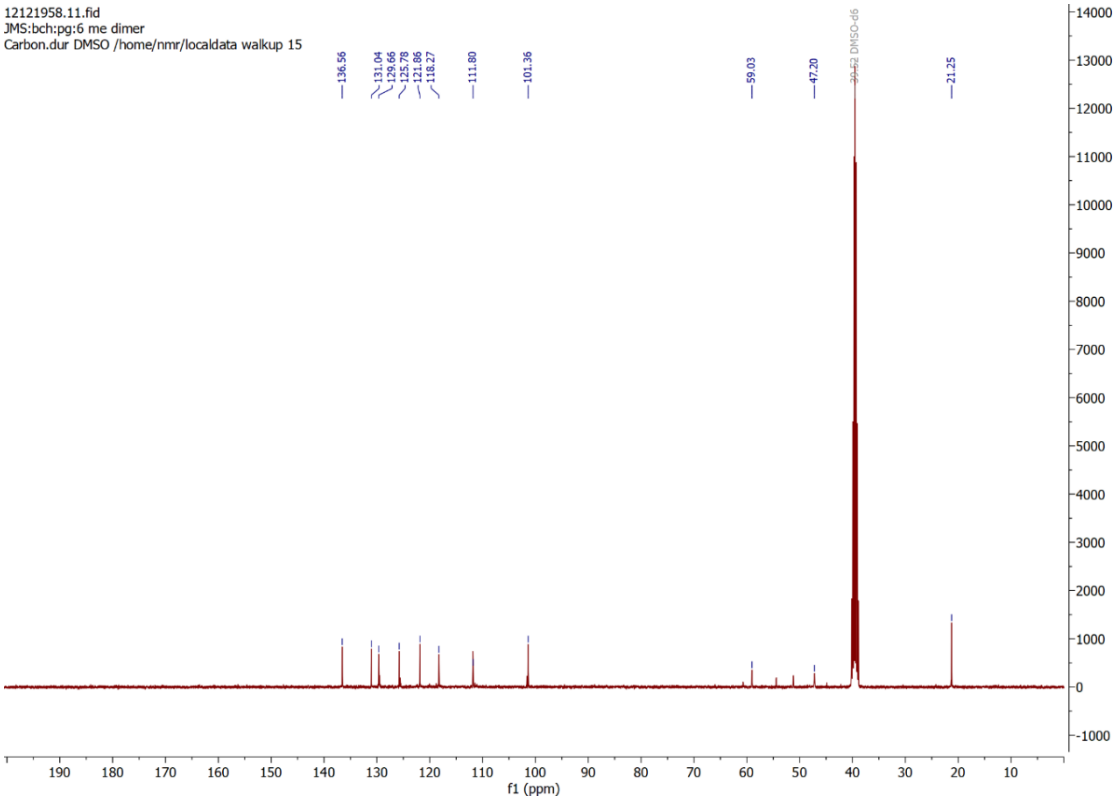
RT: 1.8808 minutes, Scan 211, 1: MS ES+ c (100.0-2000.0), NL 6.61e+7





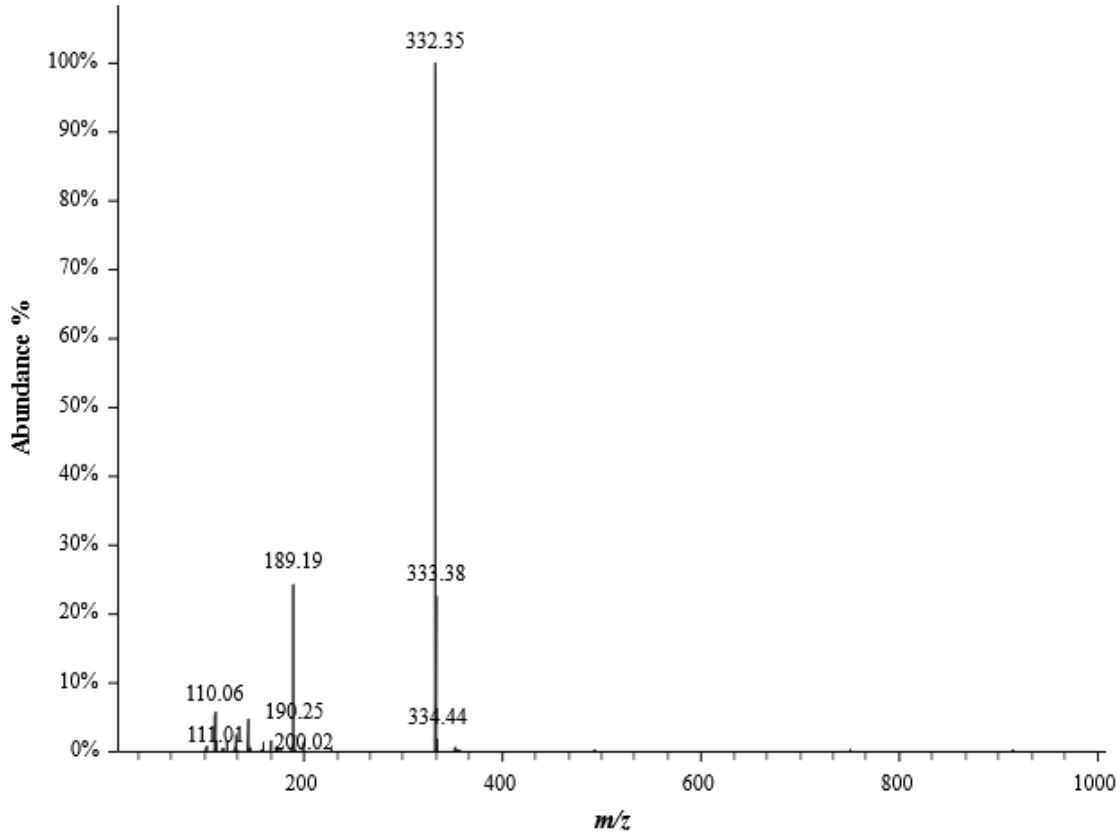
B.6.g. N,N-dimethyl-1-(6-methyl-1H-indol-3-yl)-N-((6-methyl-1H-indol-3-yl)methyl)methanaminium iodide (**6g**)

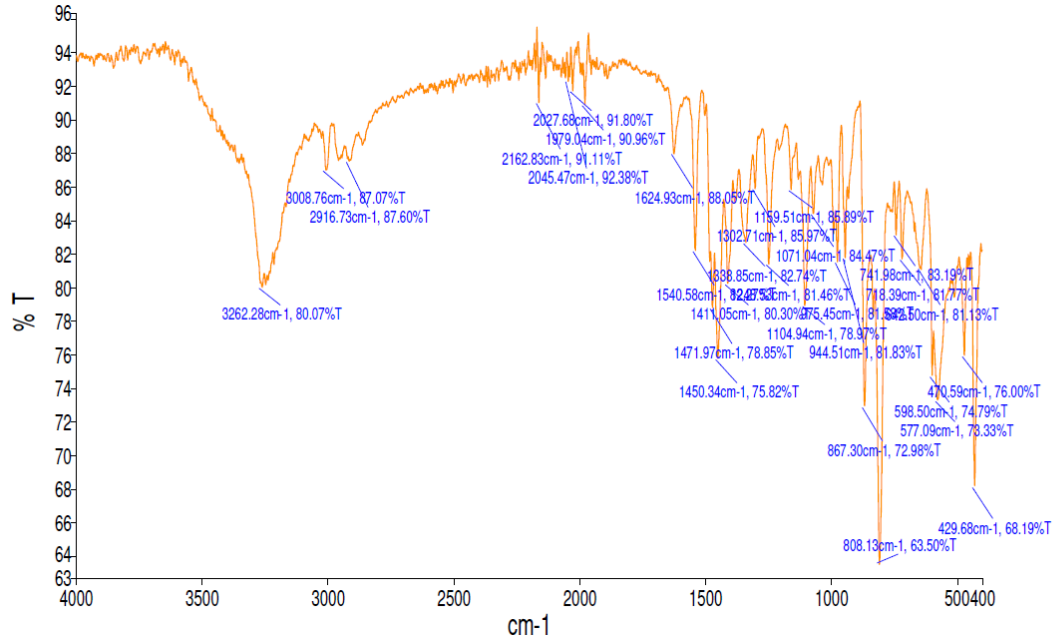




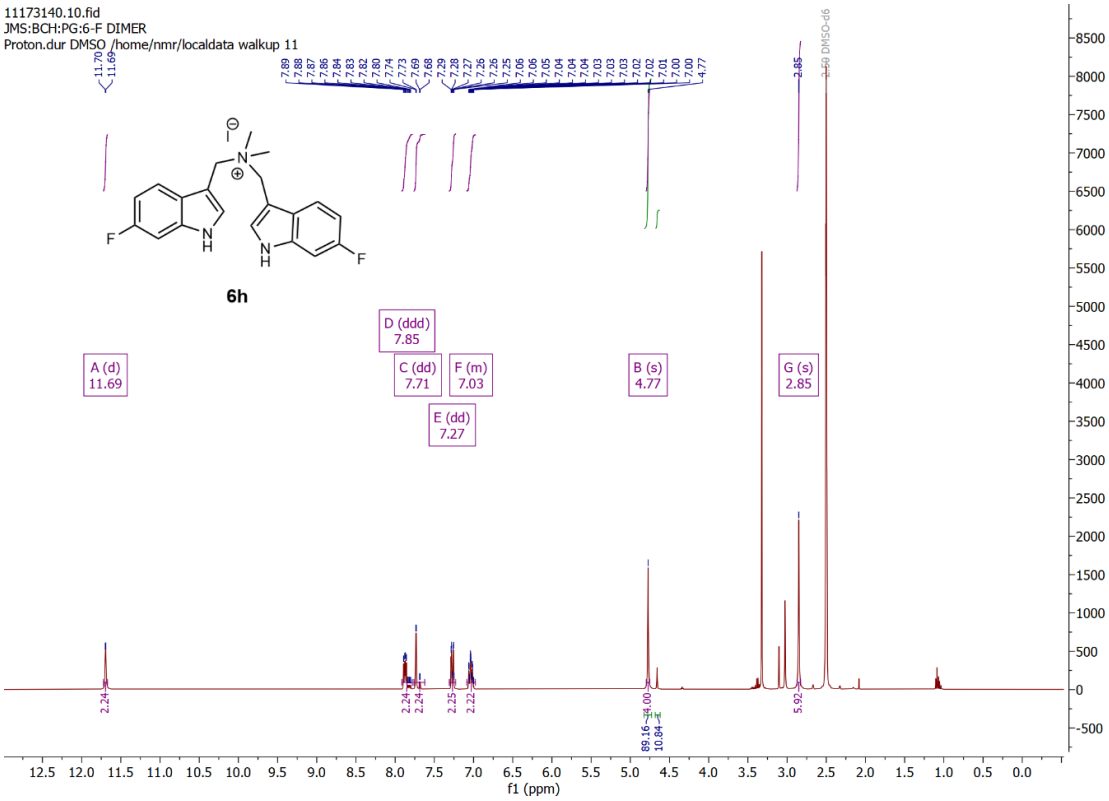
BCRH_6MeDimer_210393

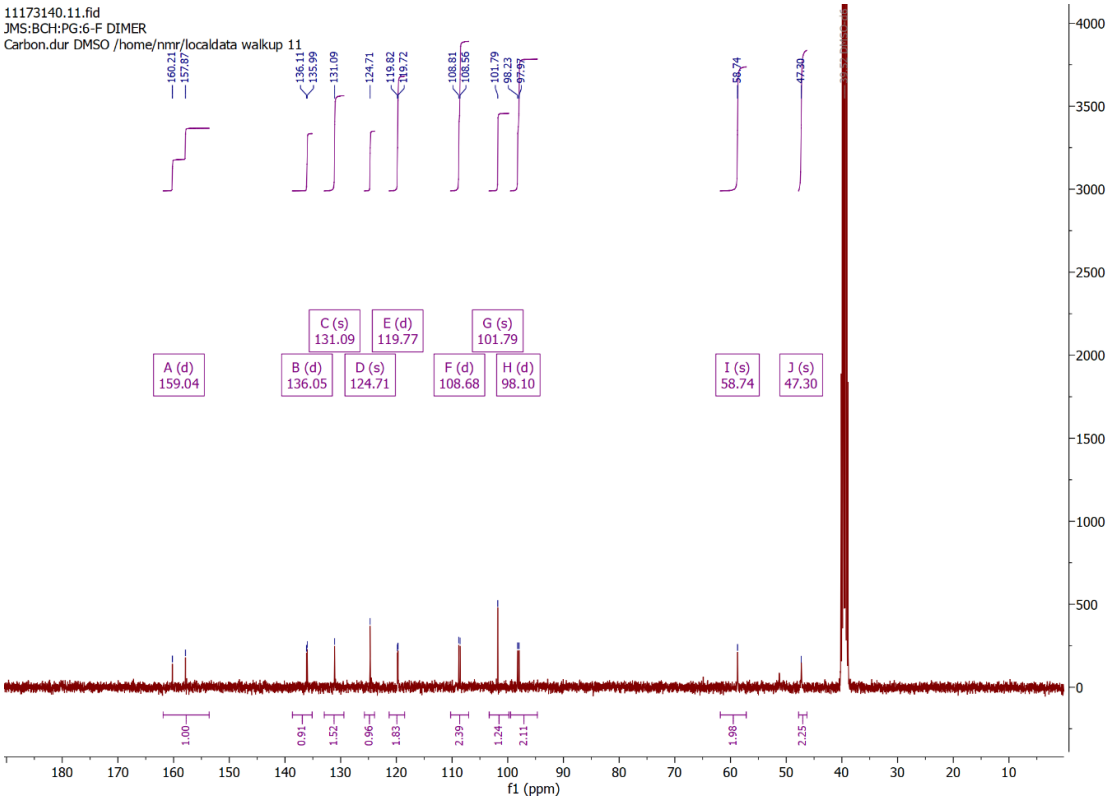
RT: 1.9155 minutes, Scan 215, 1: MS ES+ c (100.0-2000.0), NL 3.45e+7

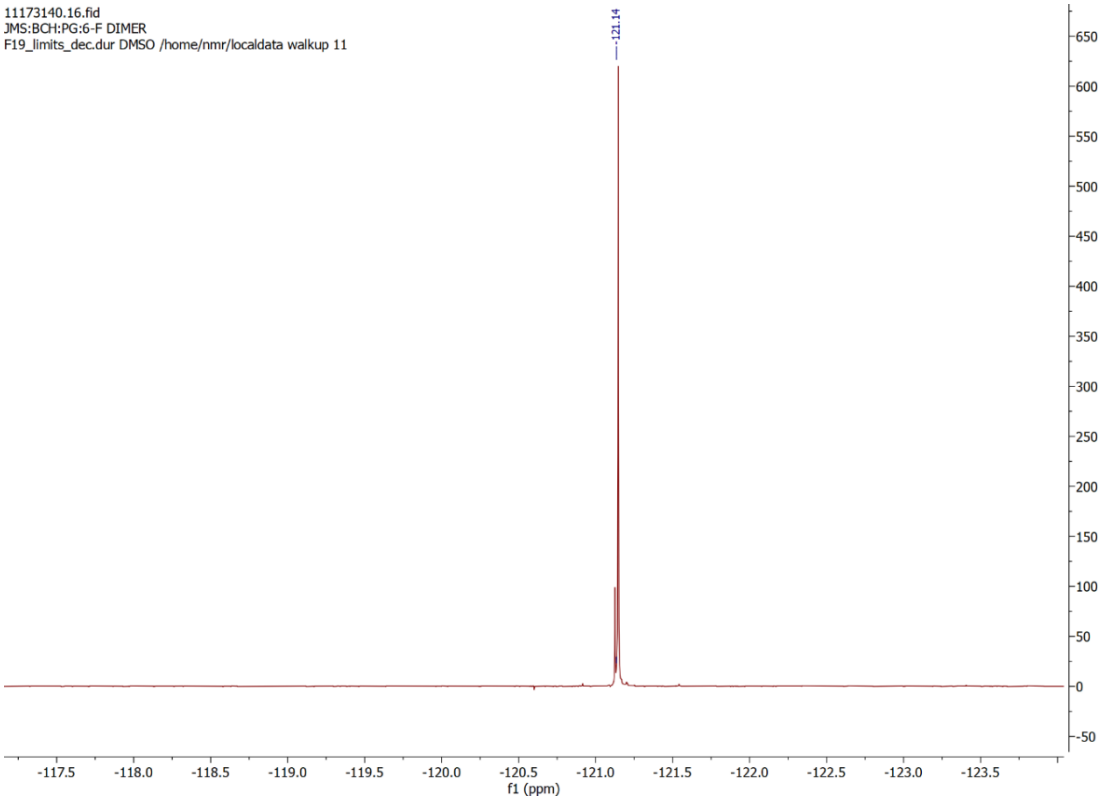




B.6.h. 1-(6-fluoro-1H-indol-3-yl)-N-((6-fluoro-1H-indol-3-yl)methyl)-N,N-dimethylmethanaminium iodide (**6h**)

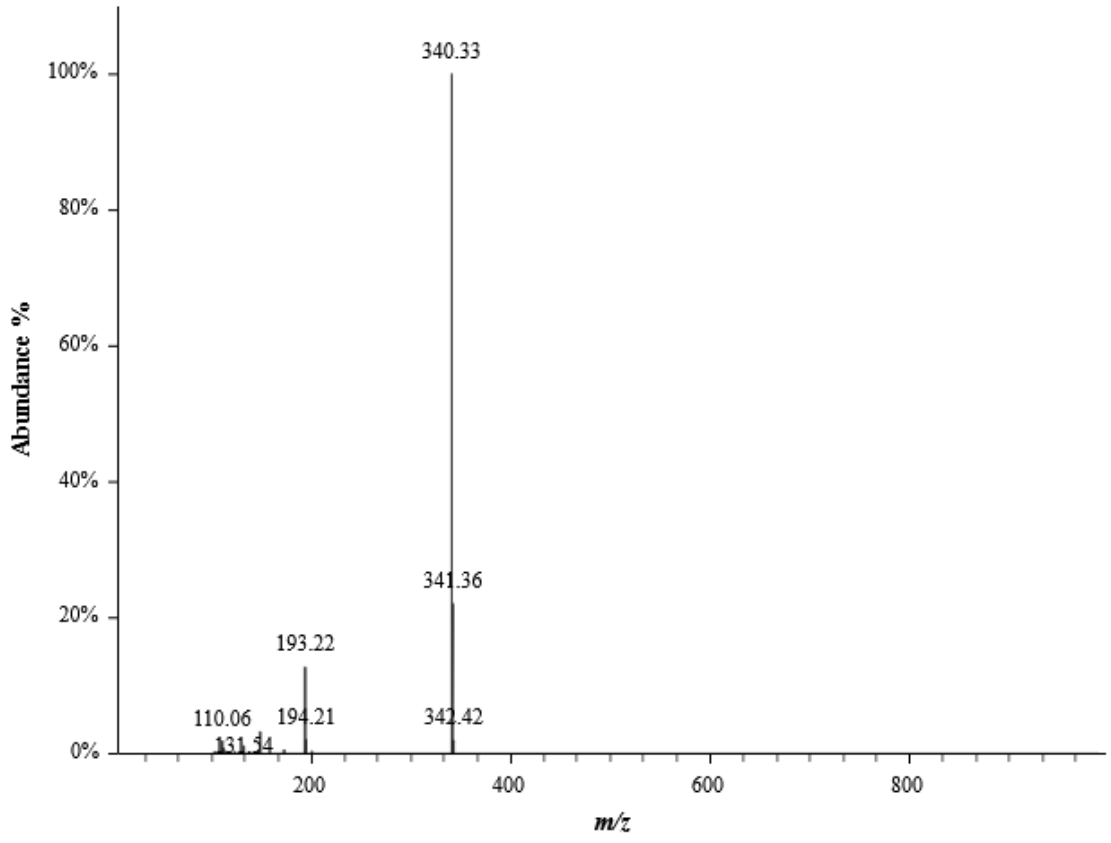


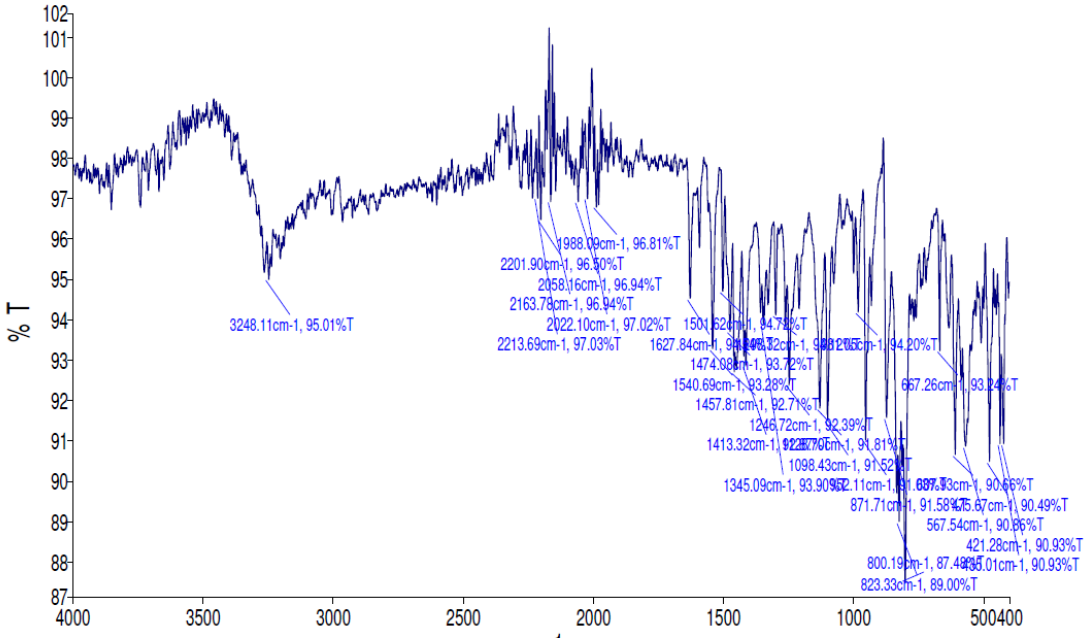




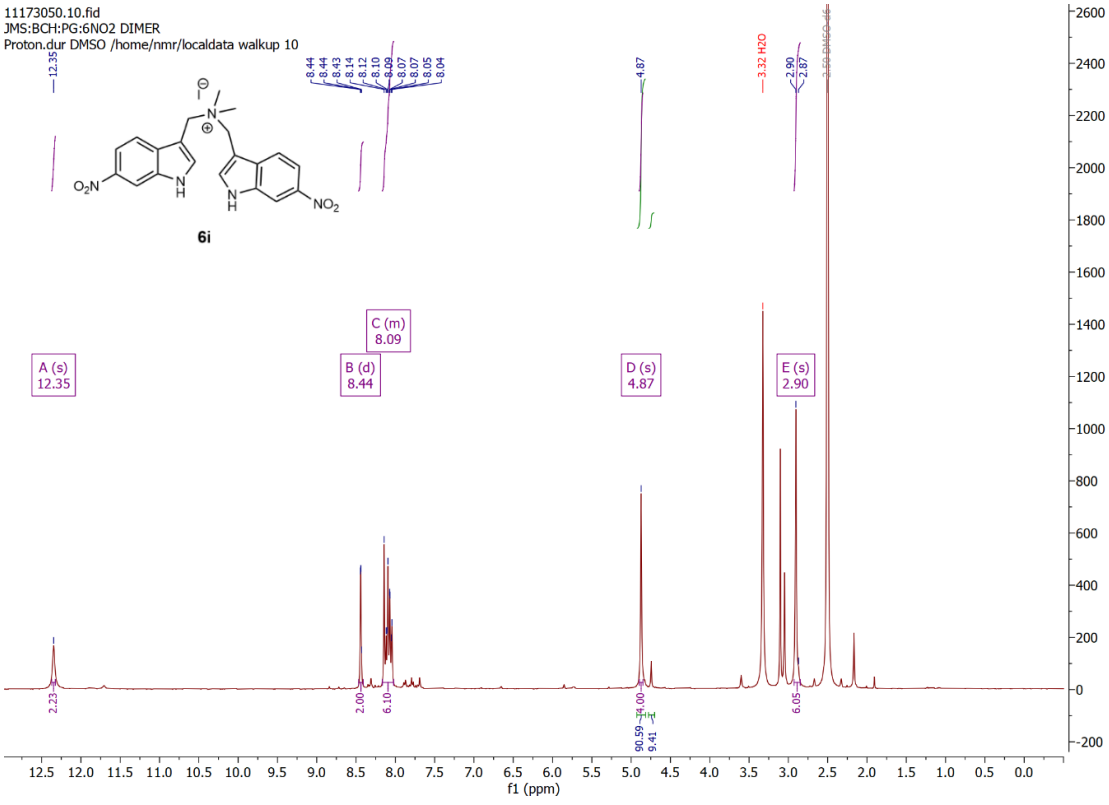
BCRH_6FDimer_210508

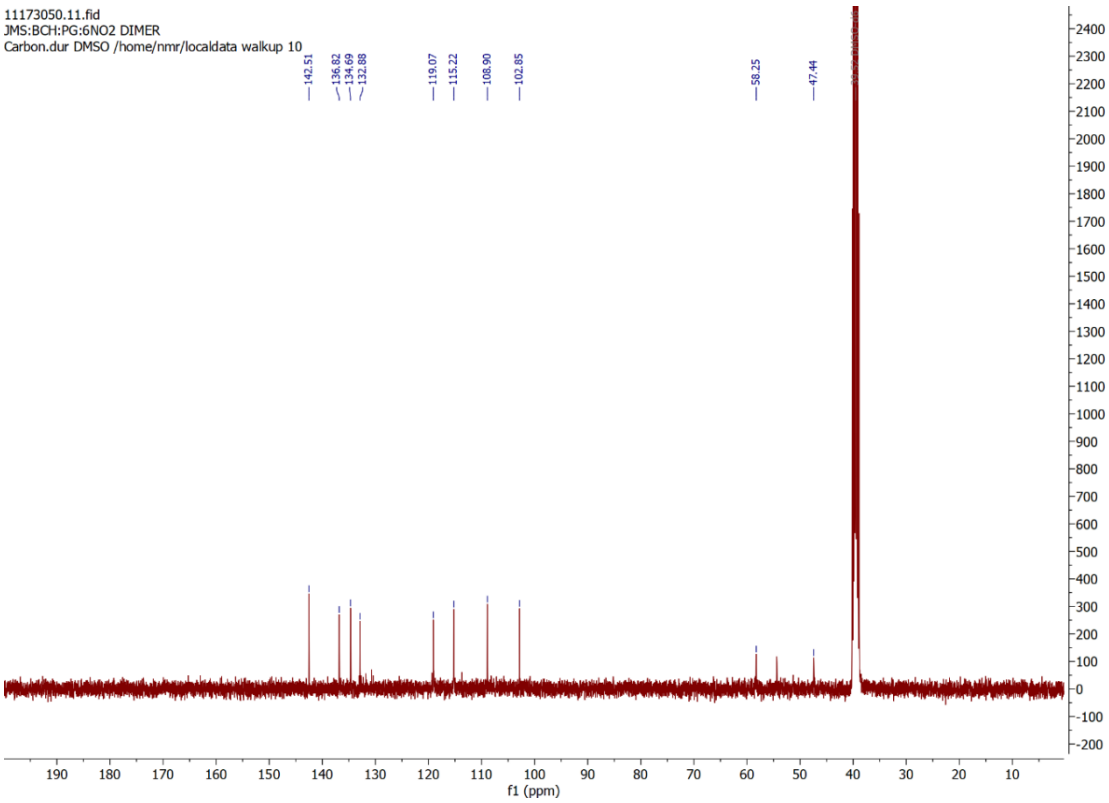
RT: 1.7246 minutes, Scan 193, 1: MS ES+ c (100.0-2000.0), NL 6.56e+7





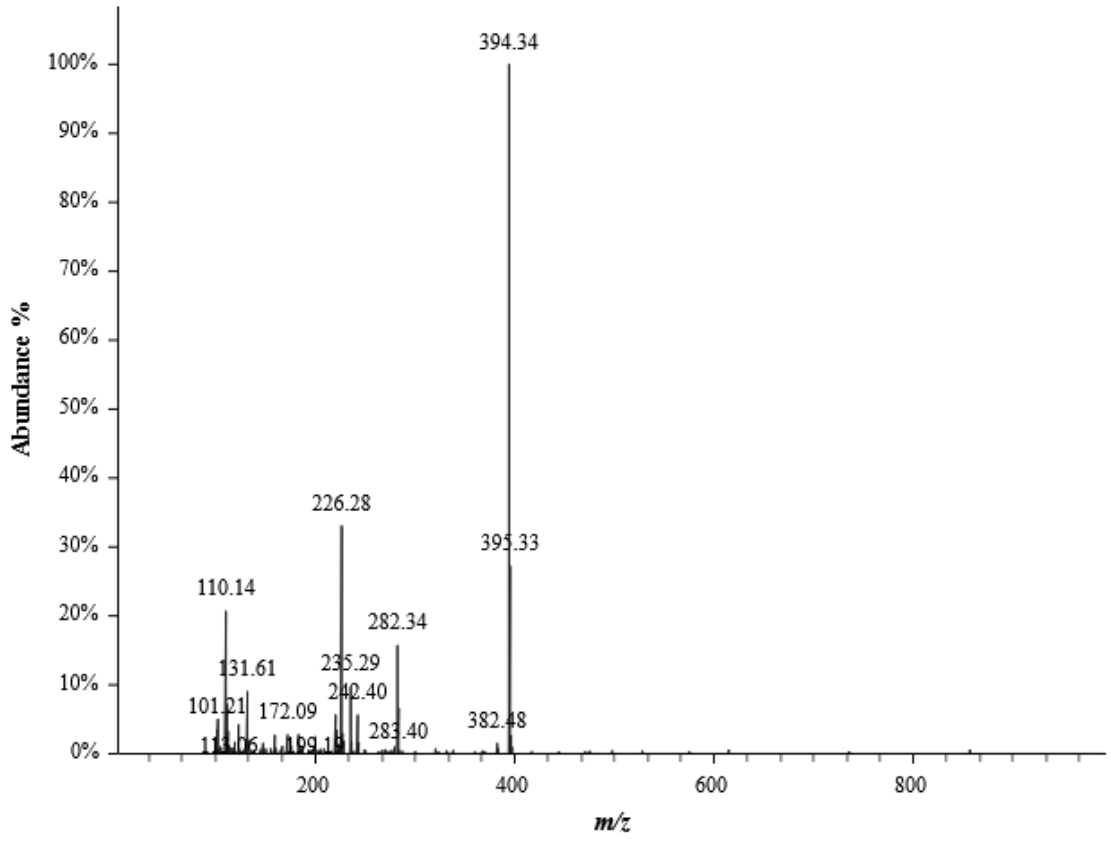
B.6.i. 1-(6-Nitro-1H-indol-3-yl)-N-((6-Nitro-1H-indol-3-yl)methyl)-N,N-dimethylmethanaminium iodide (**6i**)

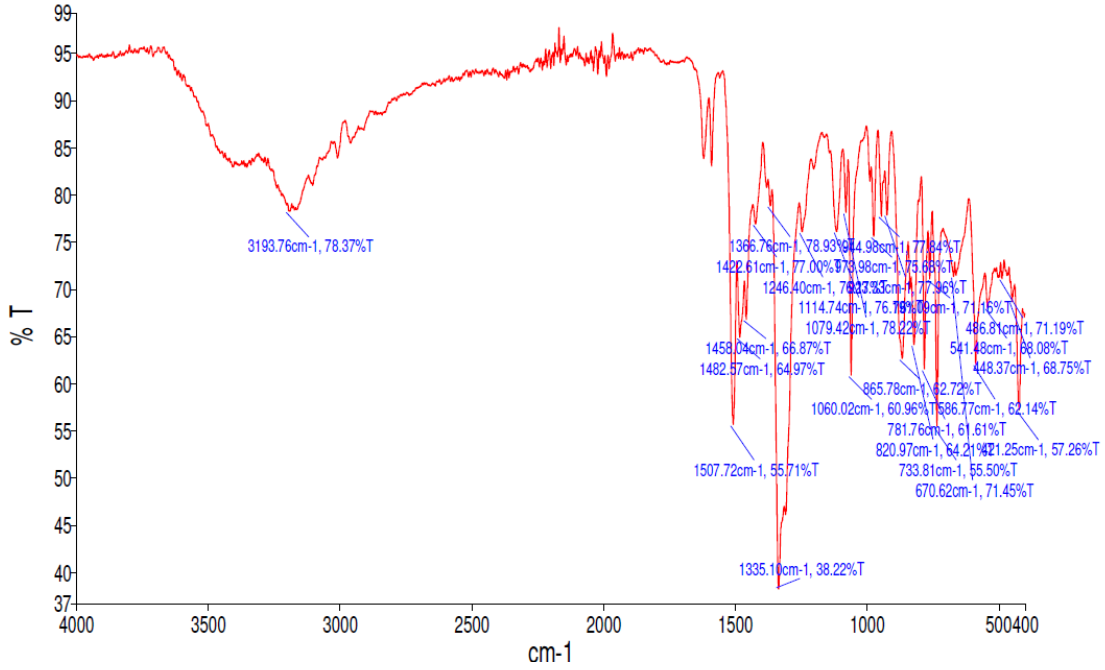




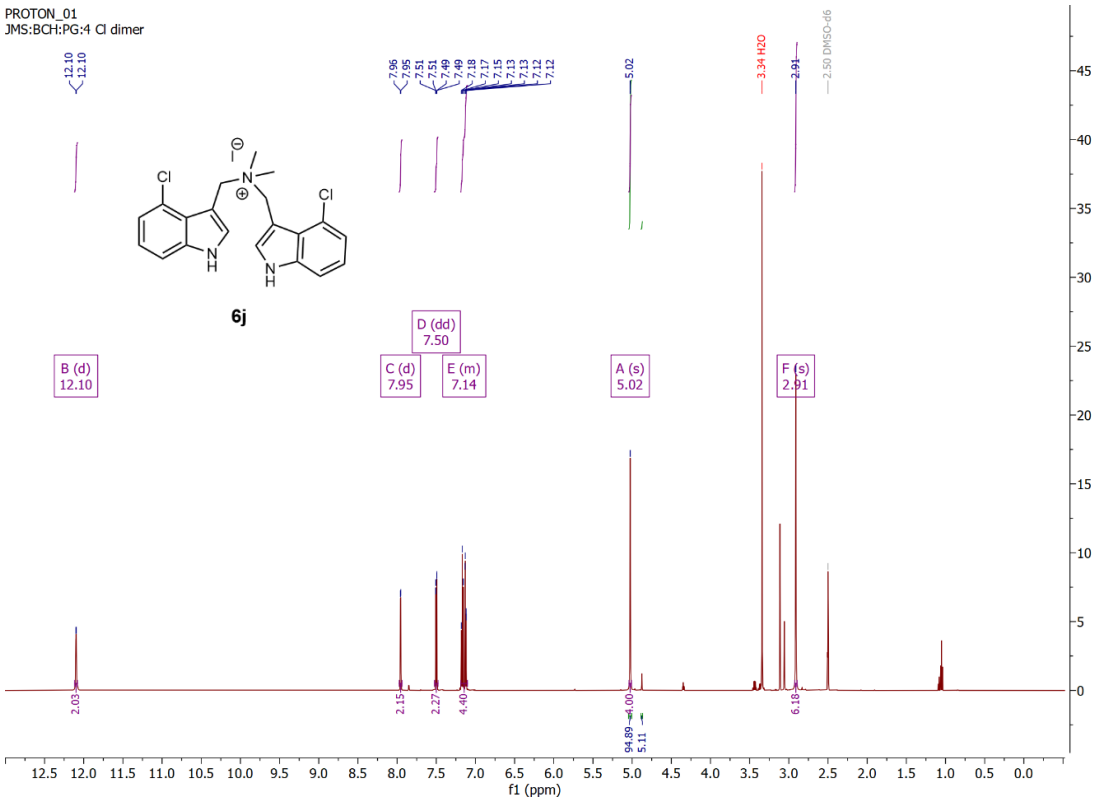
BCRH_6NO2_Dimer_212067

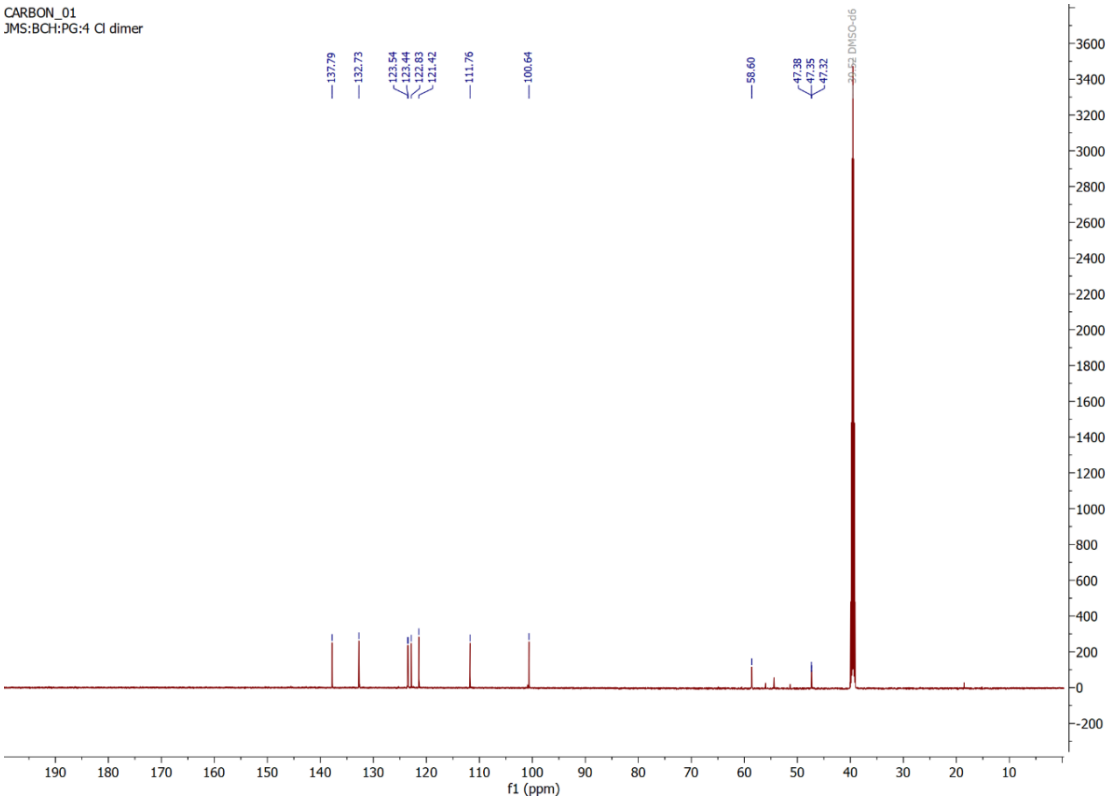
RT: 1.6898 minutes, Scan 189, 1: MS ES+ c (100.0-2000.0), NL 1.07e+7





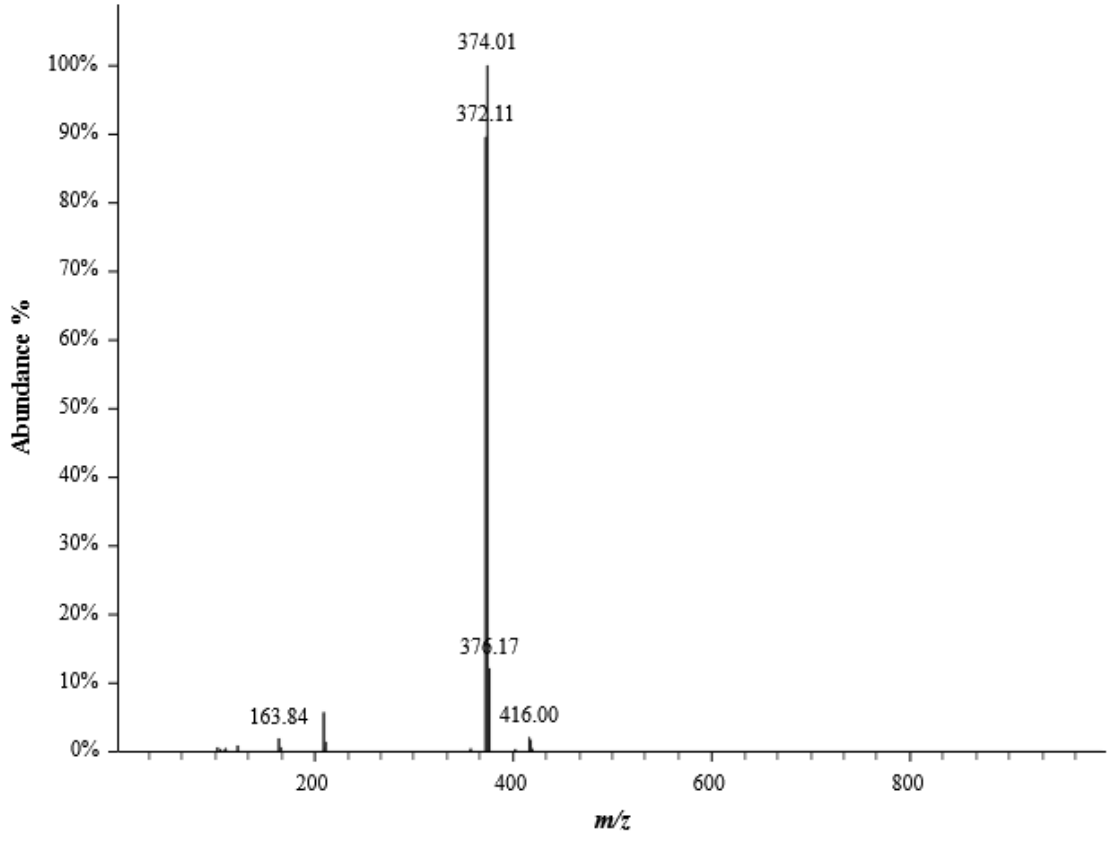
B.6.j. 1-(4-chloro-1H-indol-3-yl)-N-((4-chloro-1H-indol-3-yl)methyl)-N,N-dimethylmethanaminium iodide (**6j**)

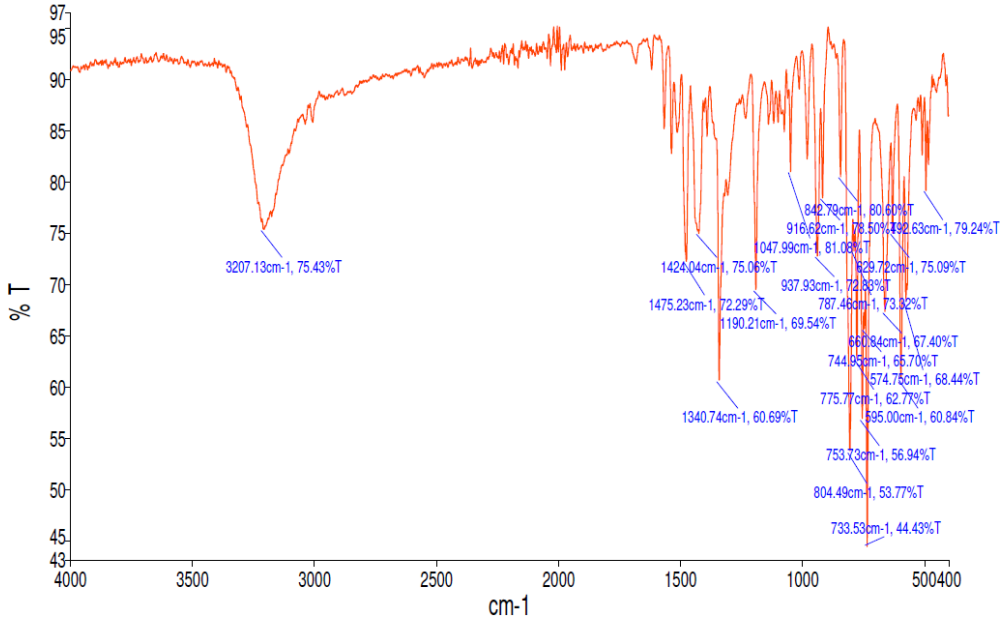




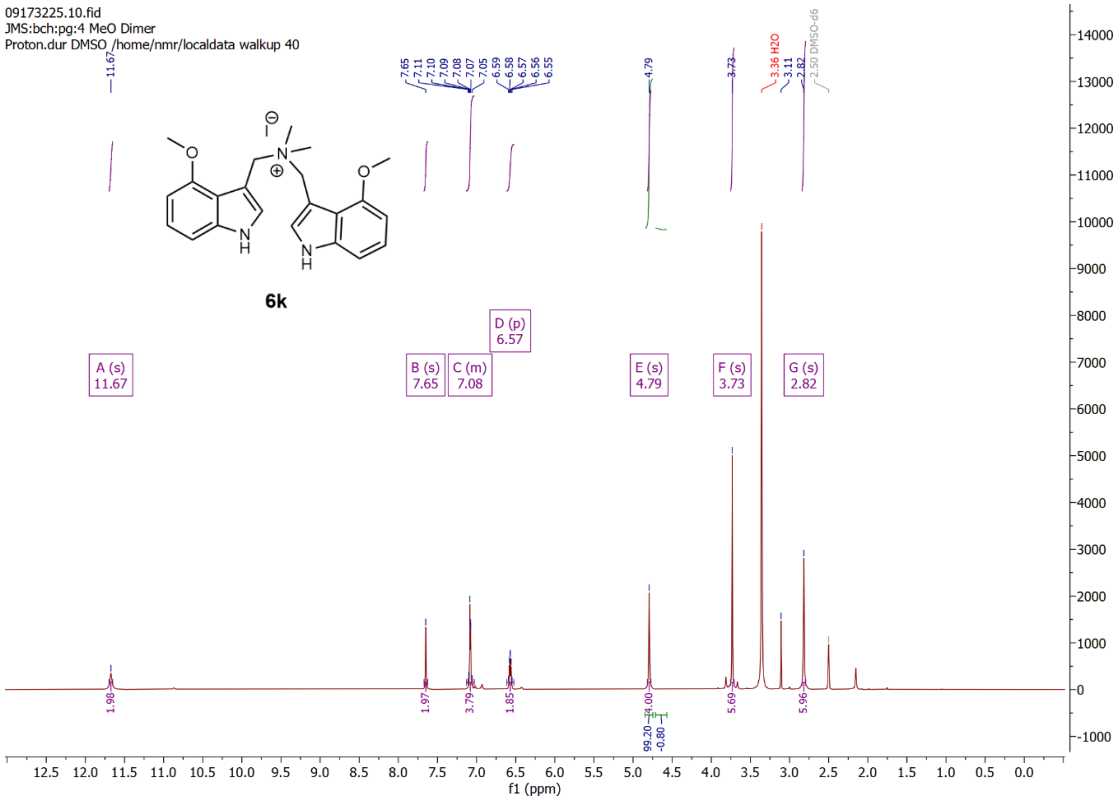
BCRH_4_Cl_dimer_222787

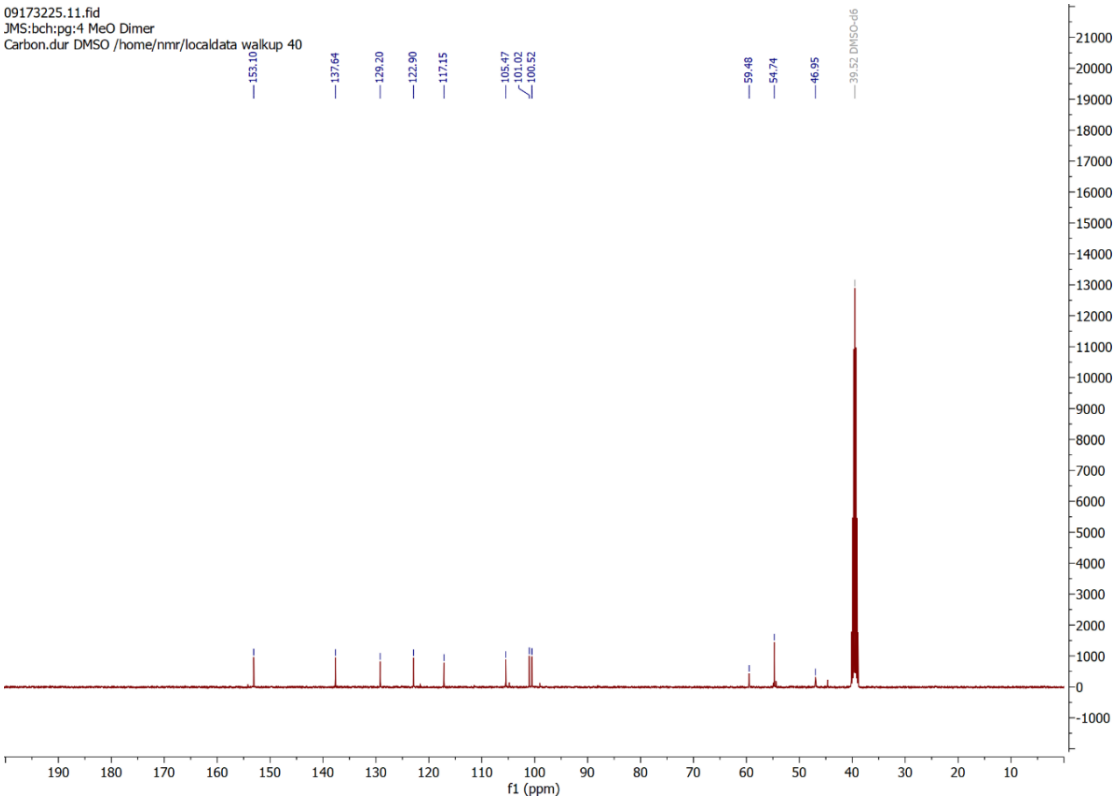
RT: 2.0196 minutes, Scan 227, 1: MS ES+ c (100.0-2000.0), NL 2.15e+8





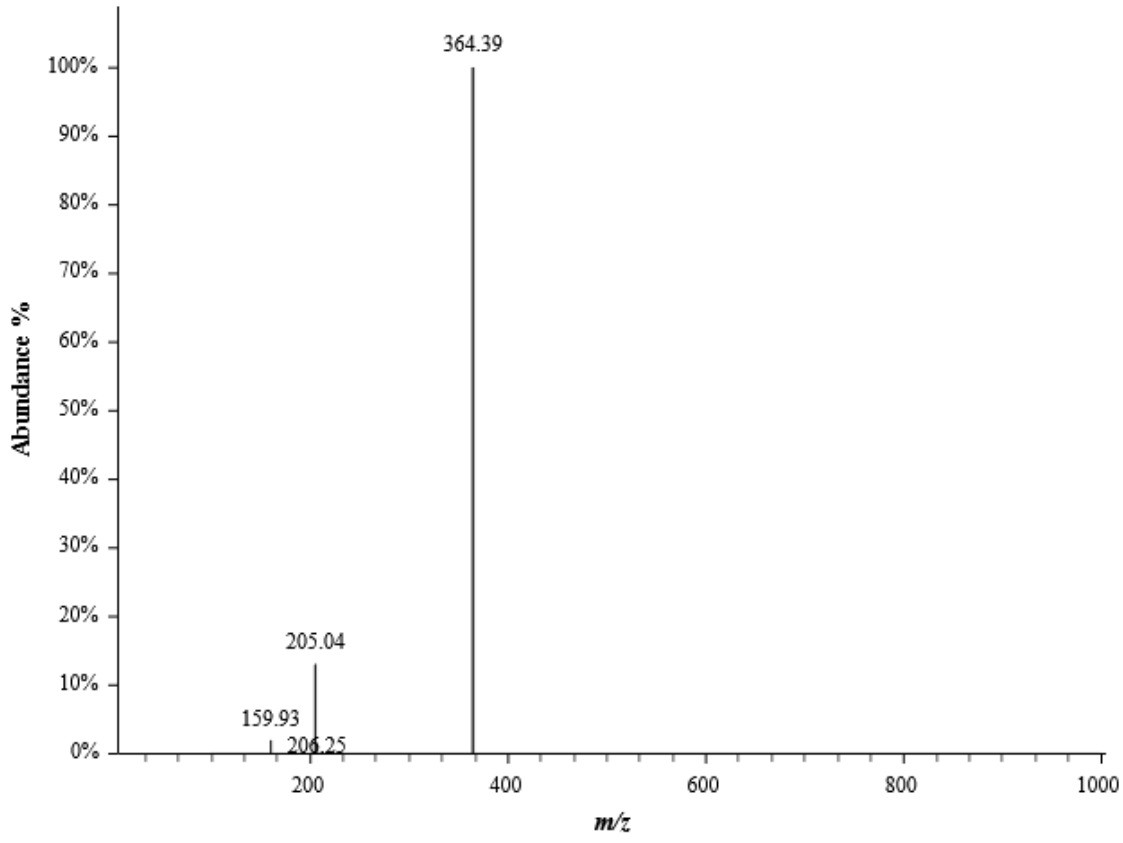
B.6.k. 1-(4-Methoxy-1H-indol-3-yl)-N-((4-methoxy-1H-indol-3-yl)methyl)-N,N-dimethylmethanaminium (**6k**)

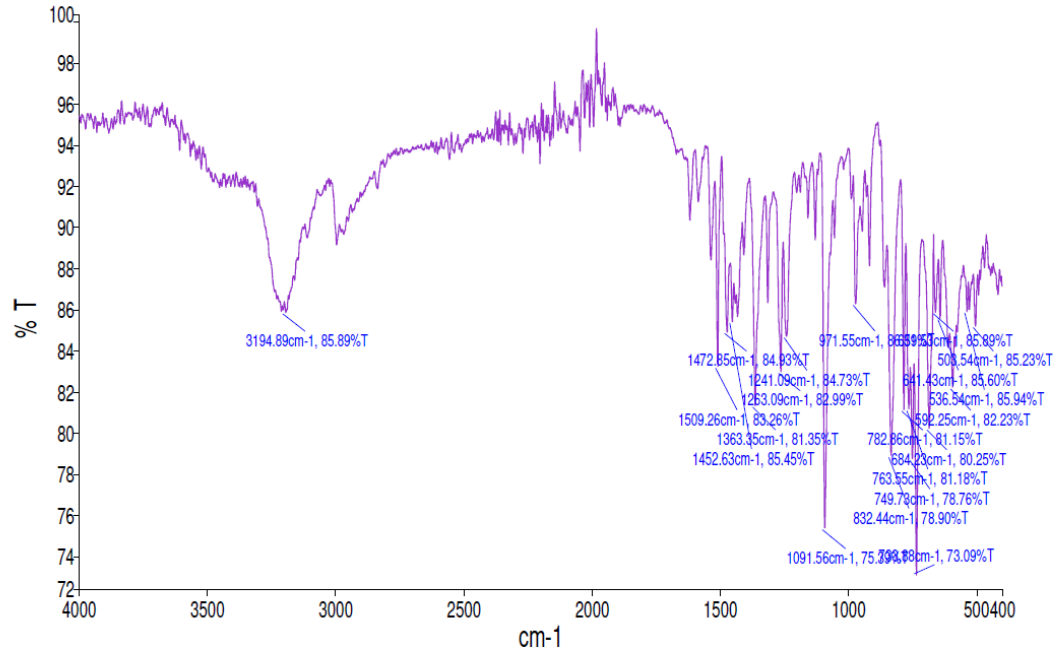




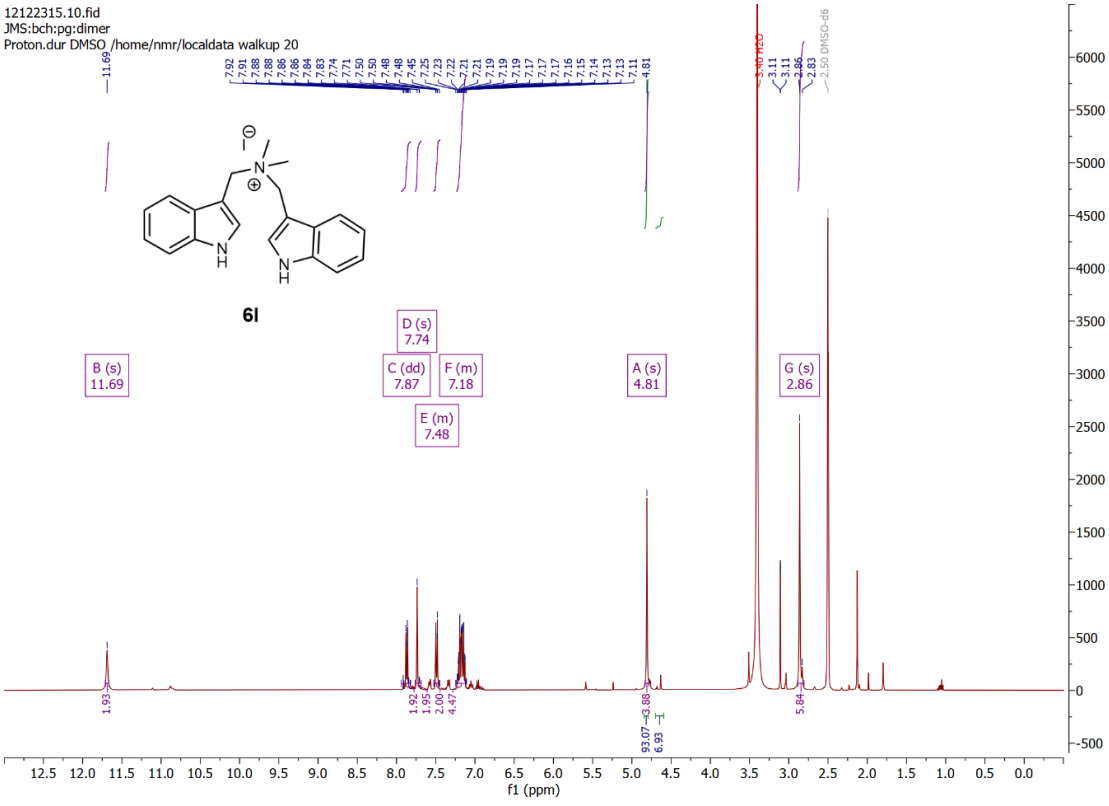
BCRH_4MeO_dimer_222788

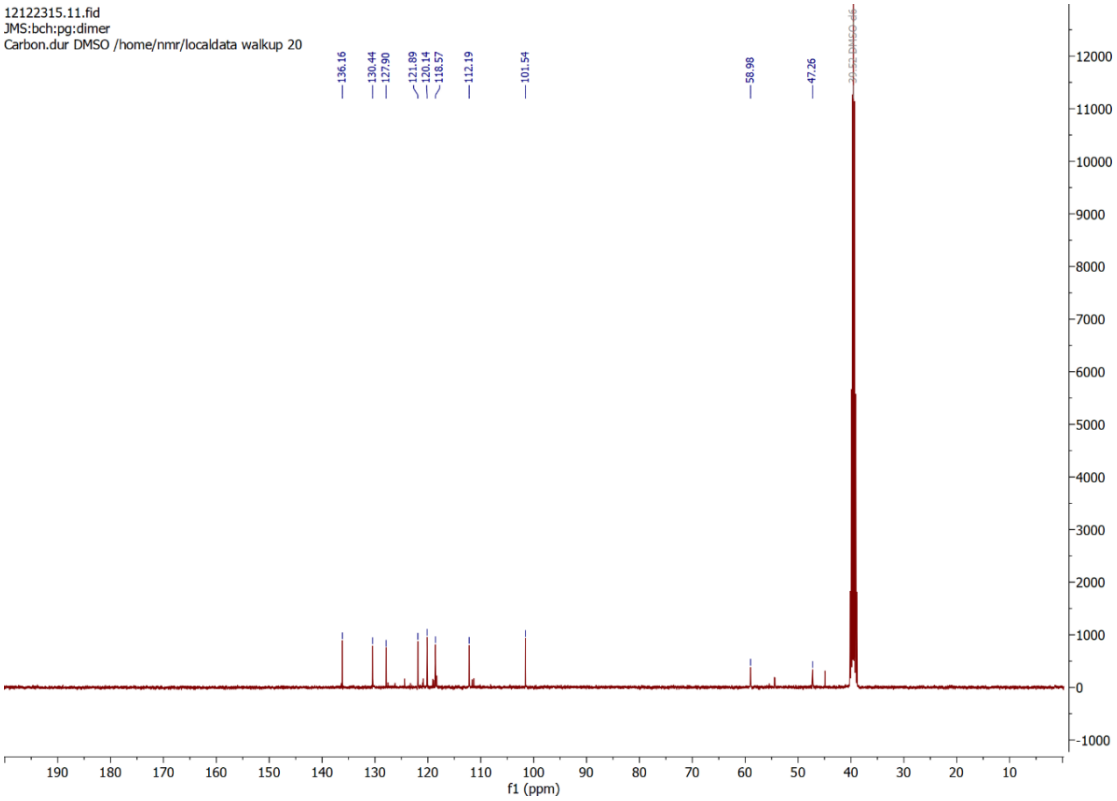
RT: 1.8287 minutes, Scan 205, 1: MS ES+ c (100.0-2000.0), NL 3.58e+8





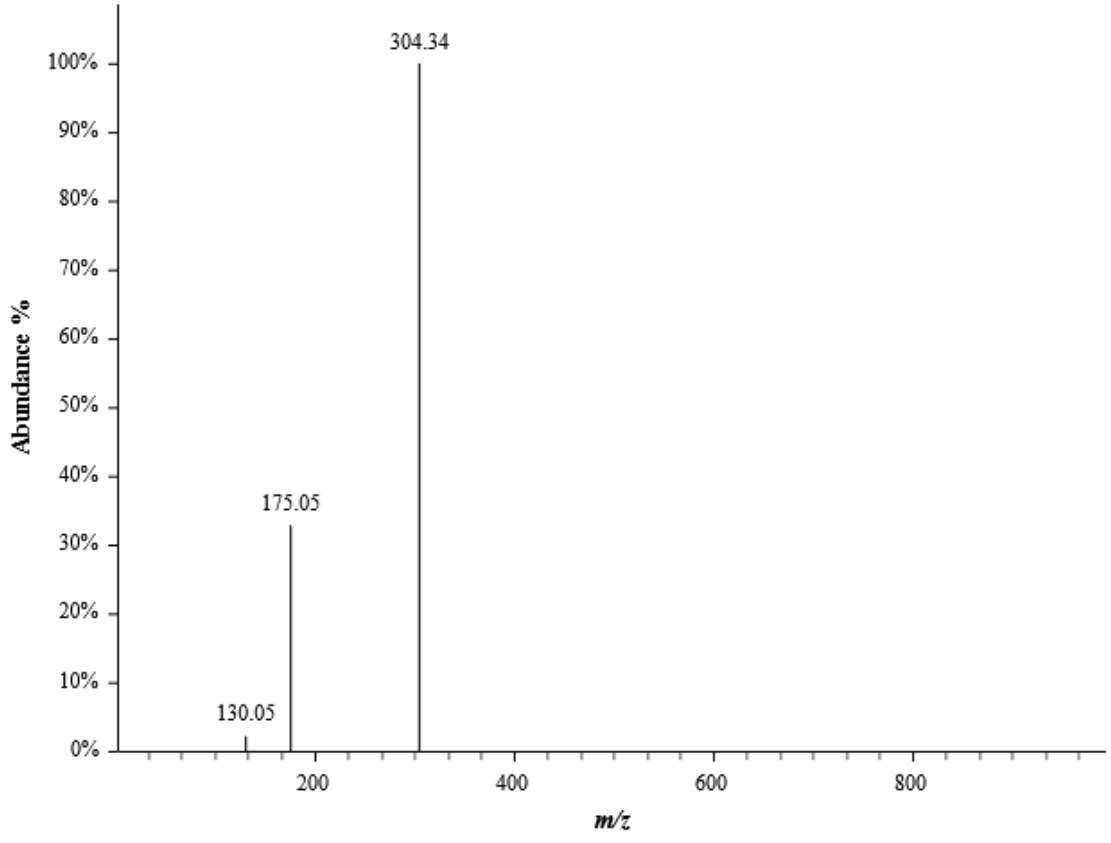
B.6.l. Bis(indol-3-ylmethyl)dimethylammonium iodide (**6l**)

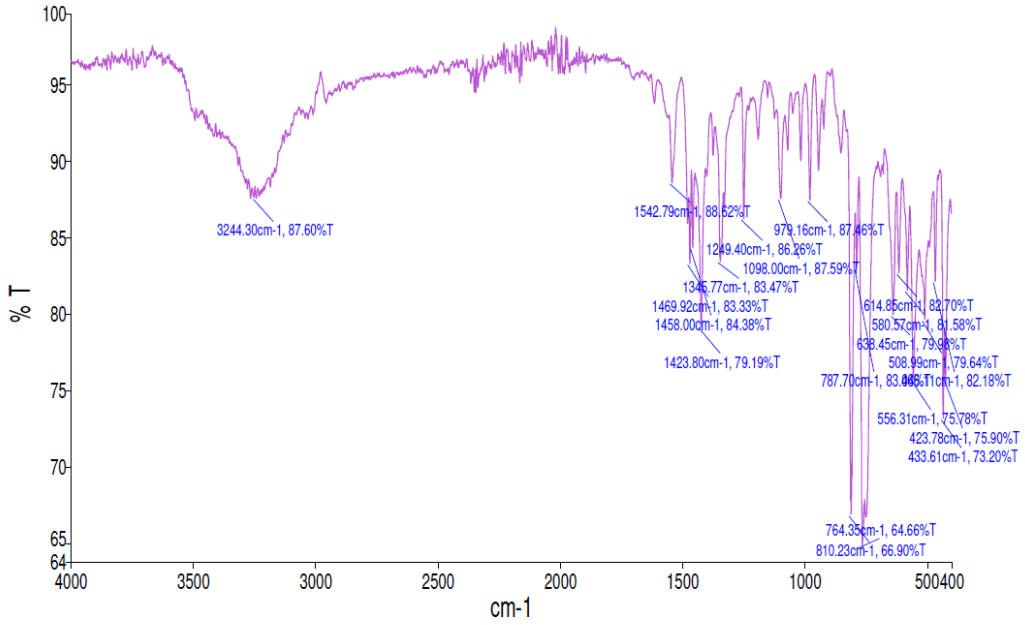




BCRH_dimer_222772

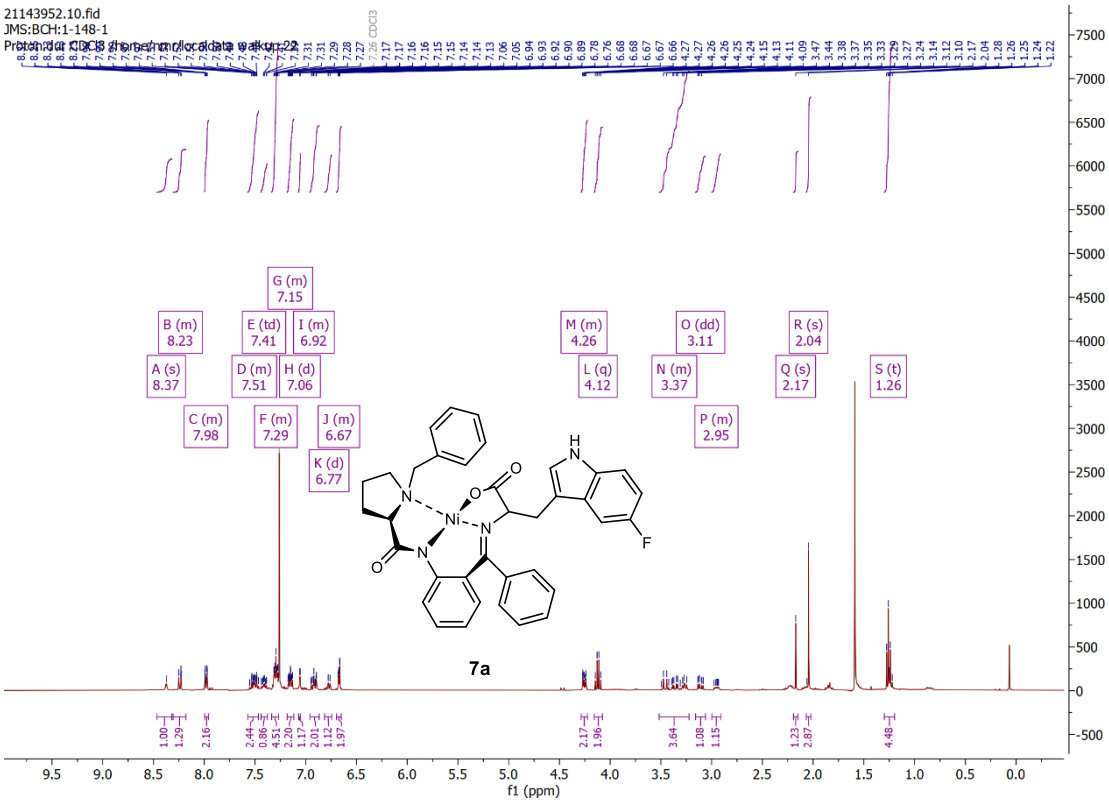
RT: 1.6030 minutes, Scan 179, 1: MS ES+ c (100.0-2000.0), NL 3.33e+8





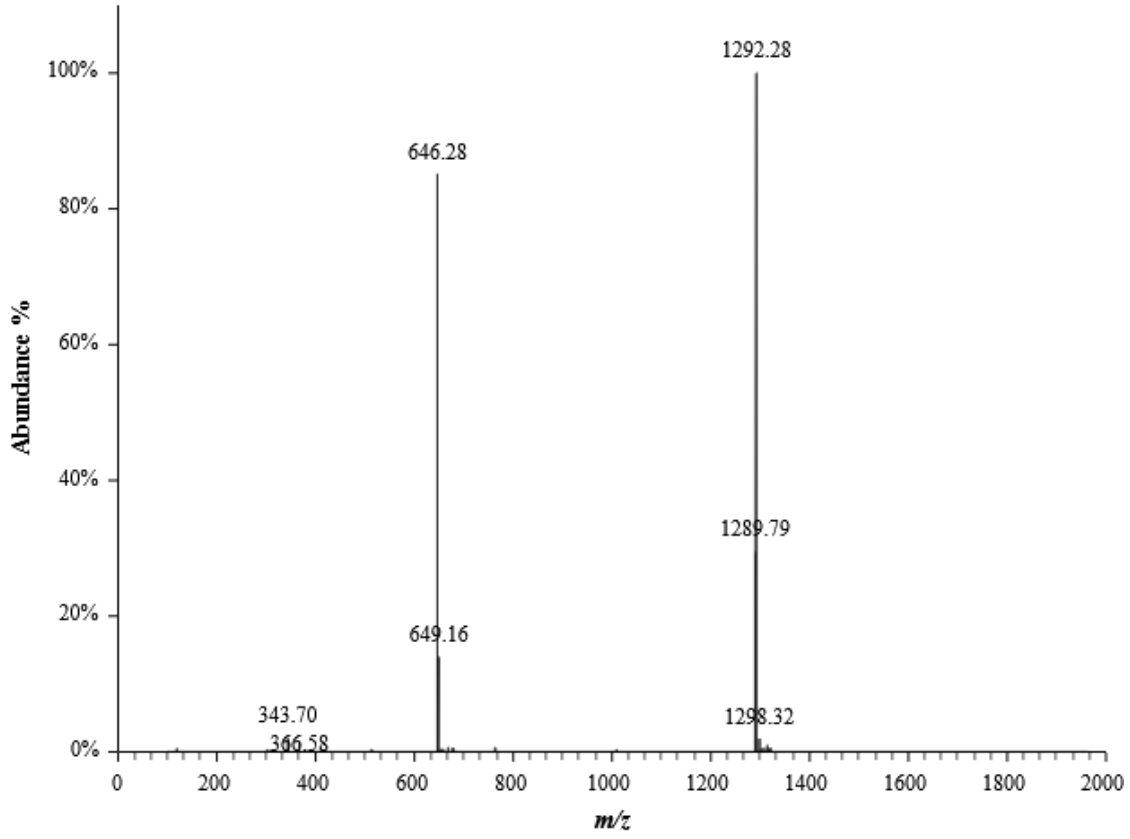
B.7. Indole BPB nickel complexes

B.7.a. 5-Fluoro-tryptophan-Ni-(S)-2-[N-(N-benzylpropyl) amino]benzophenone (**7a**)

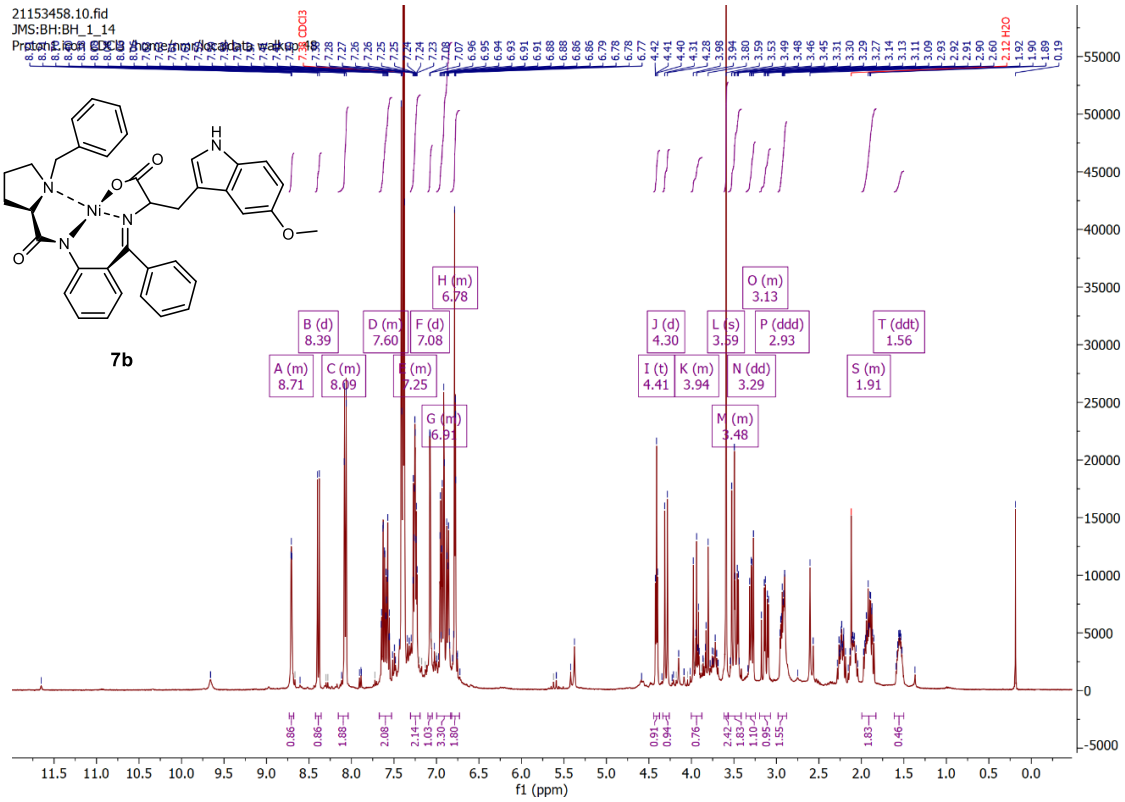


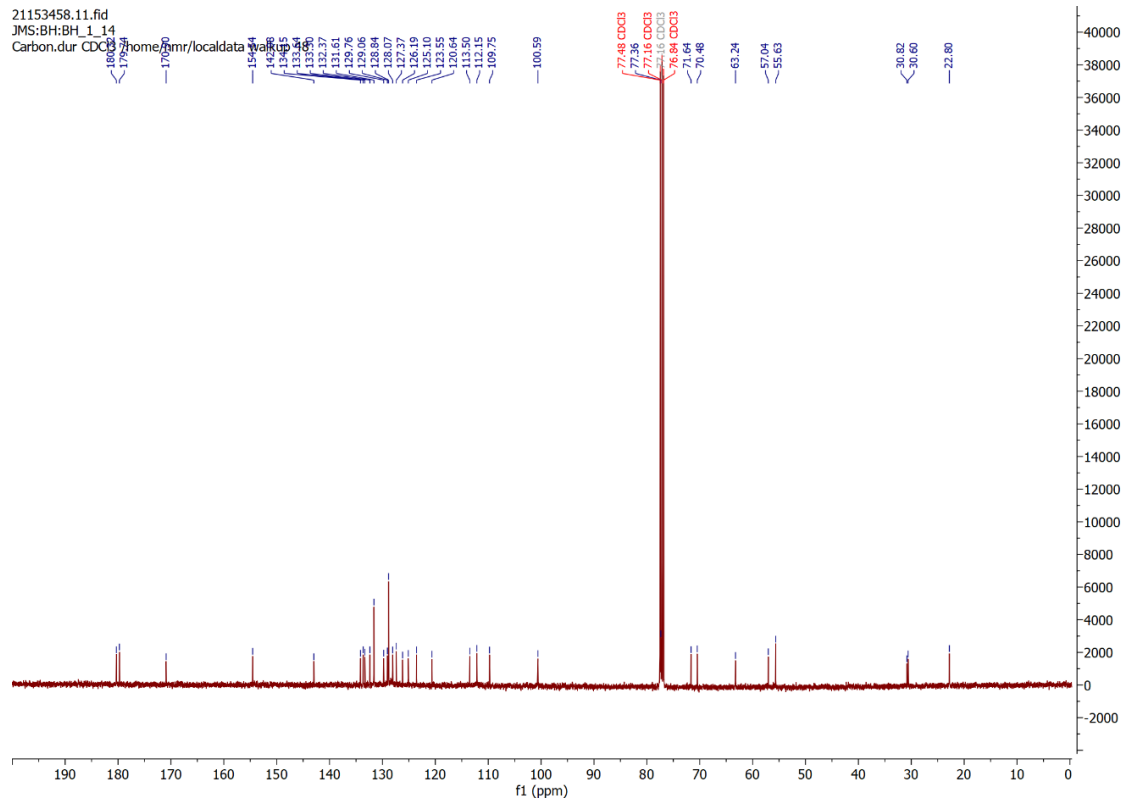
BCRH_1_148_1_165080

RT: 3.6350 minutes, Scan 279, 1: MS (100-2000) ES+, NL 1.52e+8



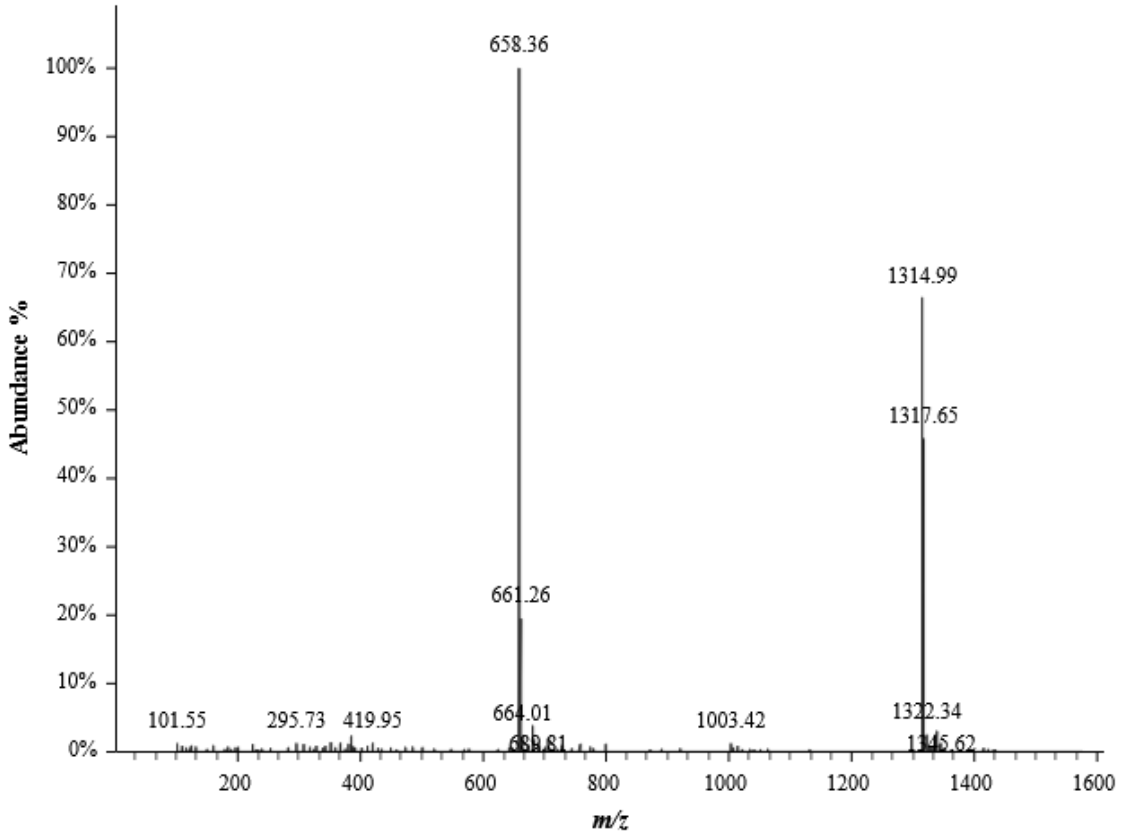
B.7.b. 5-Methoxy-Tryptophan-Ni-(S)-2-[N-(N-benzylpropyl) amino]benzophenone (**7b**)



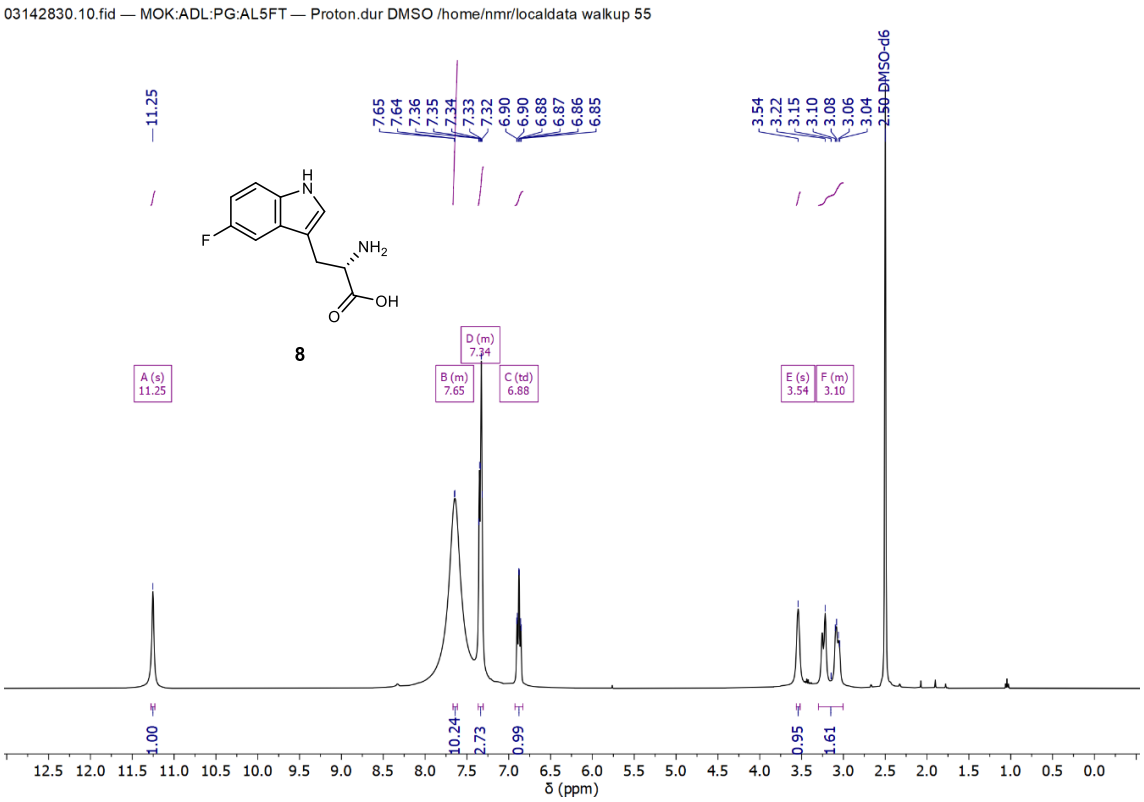


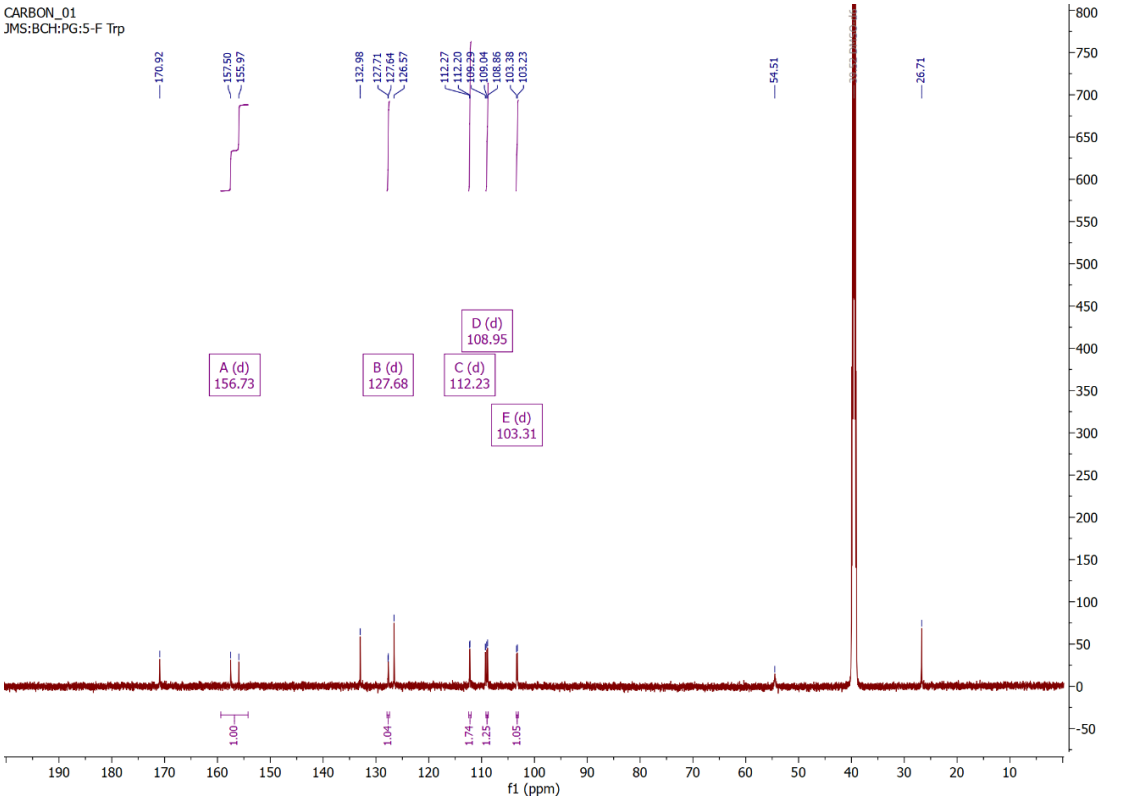
BCRH_BH0158_160031

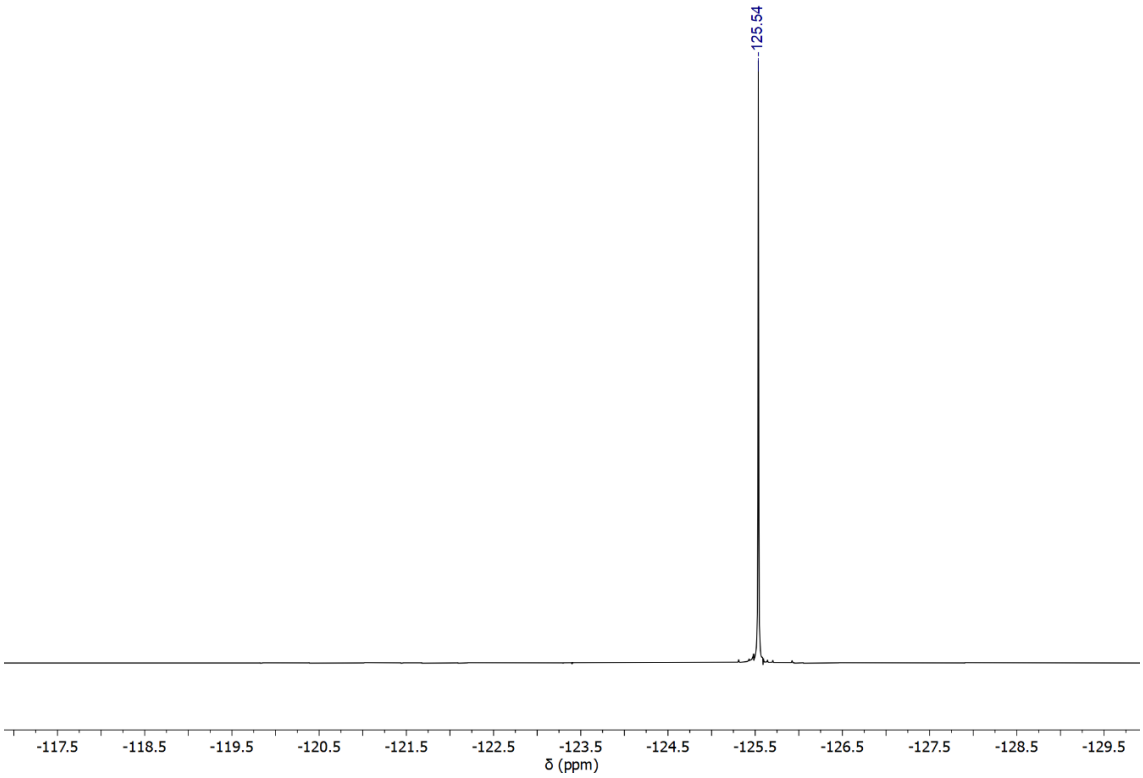
RT: 2.5812 minutes, Scan 197, 1: MS (100-2000) ES+, NL 3.11e+7



B.8. 5-Fluorotryptophan (**8**)

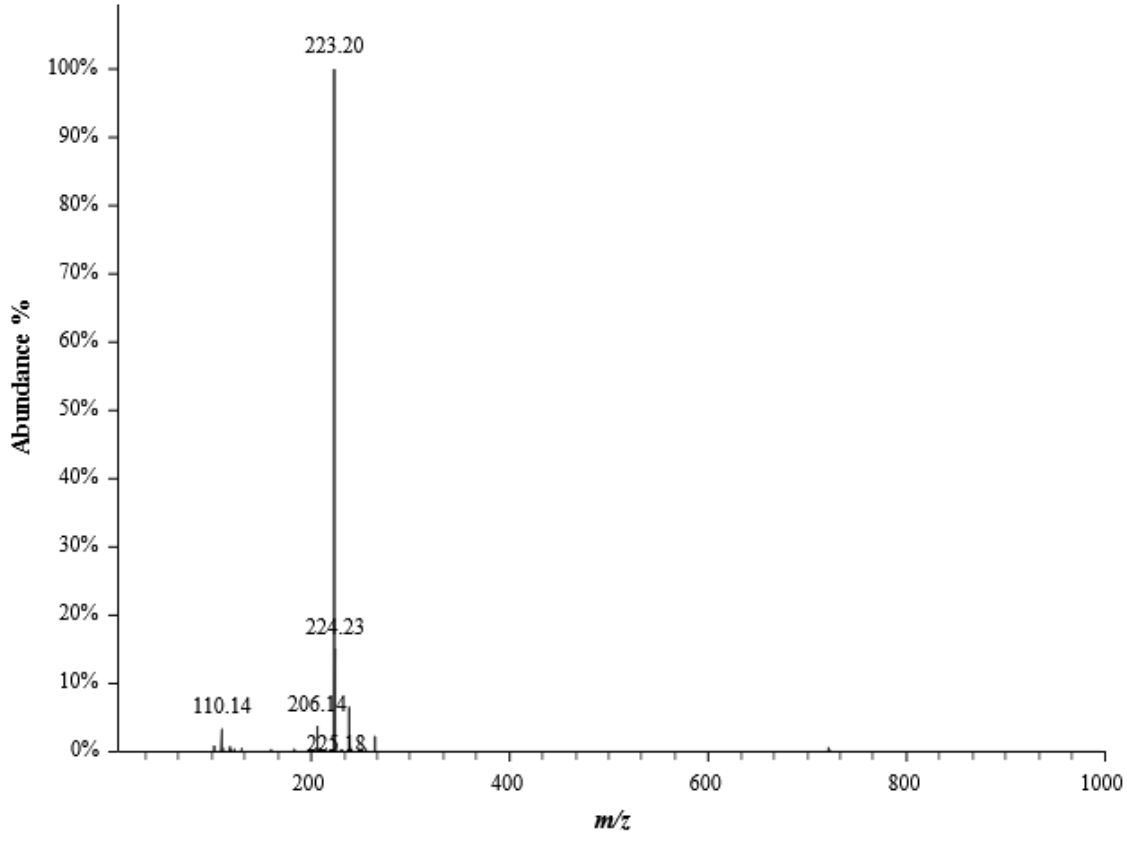




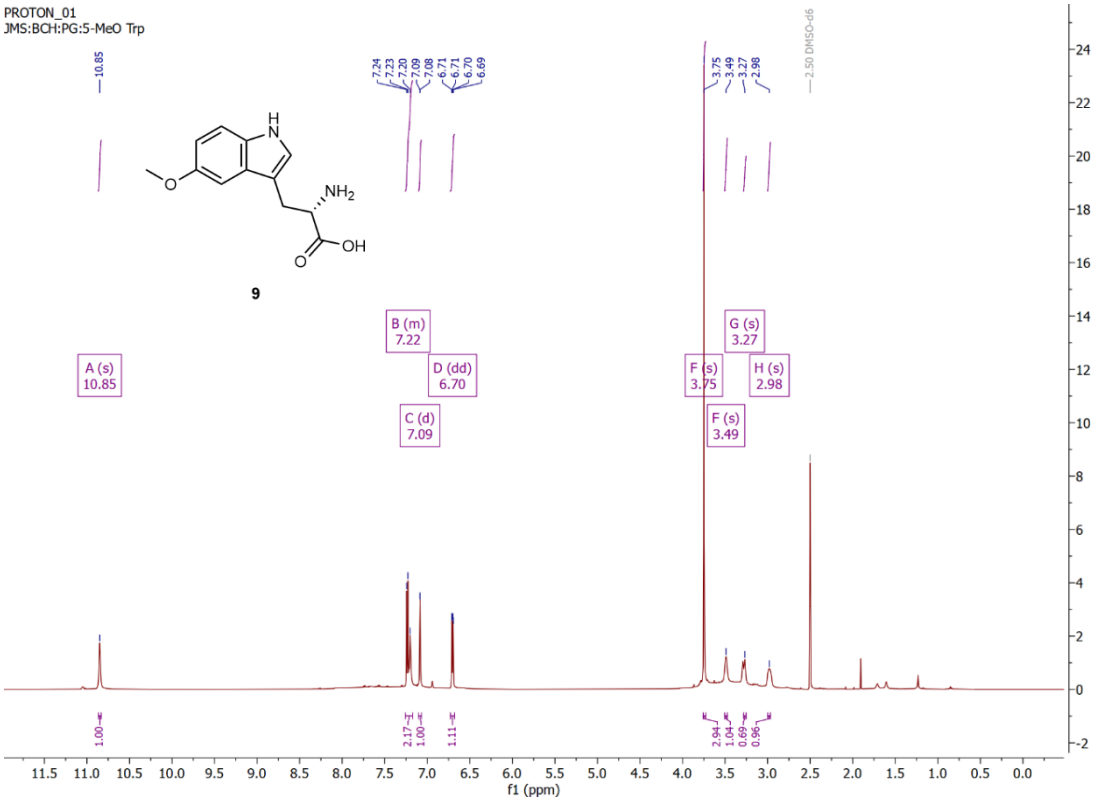


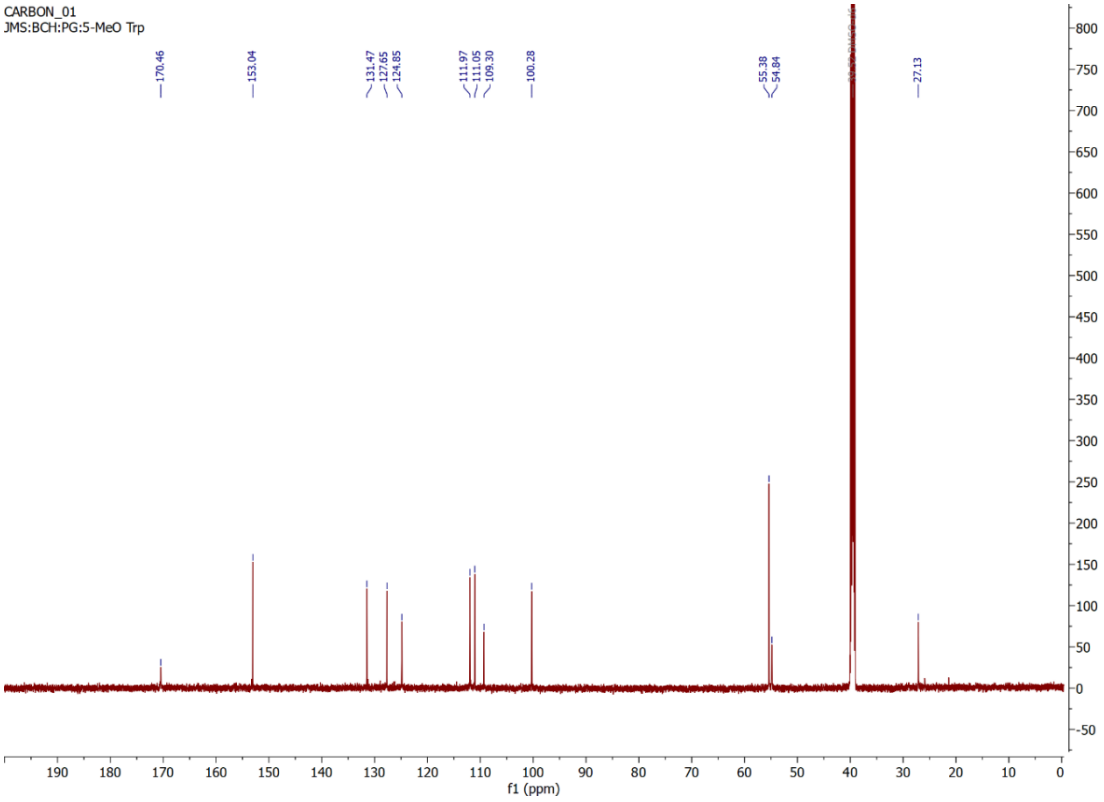
BCRH_5F_Trp_215707

RT: 0.8741 minutes, Scan 95, 1: MS ES+ c (100.0-2000.0), NL 5.11e+7



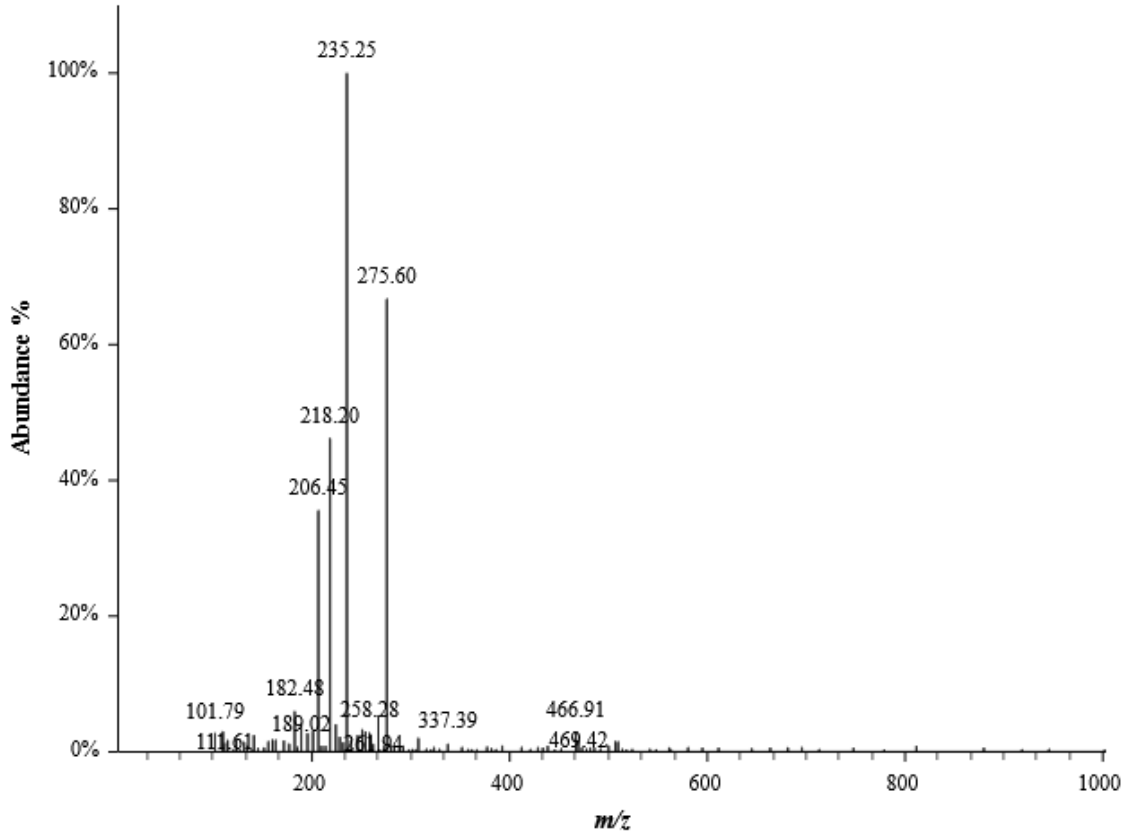
B.9. 5-Methoxytryptophan (**9**)

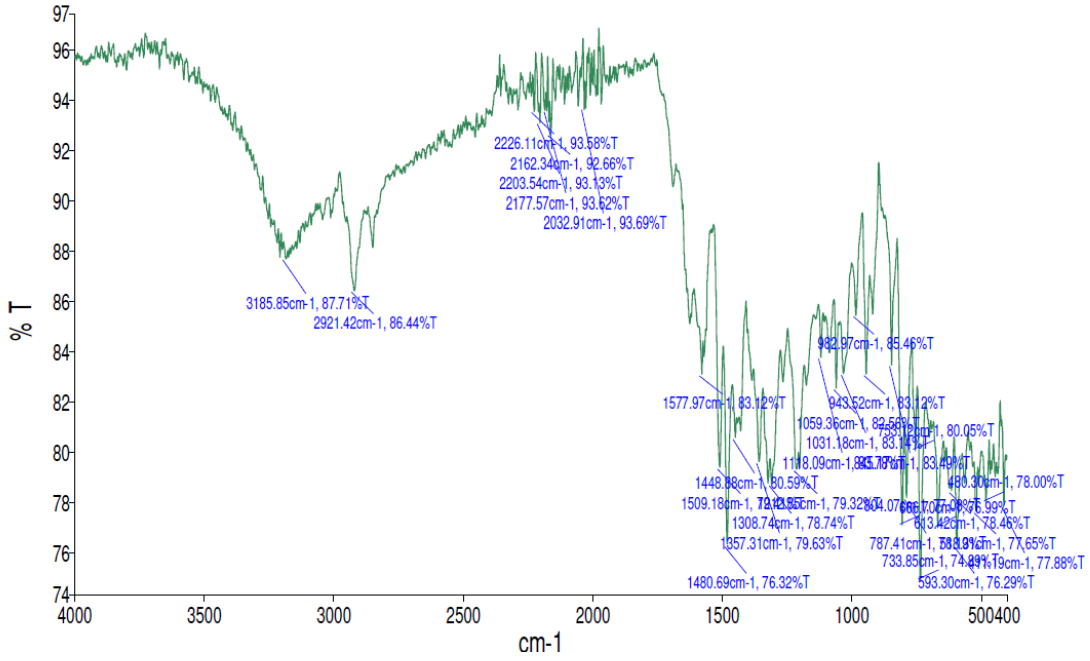




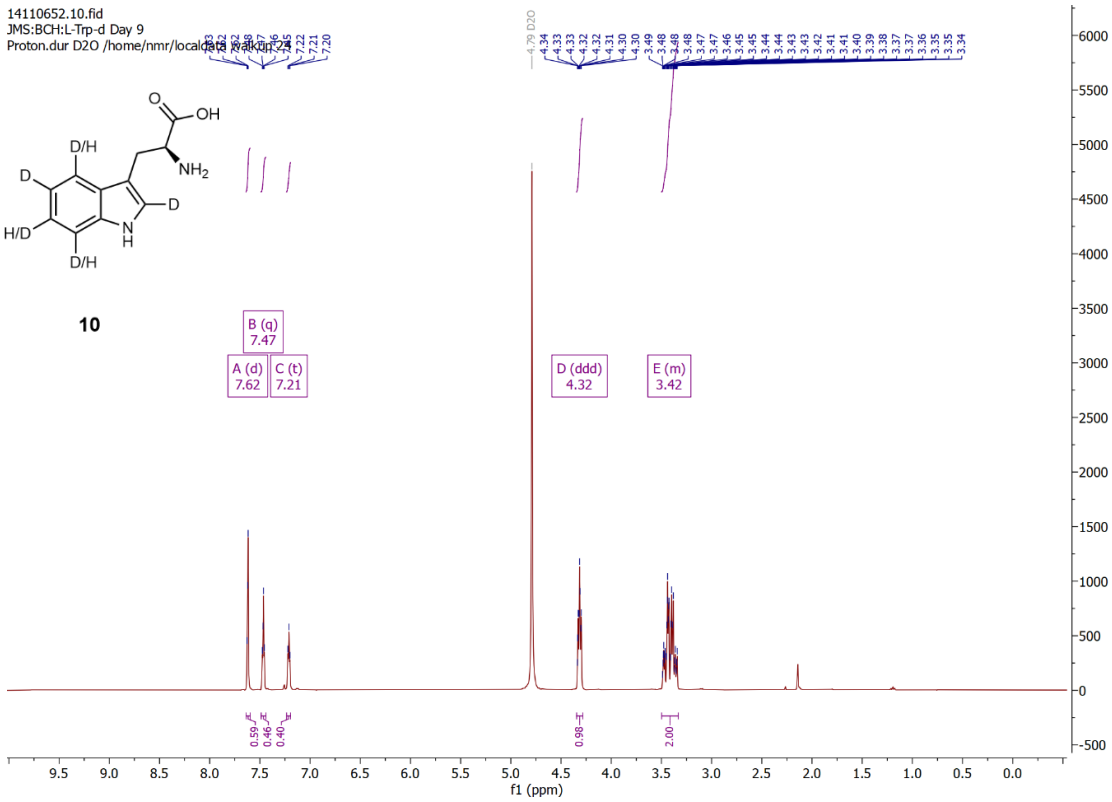
BCRH_BH01MeOTrp_160005

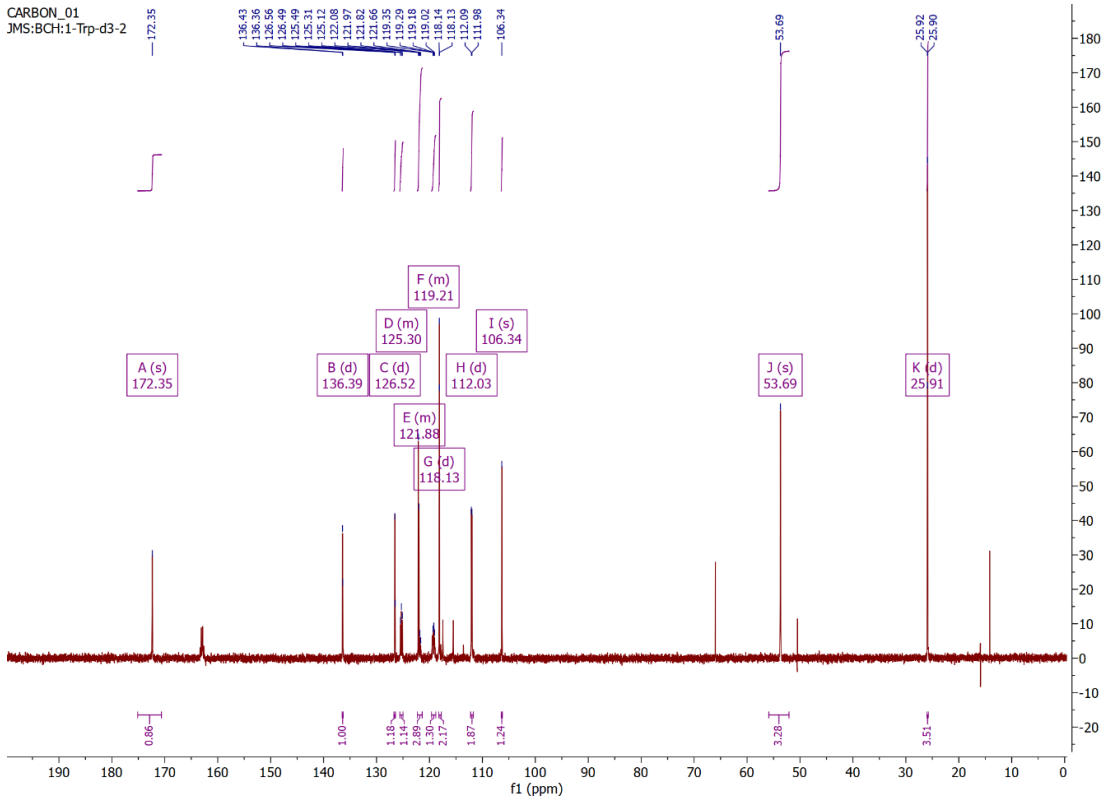
RT: 0.2939 minutes, Scan 19, 1: MS (100-2000) ES+, NL 2.65e+7

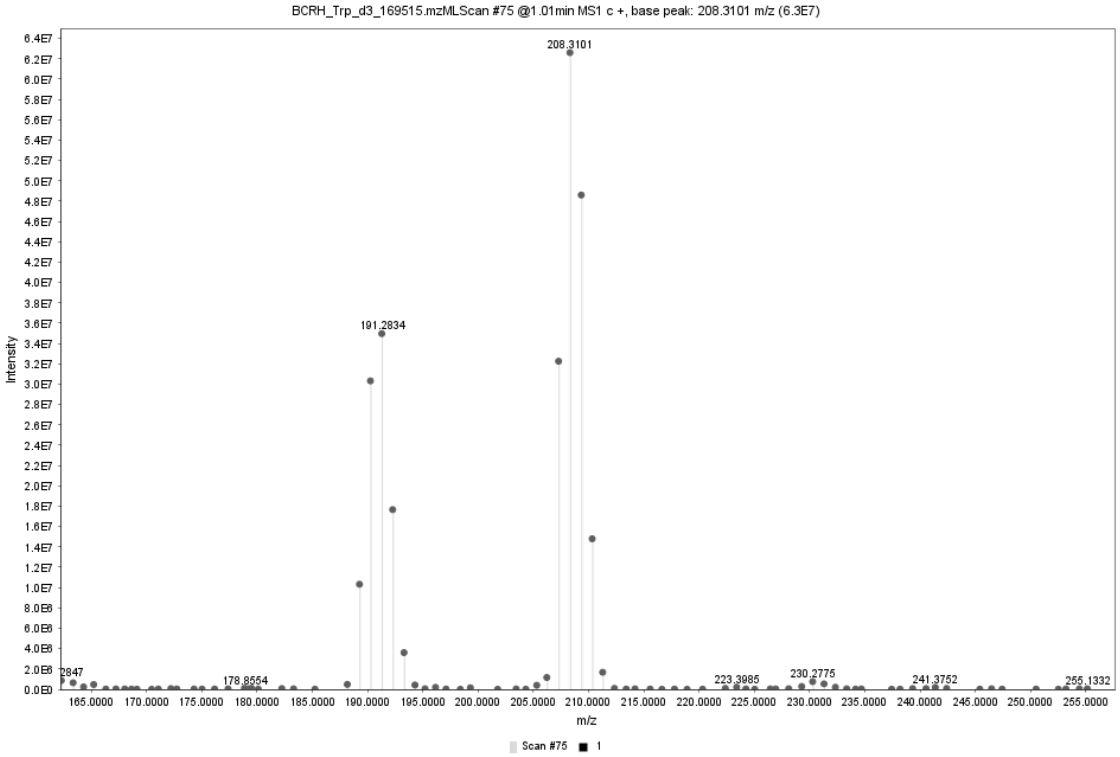


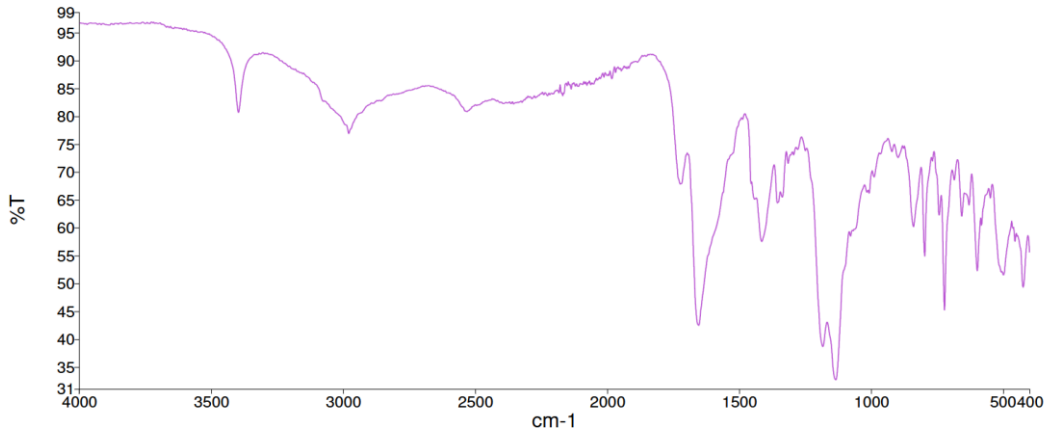


B.10. L-Tryptophan-d_3 (**10**)

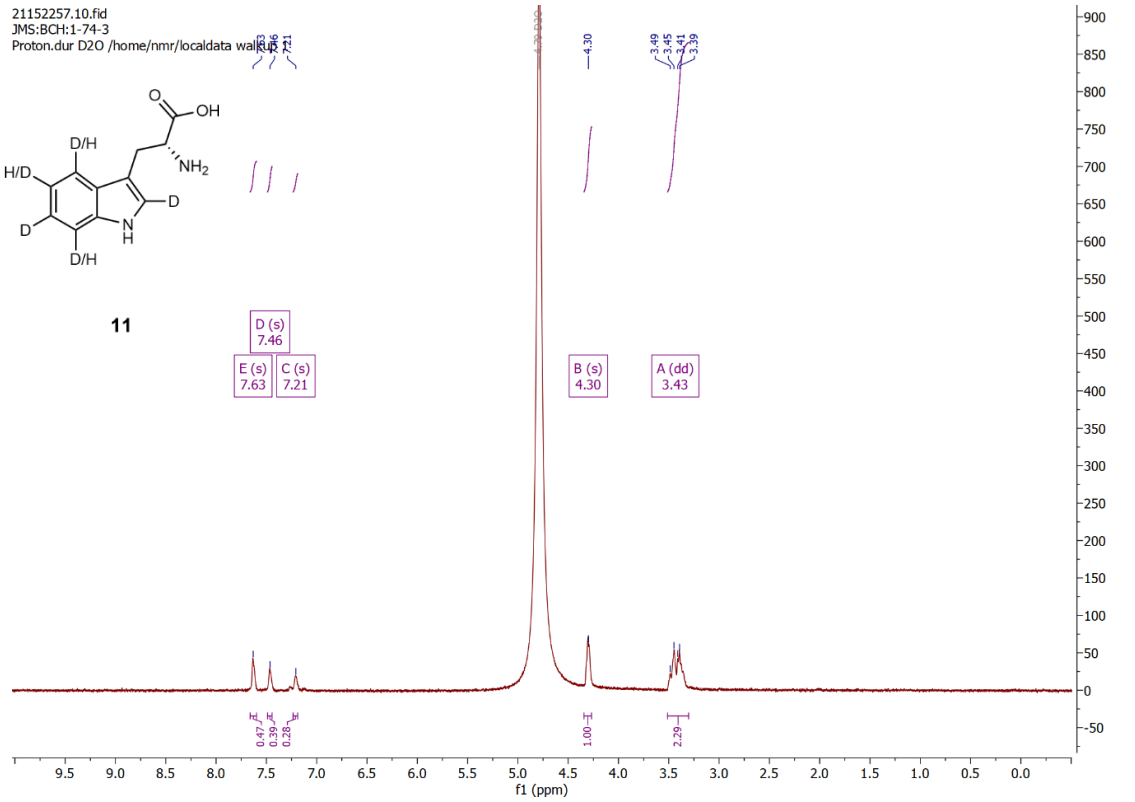


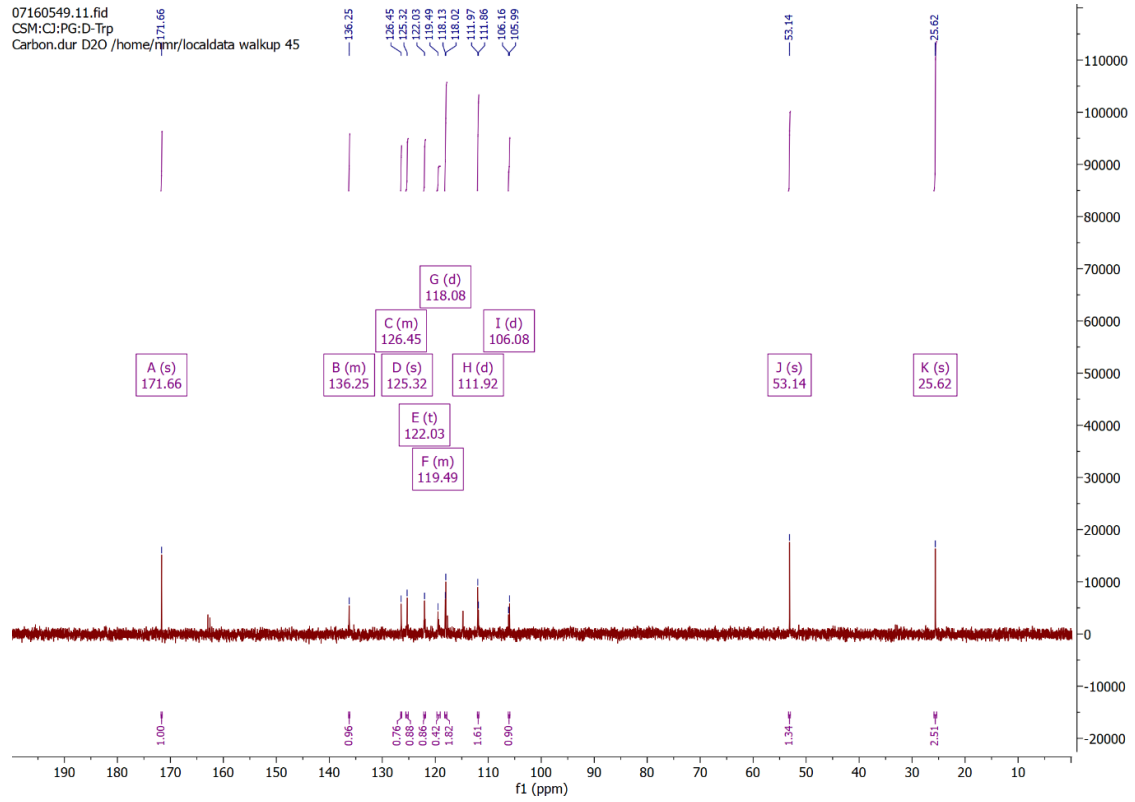


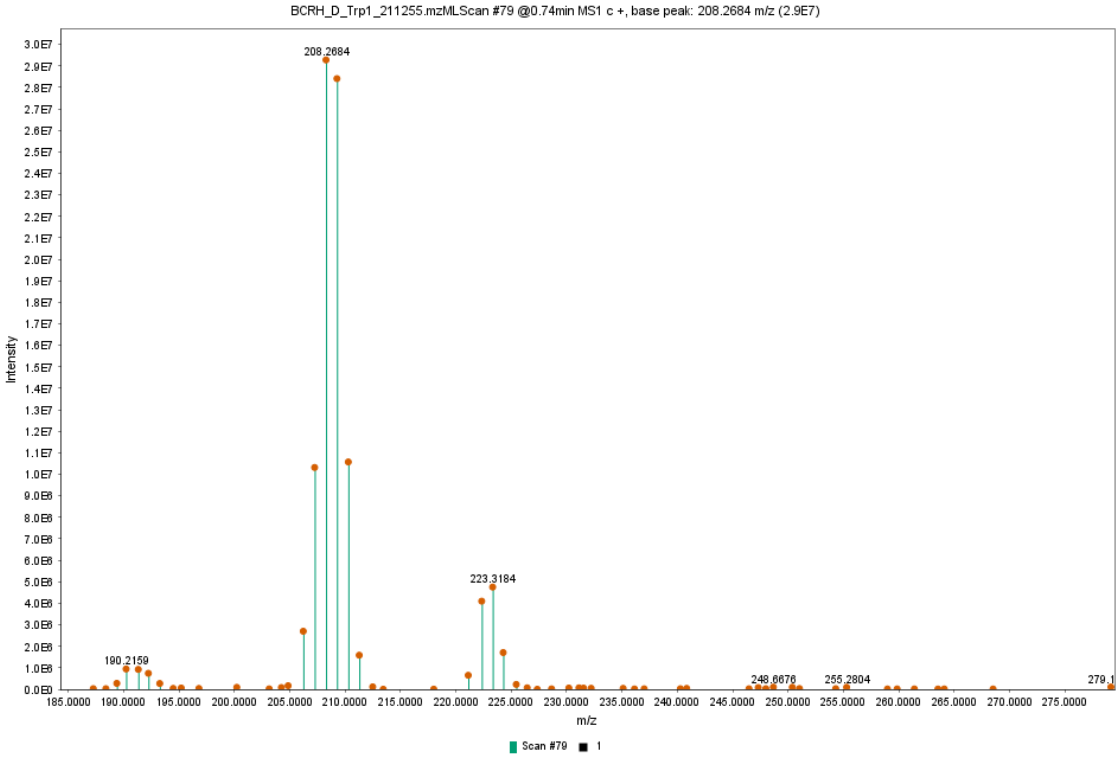

BCRH_Trp_d3_169515.mzMLScan #75 @1.01min MS1 c +, base peak: 208.3101 m/z (6.3E7)
Intensity
m/z
208.3101
191.2834
2847
178.8554
223.3985
230.2775
241.3752
255.1332
Scan #75
1

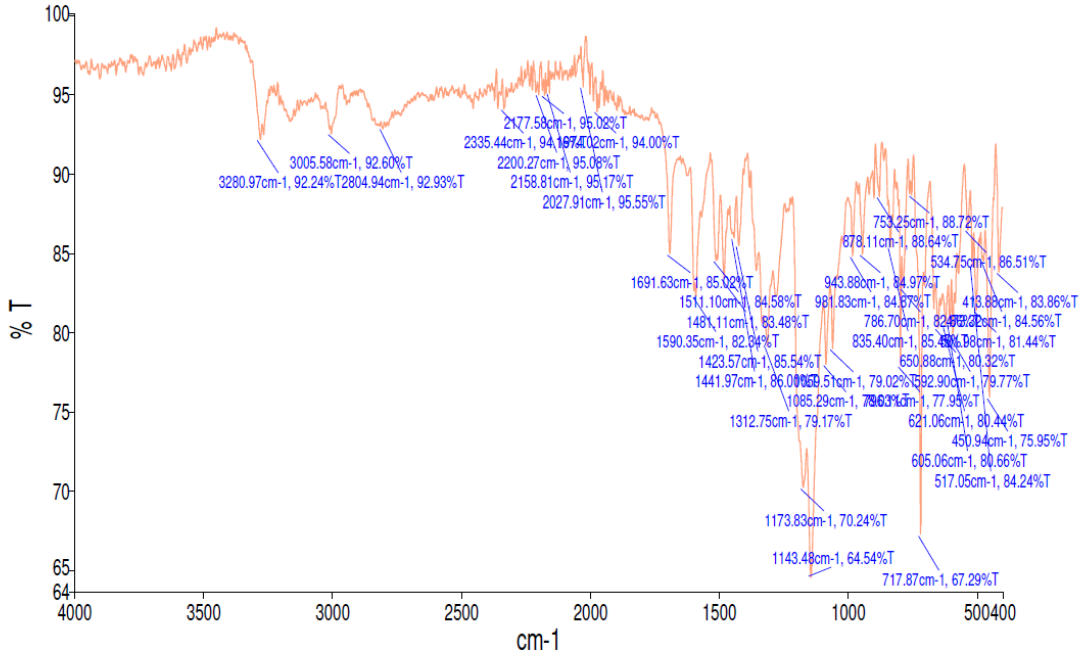

%T
cm-1

B.11. D-Tryptophan-d_3 (**11**)

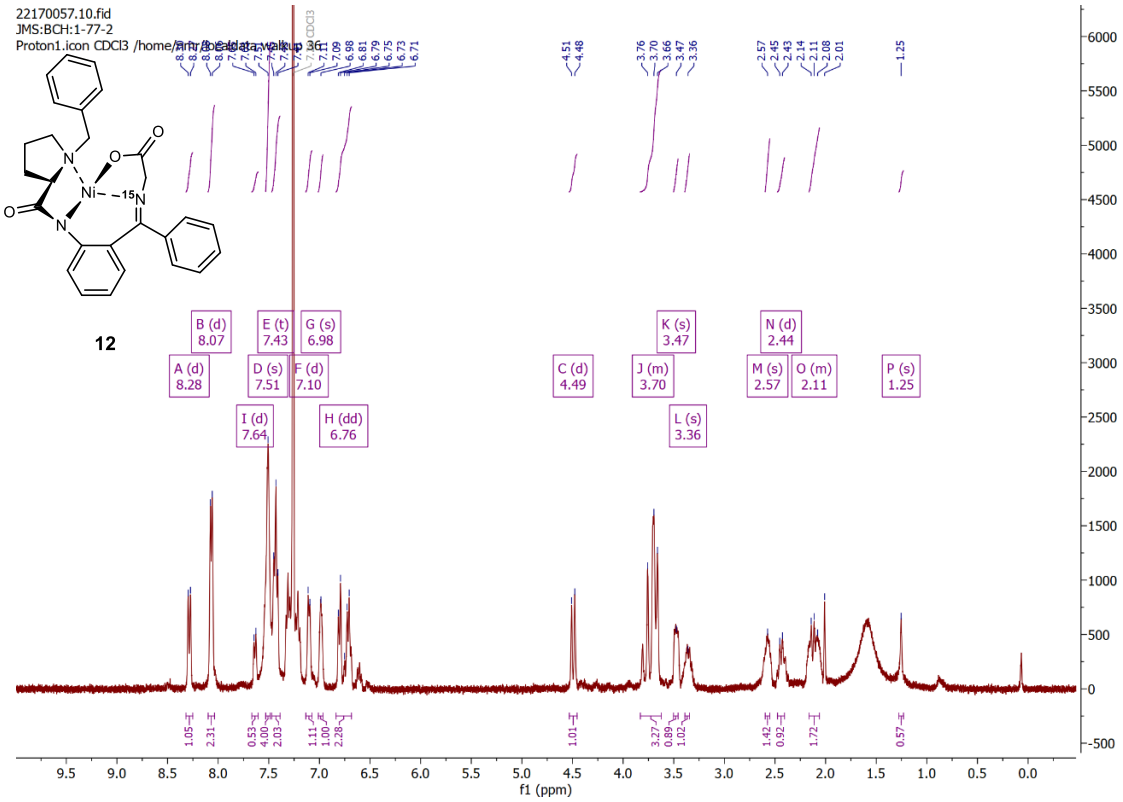





BCRH_D_Trp1_211255.mzMLScan #79 @0.74min MS1 c +, base peak: 208.2684 m/z (2.9E7)
Intensity
m/z
208.2684
223.3184
190.2159
248.6676
255.2804
279.1
Scan #79
1

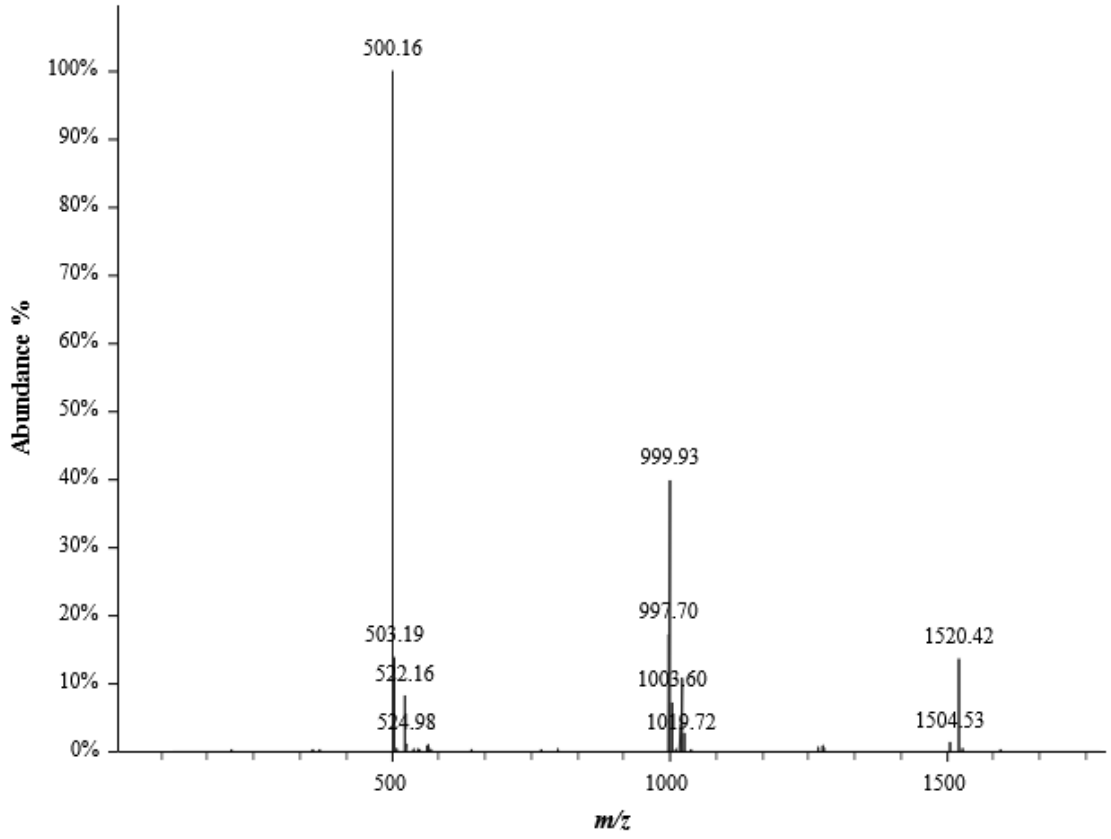

% T
cm-1
3280.97cm-1, 92.24%T
3005.58cm-1, 92.60%T
2804.94cm-1, 92.93%T
2177.58cm-1, 95.02%T
2335.44cm-1, 94.18%T
2200.27cm-1, 95.06%T
2158.81cm-1, 95.17%T
2027.91cm-1, 95.55%T
1691.63cm-1, 85.02%T
1511.10cm-1, 84.58%T
1481.11cm-1, 83.48%T
1590.35cm-1, 82.34%T
1423.57cm-1, 85.54%T
1441.97cm-1, 86.09%T
1312.75cm-1, 79.17%T
1173.83cm-1, 70.24%T
1143.48cm-1, 64.54%T
717.87cm-1, 67.29%T
753.25cm-1, 88.72%T
878.11cm-1, 88.64%T
534.75cm-1, 86.51%T
943.88cm-1, 84.97%T
981.83cm-1, 84.87%T
786.70cm-1, 84.56%T
835.40cm-1, 81.44%T
650.88cm-1, 80.32%T
582.90cm-1, 79.77%T
621.06cm-1, 80.44%T
450.94cm-1, 75.95%T
605.06cm-1, 80.66%T
517.05cm-1, 84.24%T

B.12. ^{15}N-Gly-Ni-(S)-2-[N-(N-benzylpropyl)amino]benzophenone (**12**)

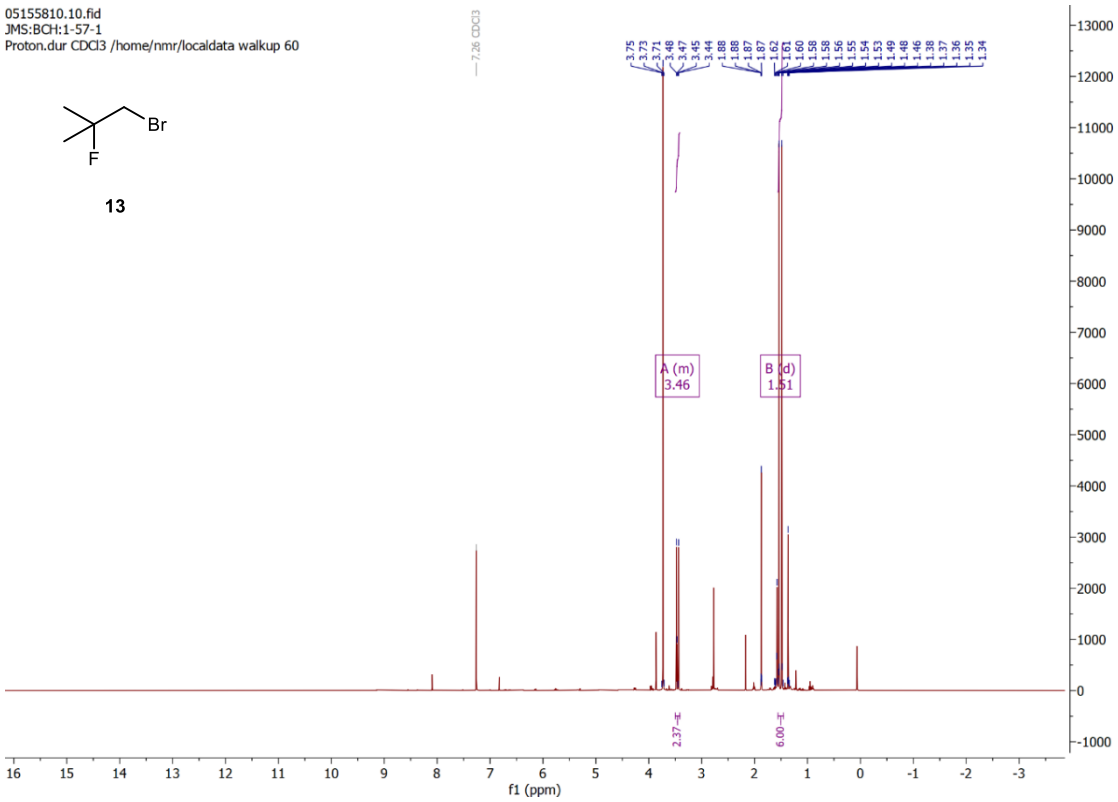


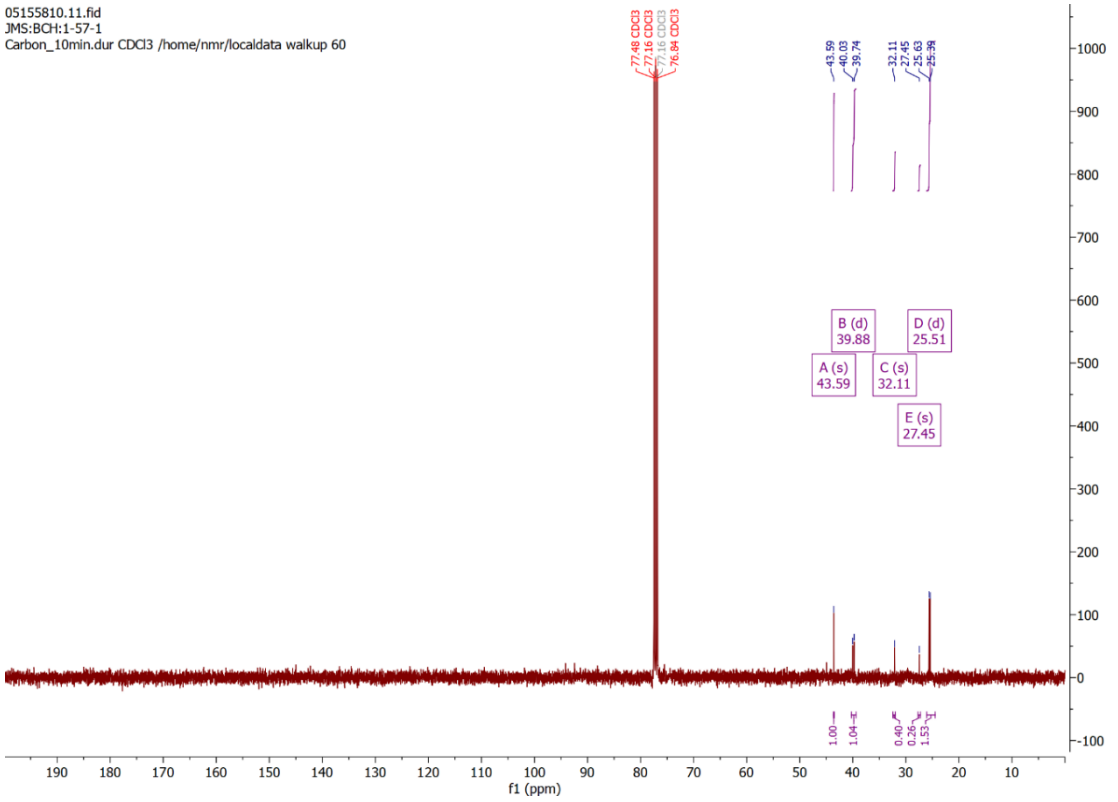
BCRH_BH0177_160029

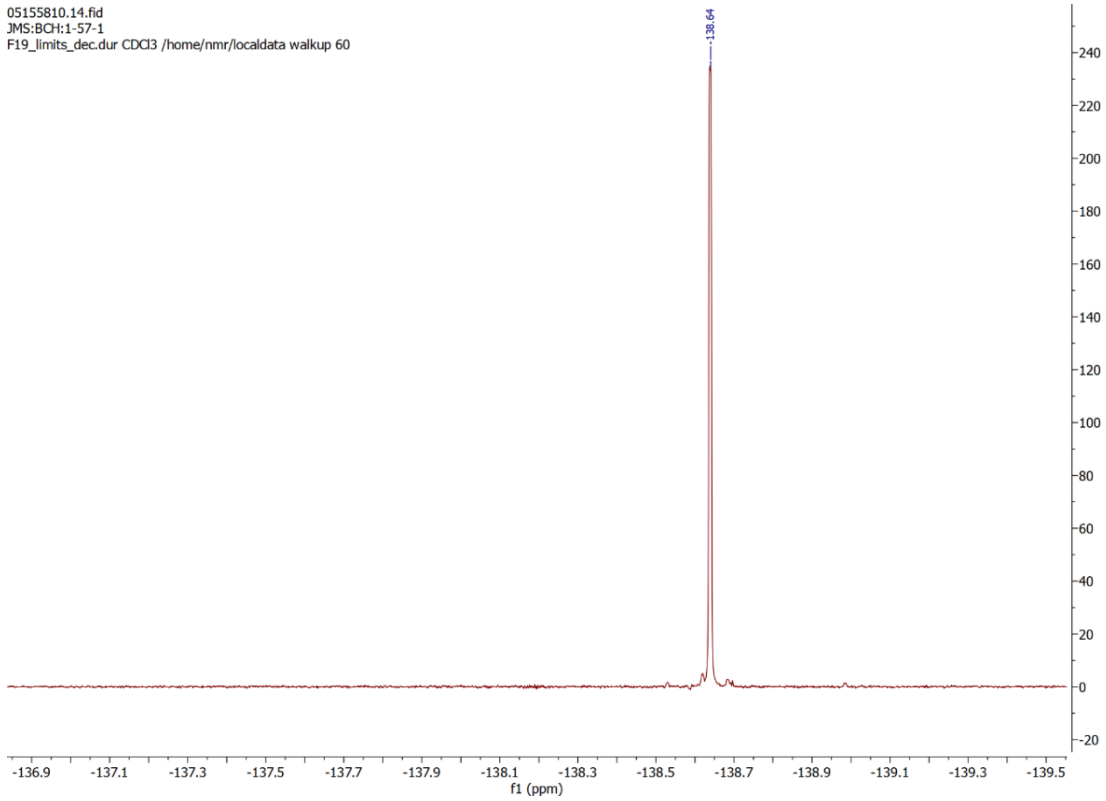
RT: 2.9667 minutes, Scan 227, 1: MS (100-2000) ES+, NL 7.62e+7



B.13. 1-Bromo-2-fluoro-2-methyl propane (**13**)

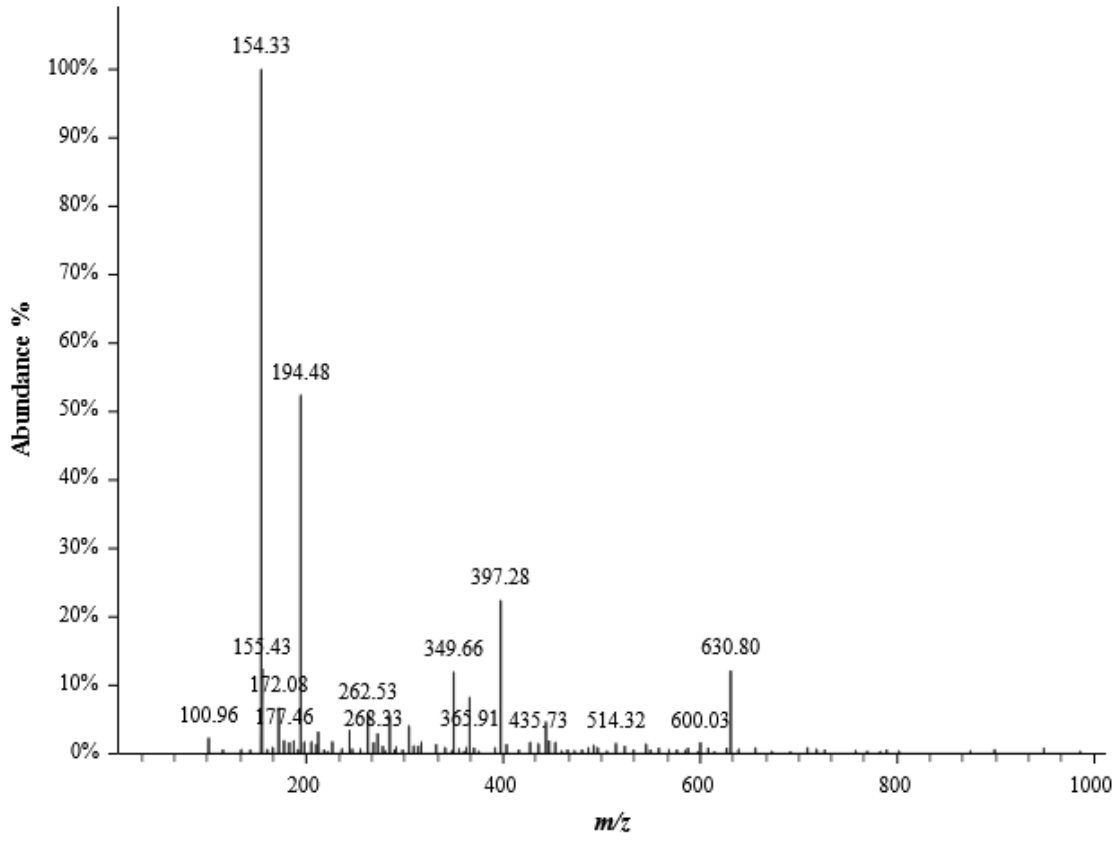


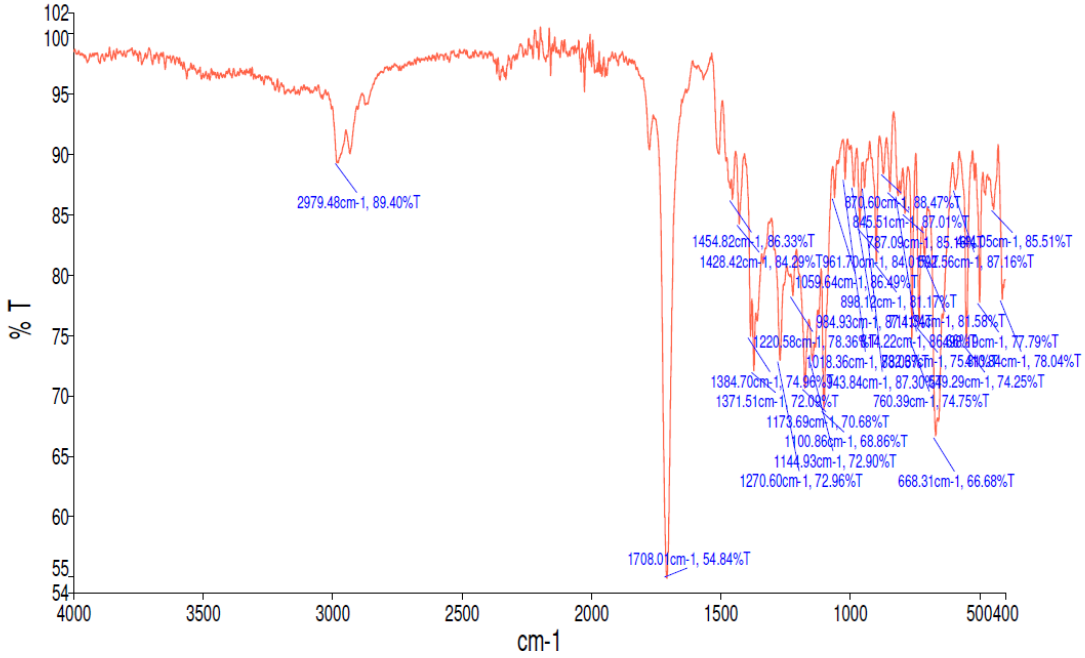




BCRH_BH0157TEST_159696

RT: 0.9107 minutes, Scan 67, 1: MS (100-2000) ES+, NL 1.62e+7

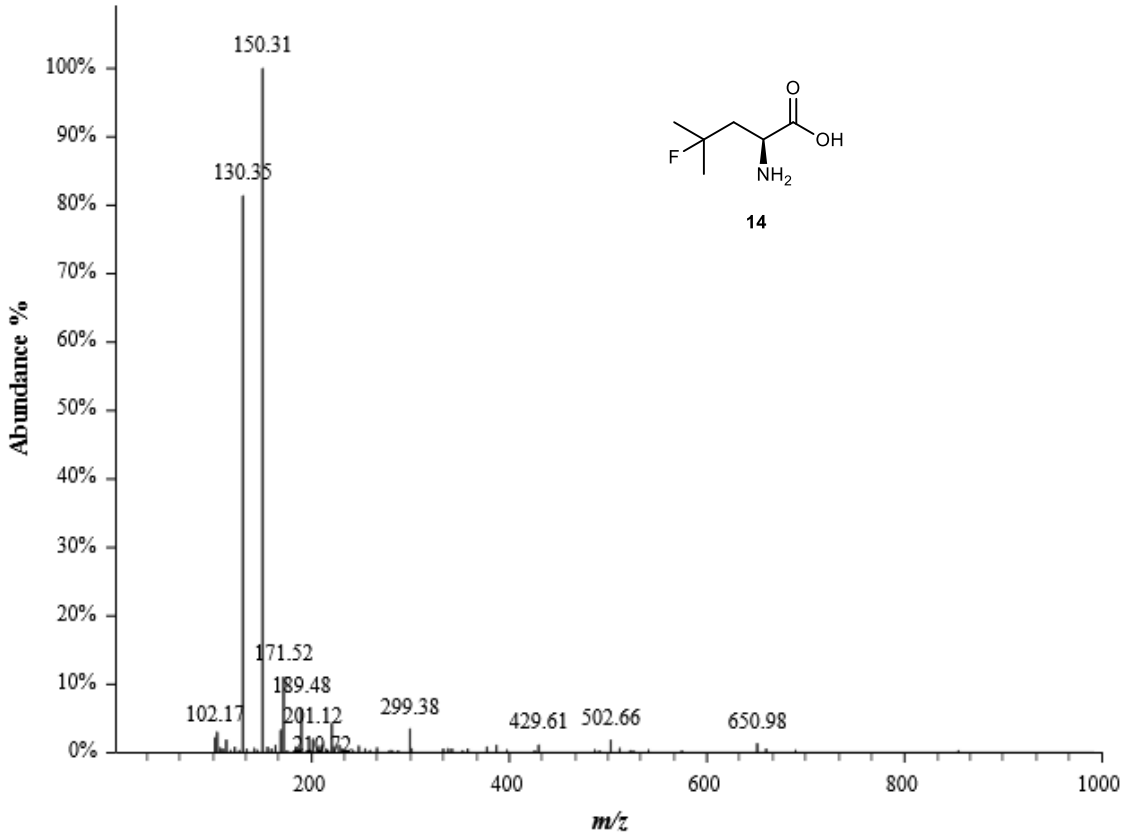




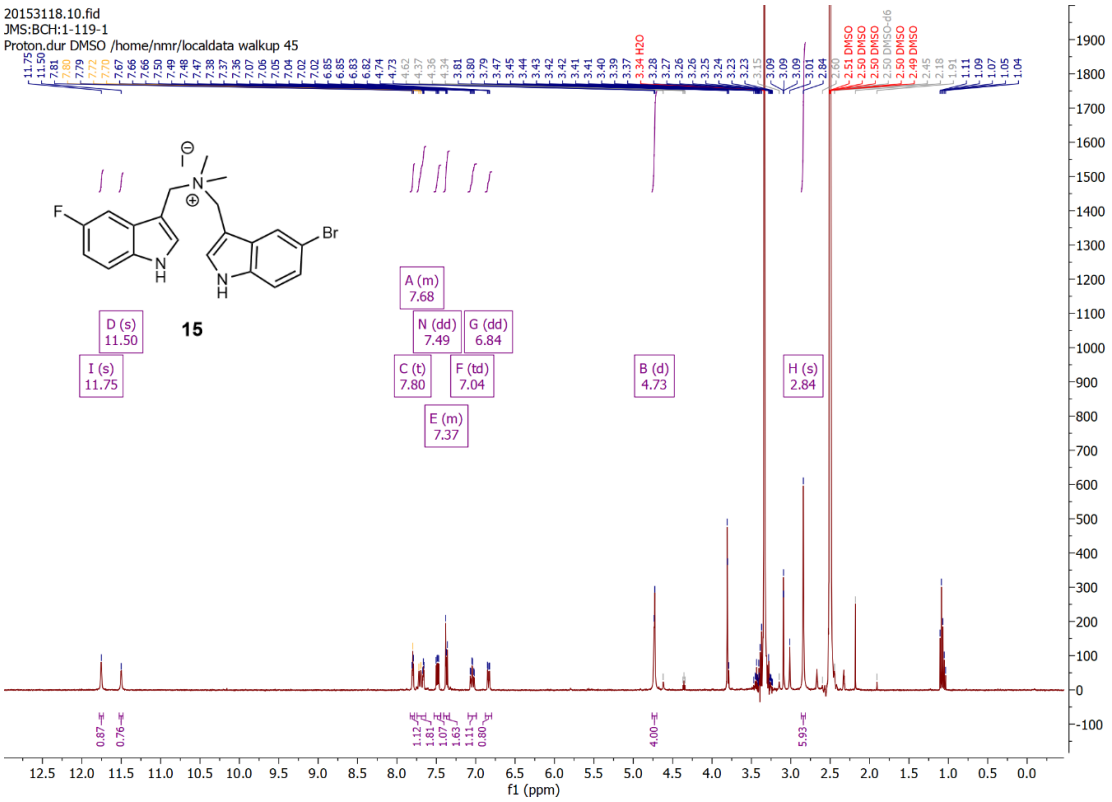
B.14. F-Leucine (**14**)

BCRH_BH0110701_163225

RT: 0.3709 minutes, Scan 25, 1: MS (100-2000) ES+, NL 9.51e+7

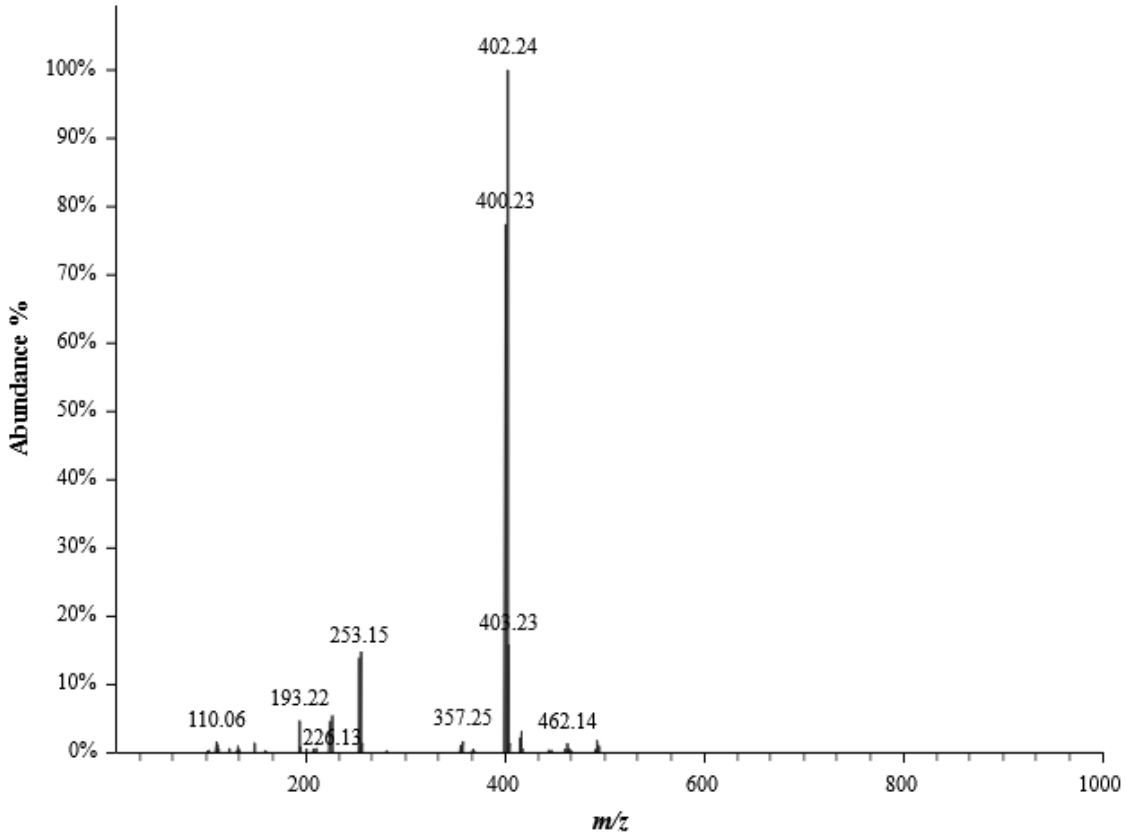


B.15. 1-(5-bromo-1H-indol-3-yl)-N-((5-fluoro-1H-indol-3-yl)methyl)-N,N-dimethylmethanaminium iodide (**15**)



BCRH_mixed_dimer_fraction_4_205625

RT: 1.8634 minutes, Scan 209, 1: MS ES+ c (100.0-2000.0), NL 6.08e+7



Appendix C: Detection of Candidate Metabolites and Plant Hormones Across Soybean Tissues

Table C1: *Presence of candidate hormones and tryptophan metabolites in soybean tissues under positive and negative ionisation mode. Each row lists a compound of interest, and each column corresponds to a plant tissue analysed under either positive (+ve) or negative (-ve) ionisation mode. Ticks indicate detection above intensity threshold based on m/z values; crosses indicate absence. Data were compiled using R (Appendix D.4) from full-scan LC-MS spectra.*

	Meri-stem -ve	**Meri-stem +ve**	**Stem -ve**	**Stem +ve**	**T1 -ve**	**T1 +ve**	**T2 spot -ve**	**T2 spot +ve**	**T2uw -ve**	**T2uw +ve**	**T2w -ve**	**T2w +ve**	**T3 -ve**	**T3 +ve**
5-F Tryptamine	✓	✓	⊗	⊗	✓	✓	✓	✓	⊗	✓	✓	✓	⊗	✓
Tryptamine	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	⊗
5-OH Tryptamine	✓	⊗	✓	✓	✓	⊗	✓	⊗	✓	⊗	✓	⊗	✓	⊗
5-OCH_3 Tryptamine	⊗	✓	✓	⊗	⊗	✓	✓	✓	⊗	✓	✓	✓	✓	⊗

	Meri-stem -ve	Meri-stem +ve	Stem -ve	Stem +ve	T1 -ve	T1 +ve	T2 spot -ve	T2 spot +ve	T2uw -ve	T2uw +ve	T2w -ve	T2w +ve	T3 -ve	T3 +ve
N-Ac Tryptamine	✓	✓	✓	✓	✓	⊗	✓	✓	✓	✓	✓	✓	⊗	✓
N-Ac-5-OH Tryptamine	⊗	✓	⊗	✓	⊗	✓	⊗	✓	⊗	✓	✓	✓	⊗	✓
3-OH Anthranilic acid	⊗	⊗	⊗	⊗	⊗	⊗	✓	⊗	✓	✓	⊗	✓	⊗	⊗
3-OH Kynurenine	✓	✓	✓	✓	⊗	⊗	⊗	✓	✓	⊗	✓	✓	⊗	✓
Anthranilic acid	✓	✓	⊗	✓	⊗	✓	⊗	✓	⊗	✓	⊗	✓	⊗	⊗
Kynurenic acid	✓	⊗	✓	✓	✓	✓	✓	✓	✓	⊗	✓	✓	✓	✓
Kynurenine	⊗	✓	⊗	✓	✓	✓	⊗	✓	✓	✓	✓	⊗	⊗	⊗

	Meri-stem -ve	Meri-stem +ve	Stem -ve	Stem +ve	T1 -ve	T1 +ve	T2 spot -ve	T2 spot +ve	T2uw -ve	T2uw +ve	T2w -ve	T2w +ve	T3 -ve	T3 +ve
Trp ethyl ester	✓	⊗	✓	⊗	⊗	✓	✓	✓	✓	⊗	⊗	✓	⊗	⊗
5-OH Trp	⊗	⊗	✓	✓	⊗	✓	✓	✓	✓	✓	✓	✓	⊗	⊗
Melatonin	✓	⊗	✓	⊗	⊗	⊗	✓	✓	✓	⊗	⊗	✓	✓	⊗
L-Trp	✓	⊗	⊗	✓	✓	✓	⊗	✓	✓	⊗	⊗	⊗	✓	⊗
Indole-3-acetamide (IAM)	✓	⊗	⊗	⊗	✓	⊗	✓	⊗	⊗	⊗	⊗	⊗	⊗	⊗
Indole-3-pyruvic acid (IPyA)	⊗	✓	✓	⊗	⊗	✓	✓	✓	⊗	✓	✓	✓	✓	⊗

	Meri-stem -ve	Meri-stem +ve	Stem -ve	Stem +ve	T1 -ve	T1 +ve	T2 spot -ve	T2 spot +ve	T2uw -ve	T2uw +ve	T2w -ve	T2w +ve	T3 -ve	T3 +ve
Tryptophol (Indole-3-ethanol)	✓	⊗	⊗	⊗	⊗	⊗	⊗	✓	⊗	✓	⊗	✓	⊗	⊗
Serotonin	✓	⊗	✓	✓	✓	⊗	✓	⊗	✓	⊗	✓	⊗	✓	⊗
Indole-3-acetic acid (IAA)	⊗	⊗	⊗	⊗	⊗	⊗	✓	⊗	✓	⊗	⊗	⊗	⊗	⊗
Indole-3-butyric acid (IBA)	✓	⊗	✓	✓	✓	✓	✓	⊗	✓	⊗	✓	⊗	✓	⊗
Naphthalene acetic acid (NAA)	✓	✓	⊗	✓	✓	⊗	✓	✓	✓	✓	✓	⊗	⊗	✓
Trans-zeatin	NA	✓	NA	⊗	NA	⊗	NA	✓	NA	✓	NA	⊗	NA	⊗

	Meri-stem -ve	Meri-stem +ve	Stem -ve	Stem +ve	T1 -ve	T1 +ve	T2 spot -ve	T2 spot +ve	T2uw -ve	T2uw +ve	T2w -ve	T2w +ve	T3 -ve	T3 +ve
Kinetin	NA	⊗	NA	✓	NA	⊗	NA	⊗	NA	✓	NA	✓	NA	⊗
6-Benzylaminopurine (BAP)	NA	⊗	NA	✓	NA	✓	NA	✓	NA	✓	NA	✓	NA	⊗
Gibberellic acid (GA3)	⊗	✓	✓	✓	✓	⊗	✓	⊗	✓	✓	✓	⊗	⊗	⊗
Gibberellin A1 (GA1)	✓	⊗	✓	✓	⊗	⊗	⊗	✓	✓	⊗	✓	✓	⊗	⊗
Gibberellin A4 (GA4)	⊗	✓	✓	⊗	✓	⊗	⊗	⊗	⊗	⊗	⊗	✓	⊗	⊗
Abscisic acid (ABA)	✓	✓	✓	⊗	✓	✓	✓	✓	✓	✓	✓	✓	⊗	⊗

	Meri-stem -ve	Meri-stem +ve	Stem -ve	Stem +ve	T1 -ve	T1 +ve	T2 spot -ve	T2 spot +ve	T2uw -ve	T2uw +ve	T2w -ve	T2w +ve	T3 -ve	T3 +ve
Jasmonic acid (JA)	✓	✓	⊗	✓	✓	✓	✓	✓	✓	✓	✓	✓	⊗	✓
Jasmonoyl-isoleucine (JA-Ile)	✓	⊗	✓	⊗	⊗	✓	✓	⊗	✓	✓	✓	⊗	⊗	⊗
Brassinolide	✓	✓	✓	✓	✓	⊗	✓	⊗	⊗	✓	✓	✓	⊗	⊗
Castasterone (24-epicastasterone)	✓	✓	✓	✓	⊗	⊗	✓	⊗	✓	⊗	✓	⊗	⊗	⊗
1-Aminocyclo-propane -1-carboxylic acid (ACC)	⊗	✓	⊗	⊗	⊗	⊗	⊗	✓	⊗	✓	⊗	✓	⊗	⊗

Appendix D: R Script

D.1. Comparing indices in MRM script

```r
# Load required libraries
library(MSnbase)
library(tidyverse)
library(pracma)
library(fs)

# --- Gaussian fitting function ---
fit_gaussian_model <- function(time, intensity, min_peak_diff = 1) {
  mean_intens <- mean(intensity, na.rm = TRUE)
  sd_intens <- sd(intensity, na.rm = TRUE)
  inlier_idx <- which(intensity < mean_intens + sd_intens & intensity > mean_intens - sd_intens)
  ref_intens <- mean(intensity[inlier_idx], na.rm = TRUE)
  sd_intens_out <- sd(intensity[inlier_idx], na.rm = TRUE)

  A0 <- max(intensity, na.rm = TRUE) - ref_intens
  baseline0 <- ref_intens
  mu0 <- time[which.max(intensity)]
  sigma0 <- 0.2
  t.window <- 0.2

  fit <- tryCatch({
    nls(intensity ~ baseline + A * exp(-((time - mu)^2) / (2 * sigma^2)),
        start = list(A = A0, mu = mu0, sigma = sigma0, baseline = baseline0),
        algorithm = "port",
        lower = c(A = 0.1 * A0, mu = mu0 - t.window, sigma = 0.001, baseline = baseline0 - 2 * sd_intens_out),
        upper = c(A = 2 * A0, mu = mu0 + t.window, sigma = 0.4, baseline = baseline0 + 2 * sd_intens_out))
  }, error = function(e) {
    message("⚠ Gaussian fitting failed: ", e$message)
    return(NULL)
  })

  if (is.null(fit)) return(list(auc = NA, params = NA, fitCurve = NULL))
  params <- coef(fit)
  if (params["sigma"] > (max(time) - min(time))) {
    return(list(auc = 0, params = params, fitCurve = NULL))
  }

  auc <- params["A"] * params["sigma"] * sqrt(2 * pi)
  list(auc = auc, params = params, fitCurve = NULL)
}

# --- File setup ---
data_dir <- "."  # update with actual directory if needed
all_files <- dir_ls(data_dir, regexp = "\\.mzML$")

# --- Process files and extract AUCs ---
results <- data.frame(File = character(), Index = integer(), AUC = numeric(), stringsAsFactors = FALSE)

for (file in all_files) {
  message("🗁 Reading file: ", path_file(file))
  srm_data <- tryCatch(readSRMData(file), error = function(e) return(NULL))

  if (!is.null(srm_data) && length(srm_data) >= 4) {
    for (i in 1:4) {
      chr <- srm_data[[i]]
      fit <- fit_gaussian_model(chr@rtime, chr@intensity)
      auc_val <- fit$auc
```

```
      results <- rbind(results, data.frame(File = path_file(file), Index = i + 1, AUC =
auc_val))
    }
  } else {
    message("⛔ Skipping file: ", path_file(file), " (unreadable or <4 chromatograms)")
  }
}

if (nrow(results) == 0) stop("❌ No usable AUCs calculated.")

# --- Wide format: each row = sample, columns = AUC for Index_2 to Index_5 ---
results_wide <- results%>%
  pivot_wider(names_from = Index, values_from = AUC, names_prefix = "Index_")

# --- Calculate cross-precursor ratios ---
cross_ratios <- results_wide%>%
  mutate(
    Ratio_2_to_4 = Index_2 / Index_4,
    Ratio_3_to_5 = Index_3 / Index_5
  )%>%
  filter(!is.na(Ratio_2_to_4), !is.na(Ratio_3_to_5), is.finite(Ratio_2_to_4),
is.finite(Ratio_3_to_5))

# --- Calculate CV between the two ratios (how different are they?) ---
cross_ratios_cv <- cross_ratios%>%
  mutate(
    Ratio_CV = abs(Ratio_2_to_4 - Ratio_3_to_5) / ((Ratio_2_to_4 + Ratio_3_to_5) / 2)
  )%>%
  filter(Ratio_CV <= 2)  # optional threshold to remove extreme outliers

# --- Print summary to console ---
cv_summary <- cross_ratios_cv%>%
  summarise(
    Mean   = mean(Ratio_CV, na.rm = TRUE),
    Median = median(Ratio_CV, na.rm = TRUE),
    SD     = sd(Ratio_CV, na.rm = TRUE),
    Min    = min(Ratio_CV, na.rm = TRUE),
    Max    = max(Ratio_CV, na.rm = TRUE),
    N      = n()
  )
print(cv_summary)

```

D.2. Hybrid AUC Calculation for MRM Data (Indices 2 and 4) with Robust Baseline Correction

```
# --- Required Libraries ---
library(mzR)        # For reading mzML mass spec files
library(ggplot2)    # For plotting Gaussian fits
library(dplyr)      # For data summarisation

# --- Set Directories and Parameters ---
base_dir <- "E:/All MS Data/mzml converts/MRM/D-Trp pH 7"
out_dir <- file.path(base_dir, "choices1")
if (!dir.exists(out_dir)) dir.create(out_dir, recursive = TRUE)

output_csv <- file.path(out_dir, "Hybrid_Indices_Results.csv")
output_avg_csv <- file.path(out_dir, "Averages.csv")
output_plot_dir <- file.path(out_dir, "Hybrid_Plots_2_4")
if (!dir.exists(output_plot_dir)) dir.create(output_plot_dir, recursive = TRUE)

# Retention time window and peak detection settings
rt_min <- 3
```

```r
rt_max <- 5
peak_fraction <- 0.1
min_peak_diff <- 1

# Toggle to optionally save plots of the Gaussian fits
write_plots <- TRUE

# --- Helper Functions ---

# Area under curve using trapezoidal rule
trapz_area <- function(x, y) {
  idx <- order(x)
  x <- x[idx]
  y <- y[idx]
  sum(diff(x) * (head(y, -1) + tail(y, -1))) / 2
}

# Subset chromatogram to defined RT window
subset_chromatogram <- function(df, rt_min, rt_max) {
  df[df$time >= rt_min & df$time <= rt_max, ]
}

# Gaussian fitting with baseline estimation and bounds
fit_gaussian_model <- function(time, intensity, min_peak_diff = 1) {
  mean_intens <- mean(intensity, na.rm = TRUE)
  sd_intens <- sd(intensity, na.rm = TRUE)
  inlier_idx <- which(intensity < mean_intens + sd_intens & intensity > mean_intens - sd_intens)
  ref_intens <- mean(intensity[inlier_idx], na.rm = TRUE)
  sd_intens_out <- sd(intensity[inlier_idx], na.rm = TRUE)

  A0 <- max(intensity, na.rm = TRUE) - ref_intens
  baseline0 <- ref_intens
  mu0 <- time[which.max(intensity)]
  sigma0 <- 0.2
  t.window <- 0.2

  fit <- tryCatch({
    nls(intensity ~ baseline + A * exp(-((time - mu)^2) / (2 * sigma^2)),
        start = list(A = A0, mu = mu0, sigma = sigma0, baseline = baseline0),
        algorithm = "port",
        lower = c(A = 0.1 * A0, mu = mu0 - t.window, sigma = 0.001, baseline = baseline0 - 2 * sd_intens_out),
        upper = c(A = 2 * A0, mu = mu0 + t.window, sigma = 0.4, baseline = baseline0 + 2 * sd_intens_out))
  }, error = function(e) {
    message("Gaussian fitting failed: ", e$message)
    return(NULL)
  })

  if (is.null(fit)) return(list(auc = NA, params = NA, fitCurve = NULL))

  params <- coef(fit)
  if (params["sigma"] > (max(time) - min(time))) {
    message("Fitted sigma exceeds time range; returning AUC = 0.")
    return(list(auc = 0, params = params, fitCurve = NULL))
  }

  auc <- params["A"] * params["sigma"] * sqrt(2 * pi)
  time_seq <- seq(min(time), max(time), length.out = 200)
  fitted_intensity <- params["baseline"] + params["A"] *
    exp(-((time_seq - params["mu"])^2) / (2 * params["sigma"]^2))

  fitCurve <- data.frame(time = time_seq, intensity = fitted_intensity)
  list(auc = auc, params = params, fitCurve = fitCurve)
}

# Combines Gaussian and fallback trapezoidal integration
get_peak_area <- function(time, intensity, min_peak_diff = 1) {
```

```r
  gauss_result <- fit_gaussian_model(time, intensity, min_peak_diff)
  if (is.null(gauss_result$auc) || is.na(gauss_result$auc) || gauss_result$auc <= 0) {
    area <- trapz_area(time, intensity)
    return(list(method = "trapezoidal", auc = area, fitCurve = NULL))
  } else {
    return(list(method = "gaussian", auc = gauss_result$auc, fitCurve =
gauss_result$fitCurve))
  }
}

# Optional plotting of Gaussian fit against raw data
plot_gaussian_fit <- function(raw_data, fit_curve, index_label, file_name, out_dir) {
  p <- ggplot(raw_data, aes(x = time, y = intensity)) +
    geom_point(size = 0.8) +
    labs(title = paste("Gaussian Fit -", index_label),
         subtitle = file_name,
         x = "Time (min)", y = "Intensity") +
    theme_minimal()

  if (!is.null(fit_curve)) {
    p <- p + geom_line(data = fit_curve, aes(x = time, y = intensity), colour = "blue",
size = 1)
  }

  plot_path <- file.path(out_dir, paste0(tools::file_path_sans_ext(file_name), "_",
index_label, ".pdf"))
  ggsave(plot_path, plot = p, width = 6, height = 4)
}

# Process a single mzML file for index 2 and 4
process_single_file <- function(file_path, rt_min = 3, rt_max = 5, peak_fraction = 0.1) {
  ms_data <- tryCatch({ openMSfile(file_path) },
                      error = function(e) { message("Error opening file ",
basename(file_path), ": ", e$message); return(NULL) })
  if (is.null(ms_data)) return(NULL)

  available_indices <- seq_along(chromatogram(ms_data))
  if (max(c(2,4)) > length(available_indices)) {
    message("File ", basename(file_path), " does not have indices 2 and 4. Skipping.")
    return(NULL)
  }

  raw205 <- data.frame(time = chromatogram(ms_data, 2)[, 1],
                       intensity = chromatogram(ms_data, 2)[, 2])
  raw208 <- data.frame(time = chromatogram(ms_data, 4)[, 1],
                       intensity = chromatogram(ms_data, 4)[, 2])
  raw205 <- subset_chromatogram(raw205, rt_min, rt_max)
  raw208 <- subset_chromatogram(raw208, rt_min, rt_max)
  if (nrow(raw205) < 2 || nrow(raw208) < 2) {
    message("Insufficient data in RT window for file ", basename(file_path))
    return(NULL)
  }

  peak205 <- get_peak_area(raw205$time, raw205$intensity, min_peak_diff)
  peak208 <- get_peak_area(raw208$time, raw208$intensity, min_peak_diff)

  if (write_plots) {
    plot_gaussian_fit(raw205, peak205$fitCurve, "Index 2", basename(file_path),
output_plot_dir)
    plot_gaussian_fit(raw208, peak208$fitCurve, "Index 4", basename(file_path),
output_plot_dir)
  }

  list(file_name = basename(file_path),
       peak205 = peak205,
       peak208 = peak208)
}

# --- Main Execution ---
```

```
147.
148. # Initialise results dataframe
149. results_df <- data.frame(
150.   file_name = character(),
151.   plant_part = character(),
152.   timepoint = character(),
153.   method_index2 = character(),
154.   auc_index2 = numeric(),
155.   method_index4 = character(),
156.   auc_index4 = numeric(),
157.   perc_index4 = numeric(),
158.   stringsAsFactors = FALSE
159. )
160.
161. # Locate all mzML files in subfolders
162. mzml_files <- list.files(base_dir, pattern = "\\.mzML$", recursive = TRUE, full.names =
TRUE)
163. cat("Found", length(mzml_files), "mzML files.\n")
164.
165. # Process each file
166. for (f in mzml_files) {
167.   rel_path <- sub(paste0(base_dir, "/?"), "", f)
168.   path_parts <- strsplit(rel_path, "/")[[1]]
169.   if (length(path_parts) < 3) {
170.     cat("Skipping file not in expected structure:", f, "\n")
171.     next
172.   }
173.   plant_part <- path_parts[1]
174.   timepoint <- path_parts[2]
175.
176.   res <- process_single_file(f, rt_min, rt_max, peak_fraction)
177.   if (is.null(res)) next
178.
179.   sum_auc <- res$peak205$auc + res$peak208$auc
180.   perc_index4 <- ifelse(sum_auc > 0, res$peak208$auc / sum_auc * 100, NA)
181.
182.   results_df <- rbind(results_df, data.frame(
183.     file_name = res$file_name,
184.     plant_part = plant_part,
185.     timepoint = timepoint,
186.     method_index2 = res$peak205$method,
187.     auc_index2 = res$peak205$auc,
188.     method_index4 = res$peak208$method,
189.     auc_index4 = res$peak208$auc,
190.     perc_index4 = perc_index4,
191.     stringsAsFactors = FALSE
192.   ))
193. }
194.
195. # Write full results and averages
196. write.csv(results_df, output_csv, row.names = FALSE)
197. cat("Results saved to:", output_csv, "\n")
198.
199. averages_df <- results_df%>%
200.   group_by(plant_part, timepoint)%>%
201.   summarise(
202.     avg_perc_index4 = mean(perc_index4, na.rm = TRUE),
203.     sd_perc_index4 = sd(perc_index4, na.rm = TRUE),
204.     n = n(),
205.     .groups = "drop"
206.   )
207. write.csv(as.data.frame(averages_df), output_avg_csv, row.names = FALSE)
208. cat("Averages saved to:", output_avg_csv, "\n")
209.
```

D.3. Identify unique peaks in treatment files

```
1. library(dplyr)
```

```r
library(fuzzyjoin)
library(readr)
library(stringr)

# ------------------------------------------------------------------
# FUNCTION: Compare Peaks for a Single Plant Part
# ------------------------------------------------------------------
compare_peaks_for_plant_part <- function(control_path, treatment_path,
                                         file_name = "negative_common_peaks.csv",
                                         threshold = 1) {
  control_file <- file.path(control_path, file_name)
  treatment_file <- file.path(treatment_path, file_name)

  if (!file.exists(control_file) || !file.exists(treatment_file)) {
    cat("Missing file for comparison:", basename(control_path), "\n")
    return(NULL)
  }

  control_df <- read.csv(control_file, stringsAsFactors = FALSE)
  treatment_df <- read.csv(treatment_file, stringsAsFactors = FALSE)

  merged_df <- fuzzy_left_join(
    treatment_df, control_df,
    by = "avg_mz",
    match_fun = function(treat, ctrl) abs(treat - ctrl) <= threshold
  )

  unique_treatment <- merged_df%>%
    filter(is.na(avg_mz.y))%>%
    rename(avg_mz = avg_mz.x,
           count = count.x,
           replicates = replicates.x,
           avg_intensity = avg_intensity.x)%>%
    select(avg_mz, count, replicates, avg_intensity)

  return(unique_treatment)
}

# ------------------------------------------------------------------
# FUNCTION: Run Comparison Across All Plant Parts and Timepoints
# ------------------------------------------------------------------
compare_control_treatment <- function(control_dir, treatment_dir,
                                      file_name = "negative_common_peaks.csv",
                                      threshold = 1,
                                      comparison_output_dir = "path/to/output") {
  if (!dir.exists(comparison_output_dir)) {
    dir.create(comparison_output_dir, recursive = TRUE)
  }

  control_timepoints <- list.dirs(control_dir, recursive = FALSE, full.names = TRUE)
  results <- list()

  for (tp in control_timepoints) {
    tp_name <- gsub(" ", "_", basename(tp))
    treatment_tp <- file.path(treatment_dir, basename(tp))
    if (!dir.exists(treatment_tp)) next

    control_parts <- list.dirs(tp, recursive = FALSE, full.names = TRUE)

    for (pp in control_parts) {
      pp_name <- gsub(" ", "_", basename(pp))
      treatment_pp <- file.path(treatment_tp, basename(pp))
      if (!dir.exists(treatment_pp)) next

      cat("Comparing:", tp_name, "-", pp_name, "\n")
      unique_peaks <- compare_peaks_for_plant_part(pp, treatment_pp, file_name,
threshold)

      if (!is.null(unique_peaks) && nrow(unique_peaks) > 0) {
```

```
        output_file_name <- paste0(tp_name, "_", pp_name, "_", file_name%>%
                                     str_replace("_common_peaks.csv",
"_unique_peaks.csv"))
        output_path <- file.path(comparison_output_dir, output_file_name)
        write.csv(unique_peaks, output_path, row.names = FALSE)
        cat("Saved:", output_path, "\n")
      } else {
        cat("No unique peaks found for:", tp_name, pp_name, "\n")
      }

      results[[paste(tp_name, pp_name, sep = "_")]] <- unique_peaks
    }
  }

  return(results)
}

# ------------------------------------------------------------------
# MAIN EXECUTION
# ------------------------------------------------------------------

# User-defined input/output directories (replace with appropriate paths)
control_dir <- "path/to/control"
treatment_dir <- "path/to/treatment"
comparison_output_dir <- "path/to/output"

# -------------------------
# 1. Compare NEGATIVE MODE
# -------------------------
comparison_results_negative <- compare_control_treatment(
  control_dir = control_dir,
  treatment_dir = treatment_dir,
  file_name = "negative_common_peaks.csv",
  threshold = 1,
  comparison_output_dir = comparison_output_dir
)

cat("\nNegative mode comparison complete.\n")

# -------------------------
# 2. Compare POSITIVE MODE
# -------------------------
comparison_results_positive <- compare_control_treatment(
  control_dir = control_dir,
  treatment_dir = treatment_dir,
  file_name = "positive_common_peaks.csv",
  threshold = 1,
  comparison_output_dir = comparison_output_dir
)

cat("\nPositive mode comparison complete.\n")
```

D.4. Metabolomics Screening Script (comparison with known list)

```
library(dplyr)

# ----------------------------
# USER SETTINGS
# ----------------------------
# Folder containing exported CSV files for a specific plant part
plant_folder <- "path/to/plant_part_folder"   # <-- Update this path as needed

# Tolerance for matching m/z values (in Da)
tol_da <- 0.5

# ----------------------------
# COMPOUND DATABASE
```

```r
# Includes deuterated and non-deuterated tryptophan derivatives,
# phytohormones, and related metabolites with their m/z values.
# ----------------------------
compounds_full <- data.frame(
  Compound = c("5-F Tryptamine", "Tryptamine", "5-OH Tryptamine", "5-OCH3 Tryptamine",
               "N-Ac Tryptamine", "N-Ac-5-OH Tryptamine", "3-OH Anthranilic acid",
               "3-OH Kynurenine", "Anthranilic acid", "Kynurenic acid", "Kynurenine",
               "Tryptophan ethyl ester", "5-OH Tryptophan", "Melatonin", "L-Tryptophan",
               "Indole-3-acetamide (IAM)", "Indole-3-pyruvic acid (IPyA)",
               "Tryptophol (Indole-3-ethanol)", "Serotonin", "Indole-3-acetic acid
(IAA)",
               "Indole-3-butyric acid (IBA)", "Naphthalene acetic acid (NAA)",
               "Trans-zeatin", "Kinetin", "6-Benzylaminopurine (BAP)",
               "Gibberellic acid (GA3)", "Gibberellin A1 (GA1)", "Gibberellin A4 (GA4)",
               "Abscisic acid (ABA)", "Jasmonic acid (JA)", "Jasmonoyl-isoleucine (JA-
Ile)",
               "Brassinolide", "Castasterone (24-epicastasterone)",
               "1-Aminocyclopropane-1-carboxylic acid (ACC)"),
  mz_neg = c(177.2, 159.2, 175.2, 189.2, 201.2, 217.3, 152.1, 223.2, 136.1, 188.2, 207.2,
             231.2, 219.2, 231.3, 203.2, 162.2, 189.2, 160.2, 175.2, 174.1, 202.2, 185.2,
             NA, NA, NA, 345.4, 331.4, 329.4, 263.3, 209.3, 322.4, 479.7, 465.7, 100.1),
  mz_pos = c(179.2, 161.2, 177.2, 191.2, 203.2, 219.3, 154.1, 225.2, 138.1, 190.2, 209.2,
             233.2, 221.2, 233.3, 205.2, 164.2, 191.2, 162.2, 177.2, 176.1, 204.2, 187.2,
             220.2, 216.2, 226.2, 347.4, 333.4, 331.4, 265.3, 211.3, 324.4, 481.7, 467.7,
102.1),
  Preferred = c("Positive", "Positive", "Positive", "Positive", "Negative", "Negative",
                "Negative", "Negative", "Negative", "Negative", "Negative", "Positive",
                "Negative", "Positive", "Positive", "Positive", "Negative", "Positive",
                "Positive", "Positive", "Negative", "Negative", "Positive", "Positive",
                "Positive", "Negative", "Negative", "Negative", "Negative", "Negative",
                "Positive", "Positive", "Negative", "Positive"),
  stringsAsFactors = FALSE
)

# ----------------------------
# FUNCTION: Process a Single CSV File (Positive or Negative)
# ----------------------------
process_csv_file <- function(file_path, scan_type, tol_da) {
  df <- read.csv(file_path, stringsAsFactors = FALSE)
  timepoint_cols <- setdiff(names(df), "merged_mz")
  res_df <- data.frame(Compound = compounds_full$Compound, stringsAsFactors = FALSE)

  for (tp in timepoint_cols) {
    res_df[[tp]] <- NA
  }

  for (i in 1:nrow(compounds_full)) {
    ref_mz <- if (scan_type == "Positive") compounds_full$mz_pos[i] else
compounds_full$mz_neg[i]
    if (is.na(ref_mz)) {
      res_df[i, timepoint_cols] <- NA
      next
    }

    mz_range <- df$merged_mz >= (ref_mz - tol_da) & df$merged_mz <= (ref_mz + tol_da)

    for (tp in timepoint_cols) {
      if (!any(mz_range)) {
        res_df[i, tp] <- "Not Found"
      } else {
        intensities <- sapply(df[mz_range, tp], function(x) {
          if (tolower(x) == "no") return(0)
          val <- suppressWarnings(as.numeric(x))
          if (is.na(val)) return(0)
          return(val)
        })
        res_df[i, tp] <- if (sum(intensities, na.rm = TRUE) > 0) "Found" else "Not Found"
      }
    }
```

```
  }
  return(res_df)
}

# ----------------------------
# MAIN PROCESSING BLOCK
# ----------------------------
csv_files <- list.files(plant_folder, pattern = "\\.csv$", full.names = TRUE)
cat("Found", length(csv_files), "CSV files in plant folder.\n")

# Identify positive and negative scan files
pos_file <- csv_files[grepl("positive", tolower(csv_files))]
neg_file <- csv_files[grepl("negative", tolower(csv_files))]

# Process positive scan file
if (length(pos_file) > 0) {
  cat("Processing positive scan:", pos_file[1], "\n")
  pos_results <- process_csv_file(pos_file[1], "Positive", tol_da)
} else {
  pos_results <- NULL
  cat("No positive CSV file found.\n")
}

# Process negative scan file
if (length(neg_file) > 0) {
  cat("Processing negative scan:", neg_file[1], "\n")
  neg_results <- process_csv_file(neg_file[1], "Negative", tol_da)
} else {
  neg_results <- NULL
  cat("No negative CSV file found.\n")
}

# Combine results into final output table
detailed_df <- data.frame(
  Compound = compounds_full$Compound,
  Positive_ref_mz = compounds_full$mz_pos,
  Negative_ref_mz = compounds_full$mz_neg,
  stringsAsFactors = FALSE
)

if (!is.null(pos_results)) {
  pos_timepoints <- setdiff(names(pos_results), "Compound")
  detailed_df <- cbind(detailed_df, pos_results[, pos_timepoints, drop = FALSE])
}

if (!is.null(neg_results)) {
  neg_timepoints <- setdiff(names(neg_results), "Compound")
  detailed_df <- cbind(detailed_df, neg_results[, neg_timepoints, drop = FALSE])
}

# Output
print(detailed_df)

output_file <- file.path(plant_folder, "Detailed_PlantPart_Results.csv")
write.csv(detailed_df, output_file, row.names = FALSE)
cat("\nDetailed results saved to:", output_file, "\n")

```

D.5. EIC Plotting of targeted compound *m/z* values

```
# Load required packages
required_packages <- c("mzR", "ggplot2", "dplyr", "tools")
for (pkg in required_packages) {
  if (!requireNamespace(pkg, quietly = TRUE)) {
    if (pkg == "mzR" && !requireNamespace("BiocManager", quietly = TRUE)) {
```

```
      install.packages("BiocManager")
      BiocManager::install("mzR")
    } else {
      install.packages(pkg)
    }
  }
  library(pkg, character.only = TRUE)
}

# ------------------------------
# USER INPUT
# ------------------------------
input_folders <- c("path/to/folder1", "path/to/folder2", ...)  # Add folders containing .mzML files
output_folder <- "path/to/output_folder"
if (!dir.exists(output_folder)) dir.create(output_folder, recursive = TRUE)

narrow_tol <- 1       # Tolerance for best-scan selection (±0.5 Da)
iso_window <- 60      # Range for isotopic envelope plotting (±60 Da)
intensity_thresh <- 1e5  # Intensity threshold for peak display

positive_targets <- c(209, 208, 182, 165, 180, 194, 207, 222, 156, 228, 142, 193, 212, 237, 224, 236)
negative_targets <- positive_targets - 2

# ------------------------------
# FUNCTIONS
# ------------------------------
extract_best_scan <- function(ms, hdr, ms1_indices, target, tol) {
  best_scan <- NA
  best_intensity <- 0
  for (scan in ms1_indices) {
    sp <- peaks(ms, scan)
    if (nrow(sp) == 0) next
    matched <- sp[sp[,1] >= (target - tol) & sp[,1] <= (target + tol), , drop = FALSE]
    tot_int <- sum(matched[,2])
    if (tot_int > best_intensity) {
      best_intensity <- tot_int
      best_scan <- scan
    }
  }
  return(best_scan)
}

process_file_for_target <- function(file, target, tol, window, intensity_thresh) {
  ms <- openMSfile(file)
  hdr <- header(ms)
  ms1_indices <- which(hdr$msLevel == 1)
  if (length(ms1_indices) == 0) return(NULL)

  best_scan <- extract_best_scan(ms, hdr, ms1_indices, target, tol)
  if (is.na(best_scan)) return(NULL)

  spec <- as.data.frame(peaks(ms, best_scan))
  colnames(spec) <- c("mz", "intensity")
  close(ms)

  spec_sub <- spec%>%
    filter(mz >= (target - window) & mz <= (target + window) & intensity >= intensity_thresh)
  if (nrow(spec_sub) == 0) return(NULL)

  return(list(spec = spec_sub, scan = best_scan, rt = hdr$retentionTime[best_scan]))
}

# ------------------------------
# MAIN LOOP
# ------------------------------
all_files <- unlist(lapply(input_folders, function(folder) {
```

```
  list.files(folder, pattern = "\\.mzML$", full.names = TRUE, recursive = TRUE)
}))

for (file in all_files) {
  base_name <- file_path_sans_ext(basename(file))
  marker <- file.path(output_folder, paste0(base_name, "_processed.txt"))
  if (file.exists(marker)) next

  for (mode in c("positive", "negative")) {
    targets <- if (mode == "positive") positive_targets else negative_targets
    color <- if (mode == "positive") "blue" else "red"

    for (target in targets) {
      res <- process_file_for_target(file, target, narrow_tol, iso_window,
intensity_thresh)
      if (is.null(res)) next

      p <- ggplot(res$spec, aes(x = mz, y = intensity)) +
        geom_segment(aes(xend = mz, yend = 0), color = color) +
        geom_point(color = color, size = 2.5) +
        labs(x = "m/z", y = "Intensity",
             title = paste("Detected m/z ~", target, "(", mode, ")"),
             subtitle = paste("Scan:", res$scan, "| RT:", round(res$rt, 2), "sec")) +
        theme_minimal()

      ggsave(filename = file.path(output_folder, paste0(base_name, "_", target, "_",
mode, ".png")),
             plot = p, width = 8, height = 6)
    }
  }

  writeLines("Processed", marker)
}
cat("All files processed. Spectra saved to:", output_folder, "\n")
```

D.6. General TIC and EIC plotting script

```
# Load necessary libraries
if (!require("mzR")) {
  if (!require("BiocManager")) install.packages("BiocManager")
  BiocManager::install("mzR")
}
if (!require("ggplot2")) install.packages("ggplot2")
if (!require("dplyr")) install.packages("dplyr")
if (!require("tools")) install.packages("tools")

library(mzR)
library(ggplot2)
library(dplyr)
library(tools)

# ------------------------------------------------------------
# Automatically find all .mzML files in the directory tree
# ------------------------------------------------------------
mzml_files <- list.files(path = ".", pattern = "\\.mzML$", recursive = TRUE, full.names =
TRUE)

# ------------------------------------------------------------
# Loop over each mzML file and extract + plot data
# ------------------------------------------------------------
for (file_path in mzml_files) {
  message("Processing file: ", file_path)

  mzml_data <- openMSfile(file_path)
  scan_headers <- header(mzml_data)
```

```r

  # Filter for positive-mode scans (polarity == 1)
  positive_scan_indices <- which(scan_headers$polarity == 1)

  if (length(positive_scan_indices) == 0) {
    message("  No positive scans found. Skipping file.")
    close(mzml_data)
    next
  }

  # Extract m/z and intensity for each positive scan
  positive_data <- do.call(rbind, lapply(positive_scan_indices, function(index) {
    mz_intensity <- peaks(mzml_data, index)
    rt <- scan_headers$retentionTime[index]
    data.frame(time = rt, mz = mz_intensity[, 1], intensity = mz_intensity[, 2])
  }))

  # ------------------------------------------------------------
  # Save extracted scan data
  # ------------------------------------------------------------
  base_name <- file_path_sans_ext(basename(file_path))
  output_dir <- file.path(dirname(file_path), paste0(base_name, "_R_output"))
  dir.create(output_dir, showWarnings = FALSE)

  csv_file <- file.path(output_dir, "positive_data.csv")
  write.csv(positive_data, file = csv_file, row.names = FALSE)

  # ------------------------------------------------------------
  # Generate profile plot for the first positive scan
  # ------------------------------------------------------------
  first_scan <- peaks(mzml_data, positive_scan_indices[1])
  profile_df <- data.frame(mz = first_scan[, 1], intensity = first_scan[, 2])

  profile_plot <- ggplot(profile_df, aes(x = mz, y = intensity)) +
    geom_line(color = "blue") +
    labs(title = paste("Profile Mode -", base_name),
         x = "m/z", y = "Intensity") +
    theme_minimal()

  ggsave(filename = file.path(output_dir, "positive_profile_plot.svg"),
         plot = profile_plot, width = 8, height = 6)

  # ------------------------------------------------------------
  # Generate centroid-style plot (intensity > threshold)
  # ------------------------------------------------------------
  threshold <- 5000
  centroid_data <- positive_data%>% filter(intensity > threshold)

  centroid_plot <- ggplot(centroid_data, aes(x = mz, y = intensity)) +
    geom_segment(aes(xend = mz, yend = 0), color = "red") +
    labs(title = paste("Centroid Mode -", base_name),
         x = "m/z", y = "Intensity") +
    theme_minimal()

  ggsave(filename = file.path(output_dir, "positive_centroid_plot.svg"),
         plot = centroid_plot, width = 8, height = 6)

  # ------------------------------------------------------------
  # Generate TIC (Total Ion Current) plots
  # ------------------------------------------------------------
  if ("totalIonCurrent"%in% names(scan_headers)) {
    tic_values <- scan_headers$totalIonCurrent[positive_scan_indices]
  } else if ("totIonCurrent"%in% names(scan_headers)) {
    tic_values <- scan_headers$totIonCurrent[positive_scan_indices]
  } else {
    warning("No TIC field found in header.")
    tic_values <- rep(NA, length(positive_scan_indices))
  }

```

```
  tic_data <- data.frame(
    time_sec = scan_headers$retentionTime[positive_scan_indices] * 60,
    TIC = tic_values
  )

  # TIC profile plot
  tic_profile <- ggplot(tic_data, aes(x = time_sec, y = TIC)) +
    geom_line(color = "blue") +
    labs(title = paste("TIC Profile -", base_name),
         x = "Retention Time (sec)", y = "Total Ion Current") +
    theme_minimal()

  ggsave(filename = file.path(output_dir, "tic_profile_plot.svg"),
         plot = tic_profile, width = 8, height = 6)

  # TIC centroid-style plot
  tic_centroid <- ggplot(tic_data, aes(x = time_sec, y = TIC)) +
    geom_segment(aes(xend = time_sec, yend = 0), color = "red") +
    labs(title = paste("TIC Centroid -", base_name),
         x = "Retention Time (sec)", y = "Total Ion Current") +
    theme_minimal()

  ggsave(filename = file.path(output_dir, "tic_centroid_plot.svg"),
         plot = tic_centroid, width = 8, height = 6)

  # Clean up
  close(mzml_data)
  message("Finished processing file: ", file_path, "\n")
}

```

D.7. Auxin EIC plotting Script

```
# Clear environment
rm(list = ls())

# Required functions
roundUpNice <- function(x, nice = c(1, 1.6, 2, 2.6, 3, 3.6, 4, 4.6, 5, 5.6, 6, 6.6, 7,
7.6, 8, 8.6, 9, 9.6, 10)) {
  10^floor(log10(x)) * nice[[which(x <= 10^floor(log10(x)) * nice)[[1]]]]
}

rounder <- function(x, y) {
  if (y >= 0) x + (y - x%% y) else x - (x%% abs(y))
}

# ------------------------------
# USER SETTINGS
# ------------------------------
process.data <- "y"       # Process raw mzXML files using xcms
read.data <- "n"          # Read combined EIC CSV instead of processing raw data
generate.plots <- "y"     # Create PDF of EIC plots

EIC.mz.range <- c(427.45, 427.45)
EIC.mz.tol <- 0.25
EIC.range <- c(EIC.mz.range[1] - EIC.mz.tol, EIC.mz.range[2] + EIC.mz.tol)

output.dir <- "path/to/output"             # Output directory for CSVs and plots
experiment.name <- "experiment_id"         # ID used in output filenames

# ------------------------------
# PROCESS RAW DATA
# ------------------------------
if (process.data == "y") {
  rm(list = setdiff(ls(), c("roundUpNice", "rounder", "process.data", "read.data",
"generate.plots",
                            "output.dir", "experiment.name", "EIC.mz.range",
"EIC.mz.tol", "EIC.range")))
```

```r

  input.list <- dir("path/to/mzxml_files", pattern = ".mzXML", full.names = TRUE)

  for (a in seq_along(input.list)) {
    file.ID <- tools::file_path_sans_ext(basename(input.list[a]))
    sample.ID.parts <- unlist(strsplit(file.ID, "_"))
    sample.ID <- paste(sample.ID.parts[(length(sample.ID.parts)-
1):length(sample.ID.parts)], collapse = "_")

    inputspec <- xcms::xcmsRaw(input.list[a], profstep = 0, includeMSn = FALSE)
    EIC.spec <- xcms::getEIC(inputspec, mzrange = EIC.range, step = 0.1)
    file.MAT <- EIC.spec@eic$xcmsRaw[[1]]
    colnames(file.MAT) <- c("time", sample.ID)

    # Optional: Save individual EIC CSV
    output.raw <- file.path(output.dir, paste0(file.ID, "_", EIC.mz.range[1],
"_EIC.csv"))
    write.csv(file.MAT, output.raw, row.names = FALSE)

    # Optional: Baseline correction
    scanrange.bcorr <- c(
      which(file.MAT[,1] > 110 & file.MAT[,1] < 120)[1],
      which(file.MAT[,1] > 240 & file.MAT[,1] < 250)[1]
    )
    bcorrval <- median(file.MAT[which(file.MAT[scanrange.bcorr[1]:scanrange.bcorr[2],2] >
0) + scanrange.bcorr[1] - 1, 2])
    file.MAT[,2] <- file.MAT[,2] - bcorrval

    assign(paste0("xy_EIC_", sample.ID), file.MAT)

    # Track TIC for normalisation
    if (a == 1) {
      file.TIC <- data.frame(Sample = sample.ID, TIC = sum(inputspec@tic))
    } else {
      file.TIC <- rbind(file.TIC, c(sample.ID, sum(inputspec@tic)))
    }
  }

  # ------------------------------
  # NORMALISATION (by TIC)
  # ------------------------------
  normalise.intensity <- "y"
  if (normalise.intensity == "y") {
    normalise.value <- as.numeric(file.TIC[1,2])
    MAT.files <- ls(pattern = "xy_EIC_")
    EIC.files <- sub("xy_EIC_", "", MAT.files)

    for (b in seq_along(MAT.files)) {
      match.pos <- match(EIC.files[b], file.TIC[,1])
      TIC.value <- as.numeric(file.TIC[match.pos, 2])
      assign(MAT.files[b], cbind(get(MAT.files[b])[,1], get(MAT.files[b])[,2] *
(normalise.value / TIC.value)))
    }
  }

  # ------------------------------
  # WRITE COMBINED DATA
  # ------------------------------
  write.comb.tic.data <- "y"
  if (write.comb.tic.data == "y") {
    output.mat.comb <- do.call(cbind, lapply(ls(pattern = "xy_EIC_"), function(obj)
get(obj)))
    colnames(output.mat.comb) <- unlist(lapply(ls(pattern = "xy_EIC_"), function(obj)
c("time", obj)))
    output.comb.file <- file.path(output.dir, paste0(experiment.name, "_all_EIC_",
EIC.mz.range[1], ".csv"))
    write.csv(output.mat.comb, output.comb.file, row.names = FALSE)
  }
}
```

```
 95.
 96. # ------------------------------
 97. # READ PRE-COMBINED DATA (Optional)
 98. # ------------------------------
 99. if (read.data == "y") {
100.   output.comb.file <- file.path(output.dir, paste0(experiment.name, "_all_EIC_",
EIC.mz.range[1], ".csv"))
101.   output.mat.comb <- read.csv(output.comb.file)
102.   for (k in seq(2, ncol(output.mat.comb), by = 2)) {
103.     assign(colnames(output.mat.comb)[k], output.mat.comb[ , (k-1):k])
104.   }
105.   MAT.files <- ls(pattern = "xy_EIC_")
106. }
107.
108. # ------------------------------
109. # GENERATE EIC PLOTS (PDF)
110. # ------------------------------
111. if (generate.plots == "y") {
112.   plot.output <- file.path(output.dir, paste0(experiment.name, "_all_EICS_",
EIC.mz.range[1], ".pdf"))
113.   pdf(file = plot.output, width = 9.736, height = 6.491, pointsize = 12, family =
"Helvetica")
114.
115.   MAT.files <- ls(pattern = "xy_EIC_")
116.   y.max <- max(sapply(MAT.files, function(f) max(get(f)[,2])))
117.   y.min <- min(sapply(MAT.files, function(f) min(get(f)[,2])))
118.   y.scale.factor <- floor(log10(y.max)) - 1
119.
120.   yrange <- c(-0.05, roundUpNice(y.max) / (10^y.scale.factor))
121.   xrange <- c(0, 390)
122.
123.   datacol.list <- c('#e6194b', '#3cb44b', '#4363d8', '#f58231', '#911eb4', '#46f0f0',
'#f032e6',
124.                     '#bcf60c', '#fabebe', '#008080', '#e6beff', '#9a6324')
125.
126.   par(mar = c(7,8,2,2) + 0.5, oma = c(1,1.5,0,1), lwd = 4, mgp = c(8,2.6,0),
127.       cex.lab = 2.5, cex.axis = 1.95, cex.main = 3, xaxs = "i", yaxs = "i")
128.
129.   for (e in seq_along(MAT.files)) {
130.     datacol <- datacol.list[e%% length(datacol.list) + 1]
131.     data <- get(MAT.files[e])
132.     scaled_intensity <- data[,2] / (10^y.scale.factor)
133.     if (e == 1) {
134.       plot(data[,1], scaled_intensity, type = "l", col = datacol, lwd = 2,
135.            xlab = "", ylab = "", xlim = xrange, ylim = yrange)
136.     } else {
137.       lines(data[,1], scaled_intensity, col = adjustcolor(datacol, alpha = 0.5), lwd = 2)
138.     }
139.   }
140.
141.   mtext("Retention Time (s)", 1, outer = TRUE, cex = 2.85, line = -1.64)
142.   axis_label <- bquote(paste("10"^.-(y.scale.factor), " × EIC Intensity (arb. units)"))
143.   mtext(axis_label, 2, srt = 90, outer = TRUE, cex = 2.85, line = -2.42)
144.
145.   dev.off()
146. }
147.
```

Appendix E: Mass Spectra for plant analysis

E.1. Peak intensity-based calculation of deuterated tryptophan in MRM data

E.1.a. Control samples (Experiment 1)

Extract Code	Treatment	Time point	Intentisty of peak at m/z 205	Intentisty of peak at m/z 208	deuterated as percentage of all Trp
1_1_MS	Control - water	4h	2041	122	5.640314378
1_1_STEM	Control - water	4h	1448	279	16.1551824
1_1_T2	Control - water	4h	1642	112	6.385404789
1_1_T3_washings	Control - water	4h	4149	218	4.991985345
1_1_T3spot	Control - water	4h	22345	280	1.237569061
1_1_T3w	Control - water	4h	2671	249	8.52739726
1_1_T4	Control - water	4h	4882	245	4.778622976
1_2_MS	Control - water	4h	2565	173	6.318480643
1_2_STEM	Control - water	4h	2196	153	6.513409962
1_2_T2	Control - water	4h	2270	274	10.77044025
1_2_T3_washings	Control - water	4h	3715	103	2.697747512
1_2_T3spot	Control - water	4h	514	131	20.31007752
1_2_T3uw	Control - water	4h	3655	228	5.871748648
1_2_T3w	Control - water	4h	1053	154	12.75890638
1_2_T4	Control - water	4h	4882	255	4.963986763
1_3_MS	Control - water	4h	8048	254	3.059503734
1_3_STEM	Control - water	4h	4070	271	6.242801198
1_3_T2	Control - water	4h	1514	233	13.3371494
1_3_T3_washings	Control - water	4h	5379	233	4.151817534
1_3_T3spot	Control - water	4h	689	219	24.11894273
1_3_T3uw	Control - water	4h	2083	256	10.94484823
1_3_T3w	Control - water	4h	2773	346	11.09329913
1_3_T4	Control - water	4h	1669	10	0.595592615
1_5_MS	Control - water	24h	284624	2527	0.880024795
1_5_STEM	Control - water	24h	103084	4085	3.811736603
1_5_T2	Control - water	24h	31940	3802	10.63734542
1_5_T3spot	Control - water	24h	1156	2163	65.170232
1_5_T3uw	Control - water	24h	70244	2210	3.050211168
1_5_T3w	Control - water	24h	38000	3056	7.443491816
1_5_T4	Control - water	24h	62472	3139	4.784258737
1_6_MS	Control - water	24h	307232	1761	0.569915823
1_6_STEM	Control - water	24h	49328	2041	3.973213417
1_6_T2	Control - water	24h	45764	2697	5.565299932
1_6_T3spot	Control - water	24h	794	2210	73.56857523
1_6_T3uw	Control - water	24h	39808	2068	4.938389531
1_6_T3w	Control - water	24h	24293	3019	11.05374927
1_6_T4	Control - water	24h	114128	1448	1.252855264
1_1_Blank_average	Averages	1_1,1_2,1_3	59	272.5	-
1_5_blank_average	Averages	1_5,1_6	1286	3637	-

E.1.b. Experiment 2

Extract Code	Treatment	Time point	M/z 205	Intentisty of peak at m/z 208	deuterated as percentage of all Trp
2_15_MS	L-Tryptophan-d3 - pH 1	4h	8881	213	2.342203651
2_15_STEM	L-Tryptophan-d3 - pH 1	4h	6698	226	3.264009243
2_15_T2	L-Tryptophan-d3 - pH 1	4h	6166	236	3.686348016
2_15_T3spot	L-Tryptophan-d3 - pH 1	4h	1571	36412	95.86393913
2_15_T3uw	L-Tryptophan-d3 - pH 1	4h	6635	12003	64.40068677
2_15_T3w	L-Tryptophan-d3 - pH 1	4h	3166	755	19.25529202
2_42_MS	6 mM L-Tryptophan-d5 - pH 1	24h	15795	197	1.231865933
2_42_STEM	6 mM L-Tryptophan-d5 - pH 1	24h	4483	149	3.216753022
2_42_T2	6 mM L-Tryptophan-d5 - pH 1	24h	6002	43	0.711331679
2_42_T3spot	6 mM L-Tryptophan-d5 - pH 1	24h	120	26288	99.54559224
2_42_T3uw	6 mM L-Tryptophan-d5 - pH 1	24h	3854	14755	79.28959106
2_42_T3w	6 mM L-Tryptophan-d5 - pH 1	24h	6343	13988	68.80133786

E.1.c. L-Tryptophan pH 7

Extract Code	Treatment	Time point	M/z 205	Intentisty of peak at m/z 208	deuterated as percentage of all Trp
4_1_MS	L-Tryptophan-d3 - pH adjusted	4h	136340	21504	13.62357771
4_1_STEM	L-Tryptophan-d3 - pH adjusted	4h	21736	7044	24.47533009
4_1_T2	L-Tryptophan-d3 - pH adjusted	4h	13414	6341	32.09820299
4_1_T3spot	L-Tryptophan-d3 - pH adjusted	4h	464	35442	98.70773687
4_1_T3uw	L-Tryptophan-d3 - pH adjusted	4h	13134	19367	59.58893573
4_1_T3w	L-Tryptophan-d3 - pH adjusted	4h	36798	20473	35.74758604
4_2_MS	L-Tryptophan-d3 - pH adjusted	4h	8476	134	1.556329849
4_2_STEM	L-Tryptophan-d3 - pH adjusted	4h	5480	312	5.386740331
4_2_T3spot	L-Tryptophan-d3 - pH adjusted	4h	117	22960	99.49300169
4_2_T3uw	L-Tryptophan-d3 - pH adjusted	4h	6195	7654	55.26752834
4_2_T3w	L-Tryptophan-d3 - pH adjusted	4h	4526	3414	42.99748111
4_3T_MS	L-Tryptophan-d3 - pH adjusted	4h	206352	713	0.34433632
4_3T_STEM	L-Tryptophan-d3 - pH adjusted	4h	136896	10010	6.813880985
4_3T_T2	L-Tryptophan-d3 - pH adjusted	4h	118256	1147	0.96061238
4_3T_T3spot	L-Tryptophan-d3 - pH adjusted	4h	6120	414480	98.54493581
4_3T_T3w	L-Tryptophan-d3 - pH adjusted	4h	108980	38060	25.88411317
4_3T_T4	L-Tryptophan-d3 - pH adjusted	4h	204768	939	0.4564745
4_3T3uw	L-Tryptophan-d3 - pH adjusted	4h	115600	136192	54.08908941
4_4T_MS	L-Tryptophan-d3 - pH adjusted	4h	86324	714	0.82033135
4_4T_STEM	L-Tryptophan-d3 - pH adjusted	4h	105852	382	0.359583561
4_4T_T2	L-Tryptophan-d3 - pH adjusted	4h	59416	139	0.2333977
4_4T_T3spot	L-Tryptophan-d3 - pH adjusted	4h	107	5386	98.05206627
4_4T_T3w	L-Tryptophan-d3 - pH adjusted	4h	142480	46216	24.49230508
4_4T_T4	L-Tryptophan-d3 - pH adjusted	4h	111092	346	0.310486549
4_4T3uw	L-Tryptophan-d3 - pH adjusted	4h	77048	10007	11.49503188
4_5_MS	L-Tryptophan-d3 - pH adjusted	24h	26025	6939	21.05023662
4_5_STEM	L-Tryptophan-d3 - pH adjusted	24h	933	5337	85.11961722
4_5_T2	L-Tryptophan-d3 - pH adjusted	24h	60406	5261	8.011634459
4_5_T3spot	L-Tryptophan-d3 - pH adjusted	24h	611	49838	98.78887589
4_5_T3uw	L-Tryptophan-d3 - pH adjusted	24h	7617	6320	45.34691828
4_5_T3w	L-Tryptophan-d3 - pH adjusted	24h	9265	6474	41.13349006
4_6T_MS	L-Tryptophan-d3 - pH adjusted	24h	135488	223	0.164319768
4_6T_STEM	L-Tryptophan-d3 - pH adjusted	24h	176272	1021	0.575882861
4_6T_T2	L-Tryptophan-d3 - pH adjusted	24h	135504	915	0.670727685
4_6T_T3spot	L-Tryptophan-d3 - pH adjusted	24h	279	56636	99.50979531
4_6T_T3uw	L-Tryptophan-d3 - pH adjusted	24h	168832	132928	44.05090138
4_6T_T3w	L-Tryptophan-d3 - pH adjusted	24h	323072	61944	16.08868203
4_6T_T4	L-Tryptophan-d3 - pH adjusted	24h	253424	406	0.159949573
4_7T_MS	L-Tryptophan-d3 - pH adjusted	24h	119120	1100	0.914989186
4_7T_STEM	L-Tryptophan-d3 - pH adjusted	24h	67520	813	1.189761901
4_7T_T2	L-Tryptophan-d3 - pH adjusted	24h	71772	1549	2.112628033
4_7T_T3spot	L-Tryptophan-d3 - pH adjusted	24h	27	82780	99.96739406
4_7T_T3w	L-Tryptophan-d3 - pH adjusted	24h	133696	47784	26.33017412
4_7T_T4	L-Tryptophan-d3 - pH adjusted	24h	175616	1112	0.629215518
4_7T3uw	L-Tryptophan-d3 - pH adjusted	24h	118736	101048	45.97604921
4_10T_MS	L-Tryptophan-d3 - pH adjusted	48h	221232	808	0.363898397
4_10T_STEM	L-Tryptophan-d3 - pH adjusted	48h	157152	1271	0.802282497
4_10T_T2	L-Tryptophan-d3 - pH adjusted	48h	78000	324	0.413666309
4_10T_T3spot	L-Tryptophan-d3 - pH adjusted	48h	64	59388	99.89235013
4_10T_T3w	L-Tryptophan-d3 - pH adjusted	48h	127168	24996	16.42701296
4_10T_T4	L-Tryptophan-d3 - pH adjusted	48h	206864	407	0.196361285

4_10T3uw	L-Tryptophan-d3 - pH adjusted	48h	151264	146224	49.152907
4_11T_MS	L-Tryptophan-d3 - pH adjusted	48h	166816	275	0.1645809
4_11T_STEM	L-Tryptophan-d3 - pH adjusted	48h	158928	1017	0.6358435
4_11T_T2	L-Tryptophan-d3 - pH adjusted	48h	88920	977	1.0867993
4_11T_T3spot	L-Tryptophan-d3 - pH adjusted	48h	19	58760	99.967675
4_11T_T3w	L-Tryptophan-d3 - pH adjusted	48h	117152	11995	9.2878657
4_11T_T4	L-Tryptophan-d3 - pH adjusted	48h	102084	227	0.2218725
4_11T3uw	L-Tryptophan-d3 - pH adjusted	48h	185536	180992	49.380129
4_9T_MS	L-Tryptophan-d3 - pH adjusted	48h	244576	597	0.2435015
4_9T_STEM	L-Tryptophan-d3 - pH adjusted	48h	132560	3307	2.4339979
4_9T_T2	L-Tryptophan-d3 - pH adjusted	48h	100668	903	0.8890332
4_9T_T3spot	L-Tryptophan-d3 - pH adjusted	48h	201	2408	92.295898
4_9T_T3w	L-Tryptophan-d3 - pH adjusted	48h	113008	19694	14.840771
4_9T_T4	L-Tryptophan-d3 - pH adjusted	48h	192048	453	0.2353234
4_9T3uw	L-Tryptophan-d3 - pH adjusted	48h	135744	95888	41.396698
3_13T_T3spot	L-Tryptophan-d3 - pH adjusted	72h	137	71256	99.808104
4_13T_MS	L-Tryptophan-d3 - pH adjusted	72h	62484	200	0.3190606
4_13T_STEM	L-Tryptophan-d3 - pH adjusted	72h	27980	178	0.632147
4_13T_T2	L-Tryptophan-d3 - pH adjusted	72h	8851	19	0.2142051
4_13T_T3uw	L-Tryptophan-d3 - pH adjusted	72h	34568	40512	53.958444
4_13T_T3w	L-Tryptophan-d3 - pH adjusted	72h	37016	3762	9.2255628
4_13T_T4	L-Tryptophan-d3 - pH adjusted	72h	113936	40	0.0350951
4_17_MS	L-Tryptophan-d3 - pH adjusted	1 week	36666	6690	15.430390
4_17_STEM	L-Tryptophan-d3 - pH adjusted	1 week	30601	3057	9.0825360
4_17_T2	L-Tryptophan-d3 - pH adjusted	1 week	15449	3718	19.397923
4_17_T3spot	L-Tryptophan-d3 - pH adjusted	1 week	3890	106184	96.466013
4_17_T3uw	L-Tryptophan-d3 - pH adjusted	1 week	9579	10771	52.928746
4_17_T3w	L-Tryptophan-d3 - pH adjusted	1 week	34122	8407	19.767687
4_22T_MS	L-Tryptophan-d3 - pH adjusted	4 weeks	4010	158	3.7907869
4_22T_Pod	L-Tryptophan-d3 - pH adjusted	4 weeks	15660	211	1.3294688
4_22T_STEM	L-Tryptophan-d3 - pH adjusted	4 weeks	4805	322	6.2804759
4_22T_T2	L-Tryptophan-d3 - pH adjusted	4 weeks	10057	179	1.7487299
4_22T_T3spot	L-Tryptophan-d3 - pH adjusted	4 weeks	93	293	75.906735
4_22T_T3uw	L-Tryptophan-d3 - pH adjusted	4 weeks	5790	210	3
4_22T_T3w	L-Tryptophan-d3 - pH adjusted	4 weeks	9515	200	2.0586721
4_23T_MS	L-Tryptophan-d3 - pH adjusted	4 weeks	8903	371	4.0004313
4_23T_Pod	L-Tryptophan-d3 - pH adjusted	4 weeks	132384	173	0.1305098
4_23T_STEM	L-Tryptophan-d3 - pH adjusted	4 weeks	37068	67	0.1804227
4_23T_T2	L-Tryptophan-d3 - pH adjusted	4 weeks	33636	212	0.6263294
4_23T_T3spot	L-Tryptophan-d3 - pH adjusted	4 weeks	1176	182	13.402061
4_23T_T3uw	L-Tryptophan-d3 - pH adjusted	4 weeks	20093	2447	10.856255
4_23T_T3uw2	L-Tryptophan-d3 - pH adjusted	4 weeks	11468	3138	21.484321
4_23T_T3w	L-Tryptophan-d3 - pH adjusted	4 weeks	14651	356	2.3722262
4_23T_T3w2	L-Tryptophan-d3 - pH adjusted	4 weeks	20730	2230	9.7125435
4_23T_T4	L-Tryptophan-d3 - pH adjusted	4 weeks	19882	229	1.1386803
4_24T_MS	L-Tryptophan-d3 - pH adjusted	4 weeks	64988	345	0.5280639
4_24T_pod	L-Tryptophan-d3 - pH adjusted	4 weeks	136032	349	0.2559007
4_24T_STEM	L-Tryptophan-d3 - pH adjusted	4 weeks	59652	388	0.6462358
4_24T_T2	L-Tryptophan-d3 - pH adjusted	4 weeks	38772	309	0.7906655
4_24T_T3spot	L-Tryptophan-d3 - pH adjusted	4 weeks	2415	152	5.9213089
4_24T_T3uw	L-Tryptophan-d3 - pH adjusted	4 weeks	35480	437	1.2166940
4_24T_T3w	L-Tryptophan-d3 - pH adjusted	4 weeks	8691	335	3.7115001
4_24T_T4	L-Tryptophan-d3 - pH adjusted	4 weeks	56432	238	0.4199752

E.1.d. Blanks

Extract Code	Treatment	Close to:	Intentisty of peak at m/z 205	Intentisty of peak at m/z 208
3-48 blank	blank	-	1010	6891
4_24T_Blank10	blank		102	262
4_24T_Blank11	blank		263	315
MeOH 2	blank	4_1,4_5,4_17,3_48	311	5731
MeOH 3	blank	4_1,4_5,4_17,3_48	418	7456
Control	blank	4_1,4_5,4_17,3_48	3027	3272
MeOH	blank	4_1,4_5,4_17,3_48	1133	6891
4_10T_Blank1	blank	4_10,4_11,4_4,4_9,4_6,4_7	169	139
4_10T_Blank2	blank	4_10,4_11,4_4,4_9,4_6,4_7	186	111
4_10T_Blank3	blank	4_10,4_11,4_4,4_9,4_6,4_7	289	139
4_10T_Blank4	blank	4_10,4_11,4_4,4_9,4_6,4_7	173	177
4_10T_Blank5	blank	4_10,4_11,4_4,4_9,4_6,4_7	138	260
4_10T_Blank6	blank	4_10,4_11,4_4,4_9,4_6,4_7	201	203
4_10T_Blank7	blank	4_10,4_11,4_4,4_9,4_6,4_7	228	184
4_10T_Blank8	blank	4_10,4_11,4_4,4_9,4_6,4_7	231	155
4_10T_Blank9	blank	4_10,4_11,4_4,4_9,4_6,4_7	317	205
4_10T_Blank10	blank	4_10,4_11,4_4,4_9,4_6,4_7	159	255
4_10T_Blank11	blank	4_10,4_11,4_4,4_9,4_6,4_7	290	186
4_10T_Blank12	blank	4_10,4_11,4_4,4_9,4_6,4_7	178	145
4_10T_Blank13	blank	4_10,4_11,4_4,4_9,4_6,4_7	104	155
4_10T_Blank14	blank	4_10,4_11,4_4,4_9,4_6,4_7	120	155
4_10T_Blank15	blank	4_10,4_11,4_4,4_9,4_6,4_7	72	116
BLANK	blank	-	1069	3339
Blank13	blank	-	104	155
Blank14	blank	-	35	32
Blank15	blank	-	29	73
Blank15(2)	blank	-	34	22
Blank16	blank	-	26	24
BLANK2	blank	-	1314	3974
BLANK3	blank	-	774	2234
BLANK4	blank	-	999	2012
Blank 0	blank	1_1,1_2,1_3	51	222
Blank 1	blank	1_1,1_2,1_3	67	323
4_13T_Blank1	blank	-	502	38
4_13T_Blank2	blank	-	263	136
4_13T_Blank3	blank	-	198	10
4_13T_Blank4	blank		277	369
4_23T_Blank5	blank		233	107
4_23T_Blank6	blank	-	417	26
BLANK	blank	1_5,1_6	1264	3300
BLANK 2	blank	1_5,1_6,4_3	1308	3974
BLANK 3	blank	4_3	1258	2434
4_23T_Blank7	blank		383	252
4_23T_Blank8	blank		287	243
4_24T_Blank9	blank		102	203
4_24T_Blank12	blank		386	369
4_24T_Blank13	blank		406	251

E.1.e. Blank averages

Extract Code	Treatment	Close to:	m/z 205	Intentisty of peak at m/z 208
Blank T2 average	Averages	4_10,4_11,4_4,4_9,4_6,4_7	204.25	129.6666667
Blank T4 average	Averages	4_10,4_11,4_4,4_9,4_6,4_7	200	192
Blank STEM average	Averages	4_10,4_11,4_4,4_9,4_6,4_7	189	215.6666667
Blank MS average	Averages	4_10,4_11,4_4,4_9,4_6,4_7	258.6667	181.3333333
Blank T3w average	Averages	4_10,4_11,4_4,4_9,4_6,4_7	255.3333	215.3333333
Blank T3uw average	Averages	4_10,4_11,4_4,4_9,4_6,4_7	234	165.5
Blank T3spot average	Averages	4_10,4_11,4_4,4_9,4_6,4_7	98.66667	142
4_1_blank_average	Averages	4_1,4_5,4_17,3_48	1222.25	6692.666667
1_1_Blank_average	Averages	1_1,1_2,1_3	59	272.5
4_13T_Blank_average	Averages		310	138.25
1_5_blank_average	Averages	1_5,1_6	1286	3637
4_3_blank_average	Averages	4_3	1283	3204
4_23T_Blank_average	Averages	4_23	283.1667	200
4_24T_Blankaverage		4_24	257.6667	273.8333333
4_21C_blanks	averages	4_21C	107.75	141
4_22T_blanks	averages	4_22T	95.16667	195.6666667

E.1.f. 5-Fluoro-L-tryptophan samples

Extract Code	Treatment	Close to:	Intentisty of peak at m/z 223	Notes
BCRH_5_10F_MS	5-F Trp	1 week	72	
BCRH_5_10F_Pod	5-F Trp	1 week	22	
BCRH_5_10F_STEM	5-F Trp	1 week	noise	
BCRH_5_10F_T2	5-F Trp	1 week	28	
BCRH_5_10F_T3w	5-F Trp	1 week	35	
BCRH_5_10F_T4	5-F Trp	1 week	noise	
BCRH_5_10F_T3uw	5-F Trp	1 week	78	
BCRH_5_11F_MS	5-F Trp	1 week	33	
BCRH_5_11F_STEM	5-F Trp	1 week	82	
BCRH_5_11F_T2	5-F Trp	1 week	22	
BCRH_5_11F_T3w	5-F Trp	1 week	480	
BCRH_5_11F_T4	5-F Trp	1 week	noise	
BCRH_5_11F_T3uw	5-F Trp	1 week	3084	
BCRH_5_12F_MS	5-F Trp	1 week	46	
BCRH_5_12F_Pod	5-F Trp	1 week	noise	
BCRH_5_12F_STEM	5-F Trp	1 week	40	
BCRH_5_12F_T2	5-F Trp	1 week	43	
BCRH_5_12F_T3w	5-F Trp	1 week	265	
BCRH_5_12F_T4	5-F Trp	1 week	noise	
BCRH_5_12F_T3uw	5-F Trp	1 week	4763	
BCRH_5_13F_MS	5-F Trp	4 weeks	41	
BCRH_5_13F_Pod	5-F Trp	4 weeks	60	weak but clear peak showing presence above noise!
BCRH_5_13F_STEM	5-F Trp	4 weeks	noise	
BCRH_5_13F_T2	5-F Trp	4 weeks	noise	
BCRH_5_13F_T3w	5-F Trp	4 weeks	260	
BCRH_5_13F_T4	5-F Trp	4 weeks	noise	
BCRH_5_13F_T3uw	5-F Trp	4 weeks	719	
BCRH_5_14F_MS	5-F Trp	4 weeks	37	
BCRH_5_14F_STEM	5-F Trp	4 weeks	noise	
BCRH_5_14F_T2	5-F Trp	4 weeks	48	
BCRH_5_14F_T3w	5-F Trp	4 weeks	288	
BCRH_5_14F_T4	5-F Trp	4 weeks	noise	
BCRH_5_14F_T3uw	5-F Trp	4 weeks	690	
BCRH_5_15F_MS	5-F Trp	4 weeks	noise	
BCRH_5_15F_STEM	5-F Trp	4 weeks	noise	
BCRH_5_15F_T2	5-F Trp	4 weeks	40	
BCRH_5_15F_T3w	5-F Trp	4 weeks	282	
BCRH_5_15F_T4	5-F Trp	4 weeks	noise	
BCRH_5_15F_T3uw	5-F Trp	4 weeks	3252	

E.2. Area under the curve full data

E.2.a. L-Trp pH 7 rate of uptake data

file_name	plant_part	timepoint	peak205_auc	method205	peak208_auc	method208	percentages
BCRH_4_17_T3_spot_179283.mzML	T3spot	1 week	9.57014985	gaussian	1009.191817	gaussian	99.06
BCRH_4_5_T3_spot_179277.mzML	T3spot	24 hours	13.17564829	gaussian	459.7270248	gaussian	97.21
BCRH_4_6T_T3spot_201829.mzML	T3spot	24 hours	16.48860007	gaussian	5433.28969	gaussian	99.70
BCRH_4_7T_T3spot_201830.mzML	T3spot	24 hours	8.436447304	gaussian	8092.38056	gaussian	99.90
BCRH_4_2_T3spot_184717.mzML	T3spot	4 hours	3.79070012	gaussian	1455.636204	gaussian	99.74
BCRH_4_3T_T3spot_200115.mzML	T3spot	4 hours	586.6864724	gaussian	34188.69148	gaussian	98.31
BCRH_4_4T_T3spot_201827.mzML	T3spot	4 hours	3.084867256	gaussian	1018.289574	gaussian	99.70
BCRH_4_22T_T3spot_183633.mzML	T3spot	4 weeks	2.312527955	gaussian	0.591478012	gaussian	20.37
BCRH_4_23T_T3spot_209683.mzML	T3spot	4 weeks	191.0129512	gaussian	0.892310454	gaussian	0.46
BCRH_4_24T_T3spot_209695.mzML	T3spot	4 weeks	567.1000233	gaussian	0.865563394	gaussian	0.15
BCRH_4_10T_T3spot_201825.mzML	T3spot	48 hours	16.16486836	gaussian	6017.311168	gaussian	99.73
BCRH_4_11T_T3spot_201826.mzML	T3spot	48 hours	7.470001437	gaussian	5792.344091	gaussian	99.87
BCRH_4_9T_T3spot_201828.mzML	T3spot	48 hours	12.92222812	gaussian	4477.751623	gaussian	99.71
BCRH_4_13T_T3spot_209671.mzML	T3spot	72 hours	17.9668878	gaussian	9814.98471	gaussian	99.82
BCRH_4_14T_T3spot_216889.mzML	T3spot	72 hours	41.08435984	trapezoidal	2991.142411	gaussian	98.65
BCRH_4_15T_T3spot_216888.mzML	T3spot	72 hours	23.20328599	gaussian	2804.605862	gaussian	99.18
BCRH_4_17_T3_uw_179284.mzML	T3uw	1 week	79.29845896	gaussian	105.091701	gaussian	56.99
BCRH_4_5_T3_uw_179278.mzML	T3uw	24 hours	47.5901304	gaussian	48.13524337	gaussian	50.28
BCRH_4_6T_T3uw_201820.mzML	T3uw	24 hours	14284.48739	gaussian	10949.10973	gaussian	43.39
BCRH_4_7T_T3uw_201821.mzML	T3uw	24 hours	9502.842187	gaussian	8135.644992	gaussian	46.12
BCRH_4_2_T3uw_184710.mzML	T3uw	4 hours	335.6773089	gaussian	434.5823278	gaussian	56.42
BCRH_4_3T3uw_200111.mzML	T3uw	4 hours	8636.642778	gaussian	10233.92302	gaussian	54.23
BCRH_4_4T_T3uw_201818.mzML	T3uw	4 hours	6819.024992	gaussian	12783.07831	gaussian	65.21
BCRH_4_22T_T3uw_183632.mzML	T3uw	4 weeks	424.8929042	gaussian	0.899037283	gaussian	0.21
BCRH_4_23T_T3uw_209681.mzML	T3uw	4 weeks	3617.6953	gaussian	374.9743415	gaussian	9.39
BCRH_4_23T_T3uw2_209682.mzML	T3uw	4 weeks	1830.011607	gaussian	543.3603342	gaussian	22.89
BCRH_4_24T_T3uw_209694.mzML	T3uw	4 weeks	8825.008943	gaussian	17.32463835	gaussian	0.20
BCRH_4_10T_T3uw_201816.mzML	T3uw	48 hours	11754.64405	gaussian	12188.98971	gaussian	50.91
BCRH_4_11T_T3uw_201817.mzML	T3uw	48 hours	15242.18655	gaussian	15205.46084	gaussian	49.94
BCRH_4_9T_T3uw_201819.mzML	T3uw	48 hours	10997.32894	gaussian	7781.185681	gaussian	41.44
BCRH_4_13T_T3uw_209670.mzML	T3uw	72 hours	4432.859522	gaussian	5828.394897	gaussian	56.80
BCRH_4_14T_T3uw_216881.mzML	T3uw	72 hours	3226.611012	gaussian	866.5622376	gaussian	21.17
BCRH_4_15T_T3uw_216883.mzML	T3uw	72 hours	5572.716216	gaussian	3773.565226	gaussian	40.38
BCRH_4_17_T3_w_179285.mzML	T3w	1 week	284.6043026	gaussian	61.96585549	gaussian	17.88
BCRH_4_17_T3_w_DCM_179293.mzML	T3w	1 week	2.562959454	gaussian	7.739603509	gaussian	75.12
BCRH_4_5_T3_w_179279.mzML	T3w	24 hours	78.91034538	gaussian	34.48667809	gaussian	30.41
BCRH_4_5_T3_w_DCM_179292.mzML	T3w	24 hours	1.55859095	gaussian	1.32177461	gaussian	45.89
BCRH_4_6T_T3w_201812.mzML	T3w	24 hours	27986.4243	gaussian	4982.10439	gaussian	15.11
BCRH_4_7T_T3w_201813.mzML	T3w	24 hours	11337.91439	gaussian	4403.543059	gaussian	27.97
BCRH_4_2_T3w_184703.mzML	T3w	4 hours	258.1948496	gaussian	171.5549506	gaussian	39.92
BCRH_4_3T_T3w_200109.mzML	T3w	4 hours	8299.781941	gaussian	2660.989872	gaussian	24.28
BCRH_4_4T_T3w_201810.mzML	T3w	4 hours	11691.99127	gaussian	3579.962165	gaussian	23.44
BCRH_4_22T_T3w_183631.mzML	T3w	4 weeks	723.6603294	gaussian	0.554074462	gaussian	0.08
BCRH_4_23T_T3w_209679.mzML	T3w	4 weeks	2248.419204	gaussian	20.51960486	gaussian	0.90
BCRH_4_23T_T3w2_209680.mzML	T3w	4 weeks	3022.47386	gaussian	327.728653	gaussian	9.78
BCRH_4_24T_T3w_209693.mzML	T3w	4 weeks	2030.919756	gaussian	3.858801761	gaussian	0.19
BCRH_4_10T_T3w_201808.mzML	T3w	48 hours	9614.044062	gaussian	1671.97082	gaussian	14.81
BCRH_4_11T_T3w_201809.mzML	T3w	48 hours	9077.860202	gaussian	999.8901223	gaussian	9.92
BCRH_4_9T_T3w_201811.mzML	T3w	48 hours	8774.249373	gaussian	1557.108253	gaussian	15.07
BCRH_4_13T_T3w_209669.mzML	T3w	72 hours	3982.973057	gaussian	406.8143077	gaussian	9.27
BCRH_4_15T_T3w_216877.mzML	T3w	72 hours	5291.856525	gaussian	2781.466312	gaussian	34.45

E.2.b. L-Trp pH 7 AUC data

file_name	plant_part	timepoint	peak205_auc	method205	peak208_auc	method208	Percentage
BCRH_4_17_MS_179286.mzML	MS	1 week	417.4231872	gaussian	40.72885449	gaussian	8.89
BCRH_4_5_MS_179280.mzML	MS	24 hours	506.7051978	gaussian	19.82613358	gaussian	3.77
BCRH_4_6T_MS_201804.mzML	MS	24 hours	10462.63557	gaussian	16.12506931	gaussian	0.15
BCRH_4_7T_MS_201805.mzML	MS	24 hours	9614.529158	gaussian	84.85178747	gaussian	0.87
BCRH_4_2_MS_184696.mzML	MS	4 hours	454.7765532	gaussian	1.070449104	gaussian	0.23
BCRH_4_3T_MS_200106.mzML	MS	4 hours	15849.95706	gaussian	57.95397482	gaussian	0.36
BCRH_4_4T_MS_201802.mzML	MS	4 hours	6584.893558	gaussian	57.91804586	gaussian	0.87
BCRH_4_22T_MS_183629.mzML	MS	4 weeks	303.0345469	gaussian	0.590925935	gaussian	0.19
BCRH_4_23T_MS_209677.mzML	MS	4 weeks	1244.639565	gaussian	2.401890725	gaussian	0.19
BCRH_4_24T_MS_209691.mzML	MS	4 weeks	14319.82081	gaussian	27.93847557	gaussian	0.19
BCRH_4_10T_MS_201800.mzML	MS	48 hours	17514.55014	gaussian	63.2672979	gaussian	0.36
BCRH_4_11T_MS_201801.mzML	MS	48 hours	12865.86214	gaussian	14.51963641	gaussian	0.11
BCRH_4_9T_MS_201803.mzML	MS	48 hours	20453.02241	gaussian	60.99300147	gaussian	0.30
BCRH_4_13T_MS_209667.mzML	MS	72 hours	7723.181381	gaussian	6.250807452	gaussian	0.08
BCRH_4_13T_MS_216872.mzML	MS	72 hours	7020.985016	gaussian	8.739617981	gaussian	0.12
BCRH_4_15T_MS_216873.mzML	MS	72 hours	7590.253522	gaussian	31.74469039	gaussian	0.42
BCRH_4_22T_Pod_183625.mzML	Pod	4 weeks	1356.073293	gaussian	0.985326278	gaussian	0.07
BCRH_4_23T_pod_209685.mzML	Pod	4 weeks	24729.46408	gaussian	7.056115299	gaussian	0.03
BCRH_4_24T_pod_209697.mzML	Pod	4 weeks	34281.51283	gaussian	10.81888036	gaussian	0.03
BCRH_4_17_STEM_179287.mzML	STEM	1 week	263.4580547	gaussian	49.23206229	gaussian	15.74
BCRH_4_5_STEM_179281.mzML	STEM	24 hours	8.664777159	gaussian	2.92306877	gaussian	25.23
BCRH_4_6T_STEM_201796.mzML	STEM	24 hours	13547.15357	gaussian	61.22077053	gaussian	0.45
BCRH_4_7T_STEM_201797.mzML	STEM	24 hours	8952.863226	gaussian	55.86162072	gaussian	0.62
BCRH_4_2_STEM_184685.mzML	STEM	4 hours	299.2859639	gaussian	18.05684454	gaussian	5.69
BCRH_4_3T_STEM_200105.mzML	STEM	4 hours	10498.03359	gaussian	789.6488761	gaussian	7.00
BCRH_4_4T_STEM_201794.mzML	STEM	4 hours	7790.319064	gaussian	25.96760249	gaussian	0.33
BCRH_4_22T_STEM_183630.mzML	STEM	4 weeks	364.895932	gaussian	1.078539237	gaussian	0.29
BCRH_4_23T_STEM_209676.mzML	STEM	4 weeks	5369.125364	gaussian	0.878372541	gaussian	0.02
BCRH_4_24T_STEM_209690.mzML	STEM	4 weeks	12229.70826	gaussian	19.16940184	gaussian	0.16
BCRH_4_10T_STEM_201792.mzML	STEM	48 hours	12140.54255	gaussian	83.69270631	gaussian	0.68
BCRH_4_11T_STEM_201793.mzML	STEM	48 hours	12475.41205	gaussian	89.50058184	gaussian	0.71
BCRH_4_9T_STEM_201795.mzML	STEM	48 hours	10695.94213	gaussian	232.0636791	gaussian	2.12
BCRH_4_13T_STEM_209666.mzML	STEM	72 hours	3327.404159	gaussian	6.594205935	gaussian	0.20
BCRH_4_14T_STEM_216864.mzML	STEM	72 hours	4080.270981	gaussian	9.654125627	gaussian	0.24
BCRH_4_15T_STEM_216863.mzML	STEM	72 hours	1617.670847	gaussian	26.63642669	gaussian	1.62
BCRH_4_17_T2_179288.mzML	T2	1 week	154.2408127	gaussian	10.35110991	gaussian	6.29
BCRH_4_5_T2_179282.mzML	T2	24 hours	475.9201582	gaussian	12.3062936	gaussian	2.52
BCRH_4_6_T2_201780.mzML	T2	24 hours	11046.88994	gaussian	54.23051953	gaussian	0.49
BCRH_4_7_T2_201781.mzML	T2	24 hours	5534.274555	gaussian	123.2479176	gaussian	2.18
BCRH_4_3T_T2_200107.mzML	T2	4 hours	9709.644767	gaussian	24.91152519	gaussian	0.26
BCRH_4_4T_T2_201778.mzML	T2	4 hours	4481.274973	gaussian	1.525440827	gaussian	0.03
BCRH_4_22T_T2_183626.mzML	T2	4 weeks	650.5257939	gaussian	16.79393843	trapezoidal	2.52
BCRH_4_23T_T2_209674.mzML	T2	4 weeks	4341.089064	gaussian	13.6634569	gaussian	0.31
BCRH_4_24T_T2_209688.mzML	T2	4 weeks	7354.836868	gaussian	12.3350096	gaussian	0.17
BCRH_4_10T_T2_201776.mzML	T2	48 hours	5987.022358	gaussian	22.49312518	gaussian	0.37
BCRH_4_11T_T2_201777.mzML	T2	48 hours	6865.117716	gaussian	59.01211925	gaussian	0.85
BCRH_4_9T_T2_201779.mzML	T2	48 hours	7824.860815	gaussian	75.83477904	gaussian	0.96
BCRH_4_13T_T2_209664.mzML	T2	72 hours	10886.97327	gaussian	56.38147799	gaussian	0.52
BCRH_4_15T_T2_216860.mzML	T2	72 hours	5498.838408	gaussian	30.38977665	gaussian	0.55
BCRH_4_17_T3_spot_179283.mzML	T3spot	1 week	9.57014985	gaussian	1009.191817	gaussian	99.06
BCRH_4_5_T3_spot_179277.mzML	T3spot	24 hours	13.17564829	gaussian	459.7270248	gaussian	97.21
BCRH_4_6T_T3spot_201829.mzML	T3spot	24 hours	16.48860007	gaussian	5433.28969	gaussian	99.70
BCRH_4_7T_T3spot_201830.mzML	T3spot	24 hours	8.436447304	gaussian	8092.38056	gaussian	99.90
BCRH_4_2_T3spot_184717.mzML	T3spot	4 hours	3.79070012	gaussian	1455.636204	gaussian	99.74
BCRH_4_3T_T3spot_200115.mzML	T3spot	4 hours	586.6864724	gaussian	34188.69148	gaussian	98.31
BCRH_4_4T_T3spot_201827.mzML	T3spot	4 hours	3.084867256	gaussian	1018.289574	gaussian	99.70
BCRH_4_22T_T3spot_183633.mzML	T3spot	4 weeks	2.312527955	gaussian	0.591478012	gaussian	20.37
BCRH_4_23T_T3spot_209683.mzML	T3spot	4 weeks	191.0129512	gaussian	0.892310454	gaussian	0.46
BCRH_4_24T_T3spot_209695.mzML	T3spot	4 weeks	567.1000233	gaussian	0.865563394	gaussian	0.15

BCRH_4_10T_T3spot_201825.mzML	T3spot	48 hours	16.16486836	gaussian	6017.311168	gaussian	99.73
BCRH_4_11T_T3spot_201826.mzML	T3spot	48 hours	7.470001437	gaussian	5792.344091	gaussian	99.87
BCRH_4_9T_T3spot_201828.mzML	T3spot	48 hours	12.92222812	gaussian	4477.751623	gaussian	99.71
BCRH_4_13T_T3spot_209671.mzML	T3spot	72 hours	17.9668878	gaussian	9814.98471	gaussian	99.82
BCRH_4_14T_T3spot_216889.mzML	T3spot	72 hours	41.08435984	trapezoidal	2991.142411	gaussian	98.65
BCRH_4_15T_T3spot_216888.mzML	T3spot	72 hours	23.20328599	gaussian	2804.605862	gaussian	99.18
BCRH_4_17_T3_uw_179284.mzML	T3uw	1 week	79.29845896	gaussian	105.091701	gaussian	56.99
BCRH_4_5_T3_uw_179278.mzML	T3uw	24 hours	47.5901304	gaussian	48.13524337	gaussian	50.28
BCRH_4_6T_T3uw_201820.mzML	T3uw	24 hours	14284.48739	gaussian	10949.10973	gaussian	43.39
BCRH_4_7T_T3uw_201821.mzML	T3uw	24 hours	9502.842187	gaussian	8135.644992	gaussian	46.12
BCRH_4_2_T3uw_184710.mzML	T3uw	4 hours	335.6773089	gaussian	434.5823278	gaussian	56.42
BCRH_4_3T3uw_200111.mzML	T3uw	4 hours	8636.642778	gaussian	10233.92302	gaussian	54.23
BCRH_4_4T_T3uw_201818.mzML	T3uw	4 hours	6819.024992	gaussian	12783.07831	gaussian	65.21
BCRH_4_22T_T3uw_183632.mzML	T3uw	4 weeks	424.8929042	gaussian	0.899037283	gaussian	0.21
BCRH_4_23T_T3uw_209681.mzML	T3uw	4 weeks	3617.6953	gaussian	374.9743415	gaussian	9.39
BCRH_4_23T_T3uw2_209682.mzML	T3uw	4 weeks	1830.011607	gaussian	543.3603342	gaussian	22.89
BCRH_4_24T_T3uw_209694.mzML	T3uw	4 weeks	8825.008943	gaussian	17.32463835	gaussian	0.20
BCRH_4_10T_T3uw_201816.mzML	T3uw	48 hours	11754.64405	gaussian	12188.98971	gaussian	50.91
BCRH_4_11T_T3uw_201817.mzML	T3uw	48 hours	15242.18655	gaussian	15205.46084	gaussian	49.94
BCRH_4_9T_T3uw_201819.mzML	T3uw	48 hours	10997.32894	gaussian	7781.185681	gaussian	41.44
BCRH_4_13T_T3uw_209670.mzML	T3uw	72 hours	4432.859522	gaussian	5828.394897	gaussian	56.80
BCRH_4_14T_T3uw_216881.mzML	T3uw	72 hours	3226.611012	gaussian	866.5622376	gaussian	21.17
BCRH_4_15T_T3uw_216883.mzML	T3uw	72 hours	5572.716216	gaussian	3773.565226	gaussian	40.38
BCRH_4_17_T3_w_179285.mzML	T3w	1 week	284.6043026	gaussian	61.96585549	gaussian	17.88
BCRH_4_17_T3_w_DCM_179293.mzML	T3w	1 week	2.562959454	gaussian	7.739603509	gaussian	75.12
BCRH_4_5_T3_w_179279.mzML	T3w	24 hours	78.91034538	gaussian	34.48667809	gaussian	30.41
BCRH_4_5_T3_w_DCM_179292.mzML	T3w	24 hours	1.55859095	gaussian	1.32177461	gaussian	45.89
BCRH_4_6T_T3w_201812.mzML	T3w	24 hours	27986.4243	gaussian	4982.10439	gaussian	15.11
BCRH_4_7T_T3w_201813.mzML	T3w	24 hours	11337.91439	gaussian	4403.543059	gaussian	27.97
BCRH_4_2_T3w_184703.mzML	T3w	4 hours	258.1948496	gaussian	171.5549506	gaussian	39.92
BCRH_4_3T_T3w_200109.mzML	T3w	4 hours	8299.781941	gaussian	2660.989872	gaussian	24.28
BCRH_4_4T_T3w_201810.mzML	T3w	4 hours	11691.99127	gaussian	3579.962165	gaussian	23.44
BCRH_4_22T_T3w_183631.mzML	T3w	4 weeks	723.6603294	gaussian	0.554074462	gaussian	0.08
BCRH_4_23T_T3w_209679.mzML	T3w	4 weeks	2248.419204	gaussian	20.51960486	gaussian	0.90
BCRH_4_23T_T3w2_209680.mzML	T3w	4 weeks	3022.47386	gaussian	327.728653	gaussian	9.78
BCRH_4_24T_T3w_209693.mzML	T3w	4 weeks	2030.919756	gaussian	3.858801761	gaussian	0.19
BCRH_4_10T_T3w_201808.mzML	T3w	48 hours	9614.044062	gaussian	1671.97082	gaussian	14.81
BCRH_4_11T_T3w_201809.mzML	T3w	48 hours	9077.860202	gaussian	999.8901223	gaussian	9.92
BCRH_4_9T_T3w_201811.mzML	T3w	48 hours	8774.249373	gaussian	1557.108253	gaussian	15.07
BCRH_4_13T_T3w_209669.mzML	T3w	72 hours	3982.973057	gaussian	406.8143077	gaussian	9.27
BCRH_4_15T_T3w_216877.mzML	T3w	72 hours	5291.856525	gaussian	2781.466312	gaussian	34.45
BCRH_4_6T_T4_201788.mzML	T4	24 hours	20385.81801	gaussian	24.86763619	gaussian	0.12
BCRH_4_7T_T4_201789.mzML	T4	24 hours	13867.34097	gaussian	75.4971653	gaussian	0.54
BCRH_4_3T_T4_200108.mzML	T4	4 hours	15150.70956	gaussian	55.44783986	gaussian	0.36
BCRH_4_4T_T4_201786.mzML	T4	4 hours	8622.325222	gaussian	20.98244018	gaussian	0.24
BCRH_4_23T_T4_209675.mzML	T4	4 weeks	2985.804612	gaussian	12.17992746	gaussian	0.41
BCRH_4_24T_T4_209689.mzML	T4	4 weeks	12066.73286	gaussian	15.07175551	gaussian	0.12
BCRH_4_10T_T4_201784.mzML	T4	48 hours	15567.56062	gaussian	31.3316154	gaussian	0.20
BCRH_4_11T_T4_201785.mzML	T4	48 hours	7740.967992	gaussian	17.14950191	gaussian	0.22
BCRH_4_9T_T4_201787.mzML	T4	48 hours	15304.10097	gaussian	31.39772661	gaussian	0.20
BCRH_4_13T_T4_209665.mzML	T4	72 hours	14287.72161	gaussian	15.71075045	gaussian	0.11
BCRH_4_14T_T4_216857.mzML	T4	72 hours	6489.168823	gaussian	21.50059761	gaussian	0.33
BCRH_4_15T_T4_216854.mzML	T4	72 hours	11100.40144	gaussian	178.3492371	gaussian	1.58

E.2.c. L-Trp pH 1 AUC data

plant_part	timepoint	method_index2	auc_index2	method_index4	auc_index4	perc_index4
MS	4 hours	trapezoidal	528.055115	trapezoidal	47.1811136	8.202041404
MS	4 weeks	trapezoidal	2453.19912	trapezoidal	116.40404	4.530039572
MS	4 weeks	trapezoidal	1277.09498	trapezoidal	88.2733463	6.465167268
MS	4 weeks	trapezoidal	4770.32951	trapezoidal	122.930987	2.512251033
MS	4 weeks	trapezoidal	1321.02309	trapezoidal	728.605988	35.54818747
STEM	4 hours	trapezoidal	404.076477	trapezoidal	49.9601406	11.00354876
STEM	4 weeks	trapezoidal	1459.86146	trapezoidal	502.667488	25.61325213
T2	4 hours	trapezoidal	383.976007	trapezoidal	59.7593632	13.46734274
T2	4 weeks	trapezoidal	5813.53826	trapezoidal	150.645552	2.525836847
T2	4 weeks	trapezoidal	5062.26832	trapezoidal	118.236989	2.282344719
T2	4 weeks	trapezoidal	7979.02735	trapezoidal	167.684301	2.058306568
T2	4 weeks	trapezoidal	1430.64531	trapezoidal	498.090322	25.82470684
T3spot	4 hours	trapezoidal	129.027483	trapezoidal	2455.69162	95.00806556
T3spot	4 weeks	trapezoidal	87.7721949	trapezoidal	735.080415	89.33318144
T3spot	4 weeks	trapezoidal	81.2838756	trapezoidal	517.710984	86.42995441
T3spot	4 weeks	trapezoidal	78.7895857	trapezoidal	545.640495	87.38216045
T3spot	4 weeks	trapezoidal	163.629418	trapezoidal	872.186954	84.20285462
T3uw	4 hours	trapezoidal	449.477977	trapezoidal	1541.13719	77.42014703
T3uw	4 weeks	trapezoidal	3782.5833	trapezoidal	856.360276	18.4602434
T3uw	4 weeks	trapezoidal	5159.04427	trapezoidal	1144.06455	18.150798
T3uw	4 weeks	trapezoidal	6323.25481	trapezoidal	1223.49396	16.2121994
T3uw	4 weeks	trapezoidal	429.944196	trapezoidal	467.630187	52.09932415
T3w	4 hours	trapezoidal	173.716959	trapezoidal	99.2023959	36.34861148
T3w	4 weeks	trapezoidal	6415.8735	trapezoidal	946.2422	12.85285696
T3w	4 weeks	trapezoidal	9376.97082	trapezoidal	688.612649	6.841259134
T3w	4 weeks	trapezoidal	6892.97703	trapezoidal	1045.63556	13.17151512
T3w	4 weeks	trapezoidal	891.450742	trapezoidal	542.395813	37.82802358
T4	4 weeks	trapezoidal	6980.82824	trapezoidal	155.891609	2.18435937
T4	4 weeks	trapezoidal	6547.13172	trapezoidal	130.449323	1.953541594
T4	4 weeks	trapezoidal	6695.70368	trapezoidal	148.222029	2.165745729
T4	4 weeks	trapezoidal	976.460931	trapezoidal	503.51799	34.02197036

E.2.d. D-Trp pH 7 AUC data

file_name	plant_part	timepoint	peak205_auc	method205	peak208_auc	method208
BCRH_5_10D_MS_219874.mzML	MS	1 week	-0.508731113	trapezoidal	-1.048498886	trapezoidal
BCRH_5_11D_MS_219875.mzML	MS	1 week	24.37011408	gaussian	-2.122211957	trapezoidal
BCRH_5_1D_MS_219871.mzML	MS	4 hours	20.54922653	gaussian	0.795934745	gaussian
BCRH_5_2D_MS_219872.mzML	MS	4 hours	78.67933669	gaussian	0.718660364	gaussian
BCRH_5_3D_MS_219873.mzML	MS	4 hours	38.12826474	gaussian	-3.540141283	trapezoidal
BCRH_5_10D_Pod_219901.mzML	Pod	1 week	65.14028123	gaussian	0.709262601	gaussian
BCRH_5_11D_Pod_219902.mzML	Pod	1 week	3.786264114	gaussian	1.511024027	gaussian
BCRH_5_1D_stem_219865.mzML	STEM	4 hours	6.596294889	gaussian	0.106985718	trapezoidal
BCRH_5_2D_stem_219866.mzML	STEM	4 hours	15.05258275	gaussian	0.87095904	gaussian
BCRH_5_3D_stem_219867.mzML	STEM	4 hours	17.26613857	gaussian	2.079373675	gaussian
BCRH_5_10D_stem_219868.mzML	STEM	1 week	5.791513017	gaussian	3.854242435	gaussian
BCRH_5_11D_stem_219869.mzML	STEM	1 week	4.399948309	gaussian	1.015058317	gaussian
BCRH_5_1D_T2_219883.mzML	T2	4 hours	87.73419395	gaussian	8.650601138	trapezoidal
BCRH_5_2D_T2_219884.mzML	T2	4 hours	12.25369418	gaussian	3.474102505	gaussian
BCRH_5_3D_T2_219885.mzML	T2	4 hours	48.23918895	gaussian	1.414793492	gaussian
BCRH_5_10D_T2_219886.mzML	T2	1 week	43.72658241	gaussian	-1.211147706	trapezoidal
BCRH_5_11D_T2_219887.mzML	T2	1 week	26.73706433	gaussian	-9.644763598	trapezoidal
BCRH_5_1D_T3uw_219895.mzML	T3uw	4 hours	89.65153262	gaussian	120.5137473	gaussian
BCRH_5_2D_T3uw_219896.mzML	T3uw	4 hours	175.8660365	gaussian	423.7272229	gaussian
BCRH_5_3D_T3uw_219897.mzML	T3uw	4 hours	360.599163	gaussian	809.2592934	gaussian
BCRH_5_10D_T3uw_219898.mzML	T3uw	1 week	67.45891157	gaussian	37.82167672	gaussian
BCRH_5_11D_T3uw_219899.mzML	T3uw	1 week	238.0451976	gaussian	11.66696464	gaussian
BCRH_5_1D_T3w_219889.mzML	T3w	4 hours	208.5138686	gaussian	287.9248121	gaussian
BCRH_5_2D_T3w_219890.mzML	T3w	4 hours	194.1751102	gaussian	179.7567768	gaussian
BCRH_5_3D_T3w_219891.mzML	T3w	4 hours	130.6108612	gaussian	163.3934736	gaussian
BCRH_5_10D_T3w_219892.mzML	T3w	1 week	184.3532207	gaussian	104.8471662	gaussian
BCRH_5_11D_T3w_219893.mzML	T3w	1 week	77.55912936	gaussian	2.281860089	gaussian
BCRH_5_1D_T4_219877.mzML	T4	4 hours	54.11882589	gaussian	1.148350181	gaussian
BCRH_5_2D_T4_219878.mzML	T4	4 hours	110.108828	gaussian	1.811127032	gaussian
BCRH_5_3D_T4_219879.mzML	T4	4 hours	40.1444112	gaussian	3.153676579	trapezoidal
BCRH_5_10D_T4_219880.mzML	T4	1 week	68.06907513	gaussian	-1.104410489	trapezoidal
BCRH_5_11D_T4_219881.mzML	T4	1 week	70.47428509	gaussian	-1.01659366	trapezoidal

Appendix F: Statistics

F.1. t-Test: two-sample assuming unequal variances for dry mass at 4 hours

	Control	**L-Trp pH 7**
Mean	1.7419	2.2103
Variance	0.3952	0.0308
Observations	4	4
Hypothesized Mean Diff	0	
Degrees of Freedom (df)	3	
t Stat	-1.435	
P(T<=t) two-tail	0.2468	
t Critical two-tail	3.1824	

F.2. t-Test: two-sample assuming equal variances for dry mass at 24 hours

	Control	**L-Trp pH 7**
Mean	2.33	2.45
Variance	0.619	0.172
Observations	7	5
Pooled Variance	0.44	
Hypothesized Mean Diff	0	
Degrees of Freedom (df)	10	
t Stat	-0.306	
P(T<=t) two-tail	0.766	

t Critical two-tail	2.228	

F.3. t-Test: two-sample assuming equal variances for dry mass at 48 hours

	Control	**L-Trp pH 7**
Mean	2.86	3.35
Variance	0.0616	0.034
Observations	3	3
Pooled Variance	0.0478	
Hypothesized Mean Diff	0	
Degrees of Freedom (df)	4	
t Stat	-2.77	
P(T<=t) two-tail	0.051	
t Critical two-tail	2.776	

F.4. t-Test: two-sample assuming equal variances for dry mass at 72 hours

	Control	**L-Trp pH 7**
Mean	3.75	3.93
Variance	0.0301	0.3372
Observations	3	3
Hypothesized Mean Diff	0	
Degrees of Freedom (df)	2	

t Stat	-0.503	
P(T<=t) two-tail	0.665	
t Critical two-tail	4.303	

F.5. t-Test: two-sample assuming equal variances for dry mass at 1 week

	Control	**L-Trp pH 7**
Mean	5.4353	5.52
Variance	0.2979	1.8432
Observations	3	2
Hypothesized Mean Diff	0	
Degrees of Freedom (df)	1	
t Stat	-0.0838	
P(T<=t) two-tail	0.9468	
t Critical two-tail	12.7062	

F.6. ANOVA: Single Factor analysis for dry mass at 4 weeks

Group	**Count**	**Sum**	**Average**	**Variance**
Control 4 weeks	3	46.745	15.582	1.6046
L-Trp 4 weeks	3	35.73	11.91	3.3252
D-Trp 4 weeks	2	13.4	6.7	0.0722
L-Trp pH 1 4 weeks	3	21.07	7.023	0.2529
5-F Trp	3	38.67	12.89	0.7869

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	160.411	4	40.103	30.048	3.22×10^{-5}	3.6331
Within Groups	12.011	9	1.335			
Total	**172.422**	13				

F.7. ANOVA: Single factor analysis for dry mass at 4 hour

Group	Count	Sum	Average (g)	Variance
Control 4h	4	6.9678	1.7419	0.3952
L-Trp 4h	4	8.841	2.2103	0.0308
D-Trp pH 7 4h	3	5.78	1.9267	0.0901
5-MeO Trp	3	5.99	1.9967	0.0816
5-F Trp	3	6.36	2.12	0.0048

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	0.5023	4	0.1256	0.9239	0.4818	3.2592
Within Groups	1.6311	12	0.1359			
Total	**2.1335**	16				

F.8. t-Test: two-sample assuming unequal variances for pod number at 1 week

	Control	L-Trp
Mean	2.5	4
Variance	0.5	2
Observations	2	2
Hypothesized Mean Diff	0	

Degrees of Freedom (df)	1	
t Stat	-1.342	
P(T≤t) one-tail	0.204	
t Critical one-tail	6.314	
P(T≤t) two-tail	0.408	
t Critical two-tail	12.706	

F.9. t-Test: two-sample assuming unequal variances for pod number at 4 weeks

	Control	**L-Trp**
Mean	9	12
Variance	57	13
Observations	3	3
Hypothesized Mean Diff	0	
Degrees of Freedom (df)	3	
t Stat	-0.621	
P(T≤t) one-tail	0.289	
t Critical one-tail	2.353	
P(T≤t) two-tail	0.579	
t Critical two-tail	3.182	

Anomalous data point removed:

	Variable 1	**Variable 2**
Mean	13	12
Variance	18	13
Observations	2	3
Pooled Variance	14.667	
Hypothesized Mean Diff	0	
Degrees of Freedom (df)	3	
t Stat	0.286	
P(T≤t) one-tail	0.397	
t Critical one-tail	2.353	
P(T≤t) two-tail	0.793	
t Critical two-tail	3.182	