Detecting Trace Pollutants used on Kiwifruit Orchards.

A thesis presented in partial fulfilment of the requirements of the degree of

Master of Philosophy

in

Chemistry

at Massey University, Manawatu, New Zealand

Jacob Scott

2023

I dedicate this thesis to my loving Father in Heaven. Your blessings have been my strength, and your Word has been a lamp unto my feet.

Abstract

Trace pollutants from sprays used on kiwifruit orchards are difficult to detect, often requiring sophisticated machinery and sample pre-treatment to detect them at very low levels. There are new techniques being created, but a significant portion of these are expensive, complicated, or are sensitive to interferents. Most techniques use some form of chromatography, but this requires expensive equipment and is slow, however, they can successfully detect the pollutants at very low levels. The use of supramolecular cages to bind and detect these pollutants could remove the need for expensive equipment and speed up the detection process. Supramolecular cages can be tailored to specific pollutants, minimizing the potential for interferences. They can also have a variety of detection mechanisms by adding a detection unit to the cage, allowing for simple testing and the use of cheaper equipment. This project will explore the use of supramolecular cages to bind and detect typical pollutants used in the New Zealand kiwifruit industry.

Contributions

All the work in this thesis was completed by Jacob Scott

Except:

- X-ray crystallography data was collected by Tyson Dias

Acknowledgements

I would like to take the opportunity to thank the large number of people who helped and supported me throughout my MSc journey. Firstly, I would like to thank my supervisors, Paul Plieger, Gareth Rowlands and Mark Waterland, for providing me the opportunity to do this project. I have greatly appreciated all of your support, guidance, patience, and knowledge. Thank you for always answering any questions I had and for being fantastic supervisors.

I would like to thank Octavio Perez-Garcia, Callaghan Innovation and Zespri International for the financial support that made this project possible.

I would like to thank Tyson Dias for running any crystals I grew on the X-ray, and David Lun for all the help with the mass spectrometer. I would also like to thank Sidney Woodhouse, Marryllyn Donaldson, Brodie Matheson and all other past Plieger Group members for their technical advice and scientific discussions.

I would like to thank my family and friends for their continual support and encouragement throughout this project. I would like to especially thank my amazing wife, Victoria-Jayne, for her love and support. Thank you for standing by me, for your sacrifices and motivating me through the grinds.

Lastly, I would like to thank my heavenly Father; for bringing me to where I am, for giving me strength when I was weak, and for never leaving my side.

Abbreviations

AMPA – Aminomethylphosphonic Acid

ATP – Adenosine Triphosphate

Boc2O – Di-tert-butyl Dicarbonate Anhydride

Cbz – Benzyl Chloroformate

CE – Capillary Electrophoresis

ELISA – Enzyme Linked Immune-Sorbent Assays

EPSPS – 5-Enolpyruvylshikimate-3-Phosphate Synthase

FRET – Fluorescence Resonance Energy Transfer

GC – Gas Chromatography

HPLC – High Pressure Liquid Chromatography

IC – Ion Chromatography

LC – Liquid Chromatography

MOF – Metal Organic Framework

NADH – Nicotinamide adenine nucleotide

NTP – 4-Nitrothiophenol

PET – Photo-induced Electron Transfer

QD – Quantum Dot

SERS – Surface Enhanced Raman Spectroscopy

SLIPS – Slippery Liquid-Infused Porous Surface

TEA – Triethyl Amine

TLC – Thin Layer Chromatography

TMB – Tetramethylbenzidine

Table of Contents

Abbreviations viii

Chapter 1 - Introduction 1

1.1 The New Zealand Kiwifruit Industry 1

1.2 Fundamentals of Cage Encapsulation 2

1.3 Detection Techniques 8

1.4 Research Overview 14

Chapter 2 – Python Code 17

2.1 Overview 17

2.2 Flowcharts 21

2.3 Breakdown of the Main Functions 24

2.4 Conclusion 26

Chapter 3 – Nitrate and Phosphate 28

3.1 A Brief Background of Fertilizers 28

3.2 Reasons to Detect Nitrate and Phosphate 29

3.3 Current Methods for Detecting Nitrate and Phosphate 30

3.4 Target Cages and Design Features 31

3.5 Synthesis and Analysis 35

3.6 Conclusion 53

Chapter 4 – Glyphosate 55

4.1 History of Glyphosate 55

4.2 Reasons to Detect Glyphosate 59

4.3 Current Methods for Detecting Glyphosate 60

4.4 Target Cages and Design Features 63

4.5 Synthesis and Analysis 67

4.6 Conclusion 78

Chapter 5 – Hydrogen Cyanamide 80

5.1 History of Hydrogen Cyanamide 80

5.2 Reasons to Detect Hydrogen Cyanamide 83

5.3 Current Methods for Detecting Hydrogen Cyanamide 83

5.4 Target Cages and Design Features 83

5.5 Synthesis and Analysis 86

5.6 Conclusion 93

Chapter 6 – References 95

Appendix 1 – Copper Hydroxide 109

A1.1 Copper Hydroxide Background 109

A1.2 Initial Scope 109

A1.3 Completed Work 110

Appendix 2 – Experimental Details 113

A2.1 Chapter 3 experimental details 113
A2.2 Chapter 4 experimental details 119
A2.3 Chapter 5 experimental details 123
A2.4 Appendix 1 experimental details 125
A2.5 X-ray collection data 126
Appendix 3 – Experimental Spectra 132
Appendix 4 – Raw Python Code 162

Chapter 1 - Introduction

1.1 The New Zealand Kiwifruit Industry

Kiwifruit was first brought to New Zealand in 1904 by Isabel Fraser and in 1928 the green cultivar known as “Hayward” was bred, which remains as one of the most widely grown green cultivars worldwide. The first kiwifruit exported from New Zealand, went to England in 1959 and since then the industry has grown massively. The foundation of the modern kiwifruit industry was built as a result of problems faced in the 1980s. The main problem was caused by multiple factors including the rise in the New Zealand Dollar, inflation, over-production of kiwifruit and multiple exporters exporting to the same markets. This led to the New Zealand Kiwifruit Marketing Board being created to end the multi-exporter regime and replace it with a single desk marketing structure.[1]

The New Zealand Kiwifruit Marketing Board fell on hard times in the early 1990s partly due to bad management and governance. From this, the New Zealand Kiwifruit Growers Incorporated became operational to give kiwifruit growers their own organization to develop a secure and stable kiwifruit industry and Zespri was created as a separate marketing and sales organization. This allowed New Zealand kiwifruit growers to invest in branding, marketing, quality control, and new research and development, which in turn has led to the industry growing into a global business.

A major focus for Zespri is sustainability and the impact that the kiwifruit industry has on the environment. This can be broken down into three major categories: packaging, climate change and water quality. To improve their packaging, Zespri is planning to make their packaging 100% recyclable, reusable or compostable. Zespri is also planning to ensure that at least 30% of the plastic they use is from recycled plastic and they are trying to reduce the carbon footprint of the packaging.[1] Thus far, Zespri has managed to make 88% of their packaging recyclable, reusable or compostable and 26% of the plastic used is from recycled plastics as described in their 2023 yearly report.[2]

With regards to climate change, Zespri’s research mainly focusses on mitigating and managing the impacts of climate change. From assessing the carbon footprint of exporting kiwifruit, to understanding the impact of changes in weather on fruit yield, the water requirements of kiwifruit vines and investigating soil carbon storage.[1]

The research into the water quality can be broken down further to protecting the waterways, enhancing the water quality, how the kiwifruit industry uses water and how Zespri can measure the progress and improvements. Zespri’s focus is on efficient water use, improving soil health and on nutrient management. Part of the nutrient management focus is around monitoring the nutrients and sprays used on the kiwifruit orchards. This will help build guidelines as to how much nutrients or sprays are required and how much will be washed away into the waterways.[1] There are multiple techniques that detect the nutrients or sprays used on orchards, however Zespri favours those that display a greater level of precision to detect and measure low levels of chemicals. Detecting trace chemicals assists in diagnosing areas for improvement in soil management, spray regimes and water management. This will result in reducing the amount that contaminates the waterways. It will also help with food safety by detecting the residues on the fruit. It has been identified that discrete molecular cages could be a tool to encapsulate chemicals of interest and accurately detect or quantify the amounts present and this project aims to further the research in this area.

1.2 Fundamentals of Cage Encapsulation

Discrete molecular cages are small molecules, only a few nanometres in size, that possess a cavity capable of molecular encapsulation. With appropriate modifications, either synthetically or through external perturbations, these cages can be tailored to a specific objective. The two main types of discrete molecular cages are non-metallic cages and metallic cages. Non-metallic cages are cages made from organic molecules, either covalently bonded together or intramolecularly bound with non-covalent interactions. Although non-metallic cages are interesting, they are not the focus of this study and will not be mentioned again. For more information on non-metallic cages, Little *et al*.[3] have published a general overview of non-metallic cages.

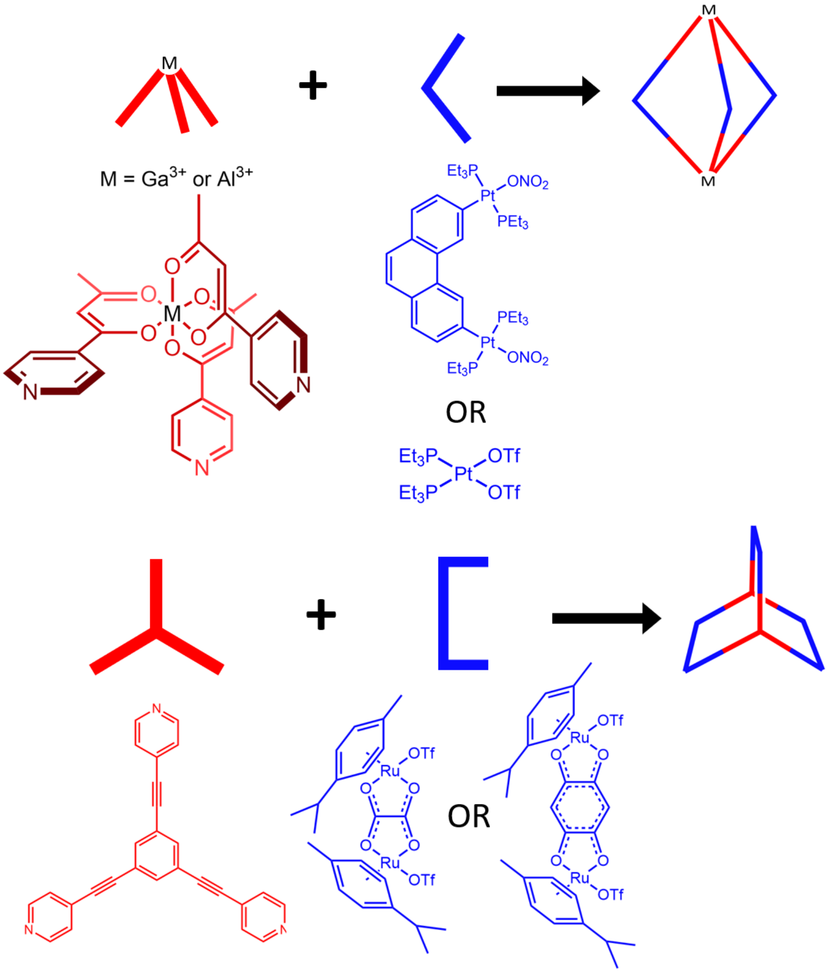
Metallic cages are made from metal centres and organic ligands bound together with dative bonds. Metallic cages self-assemble from multiple components and this allows simple variation of the structure.[4] The variations could be in the metals or the ligand, allowing a plethora of applications become viable, including being used for drug delivery,[5,6] catalysis[7,8] and sensors.[9-12] These applications utilise the cavity inside the cage in various ways that have been programmed into the components of the cage before the cage is assembled. This often entails an analyte or molecule being encapsulated into the cavity using a variety of different non-covalent interactions including electrostatic, π-π, ionic, hydrogen bonding and dipole-dipole interactions.[13] Also, because metallic cages are well understood, the behaviour of the cages can easily be predicted. The converging synthesis of metallic cages enable variations to be easily introduced.

For drug delivery, the cages used are generally rigid and designed to be resistant to harsh physiological conditions that may be encountered in the human body. These cages have the potential to encapsulate small drug molecules and be functionalized to add an antibody or some targeting functionality before or after encapsulation. This allows the drug molecule to be transported to the desired area, whilst being protected from the intervening environment.[5,6]

In catalytic cages, the cavity provides a microenvironment where the relative concentration of the reactants is greatly increased and the cage can potentially reduce the activation energy of the reaction by stabilising the transition state. The ligands being used could also have catalytic components incorporated that could be utilised when the cage has encapsulated the molecule.[7,8]

Using a cage as a sensor device, requires the cage to exhibit a detectable change in an inherent property (chemical or physical) upon binding a specific molecule.[9-12] For sensing purposes, there are multiple techniques that can be utilised and this project aimed to use both spectroscopic and electrochemical means of detection (UV/visible spectroscopy, fluorescent spectroscopy, Raman spectroscopy and cyclic voltammetry).

Some cages used for sensors are static assemblies as reported by Wang *et al.*,[9] where they described a trigonal-bipyramidal cage and a trigonal-prismatic cage, as seen in Scheme 1.1. These cages have electron rich fluorescent ligands that make the cage selectively encapsulate electron deficient nitroaromatic compounds, such as 4-nitrophenol, 2,4,6-trinitrotoluene and picric acid. These nitroaromatic compounds quench the fluorescence of the cages when encapsulated making these cages turn-off fluorescent sensors for nitroaromatic compounds.



Scheme 1.1: The trigonal-bipyramidal cage (top) and trigonal-prismatic cage (bottom) described by Wang *et al.*[9]

Another example was reported by Liu *et al.*,[10] and used static tetrahedral europium(III) cages, as seen in Figure 1.1, as sensors for copper(II) and iodide ions. Liu *et al.* found that the cage could act as a fluorescent turn-on sensor for copper(II) or a fluorescent turn-off sensor for iodide ions. The complex was both a turn-on and turn-off sensor due to the electron deficient triazine ring at the centre of the ligand that acts as a sensitising unit. When iodide ions interact with the triazine core of the ligands, they quench the fluorescence by adding electron density to the middle of the ligands. This allows the excitation energy to thermally dissipate rather than be emitted as fluorescence. Copper(II) ions have the opposite effect, where they remove electron density and therefore make the fluorescence more pronounced.

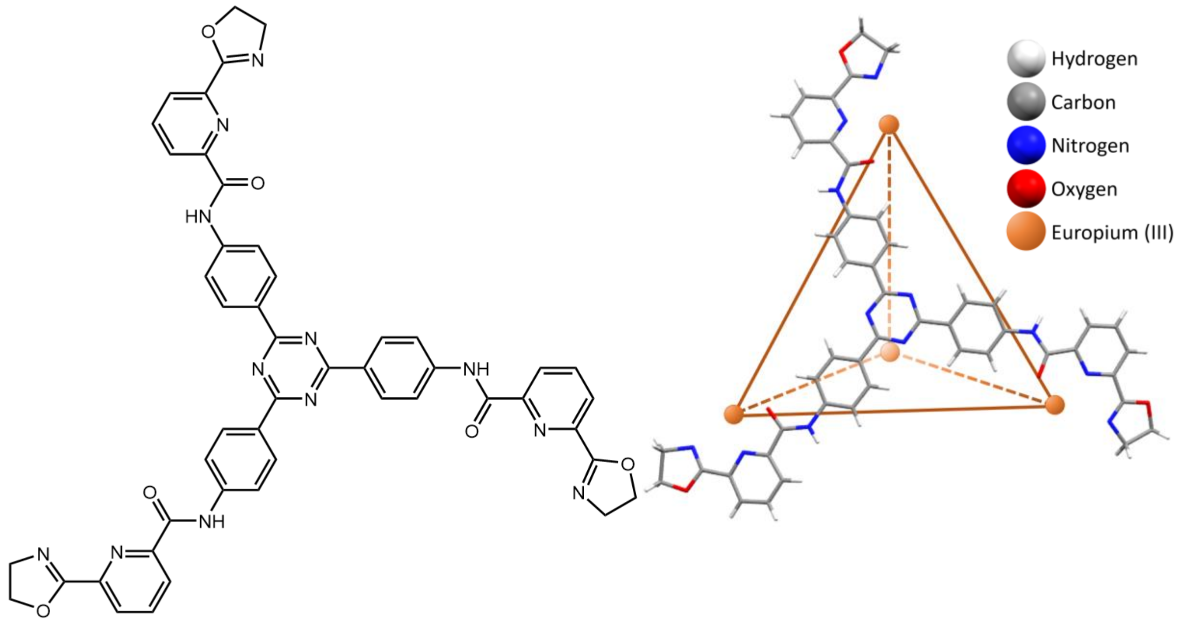


Figure 1.1: The ligand (left) and the europium(III) cage described by Liu *et al.*[10]

Other cages used for sensors are dynamic assemblies as demonstrated by the studies of Wenzel *et al.*[11] Here, a bimetallic helicate encapsulated and sensed sulfate ions. Wenzel *et al.* designed the ligand shown in Figure 1.2 to have a helical twist, protonatable amine groups that position the proton in the cavity of the cage and two square planar copper(II) binding sites. The protonated amine groups and the free coordination sites of the bound copper centres target anions while the helical twist selected for the size of the anions. The cage also had hexamethylene linkages between the amines to give the cage flexibility, leading to a large structural change depending on the state of the cage. One example of this variability is the copper-copper distance, which can change from 7.3 Å when tetrafluoroborate is coordinated to 6.4 Å when sulfate is coordinated. This led to the cage having a clear preference for sulfate ions over phosphate, perchlorate, tetrafluoroborate and nitrate ions.

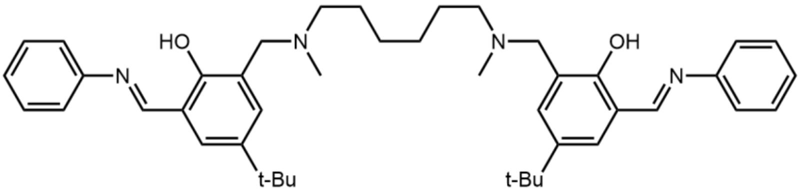


Figure 1.2: The ligand reported by Wenzel *et al.*[11]

The same research group reported a second dynamic cage that differed by the substitution of the imine with an oxime and the inclusion of an aromatic linker between the two phenols, as seen in Figure 1.3. These changes reduced the conformational freedom of the cage, resulting in it being more selective for coordinating anions rather than non-coordinating anions. This is observed in its preference for sulfate ions over perchlorate, nitrate and bromide ions compared to the ligand shown in Figure 1.2.[11,12]

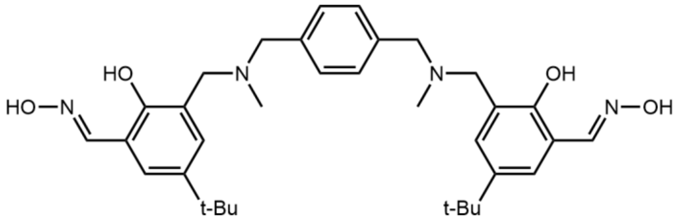


Figure 1.3: The ligand reported by Stevens *et al.*[12]

De Silva *et al.*[14] studied how the rigidity of the metal binding site influenced the anion binding capabilities. They found that switching between a phenolic oxime (top of Figure 1.4) and a phenolic imine (bottom of Figure 1.4) reduced the planarity of the metal binding site and this led to a larger binding cavity. The cage with the oxime ligand could encapsulate smaller anions, like bromide, but not larger anions, like tetrafluoroborate. On the other hand, the cage with the imine ligand could encapsulate both bromide and tetrafluoroborate.

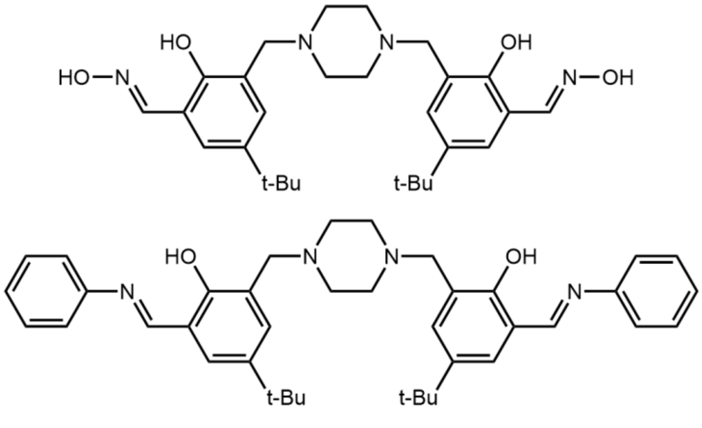


Figure 1.4: The phenolic oxime (top) and phenolic imine (bottom) ligands reported by De Silva *et al.*[14]

The focus of this project is on dynamic cages being used as sensors to detect trace pollutants used on kiwifruit orchards. Such pollutants come from fertilizers, herbicides, and pesticides. The pollutants are detrimental to human health and can leach into waterways affecting a wider area of plant life causing further exposure downstream.[15-18] Having the ability to detect very low levels of pollutants from environmental samples in real-time would be beneficial. It would help to identify the pollutants leaching into the waterways, which will aid in understanding and mitigate the process of pollutant leaching.

There are many different pollutants that could be present in the waterways and some examples are seen in Table 1.1. The table shows chemicals grouped in their main uses, the probability of a waterway containing the chemicals and the average concentration in the samples containing the molecules. The pollutants that are the focus of this project are glyphosate, hydrogen cyanamide, nitrate and phosphate. These molecules will be discussed in detail in subsequent chapters.

Table 1.1: Examples of molecules likely to be found in waterways.[19-23]

Category	Chemical Name	Concentration (μg L^{-1})	Probability (%)
Analgesic	Acetaminophen	40	23.8
	Codeine	0.2	10.6
Antacid	Cimetidine	0.074	9.5
	Ranitidine	0.01	1.2
Antibacterial	Sulfamethazine	0.2	4.8
Antibacterial	Trimethoprim	0.075	27.4
Antibiotic	Chlortetracycline	0.42	2.4
	Ciprofloxacin	0.1	2.6
	Erythromycin	0.95	21.5
	Lincomycin	0.06	19.2
	Norfloxacin	0.12	0.9
	Roxithromycin	0.05	4.8
	Sulfamethoxazole	0.18	19
	Tetracycline	0.33	1.2
	Trimethoprim	0.013	27.4
	Tylosin	0.02	13.5
Antidepressant	Fluoxetine	0.012	1.2
Antihypertensive	Diltiazem	0.021	13.1
Anti-Inflammatory	Ibuprofen	1100	9.5

Psychotropic	Caffeine	150	70.6
Steroid	Cholesterol	1	75.3
	Coprostanol	0.7	35.3
Metals	As	1.1	49.1
	Ba	22	70.1
	Cd	0.34	51.5
	Cr	3.9	68.1
	Cu	13	91.5
	Fe	650	91.6
	Ni	4.4	69.4
	Pb	7.3	73.8
	Zn	76	96.2
Herbicide	Glyphosate	0.20	36.0
Plant Growth Regulator	Hydrogen Cyanamide	0.03	Area Dependent
Fertilizer	Nitrate	2080	≈ 100
	Phosphate	350	≈ 100

Some molecules, like glyphosate, are difficult to detect without laborious sample pre-treatment, expensive equipment or the detection methods are prone to interferences.[24-27] This is where the dynamic cages could prosper. If they are tailored to a specific molecule, then there should not be any need for pre-treatment of the sample. Also, by selecting an appropriate detection technique, the cage has potential to detect very low levels of the analyte in real-time with minimal interferences.

1.3 Detection Techniques

Due to the synthesis of the cages having multiple components, it is easy to tune the cage for a specific detection technique. This can be done by designing a portion of the ligands to have a detectable group or by attaching the detectable group to the terminal of the cage. This project intended to use UV/visible spectroscopy, fluorescent spectroscopy, Raman spectroscopy and cyclic voltammetry to monitor chemical/physical changes of the cages.

Using these techniques to determine how many of the cages have encapsulated the target molecules will give the concentration of the target molecules in the solution.

UV/visible spectroscopy investigates how much electromagnetic radiation the molecule absorbs. Molecules absorb light when the energy of the photon matches the energy gap between occupied and unoccupied molecular orbitals, with the main absorption peak corresponding to the energy gap between the highest occupied molecular orbital (HOMO) to the lowest unoccupied molecular orbital (LUMO). When a molecule absorbs a photon, an electron is promoted from a basal molecular orbital/energy level to an excited molecular orbital/energy level. For the electron to return to the relaxed state, it needs to release the energy, generally by heat or a photon, as seen in Figure 1.5. Many factors impact the energy gaps between molecular orbitals, including the conjugation of the molecule, what atoms make up the molecule, how the atoms are arranged, the solvent and intramolecular interactions.[28]

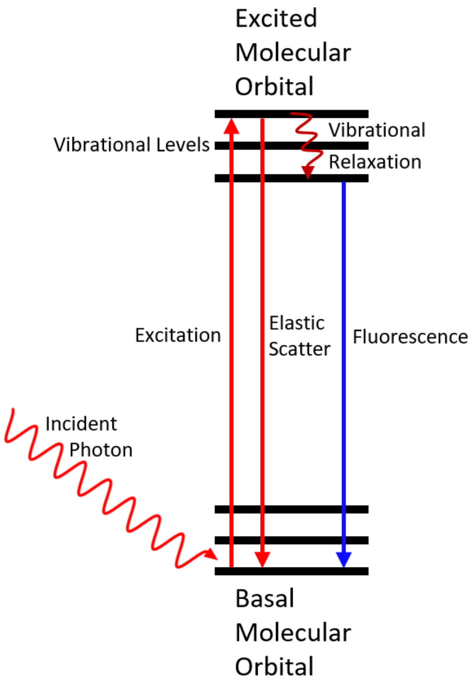


Figure 1.5: Electronic excitation diagram.

The UV/visible spectrophotometer works by irradiating the sample with a monochromatic light, that is scanned through a defined wavelength range. What radiation that is not absorbed by the molecule is transmitted to the detector, which analyses the number of photons that the sample has absorbed compared to a reference sample, as shown in Figure 1.6. The data collected from a spectrophotometer can be displayed as a graph showing the absorption vs wavelength to get a visual of what wavelengths are being absorbed.[28]

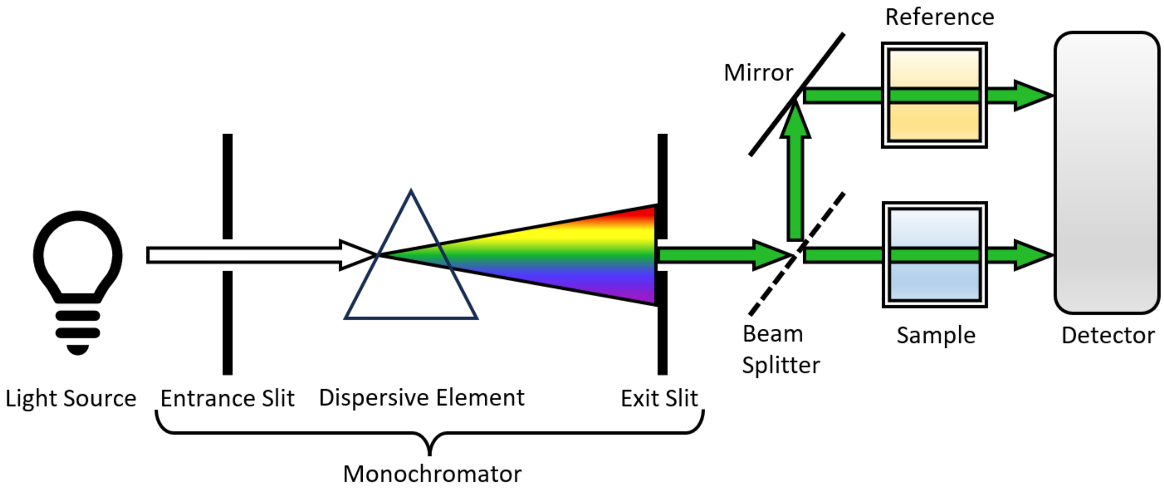


Figure 1.6: Diagram of how a UV/visible spectrophotometer functions.

When using UV/visible spectroscopy for detecting if a cage has encapsulated a target molecule, the wavelength of the absorption maximum would be monitored for variations. If the wavelength of the absorption maximum shifts, the electronic properties of the cage are changed. This indicates that the cage has encapsulated a molecule. For quantifying the level of bound molecules and therefore the concentration in solution, a calibration curve could be used. The main limitations with using UV/visible spectroscopy are the sensitivity of the instrument and the number of molecules that can interfere, which is any compound that absorbs light in the area of interest.[11]

Fluorescence spectroscopy is like UV/visible spectroscopy but rather than measuring the amount of light absorbed as a function of wavelength, it is a measure of the light emitted from the molecule. Fluorescence occurs when a photon of energy excites an electron, followed by the emission of a photon as the electron relaxes. The energy of the emitted photon will correspond to the energy gaps between molecular orbitals. The fluorescence spectrophotometer functions very similarly to a UV/visible spectrophotometer, with one key difference, which is where the detector is positioned. In a UV/visible spectrophotometer the detector is in line with the light source, so if no sample is inserted, the light from the light source would shine directly into the detector. In a fluorescence spectrophotometer, the detector is perpendicular to the light sources' line, so the light from the light source will not go to the detector, as seen in Figure 1.7.[29,30]

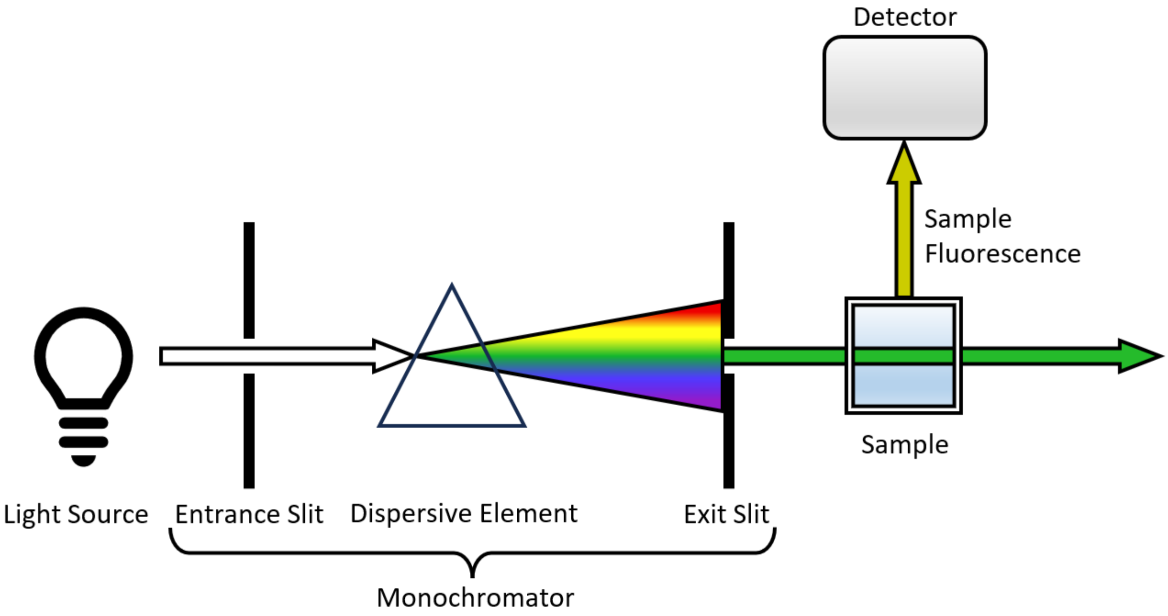


Figure 1.7: Diagram of how a fluorescence spectrophotometer functions.

Fluorescence spectroscopy is more accurate than UV/visible spectroscopy as it measures light against a dark background (without a reference) compared to UV/visible spectroscopy with that measures the difference between the incident light (reference sample) and the light after passing through the sample. Also, there are less molecules that can fluoresce compared to how many molecules absorb UV/visible light, reducing the potential for interferants.[29,30]

Raman spectroscopy detects Raman scattering and is slightly different to UV/visible spectroscopy and fluorescence spectroscopy in that it looks at the difference between the excitation wavelength and the emitted wavelength. Raman spectroscopy works by exciting an electron from an initial vibrational state to a virtual state, simultaneously transitioning back to the final vibrational state, emitting a photon. Rayleigh scattering is not a form of Raman scattering and occurs when the initial and final states are the same vibrational state. Raman scattering occurs if the initial state is a different level to the final state. When the initial state is lower than the final state, the difference is known as the Stokes shift and when the final state is lower than the initial state, the difference is known as the Anti-Stokes shift, as seen in the top of Figure 1.8. A Raman spectrophotometer is set up similar to a fluorescence spectrophotometer, but the light going to the detector, goes back towards the light source before being reflected away to the detector, as seen in the bottom of Figure 1.8.[31]

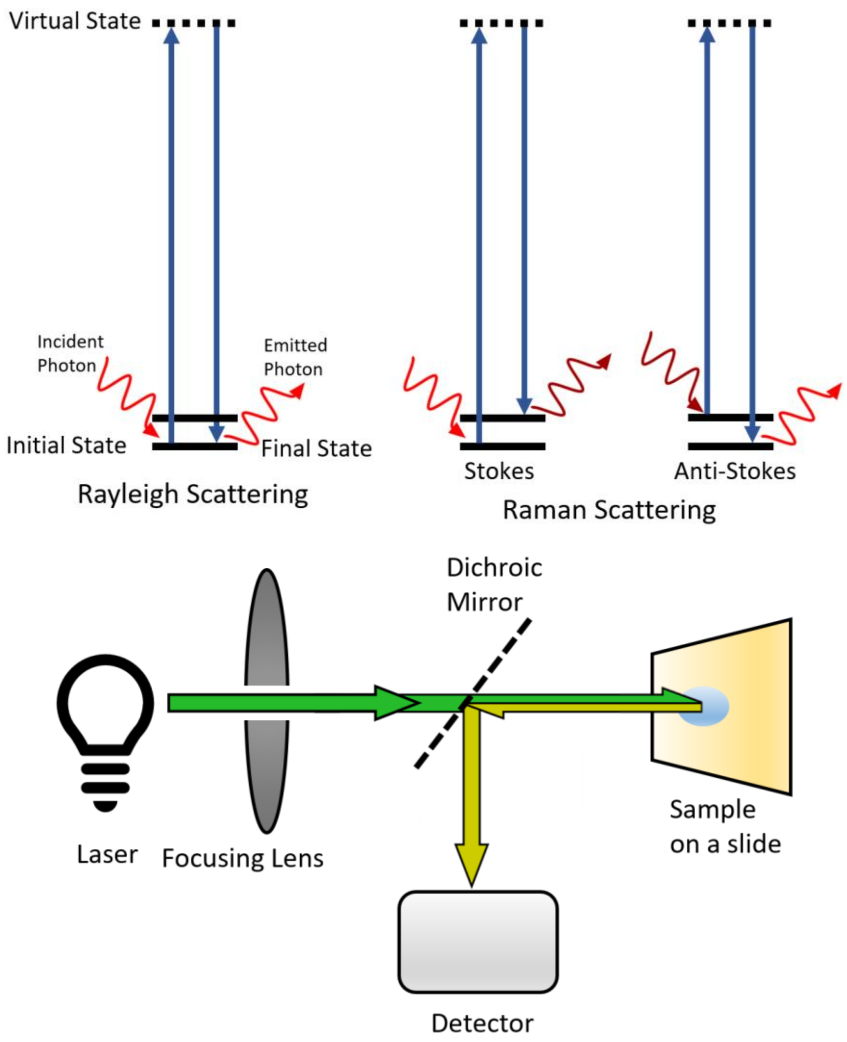


Figure 1.8: Rayleigh and Raman scattering (top) and a diagram of how a Raman spectrophotometer functions (bottom).

An advantage of using Raman spectroscopy is how it produces a fingerprint region, like that in infrared (IR) spectroscopy, allowing a more specific analysis of what is present in the sample. Another advantage is the ability to enhance Raman signals using the Surface Enhanced Raman Spectroscopy (SERS) effect, which utilises surface plasmons (collective electron oscillations) to enhance Raman scattering. Generally, metallic nanoparticles are used as a source of surface plasmons to generate large electric fields at the surface of the nanoparticles, basically by focusing the incident radiation field at the nanoparticle. The enhanced electric field will generate stronger polarisation, leading to more intense Raman scattering. The enhanced electric field only extends a few nanometers from the nanoparticles surface. At the junction between two nanoparticles ("hot spot"), the effect of these enhanced electric fields interacting can enhance the intensity of the Raman signals up to 5 or 6 orders of magnitude. By creating a way to have the analytes focused in the "hot spots", Raman spectroscopy has the potential to detect down to attomolar concentrations.[31,32]

There are three big problems with using Raman spectroscopy for detection of trace pollutants. The first is that strongly fluorescent molecules can mask any Raman scattering. The second is the effect concentration has on the signal strength and the third problem is the low number of "hot spots" in solution. To remove the negative effects of solution-based Raman, a slippery liquid-infused porous surface (SLIPS) can be used to evaporate the solvent from a drop of sample into a single spot rather than a larger circle on a normal surface.[33]

Cyclic voltammetry differs from the previous techniques as it is an electrochemical approach rather than a spectroscopic technique. Cyclic voltammetry provides a measurement of the redox reactions occurring near the surface of the electrode. Cyclic voltammetry requires three electrodes, a working electrode, a reference electrode and a counter electrode (middle of Figure 1.9). Cyclic voltammetry works by applying a cathodic or anodic potential to the working electrode that slowly changes as a function of time (left of Figure 1.9). The system will then record the current flowing through the working electrode and plot this as a function of the applied potential (right of Figure 1.9). The potential initially applied is generally one that will not initially oxidise or reduce the molecules under investigation. The potential is then swept linearly to a second potential and then back to the initial potential, assuming the molecule of interest undergoes a reversible redox reaction. The second potential is generally selected so the interval between the two potentials contains the oxidation or reduction of interest.[34-36] Cyclic voltammetry can be used with cages if the signal of the cage is strong enough to visualise the changes that occur when the cage encapsulates the molecule of interest.

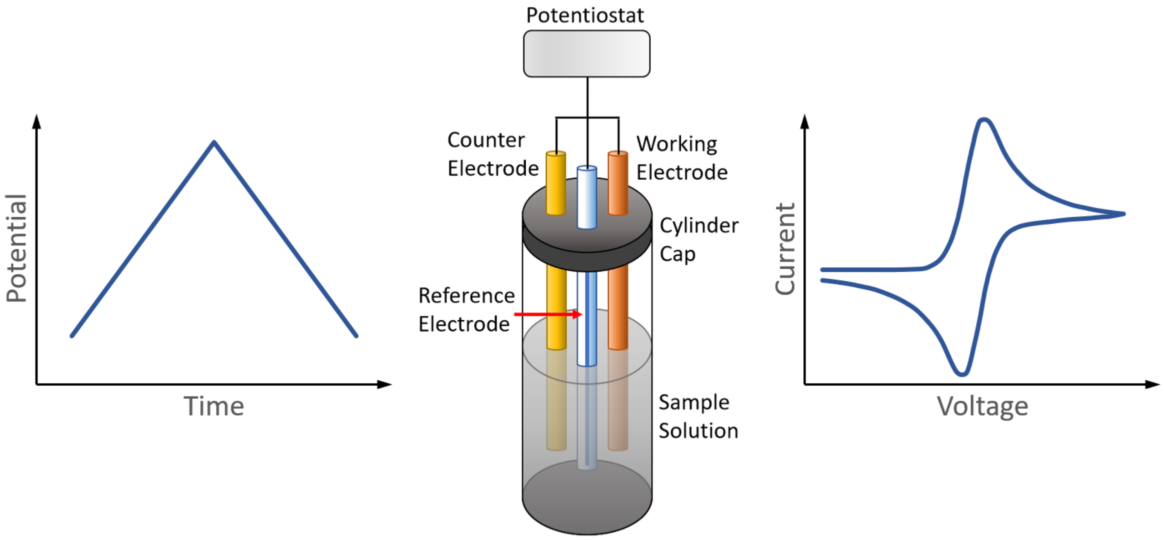


Figure 1.9: The sweeping potential applied to the working electrode (left), general cyclic voltammetry setup (middle) and an example of the measured current response (right).

Cyclic voltammetry can be very sensitive, detecting to nanomolar concentrations and very selective if the system it tuned correctly.[37] The main issue with cyclic voltammetry is the tuning of the system as there are many aspects that can be varied. These include what electrodes are used, the solvent, the electrolytes present in the system, the atmosphere in the reaction vessel, the reversibility of the redox reaction and any potential side reactions.[36]

1.4 Research Overview

The aim of this project was to develop a sensor using discrete molecular cages that can be used to detect trace pollutants on kiwifruit orchards. During this project, the main goals were to design and synthesise molecular cages that will be able to encapsulate glyphosate, hydrogen cyanamide, nitrate and phosphate. The detection capabilities of the cages would then be tested and the specificity of the individual cages using a variety of techniques. The final designs can be seen in Figure 1.10.

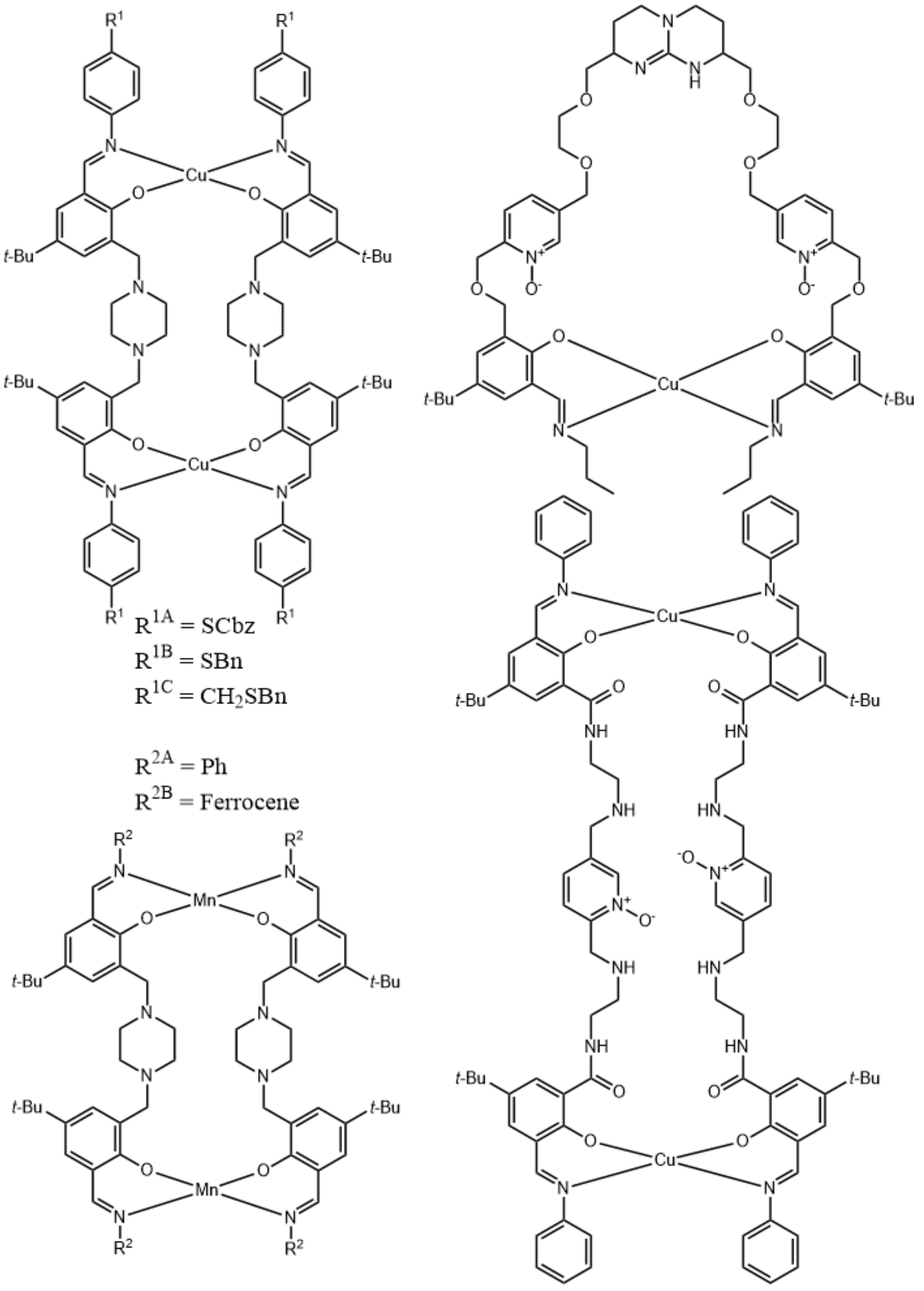


Figure 1.10: Final designs of the molecular cages targeting nitrate and phosphate (top left, R^{1A-1C}), glyphosate (right) and hydrogen cyanamide (bottom left, R^{2A-2B}).

Three cages were designed to target nitrate and phosphate that all shared the same core structure, but differed at the terminal groups, as seen in Figure 1.10. The synthesis of the cage using R^{1A} has been completed and analysed. The synthesis of the cage using R^{1B} has been completed but due to time constraints, has yet to be analysed, while the final cage, using R^{1C}, was not completed. The terminal groups of each of these cages contained sulfur, in order to attach the cage to nanoparticles, and allow Raman spectroscopy to be used as the detection technique.

Two cages have been designed to target glyphosate, one monometallic cage and one bimetallic cage, as seen in Figure 1.10. Both cages have not been completely synthesised due to time constraints. Due to the complexity of the synthesis, basic terminal groups have been used to ensure the cage could be synthesised successfully and that the cage encapsulates glyphosate effectively.

Two cages were also designed to target hydrogen cyanamide, both with the same core structure as the cages targeting nitrate and phosphate. The differences being manganese replacing the copper centers and different terminal groups as seen in Figure 1.10. The synthesis of the cage with R^{2A} has been completed, but it has not been analysed due to time constraints. The synthesis of the cage with R^{2B} has not been completed due to time constraints and was designed with ferrocene on the terminals to use cyclic voltammetry as the desired detection technique.

None of the final cage designs contain a fluorophore that can be detected in fluorescence spectroscopy. Given more time, it would have been investigated and some work was done towards this that can be found in Appendix 1. More details are discussed on each of the cages in Figure 1.10 and the next steps if this project was continued in Chapter 3, Chapter 4 and Chapter 5.

Chapter 2 – Python Code

2.1 Overview

Comparing and analysing multiple datasets at one time with varying layouts and document types is difficult on a large scale. Using Python, a powerful, yet readily accessible programming language, it is easy to achieve numerous analyses of multiple datasets. Raman spectroscopy is an example of a measurement technique that generates multiple datasets with many variables that can be altered. When running a sample on the Raman spectrometer, the main variables that can be adjusted includes the focal point (to focus the laser on the sample), the time per frame (how long the shutter stays open to collect data) and how many frames will be collected per scan. Each scan exported from the Raman spectrometer comes as a separate ".csv" file with the data for each frame stacked in a vertical column. Python is a mature language, and there are already numerous examples of code written to handle this type of data involving multiple files and the multiple frames per file. In this project a piece of code was utilised which was designed to easily handle the averaging of multiframed outputs, filtering bad frames, correcting the baseline of the spectra and plotting the data in a graph with the option of labelling peaks if needed. This code was a tool to help productivity and could handle averaging multiple files at once but was limited as it could only work with a single dataset. A second programme was developed to visualise multiple files at once, with baselining capabilities as well.

The main issues from the inherited sets of code, while functional, was that they were not very efficient being long, outdated and crudely written. In addition, they were only made to handle data from the Raman spectrometer. The first step towards improving the code was to understand each section and identify what could be optimised. The second step was to streamline the code and allow the option for analysing different data types, which took a few rounds of improvement to optimise. The next step was to improve the ease of formatting for the input and output directories, as the inherited code used redundant functions due to its age.

The first function (`dir_path_setup`) that is now redundant, was written as follows:

```
def dir_path_setup(x, folder_list, dir_paths, output_dir_path = 'a',
number_of_main_dirs = 0):
  '''Always put the initial output directory first followed by all the
subdirecteries. If you have more than one main directory being created and want
subdirecteries in them, create an if/else line in the loop to change the output
directory like so:
  for x in range(len(folder_list)):
    if x == 2:'''
  if number_of_main_dirs > x or number_of_main_dirs == x:
    if not os.path.exists(os.path.join(home, folder_list[x])):
      os.makedirs(os.path.join(home, folder_list[x]))
```

```
    dir_paths[x] = os.path.join(home, folder_list[x])
    return (dir_paths[x])
  else:
    if not os.path.exists(os.path.join(output_dir_path, folder_list[x])):
      os.makedirs(os.path.join(output_dir_path, folder_list[x]))
    dir_paths[x] = os.path.join(output_dir_path, folder_list[x])
    return (dir_paths[x])
```

This function sets up the output directories and works through the list given (`folder_list`) and checks if there is a matching folder in the `home` directory and if the folder does not exist, the function will create the folder and then the function will add the full folder directory to another list (`dir_list`). This is similar to the next function (`folder_dir_setup`), which is as follows:

```
def folder_dir_setup(x, folder_list, dir_list):
  if int(folder_list[x][-3:]) < 26:
    variable_name = 'JS001 - JS025'
  elif int(folder_list[x][-3:]) > 25 and int(folder_list[x][-3:]) < 51:
    variable_name = 'JS026 - JS050'
  else:
    variable_name = 'JS051 - JS075'
  dir_list[x] = os.path.join(home, variable_name, folder_list[x])
  return dir_list[x]
```

This function is similar to the previous one, where it works through a list of folders (`folder_list`) and assigns the full folder directory to another list (`dir_list`). The second function was tailored to the folder structure that was being used to store data at that time.

The next step after this was to combine the two functions (`dir_path_setup` and `folder_dir_setup`) into one function that can handle any data storage method and pull the files and their directories into an easily accessible dictionary. The resulting function was called `print_files`, which would walk through all of the directories in the directory tree and check for any files that matched the input parameters. If any files matched the parameters, the function would save them to a dictionary with the file name being the key for the file path.

The next focus was analysing the data, which is where `process_data` comes in. This function is made to be able to handle large amounts of data from IR, UV/vis, Raman or fluorescent spectroscopy. It does this by going through each file in a list, opening the file and extracting the data before saving it into a dictionary with the file name being the key that corresponds to all the intensities contained in that file. The wavelengths/wavenumbers are also saved into the same dictionary with the key being the file name plus “ WN” or “ WL” attached as a suffix. This allowed the data to be easily accessed later in the code if necessary.

The final step was to work on an easy way to visualise the data. This can be done by the `quick_plot` function, which takes the data from the dictionary where the `process_data` has placed the analysed data and can plot multiple data sets on separate plots as long as the data sets are all sourced from the same technique. The `quick_plot` function also has the capability to add peak labels to the graphs and to save the output figure if so desired.

The main three functions (`print_files`, `process_data` and `quick_plot`) utilise a set of smaller functions outlined in Table 2.1 and a brief flowchart of each of the main functions is displayed in Figure 2.1, Figure 2.2 and Figure 2.3.

Table 2.1: Overview of the minor functions

Function Name	Function Description
`comparing_lists`	Requires two lists and compares one list to another. If there are any matches between the two lists, it will return true.
`check_file_type`	Requires a file list, the address of the files, a file type and a string to look for at the start of the files. The function then opens each file given and compares the first line of the file with the input string and returns true if they match.
`investigate`	Requires a directory and file type, the string the files will start with and a list of NMR files without a file extension. The function will open the folders in the given directory, then opens any folders contained within that folder. If there are any files contained therein, then the function will open the files individually and compare the first line to the input string and if any match the input string, the function will return true.
`update_progress`	Requires an input value between 0-1. The function will clear everything printed and then print the progress bar, which is made of solid and blank spaces, along with a percentage at the end. The number of solid spaces is based upon the input value.
`xlsx_analysis`	Requires a list of files and two dictionaries, one with the file addresses, and a second to save the data into. The function creates a sub-dictionary in the data dictionary, using the file name as the key. It will then add the whole scanning fluorescence data set, excitation wavelengths, absorbance wavelengths, and individual data sets for each excitation wavelength into the sub-dictionary.
`baseline_als`	Requires the “y” values (e.g. Raman intensities or UV/vis absorbances), a parameter (p) for asymmetric analysis, a parameter (λ) for the smoothness of the baseline, and the number of iterations through

	the analysis. The function analyses the data and returns a list of values that represent the baseline for the data given.[38]
`peak_labels`	Requires the x and y values, the plot and the technique that is being plotted. The function will baseline the data and then find the peaks using another premade function with different set variables based on the technique. The parameters that change are the width, prominence, height, threshold and distance. The prominence, height and threshold do similar things about the height of the peaks compared to the baseline. The width is the width of the peaks and the distance is the distance between peaks. The function will then annotate the peak labels on the graph in the correct position.[39]
`regr_line`	Requires the x and y values. It can also take an axis if one is plotted, x and y labels, the colour of the line on the plot and whether to show the legend. The function will use a linear regression calculation to create a line of best fit and then plot it on the active plot.
`my_r2_score`	Requires the x and y values and a colour. Calculates the R^2 value of a scatter plot and returns the value. If there is a graph plotted, this function will also plot the linear line of best fit in the colour specified.

2.2 Flowcharts

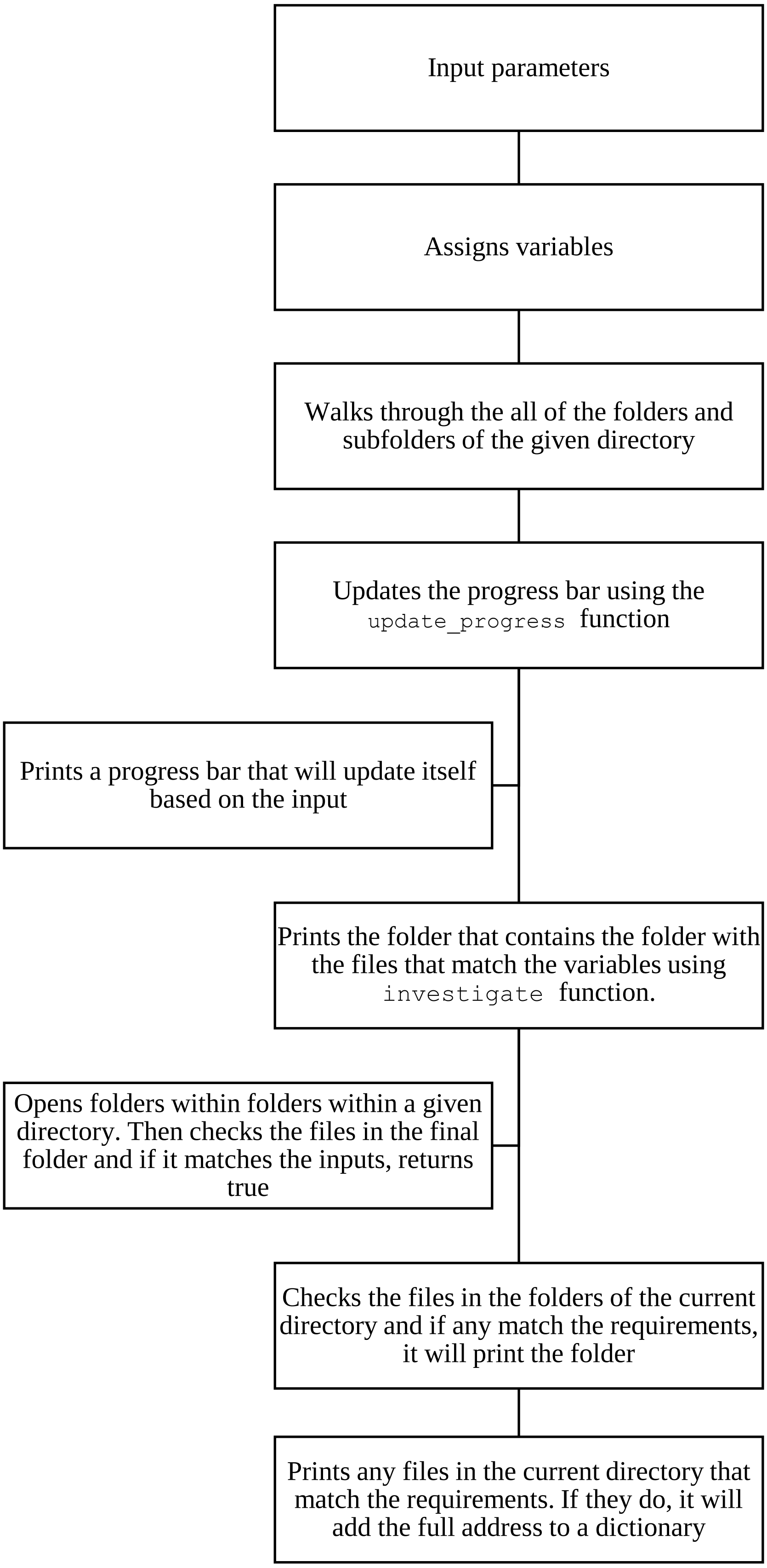


Figure 2.1: A flowchart briefly explaining the `print_files` function

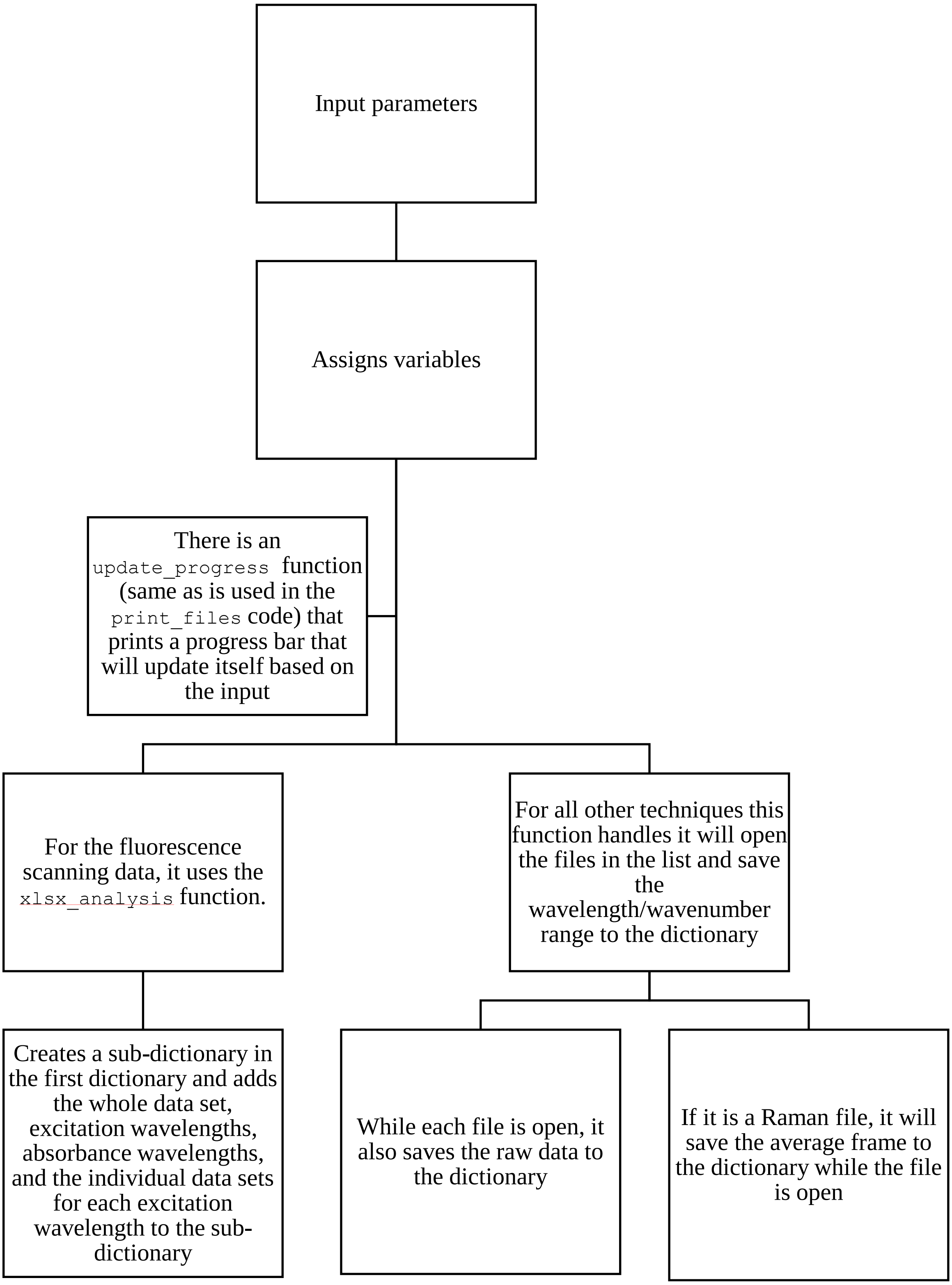


Figure 2.2: A flowchart briefly explaining the `process_data` function

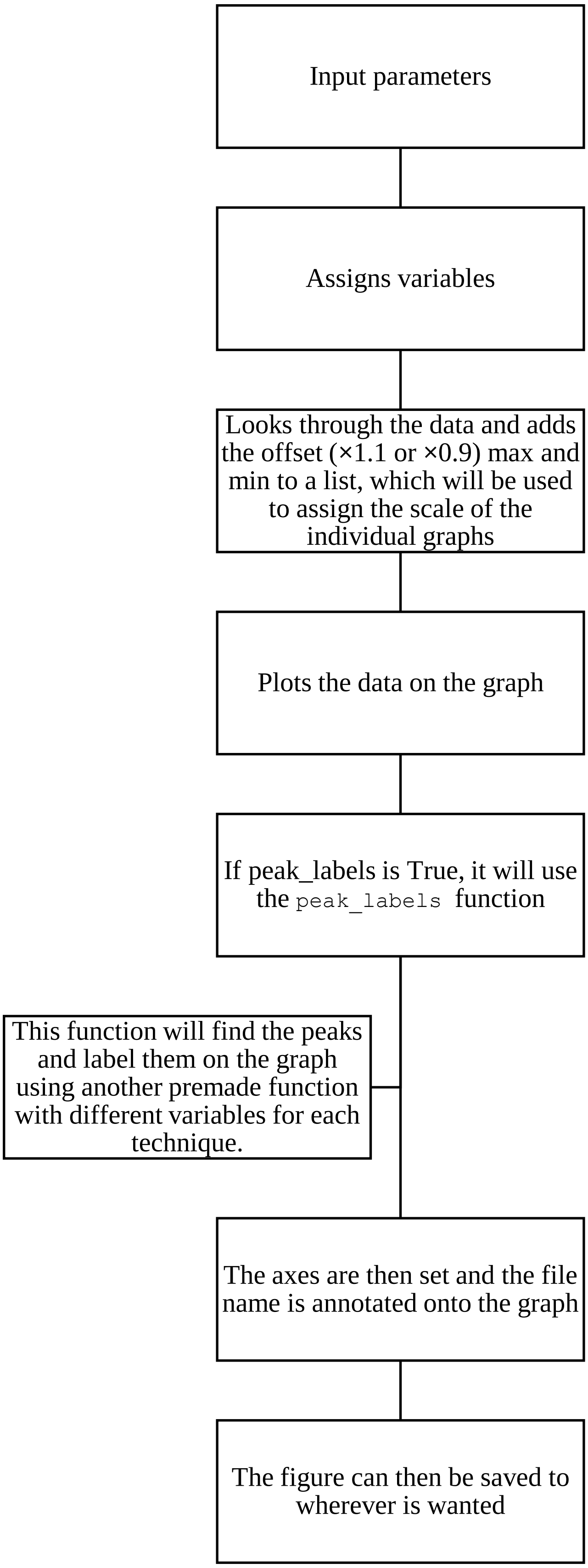


Figure 2.3: A flowchart to briefly explain the `quick_plot` function

2.3 Breakdown of the Main Functions

The function `print_files` was written to visualise the files that are from certain techniques and where those files are located, so that the data from multiple files can easily be extracted and manipulated, regardless of the file storage structure that is being used. `print_files` is started by setting some variables to help simplify the future steps and then walks through all the folders of the directory tree. Then using `investigate`, which checks the files contained two directory levels down from the current directory, it prints the folder if it returns true. Following this, `print_files` will look into each folder and print out the folders that contain a file matching the requirements and then prints out those files as well as adding the full directory path to a dictionary with the key being the file name. This allows the files to be opened and worked on very easy, but creating this function presented some problems that had to be worked through.

One consistent problem was to filter out any NMR files and folders, as some files have no file extension or may cause issues with the progress bar. This made it necessary to explicitly filter out the NMR files and folders at multiple times along the way, to ensure that they will not interfere with the functionality of the code.

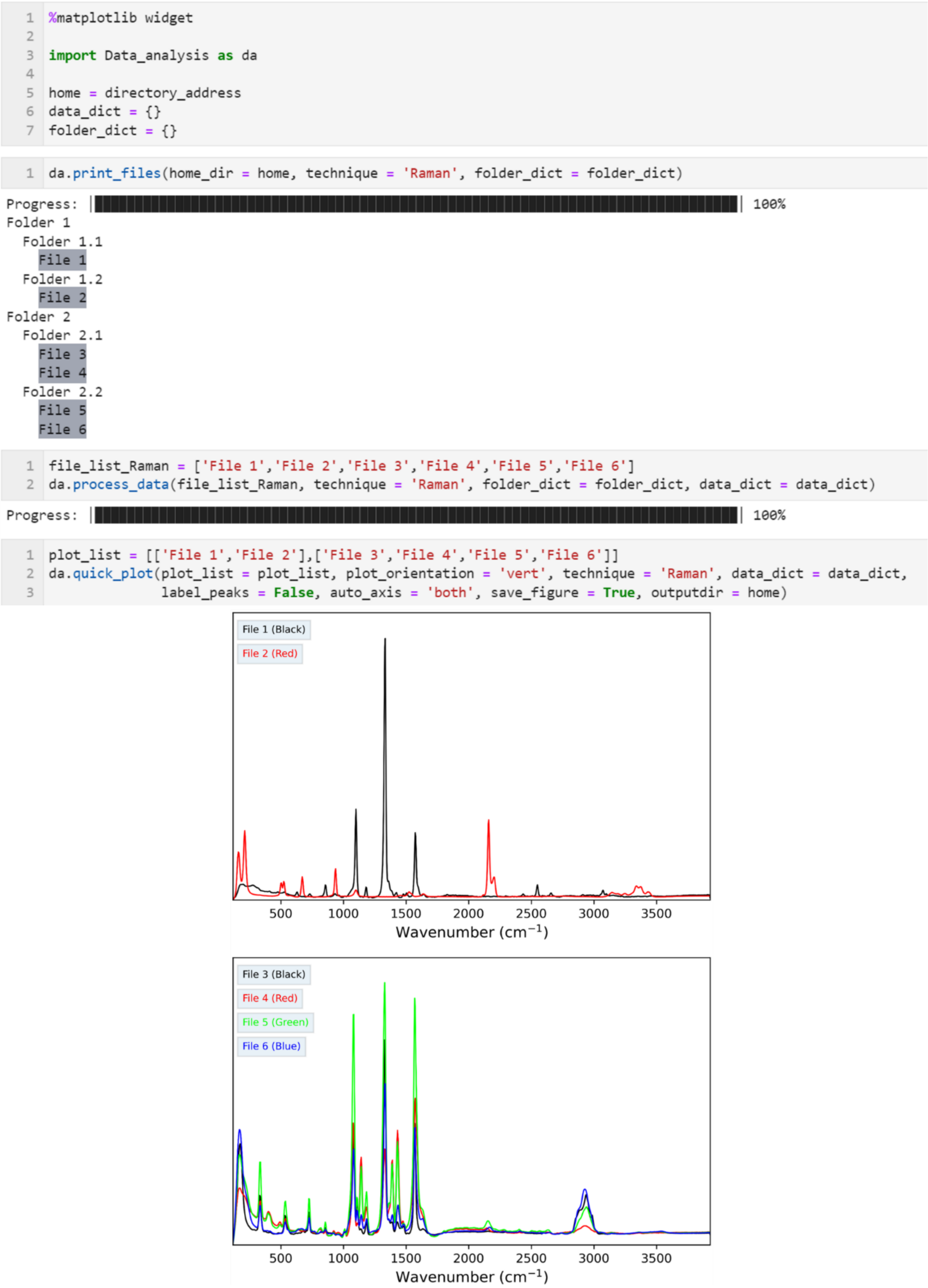
Initially the `print_files` function would print out the folder that contains the file and not the folder containing the folder with the file. Integrating the extra step was complicated, which is where the `investigate` function originated from. The `investigate` function has multiple "if" statements to filter through different aspects of the folder setup and the different files that could interfere. The main files that interfere with this, were associated with NMR data as those files do not have a file extension and so cannot be filtered out that way. The next hurdle was connected to identifying which files are from which technique. So, `investigate` opens up the files that have the correct file extensions and compares the first character of the files to pre-set variables, to determine if the file is from the correct technique.

The last of the big issues to work through was to integrate the `update_progress` function into the code to help monitor the progress. The main reason this was an issue was that in the `update_progress` function, there is a line that will clear the printed text, which allows the progress bar to update. However, this line removed all of the files and folders that were printed out as well. To overcome this, instead of printing the folders and files as they are found, they are added to a list and then before the walk function moves on, `print_files` will go through the list and print out all of the folders and files individually.

After visualising the files from varying techniques and recording where those files are located in a dictionary, the next step is to process the data using `process_data`. `process_data` starts by setting variables to simplify the code and shorten the code. The simplest data sets to analyse come from IR, UV/visible spectroscopy and fluorescent spectra from a single excitation wavelength. For these data sets, the function will open each file and save the wavelength/wavenumber values and the intensity as two separate entries to a dictionary. For the Raman data if the filtering spectrums is false, `process_data` will first average all the frames of the data before it saves the wavenumbers and intensities to the dictionary. If the filtering spectrums is true, `process_data` will plot the array of data from one Raman file, then ask for a wavelength to focus on. `process_data` will then plot a histogram with the intensities of the data at the wavelength specified and based on the inputs, certain outliers will be filtered out before the remaining data is averaged and saved to the dictionary. For the fluorescent scans, the function `xlsx_analysis` was written to open the Excel spreadsheet and save the full data set, excitation wavelengths, absorption wavelengths and the intensities for each excitation wavelength in a sub-dictionary.

The final main function, `quick_plot`, was designed to quickly visualise the data by plotting it. This function takes the data from the dictionary that came from the `process_data` function and then plots the selected data sets on a graph. There are multiple features that have been integrated into this function, which include changing the orientation of the plots, the number of plots, what technique (IR, Raman, UV/vis or fluorescence) is being plotted, peak labelling, auto-scaling and the option to save the figure. The problems that arose during the writing of this function were all simple and easy to resolve. One of the more complex problems was from setting the axes, as the different spectra had different scales. When plotting the graphs on a single plot, sometimes the axes would be set to the final graph rather than the largest/smallest point, which in some cases cut out some parts of other graphs. So, a few lines of code were added to set up the vertical limits automatically. To start with, the code will determine the maximum and minimum “y” values and then multiplies them by 0.9 or 1.1 depending on if they are positive or negative so the plots will not cut off any of the graphs plotted and appends those values to a list. Following this, the function will set the axes after all graphs have been plotted based on the maximum and minimum values in the list.

2.4 Conclusion



```
%matplotlib widget

import Data_analysis as da

home = directory_address
data_dict = {}
folder_dict = {}
```

```
da.print_files(home_dir = home, technique = 'Raman', folder_dict = folder_dict)
```

```
Progress: |████████████████████████████████████████████████████████████████████████| 100%
Folder 1
  Folder 1.1
    File 1
  Folder 1.2
    File 2
Folder 2
  Folder 2.1
    File 3
    File 4
  Folder 2.2
    File 5
    File 6
```

```
file_list_Raman = ['File 1','File 2','File 3','File 4','File 5','File 6']
da.process_data(file_list_Raman, technique = 'Raman', folder_dict = folder_dict, data_dict = data_dict)
```

```
Progress: |████████████████████████████████████████████████████████████████████████| 100%
```

```
plot_list = [['File 1','File 2'],['File 3','File 4','File 5','File 6']]
da.quick_plot(plot_list = plot_list, plot_orientation = 'vert', technique = 'Raman', data_dict = data_dict,
              label_peaks = False, auto_axis = 'both', save_figure = True, outputdir = home)
```

Figure 2.4: An example of the inputs (grey boxes) and outputs (between the grey boxes) for the three main functions. The data is from work done in Chapter 3 and Chapter 5.

Overall, there are three main functions that have been written to first find the files and visualise them in a list (`print_files`), then the second to process the data into a simple but usable format (`process_data`) and finally to graph the data easily in a plot (`quick_plot`). An example of the inputs and outputs of these three functions can be seen in Figure 2.4. It can be seen that the inputs and outputs are simple and easily varied for different datasets.

These three main functions utilise multiple smaller functions, as outlined in Table 2.1, to simplify and streamline the main functions and this code does not require any third-party software (that often have high subscription fees). However, as these minor functions are separate from the main functions, they can be utilised in other sets of code. The next step in improving the code would be adding different techniques to the main functions to allow the code to be more versatile.

Chapter 3 – Nitrate and Phosphate

3.1 A Brief Background of Fertilizers

Due to ever increasing human population, there remains a need for more intensive agricultural areas. One solution to this is to intensify agricultural use, leading in turn to more fertilizer being applied to farms and orchards, as seen in Figure 3.1. The most common nutrient elements applied to most farms or orchards are nitrogen as ammonium or nitrate, phosphorus as a form of phosphate, and potassium, as they are required in larger quantities than other nutrients.[40]

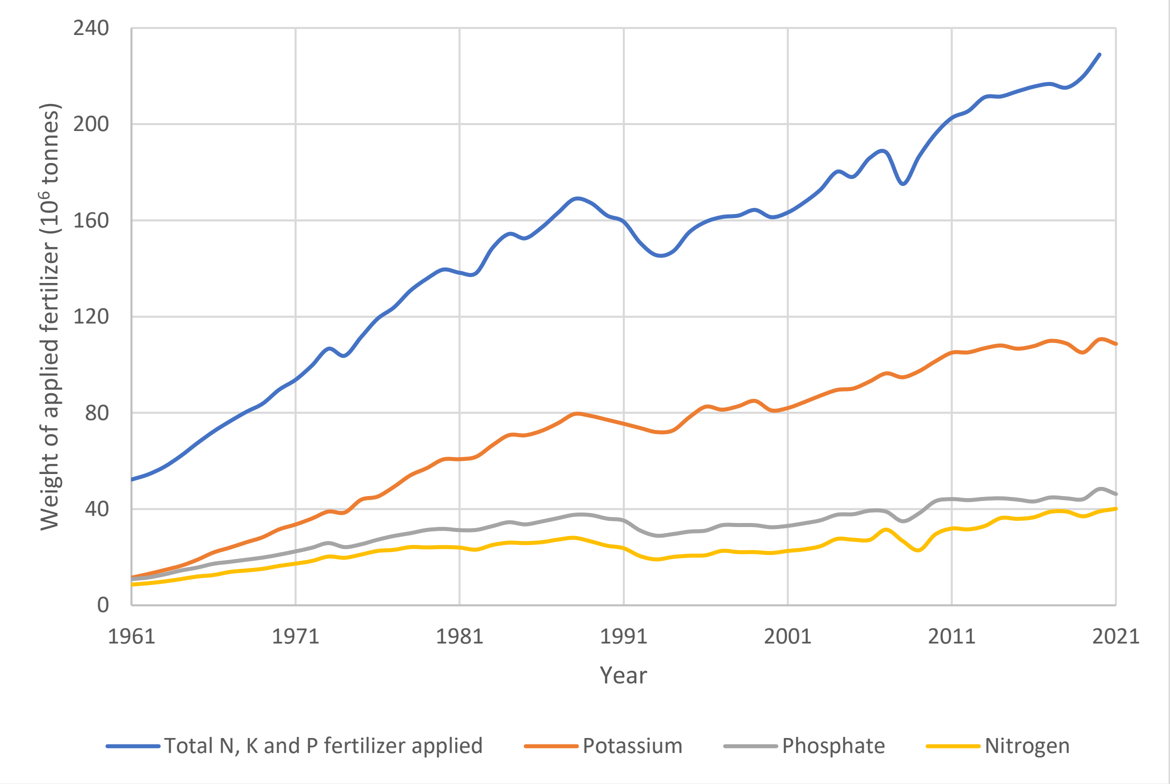


Figure 3.1: Weight of applied potash, phosphate, nitrogen fertilizers[41] and the total weight of applied nitrogen, potassium and phosphorus fertilizers.[42]

Potassium is important for many physiological processes in plant cells including the regulation and movement of water, nutrients and carbohydrates, ion homeostasis and enzyme activation.[43,44]

Nitrogen and phosphorus are key components in DNA bases and amino acids, making them necessary for plant growth.[45,46] Nitrogen compounds can come from natural sources like igneous rocks, atmospheric deposition and symbioses of some plants, along with cyanobacteria and some heterotrophs.[44,47] Nitrogen compounds can also come from anthropogenic sources, including nitrogen-rich fertilizers used for agricultural purposes, discharge of poorly treated domestic, and industrial wastewaters, livestock manure, and leachate from landfill sites.[47,48] Phosphorus compounds can come from natural sources like plant decomposition or atmospheric deposition. They can also come from untreated sewage, recycling of organic wastes and the biggest contributor is phosphate-rich fertilizer.[49]

Although there are many different chemicals used in fertilizers, nitrate and phosphate are the focus of this chapter because they can be detrimental to the environment and human health. Nitrate is an anion comprised of one nitrogen atom bound to three oxygen atoms and phosphate is comprised of one phosphorus atom bound to four oxygen atoms. Both nitrate and phosphate are used extensively in fertilizers, as outlined above, and nitrate is also used as a pesticide.

3.2 Reasons to Detect Nitrate and Phosphate

As nitrate and phosphate are vital to growing crops, they are both applied extensively in agriculture to try and maximise the crop production as previously mentioned. This leads to large amounts of nitrate and phosphate running off the farms or orchards or leaching into groundwater and waterways, as they are not completely absorbed by the plant or in the soil.

Nitrate is generally considered to be the most widespread groundwater pollutant due to its high-water solubility and its lack of soil interaction, so it can leach into groundwater easily.[17] Phosphate does not leach into groundwater as readily as nitrate. This is due to phosphate being immobilised in the soil initially, unavailable to plants leading to excess phosphate being applied. The excess phosphate will then start leaching into the groundwater.[50] Phosphate runoff is more of a concern than phosphate leaching.

When nitrate and phosphate leach into groundwater or waterways, they both have toxic effects. Both nitrate and phosphate can also cause eutrophication in waterways, which is the build-up of nutrients, leading to algal blooms, that are harmful to marine environments.[51,52] Eutrophication is harmful to marine environments because the algae consumes oxygen causing fish and other marine organisms to suffocate and causes the water quality to go down. It also makes the water more favourable for the rapid growth of toxin-producing cyanobacteria, destabilising the ecosystem.[53] Excess nitrate can also cause methemoglobinemia (blue baby syndrome) in infants and stomach cancer in humans,[18,54] whereas excess phosphate can also accelerate aging and harm blood vessels in humans.[55]

3.3 Current Methods for Detecting Nitrate and Phosphate

Nitrate and phosphate have both been extensively studied, but there are few methods that use cages as the detection method. There have been many reviews published outlining detection methods for nitrate and phosphate.[55-61]

Nitrate and phosphate can be detected by a variety of methods including electrochemical (voltammetry, potentiometry), biological sensors, chromatography (high pressure liquid chromatography, ion chromatography, gas chromatography) and spectroscopic techniques (colorimetry, fluorometry, chemiluminescence, Raman, infrared).[55-60]

Electrochemical sensors are a promising prospect as they have high sensitivity, swift response times and are often user-friendly. The electrochemical devices are low powered devices that require minimal sample pre-treatment. Currently these devices are too large for environmental applications, but there is research going into the miniaturisation. The electrodes used are often modified, as bare electrodes can get easily poisoned by interferants.[56-59]

Biological sensors are another option for detecting nitrates, often based on the nitrate reductase enzyme. These sensors are generally highly specific and sensitive at neutral pH as they are enzymes that specifically target nitrate. As they require enzymes, they are expensive and require low storage temperatures. The repeatability, the sensitivity to oxygen interfering with the enzyme, and the low electron transfer efficiency of the redox cofactor used to reactivate the enzyme, limit the applicability of them in environmental application.[56,57,62]

Chromatography techniques require relatively small volumes of samples, are very selective based on the column derivatization and generally have a greater sensitivity than electrochemical methods. However, they are slow and require specialised equipment, making chromatography techniques less applicable to real-time environmental testing.[61]

Spectroscopic techniques are fast and require relatively small volumes of samples and generally have a greater sensitivity than electrochemical methods. However, spectroscopic techniques often require derivatization procedures, sample preparation, and require specialised equipment, making spectroscopic techniques less useful for *in situ* environmental testing. In general, spectroscopic techniques can be classified in three main categories: Direct detection, catalytic detection, and complexation. Direct detection of nitrate using spectroscopic techniques usually requires the most sample pretreatment. Catalytic detection monitors a reaction that nitrate will be involved in, often the oxidation of indicators or organic dyes. Detection by coordinating nitrate to a complexing agent requires monitoring the signal of the complexing agent. There are multiple types of complexing agents including polymers, membranes, and in the case of this project, cages.[55-57,60]

3.4 Target Cages and Design Features

All the cages in this section, shown in Figure 3.2, were designed with a common core to facilitate synthesis. Any variation occurred at the end of each molecule. All three target cages contain a piperazine bridge, metal binding sites with a coordinated copper and a varying protected sulfur group at the ends of the ligands.

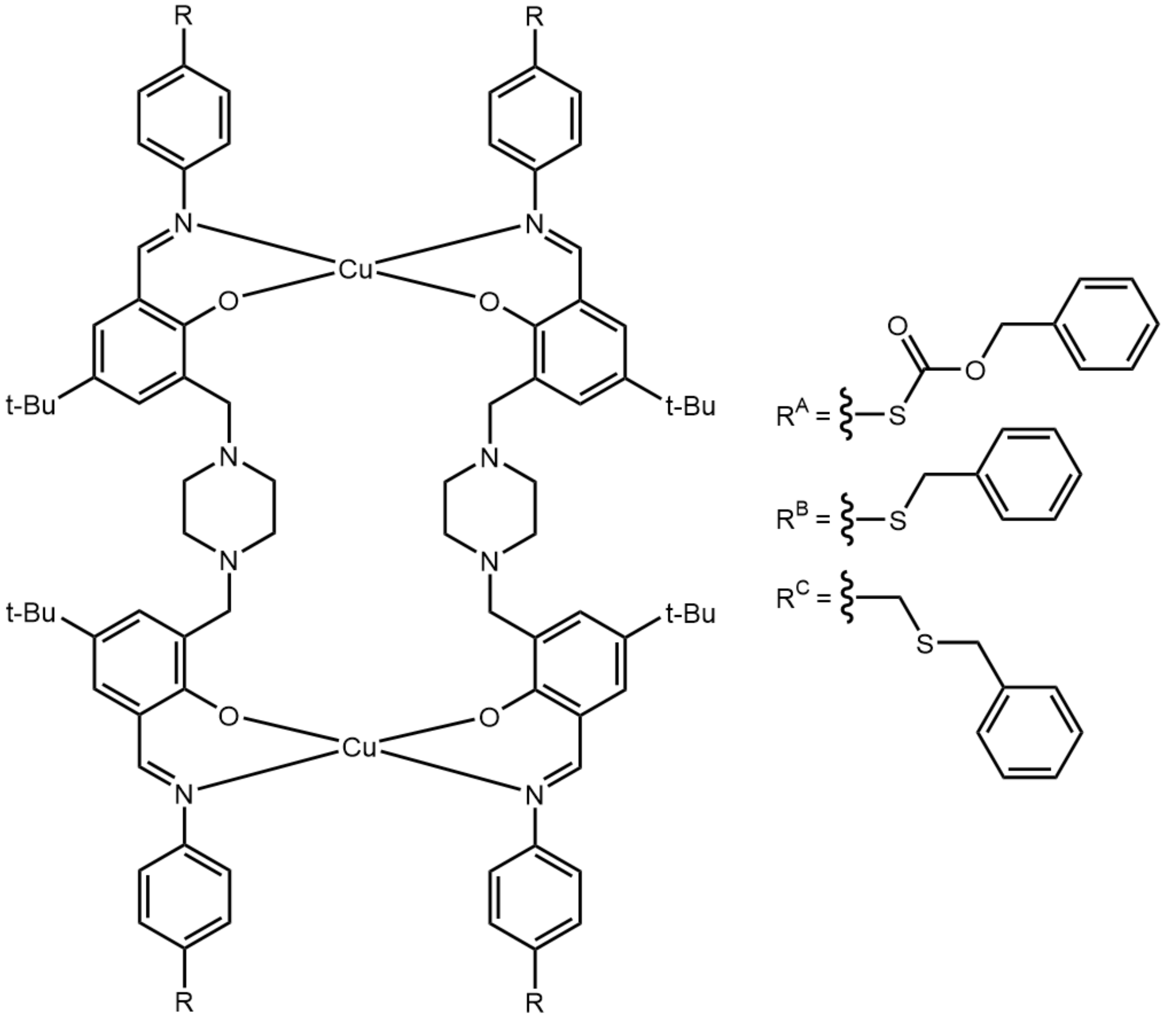


Figure 3.2: The main target cages ([Cu_2(**1**)$_2$]) to capture nitrate and phosphate.

The piperazine bridges were chosen as they are short leading to a small binding cavity, to exclude large molecules from the cavity. The nitrogen atoms of the piperazine linker can be protonated allowing it to ionically bond either nitrate or phosphate. This protonation can be reversed based on the pH, changing the chemical environment of the cavity. This would make the cavity unfavourable to bound anions, releasing them from the cavity, allowing the cage to be used again. Copper was chosen because it can switch between square planar, square pyramidal and octahedral coordination modes depending on the need of the complex, allowing the anions to coordinate to the metal centres. The metal binding sites are based on an oxime group, which is known to bind strongly to copper. The sulfur groups on the end of the ligands are to be attached to nanoparticles after the cage has been synthesised to utilise the Surface Enhanced Raman Spectroscopy (SERS) effect to enhance the Raman signal of the cage. So the sulfur group does not interfere with any reactions, a protecting group has been used. Initially the benzyloxycarbonyl protecting group ($\mathbf{R}^{\mathbf{A}}$) was studied but this had to be changed to the benzyl protecting group ($\mathbf{R}^{\mathbf{B}}$) after it was discovered that the starting materials had decomposed. Unfortunately, time constraints prevented the obtainment of more of the desired benzyl chloroformate.

The sulfur atoms at the end of the cage, with the ligand $\mathbf{R}^{\mathbf{A}}$, point perpendicular to the metal-metal axis. This could diminish the binding efficiency of the cage to a nanoparticle because of how close the sulfur is to the rest of the cage. To try to rectify this, a methyl linker has been added between the sulfur and benzene ring to orient the sulfur along the metal-metal axis. This change ($\mathbf{R}^{\mathbf{C}}$) should position the sulfur away from the rest of the cage giving it more freedom and allowing it to bind to the nanoparticle easier as seen in Figure 3.3.

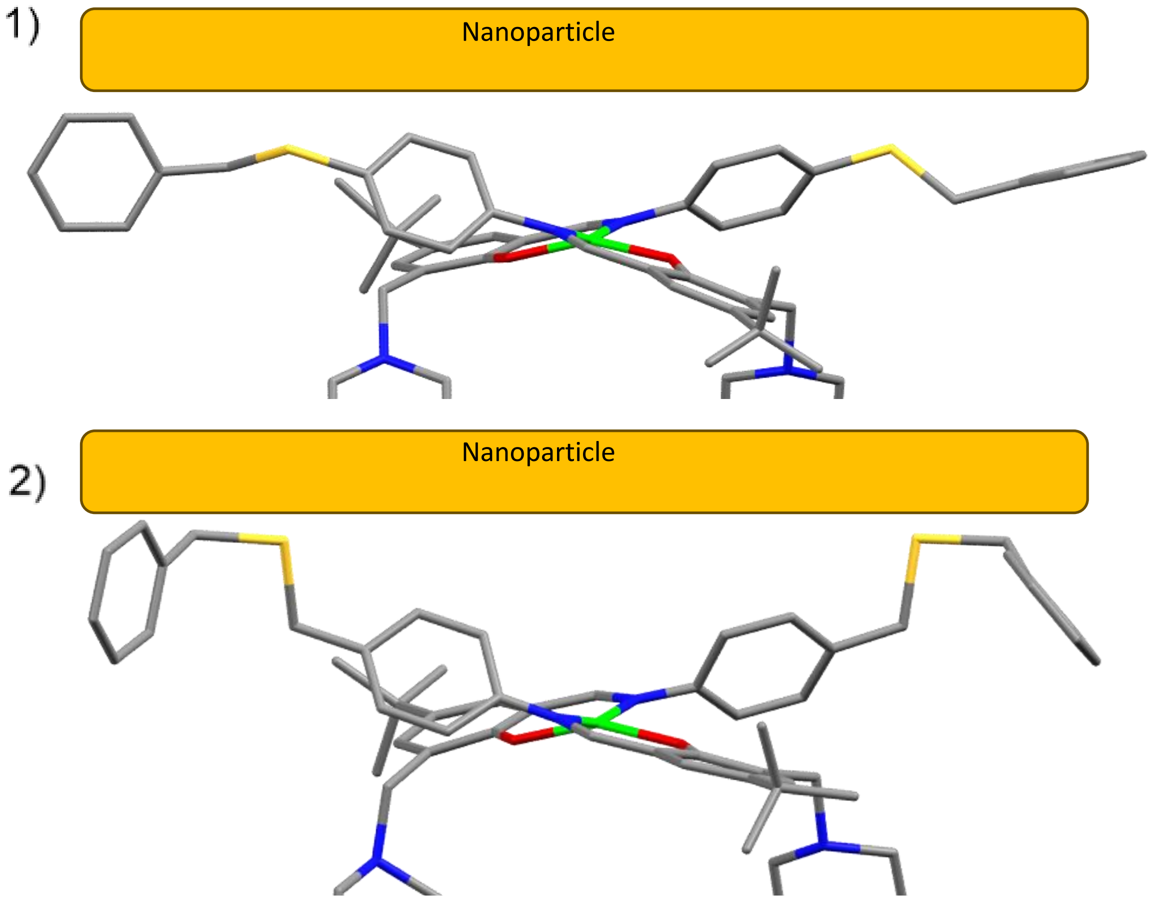


Figure 3.3: A model of the complexes with $\mathbf{R}^{\mathbf{A}}$ (1) and $\mathbf{R}^{\mathbf{B}}$ (2) interacting with nanoparticles

The binding of nitrate and phosphate in the cage system has been modelled using a simple geometry optimising calculation, which performs an iterative series of single point calculations until it reaches an energy minimum. This is visualised in Figure 3.4, and displays how nitrate and phosphate could interact with the cage. In both cases, the protonated piperazine nitrogens and the copper centres are interacting with nitrate and phosphate.

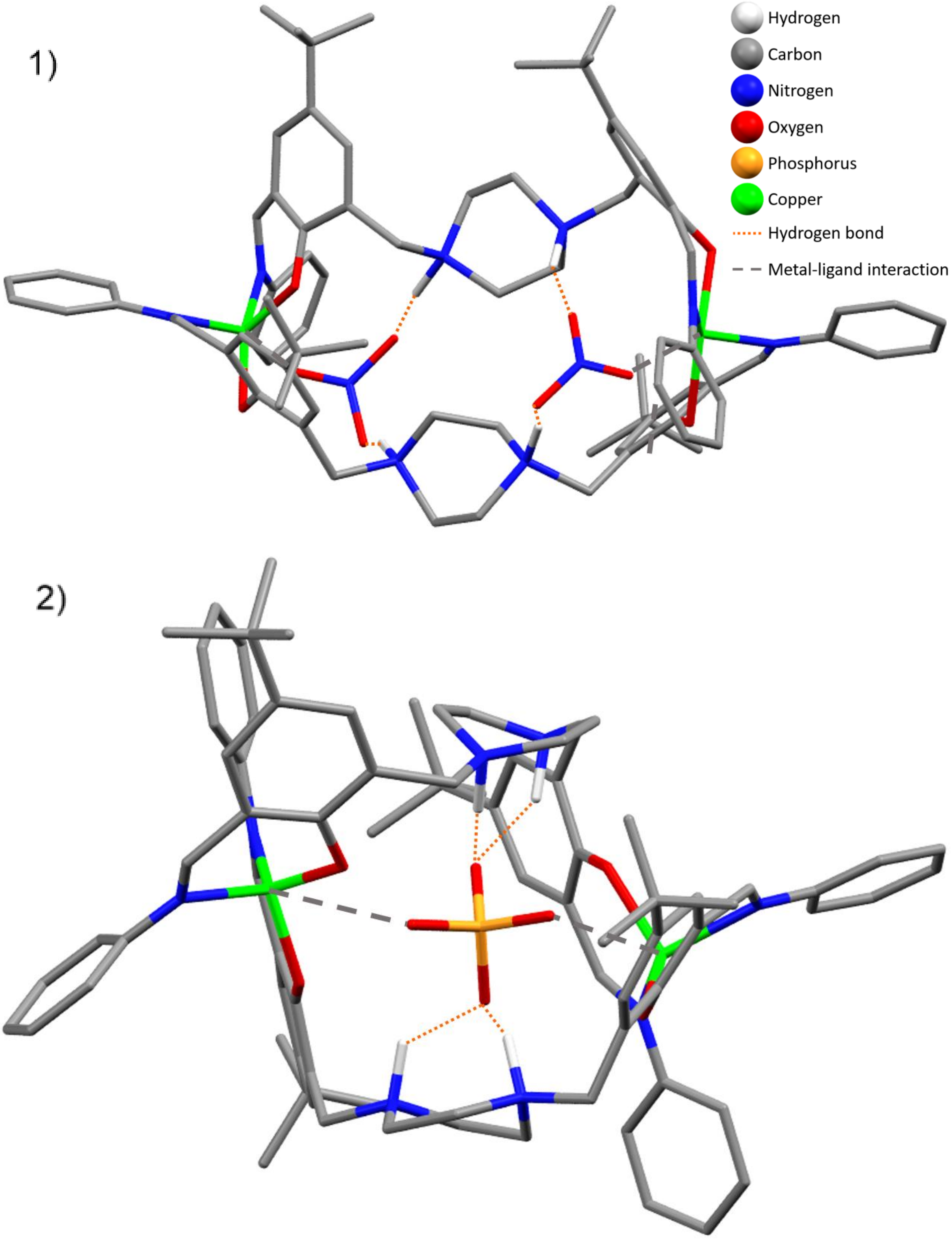
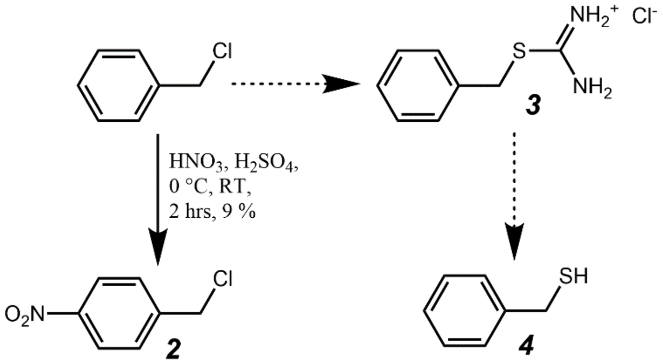


Figure 3.4: The target cage interacting with 1) nitrate and 2) phosphate. The terminal groups and non-interacting hydrogens are removed for clarity.

3.5 Synthesis and Analysis

The first cage that was synthesised ($[Cu_2(\mathbf{1}^{\mathbf{A}})_2]$) was based on a cage prepared by a previous student from the Plieger group (Josh Blazek).[63] There are two parts to the synthesis, the terminus containing a protected sulfur and the bridging unit, with the copper binding sites connected by a piperazine unit. The second cage is prepared in an analogous fashion and only differs by use of benzyl chloride instead of benzyl chloroformate. The third cage was like the second, but with a methyl bridge between the sulfur and the benzene ring. The third cage synthesis has three steps before following the same synthetic steps as the previous two cages, as seen in Scheme 3.1.

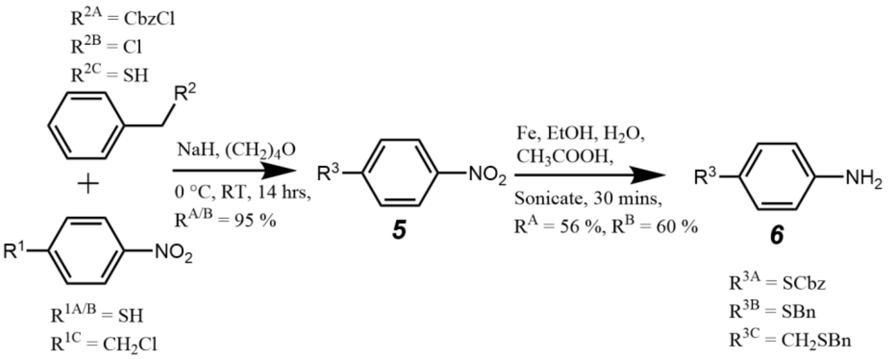


Scheme 3.1: First synthetic step of the third cages' terminal group.

The three reactions are split between two pathways. The first is a nitration of benzyl chloride and the second pathway is a two-step sulfation from benzyl chloride. The nitration reaction was achieved by mixing benzyl chloride, nitric acid and sulfuric acid together at 0 °C, before stirring at room temperature for 2 hours. The crude material was then purified by column chromatography eluting with ethyl acetate and hexane. This produced the pure nitro aromatic (**2**), confirmed by ^{1}H NMR. The ^{1}H NMR showed a shift in the two aromatic hydrogens *ortho* to the nitro group to 8.23 ppm. The yield was low for this reaction, likely due to alternate nitration products and a portion of starting material remaining unreacted. This was as far as the synthesis of the ligand $\mathbf{1}^{\mathbf{C}}$ reached.

The thiol (**4**) was to be prepared by modification of a procedure from Sienkiewicz-Gromiuk *et al.* This would be achieved by adding thiourea to benzyl chloride in ethanol and water and refluxing the solution for 3 hours before cooling to 0 °C, producing a precipitate of benzylisothiourea hydrochloride salt (**3**). The salt would then be dissolved in sodium hydroxide and refluxed, before acidifying the solution and extracting benzyl mercaptan (**4**) with DCM.[64]

The synthesis of the terminal group, as seen in Scheme 3.2, starts with the protection of the sulfur group with benzyl chloroformate (R^A), benzyl chloride (R^B) or nitrobenzyl chloride (R^C). This is so that the sulfur does not interfere with subsequent reactions or the complexation and can be deprotected after complexation to bind to nanoparticles for detection. The protection reaction was achieved by adding sodium hydride to the thiol in THF at 0 °C, before adding the protecting group (benzyl chloroformate, benzyl chloride or nitrobenzyl chloride) and stirring at room temperature overnight. The crude product was purified by recrystallisation with ethanol to give the protected sulfur product (**5**). **5^A** and **5^B** were confirmed by ^{1}H NMR, which showed a new peak at 5.33 and 4.29 ppm respectively for the two methylene hydrogens and X-ray crystallography as seen in Figure 3.5. The details about the X-ray collection data can be found in Appendix 2. The reaction to produce the product, **5^C**, was not attempted due to time constraints.



Scheme 3.2: The synthesis of the sulfur containing terminal groups.

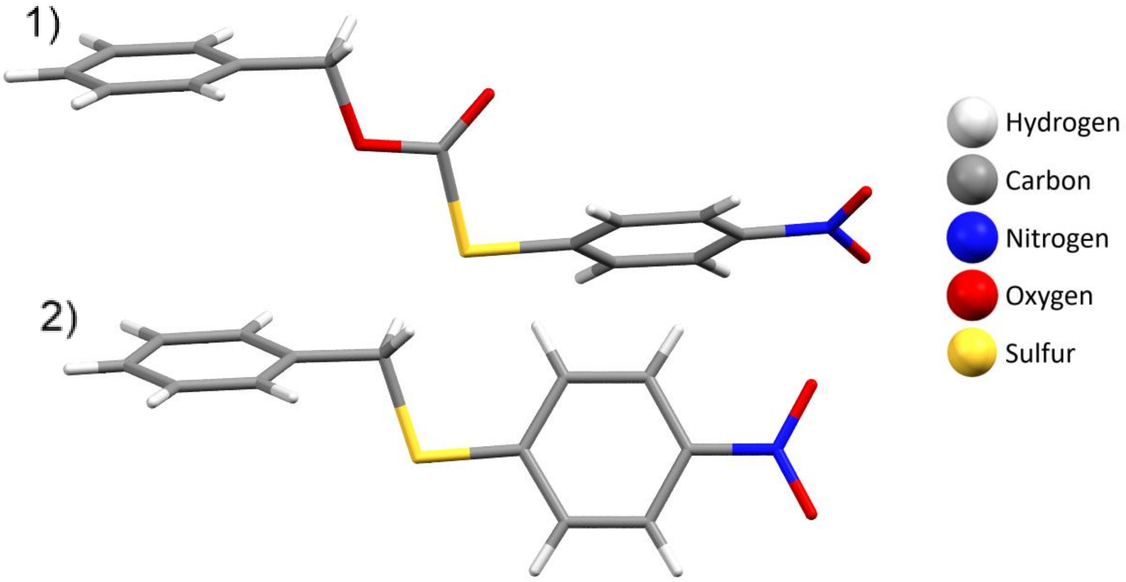


Figure 3.5: X-ray structures of 1) **5^A** and 2) **5^B**.

The crystals were both grown by dissolving the product in ethanol and left to produce the crystals. The molecule **6**^{A} crystallises in the space group $P2_1/c$, whereas the molecule **6**^{B} crystallises in the space group $P2_1/n$ with the asymmetric unit for both molecules containing one full molecule. As seen in Figure 3.6, the packing of **5**^{A} along the c axis is in a zig-zag arrangement, where the methyl bridge, between the benzene and the ester groups, are overlapping. This puts two nitrated benzene rings in proximity (3.578(2) Å) interacting through π-π interactions.

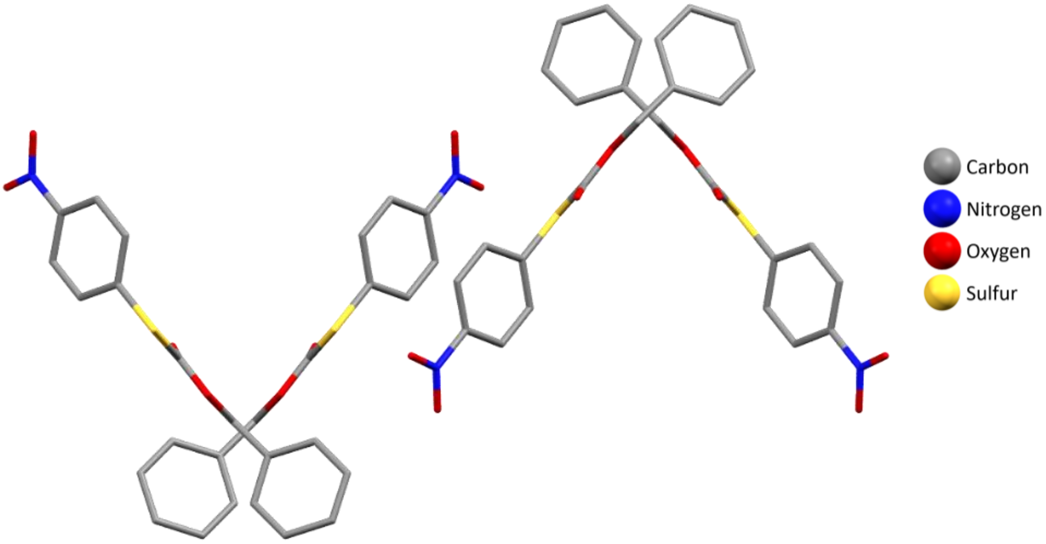


Figure 3.6: The crystal packing of the protected sulfur (**5**^{A}) ligand along the c axis. The hydrogens have been removed for clarity.

The packing of **5**^{B} is a stacked end-to-end fashion along the b axis as seen in Figure 3.7. The benzene rings are not in close enough proximity to each other for π-π interactions, as they sit between 3.675(2) and 3.876(2) Å at their closest points.

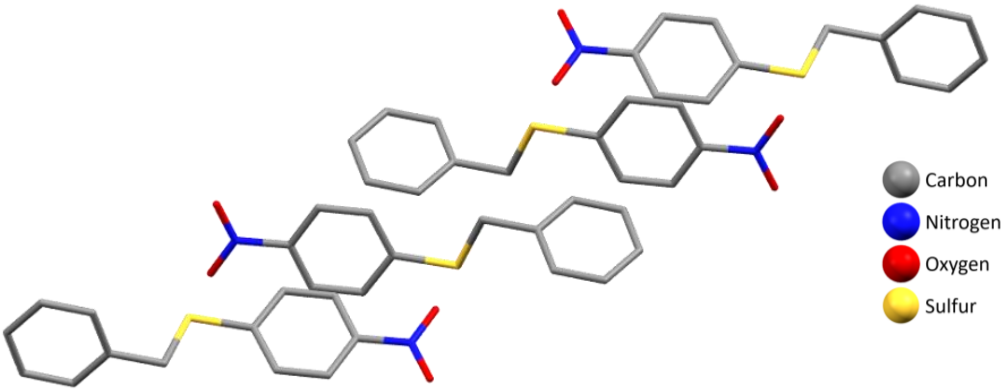
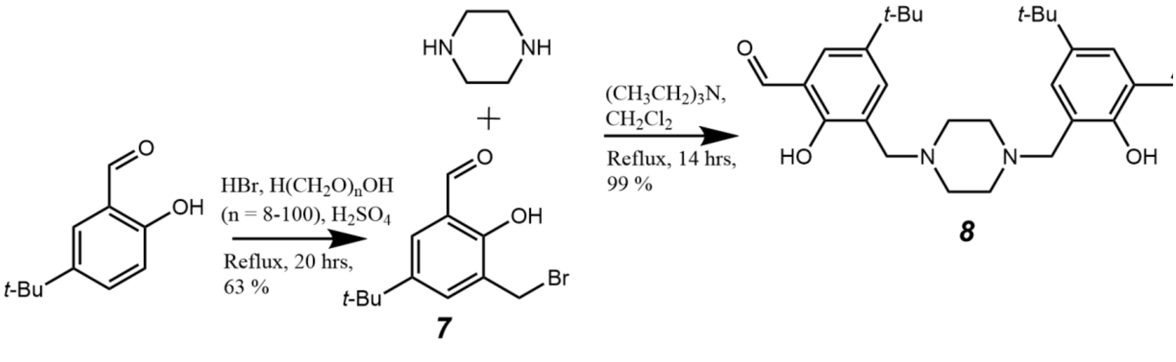


Figure 3.7: The crystal packing of the protected sulfur (**5**^{B}) ligand along the b axis. The hydrogens have been removed for clarity.

The final step for the terminal unit is reduction of the nitro group to an amine. This was achieved by sonicating the protected nitro-thiol (**5**) and iron powder in ethanol, water and glacial acetic acid. The crude mixture was purified by recrystallising the product with ethanol, resulting in the pure amine (**6**). **6^{A}** and **6^{B}** were confirmed by ^{1}H NMR, which agreed with literature.

There are two steps for the synthesis of the bridging unit, shown in Scheme 3.3. The first involves formation of a benzyl bromide (**7**) from 5-*tert*-butyl-2-hydroxybenzaldehyde. This was achieved by stirring 5-*tert*-butyl-2-hydroxybenzaldehyde and paraformaldehyde in hydrogen bromide before sulfuric acid was added and the solution was refluxed. The crude material was extracted and purified by recrystallisation in pentane giving the brominated product (**7**), which was confirmed by ^{1}H NMR, which showed a new peak integrating to two protons at 4.57 ppm.

The second reaction was combining the benzylic bromide (**7**) with piperazine. This was performed by dissolving triethylamine and piperazine in DCM before slowly adding the benzylic bromide (**7**) and heating the solution to reflux overnight. The solution was purified by recrystallisation by dissolving the crude material in hot chloroform, then hot ethanol was added to induce crystallisation. This produced the pure linker product (**8**) and was confirmed by X-ray crystallography and ^{1}H NMR, which showed the combination of the 8 piperazine hydrogens at 2.69 ppm, and the 18 *tert*-butyl hydrogens at 1.32 ppm.



Scheme 3.3: Synthesis of the bridging unit.

The X-ray crystallography of the linker, **7**, is shown in Figure 3.8 and the details about the X-ray collection data can be found in Appendix 2. The molecule crystallises in the space group P-1 and the asymmetric unit contains half of the molecule, with the full molecule generated by a reflection operation between the carbons in the piperazine ring.

Figure 3.8: X-ray structure of the piperazine linker **7,** with non-interacting hydrogens omitted for clarity.

As seen in Figure 3.9, the packing of **7** along the c axis is a diagonally stacked arrangement, where each linker is arranged is the same orientation. The benzene rings are overlapping and are separated by a 3.395(3) Å gap, interacting through π-π interactions. The other interactions to note are the 1.67(4) Å long hydrogen bond between the phenolic hydrogen and the piperazine nitrogen and the 2.41(3) Å long interaction between the hydrogen on the carbon with the aldehyde and the phenolic oxygen.

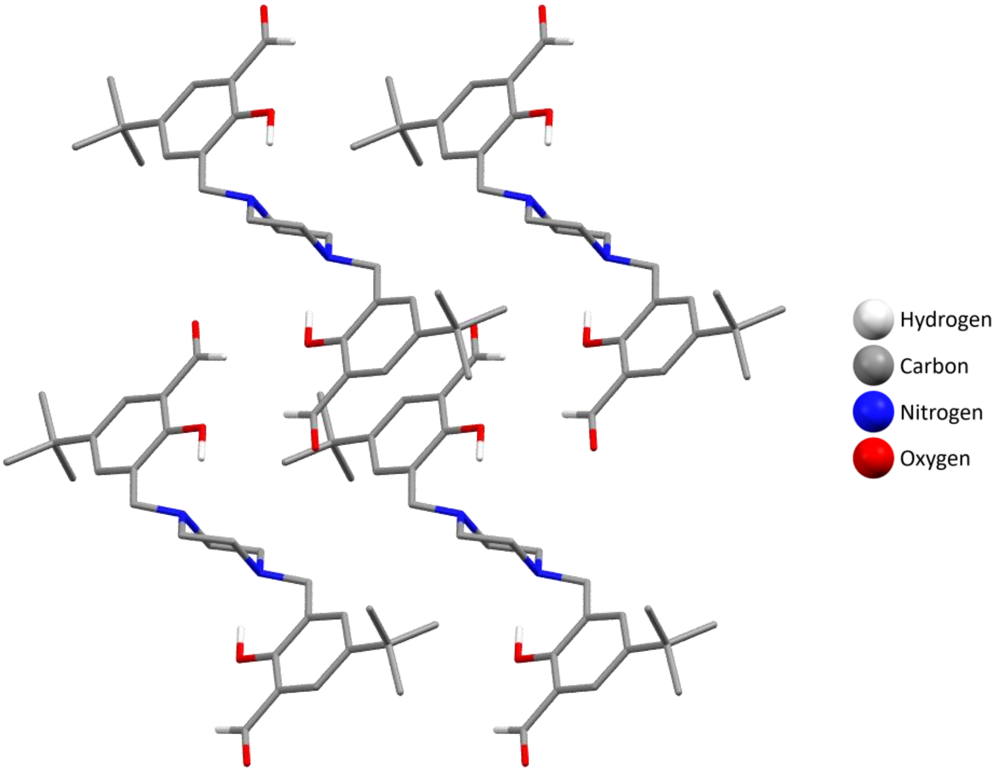
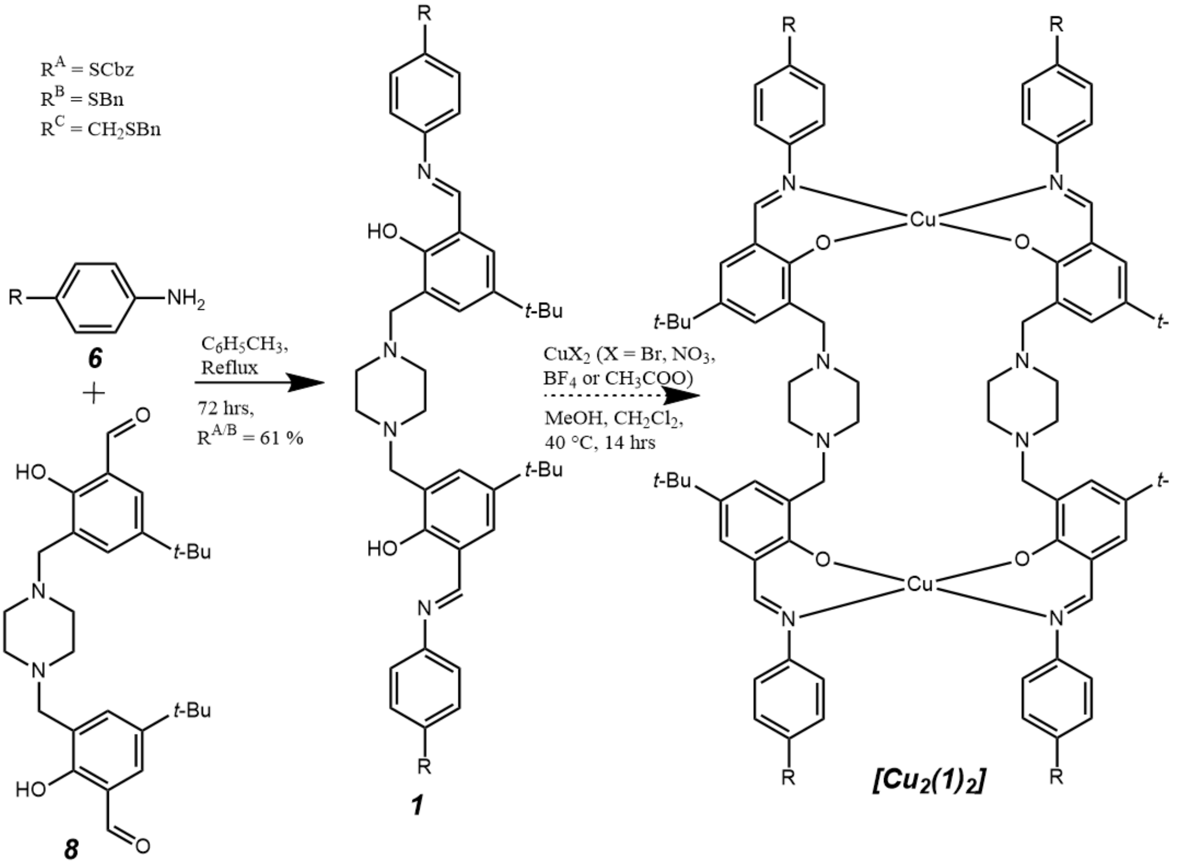


Figure 3.9: The crystal packing of the piperazine linker **7**, along the c axis. The non-interacting hydrogens have been removed for clarity.

To complete the synthesis of the ligand before complexation (Scheme 3.4), the terminal group (**6**) and the bridging unit (**8**) need to be refluxed together in toluene for 3 days, before purification using either chloroform and hot ethanol ($\mathbf{1^{A}}$) or chloroform and hexane ($\mathbf{1^{B}}$). This gave the pure products, which were confirmed by ^{1}H NMR, showing the combination of the methyl bridge from the terminal group integrating to 4 hydrogens and the *tert*-butyl groups of the bridging unit integrating to 18 hydrogens. For the ligand $\mathbf{1^{A}}$ these peaks are at 5.30 and 1.37 ppm respectively and for the ligand $\mathbf{1^{B}}$ the peaks are at 4.15 and 1.35 ppm respectively.



Scheme 3.4: Final cage formation reactions.

The final step is the complexation of each ligand with one of four copper salts (bromide, nitrate, tetrafluoroborate or acetate). This was achieved by dissolving the ligand (**1**) and metal salt in methanol and DCM and stirred at 40 °C overnight. This has been done with the ligands $\mathbf{1^{A}}$ and $\mathbf{1^{B}}$ with X-ray data confirming the formation of the complex using the first ligand ($\mathbf{1^{A}}$) and copper nitrate, as seen in Figure 3.10. The cages with the ligand $\mathbf{1^{A}}$ and the other copper salts have not been crystallised at an adequate quality for X-ray crystallographic analysis but are assumed to have complexed as they showed colour changes like the cage with copper nitrate and showed similar changes in the IR spectrum. The details about the X-ray collection data can be found in Appendix 2.

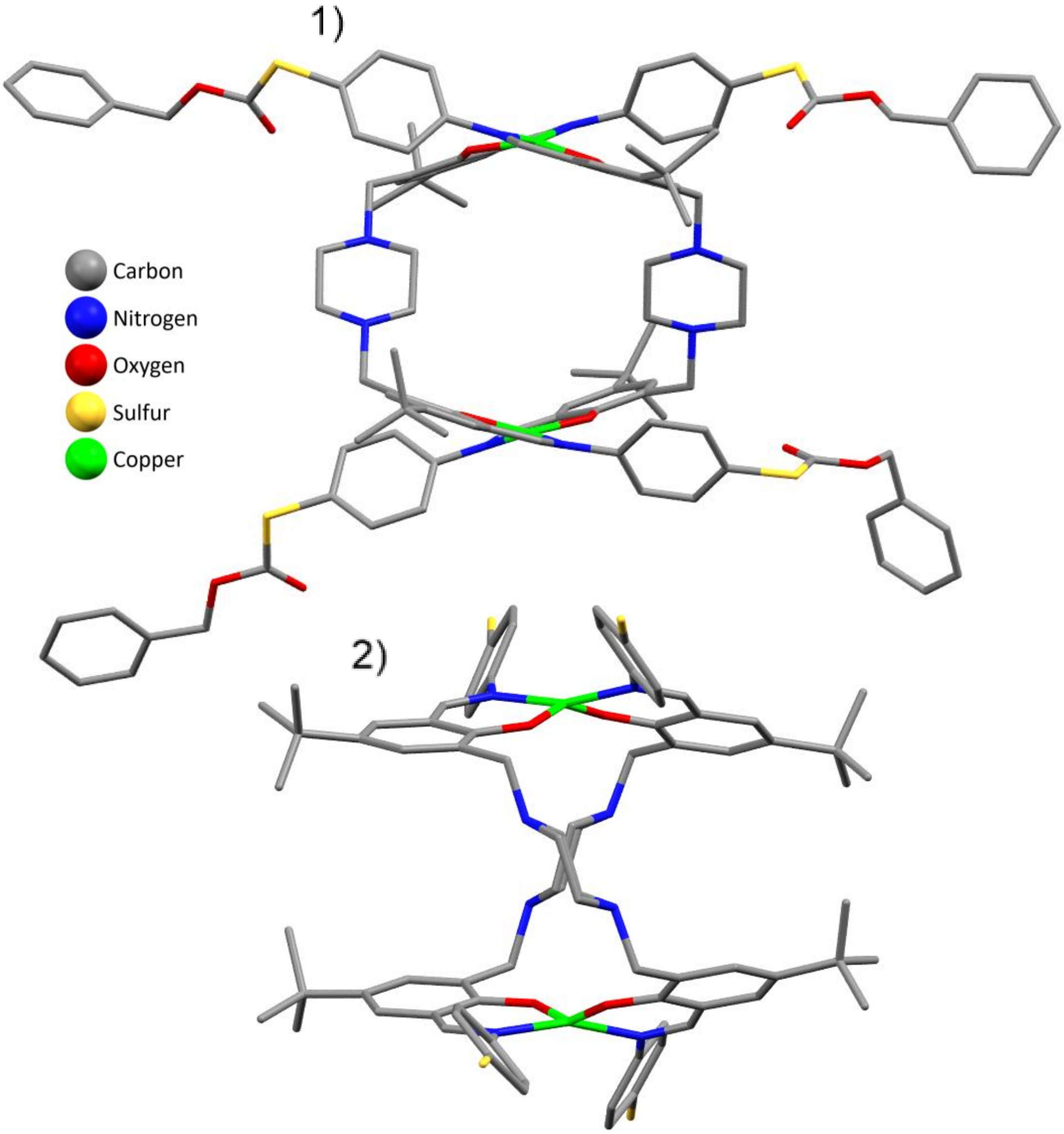


Figure 3.10: X-ray structure of the first cage targeting nitrate and phosphate with counterions and hydrogens omitted for clarity. 1) Shows the cavity in the full cage and 2) shows the metal binding sites and the angle of the ligands with the sulfur protecting groups omitted for clarity.

The crystals were grown by dissolving the product in DCM, then diethyl ether was added using a vapour diffusion technique to induce crystal growth. The X-ray data is low quality, but still show the structure of the cage clearly. The complex crystallised in a $P2_1/c$ space group and the asymmetric unit contains the complete two metal and two ligand complexes, with four charge balancing nitrate ions. The charge balancing nitrate ions are located outside the cavity, each interacting with a protonated amine in the piperazine linker. The cage in the asymmetric unit has a “M” helical twist and when the symmetry operations are applied the crystal structure reveals both enantiomers are present. The crystal structure revealed that the four piperazine nitrogens are protonated, with counter ions ionically bonding outside of the cage’s cavity. The fact that the counter ions are not inside the cavity of the cage is likely a result of how the cage acts in the solid state and does not determine that the cage will be able to function as intended in solution.[65]

The copper-oxygen bonds are 1.862(6) and 1.886(7) Å at one copper centre and 1.905(9) and 1.91(1) Å at the other copper centre. The copper-nitrogen bonds are 1.95(1) and 1.97(1) Å at one copper centre and 1.956(8) and 1.978(8) Å at the other copper centre. These bond lengths are close to similar molecules.[12,66] The metal binding site is a distorted square planar with a τ_4 value of 0.29 and 0.30 for the two binding sites.[67]

The cavity in the cage is 7.497(3) Å between the copper centres and 6.00(2) Å between the closest carbons in the piperazine linker, which is similarly sized to the calculated values. Compared to the calculated values with nitrate and phosphate anions coordinated to the cage have copper-copper distances of 11.463 and 7.448 Å respectively and the shortest distance between the carbons of the piperazine linkers are 3.484 and 7.102 Å respectively. This is promising for the potential sensoring properties of the cage. The other detail to note is the helical twist of the ligand, which can contract or extend to ideally encapsulate nitrate or phosphate anions.

The copper cage formed from ligand $\mathbf{1^A}$, packs so that the benzyl protecting group displays π- π interactions with the aniline ring of the imine as shown in Figure 3.11. The terminal benzyl group is interacting via edge-to-face π-stacking to the protected thiophenol rings and the phenolic benzenes of two separate cages.

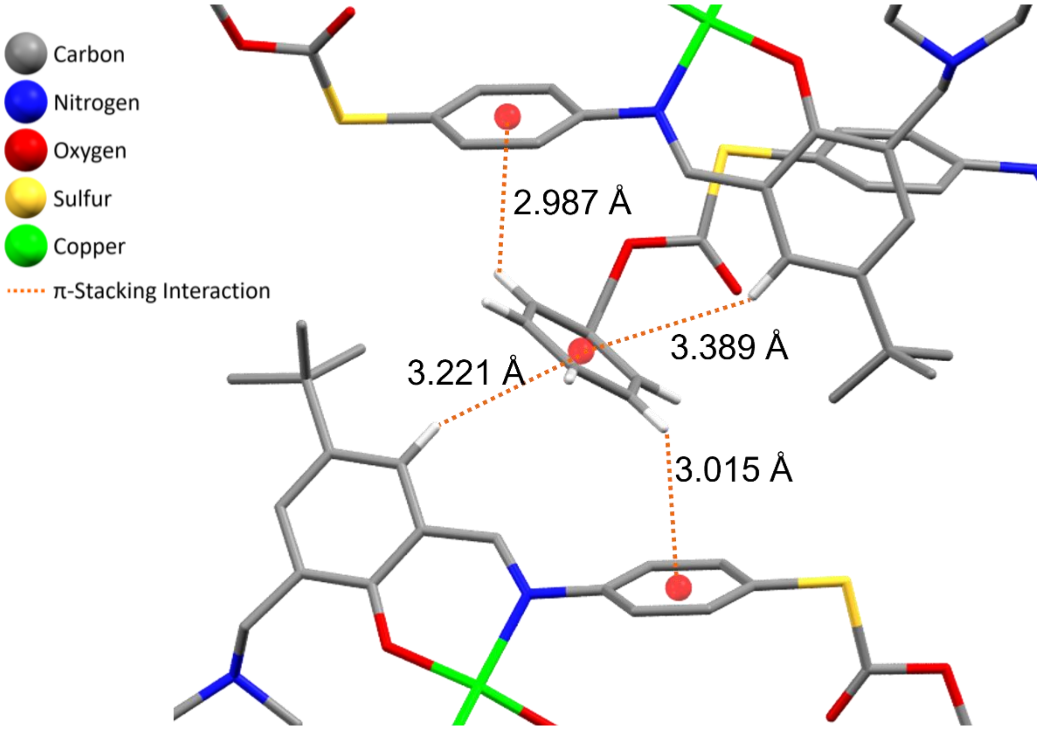


Figure 3.11: The interactions between aromatic rings of the copper cage using the ligand **1^A**.

We have yet to grow crystals of any cage formed from ligand **1^B** so other analytical techniques were required to try and confirm their successful synthesis. Unfortunately, these have been inconclusive. The mass spectrum of the complexes showed the full ligand, proving it hasn't decomposed upon exposure to the metal salts. The peaks in the IR spectra of ligand **1^B** and the complex [Cu_2(**1^B**)$_2$], as seen in Figure 3.12, are in the same positions with some slight broadening. This includes the imine (C=N) stretch peak, which is expected to shift upon complexation, but remained at 1616 cm^{-1}. The peak position suggests the ligand has not decomposed and the peak broadening suggests that the ligand has complexed successfully. These changes are consistent with the changes between the IR spectra of ligand **1^A** and the complex [Cu_2(**1^A**)$_2$], that has been confirmed via X-ray crystallography, seen in Figure 3.12.

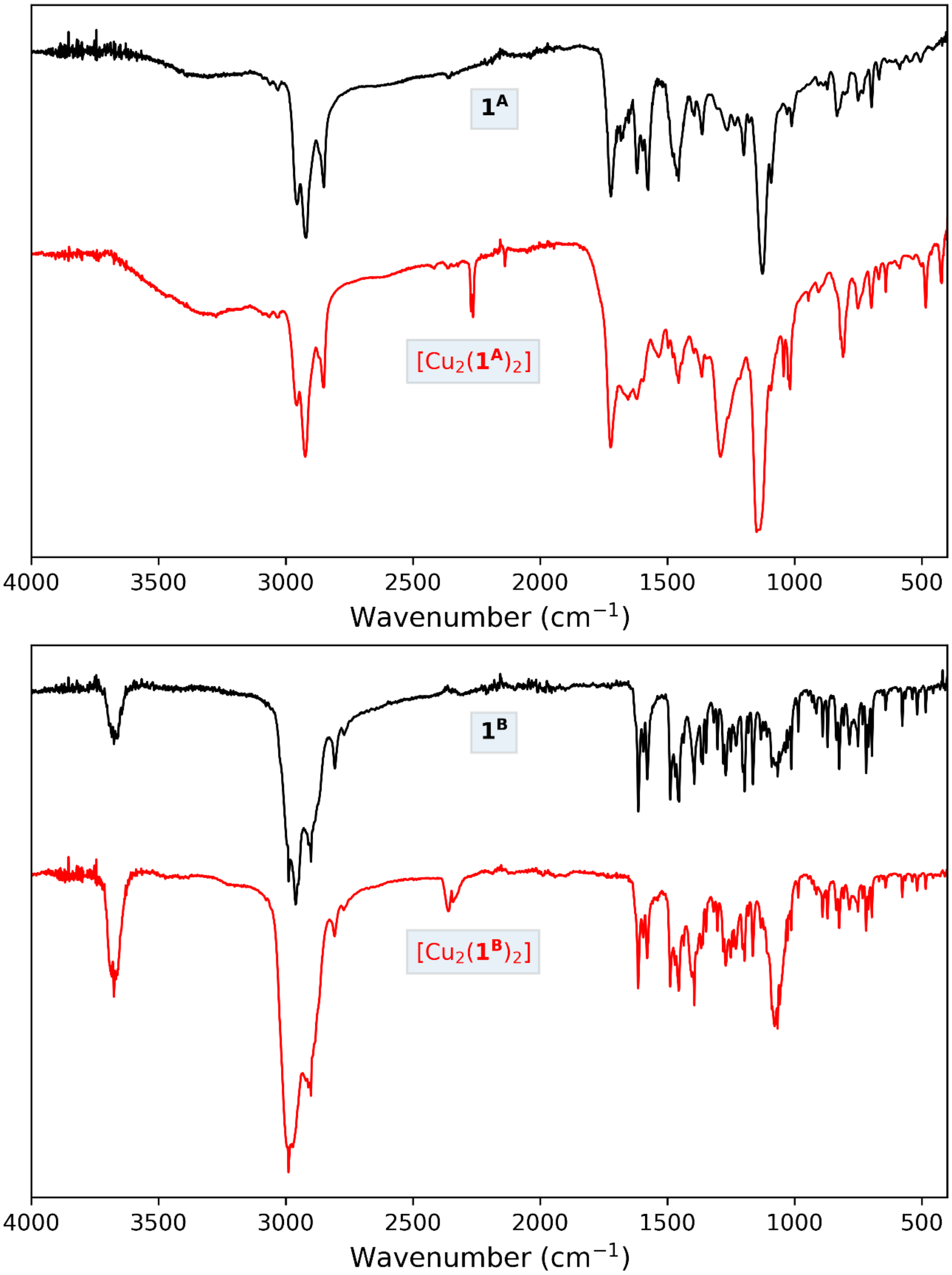


Figure 3.12: The IR of ligand **1**^{A} and complex [Cu$_2$(**1**^{A})$_2$] (top), and ligand **1**^{B} and complex [Cu$_2$(**1**^{B})$_2$] (bottom).

The sulfur group on the ends of the ligands are there to attach the cage to the surface of nanoparticles. This will hold the cages in a Raman "hot spot", enhancing the signal of the cage, which improves the potential sensitivity and limit of detection.

During the synthesis of the three cages, a brief proof of concept was undertaken to confirm that a portion of the cage could be detected at low concentrations using Raman spectroscopy. The nanoparticles that were intended to be used for the detection, were gold nanostars. In general, gold and silver are the most efficient elements for SERS because of their physical properties[68] and as gold nanoparticles are more robust to solvent conditions than silver nanoparticles, gold was the chosen metal. The reasoning behind choosing nanostars over other nanostructures is due to the increased “hot spot” concentration on the surface of the nanoparticles due to the points of the nanostars being close to each other.

The synthesis of the gold nanostars was a seed-mediated synthesis. The gold seed particles were synthesised by boiling chloroauric acid and citrate in water for 15 minutes. The citrate stabilised seed particles were generally in the size range of 11-13 nm in diameter based on the transmission electron microscopy (TEM) images of the seed particles in Figure 3.13.

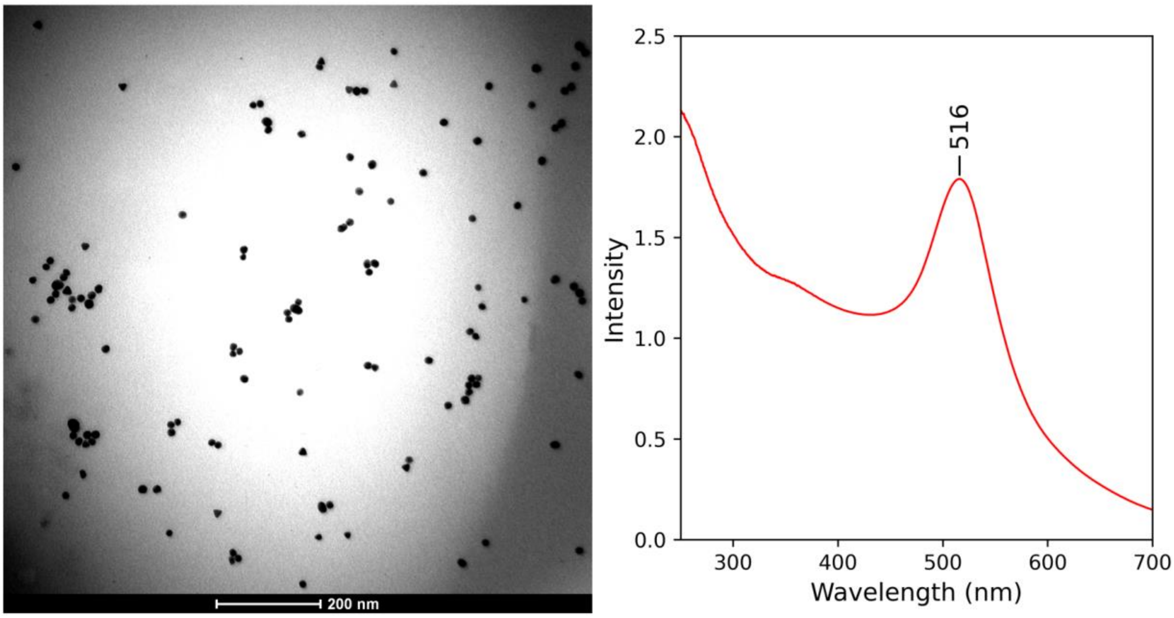


Figure 3.13: TEM (left) and UV/visible spectrum (right) of the gold seed particles.

Following the seed particle synthesis came the nanostar synthesis. This was attempted by mixing the seed particles, chloroauric acid and hydrochloric acid in water. A mixture of silver nitrate and ascorbic acid in water was then added and the solution was stirred for 30 seconds before it was centrifuged for 15 minutes. This method was based on work reported by Yuan *et al.*,[69] however after multiple attempts with minor changes, the nanoparticles would always come out of the centrifuge aggregated. The aggregation could have been caused by the seed particles reacting too quickly, leading to aggregation occurring before the reaction can be quenched.

An alternative gold nanostar synthesis was attempted that uses polyvinylpyrrolidone (PVP)-stabilised seed particles in the hope that this would prevent aggregation. To convert the citrate-stabilised seed particles to the PVP-stabilised seed particles, an aqueous solution of PVP was slowly added to the citrate-stabilised seed particles and the mixture stirred overnight. The solution was centrifuged for 90 minutes at varying speeds, to investigate the size distribution, before resuspending in ethanol. The results can be seen in the TEM images and UV/visible spectra in Figure 3.14 and Figure 3.15, which show that the first centrifugation at 4000 rpm caused the larger seed particles to form the pellet, leading to a larger size range. The second centrifugation at 5000 rpm had a more uniformed size distribution with trace amounts of abnormal particles. The final centrifugation at 7000 rpm has a higher density than the first two and a very uniform size of the seed particles. The last TEM image shows the nanoparticles left in solution after the third centrifugation and shows that few seed particles are left in the supernatant and like the final centrifugation, the seed particles are a uniform size.

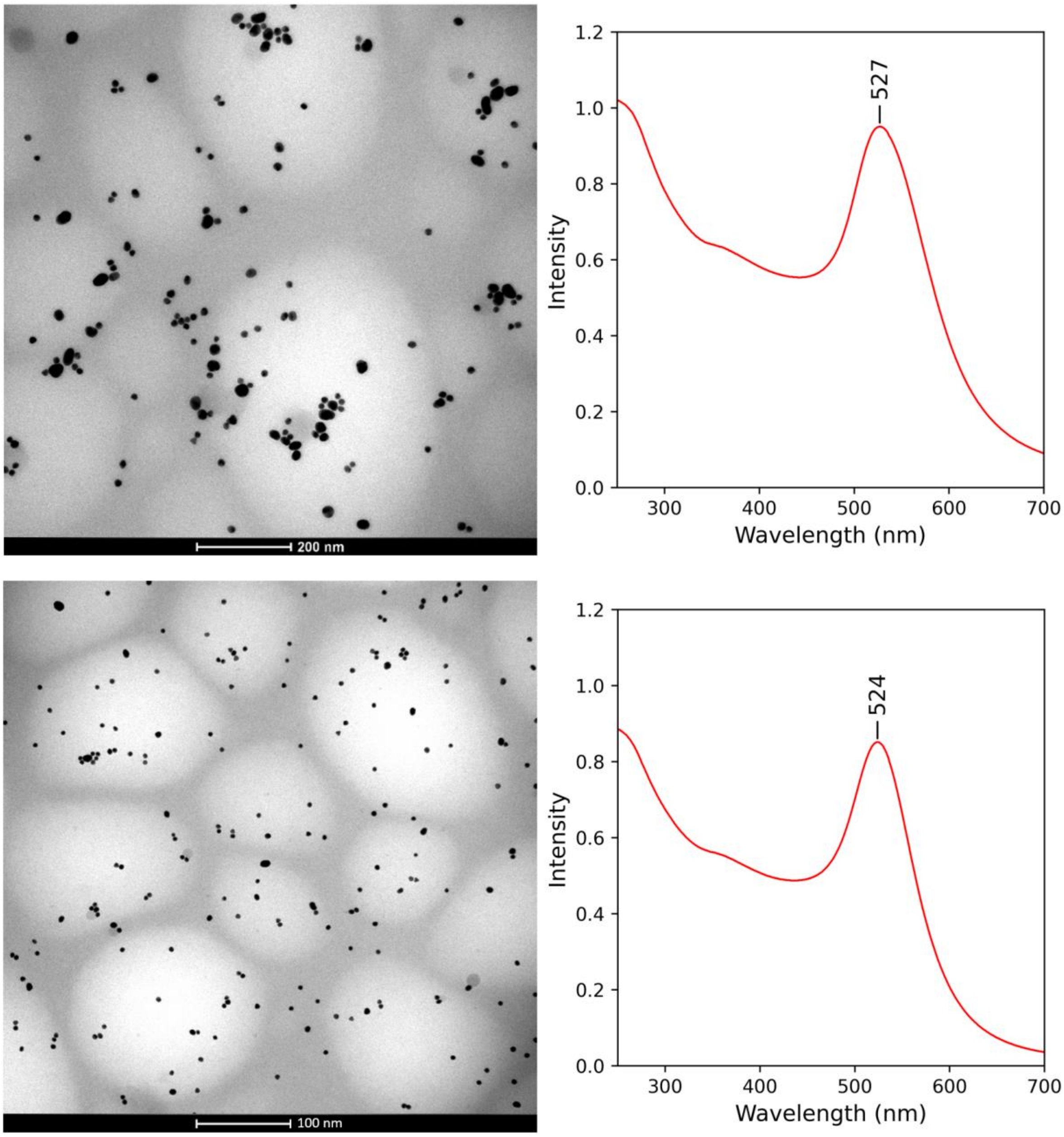


Figure 3.14: TEM (left) and UV/visible spectrum (right) of the silver nanoparticles spun at 4000 rpm (top) and 5000 rpm (bottom).

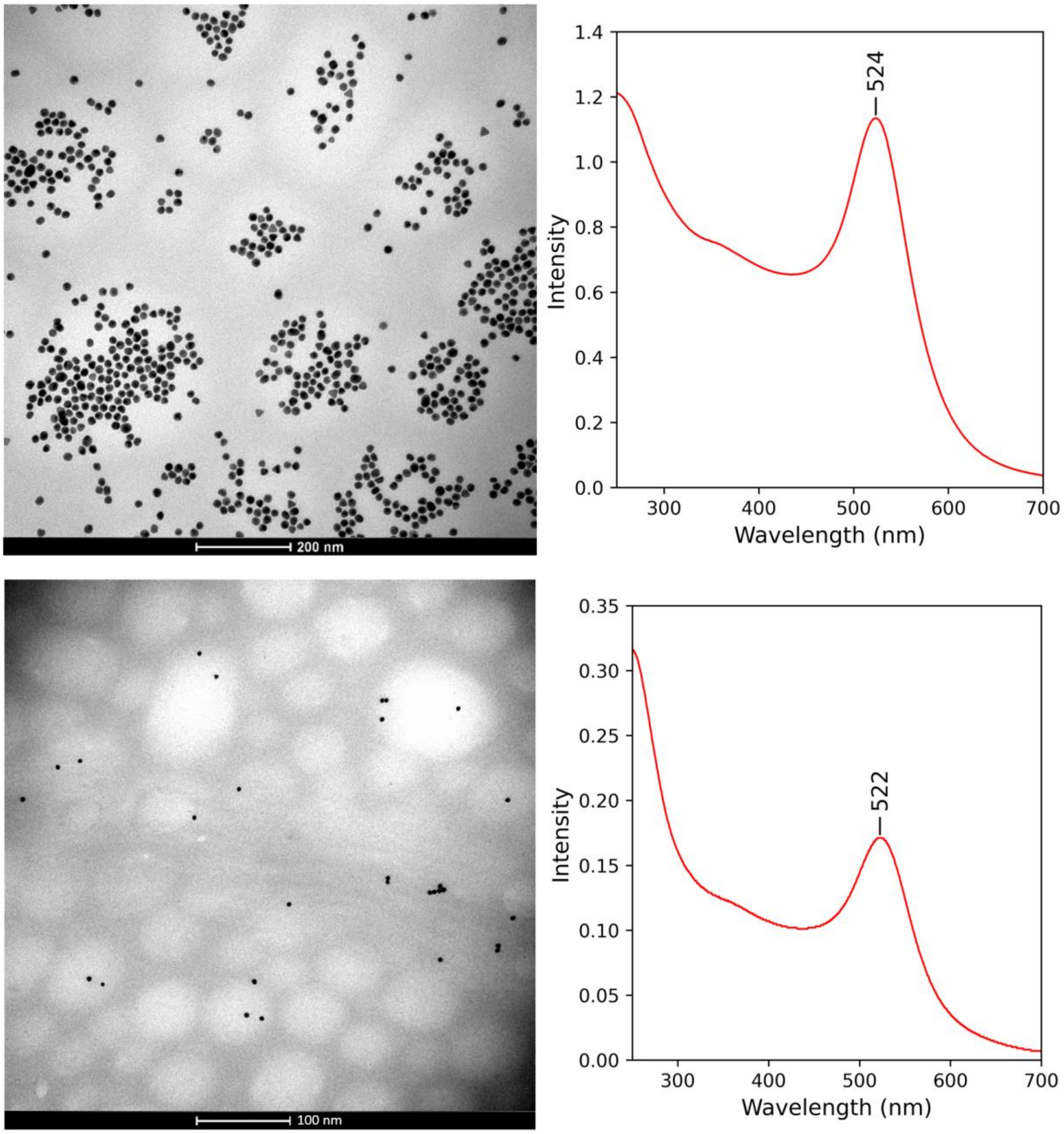


Figure 3.15: TEM (left) and UV/visible spectrum (right) of the silver nanoparticles spun at 7000 rpm (top) and the aqueous layer after all mentioned centrifugations (bottom).

For the gold nanostar synthesis using the PVP-stabilised seed particles, aqueous chloroauric acid and PVP in DMF were mixed together and the process monitored by UV/visible spectroscopy. After the peak at 325 nm disappeared, the seed particles are meant to be added and the solution was stirred until no changes are observed by UV/visible spectroscopy. The first time this was attempted, the band at 325 nm did not change significantly in the first hour of stirring, so then the seed particles were added. The UV/visible spectrum was continually monitored and the band at 325 nm diminished to a third of the original size and no other band was appearing after 5 days of stirring the solution.

Following this, the solution was sonicated, which caused the band at 325 nm to reduce further while no other peaks were forming. This was until the nanoparticles started aggregating after 6 hours of sonication. The UV/visible spectra can be seen in Figure 3.16.

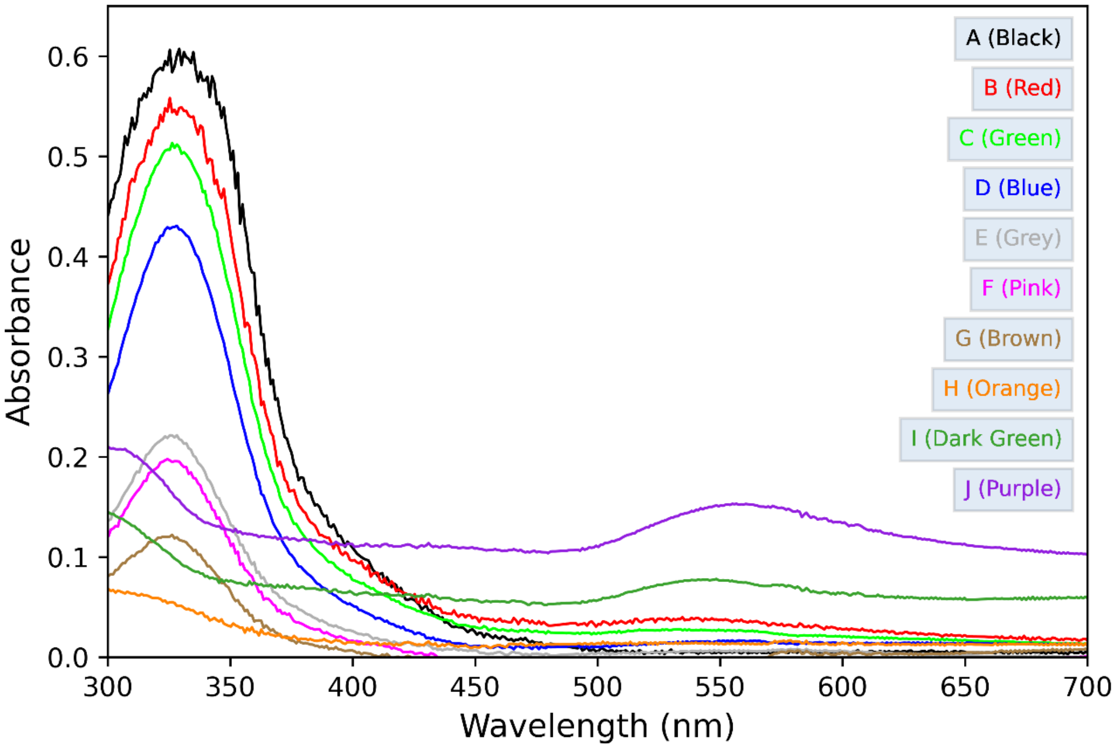


Figure 3.16: UV/visible spectrum of chloroauric acid and PVP initially (A), chloroauric acid and PVP after mixing for an hour, when the seed particles were added (B), the solution after stirring for 12 hours (C), 36 hours (D) and 5.5 days (E) and the solution after sonicating for 3 (F), 6 (G), 8 (H), 10 (I) and 12 hours (J).

A second attempt at the PVP-stabilised nanostars was conducted by mixing the aqueous chloroauric acid, PVP in DMF and the seed particles together. The solution was then subjected to sonication and the UV/visible spectrum was monitored, showing the band at 325 nm slowly decreasing, with no other peaks forming until the nanoparticles started aggregating as in the first attempt. The only difference being that the nanoparticles started to aggregate after 8 hours of sonication. The UV/visible spectra can be seen in Figure 3.17.

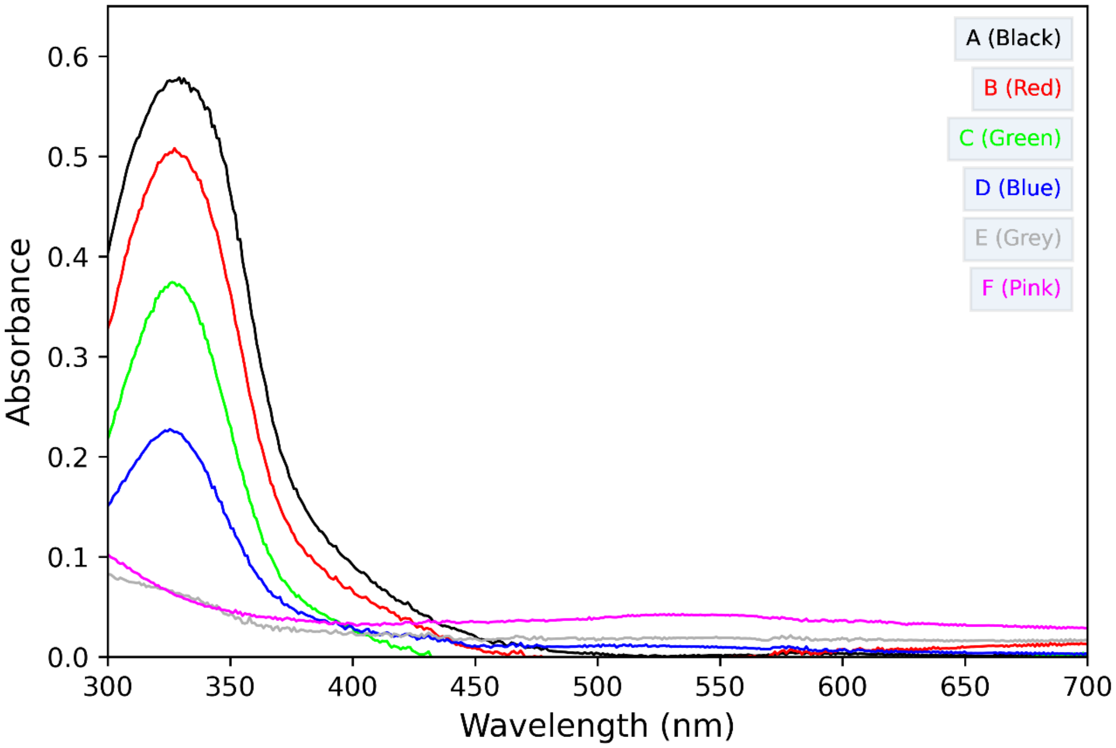


Figure 3.17: UV/visible spectrum of chloroauric acid, PVP and the seed particles initially (A) and the solution after sonicating for 2 (F), 6 (G), 8 (H), 10 (I) and 12 hours (J).

As the gold nanoparticle synthesis was problematic, silver nanoprisms were used for the proof of concept instead of the intended gold nanostars. This is because the synthesis of the silver nanoparticles was simple and well known. The synthesis is achieved by mixing sodium citrate, silver nitrate, hydrogen peroxide, potassium bromide and sodium borohydride together and waiting for 5 minutes for the nanoparticles to form. The UV/visible spectra of two samples of the silver nanoprisms can be seen in Figure 3.18.

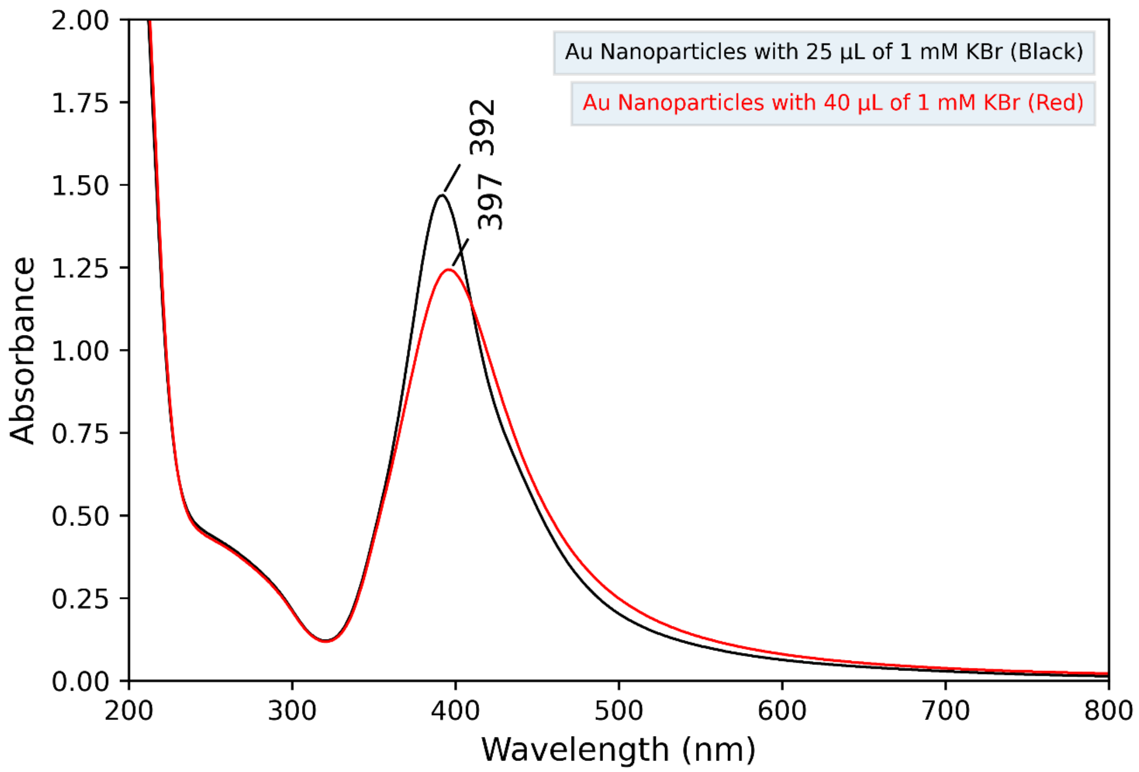


Figure 3.18: UV/visible spectrum of the silver nanoparticles synthesised with 25 and 40 µL of 1 mM potassium bromide solution.

During the synthesis of the nanoparticles, a few preliminary tests were conducted to confirm that a portion of the cage was detectable using Raman spectroscopy. All the Raman spectroscopy tests were performed using a 532 nm laser. Initially, three solution-based tests were performed with the gold seed particles and 4-nitrothiophenol (NTP). One of these tests had the seed particles mixed with a 10^{-5} M solution of NTP, another used a 10^{-8} M solution of NTP and the third used a 10^{-5} M solution of NTP combined with a small quantity of potassium chloride. The baselined Raman spectra, as seen in Figure 3.19, of the two solutions with the 10^{-5} M NTP showed noticeable peaks at 1080, 1330 and 1570 cm^{-1} corresponding to NTP, but the solution with the 10^{-8} M NTP had no noticeable peaks at these wavenumbers.

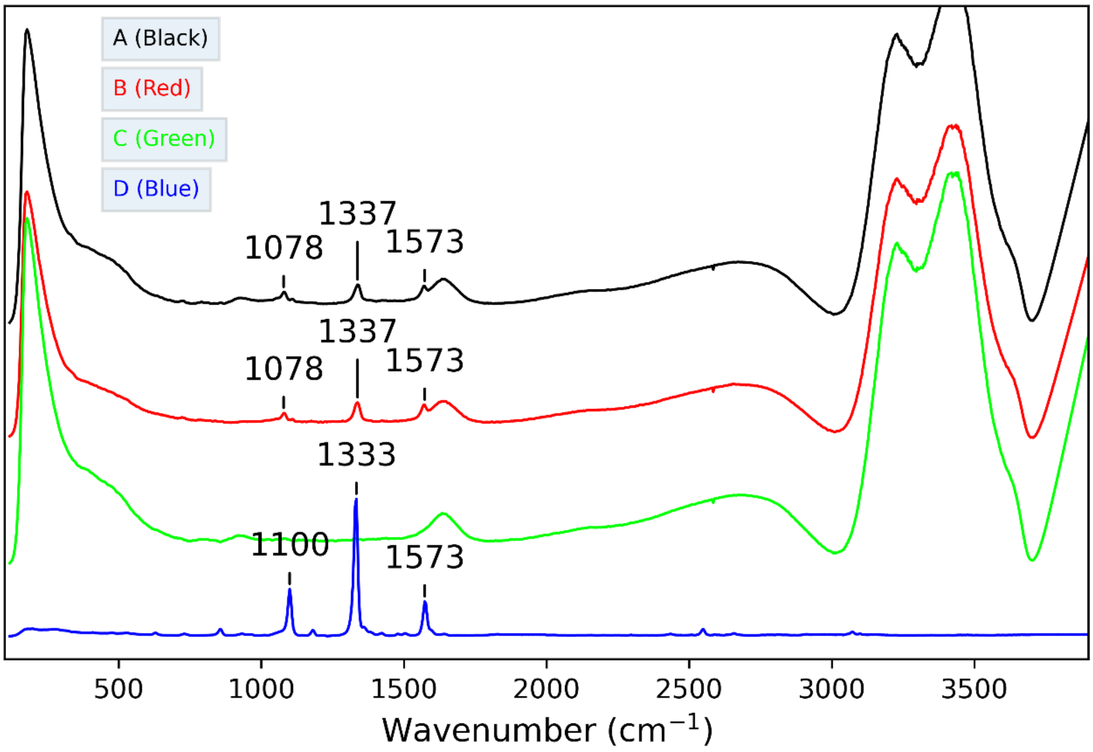


Figure 3.19: Raman spectrum of the silver nanoparticles and 10^{-5} M solution of NTP (A), with KBr (B), silver nanoparticles and 10^{-8} M solution of NTP (C) and solid NTP (D).

The fact that there were noticeable peaks corresponding to NTP at low concentrations, in solution was promising, as this suggests that minor changes in the Raman spectra of the cage could be detectable. The signal intensity can also be improved by enriching an aliquot of sample by evaporating the solvent. The samples solutions were enriched by using a slippery liquid-infused porous surface (SLIPS) to repel a drop of the sample solution, causing it to dry into a small spot of sample rather than a large spot of sample. This increases the concentration of “hot spots” and the number of NTP molecules in the “hot spots”, increasing the NTP signal intensity. The first set of tests was then looking into whether the ratio of seed particle solution to NTP solution made a difference to the spectra, when running a solid-state sample. The differences between the samples were not significant enough to favour an excess of seed particles or NTP so all further testing had an equal ratio of seed particles to NTP.

The next tests were looking into the detection limits of the system focusing on the peaks at 1080, 1330 and 1570 cm^{-1}. This was repeated three times; the first used an 80% pure NTP, as that is how it is stored, the second used purified NTP and the third had more data points. The first two tests used solutions of NTP at 10^{-5}, 10^{-6}, 10^{-7}, 10^{-8} and 10^{-9} M and the third test used solutions of NTP at 10^{-6}, $10^{-6.33}$, $10^{-6.67}$, 10^{-7}, $10^{-7.33}$, $10^{-7.67}$ and 10^{-8} M. In Figure 3.20, a standard curve can be seen using some of the points collected. Some of the data has not been added as those samples were not clean or did not have clear NTP peaks. However, looking at the standard curve of the three peaks, the preliminary results show that this could be a viable method of detection for the proposed system.

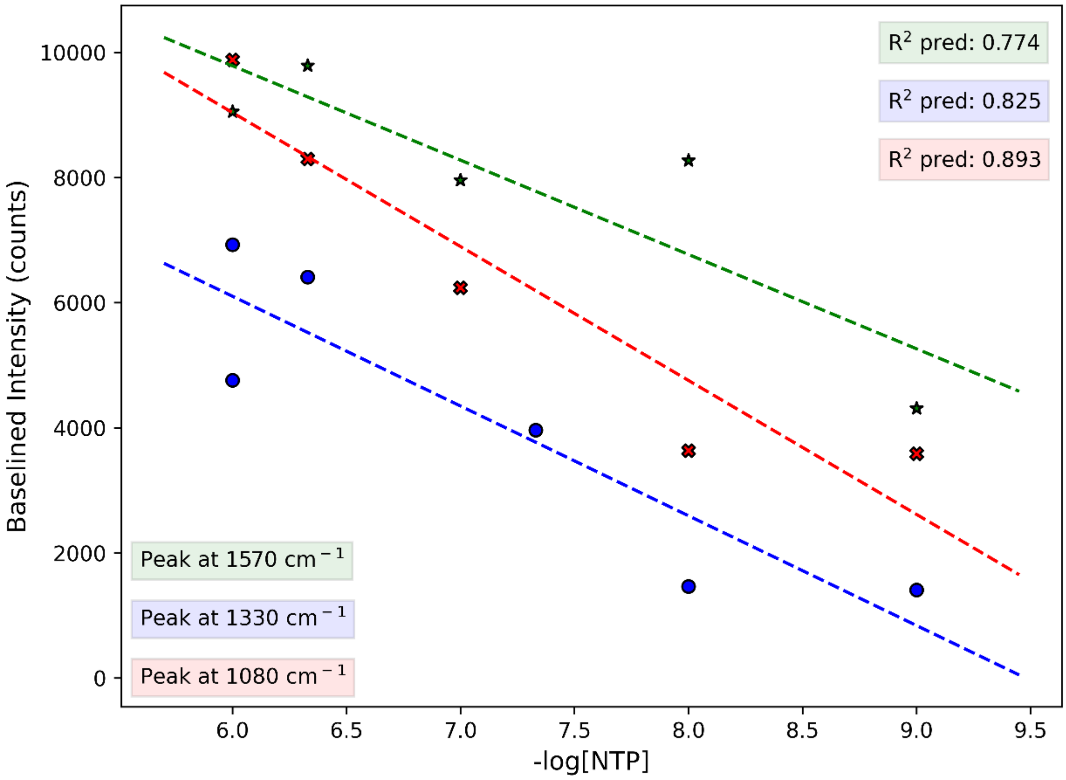


Figure 3.20: Standard curve of NTP and silver nanoparticles using Raman spectroscopy.

3.6 Conclusion

The overall progress toward the nitrate and phosphate sensing cages can be seen in Table 1.1, summarising the current state of the work performed towards the three cages.

Table 3.1: Overall progress to synthesising the cages to target nitrate and phosphate.

	Portion	**Percentage complete**
R^A-capped cage (cage 1)	Ligand synthesis	100% (5/5)
	Complexation and analysis	80%
R^B-capped cage (cage 2)	Ligand synthesis	100% (5/5)
	Complexation and analysis	50%
R^C-capped cage (cage 3)	Ligand synthesis	37.5% (3/8)
	Complexation and analysis	0%

The synthesis for the R^A and R^B-capped ligands has been completed while the R^C-capped ligand has multiple steps remaining to finish the synthesis. The remaining steps for the R^C-capped ligand follow the same procedures as the R^A and R^B-capped ligands, thus they should be simple reactions.

The complexation reaction of the R^A and R^B-capped ligands has been attempted and, in the case of the R^A-capped ligand, successful cage formation was confirmed by single crystal X-ray. Complexation of the R^B-capped ligand has not been confirmed; however, evidence suggests that the cage was formed successfully. Following the synthesis and analysis of the cages, they will be tested for their ability to encapsulate nitrate and phosphate. This testing will initially use UV/vis titration to study the binding limits or X-ray crystallography to analyse how well the analytes sit in the cavity.

Alongside the cage synthesis, a different method to synthesise the gold nanostar synthesis should be undertaken. After the cages and gold nanostars have been synthesised and analysed, the binding strength of the cage to the nanostars can be tested using Raman spectroscopy. If the binding strength is low, the sulfur groups of the ligands can be deprotected to increase the binding strength.

Chapter 4 – Glyphosate

4.1 History of Glyphosate

Glyphosate or [N-(phosphonomethyl)glycine] is an organophosphorus compound comprised of a methyl phosphate bound to the nitrogen of glycine as seen in Figure 4.1. Glyphosate is an odourless white crystalline solid with one basic amino group and three ionisable acidic sites.[70,71]

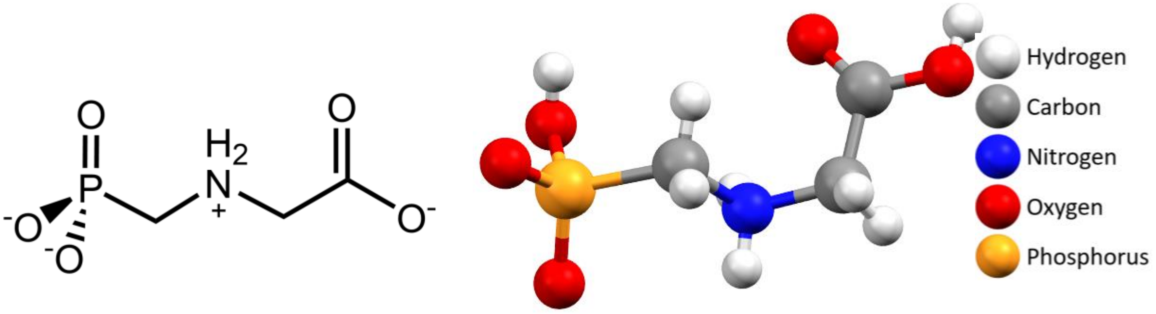
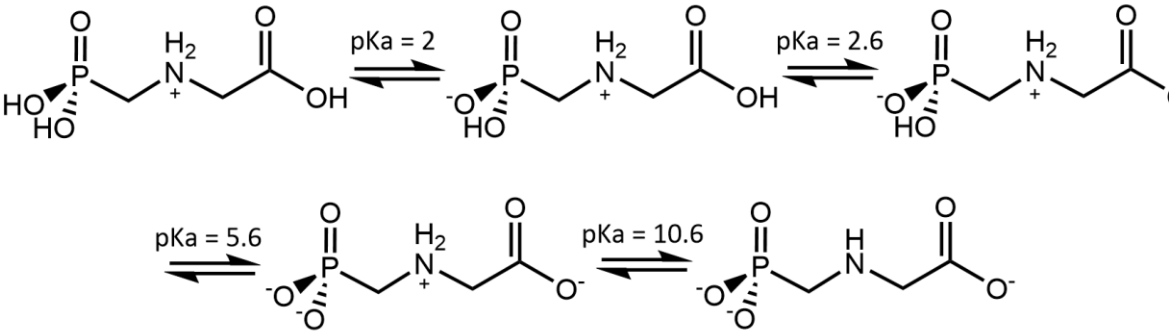


Figure 4.1: A skeletal view of glyphosate at pH 7 (left) and a solid-state X-ray structure of glyphosate with element key (right).[72]

The various ionisation states of glyphosate can be seen in Scheme 4.1 along with the pKa values. Glyphosate mainly found as monoanion or dianion depending on the pH of the solution. Glyphosate is the most widely used broad-spectrum herbicide, which acts non selectively. A well-known product that contains glyphosate is "Roundup^{TM}", a post emergence, broad spectrum herbicide, which consists of the isopropylamine salt of glyphosate and a polyoxyethylene tallow amine (POEA) as the surfactant.[26]



Scheme 4.1: The various ionisation states of glyphosate.[71]

In 1950, glyphosate was invented by Dr Henri Martin, a Swiss chemist working for a small pharmaceutical company, called Cilag. As glyphosate had no pharmaceutical application, it wasn't of any use to Dr Martin. Many years later Dr Phil Hamm with the help of Dr. John Franz, who worked as a chemist for Monsanto, rediscovered glyphosate in 1970, and glyphosate rapidly went through the herbicidal testing systems and was commercially released as Roundup^{TM} in 1974.[70,73]

Upon release, glyphosate was used to kill perennial weeds before the planting of crops or carefully sprayed on well-established orchards to control the vegetation on the ground, avoiding any foliage of the crops. Glyphosate is also used in non-crop situations like roadsides and railway tracks due to its effectiveness in killing most plants.[74]

Glyphosate is very effective due to its novel mode of action. Glyphosate inhibits the enzyme, 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS), which is only found in plants and microorganisms.[75] EPSPS catalyses the sixth of seven reactions of the shikimate pathway (Figure 4.2). The shikimate pathway is a metabolic pathway that produces an intermediate product (chorismite) used in the synthesis of essential aromatic amino acids (tyrosine, tryptophan and phenylalanine).[76]

Glyphosate acts as a transition state analogue of phosphoenolpyruvate reducing the activity of EPSPS. The low levels of EPSPS activity lead to a deficiency in aromatic amino acids and this in turn affects various metabolic functions of the plants. Plants treated with glyphosate are also more susceptible to diseases and plant pathogens as the plants immune response is weakened, further increasing the effectiveness of glyphosate.[15,77] Another way glyphosate acts to kill the plant is by reducing the feedback inhibition of the shikimate pathway, resulting in a lot of the carbon being converted into shikimate 3-phosphate. The final reaction to shikimate 3-phosphate also requires adenosine triphosphate (ATP). This causes carbon and ATP to be drawn into the shikimate pathway from other metabolic pathways, causing shikimate 3-phosphate to rapidly accumulate and the available levels of carbon and ATP to drop.[73]

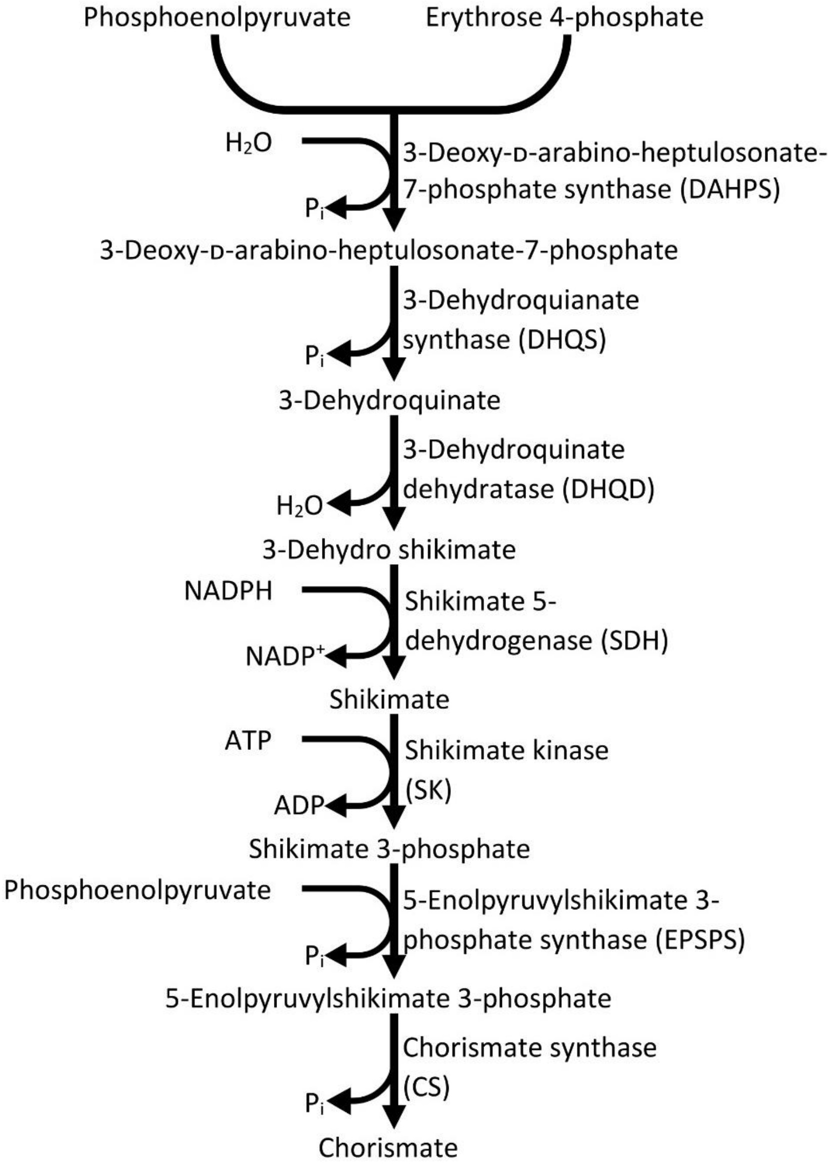


Figure 4.2: The shikimate pathway.

Glyphosate is absorbed either through the roots along with vitamins and minerals or through the leaves as glyphosate is very water soluble.[78] As glyphosate is highly charged it can only be absorbed by the leaves through the lipid bilayer via aqueous pores, slowing the absorption into the plant.[79] To help glyphosate to be absorbed, surfactants are used as emulsifiers and as wetting agents to keep glyphosate hydrated and dispersed on the plant surface. This works by increasing the spray droplet retention, increasing the permeability of glyphosate through the waxy covering on plant cuticles.[80-82]

Glyphosate quickly gained popularity in mainstream agriculture (Figure 4.3) due to multiple factors: glyphosate resistant crops becoming deregulated for farmers, the price of glyphosate decreasing as a result of the original patent expiring and new suppliers have emerged on the market and a surge in glyphosate-resistant weeds.[71]

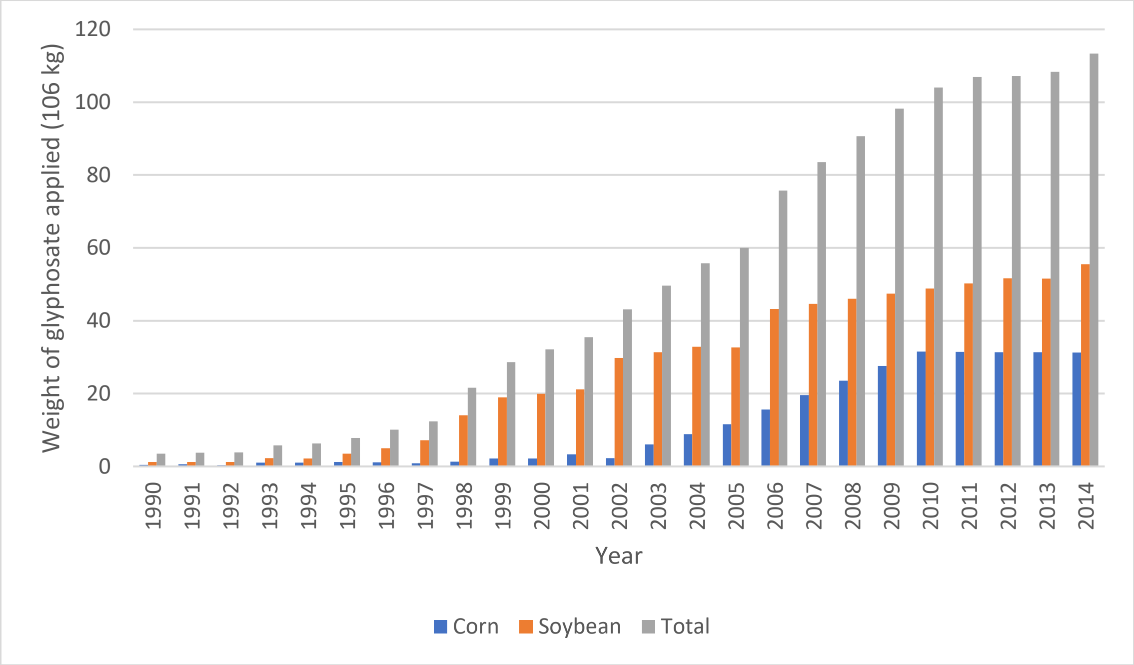


Figure 4.3: Use of glyphosate on different crops in the USA.[83]

Glyphosate resistant crops becoming deregulated allowed farmers to spray glyphosate on their crops while they were still growing as well as before planting and after harvesting to kill the weeds, with no impact on their crops. The first of the glyphosate resistant crops deregulated in the USA were soybean and canola in 1996, followed by cotton and maize in 1997 and 1998 respectively.[83]

Due to its effectiveness, farmers with glyphosate resistant crops have been mainly using glyphosate to kill all the weeds, leading to a genetic bottleneck for the weeds. This has led to the emergence of glyphosate resistant weeds and due to the widespread application of glyphosate, the number of glyphosate resistant weeds is rapidly increasing (Figure 4.4). This has in turn, reduced the effectiveness of glyphosate, causing farmers to apply excessive amounts of glyphosate to ensure the resistant weeds still die.[84]

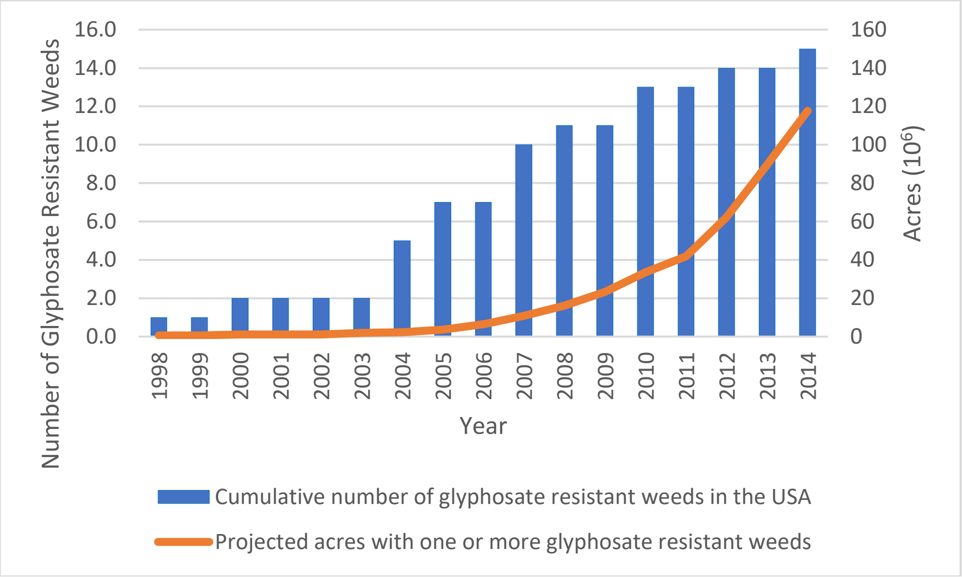


Figure 4.4: Emergence and spread of glyphosate resistant weeds.[83]

4.2 Reasons to Detect Glyphosate

The reported half-life of glyphosate is highly variable, as many factors affect how quickly glyphosate breaks down. The three main factors are the soil temperature, soil moisture and the fungi and/or bacteria that are present in the soil. Glyphosate is generally broken down by various soil microorganisms (fungi and bacteria) through either C-N bond cleavage, forming aminomethylphosphonic acid (AMPA) and glyoxylate or through C-P bond cleavage, forming sarcosine and phosphate.[85,86] Glyoxylate, sarcosine and phosphate are all readily used in biochemical processes so these are harmless molecules. Glyoxylate can be used in the tricarboxylic acid cycle or broken down further to glycine. Sarcosine can easily be converted to glycine and formaldehyde, which like phosphate and glyoxylate, can be utilised in many different areas of the cell and therefore, are non-toxic. AMPA on the other hand is more toxic than glyphosate and more resistant to degradation. This is shown in the molecules’ reported half-life in soil, which is 1 to 197 days for glyphosate[87,88] and 23 to 958 days for AMPA.[88]

Glyphosate was considered safe in terms of mammalian toxicity for many years because animals do not have the EPSPS enzyme, so the main activity of glyphosate is void in mammals.[89] A review by Williams *et al.*,[90] published in 2000, details the potential safety risks glyphosate poses to humans. Williams *et al.* found that there was no potential for glyphosate and AMPA to pose health risks to humans, whether it is acute or prolonged exposure. However, recently the overuse of glyphosate and its different formulations (Roundup, Bioforce, Grands Travaux, etc.) have been shown to poses a threat to aquatic and terrestrial organisms.[83] Other more recent papers have reported that glyphosate has toxic effects at very low levels over a prolonged exposure.[91,92] Another paper published in 2022 discusses how glyphosate has various types of neurotoxic effects in many different animals including humans.[93] For more details on the toxic effects, Gill *et al.*[15] goes into detail about the effect of glyphosate on multiple groups of organisms, including humans.

Due to the recent studies concluding that glyphosate poses a risk to humans,[15,93] there has been an increased interest in detecting glyphosate and AMPA. There is little interest in detecting any other derivatives as they are short lived or common organic products, like glycine, which present no toxicity issues. Regardless of the interest, there are big challenges to overcome to detect glyphosate. These include glyphosate’s similarity to amino acids, ionic character, low volatility, low molecular mass, lack of chromophore/fluorophore, high polarity, high water solubility, poor organic solubility, high absorption to soil and glyphosate’s ability to complex to metals.[94-97]

4.3 Current Methods for Detecting Glyphosate

Multiple articles have been published recently reviewing the detection techniques of glyphosate.[26,27,78,94,98] Each review describes multiple techniques including liquid chromatography (LC), high pressure liquid chromatography (HPLC), ion chromatography (IC), gas chromatography (GC), capillary electrophoresis (CE), spectrophotometric analysis, immunogenic analysis, nuclear magnetic resonance (NMR) and electrochemical techniques. Many chromatographic techniques are generally coupled to mass spectrometry or require derivatization of glyphosate to be detectable (often with UV/visible spectroscopy for detection). LC and HPLC are fast, sensitive, and repeatable, whilst requiring large expensive equipment. IC is simple, rapid, reliable, and requires inexpensive equipment, but IC is insensitive with high detection limits. GC is highly sensitive and selective, yet time consuming, laborious, requires a large amount of sample manipulation and specialised high-end equipment. CE is cheap and quick, but prone to interferences from high levels of salt. CE is also not a very commonly reported technique. Spectrophotometric techniques struggle due to the lack of a fluorophore or chromophore in glyphosate. Immunogenic analyses, such as enzyme linked immune-sorbent assays (ELISA), are very sensitive with good resolution, however they are tedious and labor-intensive with a short signal stability. NMR is a very quick technique that struggles with multiple interferants. Electrochemical techniques are variable, simple, fast and are prone to less interferants, but require extensive calibration.[26,27,78,94,98]

As time has gone by, there have been more detection methods published, using unconventional techniques. One such method is a paper-based chemosensor array device reported by Zhang *et al.*,[99] that uses the competitive binding activity of zinc (II) ions to oxyanions and catechol dyes (alizarin red S, bromopyrogallol red, pyrogallol red, and pyrocatechol violet as seen in Figure 4.5). Initially, the zinc ions are coordinated to the catechol dyes, but when presented with oxyanions, like glyphosate, a portion of the zinc ions dissociate from the dyes and coordinate to the oxyanions. This produces a measurable colour change, which, using automatic image processing software to characterise the fingerprint-like response patterns, corresponds to the oxyanion concentration and structure. However, this method lacks accuracy when presented with samples containing more zinc or when a multitude of different molecules are present (environmental samples).

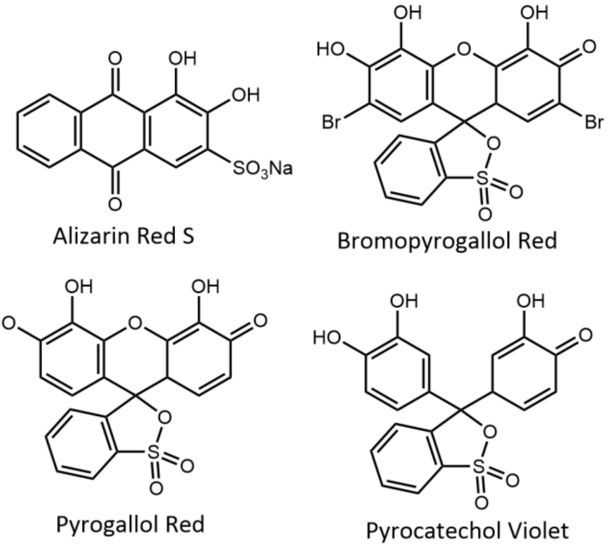
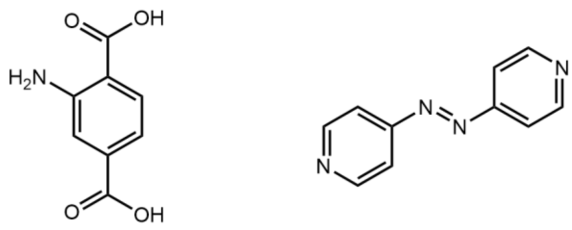


Figure 4.5: Diagram of the catechol dyes used by Zhang *et al.*[99]

Two recently published methods use simple metal-glyphosate interactions for detection.[100,101] One method uses a carbon dot-iron(III) system as a turn-on fluorescent sensor. The iron cations quench the carbon dot fluorescence, however glyphosate coordinates to the iron cations through the phosphate or carboxylate functional group. This leads to more carbon dot fluorescence, as there are fewer iron cations available to quench the emission.[100] The other method monitors the activity of a gold-platinum (Au-Pt) nanozyme, which oxidises 3,3′,5,5′-tetramethylbenzidine (TMB) to accurately measure glyphosate levels. L-Cysteine is also present, to inhibit the nanozyme activity, along with copper(II) ions. The glyphosate in a sample coordinates to the copper ions, which then coordinate to the L-cysteine molecules. This causes the activity of the Au-Pt nanozyme to increase, due to the reduction of available L-cysteine to inhibit it.[101] These methods are able to detect very low levels of glyphosate at 9 and 5 μg L^{-1} respectively, but they do not show any specificity. Any interferent molecule that can coordinate to a metal will give a positive result, making these methods unusable for many studies.

Other methods have used nanoparticles to interact with glyphosate causing the nanoparticles to aggregate.[102,103] In one report, the level of the quenching of fluorescence from carbon nanodots by silver nanoparticles quantifies the glyphosate levels in a sample. The silver nanoparticles are positively charged and at pH 4, glyphosate has an overall negative charge, causing ionic interactions between the nanoparticles and glyphosate. Once the nanoparticles and glyphosate have aggregated, the fluorescence from the carbon nanodots is detectable.[102] Guo *et al.*[103] describes a similar procedure in their paper, using the thioglycolate on the cadmium-tellurium quantum dots (QDs) to interact ionically with the cysteamine on the gold nanoparticles. While the QDs are interacting with the nanoparticles, fluorescence resonance energy transfer (FRET) occurs, quenching the QD fluorescence. Once glyphosate is added, the nanoparticles aggregate and the QD emission becomes detectable. As with the metal-glyphosate interaction methods, these methods will lose their effectiveness in the presence of other molecules capable of aggregating nanoparticles.

New methods are continually being developed, for example, Wiwasuku *et al*.[104] reported the use of a metal organic framework (MOF) to detect chromium(III) and glyphosate. This MOF is made of cadmium(II), 2-amino-1,4-benzenedicarboxylic acid and 4,4′-azopyridine (Figure 4.6) in dimethylformamide (DMF). The azopyridine is used as a fluorescent probe that is quenched by the free amine of 2-amino-1,4-benzenedicarboxylic acid by a photo-induced electron transfer (PET) process. In the presence of chromium ions, the MOF starts to breakdown due to the competitive binding of the metals to the MOF ligands. This frees a portion of the azopyridine, which is now able to fluoresce and this fluorescence becomes detectable to the naked eye. A similar process happens in the presence of glyphosate, as glyphosate competitively coordinates to the cadmium ions displacing the ligands that form the MOF. Again, this breaks the MOF down releasing a portion of the azopyridine. Wiwasuku *et al.* looked into whether other pesticides (that do not contain metal coordinating sites) and cationic metals (including calcium(II), magnesium(II) and zinc(II)) are interferants and found that this MOF is highly specific to glyphosate and chromium ions compared to other herbicides and metals. However, this selectivity for glyphosate is due to the other herbicides not being able to coordinate to cadmium strongly. This means that a strong chelating ligand, like picolinic acid (2-pyridinecarboxylic acid) a natural product made in the kynurenine pathway (the major tryptophan metabolism pathway),[105] will have the same effect on the MOF as glyphosate.



2-amino-1,4-benzenedicarboxylic acid 4,4'-azopyridine

Figure 4.6: Structures of the two ligands used in the MOF made by Wiwasuku *et al.*[104]

Currently, the methods to accurately test glyphosate levels in environmental samples require prior purification, derivatization, expensive equipment or have a complex methodology. Some methods may be able to detect glyphosate in the environment but not to the low levels that will be found in the samples, which will be between 0.5 to 500 μgL^{-1}.[106]

4.4 Target Cages and Design Features

There are two target cages designed to encapsulate glyphosate in this project, as shown in Figure 4.7. One is a bimetallic system while the other is a monometallic system. Both ligands are designed to bind one or two coppers centres and contain amine oxide groups, hydrogen bonding functional groups and the potential for functionalization at the copper binding terminal/s. Due to time constraints, the synthesis of these cage’s has not been completed.

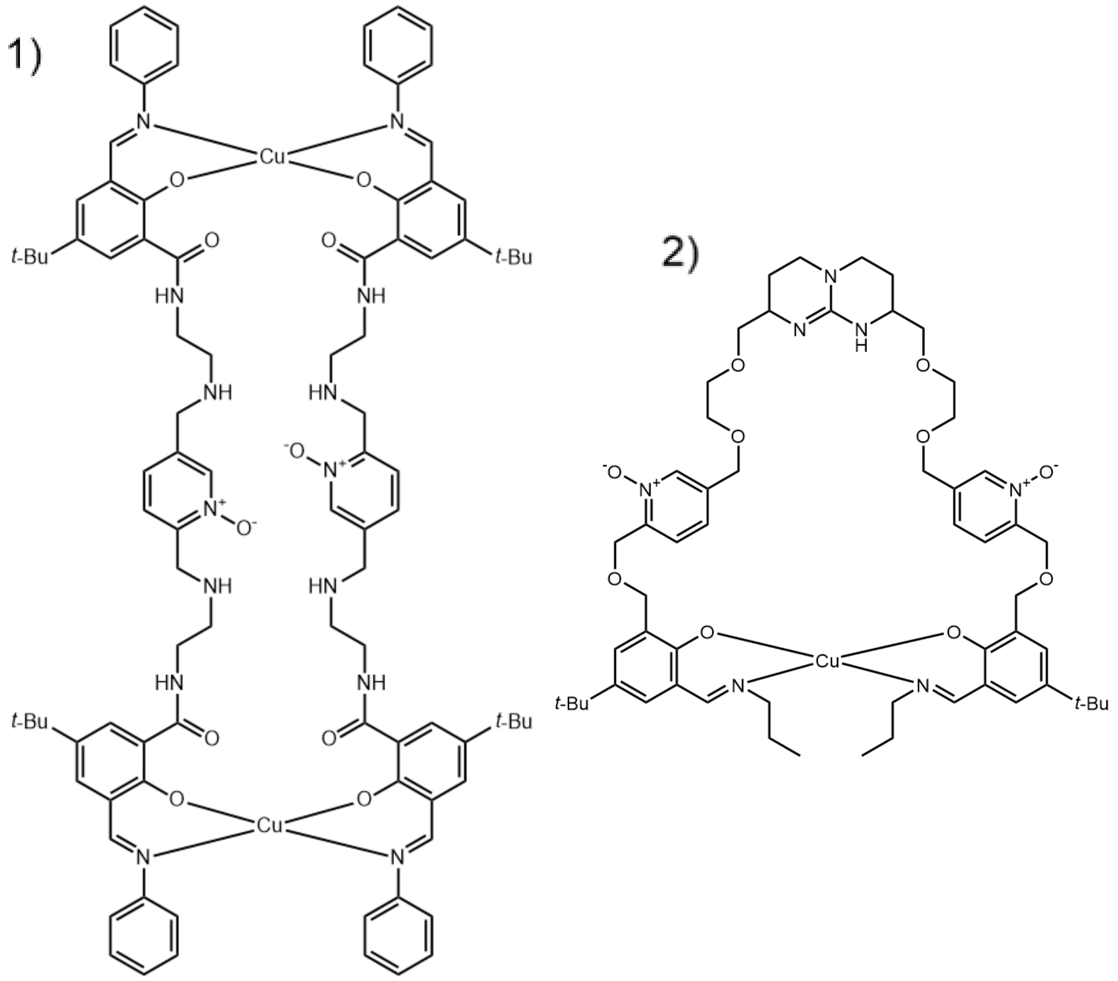


Figure 4.7: The main target cages to capture glyphosate. 1) is a bimetallic system and 2) is a monometallic system.

Copper was the metal chosen for these cages because it has the ability to switch between square planar, square pyramidal and octahedral coordination modes depending on the need of the complex, allowing the glyphosate terminals to coordinate to the metal centres. The oxime metal binding sites were then chosen because copper binds strongly to the oxime group. The imine group of the metal binding site allows for easy derivatization. This is because the imine is formed from a condensation reaction between an aldehyde and amine, allowing the terminals to be varied easily, leading to the ability to use various detection techniques. The ethylenediamine/ether linkers serve multiple roles as they determine the cavity size and they act as hydrogen bond acceptors. The ethylenediamine linkers can also act as hydrogen bond donors. The *N*-oxide groups are positioned in the middle of the complex to allow an electrostatic attraction between the ammonium cation in the middle of glyphosate and the cage. The final design feature of the monometallic cage is the guanidinium binding group. This was added because Neal *et al.*[107] found that the guanidinium moiety binds phosphate strongly and selectively.

The binding of glyphosate in the two cage systems has been modelled using a simple geometry optimising calculation, which performs an iterative series of single point calculations until it reaches an energy minimum. This is visualised in Figure 4.8, and displays how glyphosate could interact with the mono- or bimetallic systems. In both systems, the *N*-oxide groups are directed at the positively charged ammonium ion to form non-covalent interactions and the carboxylic acid of glyphosate is directed at a copper centre. In the monometallic system, the phosphate is interacting with the hydrogens of the guanidinium group and in the bimetallic system, the phosphate is directed at the second copper centre.

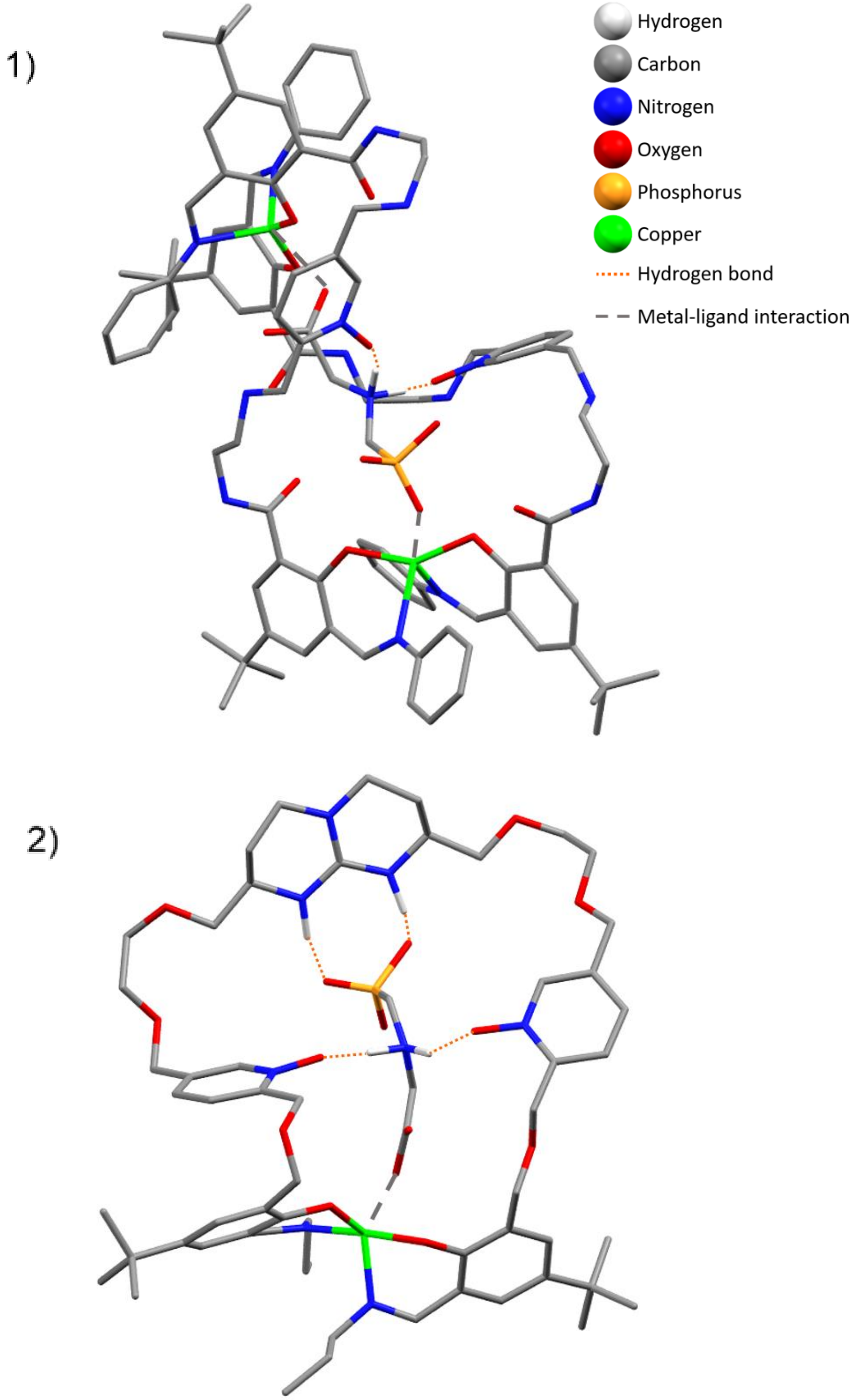
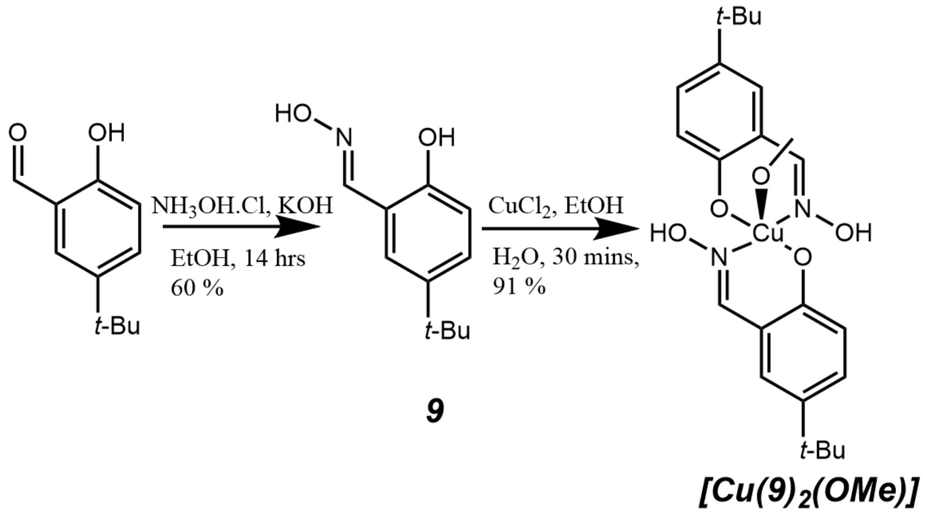


Figure 4.8: The bimetallic (1) and monometallic (2) target cages interacting with glyphosate. The non-interacting hydrogens are removed for clarity.

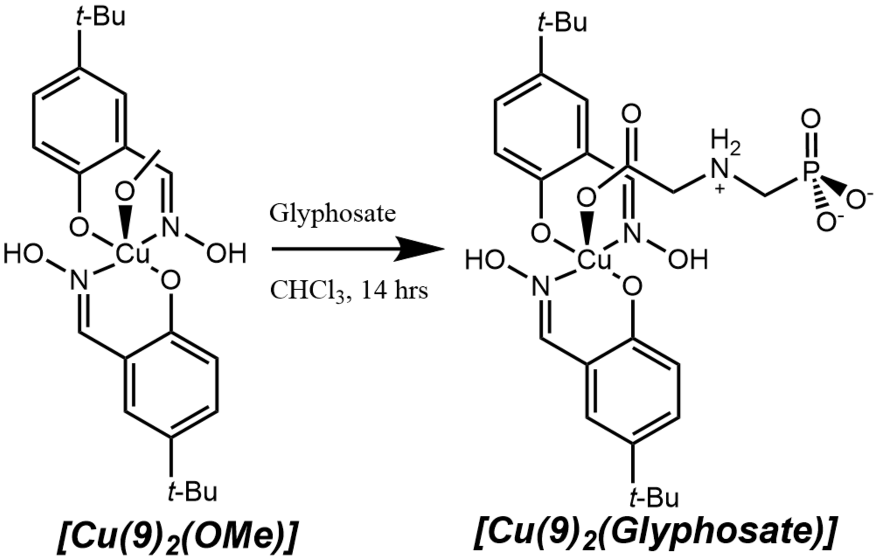
4.5 Synthesis and Analysis

The first target was a simple test system that would show if glyphosate and the coordinate copper could interact. Complex [Cu(**9**)$_2$(OMe)] was prepared in two steps, as seen in Scheme 4.2, by reacting 5-(*tert*-butyl)-2-hydroxybenzaldehyde in ethanol with hydroxylammonium chloride and potassium hydroxide to give 5-(*tert*-butyl)-2-hydroxybenzaldehyde oxime (**9**).



Scheme 4.2: The synthetic method of the copper complex, [Cu(**9**)$_2$(OMe)].

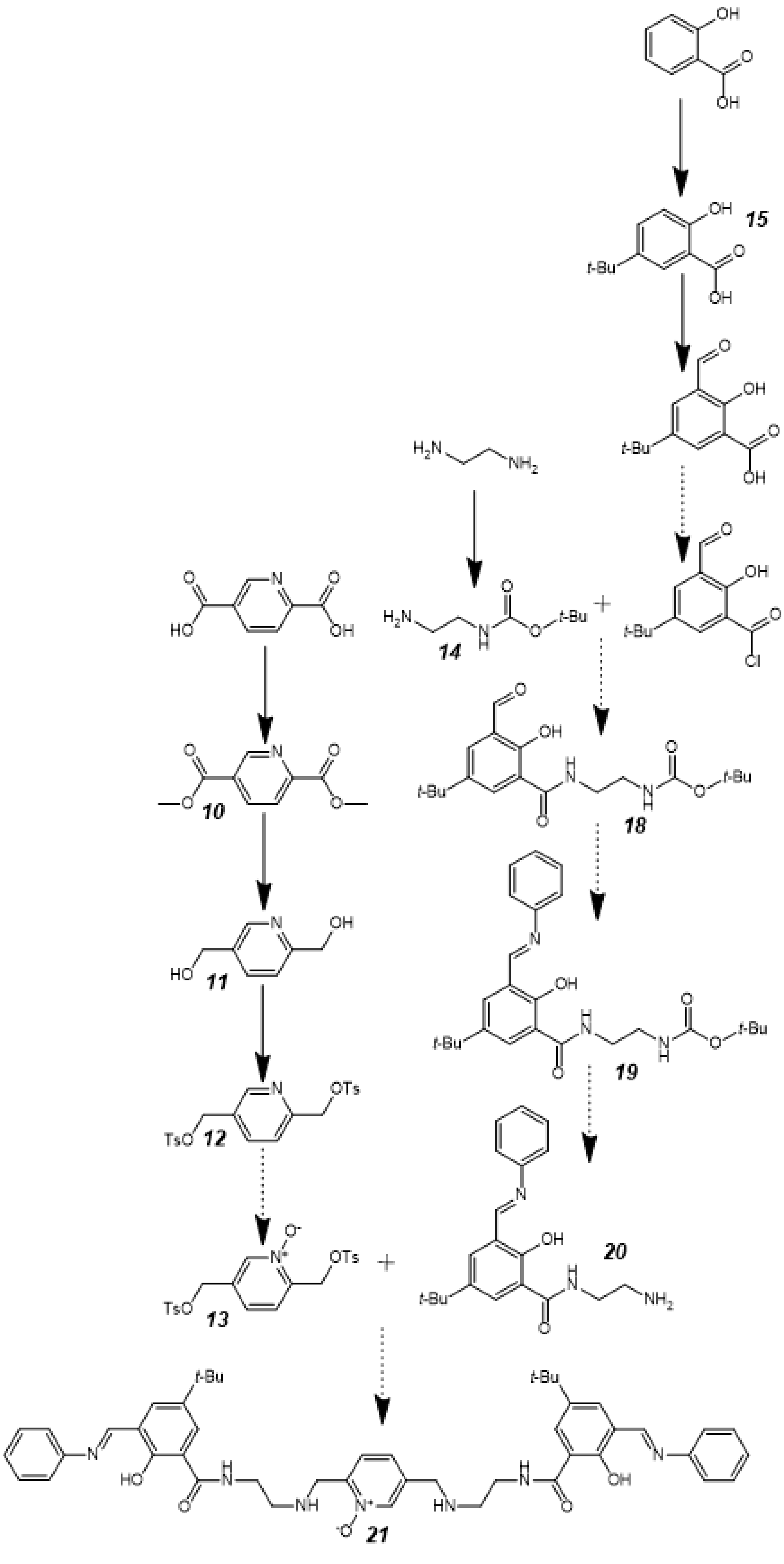
The structure of oxime (**9**) was confirmed by ^{1}H NMR. Complexation was achieved by mixing oxime (**9**) with copper chloride in 50% ethanol. The complex was washed and crystallised by vapour diffusion of diethyl ether in chloroform. Finally, the copper complex and glyphosate were mixed in chloroform before attempting crystallisation by vapour diffusion with diethyl ether (Scheme 4.3). If this was successful, any interaction between copper and glyphosate could be investigated. The final reaction may have been successful as the solution changed from a brown to a black colour, however, no adequate crystals were grown that showed glyphosate coordinating to the copper centre to confirm the successfulness of the reaction.



Scheme 4.3: The test reaction investigating if a coordinated copper and glyphosate interact.

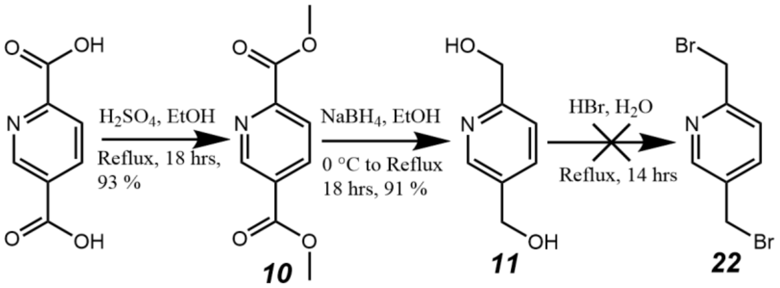
Following this, the mono- and bimetallic cages were designed. To simplify the synthesis, the terminal groups were based off the work outlined in Chapter 3. Following a few unsuccessful reactions, the initial synthetic plan had to change. The final synthetic plan for the bimetallic cage's ligand is shown in Scheme 4.4 and the monometallic cage's ligand is shown in Scheme 4.12.

The synthesis of the bimetallic cage, shown in Scheme 4.4, is simpler than the monometallic cage and therefore initially this was the focus. The ligand synthesis could be divided into two sections; preparation of the pyridine bridge and synthesis of the metal binding pocket. These fragments would then be linked at the end of the synthesis.



Scheme 4.4: Proposed synthesis of the ligand for the bimetallic cage.

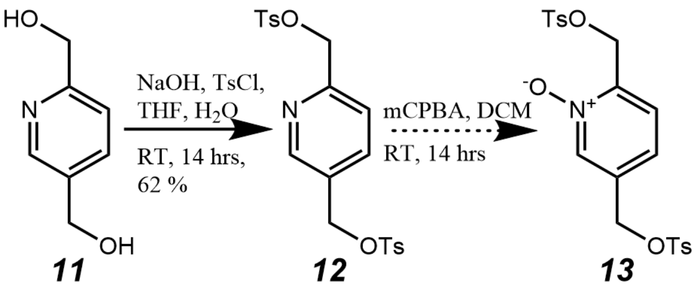
The first three steps of the synthesis of the pyridine bridge (Scheme 4.5) have been reported.[108-110] The first step is an esterification of the carboxylic acid groups, by heating 2,5-pyridinedicarboxylic acid in methanol with sulfuric acid to produce the ester (**10**). The structure was confirmed by ^{1}H NMR, which agreed with literature and showed two new peaks integrating to three protons each at 4.03 and 3.99 ppm. Reduction of **10** to 2,5-pyridinedimethanol (**11**) was achieved by reaction with sodium borohydride, refluxing in ethanol overnight. The crude material was dissolved in a mixture of acetone and aqueous potassium carbonate and heated to reflux for another hour before being separated to give the diol (**11**) as confirmed by ^{1}H NMR, which showed two new peaks integrating to two protons each at 4.62 and 4.58 ppm.



Scheme 4.5: Synthesis of 2,5-dibromopyridine (**12**).

Researchers often converted the diol to the dibromo (**22**, 2,5-bis(bromomethyl)pyridine) for further reactions as the bromo groups act as good leaving groups.[108,110] The procedure for this reaction is to dissolve the diol (**11**) in hydrogen bromide and reflux, before purifying by precipitation followed by a silica column. However, this reaction was unsuccessful, and it is likely because the product (**22**) did not precipitate out of solution in one attempt and it stuck to the silica in the column in another attempt.

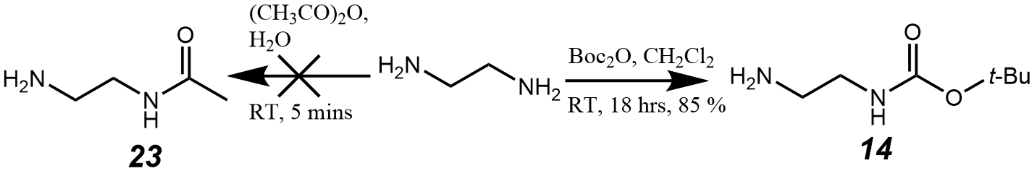
As the reaction forming the dibromo (**22**) was unsuccessful, another option was undertaken, which was to tosylate the alcohols as tosylates are good leaving groups (Scheme 4.6). This was achieved by reacting the diol (**11**) with sodium hydroxide in THF and water before adding tosyl chloride in THF. The product was extracted and purified by subjecting it to a silica column, giving the ditosylate (**12**). This was analysed by ^{1}H NMR, which showed a new peak integrating to six protons at 2.46 ppm and mass spectrometry, which showed a peak at 448.05 m/z compared to the theoretical molecular weight of 447.08. The yield for the reaction was 62% and could potentially be improved by allowing the reaction to run longer with an excess of tosyl chloride to ensure the diol is all doubly tosylated.



Scheme 4.6: Synthesis of the *N*-oxide (**14**).

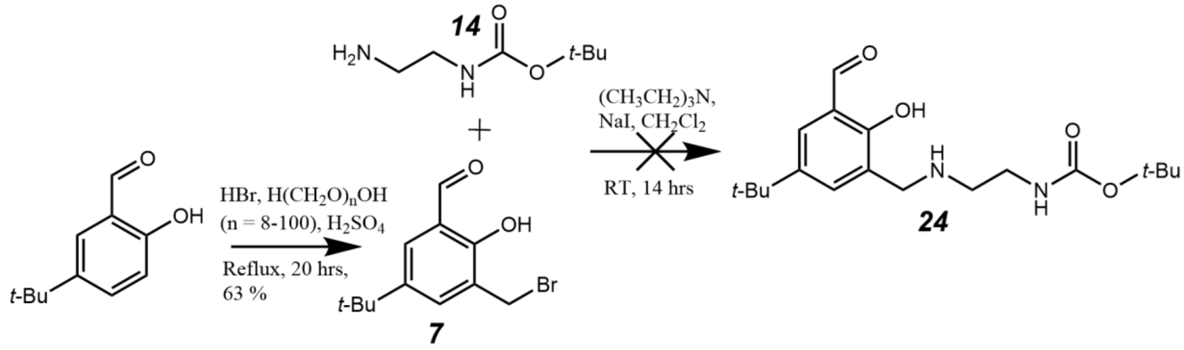
Due to time constraints the final step was not attempted, but oxidation of pyridines to *N*-oxides is a well-known reaction and there are multiple ways it could be undertaken. One such method is achieved by reacting the ditosylate (**12**) with *meta*-chloroperoxybenzoic acid in DCM. The solution would then be stirred overnight before adding diethyl ether and filtering the *N*-oxide product (**13**) off.

The synthesis of the components containing the metal binding sites were more complex than the bridging unit. The first step was to protect one amine of ethylenediamine, leaving the other free for further reactions as seen in Scheme 4.7. The first attempt at this protection was an acylation reaction (**23**), by reacting ethylenediamine with acetic anhydride in water, followed by extracting the product with ethyl acetate. However, this reaction was unsuccessful, most likely due to the product (**23**) still being water soluble and not being extracted successfully. The second attempt at protecting one side of ethylenediamine was using di-tert-butyl dicarbonate anhydride (Boc_2O). This reaction was performed by dissolving ethylene diamine and Boc_2O in DCM then stirring overnight. The crude reaction mixture was washed and dried to give the Boc-protected amine (**14**). The desired product was confirmed using ^1H NMR, which showed a new peak integrating to nine protons at 1.38 ppm.



Scheme 4.7: Reactions to protect one amine of ethylenediamine.

The next reaction was bromomethylation of 5-*tert*-butyl-2-hydroxybenzaldehyde prior to adding the Boc-ethylenediamine linker (**14**) as seen in Scheme 4.8. The bromination, as mentioned in Chapter 3, was achieved by stirring hydroxybenzaldehyde and paraformaldehyde in hydrogen bromide, before sulfuric acid was added and the solution was refluxed. The crude material was extracted and purified by recrystallisation in pentane giving the brominated product (**7**). This was confirmed by ^{1}H NMR, which showed a new peak integrating to two protons at 4.57 ppm corresponding to the methylene protons of the bromo substituent.

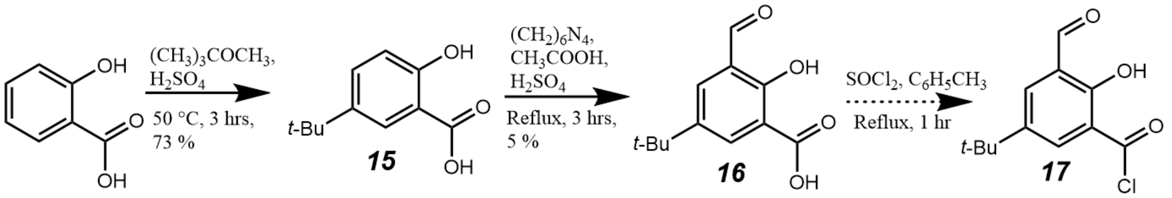


Scheme 4.8: Attempted synthesis of the secondary amine (**17**).

The following reaction between the amine (**14**) and the bromo group (**7**) was unsuccessful as the free amine underwent multiple alkylations with the reactive benzylic bromide (**7**). Numerous variations were attempted but to no avail. Initially, it was attempted by dissolving the protected amine (**14**) in DCM and adding it to a solution of TEA in DCM, followed by the slow addition of the benzylic bromide (**7**) in DCM. The solution was then refluxed overnight, washed and then partially purified by using hot chloroform and hot ethanol, giving multiple products. Mass spectrometry suggested that the unprotected amine added twice to the carbon with the bromine attached as one product and that the amine reacted with the aldehyde forming an imine as a second product.

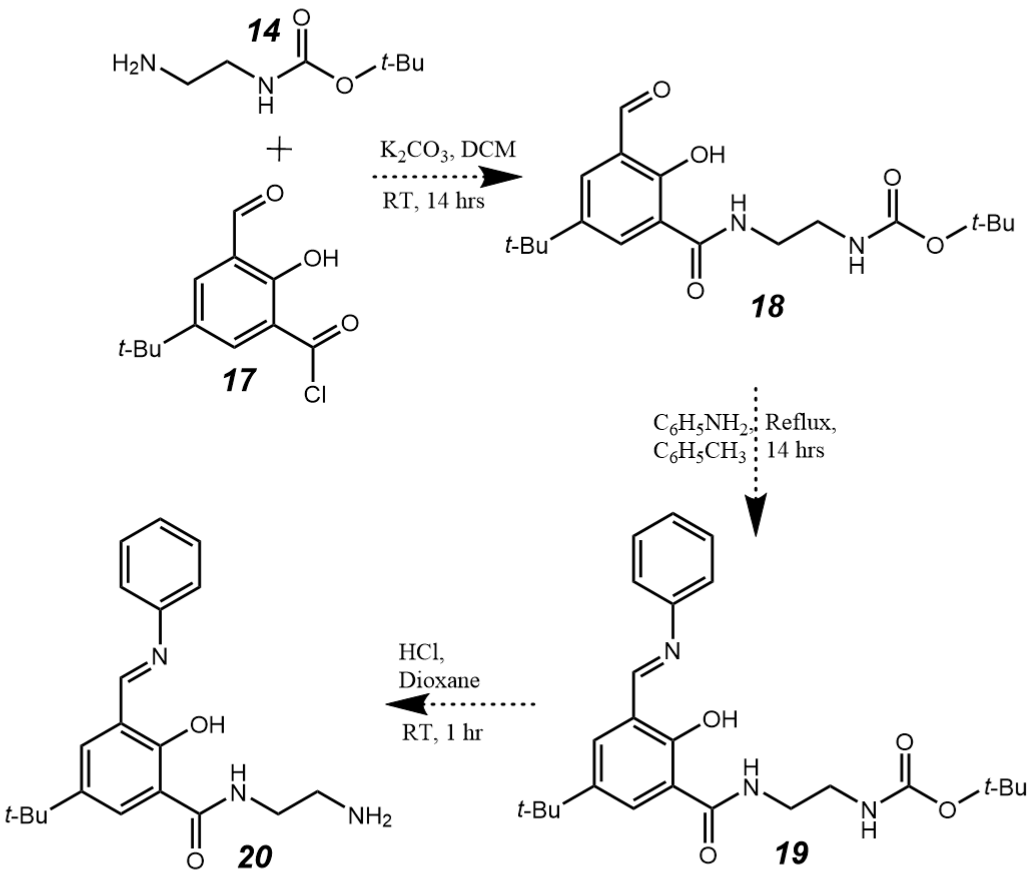
The second attempt at the reaction used a greater excess of benzylic bromide (**7**) to try to reduce the possibility of multiple additions to this moiety, however that did not appear to change anything based on the ^{1}H NMR. For the third attempt, a catalytic quantity of sodium iodide was added to benzylic bromide (**7**) before adding the amine (**14**) to attempt to convert the bromine group to an iodine group, as iodine is a better leaving group. These variations had no visible change based on the ^{1}H NMR. This led to one final attempt at this reaction and the procedure was as follows; the protected amine (**14**) was dissolved with TEA in a large volume of DCM and sodium iodide was dissolved with the bromide (**7**) in a large volume of DCM. These two solutions were then very slowly added to another volume of DCM and left to stir. The idea here was to keep the amount of available benzylic bromide (**7**) and amine (**14**) molecule as low as possible with minimal energy, to favour the intended reaction. The ^{1}H NMR of this reaction came out significantly cleaner than previous attempts, but the mass spectrograph still showed that the main product was from multiple additions of the amine to the bromo compound. As this route was unsuccessful, a new route was devised that used an amide formation reaction to try to avoid the overreaction that occurred during the initial route.

The new route required an acyl chloride to replace the methylene bromide which could be formed from salicylic acid as seen in Scheme 4.9. The first step was addition of a *tert*-butyl group, which was achieved by reaction with methyl-*tert*-butyl ether in the presence of sulfuric acid and purified by stirring in hexane overnight. This resulted in the alkylated product (**15**), confirmed by ^{1}H NMR, which showed a new peak integrating to nine protons at 1.35 ppm. Formylation of **15** involved dissolving the acid (**15**) and hexamethylenetetramine in acetic acid, then refluxing the solution for 2 hours before adding sulfuric acid and refluxing for another hour. The crude material was extracted and purified by recrystalising from methanol to give the aldehyde (**16**). This was confirmed with ^{1}H NMR, which showed the disappearance of the lowest aromatic peak. Due to time constraints, the next reaction was not attempted, but would have been a functional group conversion of the acid (**16**) to an acyl chloride (**17**). This functional group conversion would have been done by dissolving the acid (**16**) and thionyl chloride in toluene and refluxing, producing the acyl chloride (**17**).



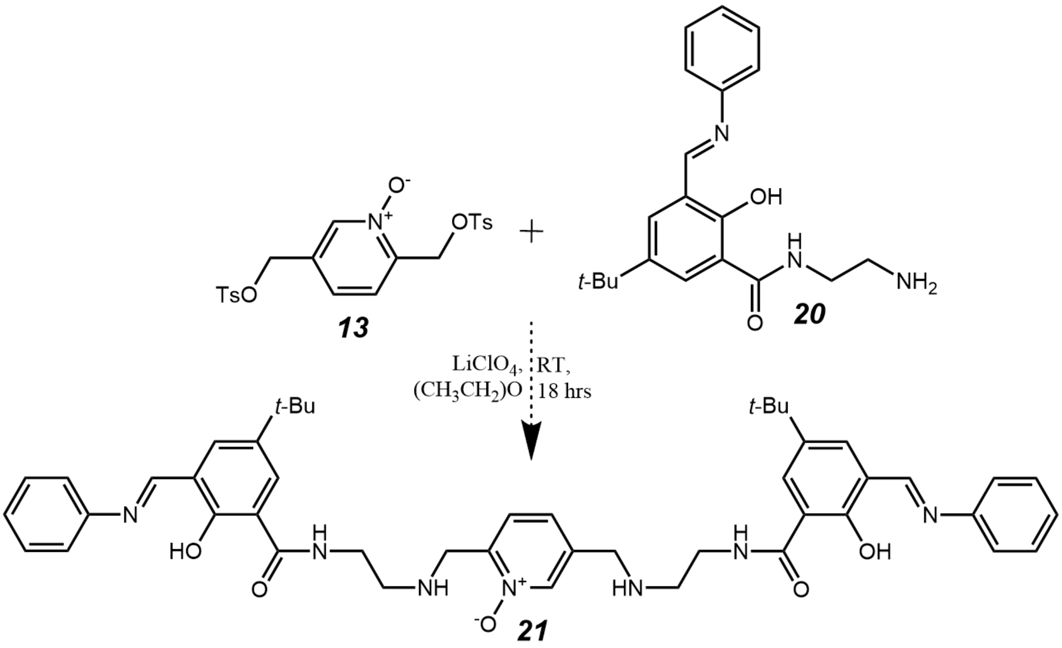
Scheme 4.9: Synthesis of the acyl chloride (**17**).

Due to time constraints the final steps of the metal binding unit was unable to be completed. The final reactions would have been completed by a sequence of amide formation, imine condensation, and deprotection as outlined in Scheme 4.10. The amide formation could be achieved by dissolving the amine (**14**) and potassium carbonate in DCM and slowly adding the acyl chloride (**17**), stirring overnight to give the amide product (**18**). The imine condensation would be based on previous work and achieved by dissolving the amide (**18**) and aniline in toluene and then refluxing to give the imine product (**19**). The final deprotection reaction would be achieved by dissolving the Boc-amine (**19**) in hydrochloric acid and dioxane and stirred for an hour to give the free amine (**20**).



Scheme 4.10: Proposed synthesis to complete the metal binding unit.

The final reaction (Scheme 4.11) to complete the ligand synthesis would be achieved by dissolving the free amine (**20**), the *N*-oxide (**13**) and lithium perchlorate in diethyl ether and leaving to stir at room temperature for 18 hours to give the complete ligand for the bimetallic cage (**21**).



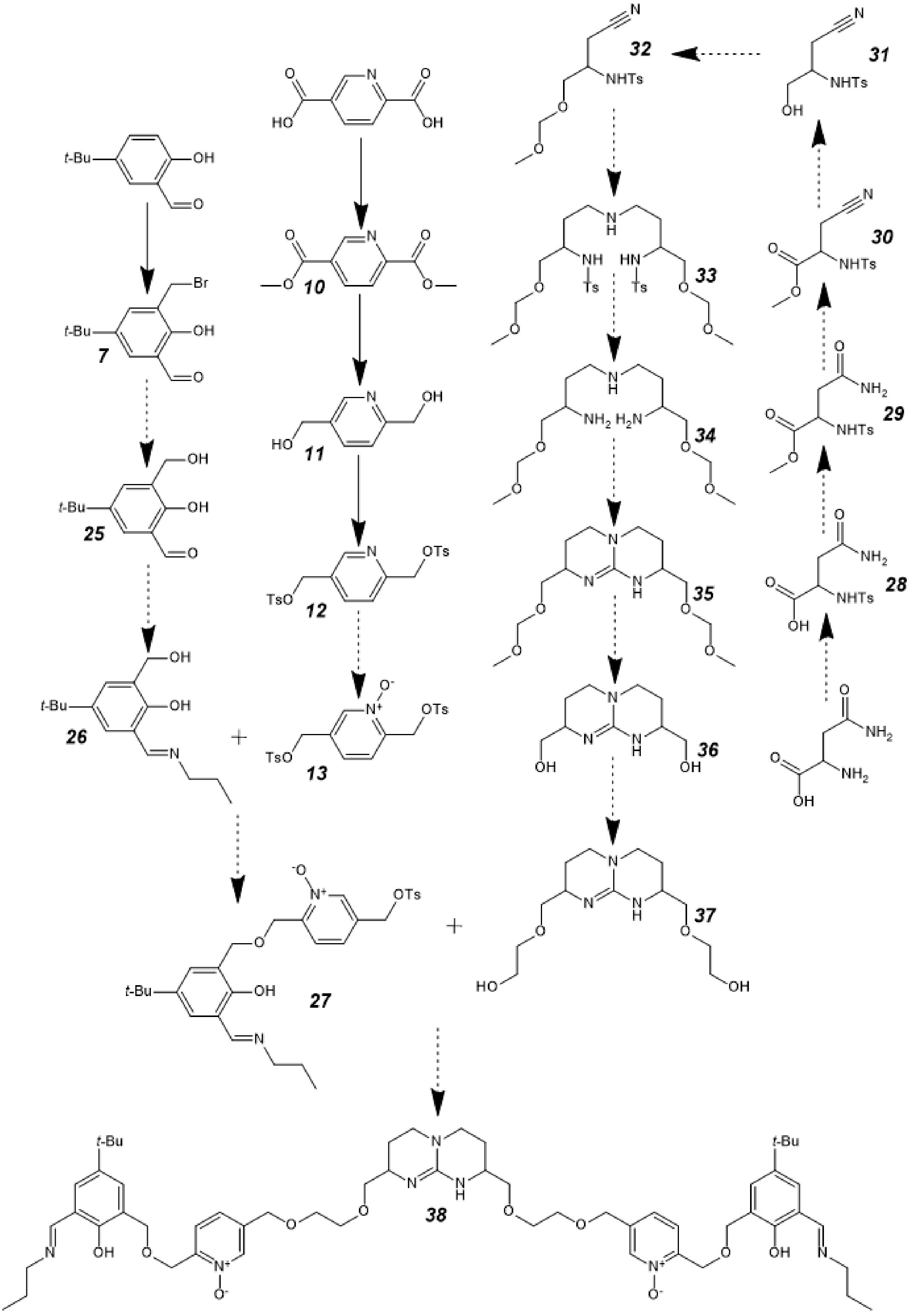
Scheme 4.11: Final reaction of the ligand for the bimetallic cage.

Synthesis of the monometallic cage is more complicated and involves a long reaction sequence. In the synthesis of the monometallic ligand, there are three main paths that converge to give the ligand: the metal binding site, the pyridine bridging unit and the guanidinium unit.

The pyridine bridging unit would be prepared as discussed previously for the bimetallic cage ligand. Synthesis of the metal binding site would follow a similar strategy (left side of Scheme 4.12) to that discussed for the bimetallic cage ligand. The first reaction would be the bromomethylation of 5-*tert*-butyl-2-hydroxybenzaldehyde to produce the benzylic bromide (**7**) as described earlier. Next is a functional group conversion of the bromine to the alcohol (**25**) by refluxing the benzylic bromide (**7**) in dioxane with calcium carbonate in water. The next step is a condensation reaction, where the aldehyde of **25** reacts with propyl amine to form the imine (**26**).

The guanidinium unit would be synthesised from asparagine by the method that was reported by Echavarren *et al.*[111] This involves a long synthetic route and was the main reason that initial studies focused on the bimetallic cage ligand. The sequence involves tosylation of the free amine of asparagine to give **28**, esterification to give **29**, conversion of the amide to cyanide **30**, ester reduction to an alcohol **31**, protection of the resulting alcohol to give **32**. Following this is catalytically hydrogenating the cyanides together to give **33**, deprotecting the tosylated amines to give **34**, ring closure using 1,1'-thiocarbonyldiimidazole to give **35**, deprotecting the alcohols to give **36**, reacting the free alcohols with bromoacetic acid to give **37**.

The final two steps in Scheme 4.12 are joining the metal binding unit (**26**), pyridine unit (**13**) and guanidinium unit (**37**) together. The first step would be joining the metal binding unit (**26**) and the pyridine unit (**13**) together by refluxing the metal binding unit, metallic sodium and the pyridine unit in THF to give the product **27**. The final reaction would follow the same conditions, by refluxing the guanidinium unit (**37**), metallic sodium and the previous product **27** in THF to give the full ligand (**38**).



Scheme 4.12: Monometallic cage's ligand synthesis.

4.6 Conclusion

The overall progress toward the glyphosate sensing cages can be seen in Table 4.1, which summarises the current state of the work performed towards the bimetallic cage and the monometallic cage.

Table 4.1: Overall progress to synthesising the cages to target glyphosate.

	Portion	Percentage complete
Bimetallic cage's ligand	Metal binding unit	43% (3/7)
	Pyridine bridging unit	75% (3/4)
Monometallic cage's ligand	Metal binding unit	33% (1/3)
	Pyridine bridging unit	75% (3/4)
	Guanidinium unit	0% (0/10)

Synthesis of the ligand for the bimetallic cage was split into two parts: the pyridine bridging unit and the metal binding unit. The synthesis of the pyridine bridging reached the penultimate step, with only formation of the *N*-oxide left to achieve. This final reaction is promising, but if it is unsuccessful then it would likely be due to the tosylated alcohol, so an alternative would be to form the *N*-oxide prior to tosylation.

The metal binding unit synthesis could be more problematic as there is more left to complete. Each reaction is viable, with the amide formation being the most likely reaction to have alternative products. If this reaction does prove to be troublesome, there are alternatives including forming the imine before forming the amide bond, but this would limit purification options. Another alternative would be to form the amide bond before adding the aldehyde, which is done in harsh conditions having the potential to break the amide bond or if these are both unsuccessful, using a secondary amine would remedy the issues but reduce the ligand's flexibility.

The monometallic cage's ligand synthesis was split into three parts: the pyridine bridging unit, the metal binding unit and the guanidinium unit. The pyridine bridging unit is the same as in the bimetallic cage's ligand. The metal binding unit synthesis is relatively simple with the bromination reaction being successful, the functional group conversion to the alcohol (**25**) has been done previously by group members successfully and the final reaction is similar to others that have been done before in this project. The first eight reactions for the guanidinium unit should be simple as it is based on the reactions reported by Echavarren *et al*.[111] and the final two reactions are likely to work as it is a deprotection and then alkylation.

Based on the work done, the synthesis of these cages are likely to be successful but they still require work. After the synthesis of these cages has been completed and analysed, they would be tested for their ability to encapsulate glyphosate. This testing would likely be done using UV/visible spectroscopy to test the binding limits or X-ray crystallography to analyse the binding strengths.

Chapter 5 – Hydrogen Cyanamide

5.1 History of Hydrogen Cyanamide

Hydrogen cyanamide is an amine bound to the carbon of a cyanide molecule. Hydrogen cyanamide is highly soluble in water and polar organic solvents. Hydrogen cyanamide can tautomerize to form a carbodiimide (left of Figure 5.1), with the cyanamide tautomer dominating, however neither tautomer can be easily isolated in their pure form.[112,113] This tautomerization was elucidated in a paper investigating the infra-red spectrum of hydrogen cyanamide in 1958.[114] Based on a structural database search, hydrogen cyanamide has been crystallised on its own (right of Figure 5.1),[115] whereas carbodiimide has not been crystallised on its own as a neutral species. Hydrogen cyanamide has been crystallised coordinating to manganese as a monoanion[116,117] and carbodiimide has been crystallised as a dianion coordinating to various metals.[118,119]

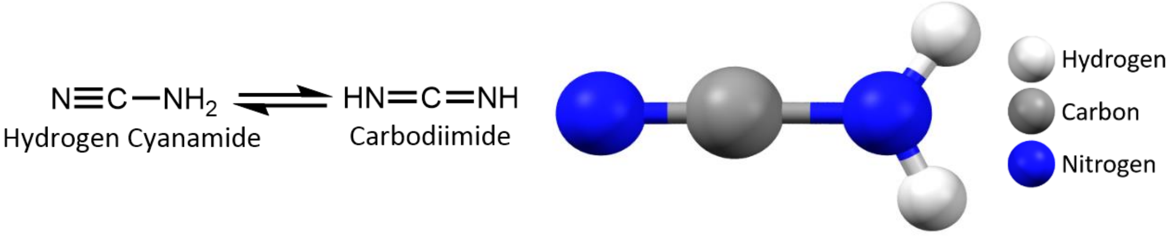
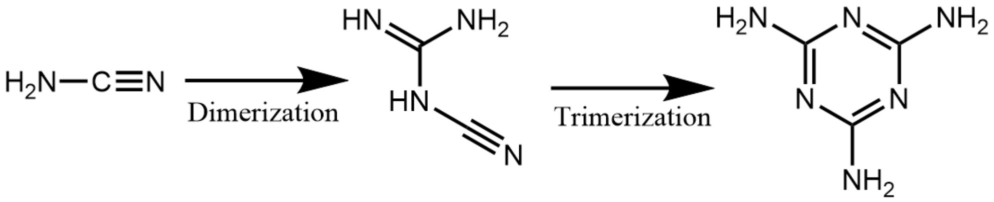


Figure 5.1: Hydrogen cyanamide tautomerization equilibrium (left) and a X-ray structure of hydrogen cyanamide (right).**[115]**

The first-time hydrogen cyanamide was reported was in 1838 by M. A. Bineau, who obtained hydrogen cyanamide by passing a gaseous stream of cyanogen chloride through an ethereal solution of ammonia. Bineau believed he had made ammonium chlorocyanate.[120] In 1851, two researchers, S. Cloez and S. Cannizzaro, repeated some of the work by Bineau and clarified the fact that Bineau had synthesised a mixture of ammonium hydrochloride and an amide species that was not ammonium chlorocyanate as Bineau claimed. Cloez and Cannizzaro were able to isolate the amide species, which they called cyanamide.[121] The synthesis and purification of pure cyanamide was originally very difficult due to hydrogen cyanamide's ability to oligomerize. Historically, the only form of cyanamide readily available was calcium cyanamide ($CaCN_2$). The synthesis of calcium cyanamide was developed by A. Frank and N. Caro in 1895, subsequently known as the Frank-Caro process. The process uses a batch oven reactor containing calcium carbide, which is heated to ~1000 °C. Nitrogen is then introduced via inlets in the reactor wall.[122] In 1918 one of the first methods to pure cyanamide was published,[123] however industrial scale synthesis and purification was still out of reach due to hydrogen cyanamide's instability. It wasn't until the 1960s that hydrogen cyanamide became readily available because researchers had developed methods to purify and stabilise hydrogen cyanamide in aqueous solution and in crystalline form.[112,113]

Hydrogen cyanamide is an amphiphilic reagent reacting as a nucleophile at the amino group and as an electrophile at the nitrile group.[113] This is very unusual for a molecule of this size, making it very useful in a variety of circumstances. Due to its amphiphilic nature, hydrogen cyanamide readily oligomerizes with itself, forming dicyandiamide (dimer) and melamine (trimer) as shown in Scheme 5.1.[124] This makes storing it difficult as it needs to be kept cool and stabilised using weak acids if in solution, because at higher pHs, hydrogen cyanamide will deprotonate and start oligomerizing. However, at very low pHs, there is a potential for hydrolysis to carbon dioxide and ammonia.[113,125]



Scheme 5.1: Hydrogen cyanamide oligomerization.

To ensure hydrogen cyanamide does not oligomerize, stabilisers are added. The stabilisers work to neutralise alkaline traces (NH_3) which can promote dimerization and trimerization. In the case of crystalline cyanamide the stabiliser is the formic acid ester, which neutralises alkaline traces, absorbs traces of moisture and continuously releases small amounts of formic acid. Phosphate buffer is added to solubilised hydrogen cyanamide to act as the stabiliser.[124]

The unique structure and reactivity of hydrogen cyanamide has allowed the hydrogen cyanamide moiety to be utilised in a wide range of applications. Hydrogen cyanamide has potential to be used as a fungicide and herbicide or as a key precursor to many important chemicals.[112,113] Hydrogen cyanamide is a precursor to other herbicides like amitrole and hexazinone, fungicides like carbendazim and benomyl and insecticides that are neonicotinoids like acetamiprid (Figure 5.2).[112] The pharmaceutical industry also uses hydrogen cyanamide to synthesise a multitude of complex molecules like cathepsin C inhibitor, which can cause tissue destruction as a treatment for cystic fibrosis, or cathepsin K inhibitor, which is used to help treat osteoporosis,[113] or cimetidine, an anti-ulcer drug.[112]

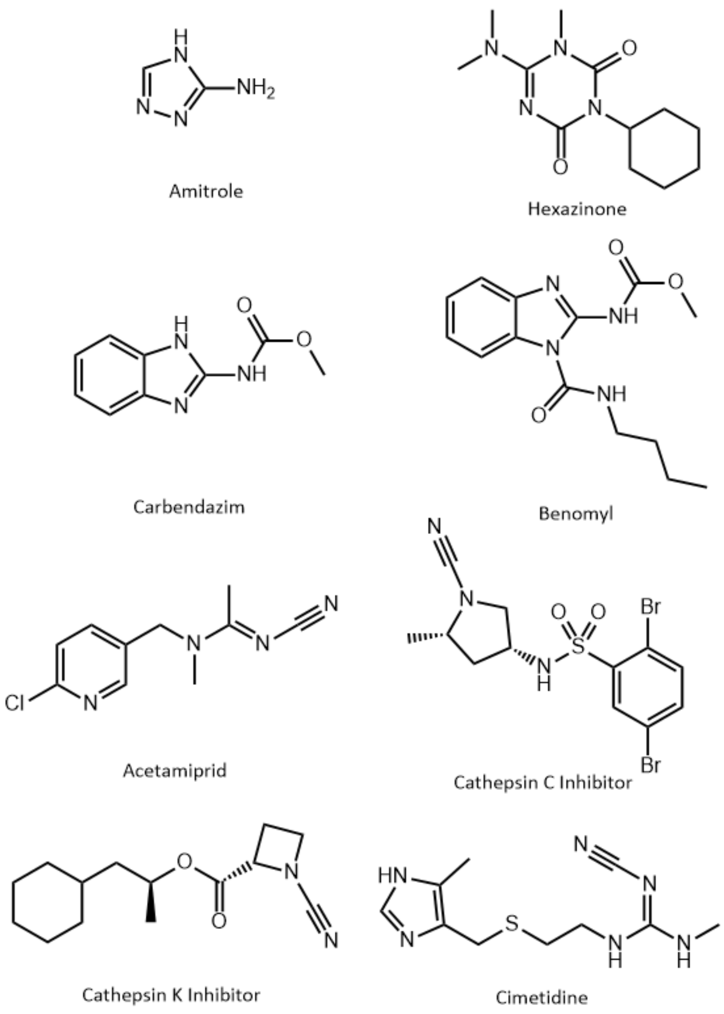


Figure 5.2: Agricultural and Pharmaceutical molecules utilising hydrogen cyanamide.

One of hydrogen cyanamide's main uses is in agriculture as a plant growth regulator and inducing budbreak (Hi-Cane and Dormex) to compensate for low winter chilling hours and to improve the budbreak of multiple crops, including kiwifruit, grapes, peaches, blueberries and apples to list a few. Budbreak refers to the leaf buds breaking open with new growth at the end of winter. The process that hydrogen cyanamide induces budbreak in fruiting plants is a complicated process that has just recently started to be understood using metabolomics. Multiple studies have been undertaken to try to understand this process in a variety of fruiting plants (grapes,[126] blueberry,[127] apple[128] and peach[129]). What these studies agree on is that there is an increase of starch hydrolysis, synthesis of ethylene and cytokinin (promotes dormancy release), increased expression of multiple proteins associated with breaking bud dormancy and with cell proliferation and an initial increase of reactive oxygen species (promotes growth initiation). All these factors caused by hydrogen cyanamide application lead to earlier budbreak.[126-130]

5.2 Reasons to Detect Hydrogen Cyanamide

The main problem with hydrogen cyanamide, is the toxicity of it. According to the European Food Safety Authority (EFSA), hydrogen cyanamide is acutely toxic to mammals via oral and dermal routes and it is known to irritate the skin and eyes, and it is a sensitiser.[23] Hydrogen cyanamide can cause respiratory irritation, contact dermatitis, headaches and gastrointestinal symptoms, including nausea vomiting or diarrhoea.[131,132] Hydrogen cyanamide is also toxic to aquatic life, other terrestrial life and some non-target plant life.[16,23] Regardless of this, hydrogen cyanamide is still used on many orchards around the world due to its effectiveness and the lack of other options. There are reports of replacements for hydrogen cyanamide (e.g., gibberellic acid, Amino25, potassium thiosulphate, urea and zinc sulfate), but none are as effective as hydrogen cyanamide when looking at budbreak, fruit ripening and fruit yield.[131,133,134]

5.3 Current Methods for Detecting Hydrogen Cyanamide

Unlike glyphosate, there are scarcely any published detection methods for hydrogen cyanamide and like glyphosate, most methods use a chromatographic technique. The variation comes with sample preparation and detection. The published methods often use tedious liquid chromatography (LC) purification methods before analysing the sample. One LC method uses multi-walled carbon nanotubes to filter the sample prior to derivatization with dansyl chloride and detection with tandem mass spectrometry.[25] Another LC method freezes the plant samples in liquid nitrogen, homogenises them, followed by extraction by centrifugation and membrane filtration, before analysis.[135] Methods using high pressure liquid chromatography (HPLC) are not very common, with most involving tedious clean-up procedures,[23] simple samples[136] or derivatization.[24] Few gas chromatography (GC) methods have been published and those that have involve complicated purification procedures prior to analysis.[137,138] So overall, there are few methods that are able to quantify hydrogen cyanamide levels in environmental samples, of which, most require expensive equipment and are unable to perform real-time analysis due to the tedious clean-up procedures. The levels that would be necessary for the environmental samples are approximately 0.1 to 50 μgL^{-1}.[139]

5.4 Target Cages and Design Features

There are two main target cages to encapsulate hydrogen cyanamide, shown in Figure 5.3. The difference between them is their terminal groups, where the left one has phenyl groups and the right one has ferrocene groups.

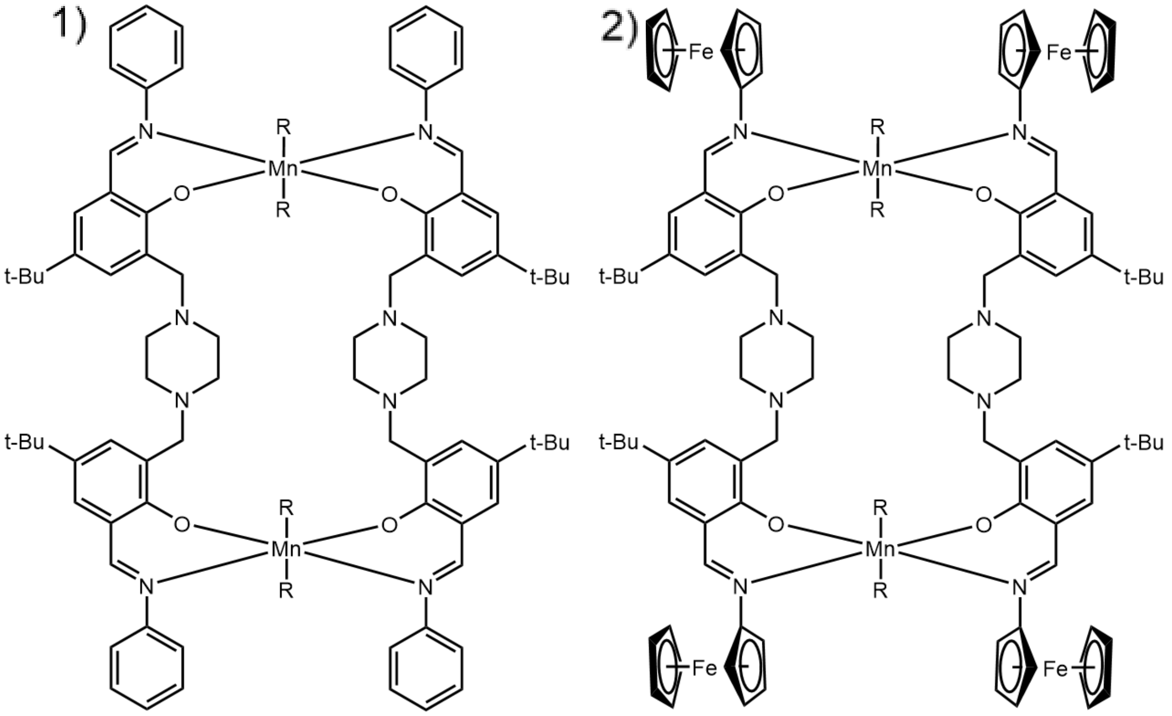


Figure 5.3: The target cages designed to capture hydrogen cyanamide. 1) is phenyl capped and 2) is ferrocene capped (R = pyridine or methanol).

As with the previous cages outlined in Chapter 3, this cage contains a piperazine linker, metal binding sites and a variable terminal group, but unlike the previous cages, these use manganese rather than copper. As previously outlined, the piperazine can be reversibly protonated allowing the release of any encapsulated molecule, meaning the cage can be reused, as well as the piperazine linker being small, stopping large molecules from being encapsulated. The metal binding site is the same as in previous cages, as it is a known synthesis. The change of metal is because manganese is known to coordinate to both ends of hydrogen cyanamide.[116,117] The phenyl terminal group is from the first cage as it is a simple synthesis to confirm if hydrogen cyanamide will interact with the manganese metal centres. The ferrocene on the terminal on the second cage it to test the capabilities of electronic detection using cyclic voltammetry.

The binding of hydrogen cyanamide in the two cage systems has been modelled using a simple geometry optimising calculation, which performs an iterative series of single point calculations until it reaches an energy minimum. This is visualised in Figure 5.4, and displays how hydrogen cyanamide could interact with the cage systems. In both systems, the piperazine groups are too far away from the hydrogen cyanamide to interact, but if there was more of a helical twist, this may not be the case. The manganese centres are interacting with both ends of the hydrogen cyanamide in both systems.

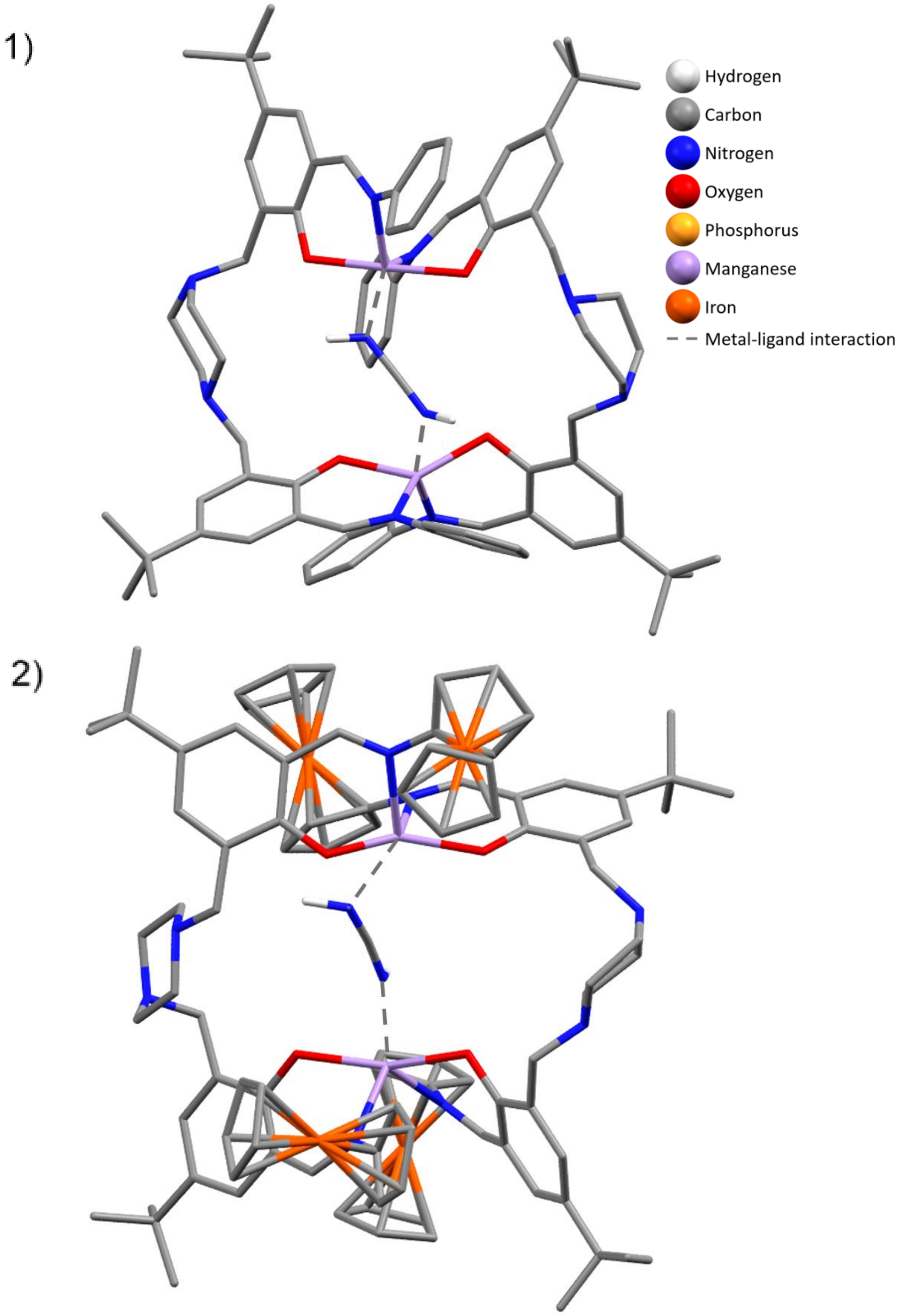
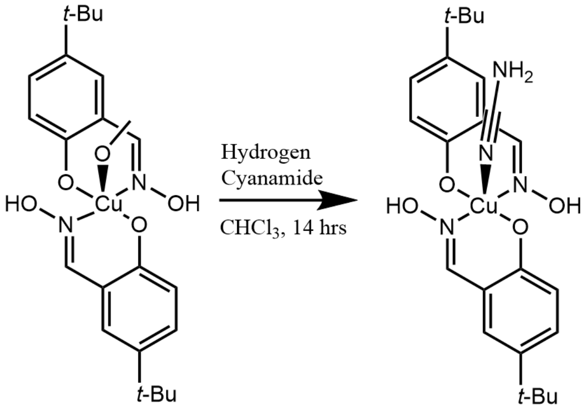


Figure 5.4: The target cages interacting with carbodiimide. 1) has the phenyl terminal groups and 2) has the ferrocene terminal groups. The hydrogens on the cages are removed for clarity

5.5 Synthesis and Analysis

Before searching the structural database for metals that hydrogen cyanamide coordinates to, a simple test system was investigated to see if hydrogen cyanamide and a copper centre could interact. The coordinate copper is from the complex [Cu(**9**)$_2$] as discussed in Chapter 4. Following the synthesis of the complex, the copper complex and hydrogen cyanamide were mixed in chloroform before crystallization by vapour diffusion with diethyl ether was attempted (Scheme 4.3).

Following the synthesis of the complex and the addition of hydrogen cyanamide (Scheme 5.2) there was an attempt to grow crystals for X-ray spectroscopy that would show an interaction between copper and hydrogen cyanamide. A colour change from brown to black suggested successful complexation but this could not be confirmed as no crystals were grown of adequate quality for X-ray crystallography.

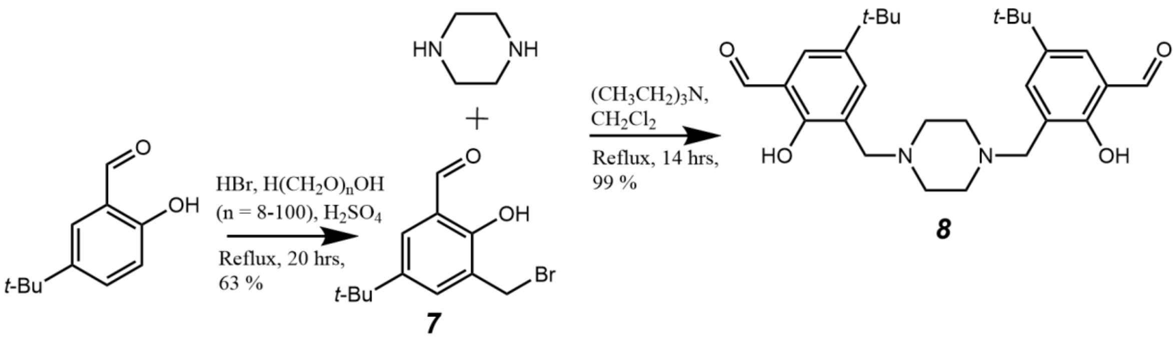


[Cu(9)$_2$(OMe)] ***[Cu(9)$_2$(Hydrogen Cyanamide)]***

Scheme 5.2: The test reaction investigating if a coordinated copper and hydrogen cyanamide interact

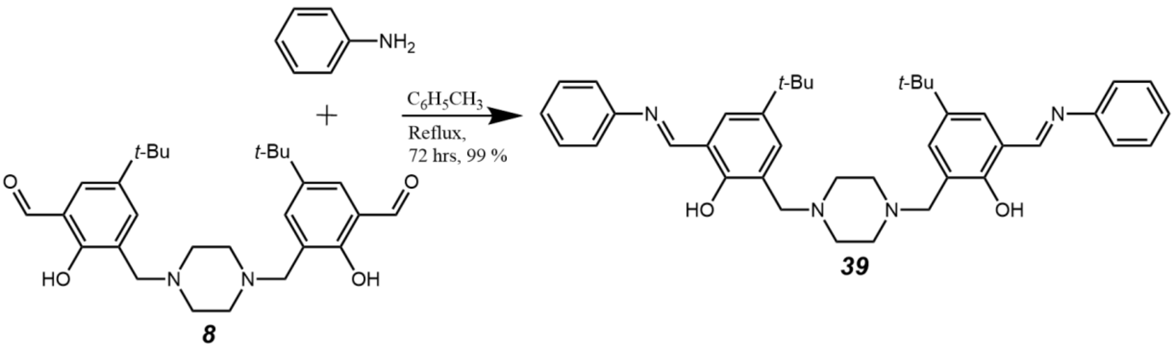
Following this, the first cage with the phenyl capped ligands was designed (left side of Figure 5.3). This cage uses the same synthetic sequence as the previous cages (Chapter 3 and Chapter 4). The synthesis of the linker for the two cages is shown in Scheme 5.3.

The first reaction was bromomethylation of 5-*tert*-butyl-2-hydroxybenzaldehyde and the second reaction was combining the benzylic bromide (**7**) with piperazine. These two reactions have been described in earlier chapters (Chapter 3 and Chapter 4).



Scheme 5.3: Synthesis of the linker for the two cages targeting hydrogen cyanamide

The final reaction, seen in Scheme 5.4, to produce the phenyl capped ligand is an imine formation reaction between the aldehyde of the piperazine linker (**8**) and aniline. This was achieved by dissolving the linker (**8**) and aniline in toluene and heating to reflux for 3 days. The crude material was purified by recrystallisation in chloroform and ethanol to give the pure product (**39**). This was confirmed by ^{1}H NMR, which showed the aldehyde hydrogen shifted from 10.25 ppm to the imine peak at 8.71 ppm and by X-ray crystallography as seen in Figure 5.5. The details about the X-ray collection data can be found in Appendix 2.



Scheme 5.4: The final reaction to the phenyl capped ligand

The crystals were grown by dissolving the product in chloroform, then methanol was added to induce crystal growth. The molecule crystallises in the C2/c space group and the asymmetric unit contains half of the ligand, with the full molecule generated by applying a reflection operation between the carbons in the piperazine ring. The full ligand structure is shown in Figure 5.5.

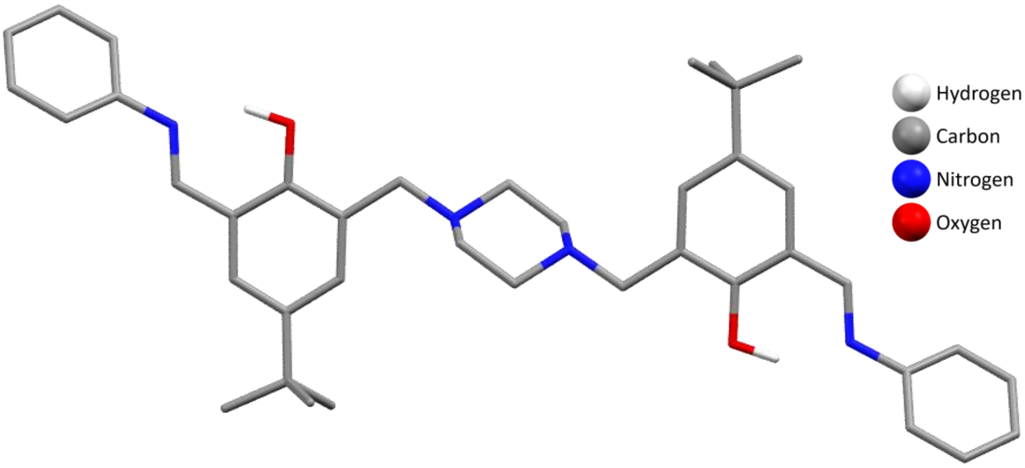


Figure 5.5: The X-ray structure of the phenyl capped ligand (**39**). The non-interacting hydrogens have been removed for clarity

The packing of the ligand, **39**, is a stacked linear 3 x 3 arrangement along the b axis as seen in Figure 5.6. The ligands have offset π-π interactions between the terminal phenyl groups of the same row holding them in that orientation with an interaction length of 3.317(3) Å. The imine bond length is 1.294(3) Å, which is within the error of closely related molecules,[140] whereas the intramolecular hydrogen bond between the phenolic hydrogen and the lone pair of electrons on the imine is 1.70(3) Å long. The hydrogen bond is moderately strong[141] and it is slightly shorter than similar molecules by 0.1-0.2 Å.[142] A search of the electron density map showed that the amines in the centre of the ligand are deprotonated.

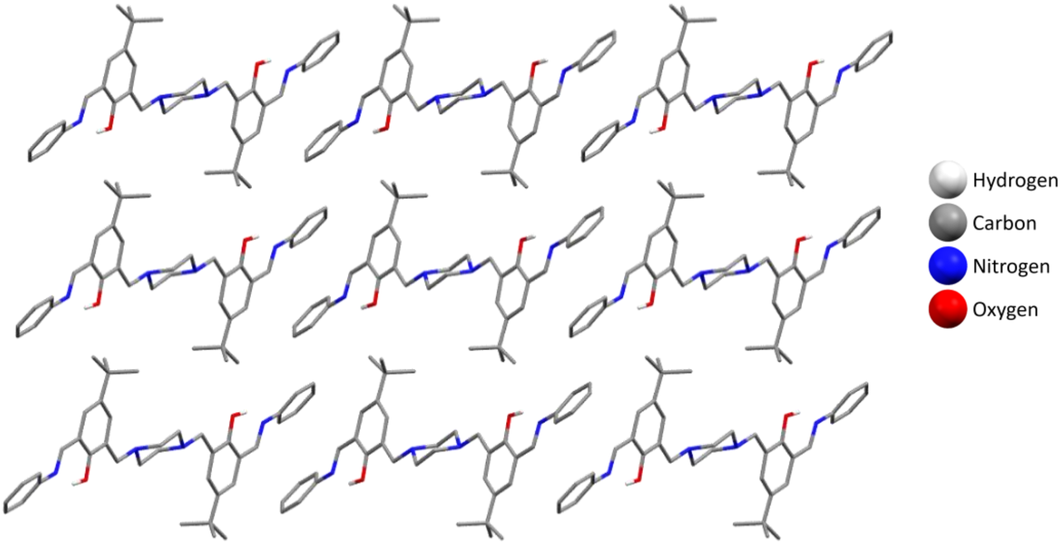
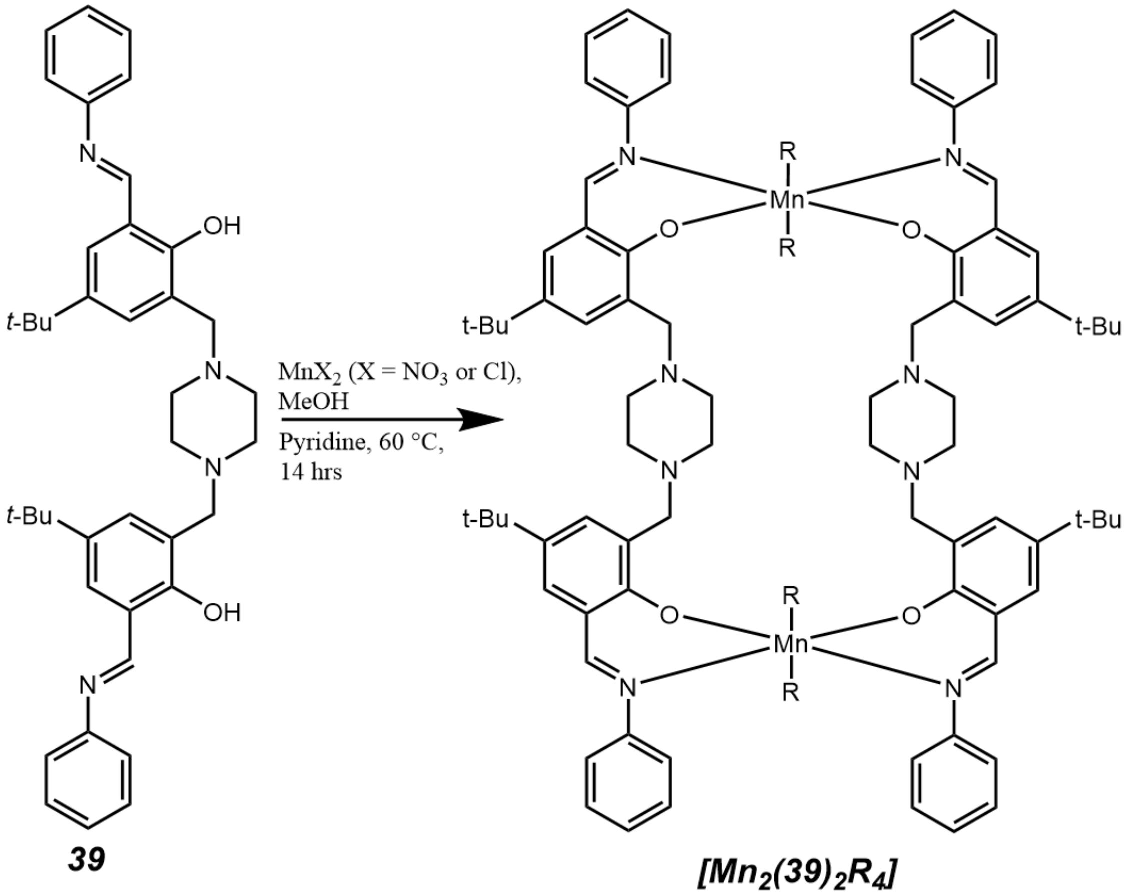


Figure 5.6: The crystal packing of the phenyl capped ligand (**39**) along the b axis. The non-interacting hydrogens have been removed for clarity

The next step was the complexation, seen in Scheme 5.5, which was achieved by dissolving the ligand (**39**, an orange solid), the nitrate or chloride manganese salt (pale pink solids) and pyridine in methanol. The solution was then stirred overnight and the solution changed from yellow to a dark brown. This colour change suggests that complexation has occurred, however no crystals have been grown that are adequate for X-ray crystallography.



Scheme 5.5: Complexation of the phenyl capped ligand (R = pyridine, H_2O or methanol)

As the complex is a neutral species with no counter ions, many simple analyses like solubility will not be able to differentiate between ligand and complex. Other techniques to try to confirm if the complex has been successfully made have been attempted, but none of these has been conclusive. The ^{1}H NMR and IR are shown in Figure 5.7. Mass spectrometry only showed peaks corresponding to the full ligand and not of the complex and the ^{1}H NMR shows broadened hydrogen peaks from the ligand. The IR shows that the ligand has not decomposed and there is some broadening of peaks, which could indicate that the complexation was successful. The imine (C=N) stretch peak is expected to shift upon complexation as the coordination bond should lengthen the imine bond and decrease the bond stretching frequency of the imine bond. However, this is not the case as the imine stretch peak remained at 1616 cm^{-1}, which is consistent with the complex in Chapter 3, that was successfully crystallised.

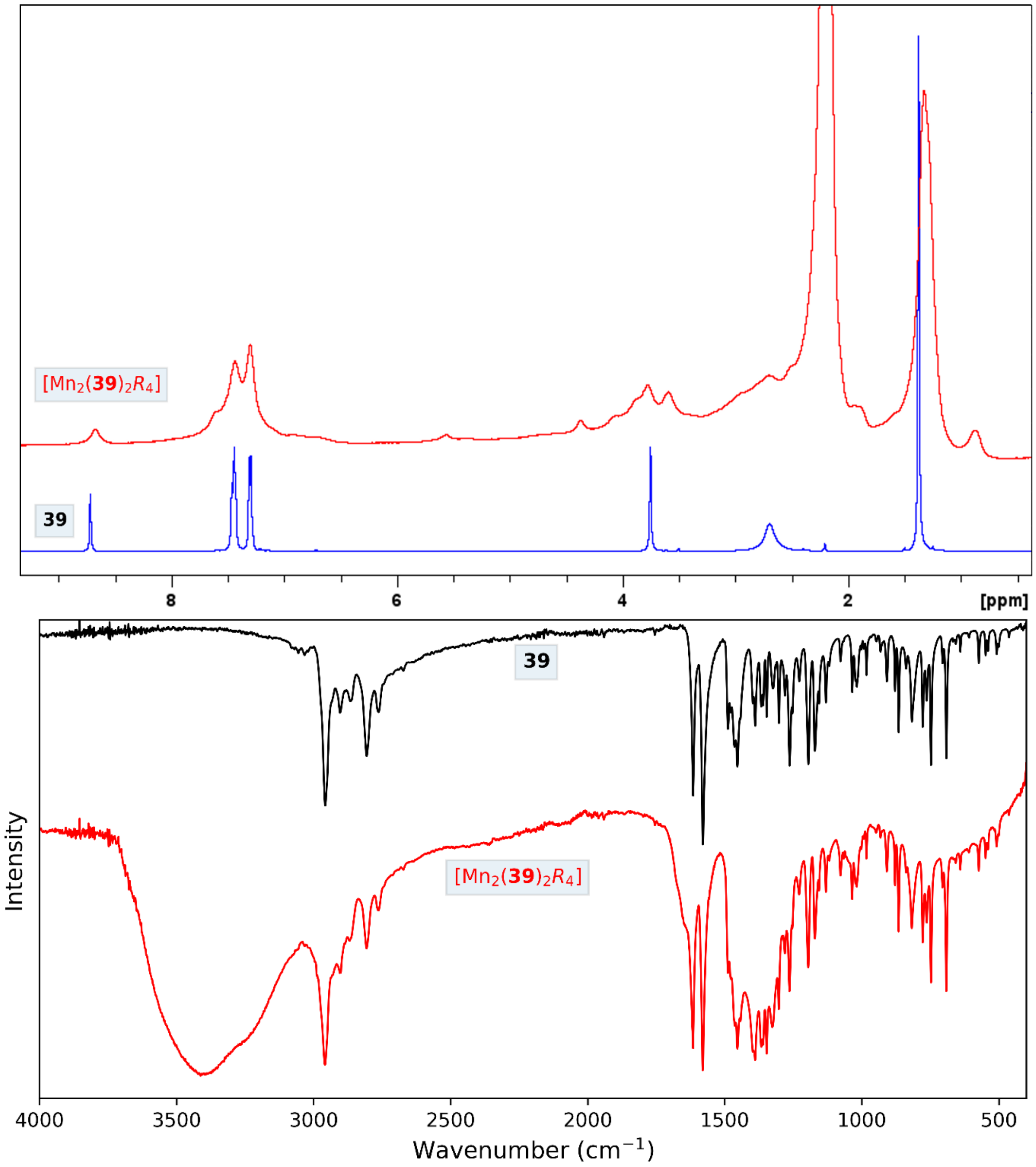
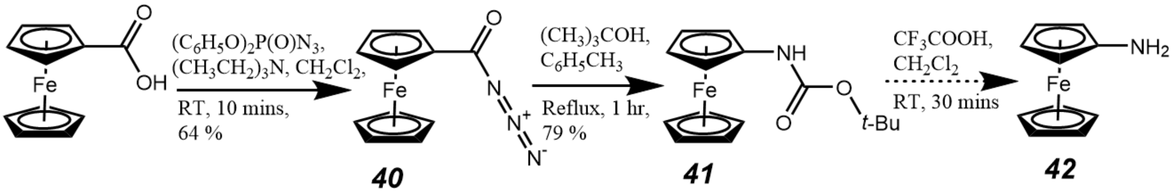


Figure 5.7: ^{1}H NMR (top) and IR (bottom) of the ligand **39** and the complex [Mn$_2$(**39**)$_2$R$_4$]

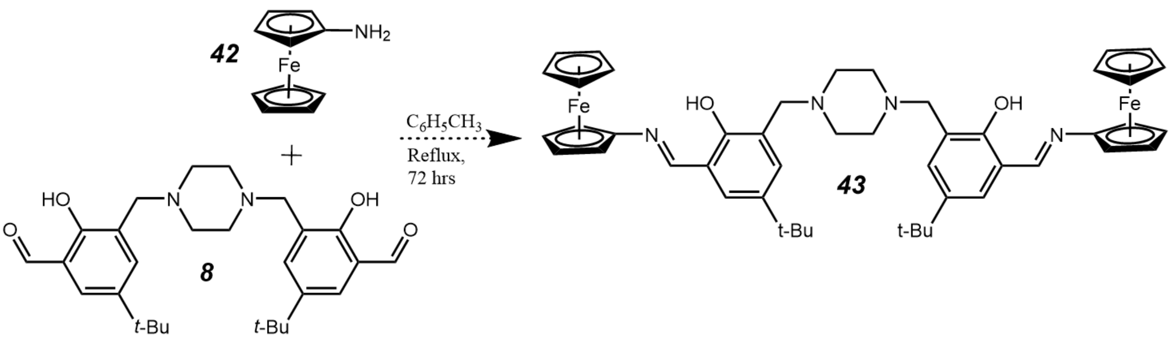
The ferrocene terminal of the second cage was synthesised from ferrocenecarboxylic acid as seen in Scheme 5.6. The first step was to synthesise an acyl azide from the acid by dissolving the acid, diphenylphosphoryl azide and triethylamine in DCM and stirring the solution for 10 minutes. The crude mixture was purified by column chromatography using ethyl acetate and hexane as the eluent, to give the pure acyl azide product (**40**). This was confirmed by ^{1}H NMR, which matched literature values. The second reaction was a Curtius rearrangement to give the Boc-protected amine (**41**). This was achieved by refluxing the acyl azide (**40**) in toluene for 30 minutes before adding *tert*-butyl alcohol and refluxing for another hour. The crude material was purified by recrystallisation in hot toluene, to give the Boc-aminoferrocene (**41**), confirmed by ^{1}H NMR, which matched literature values.

The final reaction is the deprotection of the Boc-amine (**41**) to give aminoferrocene (**42**). The deprotection reaction was achieved by suspending the Boc-amine in DCM before trifluoroacetic acid was added and the solution was stirred for 30 minutes. The crude is meant to be purified by column chromatography using ethyl acetate. However, based on the TLC, the crude appeared to be pure and due to the instability of aminoferrocene, the majority was subjected to the next reaction without further analysis. After the next reaction was started, ^{1}H NMR analysis of the remaining aminoferrocene crude was undertaken. The ^{1}H NMR showed that the crude contained some unreacted starting material, some product and some unknown materials.



Scheme 5.6: Synthesis of aminoferrocene

The next reaction was the imine formation reaction to give the full ferrocene-capped ligand (**43**), as seen in Scheme 5.7. This reaction was attempted by dissolving the linker (**8**) and the amine (**42**) in toluene and refluxing the solution for 3 days. However, upon further analysis, the crude material from the deprotection reaction was an equal mixture of multiple products that, due to time constraints, could not be identified. This showed up on the ^{1}H NMR spectra of the crude material, meaning the reaction forming the full ligand was unlikely to be successful, due to the impure aminoferrocene.



Scheme 5.7: The final reaction to the ferrocene capped ligand

During the synthesis of the hydrogen cyanamide ligands and complexes, a test was run to investigate whether hydrogen cyanamide was detectable using Raman spectroscopy without using anything to enhance the signal. The results of this test can be seen in Figure 5.8, which shows that there is no clear correlation between solid hydrogen cyanamide, a dried sample of a dilute solution of hydrogen cyanamide, and a dried sample of silver nanoparticles and hydrogen cyanamide. This shows the need for something to enhance the signal of hydrogen cyanamide to be detectable, to use Raman spectroscopy to detect hydrogen cyanamide.

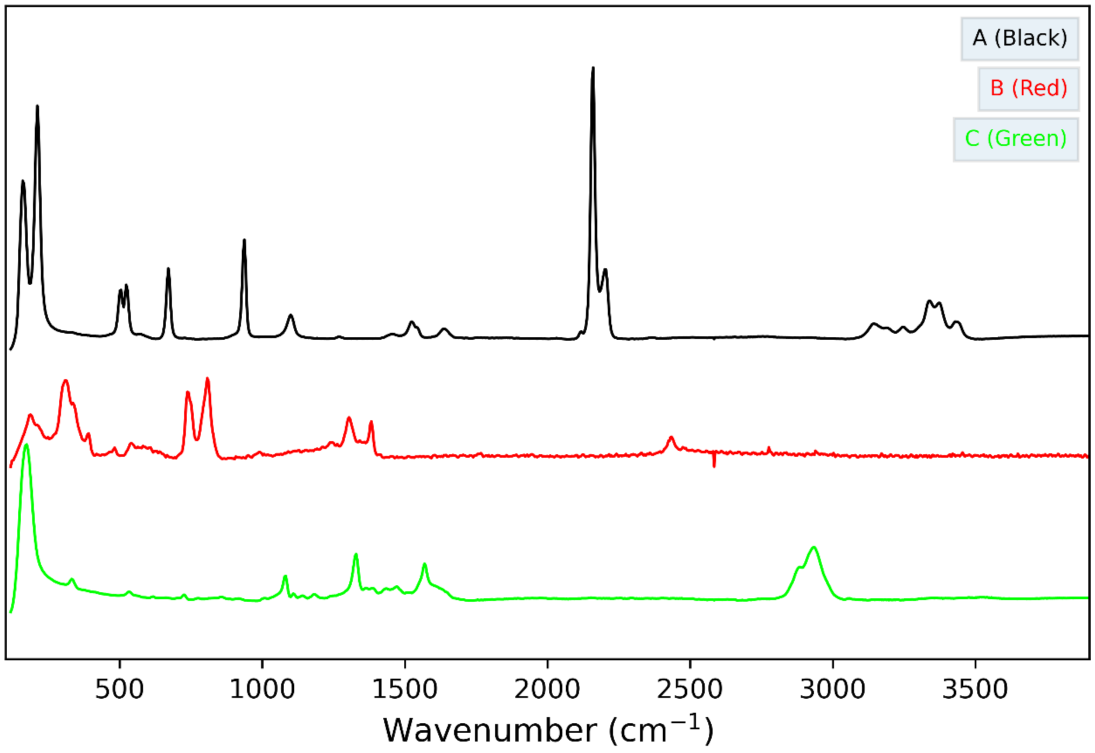


Figure 5.8: Raman spectrum of solid hydrogen cyanamide (A) and a dried solution of 10^{-5} M hydrogen cyanamide (B) with silver nanoparticles (C)

5.6 Conclusion

The overall progress toward the hydrogen cyanamide sensing cages can be seen in Table 1.1, summarising the current state of the work performed towards the phenyl capped cage and the ferrocene capped cage.

Table 5.1: Overall progress to synthesising and analysing the cages to target hydrogen cyanamide.

	Portion	**Percentage complete**
Phenyl capped cage	Ligand synthesis	100% (3/3)
	Complexation and analysis	75%
Ferrocene capped cage	Ligand synthesis	67% (4/6)
	Complexation and analysis	0%

The synthesis of the ligand for the phenyl capped cage has been completed and the complexation has been attempted twice with evidence suggesting it has been successful. With conclusive evidence of the complexation of the phenyl capped ligand, binding studies would be the next step, to investigate how strongly and selectively hydrogen cyanamide binds to the cage. If the complexation happens to not be successful, different solvents would be used to investigate if it is due to the solvent, otherwise other manganese(II) salts or manganese(III) salts could be used to achieve the complexation.

The synthesis for the ligand of the ferrocene capped cage has not been completed. There are two steps left, the deprotection of the Boc-aminoferrocene (**41**) to aminoferrocene (**42**), then the imine formation between aminoferrocene and the piperazine linker (**8**). The deprotection was not completely reacted the initial attempt, so repeating this reaction and increasing the quantity of trifluoroacetic acid used and the reaction time should help ensure the success of this reaction. Also, completely purifying the crude would aid in the success of the following reaction. The final reaction for the ferrocene capped ligand is likely to succeed with the challenge coming in the purification of the product, as imine bonds are likely to break if they are subjected to a silica column. This means that the ligand should be purified by recrystallisation, which could prove to be tough.

Following the ferrocene capped cage synthesis, the testing of the sensitivity and selectivity of the cage with cyclic voltammetry would be the next step. This will likely be after the phenyl capped cages' sensitivity and selectivity would have been tested, so this will provide an indication of how sensitive and selective the ferrocene capped cage would be.

Chapter 6 – References

1. NZKGI, *Kiwifruit Book*. **2022**.
2. Zespri *Building Back Stronger*; **2023**.
3. Little, M. A.; Cooper, A. I., The Chemistry of Porous Organic Molecular Materials. *Advanced Functional Materials* **2020,** *30* (41), 1909842.
4. Yoshizawa, M.; Klosterman, J. K.; Fujita, M., Functional Molecular Flasks: New Properties and Reactions within Discrete, Self-Assembled Hosts. *Angewandte Chemie International Edition* **2009,** *48* (19), 3418-3438.
5. Casini, A.; Woods, B.; Wenzel, M., The Promise of Self-Assembled 3D Supramolecular Coordination Complexes for Biomedical Applications. *Inorganic Chemistry* **2017,** *56* (24), 14715-14729.
6. Pöthig, A.; Casini, A., Recent Developments of Supramolecular Metal-based Structures for Applications in Cancer Therapy and Imaging. *Theranostics* **2019,** *9* (11), 3150-3169.
7. Chen, S.; Chen, L.-J., Metal–Organic Cages: Applications in Organic Reactions. *Chemistry* **2022,** *4* (2), 494-519.
8. Ham, R.; Nielsen, C. J.; Pullen, S.; Reek, J. N. H., Supramolecular Coordination Cages for Artificial Photosynthesis and Synthetic Photocatalysis. *Chemical Reviews* **2023,** *123* (9), 5225-5261.
9. Wang, M.; Vajpayee, V.; Shanmugaraju, S.; Zheng, Y.-R.; Zhao, Z.; Kim, H.; Mukherjee, P. S.; Chi, K.-W.; Stang, P. J., Coordination-Driven Self-Assembly of M_3L_2 Trigonal Cages from Preorganized Metalloligands Incorporating Octahedral Metal Centers and Fluorescent Detection of Nitroaromatics. *Inorganic Chemistry* **2011,** *50* (4), 1506-1512.
10. Liu, C.-L.; Zhang, R.-L.; Lin, C.-S.; Zhou, L.-P.; Cai, L.-X.; Kong, J.-T.; Yang, S.-Q.; Han, K.-L.; Sun, Q.-F., Intraligand Charge Transfer Sensitization on Self-Assembled Europium Tetrahedral Cage Leads to Dual-Selective Luminescent Sensing toward Anion and Cation. *Journal of the American Chemical Society* **2017,** *139* (36), 12474-12479.
11. Wenzel, M.; Bruere, S. R.; Knapp, Q. W.; Tasker, P. A.; Plieger, P. G., Zwitterionic dicopper helicates: anion encapsulation and binding studies. *Dalton Transactions* **2010,** *39* (11), 2936.
12. Stevens, J. R.; Plieger, P. G., Anion-driven conformation control and enhanced sulfate binding utilising aryl linked salicylaldoxime dicopper helicates. *Dalton Transactions* **2011,** *40* (45), 12235.

13. Liang, X.; Liang, W.; Jin, P.; Wang, H.; Wu, W.; Yang, C., Advances in Chirality Sensing with Macrocyclic Molecules. *Chemosensors* **2021,** *9* (10), 279.
14. De Silva, D. N. T.; Jameson, G. B.; Pannu, A. P. S.; Pouhet, R.; Wenzel, M.; Plieger, P. G., Piperazine linked salicylaldoxime and salicylaldimine-based dicopper(ii) receptors for anions. *Dalton Transactions* **2015,** *44* (36), 15949-15959.
15. Gill, J. P. K.; Sethi, N.; Mohan, A.; Datta, S.; Girdhar, M., Glyphosate toxicity for animals. *Environmental Chemistry Letters* **2018,** *16* (2), 401-426.
16. Eckel, W. P. Ecological Risk Assessment Problem Formulation for the Herbicide Cyanamide. *United States Environmental Protection Agency*, **2007**.
17. Abascal, E.; Gómez-Coma, L.; Ortiz, I.; Ortiz, A., Global diagnosis of nitrate pollution in groundwater and review of removal technologies. *Science of The Total Environment* **2022,** *810*, 152233.
18. Eltaweil, A. S.; Omer, A. M.; El-Aqapa, H. G.; Gaber, N. M.; Attia, N. F.; El-Subruiti, G. M.; Mohy-Eldin, M. S.; Abd El-Monaem, E. M., Chitosan based adsorbents for the removal of phosphate and nitrate: A critical review. *Carbohydrate Polymers* **2021,** *274*, 118671.
19. Lozano, I.; Pérez-Guzmán, C. J.; Mora, A.; Mahlknecht, J.; Aguilar, C. L.; Cervantes-Avilés, P., Pharmaceuticals and personal care products in water streams: Occurrence, detection, and removal by electrochemical advanced oxidation processes. *Science of The Total Environment* **2022,** *827*, 154348.
20. Pamuru, S. T.; Forgione, E.; Croft, K.; Kjellerup, B. V.; Davis, A. P., Chemical characterization of urban stormwater: Traditional and emerging contaminants. *Science of The Total Environment* **2022,** *813*, 151887.
21. Kolpin, D. W.; Furlong, E. T.; Meyer, M. T.; Thurman, E. M.; Zaugg, S. D.; Barber, L. B.; Buxton, H. T., Pharmaceuticals, Hormones, and Other Organic Wastewater Contaminants in U.S. Streams, 1999−2000: A National Reconnaissance. *Environmental Science & Technology* **2002,** *36* (6), 1202-1211.
22. Humphries, D.; Anderson, A.-M.; Byrtus, G., *Glyphosate residues in Alberta's atmospheric deposition, soils and surface waters*. Alberta Environment: **2005**.
23. Conclusion on the peer review of the pesticide risk assessment of the active substance cyanamide. *European Food Safety Authority*, **2010**.
24. Neves Junior, A. C. V.; Melo, A.; Pinho, C.; Coneglian, R. C. C.; Soares, A. G.; Ferreira, I. M. P. L. V. O., Response surface methodology for optimization of cyanamide analysis by *in situ* derivatization and dispersive liquid-liquid microextraction. *Journal of Chemometrics* **2014,** *28* (9), 716-724.

25. Cheng, C.; Di, S.; Zhang, W.; Chen, L.; Tian, Z.; Zhou, Z.; Diao, J., Determination of cyanamide residue in 21 plant-derived foods by liquid chromatography-tandem mass spectrometry. *Food Chem.* **2018,** *239,* 529-534.

26. Valle, A. L.; Mello, F. C. C.; Alves-Balvedi, R. P.; Rodrigues, L. P.; Goulart, L. R., Glyphosate detection: methods, needs and challenges. *Environmental Chemistry Letters* **2019,** *17* (1), 291-317.

27. Zambrano-Intriago, L. A.; Amorim, C. G.; Rodríguez-Díaz, J. M.; Araújo, A. N.; Montenegro, M. C. B. S. M., Challenges in the design of electrochemical sensor for glyphosate-based on new materials and biological recognition. *Science of The Total Environment* **2021,** *793*, 148496.

28. Pavia, D. L.; Lampman, G. M.; Kriz, G. S.; Vyvyan, J. R., *Introduction to Spectroscopy*. 5th ed.; Cengage Learning: Connecticut, USA, **2015**.

29. Jameson, D. M., *Introduction to Fluorescence*. Taylor & Francis Group: Florida, USA, **2014**.

30. Albani, J. R., *Principles and Applications of Fluorescence Spectroscopy*. Blackwell Science: Oxford, England, **2007**.

31. Mehta, M. Surface-Enhanced Raman Spectroscopy for Environmental and Biological Analysis. Massey University, Manawatu, New Zealand, **2019**.

32. Zhao, Q.; Hilal, H.; Kim, J.; Park, W.; Haddadnezhad, M.; Lee, J.; Park, W.; Lee, J.-W.; Lee, S.; Jung, I.; Park, S., All-Hot-Spot Bulk Surface-Enhanced Raman Scattering (SERS) Substrates: Attomolar Detection of Adsorbates with Designer Plasmonic Nanoparticles. *Journal of the American Chemical Society* **2022,** *144* (29), 13285-13293.

33. Mehta, M.; Waterland, M., Ultrasensitive surface-enhanced Raman scattering detection of biological pollutants by controlled evaporation on omniphobic substrates. *Heliyon* **2020,** *6* (6), e04317.

34. Yamada, H.; Yoshii, K.; Asahi, M.; Chiku, M.; Kitazumi, Y., Cyclic Voltammetry Part 1: Fundamentals. *Electrochemistry* **2022,** *90* (10), 102005-102005.

35. Compton, R. G.; Banks, C. E., *Understanding Voltammetry*. 3rd ed.; World Scientific Publishing Europe Ltd.: London, England, **2018**.

36. Elgrishi, N.; Rountree, K. J.; McCarthy, B. D.; Rountree, E. S.; Eisenhart, T. T.; Dempsey, J. L., A Practical Beginner's Guide to Cyclic Voltammetry. *Journal of Chemical Education* **2018,** *95* (2), 197-206.

37. Sanko, V.; Şenocak, A.; Tümay, S. O.; Orooji, Y.; Demirbas, E.; Khataee, A., An electrochemical sensor for detection of trace-level endocrine disruptor bisphenol A

using $Mo_2Ti_2AlC_3$ MAX phase/MWCNT composite modified electrode. *Environmental Research* **2022,** *212*, 113071.

38. Eilers, P. H. C., A Perfect Smoother. *Analytical Chemistry* **2003,** *75* (14), 3631-3636.

39. Retrieved from: https://github.com/scipy/scipy/blob/v1.10.1/scipy/signal/_peak_finding.py#L729-L1010

40. Bhadu, A.; Singh, B.; Gulshan, T.; Kumawat, S. N.; Choudhary, R. K.; Farooq, F., Customized Fertilizer: A Key for Enhanced Crop Production. *International Journal of Plant & Soil Science* **2022,** *34* (23), 954-964.

41. FAO. *FAOSTAT*. License: CC BY-NC-SA 3.0 IGO. Extracted from: https://www.fao.org/faostat/en/#data/RFN. Date of Access: **15-09-2023**

42. Fuglie, K.; Jelliffe, J.; Morgan, S. International Agricultural Productivity. *U.S. Department of Agriculture*, **2022**.

43. Wang, Y.; Wu, W.-H., Genetic approaches for improvement of the crop potassium acquisition and utilization efficiency. *Current Opinion in Plant Biology* **2015,** *25*, 46-52.

44. Tripathi, S.; Srivastava, P.; Devi, R. S.; Bhadouria, R., Influence of synthetic fertilizers and pesticides on soil health and soil microbiology. In *Agrochemicals Detection, Treatment and Remediation*, Majeti Narasimha Vara Prasad, Ed. Butterworth-Heinemann: **2020;** pp 25-54.

45. Chojnacka, K.; Moustakas, K.; Witek-Krowiak, A., Bio-based fertilizers: A practical approach towards circular economy. *Bioresource Technology* **2020,** *295*, 122223.

46. Sultenfuss, J. H.; Doyle, W. J., *Better Crops With Plant Food*. Potash & Phosphate Institute: Georgia, United States, **1999;** Vol. 83.

47. Gutiérrez, M.; Biagioni, R. N.; Alarcón-Herrera, M. T.; Rivas-Lucero, B. A., An overview of nitrate sources and operating processes in arid and semiarid aquifer systems. *Science of The Total Environment* **2018,** *624*, 1513-1522.

48. Tokazhanov, G.; Ramazanova, E.; Hamid, S.; Bae, S.; Lee, W., Advances in the catalytic reduction of nitrate by metallic catalysts for high efficiency and N_2 selectivity: A review. *Chemical Engineering Journal* **2020,** *384*, 123252.

49. Smil, V., Phosphorus in the Environment: Natural Flows and Human Interferences. *Annual Review of Energy and the Environment* **2000,** *25* (1), 53-88.

50. Hussain, M. Z.; Hamilton, S. K.; Robertson, G. P.; Basso, B., Phosphorus availability and leaching losses in annual and perennial cropping systems in an upper US Midwest landscape. *Scientific Reports* **2021,** *11* (1).

51. Bishayee, B.; Chatterjee, R. P.; Ruj, B.; Chakrabortty, S.; Nayak, J., Strategic management of nitrate pollution from contaminated water using viable adsorbents: An economic assessment-based review with possible policy suggestions. *Journal of Environmental Management* **2022,** *303*, 114081.
52. Liu, Y.; Zhang, X.; Wang, J., A critical review of various adsorbents for selective removal of nitrate from water: Structure, performance and mechanism. *Chemosphere* **2022,** *291*, 132728.
53. Velusamy, K.; Periyasamy, S.; Kumar, P. S.; Vo, D.-V. N.; Sindhu, J.; Sneka, D.; Subhashini, B., Advanced techniques to remove phosphates and nitrates from waters: a review. *Environmental Chemistry Letters* **2021,** *19* (4), 3165-3180.
54. Zhao, F.; Xin, J.; Yuan, M.; Wang, L.; Wang, X., A critical review of existing mechanisms and strategies to enhance N2 selectivity in groundwater nitrate reduction. *Water Research* **2022,** *209*, 117889.
55. Moumen, E.; Bazzi, L.; El Hankari, S., Metal-organic frameworks and their composites for the adsorption and sensing of phosphate. *Coordination Chemistry Reviews* **2022,** *455*, 214376.
56. Sohrabi, S.; Moraveji, M. K.; Mousavi, S., From nitrate determination using microfluidic sensors to photocatalytic process intensification. *International Journal of Environmental Analytical Chemistry* **2022,** *102* (10), 2416-2450.
57. Singh, S.; Anil, A. G.; Kumar, V.; Kapoor, D.; Subramanian, S.; Singh, J.; Ramamurthy, P. C., Nitrates in the environment: A critical review of their distribution, sensing techniques, ecological effects and remediation. *Chemosphere* **2022,** *287*, 131996.
58. Tan, J. F.; Anastasi, A.; Chandra, S., Electrochemical detection of nitrate, nitrite and ammonium for on-site water quality monitoring. *Current Opinion in Electrochemistry* **2022,** *32*, 100926.
59. Mahato, P., An overview of the molecular sensors developed for the recognition of inorganic phosphate. *Journal of the Indian Chemical Society* **2022,** *99* (5), 100414.
60. Patel, V.; Kruse, P.; Selvaganapathy, P. R., Review—Solid State Sensors for Phosphate Detection in Environmental and Medical Diagnostics. *Journal of The Electrochemical Society* **2022,** *169* (7), 077505.
61. Guo, Z.-X.; Cai, Q.; Yang, Z., Determination of glyphosate and phosphate in water by ion chromatography—inductively coupled plasma mass spectrometry detection. *Journal of Chromatography A* **2005,** *1100* (2), 160-167.

62. Kalimuthu, P.; Fischer-Schrader, K.; Schwarz, G.; Bernhardt, P. V., A sensitive and stable amperometric nitrate biosensor employing Arabidopsis thaliana nitrate reductase. *JBIC Journal of Biological Inorganic Chemistry* **2015,** *20* (2), 385-393.

63. Blazek, J. A Thiol Functionalised Salicylaldimine ligand Synthesis and Structural Analysis of Cu(II) Complexes. Massey University, **2013**.

64. Sienkiewicz-Gromiuk, J.; Tarasiuk, B.; Mazur, L., New organic single crystal of (benzylthio)acetic acid: Synthesis, crystal structure, spectroscopic (ATR-FTIR, ^{1}H and ^{13}C NMR) and thermal characterization. *Journal of Molecular Structure* **2016,** *1110*, 65-71.

65. Heyer, D.; Lehn, J.-M., Anion coordination chemistry - synthesis and anion binding features of cyclophane type macrobicyclic anion receptor molecules. *Tetrahedron Letters* **1986,** *27* (48), 5869-5872.

66. Plieger, P. G.; Parsons, S.; Parkin, A.; Tasker, P. A., Transport of metal salts; encapsulation of anions in dinuclear Cu(II) complexes $[Cu_2L_2SO_4]SO_4$ and $[Cu_2L_2BF_4](BF_4)_3$, where L = 2,2’-[1,6-hexanediylbis[(methylimino)methylene]]bis[4-tert-butyl-6-(phenylazomethinyl)phenol]. *Journal of the Chemical Society, Dalton Transactions* **2002,** (21), 3928-3930.

67. Yang, L.; Powell, D. R.; Houser, R. P., Structural variation in copper(I) complexes with pyridylmethylamide ligands: structural analysis with a new four-coordinate geometry index, τ_4. *Dalton Trans.* **2007,** (9), 955-964.

68. Guselnikova, O.; Lim, H.; Kim, H. J.; Kim, S. H.; Gorbunova, A.; Eguchi, M.; Postnikov, P.; Nakanishi, T.; Asahi, T.; Na, J.; Yamauchi, Y., New Trends in Nanoarchitectured SERS Substrates: Nanospaces, 2D Materials, and Organic Heterostructures. *Small* **2022,** *18* (25), 2107182.

69. Yuan, H.; Khoury, C. G.; Hwang, H.; Wilson, C. M.; Grant, G. A.; Vo-Dinh, T., Gold nanostars: surfactant-free synthesis, 3D modelling, and two-photon photoluminescence imaging. *Nanotechnology* **2012,** *23* (7), 075102.

70. Dill, G. M.; Sammons, R. D.; Feng, P. C. C.; Kohn, F.; Kretzmer, K.; Mehrsheikh, A.; Bleeke, M.; Honegger, J. L.; Farmer, D.; Wright, D.; Haupfear, E. A., Glyphosate: Discovery, Development, Applications, and Properties. *Glyphosate Resistance in Crops and Weeds* **2010**, 1-33.

71. Vidal, E.; Negro, A.; Cassano, A.; Zalazar, C., Simplified reaction kinetics, models and experiments for glyphosate degradation in water by the UV/H_2O_2 process. *Photochemical & Photobiological Sciences* **2015,** *14* (2), 366-377.

72. Wilson, C. J. G.; Wood, P. A.; Parsons, S., Discerning subtle high-pressure phase transitions in glyphosate. *CrystEngComm* **2023,** *25* (6), 988-997.
73. Duke, S. O.; Powles, S. B., Glyphosate: a once-in-a-century herbicide. *Pest Management Science* **2008,** *64* (4), 319-325.
74. Duke, S. O., The history and current status of glyphosate. *Pest Management Science* **2018,** *74* (5), 1027-1034.
75. Duke, S. O.; Wedge, D. E.; Cerdeira, A. L.; Matallo, M. B., Interactions of Synthetic Herbicides with Plant Disease and Microbial Herbicides. In *Novel Biotechnologies for Biocontrol Agent Enhancement and Management*, Maurizio Vurro; Jonathan Gressel, Eds. Springer: Dordrecht, **2007;** pp 277-296.
76. Herrmann, K. M.; Weaver, L. M., The Shikimate Pathway. *Annual Review of Plant Physiology and Plant Molecular Biology* **1999,** *50* (1), 473-503.
77. Reddy, K. N.; Duke, S. O., Soybean Mineral Composition and Glyphosate Use. In *Processing and Impact on Active Components in Food*, Victor Preedy, Ed. Academic Press: San Diego, **2015;** pp 369-376.
78. Singh, S.; Kumar, V.; Datta, S.; Wani, A. B.; Dhanjal, D. S.; Romero, R.; Singh, J., Glyphosate uptake, translocation, resistance emergence in crops, analytical monitoring, toxicity and degradation: a review. *Environmental Chemistry Letters* **2020,** *18* (3), 663-702.
79. Schönherr, J., A mechanistic analysis of penetration of glyphosate salts across astomatous cuticular membranes. *Pest Management Science* **2002,** *58* (4), 343-351.
80. Liu, Z., Effects of surfactants on foliar uptake of herbicides – a complex scenario. *Colloids and Surfaces B: Biointerfaces* **2004,** *35* (3), 149-153.
81. Liu, H.; Shao, B.; Long, X.; Yao, Y.; Meng, Q., Foliar penetration enhanced by biosurfactant rhamnolipid. *Colloids and Surfaces B: Biointerfaces* **2016,** *145*, 548-554.
82. Stevens, P. J. G.; Kimberley, M. O.; Murphy, D. S.; Policello, G. A., Adhesion of spray droplets to foliage: The role of dynamic surface tension and advantages of organosilicone surfactants. *Pesticide Science* **1993,** *38* (2-3), 237-245.
83. Benbrook, C. M., Trends in glyphosate herbicide use in the United States and globally. *Environmental Sciences Europe* **2016,** *28* (1).
84. Green, J. M., The rise and future of glyphosate and glyphosate-resistant crops. *Pest Management Science* **2018,** *74* (5), 1035-1039.

85. Sun, M.; Li, H.; Jaisi, D. P., Degradation of glyphosate and bioavailability of phosphorus derived from glyphosate in a soil-water system. *Water Research* **2019,** *163*, 114840.

86. Zhan, H.; Feng, Y.; Fan, X.; Chen, S., Recent advances in glyphosate biodegradation. *Applied Microbiology and Biotechnology* **2018,** *102* (12), 5033-5043.

87. Silva, V.; Montanarella, L.; Jones, A.; Fernández-Ugalde, O.; Mol, H. G. J.; Ritsema, C. J.; Geissen, V., Distribution of glyphosate and aminomethylphosphonic acid (AMPA) in agricultural topsoils of the European Union. *Science of The Total Environment* **2018,** *621*, 1352-1359.

88. Bento, C. P. M.; Yang, X.; Gort, G.; Xue, S.; Van Dam, R.; Zomer, P.; Mol, H. G. J.; Ritsema, C. J.; Geissen, V., Persistence of glyphosate and aminomethylphosphonic acid in loess soil under different combinations of temperature, soil moisture and light/darkness. *Science of The Total Environment* **2016,** *572*, 301-311.

89. Gill, J. P. K.; Sethi, N.; Mohan, A., Analysis of the glyphosate herbicide in water, soil and food using derivatising agents. *Environmental Chemistry Letters* **2017,** *15* (1), 85-100.

90. Williams, G. M.; Kroes, R.; Munro, I. C., Safety Evaluation and Risk Assessment of the Herbicide Roundup and Its Active Ingredient, Glyphosate, for Humans. *Regulatory Toxicology and Pharmacology* **2000,** *31* (2), 117-165.

91. Mesnage, R.; Defarge, N.; Spiroux de Vendômois, J.; Séralini, G. E., Potential toxic effects of glyphosate and its commercial formulations below regulatory limits. *Food and Chemical Toxicology* **2015,** *84*, 133-153.

92. Van Bruggen, A. H. C.; He, M. M.; Shin, K.; Mai, V.; Jeong, K. C.; Finckh, M. R.; Morris, J. G., Environmental and health effects of the herbicide glyphosate. *Science of The Total Environment* **2018,** *616-617*, 255-268.

93. Costas-Ferreira, C.; Durán, R.; Faro, L. R. F., Toxic Effects of Glyphosate on the Nervous System: A Systematic Review. *International Journal of Molecular Sciences* **2022,** *23* (9), 4605.

94. Koskinen, W. C.; Marek, L. J.; Hall, K. E., Analysis of glyphosate and aminomethylphosphonic acid in water, plant materials and soil. *Pest Management Science* **2016,** *72* (3), 423-432.

95. Ibáñez, M.; Pozo, Ó. J.; Sancho, J. V.; López, F. J.; Hernández, F., Re-evaluation of glyphosate determination in water by liquid chromatography coupled to electrospray tandem mass spectrometry. *Journal of Chromatography A* **2006,** *1134* (1), 51-55.

96. de Llasera, M. P. G.; Gómez-Almaraz, L.; Vera-Avila, L. E.; Peña-Alvarez, A., Matrix solid-phase dispersion extraction and determination by high-performance liquid chromatography with fluorescence detection of residues of glyphosate and aminomethylphosphonic acid in tomato fruit. *Journal of Chromatography A* **2005,** *1093* (1), 139-146.
97. Skeff, W.; Recknagel, C.; Schulz-Bull, D. E., The influence of salt matrices on the reversed-phase liquid chromatography behavior and electrospray ionization tandem mass spectrometry detection of glyphosate, glufosinate, aminomethylphosphonic acid and 2-aminoethylphosphonic acid in water. *Journal of Chromatography A* **2016,** *1475*, 64-73.
98. Gauglitz, G.; Wimmer, B.; Melzer, T.; Huhn, C., Glyphosate analysis using sensors and electromigration separation techniques as alternatives to gas or liquid chromatography. *Analytical and Bioanalytical Chemistry* **2018,** *410* (3), 725-746.
99. Zhang, Z.; Hamedpour, V.; Lyu, X.; Sasaki, Y.; Minami, T., A Printed Paper-Based Anion Sensor Array for Multi-Analyte Classification: On-Site Quantification of Glyphosate. *ChemPlusChem* **2021,** *86* (6), 798-802.
100. Hou, J.; Wang, X.; Lan, S.; Zhang, C.; Hou, C.; He, Q.; Huo, D., A turn-on fluorescent sensor based on carbon dots from Sophora japonica leaves for the detection of glyphosate. *Analytical Methods* **2020,** *12* (33), 4130-4138.
101. Ma, J.; Feng, G.; Ying, Y.; Shao, Y.; She, Y.; Zheng, L.; Abd Ei-Aty, A. M.; Wang, J., Sensitive SERS assay for glyphosate based on the prevention of L-cysteine inhibition of a Au–Pt nanozyme. *The Analyst* **2021,** *146* (3), 956-963.
102. Wang, L.; Bi, Y.; Hou, J.; Li, H.; Xu, Y.; Wang, B.; Ding, H.; Ding, L., Facile, green and clean one-step synthesis of carbon dots from wool: Application as a sensor for glyphosate detection based on the inner filter effect. *Talanta* **2016,** *160*, 268-275.
103. Guo, J.; Zhang, Y.; Luo, Y.; Shen, F.; Sun, C., Efficient fluorescence resonance energy transfer between oppositely charged CdTe quantum dots and gold nanoparticles for turn-on fluorescence detection of glyphosate. *Talanta* **2014,** *125*, 385-392.
104. Wiwasuku, T.; Boonmak, J.; Burakham, R.; Hadsadee, S.; Jungsuttiwong, S.; Bureekaew, S.; Promarak, V.; Youngme, S., Turn-on fluorescent probe towards glyphosate and Cr^{3+} based on Cd(II)-metal organic framework with Lewis basic sites. *Inorganic Chemistry Frontiers* **2021,** *8* (4), 977-988.
105. Savitz, J., The kynurenine pathway: a finger in every pie. *Molecular Psychiatry* **2020,** *25* (1), 131-147.

106. Gandhi, K.; Khan, S.; Patrikar, M.; Markad, A.; Kumar, N.; Choudhari, A.; Sagar, P.; Indurkar, S., Exposure risk and environmental impacts of glyphosate: Highlights on the toxicity of herbicide co-formulants. *Environmental Challenges* **2021,** *4*, 100149.

107. Neal, J. F.; Zhao, W.; Grooms, A. J.; Smeltzer, M. A.; Shook, B. M.; Flood, A. H.; Allen, H. C., Interfacial Supramolecular Structures of Amphiphilic Receptors Drive Aqueous Phosphate Recognition. *Journal of the American Chemical Society* **2019,** *141* (19), 7876-7886.

108. Kramer, J. J. P.; Nieger, M.; Bräse, S., Synthesis of Planar Chiral N-Heterocyclic-Substituted Pyridinophanes. *European Journal of Organic Chemistry* **2013,** *2013* (3), 541-549.

109. Warr, R. J.; Willis, A. C.; Wild, S. B., Inorganic Asymmetric Synthesis: Asymmetric Synthesis of a Two-Bladed Propeller, Octahedral Metal Complex. *Inorganic Chemistry* **2006,** *45* (21), 8618-8627.

110. Yuan, Z.; Kuang, G.-C.; Clark, R. J.; Zhu, L., Chemoselective Sequential “Click” Ligation Using Unsymmetrical Bisazides. *Organic Letters* **2012,** *14* (10), 2590-2593.

111. Echavarren, A.; Galán, A.; De Mendoza, J.; Salmerón, A.; Lehn, J.-M., Anion-receptor molecules: Synthesis of a chiral and functionalized binding subunit, a bicyclic guanidinium group derived from L- or D- asparagine. *Helvetica Chimica Acta* **1988,** *71* (4), 685-693.

112. Güthner, T.; Mertschenk, B., Cyanamides. In *Ullmann's Encyclopedia of Industrial Chemistry*, Wiley: Weinheim, **2006**.

113. Ayres, J. N. Exploring New Synthetic Routes Towards Cyanamides. Traditional Thesis, Cardiff University, **2018**.

114. Davies, M.; Jones, W. J., The infra-red spectrum and structure of cyanamide and dimethylcyanamide. *Transactions of the Faraday Society* **1958,** *54*, 1454.

115. Denner, L.; Luger, P.; Buschmann, J., X-ray structure of cyanamide at 108 K. *Acta Crystallographica Section C Crystal Structure Communications* **1988,** *44* (11), 1979-1981.

116. Yuan, M.; Gao, S.; Sun, H.-L.; Su, G., An Antiferromagnetic Mn(III) Chain Bridged by Hydrogencyanamide: $[\mathrm{Mn^{III}}(5\text{-Brsalen})(\mu_{1,3}\text{-NCNH})]_n$. *Inorganic Chemistry* **2004,** *43* (26), 8221-8223.

117. Yuan, M.; Zhao, F.; Zhang, W.; Pan, F.; Wang, Z.-M.; Gao, S., Hydrogencyanamide-Bridged One-Dimensional Polymers Built on $\mathrm{Mn^{III}}$–Schiff Base Fragments:

Synthesis, Structure, and Magnetism. *Chemistry - A European Journal* **2007,** *13* (10), 2937-2952.

118. Liu, X.; Krott, M.; Müller, P.; Hu, C.; Lueken, H.; Dronskowski, R., Synthesis, Crystal Structure, and Properties of MnNCN, the First Carbodiimide of a Magnetic Transition Metal. *Inorganic Chemistry* **2005,** *44* (9), 3001-3003.
119. Herlitschke, M.; Tchougréeff, A. L.; Soudackov, A. V.; Klobes, B.; Stork, L.; Dronskowski, R.; Hermann, R. P., Magnetism and lattice dynamics of FeNCN compared to FeO. *New Journal of Chemistry* **2014,** *38* (10), 4670-4677.
120. Bineau, M. A., On some ammoniacal combinations and on the role that ammonia plays in the reactions chemicals. In *Annales de chimie et de physique*, V. Masson: Paris, **1838;** Vol. 67, pp 225 - 272.
121. Cloez, S.; Cannizzaro, S., Research on Cranial Amides. In *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, Gauthier-VillarsGauthier-Villars: Paris, **1851;** Vol. 32, pp 62 - 64.
122. Frank, A.; Caro, N. Improvements relating to the Production of Cyanogen Compounds and to the Preparation of Material for Use therein. Patent number 88363, **1895**.
123. Mcmillan, A., Pure Cyanamide. *Journal of Industrial & Engineering Chemistry* **1918,** *10* (6), 487-487.
124. Wehrstedt, K.-D.; Wildner, W.; Güthner, T.; Holzrichter, K.; Mertschenk, B.; Ulrich, A., Safe transport of cyanamide. *Journal of Hazardous Materials* **2009,** *170* (2), 829-835.
125. Cameron, W., Cyanamides. In *Kirk-Othmer Encyclopedia of Chemical Technology*, John Wiley & Sons, Inc.: Hoboken, **2010;** pp 1 - 15.
126. Liang, D.; Huang, X.; Shen, Y.; Shen, T.; Zhang, H.; Lin, L.; Wang, J.; Deng, Q.; Lyu, X.; Xia, H., Hydrogen cyanamide induces grape bud endodormancy release through carbohydrate metabolism and plant hormone signaling. *BMC Genomics* **2019,** *20* (1).
127. Wang, H.; Xia, X.; An, L., Metabolomics Analysis Reveals the Mechanism of Hydrogen Cyanamide in Promoting Flower Bud Break in Blueberry. *Agronomy* **2021,** *11* (1), 102.
128. Sapkota, S.; Liu, J.; Islam, M. T.; Sherif, S. M., Changes in Reactive Oxygen Species, Antioxidants and Carbohydrate Metabolism in Relation to Dormancy Transition and Bud Break in Apple (Malus × domestica Borkh) Cultivars. *Antioxidants* **2021,** *10* (10).

129. Tang, L.; Chhajed, S.; Vashisth, T.; Olmstead, M. A.; Olmstead, J. W.; Colquhoun, T. A., Transcriptomic Study of Early Responses to the Bud Dormancy-breaking Agent Hydrogen Cyanamide in 'TropicBeauty' Peach. *Journal of the American Society for Horticultural Science* **2019,** *144* (4), 244-256.

130. Lin, S.-Y.; Agehara, S., Foliar Application of Defoliants before Winter Chill Accumulation Advances Budbreak and Improves Fruit Earliness of Blackberry under Subtropical Climatic Conditions. *HortScience* **2021,** *56* (2), 210-216.

131. Sassine, Y. N.; Al-Turk, S. M.; El-Sebaaly, Z.; Bachour, L.; El-Masri, I. Y., Finding alternatives for Dormex (Hydrogen Cyanamid) as dormancy breaking agent. *Fresenius Environmental Bulletin* **2019,** *28* (12A), 10214-10224.

132. Schep, L.; Temple, W.; Beasley, M., The adverse effects of hydrogen cyanamide on human health: an evaluation of inquiries to the New Zealand National Poisons Centre. *Clinical Toxicology* **2009,** *47* (1), 58-60.

133. Lin, S.-Y.; Agehara, S., Budbreak patterns and phytohormone dynamics reveal different modes of action between hydrogen cyanamide- and defoliant-induced flower budbreak in blueberry under inadequate chilling conditions. *PLOS ONE* **2021,** *16* (8), e0256942.

134. Sharif, A. F.; Fayed, M. M., Evaluation of Multiple Organ Dysfunction Score (MODS) and the Sequential Organ Failure Assessment (SOFA) score as in-hospital outcome predictors among cases of hydrogen cyanamide exposure: a cross-sectional study. *Environmental Science and Pollution Research* **2021,** *28* (31), 42161-42176.

135. Kamo, T.; Takemura, T.; Wasano, N.; Fujii, Y.; Hiradate, S., Quantification of Cyanamide in Young Seedlings of ViciaSpecies, Lens culinaris, and Robinia pseudo-acacia by Gas Chromatography-Mass Spectrometry. *Bioscience, Biotechnology, and Biochemistry* **2012,** *76* (7), 1416-1418.

136. Turowski, M.; Deshmukh, B., Direct Chromatographic Method for Determination of Hydrogen Cyanamide and Dicyandiamide in Aqueous Solutions. *Analytical Letters* **2004,** *37* (9), 1981-1989.

137. Kamo, T.; Sato, M.; Kato, K.; Hiradate, S.; Nakajima, E.; Fujii, Y.; Hirota, M., Quantification of Cyanamide Contents in Herbaceous Plants. *Bioscience, Biotechnology, and Biochemistry* **2006,** *70* (9), 2310-2312.

138. Hiradate, S.; Kamo, T.; Nakajima, E.; Kato, K.; Fujii, Y., Direct quantitative determination of cyanamide by stable isotope dilution gas chromatography–mass spectrometry. *Journal of Chromatography A* **2005,** *1098* (1-2), 138-143.

139. Australian Environment Agency, Hydrogen Cyanamide: New Zealand Environmental and Human Health Risk Assessment Report to Support Reassessment. **2020**, 1-40.

140. Arod, F.; Pattison, P.; Schenk, K. J.; Chapuis, G., Polymorphism in *N*-Salicylideneaniline Reconsidered. *Crystal Growth & Design* **2007,** *7* (9), 1679-1685.

141. Steiner, T., The Hydrogen Bond in the Solid State. *Angewandte Chemie International Edition* **2002,** *41* (1), 48-76.

142. Mason, H. E.; Yates, J. L. R.; Potts, R. J.; Gutmann, M. J.; Howard, J. A. K.; Sparkes, H. A., Structural studies of *N*-(methoxysalicylidene)-fluoroaniline, *N*-(methoxysalicylidene)-chloroaniline and *N*-(methoxysalicylidene)-bromoaniline derivatives. *Acta Crystallographica Section B Structural Science, Crystal Engineering and Materials* **2021,** *77* (6), 974-980.

143. Zhang, J.; Richardson, H. W., Copper Compounds. In *Ullmann's Encyclopedia of Industrial Chemistry*, Wiley: Weinheim, **2016;** pp 1-31.

144. Lossin, A.; Westhoff, F.-J., The production and application of cuprous oxide and cupric hydroxide. *JOM* **1997,** *49* (10), 38-39.

145. Arena, M.; Auteri, D.; Barmaz, S.; Bellisai, G.; Brancato, A.; Brocca, D.; Bura, L.; Byers, H.; Chiusolo, A.; Court Marques, D.; Crivellente, F.; De Lentdecker, C.; Egsmose, M.; Erdos, Z.; Fait, G.; Ferreira, L.; Goumenou, M.; Greco, L.; Ippolito, A.; Istace, F.; Jarrah, S.; Kardassi, D.; Leuschner, R.; Lythgo, C.; Magrans, J. O.; Medina, P.; Miron, I.; Molnar, T.; Nougadere, A.; Padovani, L.; Parra Morte, J. M.; Pedersen, R.; Reich, H.; Sacchi, A.; Santos, M.; Serafimova, R.; Sharp, R.; Stanek, A.; Streissl, F.; Sturma, J.; Szentes, C.; Tarazona, J.; Terron, A.; Theobald, A.; Vagenende, B.; Verani, A.; Villamar-Bouza, L., Peer review of the pesticide risk assessment of the active substance copper compounds copper(I), copper(II) variants namely copper hydroxide, copper oxychloride, tribasic copper sulfate, copper(I) oxide, Bordeaux mixture. *EFSA Journal* **2018,** *16* (1).

146. Dash, P. P.; Patel, D. A.; Mohanty, P.; Behura, R.; Behera, S.; Sahoo, S. K.; Jali, B. R., Advances on chromo-fluorogenic sensing of copper(II) with Schiff bases. *Inorganica Chimica Acta* **2023,** *556*, 121635.

147. Kashyap, P.; Sharma, P.; Gohil, R.; Rajpurohit, D.; Mishra, D.; Shrivastav, P. S., Progress in appended calix[4]arene-based receptors for selective recognition of copper ions. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **2023,** *303*, 123188.

148. Loya, M.; Ghosh, S.; Atta, A. K., A review on dual detection of Cu^{2+} and Ni^{2+} ions by using single fluorometric and colorimetric organic molecular probes. *Journal of Molecular Structure* **2023,** *1278,* 134949.

149. Scheiber, I. F.; Mercer, J. F. B.; Dringen, R., Metabolism and functions of copper in brain. *Progress in Neurobiology* **2014,** *116,* 33-57.

150. Karimi, B.; Masson, V.; Guilland, C.; Leroy, E.; Pellegrinelli, S.; Giboulot, E.; Maron, P.-A.; Ranjard, L., Ecotoxicity of copper input and accumulation for soil biodiversity in vineyards. *Environmental Chemistry Letters* **2021,** *19* (3), 2013-2030.

151. Van Zwieten, L.; Rust, J.; Kingston, T.; Merrington, G.; Morris, S., Influence of copper fungicide residues on occurrence of earthworms in avocado orchard soils. *Science of The Total Environment* **2004,** *329* (1-3), 29-41.

152. Grosell, M.; Blanchard, J.; Brix, K. V.; Gerdes, R., Physiology is pivotal for interactions between salinity and acute copper toxicity to fish and invertebrates. *Aquatic Toxicology* **2007,** *84* (2), 162-172.

153. Rajasekar, M.; Geetha Sree Agash, S.; Narendran, C.; Rajasekar, K., Recent trends in fluorescent-based copper (II) chemosensors and their biomaterial applications. *Inorganic Chemistry Communications* **2023,** *151*, 110609.

154. Yan, P.; Duan, G.; Ji, R.; Ge, Y., A simple and efficient synthesis of new fluorophore 4-hydroxy pyrazolo[1,5-a]pyridines through a tandem reaction. *Tetrahedron Letters* **2018,** *59* (25), 2426-2429.

Appendix 1 – Copper Hydroxide

A1.1 Copper Hydroxide Background

Copper hydroxide [$Cu(OH)_2$] is very versatile. It can be used as a fungicide, as a precursor in reactions requiring copper(II), in the production of rayon fabric and used to stabilise nylon.[143,144]

Copper hydroxide-based fungicides are popular for use in orchards and vineyard as they are not harmful in moderation. This is because copper hydroxide separates into copper(II) ions and hydroxide ions in the environment. Copper(II) ions are vital to many cellular processes such as cellular respiration, bone formation, the synthesis of connective tissue proteins and neurotransmitters, and they are used as cofactors in enzymes.[145,146]

Copper(II) ions are vital to many cell processes and it can also cause a plethora of diseases if there is an imbalance in the body.[147] Disruptions to copper(II) homeostasis in humans cause serious neurodegenerative diseases like Parkinson's disease, Menkes and Wilson’s diseases, asthma, Alzheimer’s disease, central nervous system disorders and Prion diseases.[146-149]

Elevated levels of copper(II) can also be harmful to terrestrial and aquatic invertebrates, and small fish. This can lead to reduced ecological biodiversity in orchards and vineyards containing excessive levels of copper(II).[150-152]

The potential toxicity of copper(II) leads to the need to monitor the levels of copper(II) in the environment. The traditional analytical methods that have high sensitivity like atomic absorption spectroscopy and inductively coupled plasma emission spectroscopy, have significant drawbacks. The drawbacks include time-consuming methods, the need for skilled analysts, expensive instruments and complicated sample preparations.[146,148] This caused colourimetric and fluorometric probes to gain more attention in recent years, as they are cheaper, faster and easier, but lack in sensitivity compared to traditional methods.[146-148]

A1.2 Initial Scope

The initial scope of this project was to create a ligand that could be used to detect copper(II) in environmental samples selectively, with a strong response. The ligand would also have innate spectroscopic properties that are easily monitored and would change noticeably when the ligand binds to copper(II) resulting in an easy and efficient method to detect copper(II).

Fluorescence spectroscopy shows promise as it is a simple, fast technique that can be sensitive if the ligands are tailored correctly. If the ligand has a strong fluorophore, then when the ligands bind to copper, there will be a detectable shift in the fluorescence of the fluorophore. This would allow the detection of copper(II) ion in environmental samples without the need for pre-treatment of the samples.

Fluorescent probes often have the metal binding site and the fluorophore joined together using a linking unit.[146,147] Other probes are a turn-off fluorescent sensor with the binding site splitting the fluorophore when the metal is bound.[153] The ligand (**44**) as shown in Figure A1.1, was designed to have the metal binding site and fluorophore joint together. The metal binding site chosen was based on salicylaldoxime, which is known to strongly bind to copper. The fluorophore is based on work reported by Yan *et al*.[154] This fluorophore was chosen as it has a simple one step reaction from purchasable materials and Yan *et al*. showed that slight variations can change the fluorescence output, suggesting the fluorophore can be tuneable.

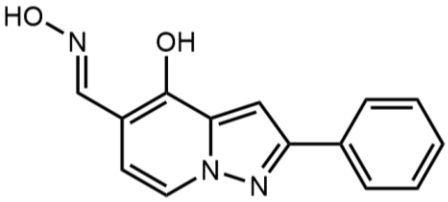


Figure A1.1: The ligand sensor (44) designed to coordinate to copper (II).

The plan for this section was to order the precursors required to synthesise the fluorophore, followed by functional group conversions to give the final ligand (**44**). However, one of the precursors was not available and had to be synthesised from available chemicals.

A1.3 Completed Work

The synthesis of the ligand can be broken into two parts, synthesis of the fluorophore and the functional group conversions as seen in Scheme A1.2 and Scheme A1.3. The first step of the synthesis was protection of oxalic acid, by dissolving oxalic acid in ethanol with trace amounts of concentrated sulfuric acid and refluxing the solution for 4 hours. This produced equimolar amounts of diethyl oxalate (**45**). This was confirmed by ^{1}H NMR (Figure A1.2), shown in with two peaks lining up with expected chemical shift and integrations.

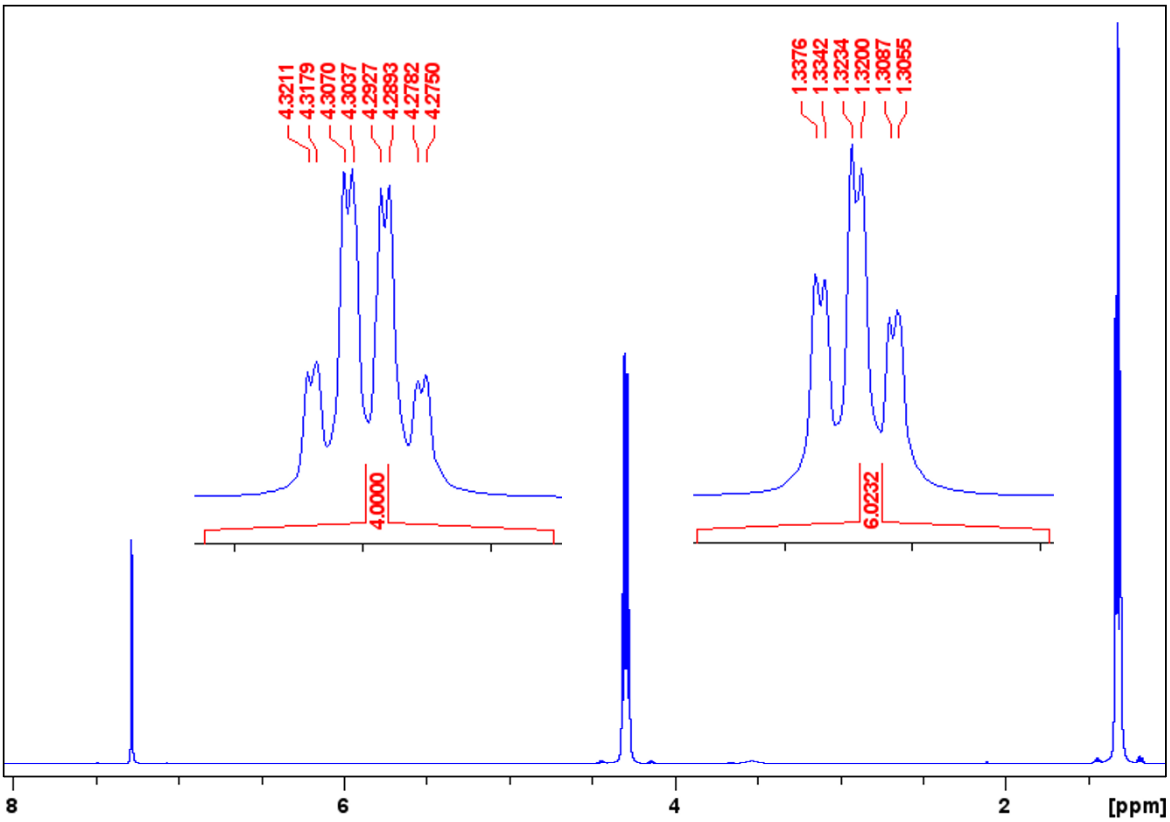
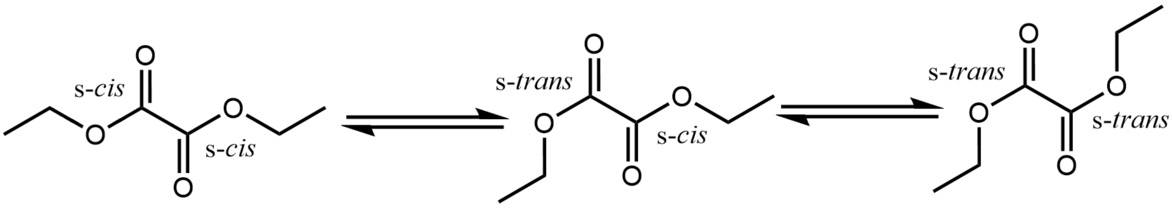


Figure A1.2: ^{1}H NMR of diethyl oxalate (45).

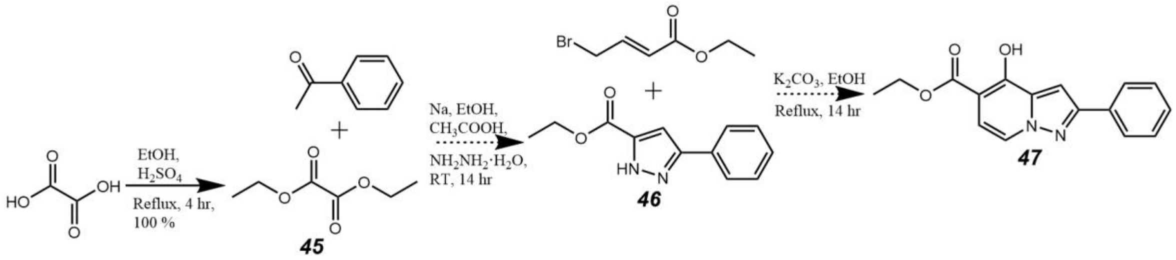
The splitting patterns of the ^{1}H NMR peaks of diethyl oxalate (**45**) are not as expected. The "-CH_3" peak showed up as a triplet of doublets and the "-CH_2-" peak showed up as a quartet of doublets. The chloroform peak was a singlet meaning that smaller splitting was not spectra wide but rather confined to the diethyl oxalate peaks. The coupling constant between the "-CH_3" and the "-CH_2-" peaks is 7.15 Hz, which is expected. The second coupling constant, 1.7 Hz, is unexpected and is hypothesised to be due to the potential for s-*cis* and s-*trans* conformers splitting the ^{1}H NMR peaks, as shown in Scheme A1.1.



Scheme A1.1: Hypothesised geometric isomers causing a second coupling constant in the ^{1}H NMR of diethyl oxalate (45)

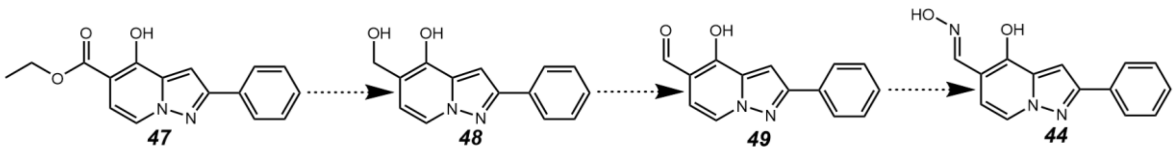
Due to time constraints, no further reactions were attempted. However, the second reaction (Scheme A1.2) to form the fluorophore would have been achieved by reacting sodium metal with diethyl oxalate (**45**) in ethanol, before adding acetophenone. Finally, addition of hydrazine hydrate would lead to the formation of ethyl 3-phenyl-1*H*-pyrazole-5-carboxylate (**46**).

The final reaction to give the fluorophore (**47**) would have been achieved by dissolving pyrazole (**46**), ethyl 4-bromobut-2-enoate and potassium carbonate in ethanol and refluxing for 14 hours.



Scheme A1.2: Synthesis of the fluorophore

The final three reactions, as shown in Scheme A1.3, were a reduction of the ester to an alcohol (**48**), oxidation to give the aldehyde (**49**), and then aldoxime formation reaction to give the ligand (**44**).



Scheme A1.3: Functional group conversions to give the ligand (44)

Appendix 2 – Experimental Details

A2.1 Chapter 3 experimental details

O,O'-dibenzyl S,S'-((((1E,1'E)-((piperazine-1,4-diylbis(methylene))bis(5-(tert-butyl)-2-hydroxy-3,1-phenylene))bis(methaneylylidene))bis(azaneylylidene))bis(4,1-phenylene)) bis(carbonothioate) (**1$^{\mathbf{A}}$**)

Both **6$^{\mathbf{A}}$** (222.9 mg, 0.860 mmol) and **8** (153.6 mg, 0.344 mmol) were dissolved together in anhydrous toluene (10 mL) and set to reflux for 72 hrs. The solvent was removed under vacuum and the resulting pale-yellow solid was recrystallised twice by dissolving in a minimum amount of chloroform and inducing precipitation with the addition of hot ethanol, which produced **1$^{\mathbf{A}}$**, as a pale-yellow solid (199.6 mg, 0.210 mmol, 61.0 %). ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.72 (s, 2H), δ 7.59 (d, J = 8.4 Hz, 4H), δ 7.41 (s, 10H), δ 7.30 (d, J = 8.4 Hz, 4H), δ 5.30 (s, 4H), δ 3.75 (s, 4H), δ 2.69 (broad, 8H), δ 1.37 (s, 18H); MS: m/z (ESI) 949.28 $[M+H]^+$.

6,6'-(piperazine-1,4-diylbis(methylene))bis(2-((E)-((4-(benzylthio)phenyl)imino)methyl)-4-(tert-butyl)phenol) (**1$^{\mathbf{B}}$**)

Both **6$^{\mathbf{B}}$** (202.5 mg, 0.941 mmol) and **8** (168.0 mg, 0.376 mmol) were dissolved together in anhydrous toluene (10 mL) and set to reflux for 72 hrs. The solvent was removed under vacuum and the resulting pale-yellow solid was recrystallised twice by dissolving in a minimum amount of chloroform and inducing precipitation with the addition of hexane, which when left open for the chloroform to slowly evaporate off (not completely) produced **1$^{\mathbf{B}}$**, as a yellow solid (202.4 mg, 0.235 mmol, 61.4 %). ^{1}H NMR (500 MHz, $CDCl_3$): δ 13.30 (s, 2H), δ 8.68 (s, 2H), δ 7.31 (m, 26H), δ 4.15 (s, 4H), δ 3.73 (broad, 4H), δ 2.68 (broad, 8H), δ 1.65 (broad, 2H), δ 1.35 (s, 16H), δ 1.31 (s, 2H), δ 1.26 (s, 4H); MS: m/z (ESI) 861.37 $[M+H]^+$.

4-Nitrobenzyl chloride (**2**)

Benzyl chloride (4.00 mL, 28.728 mmol) was cooled to 0 °C. A mixture of nitric acid (1.29 mL) and sulfuric acid (1.62 mL) was cooled to 0 °C before being slowly added to the benzyl chloride. The solution was then slowly warmed to room temperature and then left stirring for 2 hours. Water (50 mL) was then added and the crude product was extracted using ethyl acetate. The product was then purified via column chromatography, using ethyl acetate : hexane (1:19). This produced **2** (429.2 mg, 2.510 mmol, 8.7 %). ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.23 (dd, J = 6.9 and 1.95 Hz, 2H), δ 7.59 (d, J = 8.7 Hz, 2H), δ 4.67 (s, 2H).

S-(4-Nitrophenyl) *O*-(phenylmethyl) carbonothioate (**5^{A}**)

4-Nitrothiophenol (312.5 mg, 80% pure, technical grade) was purified using flash chromatography by dry loading onto silica and eluting with 1 : 19 ethyl acetate : n-hexane. The first yellow band was collected, the solvent removed and dried under high vacuum to give 4-Nitrothiophenol (250 mg, 1.611 mmol) as a pale-yellow solid. The purified 4-Nitrothiophenol was dissolved in anhydrous THF (5 mL), cooled to 0 °C and placed under an argon atmosphere. Sodium hydride (96.7 mg of a 60 % dispersion in mineral oil, 2.417 mmol) was added slowly causing the pale-yellow solution to turn red. Benzyl chloroformate (0.344 mL, 2.417 mmol) was injected in one portion causing the solution to turn a yellow/grey colour. The reaction was left to stir at room temperature overnight. Acetone (1 mL) was added and the reaction mixture poured onto an ethyl acetate / 2 M hydrochloric acid / ice partition. The organic layer was removed, washed with hydrochloric acid (2M, 2x 10 mL), water (2x 10 mL), saturated sodium bicarbonate (10 mL), dried with magnesium sulphate, filtered and the solvent removed. The solid was purified by recrystallisation with ethanol, giving o-(phenylmethyl) S-(4-nitrophenyl) carbonothioate (**5^{A}**) as a pale-yellow solid (444.4 mg, 1.536 mmol, 95.3 %). ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.24 (d, J = 8.85 Hz, 2H), δ 7.74 (d, J = 8.85 Hz, 2H), δ 7.42 (s, 5H), δ 5.33 (s, 2H).

Benzyl(4-nitrophenyl)sulfane (**5^{B}**)

4-Nitrothiophenol (2919.5 mg, 80% pure, technical grade) was purified using flash chromatography by dry loading onto silica and eluting with 1 : 19 ethyl acetate : n-hexane. The first yellow band was collected, the solvent removed and dried under high vacuum to give 4-Nitrothiophenol (2335.6 mg, 15.052 mmol) as a pale-yellow solid. The purified 4-Nitrothiophenol was dissolved in anhydrous THF (50 mL), cooled to 0 °C and placed under an argon atmosphere. Sodium hydride (903.1 mg of a 60 % dispersion in mineral oil, 22.578 mmol) was added slowly causing the pale-yellow solution to turn red. Benzyl chloride (6.287 mL, 22.578 mmol, 50% solution in toluene) was injected in one portion causing the solution to turn a yellow/grey colour. The reaction was left to stir at room temperature overnight. Acetone (10 mL) was added and the reaction mixture poured onto an ethyl acetate / 2 M hydrochloric acid / ice partition. The organic layer was removed, washed with hydrochloric acid (2M, 2x 50 mL), water (2x 50 mL), saturated sodium bicarbonate (50 mL), dried with magnesium sulphate, filtered and the solvent removed. The solid was purified by recrystallisation with ethanol, giving **5^{B}** as a pale-yellow solid (3492.8 mg, 14.239 mmol, 94.6 %). ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.11 (d, J = 8.85 Hz, 2H), δ 7.41 (d, J = 7.25 Hz, 2H), δ 7.37 (d, J = 7.5 Hz, 4H) , δ 7.32 (d, J = 7.05 Hz, 1H), δ 4.29 (s, 2H).

S-(4-Aminophenyl) *O*-(phenylmethyl) carbonothioate (**6^A**)

To a suspension of **5^A** (444.4 mg, 1.536 mmol) in ethanol : water : glacial acetic acid (10 mL : 10 mL : 5 mL) iron powder (428.9 mg, 7.680 mmol) was added and sonicated. The reaction was monitored by TLC (1:4 ethyl acetate : n-hexane and visualised under UV light) for completion. Upon completion the reaction mixture was filtered and water (20 mL) was added to the filtrate until no more precipitation was observed. The precipitate was extracted with ethyl acetate (2x 20 mL), washed with saturated sodium bicarbonate (20 mL), brine (3x 20 mL), dried with magnesium sulphate, filtered and the solvent removed. The crude solid was recrystallised from ethanol producing o-(phenylmethyl) S-(4-aminophenyl) carbonothioate (**6^A**), as a pale-yellow solid (222.9 mg, 0.860 mmol, 56.0 %). ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.39 (s, 5H), δ 7.31 (d, J = 8.6 Hz, 2H), δ 6.67 (d, J = 8.6 Hz, 2H), δ 5.26 (s, 2H), δ 3.48 (s, 2H); MS: m/z (ESI) 260.16 $[M+H]^+$.

Benzyl(4-aminophenyl)sulfane (**6^B**)

To a suspension of **5^B** (300 mg, 1.230 mmol) in ethanol : water : glacial acetic acid (9 mL : 9 mL : 4.5 mL) iron powder (343.4 mg, 6.148 mmol) was added and sonicated. The reaction was monitored by TLC (1:4 ethyl acetate : n-hexane and visualised under UV light) for completion. Upon completion the reaction mixture was filtered and water (24 mL) was added to the filtrate until no more precipitation was observed. The precipitate was extracted with ethyl acetate (2x 24 mL), washed with saturated sodium bicarbonate (24 mL), brine (3x 24 mL), dried with magnesium sulphate, filtered and the solvent removed. The crude solid was recrystallised from ethanol producing **6^B**, as a pale-yellow solid (159.4 mg, 0.740 mmol, 60.2 %). ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.28 (t, J = 7.6, 5H), δ 7.24 (d, J = 7.1 Hz, 1H), δ 7.20 (d, J = 7.1 Hz, 2H), δ 7.16 (d, J = 8.4 Hz, 2H), δ 6.59 (d, J = 8.4 Hz, 2H), δ 3.97 (s, 2H), δ 3.64 (b, 2H); MS: m/z (ESI) 216.03 $[M+H]^+$.

3-(Bromomethyl)-5-*tert*-Butyl-2-hydroxybenzaldehyde (**7**)

5-*tert*-Butyl-2-hydroxybenzaldehyde (3414.7 mg, 191.587 mmol), hydrogen bromide (77.5 mL, 1436.90 mmol of a 48% solution) and paraformaldehyde (8630.0 mg, 287.381 mmol) were mixed together at 0 °C for 1 hour. Concentrated sulfuric acid (3.4 mL) was then added to the solution dropwise and the solution was then refluxed for 20 hours at 70 °C. Cool to room temperature and add water (125 mL), then extract the product with dichloromethane (3x 125 mL). The combined organic solution was dried with magnesium sulfate, filtered, followed by removing the solvent under reduced pressure. Giving the crude product which was purified by recrystallisation in pentane (10 % w/v). This gave **7** (32662.2

mg, 120.458 mmol, 62.9 %). ^{1}H NMR (500 MHz, $CDCl_3$): δ 11.32 (s, 1H), δ 9.86 (s, 1H), δ 7.66 (d, J = 2.4 Hz, 1H) , δ 7.53 (d, J = 2.5 Hz, 1H), δ 4.57 (s, 2H), δ 1.33 (s, 9H).

3,3′-(1,4-Piperazine-bis-(methylene)-bis-(5-tert-butyl)-2-hydroxybenzaldehyde) (**8**)

Triethyl amine (5.14 mL, 36.880 mmol) was dissolved in dichloromethane (30 mL). To this solution, a solution of piperazine (0.794 mg, 9.220 mmol) dissolved in anhydrous dichloromethane (30 mL) was added followed by adding a solution of **7** (5.00 g, 18.440 mmol) dissolved in anhydrous dichloromethane (30 mL) dropwise over 2 hours. After which the reaction was left to reflux overnight. The reaction mixture was cooled, then washed with water (3x 50 mL), dried with magnesium sulphate, filtered and the solvent removed. The crude product was then recrystallised by dissolving in the minimum amount of hot chloroform and adding boiling ethanol until precipitation is induced, producing **8** as a white powder (4.257 mg, 9.130 mmol, 99.0 %). ^{1}H NMR (500 MHz, $CDCl_3$): δ 10.25 (s, 2H), δ 7.61 (d, J = 2.05 Hz, 2H), δ 7.39 (d, J = 2.05 Hz, 2H), δ 3.73 (s, 4H), δ 2.67 (broad, 8H), δ 1.31 (s, 18H); MS: m/z (ESI) 467.23 $[M+H]^+$.

Nanoparticle syntheses in thesis

Aqueous gold seed particles

Chloroauric acid (17.4 mL of a 1 mM solution) was boiled. Citrate (2.3 mL of a 1 % solution) was added and the solution was stirred vigorously stirred while boiling for 15 minutes, ensuring the volume remained constant. Cool the solution and store in the fridge. The seed particles were analysed using the TEM.

Aqueous gold nanostars

Aqueous gold seed particles (100 µL) was added to a stirred solution of chloroauric acid (10 mL of a 0.25 mM solution) with hydrochloric acid (10 µL of a 1 M solution). Silver nitrate (100 µL of a 205.2 mM solution) was add to the solution simultaneously with ascorbic acid (50 µL of a 100 mM solution). The solution was then stirred for 30 seconds before centrifuging the solution at 3000-5000 rcf. for 15 minutes. The pallet was then resuspended in deionised water. This was unsuccessful as the nanoparticles aggregated very quickly.

PVP stabilised gold seed particles

A solution of polyvinylpyrrolidone (0.454 mL, final concentration equalling 60 subunits per nm^2 of gold nanoparticles) was added dropwise to the aqueous gold seed particles (9.545 mL). The reaction was stirred overnight and then centrifuged for 90 minutes at 4000 rpm. The pellet was then resuspended in ethanol.

PVP stabilised gold nanostars

Chloroauric acid (82 µL of 50 mM solution) was mixed with polyvinylpyrrolidone in dimethyl formamide (15 mL of a 10 mM solution). After the band at 325 nm on the UV/vis spectra has diminished, the PVP stabilised gold seed particles (0.334 mL) were added and the solution was stirred until no further change on the UV/vis spectra is observed.

No changes were seen before the seed particles were added for an hour on the first run so then I added the seeds and no changes occurred still, so I sonicated it and monitored it. After many hours of sonication, the particles continued to change their UV/vis spectra and then something precipitated out with around 12 hours of sonication. The second I sonicated for hours before adding the seeds and the band at 325 nm diminished as a peak arose around 280ish nm and then something precipitated out again at around the 12-hour mark.

Silver nanoprisms

Mix sodium citrate (2 mL of a 12.5 mM solution) with silver nitrate (5 mL of a 0.375 mM solution), hydrogen peroxide (5 mL of a 50 mM solution) and potassium bromide (0, 20, 25, or 40 µL of a 1 mM solution). Add sodium borohydride (2.5 mL of a fresh 5 mM solution) to the solution and mix the solution, then let it rest as the nanoparticles grow.

The best ones for me have 25 or 40 µL of KBr (orange/yellow)

$[Cu_2(\mathbf{1}^{\mathbf{A}})_2](BF_4)_4$

$\mathbf{1}^{\mathbf{A}}$ (81.6 mg, 0.086 mmol) and copper tetrafluoroborate monohydrate (22.0 mg, 0.086 mmol) were suspended in 9 : 1 methanol : dichloromethane (4.5 mL : 0.5 mL). Dichloromethane was added till all reagents were dissolved. The reaction was then stirred at 40 °C overnight. The solution was then cooled and the solvent removed under vacuum. The crude material was recrystallised by vapour diffusion of *n*-pentane into a solution of the green crude material dissolved in tetrahydrofuran.

$[Cu_2(\mathbf{1}^{\mathbf{A}})_2](NO_3)_4$

$\mathbf{1}^{\mathbf{A}}$ (81.6 mg, 0.086 mmol) and copper nitrate trihydrate (20.8 mg, 0.086 mmol) were suspended in 9 : 1 methanol : dichloromethane (4.5 mL : 0.5 mL). Dichloromethane was added till all reagents were dissolved. The reaction was then stirred at 40 °C overnight. The solution was then cooled and the solvent removed under vacuum. The crude material was recrystallised by vapour diffusion of diethyl ether into a solution of the green crude material dissolved in dichloromethane.

$[Cu_2(\mathbf{1}^{\mathbf{A}})_2]Br_4$

$\mathbf{1}^{\mathbf{A}}$ (81.6 mg, 0.086 mmol) and copper bromide (19.2 mg, 0.086 mmol) were suspended in 9 : 1 methanol : dichloromethane (4.5 mL : 0.5 mL). Dichloromethane was added till all reagents were dissolved. The reaction was then stirred at 40 °C overnight. The solution was then cooled and the solvent removed under vacuum. The crude material was recrystallised by vapour diffusion of *n*-pentane into a solution of the green crude material dissolved in chloroform.

$[Cu_2(\mathbf{1}^{\mathbf{A}})_2](CH_3COO)_4$

$\mathbf{1}^{\mathbf{A}}$ (81.6 mg, 0.086 mmol) and copper acetate monohydrate (17.2 mg, 0.086 mmol) were suspended in 9 : 1 methanol : dichloromethane (4.5 mL : 0.5 mL). Dichloromethane was added till all reagents were dissolved. The reaction was then stirred at 40 °C overnight. The solution was then cooled and the solvent removed under vacuum. The crude material was recrystallised by vapour diffusion of diethyl ether into a solution of the green crude material dissolved in dichloromethane.

$[Cu_2(\mathbf{1}^{\mathbf{B}})_2](BF_4)_4$

$\mathbf{1}^{\mathbf{B}}$ (50.0 mg, 0.058 mmol) and copper tetrafluoroborate monohydrate (13.1 mg, 0.058 mmol) were suspended in 9 : 1 methanol : dichloromethane (4.5 mL : 0.5 mL). Dichloromethane was added till all reagents were dissolved. The reaction was then stirred at 40 °C overnight. The solution was then cooled and the solvent removed under vacuum. A purification was attempted unsuccessfully by crystallization with vapour diffusion of *n*-pentane into a solution of the product dissolved in tetrahydrofuran.

$[Cu_2(\mathbf{1}^{\mathbf{B}})_2](NO_3)_4$

$\mathbf{1}^{\mathbf{B}}$ (50.0 mg, 0.058 mmol) and copper nitrate trihydrate (14.0 mg, 0.058 mmol) were suspended in 9 : 1 methanol : dichloromethane (4.5 mL : 0.5 mL). Dichloromethane was added till all reagents were dissolved. The reaction was then stirred at 40 °C overnight. The solution was then cooled and the solvent removed under vacuum. A purification was attempted unsuccessfully by crystallization with vapour diffusion of diethyl ether into a solution of the product dissolved in dichloromethane.

$[Cu_2(\mathbf{1}^{\mathbf{B}})_2]Br_4$

$\mathbf{1}^{\mathbf{B}}$ (50.0 mg, 0.058 mmol) and copper bromide (13.0 mg, 0.058 mmol) were suspended in 9 : 1 methanol : dichloromethane (4.5 mL : 0.5 mL). Dichloromethane was added till all reagents were dissolved. The reaction was then stirred at 40 °C overnight. The solution was then cooled and the solvent removed under vacuum. A purification was attempted unsuccessfully by crystallization with vapour diffusion of *n*-pentane into a solution of the product dissolved in chloroform.

$[Cu_2(\mathbf{1^B})_2](CH_3COO)_4$

$\mathbf{1^B}$ (50.0 mg, 0.058 mmol) and copper acetate monohydrate (11.6 mg, 0.058 mmol) were suspended in 9 : 1 methanol : dichloromethane (4.5 mL : 0.5 mL). Dichloromethane was added till all reagents were dissolved. The reaction was then stirred at 40 °C overnight. The solution was then cooled and the solvent removed under vacuum. A purification was attempted unsuccessfully by crystallization with vapour diffusion of diethyl ether into a solution of the product dissolved in dichloromethane.

A2.2 Chapter 4 experimental details

5-(*tert*-Butyl)-2-hydroxybenzaldehyde oxime (**9**)

Hydroxylammonium chloride (97.5 mg, 1.403 mmol) was dissolved in ethanol (15 mL). A solution of potassium hydroxide dissolved in ethanol (15 mL) was added, while mixing. The precipitation was filtered off and the filtrate was then added dropwise to a solution of 5-*tert*-Butyl-2-hydroxybenzaldehyde dissolved in ethanol (15 mL) and stirred at room temperature overnight. The solvent was then removed under vacuum to produce a brown oil. The oil was then dissolved in chloroform, washed with water, dried with magnesium sulphate, filtered and the solvent removed under vacuum. This produced the 5-(*tert*-Butyl)-2-hydroxybenzaldehyde oxime (**9**, 163.5 mg, 0.846 mmol, 60.1 %) as a yellow oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ 9.83 (s, 1H), δ 8.25 (s, 1H), δ 7.80 (broad, 1H), δ 7.33 (dd, J = 8.65 and 2.4 Hz, 1H), δ 7.16 (d, J = 2.4 Hz, 1H), δ 6.94 (d, J = 8.65, 1H), δ 1.31 (s, 9H).

2,5-Dimethyl 2,5-pyridinedicarboxylate (**10**)

2,5-Pyridinedicarboxylic acid (200 mg, 1.197 mmol) was suspended in methanol (0.4855 mL, 11.967 mmol). Concentrated sulfuric acid (0.032 mL, 0.598 mmol) was added and the solution was stirred and refluxed for 18 hours. The solution was then neutralised with a saturated sodium bicarbonate solution and the product extracted with chloroform. The organic phase was then dried with magnesium sulfate, filtered, then the solvent was removed under reduced pressure resulting in the pure product (216.0 mg, 1.107, 92.5 %). ^{1}H NMR (500 MHz, $CDCl_3$): δ 9.30 (s, 1H), δ 8.44 (dd, J = 8.1 and 2.05 Hz, 1H), δ 8.20 (d, J = 8.1 Hz, 1H), δ 4.03 (s, 3H), δ 3.99 (s, 3H).

2,5-Pyridinedimethanol (**11**)

10 (216 mg, 1.107 mmol) was suspended in dry ethanol (3 mL) and cooled to 0 °C before sodium borohydride (167.5 mg, 4.427 mmol). The solution was then stirred at 0 °C for 1 hour, then stirred for 3 hours at room temperature, followed by refluxing the solution for 14 hours. The solvent was then removed under reduced pressure. The crude material was then dissolved in a 1:1 mixture of acetone : potassium carbonate (6 mL) and stirred at reflux for 1 hour before taking the organic layer and removing the solvent under reduced pressure giving the pure product (140.3 mg, 1.008 mmol, 91.1 %). ^{1}H NMR (500 MHz, D_2O): δ 8.35 (s, 1H), δ 7.76 (dd, J = 8.05 and 1.7 Hz, 1H), δ 7.40 (d, J = 8.0 Hz, 1H), δ 4.62 (s, 2H), δ 4.58 (s, 2H).

2,5-Pyridinedimethyl ditosylate (**12**)

11 (274.0 mg, 1.970 mmol) was dissolved in a 1 : 1 mixture of tetrahydrofuran : water (6 mL) and the solution was cooled to 0 °C. Crushed sodium hydroxide (165.5 mg, 4.138 mmol) was added, followed by adding tosyl chloride (748.7 mg, 3.941 mmol) dissolved in tetrahydrofuran (8 mL) dropwise. The solution was warmed back to room temperature and stirred overnight. The product was extracted with dichloromethane, washed with brine and dried with magnesium sulfate. The solvent was removed under reduced pressure, giving the pure product (546.9 mg, 1.223 mmol, 62.1 %). ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.40 (d, J = 1.9 Hz 1H), δ 7.83 (d, J = 8.15 Hz, 2H), δ 7.79 (d, J = 8.15 Hz, 2H), δ 7.64 (dd, J = 8.05 and 1.9 Hz, 1H), δ 7.41 (d, J = 8.1 Hz, 1H), δ 7.36 (d, J = 8.1 Hz, 2H), δ 5.13 (s, 2H), δ 5.07 (s, 2H), δ 2.46 (s, 6H).

N-Boc-ethylenediamine (**14**)

Ethylene diamine (1.83 mL, 27.49 mmol) was dissolved in dichloromethane (6.9 mL). Di-*tert*-butyl dicarbonate (1000 mg, 4.58 mmol) was dissolved in dichloromethane (55.5 mL) before being slowly added to the ethylene diamine solution over 6 hours. The solution was then left stirring for a further 12 hours before filtering any solid off and removing the solvent under reduced pressure. The crude product was then redissolved in ethyl acetate and washed with saturated sodium bicarbonate (20 mL), dried with sodium sulfate, then filtered and the solvent was removed under reduced pressure to give **14** (622.3 mg, 3.884 mmol, 84.8 %). ^{1}H NMR (500 MHz, $CDCl_3$): δ 5.20 (s, 1H), δ 3.11 (d, J = 5.35 Hz, 2H), δ 2.73 (t, J = 5.8 Hz, 2H) , δ 1.40 (s, 2H), δ 1.38 (s, 9H).

5-*tert*-Butyl-2-hydroxybenzoic acid (**15**)

Salicylic acid (200 mg, 1.449 mmol) was suspended in sulfuric acid (1 mL of an 80 % solution) and heated to 50 °C. Methyl-tert-butyl ether (0.172 mL, 1.449 mmol) was added over 2 hours while stirring. The solution was then stirred for 1 hour. The solution was then cooled and filtered, then washed with 70 °C water and dissolved ethyl acetate. The water was then separated and the ethyl acetate was dried with sodium sulphate, filtered, and the solvent was removed under reduced pressure. The crude product was stirred in hexane overnight, filtered, washed with hexane, and the solvent was removed under reduced pressure to give the pure product (208.6 mg, 1.074 mmol, 72.9 %). ^{1}H NMR (500 MHz, $CDCl_3$): δ 10.24 (s, 1H), δ 7.94 (d, J = 2.4 Hz, 1H), δ 7.62 (dd, J = 8.75 and 2.45 Hz, 1H), δ 6.99 (d, J = 8.75 Hz, 1H), δ 1.35 (s, 9H).

3-(formyl)-5-*tert*-Butyl-2-hydroxybenzoic acid (**16**)

15 (208.6 mg, 1.074 mmol) was dissolved in acetic acid (5 mL). Hexamethylenetetramine (301.1 mg, 2.148 mmol) was added and the solution was heated at reflux for 2 hours. 33% sulfuric acid (1 mL) was then added and the solution was refluxed another hour. After cooling to room temperature, the crude product was extracted with diethyl ether, washed with water and then brine, dried with magnesium sulfate, filtered, and dried under reduced pressure. The product was purified by recrystallisation from methanol (10.7 mg, 0.048 mmol, 4.5 %). ^{1}H NMR (500 MHz, $CDCl_3$): δ 10.35 (s, 1H), δ 8.32 (s, 1H), δ 8.04 (s, 1H), δ 1.37 (s, 9H).

3-(formyl)-5-*tert*-Butyl-2-hydroxybenzoyl chloride – Not attempted – (**17**)

A solution of **16** (1 eq.) and thionyl chloride (10 eq.) in toluene (10 mL per mmol of **16**) was refluxed for 1 h. The solvent was then removed under reduced pressure to give the crude product.

tert-Butyl (2-(5-(tert-butyl)-3-formyl-2-hydroxybenzamido)ethyl)carbamate – Not attempted – (**18**)

14 (1 eq. 25.0 mmol) and potassium carbonate (4 eq.) were dissolved in DCM (2 mL per mmol of potassium carbonate). A solution of **17** (1.1 eq.) in DCM (2 mL per mmol of potassium carbonate) was added slowly to the mixture and stirred under inert atmosphere at room temperature overnight. The mixture was filtered, and the filtrate was washed with water. The combined organic layer was dried over sodium sulfate. The filtrate was evaporated under reduced pressure. The crude material was purified by column chromatography eluting with ethyl acetate to give the pure product.

2,5-Bis(bromomethyl)pyridine (**22**)

11 (440.2 mg, 3.163 mmol) was dissolved in hydrogen bromide (5 mL of a 48 wt% solution) and refluxed overnight while stirring. The solution was then cooled to room temperature and then poured onto ice and neutralised with sodium hydroxide (1 M). The precipitate was then collected and columned using dichloromethane. This reaction was unsuccessful for whatever reason. It appeared to be successful but nothing dissolved in $CDCl_3$ or D_2O to be picked up on the NMR.

N-Acetylethylenediamine (**23**)

Ethylene diamine (1.00 mL, 14.983 mmol) was dissolved in water (75 mL). Acetic anhydride (1.70 mL, 17.980 mmol) was added portion wise and the solution was stirred for 5 minutes. Excess water was added and the product was extracted using ethyl acetate (2x 45 mL). The combined organic layer was then washed with sodium hydrogen carbonate, followed by water. The organic phase was then dried with sodium sulfate, filtered, then the solvent was removed under reduced pressure. The was unsuccessful as the product likely stay in the aqueous layer.

tert-Butyl (2-((5-(tert-butyl)-3-formyl-2-hydroxybenzyl)amino)ethyl)carbamate (**24**)

Triethyl amine (0.70 mL, 4.993 mmol) was dissolved in dichloromethane (6.5 mL). To this solution, a solution of **14** (200.0 mg, 1.248 mmol) dissolved in anhydrous dichloromethane (6.5 mL) was added followed by adding a solution of **7** (372.3 mg, 1.373 mmol) dissolved in anhydrous dichloromethane (6.5 mL) dropwise over 2 hours. After which the reaction was left to reflux overnight. The reaction mixture was cooled, then washed with water (3x 15 mL), dried with magnesium sulphate, filtered and the solvent removed. The crude product was then recrystallised by dissolving in the minimum amount of hot chloroform and adding boiling ethanol until precipitation is induced. This was unsuccessful as the amine over added to **7** forming imines as well as amine linkages.

Rerun was done with **14** (16 mg) with Et_3N (55.75 µL) dissolved together in DCM (400 mL). **7** (29.8 mg) and NaI (16.5 mg) were also dissolved in DCM (400 mL). Both solutions were added dropwise to a stirred round bottom flask containing DCM (200 mL) and the solution was stirred overnight at room temperature.

[Cu(**9**)$_2$(OMe)]

9 (143.7 mg, 0.744 mmol) was mixed with copper chloride (50.0 mg, 0.372 mmol) in 50 % ethanol. The solution was stirred for 30 minutes, then filtered, washed with hot water followed by methanol. The complex was crystallised by vapour diffusion of diethyl ether into a solution of the crude material dissolved in chloroform producing [Cu(**9**)$_2$(OMe)] (152.0 mg, 0.341 mmol, 90.7 %).

[Cu(**9**)$_2$(glyphosate)]

[Cu(**9**)$_2$] (50 mg, 0.112 mmol) was dissolved in chloroform (2 mL). A solution of glyphosate (37.9 mg, 0.224 mmol) dissolved in chloroform (2 mL), was then added. The two solutions where then mixed overnight and set up in a vapour diffusion apparatus, using diethyl ether. This crystallization was unsuccessful.

A2.3 Chapter 5 experimental details

6,6'-(piperazine-1,4-diylbis(methylene))bis(4-(tert-butyl)-2-((E)-(phenylimino)methyl)phenol) (**39**)

Both **8** (241.3 mg, 0.541 mmol) and aniline (125.8 mg, 1.352 mmol) were dissolved together in anhydrous toluene (20 mL) and set to reflux for 72 hrs. The solvent was removed under vacuum and the resulting pale-yellow solid was recrystallised twice by dissolving in a minimum amount of chloroform and inducing precipitation with the addition of hot ethanol, which produced **39**, as an orange solid (330.0 mg, 0.535 mmol, 99.0 %). ^{1}H NMR (500 MHz, CDCl$_3$): δ 13.25 (b, 2H), δ 8.71 (s, 2H), δ 7.44 (m, 8H), δ 7.30 (m, 6H), δ 3.75 (s, 4H), δ 2.69 (b, 8H), δ 1.37 (s, 18H); MS: m/z (ESI) 617.31 $[M+H]^+$.

Ferrocenecarbonyl azide (**40**)

Ferrocenecarboxylic acid (500.0 mg, 2.173 mmol) was dissolved in dichloromethane (5 mL) with triethylamine (1.213 mL, 8.695 mmol). Diphenylphosphoryl azide (0.468 mL, 2.173 mmol) was then added and the solution was stirred for 10 minutes. 1 M Hydrochloric acid (50 mL) was then added and the product was extracted using dichloromethane. The pooled dichloromethane was dried with magnesium sulfate, filtered, then the solvent was removed under reduced pressure. This resulted in a crude product which was then purified via column chromatography, using ethyl acetate : hexane (1:1). This gave the pure product (**40**, 352.5 mg, 1.382 mmol, 63.6 %)as the first band followed by the starting material. ^{1}H NMR (500 MHz, CDCl$_3$): δ 4.85 (s, 2H), δ 4.54 (s, 2H), δ 4.28 (s, 5H).

N-Boc-aminoferrocene (**41**)

Under an argon environment **40** (82.2 mg, 0.322 mmol) was dissolved in anhydrous toluene (2.3 mL). The solution was then refluxed for 30 minutes, then allowed to cool to room temperature. *tert*-Butyl alcohol (0.076 mL, 0.806 mmol) was added and the solution was refluxed for 1 hour. The solvent was then removed under reduced pressure before the product was purified by dissolving it in hot toluene, then allowing it to sit unstirred overnight. The solvent was then decanted leaving the product (76.9 mg, 0.256 mmol, 79.3 %). ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.72 (s, 2H), δ 7.59 (d, J = 8.4 Hz, 4H), δ 7.41 (s, 10H), δ 7.30 (d, J = 8.4 Hz, 4H), δ 5.30 (s, 4H), δ 3.75 (s, 4H), δ 2.69 (broad, 8H), δ 1.37 (s, 18H).

Aminoferrocene (**42**)

41 (696.6 mg, 2.314 mmol) was suspended in dichloromethane (5 mL) at 35 °C. Trifluoroacetic acid (0.885 mL, 11.570 mmol) was added dropwise to the reaction over 1.5 minutes. The reaction was then stirred while coming to room temperature for 30 min. The solvent was removed under reduced pressure and sodium hydroxide (20 mL of a 1 M solution) was added. The product was extracted with ethyl acetate, dried with sodium sulfate, filtered and solvent removed under reduced pressure. This resulted in a crude product which was then purified via column chromatography, using ethyl acetate giving the pure product (mg, mmol, %). ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.72 (s, 2H), δ 7.59 (d, J = 8.4 Hz, 4H), δ 7.41 (s, 10H), δ 7.30 (d, J = 8.4 Hz, 4H), δ 5.30 (s, 4H), δ 3.75 (s, 4H), δ 2.69 (broad, 8H), δ 1.37 (s, 18H).

6,6'-(piperazine-1,4-diylbis(methylene))bis(4-(tert-butyl)-2-((E)-(ferrocene-imino)methyl)phenol) (**43**)

Both **42** (252.2 mg, 1.255 mmol) and **8** (146.3 mg, 0.314 mmol) were dissolved together in anhydrous toluene (10 mL) and set to reflux for 72 hrs. The solvent was removed under vacuum and the resulting pale-yellow solid was recrystallised twice by dissolving in a minimum amount of chloroform and inducing precipitation with the addition of hot ethanol, which produced **43**, as a pale-yellow solid (mg, mmol, %). ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.72 (s, 2H), δ 7.59 (d, J = 8.4 Hz, 4H), δ 7.41 (s, 10H), δ 7.30 (d, J = 8.4 Hz, 4H), δ 5.30 (s, 4H), δ 3.75 (s, 4H), δ 2.69 (broad, 8H), δ 1.37 (s, 18H).

$[Cu(\mathbf{9})_2(\text{hydrogen cyanamide})]$

$[Cu(\mathbf{9})_2]$ (50 mg, 0.112 mmol) was dissolved in chloroform (2 mL). A solution of hydrogen cyanamide (121 mg, 2.240 mmol) dissolved in chloroform (2 mL), was then added. The two solutions where then mixed and set up in a vapour diffusion apparatus, using diethyl ether. This crystallization was unsuccessful.

$[Mn_2(\mathbf{39})_2](NO_3)_x$

39 (50 mg, 0.081 mmol) was suspended in acetonitrile (5 mL) or methanol (5 mL). A solution of manganese (II) nitrate tetrahydrate (20.4 mg, 0.081 mmol) in the same solvent (5 mL) was added to the first solution with pyridine (6.56 μL, 0.081 mmol). The solution was then stirred at 60 °C overnight and then the solvent was removed under vacuum. A purification was attempted unsuccessfully by crystallization with vapour diffusion of diethyl ether into a solution of the product dissolved in chloroform.

$[Mn_2(\mathbf{39})_2]Cl_x$

39 (50 mg, 0.081 mmol) was suspended in acetonitrile (5 mL) or methanol (5 mL). A solution of manganese (II) chloride (10.2 mg, 0.081 mmol) in the same solvent (5 mL) was added to the first solution with pyridine (6.56 µL, 0.081 mmol). The solution was then stirred at 60 °C overnight and then the solvent was removed under vacuum. A purification was attempted unsuccessfully by crystallization with vapour diffusion of diethyl ether into a solution of the product dissolved in chloroform.

A2.4 Appendix 1 experimental details

Diethyl oxalate (**45**)

Oxalic acid (200 mg, 2.221 mmol) was dissolved in ethanol (5 mL) in the presence of concentrated sulfuric acid (4 drops) and refluxed for 4 hours. Remove the solvent under reduced pressure and dissolve in a water : chloroform mixture. Extract the product with chloroform and dried with sodium sulfate, filtered and the solvent was removed under reduced pressure to give diethyl oxalate (324.5 mg, 2.220 mmol, ~ 100 %). ^{1}H NMR (500 MHz, $CDCl_3$): δ 4.30 (qd, J = 7.15 and 1.65 Hz 2H), δ 1.32 (td, J = 7.15 and 1.65 Hz, 3H).

Ethyl 3-phenyl-1*H*-pyrazole-5-carboxylate – Not attempted. (**46**)

Sodium (1.25 eq.) was dissolved in absolute ethanol (1 L per mmol of acetophenone) and the mixture was stirred under inert atmosphere. The reaction mixture was cooled to room temperature, before **45** (1.25 mol) was added. The mixture was then stirred for 10 minutes, followed by the dropwise addition of acetophenone (1 eq.) within 1.5 hours. The resulting mixture was stirred for another 4 hours before acetic acid (1.25 eq.) was added, followed by dropwise addition of aqueous hydrazine hydrate (1 eq.) and stirred overnight. The solvent was removed under reduced pressure to give a residue, which was dissolved in DCM (1.5 L per mmol of acetophenone). The organic solution was then washed with brine and water, dried over sodium sulfate, filtered and the solvent removed under reduced pressure to yield the crude product as a white solid, which was recrystallised from ethyl acetate to afford the pure product as a white solid.

Ethyl 4-hydroxy-2-phenylpyrazolo[1,5-a]pyridine-5-carboxylate – Not attempted. (**47**)

46 (1 eq.), ethyl 4-bromobut-2-enoate (2 eq.) and potassium carbonate (2.05 eq.) were dissolved in ethanol (20 mL per mmol of **46**). The mixture was refluxed for 8 hours and then were poured into water (100 mL per mmol of **46**) and extracted with DCM. The organic solution was then washed with water, dried with sodium sulfate, filtered, and the solvent was removed under reduced pressure. The crude product was purified by column chromatography to give the pure product.

A2.5 X-ray collection data

Table 2.1: X-ray collection data of $\mathbf{5^A}$ (found in Chapter 3)

Identification code	$\mathbf{5^A}$
Empirical formula	$C_{14}H_{11}NO_4S$
Formula weight	289.30
Temperature (K)	100.0
Crystal system	monoclinic
Space group	$P2_1/c$
a (Å)	7.4250(5)
b (Å)	22.1913(15)
c (Å)	8.0247(6)
α (°)	90
β (°)	92.913(2)
γ (°)	90
Volume (Å^3)	1320.52(16)
Z	4
ρ_{calc} (g/cm^3)	1.455
μ (mm^{-1})	2.310
F(000)	600.0
Crystal size (mm^3)	0.28 × 0.22 × 0.08
Radiation	CuKα (λ = 1.54178)
2Θ range for data collection (°)	7.968 to 136.432

Index ranges	$-8 \le h \le 8$, $-26 \le k \le 26$, $-8 \le l \le 9$
Reflections collected	9994
Independent reflections	2352 [R_{int} = 0.0383, R_{sigma} = 0.0473]
Data/restraints/parameters	2352/0/182
Goodness-of-fit on F^2	1.187
Final R indexes [I>=2σ (I)]	R_1 = 0.0494, wR_2 = 0.1108
Final R indexes [all data]	R_1 = 0.0502, wR_2 = 0.1116
Largest diff. peak/hole / e $Å^{-3}$	0.61/-0.41

Table 2.2: X-ray collection data of $\mathbf{5^B}$ (found in Chapter 3)

Identification code	$\mathbf{5^B}$
Empirical formula	$C_{13}H_{11}NO_2S$
Formula weight	245.29
Temperature (K)	150.00
Crystal system	monoclinic
Space group	$P2_1/n$
a (Å)	11.0161(4)
b (Å)	7.9208(3)
c (Å)	14.3348(5)
α (°)	90
β (°)	111.6920(10)
γ (°)	90
Volume ($Å^3$)	1162.22(7)
Z	4
ρ_{calc} (g/cm^3)	1.402
μ (mm^{-1})	2.384

F(000)	512.0
Crystal size (mm^3)	0.18 × 0.17 × 0.12
Radiation	CuKα (λ = 1.54178)
2Θ range for data collection (°)	8.736 to 151.838
Index ranges	$-13 \leq h \leq 13$, $-9 \leq k \leq 9$, $-18 \leq l \leq 17$
Reflections collected	11317
Independent reflections	2346 [R_{int} = 0.0316, R_{sigma} = 0.0271]
Data/restraints/parameters	2346/0/199
Goodness-of-fit on F^2	1.148
Final R indexes [I>=2σ (I)]	R_1 = 0.0380, wR_2 = 0.0960
Final R indexes [all data]	R_1 = 0.0392, wR_2 = 0.0968
Largest diff. peak/hole / e Å^{-3}	0.31/-0.33

Table 2.3: X-ray collection data of **8** (found in Chapter 3)

Identification code	**8**
Empirical formula	$C_{28}H_{38}N_2O_4$
Formula weight	466.60
Temperature (K)	100.0
Crystal system	triclinic
Space group	P-1
a (Å)	7.7548(7)
b (Å)	9.5910(9)
c (Å)	9.5917(9)
α (°)	110.608(5)
β (°)	93.971(5)
γ (°)	99.992(5)

Volume (Å^3)	651.01(11)
Z	1
ρ_{calc} (g/cm^3)	1.190
μ (mm^{-1})	0.630
F(000)	252.0
Crystal size (mm^3)	0.11 × 0.08 × 0.06
Radiation	CuKα (λ = 1.54178)
2Θ range for data collection (°)	9.952 to 122.312
Index ranges	$-8 \leq h \leq 8$, $-10 \leq k \leq 10$, $-10 \leq l \leq 10$
Reflections collected	16669
Independent reflections	1958 [R_{int} = 0.0436, R_{sigma} = 0.0268]
Data/restraints/parameters	1958/0/231
Goodness-of-fit on F^2	1.259
Final R indexes [I>=2σ (I)]	R_1 = 0.0505, wR_2 = 0.1208
Final R indexes [all data]	R_1 = 0.0542, wR_2 = 0.1222
Largest diff. peak/hole / e Å^{-3}	0.24/-0.21

Table 2.4: X-ray collection data of $[Cu_2(\mathbf{1^A})_2](NO_3)_4$ (found in Chapter 3)

Identification code	$[Cu_2(\mathbf{1^A})_2](NO_3)_4$
Empirical formula	$C_{112}H_{120}Cu_2N_{12}O_{24}S_4$
Formula weight	2210.49
Temperature (K)	100.00
Crystal system	monoclinic
Space group	$P2_1/c$
a (Å)	25.227(4)
b (Å)	23.228(4)

c (Å)	26.205(4)
α (°)	90
β (°)	116.688(14)
γ (°)	90
Volume (Å^3)	13720(4)
Z	4
ρ_{calc} (g/cm^3)	1.070
μ (mm^{-1})	1.436
F(000)	4632.0
Crystal size (mm^3)	0.07 × 0.05 × 0.01
Radiation	CuKα (λ = 1.54178)
2Θ range for data collection (°)	5.36 to 133.192
Index ranges	$-29 \leq h \leq 29, -27 \leq k \leq 26, -30 \leq l \leq 31$
Reflections collected	164861
Independent reflections	24104 [R_{int} = 0.2903, R_{sigma} = 0.2255]
Data/restraints/parameters	24104/355/1214
Goodness-of-fit on F^2	0.953
Final R indexes [I>=2σ (I)]	R_1 = 0.1238, wR_2 = 0.3281
Final R indexes [all data]	R_1 = 0.2568, wR_2 = 0.4118
Largest diff. peak/hole / e Å^{-3}	0.93/-0.37

Table 2.5: X-ray collection data of **39** (found in Chapter 5)

Table 1 Crystal data and structure refinement for 39.

Identification code	**39**
Empirical formula	$C_{40}H_{48}N_4O_2$
Formula weight	616.82

Temperature (K)	100.00
Crystal system	monoclinic
Space group	C2/c
a (Å)	18.3486(17)
b (Å)	6.5794(6)
c (Å)	28.861(3)
α (°)	90
β (°)	93.969(6)
γ (°)	90
Volume (Å^3)	3475.9(6)
Z	4
ρ_{calc} (g/cm^3)	1.179
μ (mm^{-1})	0.567
F(000)	1328.0
Crystal size (mm^3)	0.14 × 0.14 × 0.01
Radiation	CuKα (λ = 1.54178)
2Θ range for data collection (°)	6.14 to 136.712
Index ranges	$-22 \leq h \leq 22$, $-7 \leq k \leq 6$, $-34 \leq l \leq 33$
Reflections collected	9443
Independent reflections	3159 [R_{int} = 0.0584, R_{sigma} = 0.0632]
Data/restraints/parameters	3159/0/216
Goodness-of-fit on F^2	1.061
Final R indexes [I>=2σ (I)]	R_1 = 0.0607, wR_2 = 0.1562
Final R indexes [all data]	R_1 = 0.0939, wR_2 = 0.1772
Largest diff. peak/hole / e Å^{-3}	0.28/-0.29

Appendix 3 – Experimental Spectra

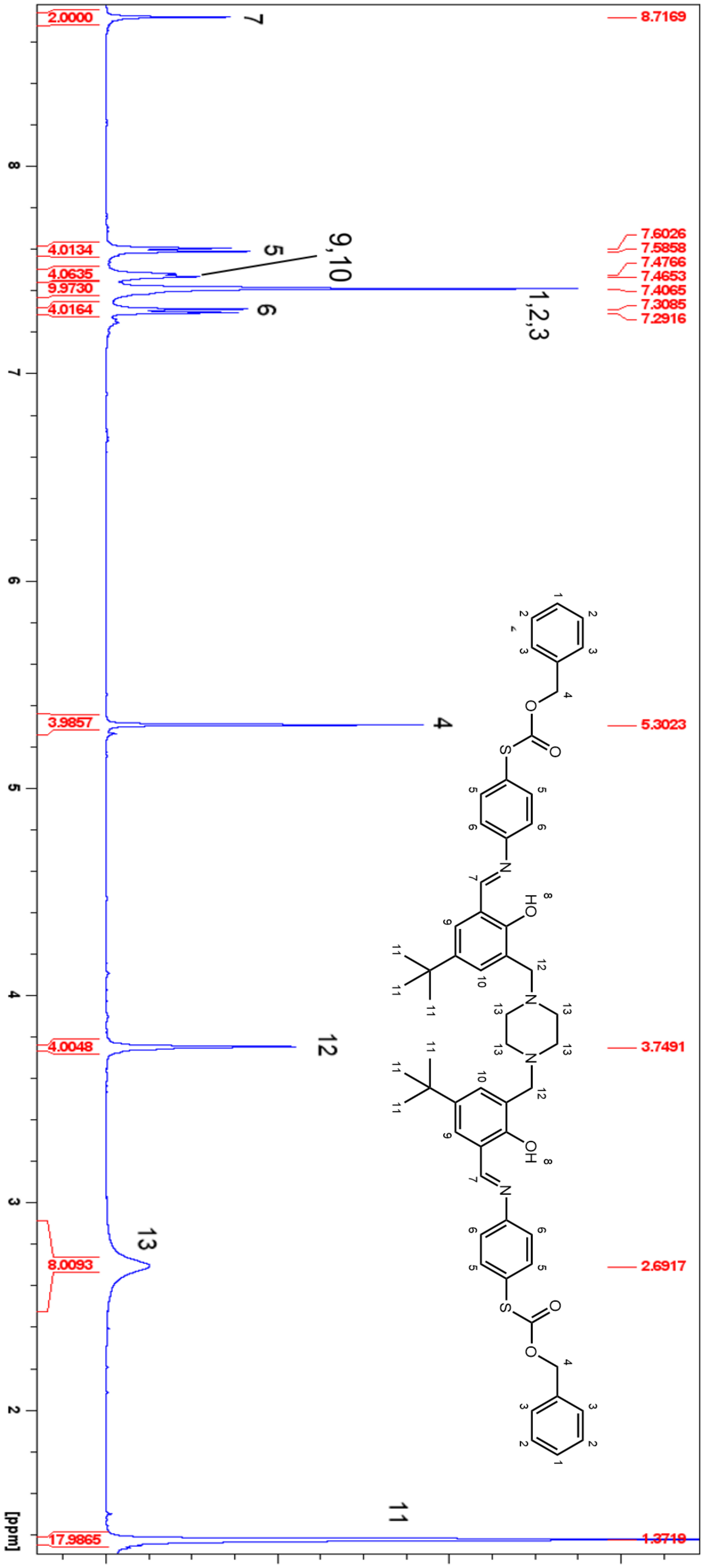


Figure A3.1: ^{1}H NMR of **1^{A}**.

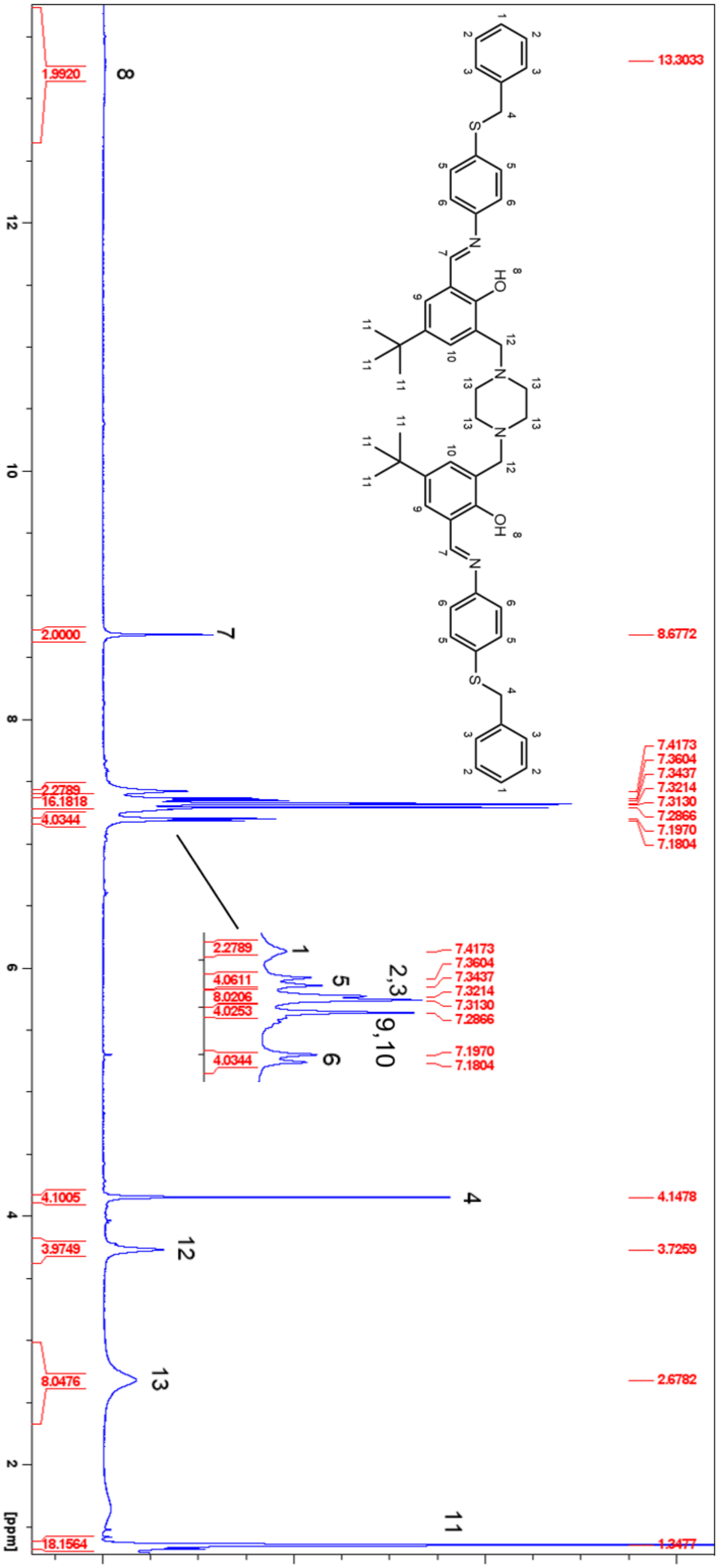


Figure A3.2: ^{1}H NMR of $\mathbf{1^B}$.

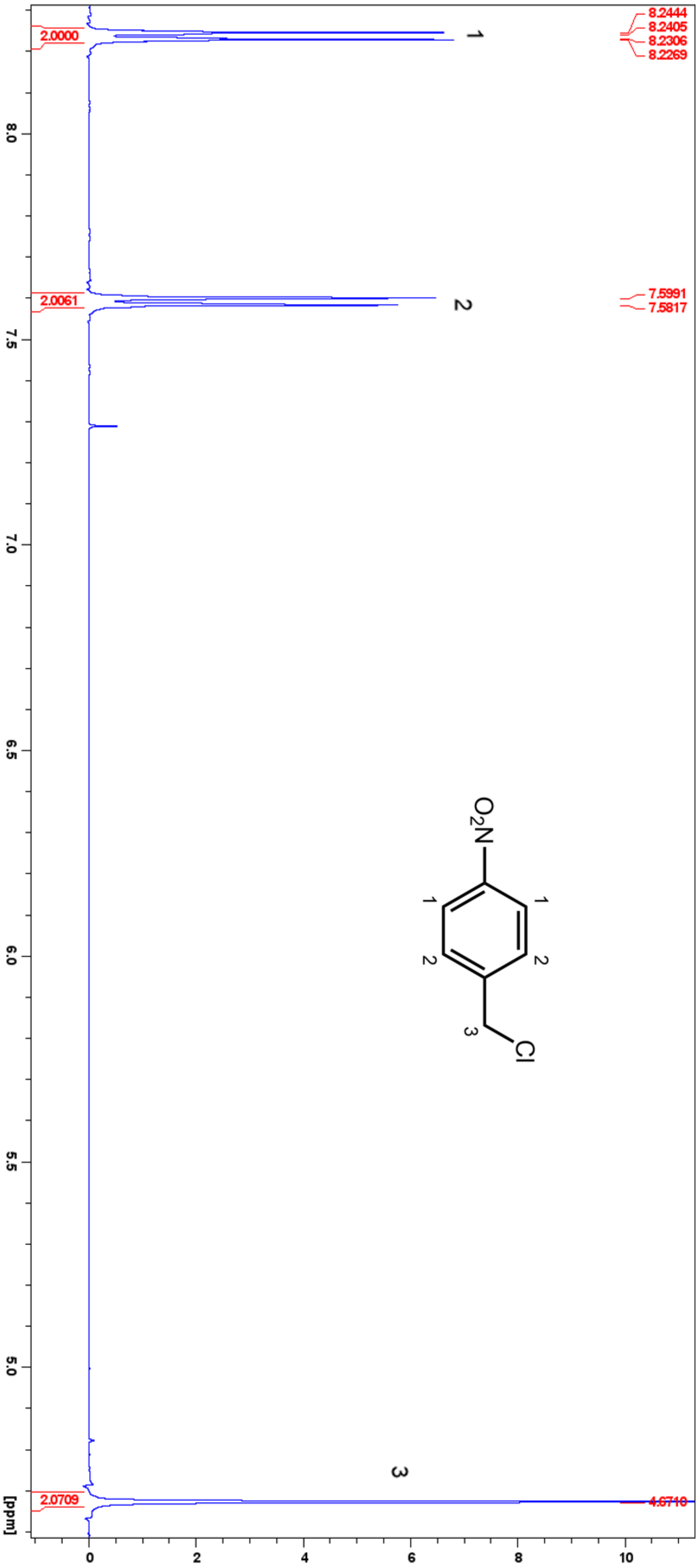


Figure A3.3: ^{1}H NMR of **2**.

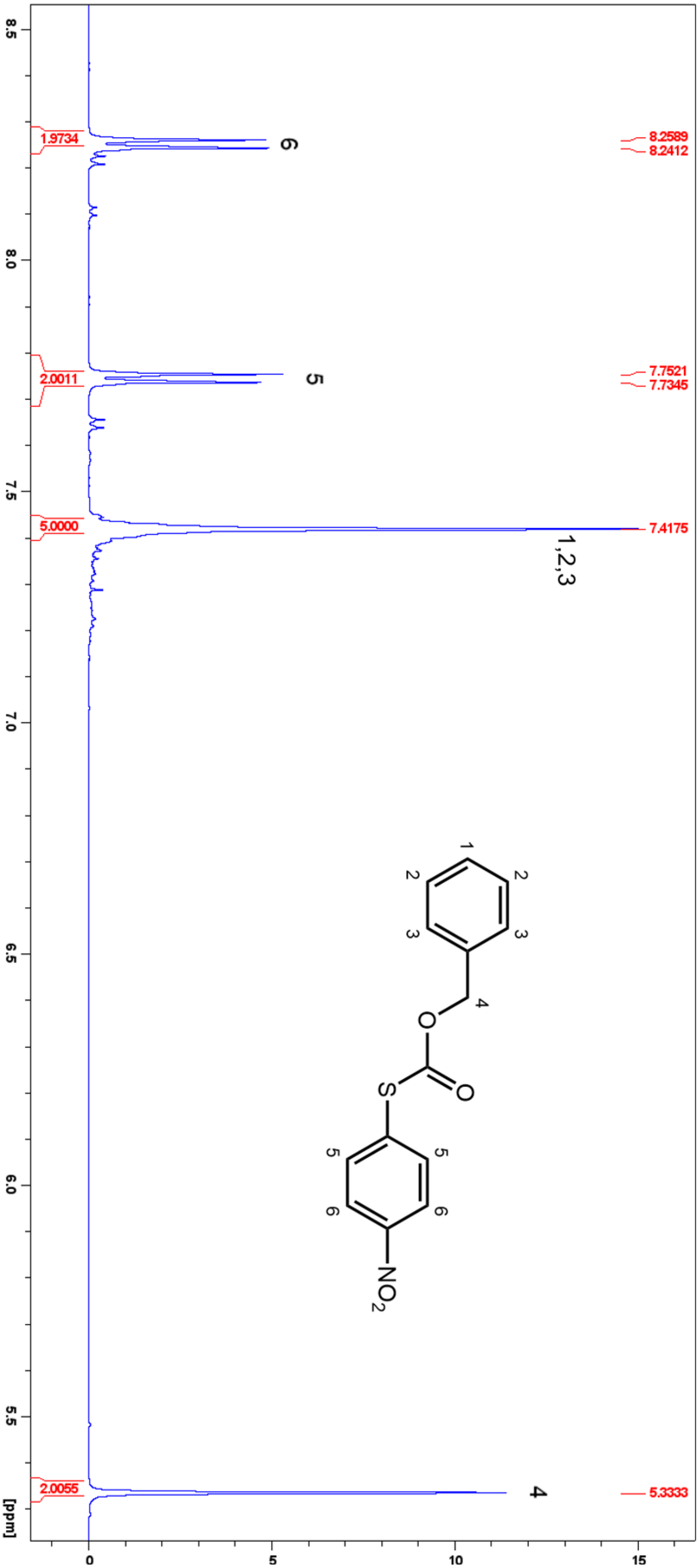


Figure A3.4: ^{1}H NMR of 5^A.

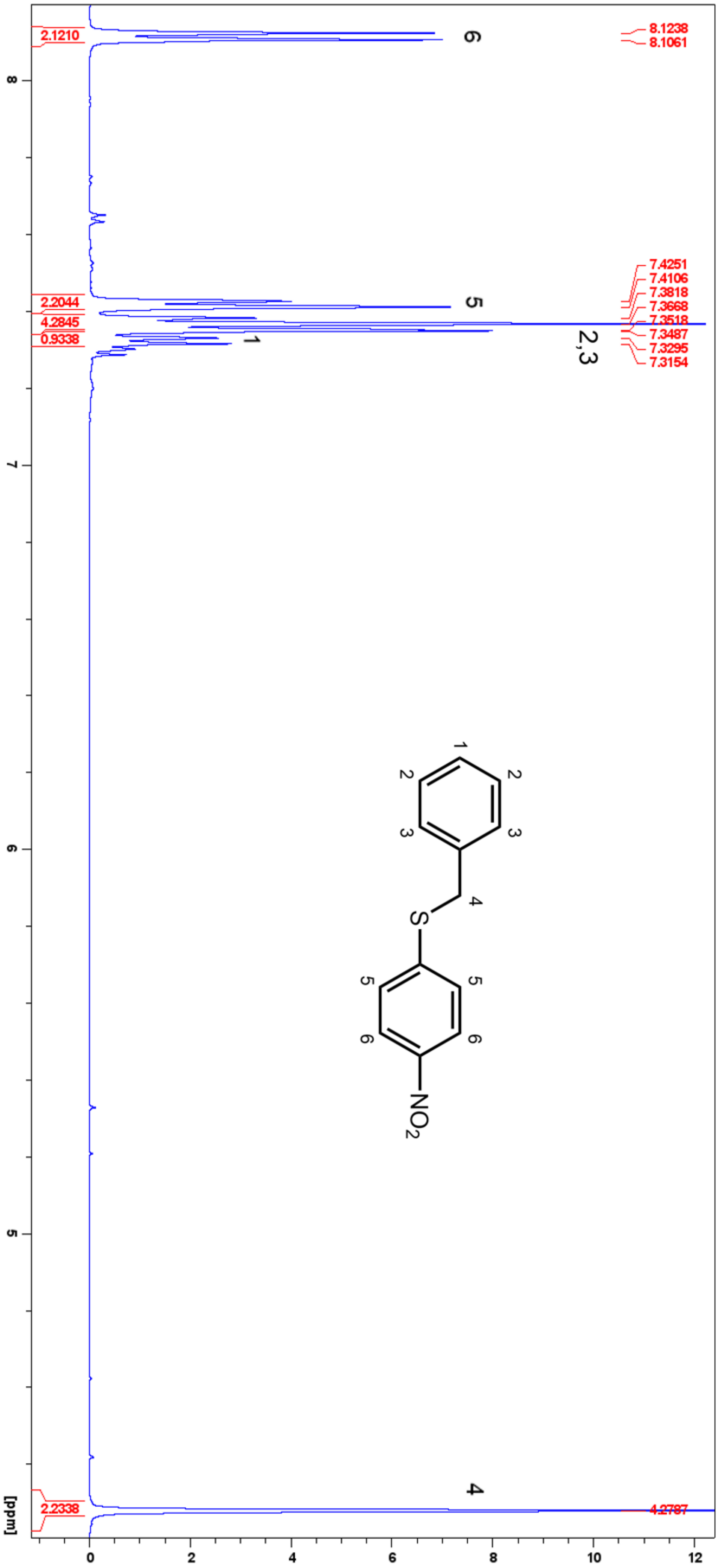


Figure A3.5: ^{1}H NMR of **5^B**.

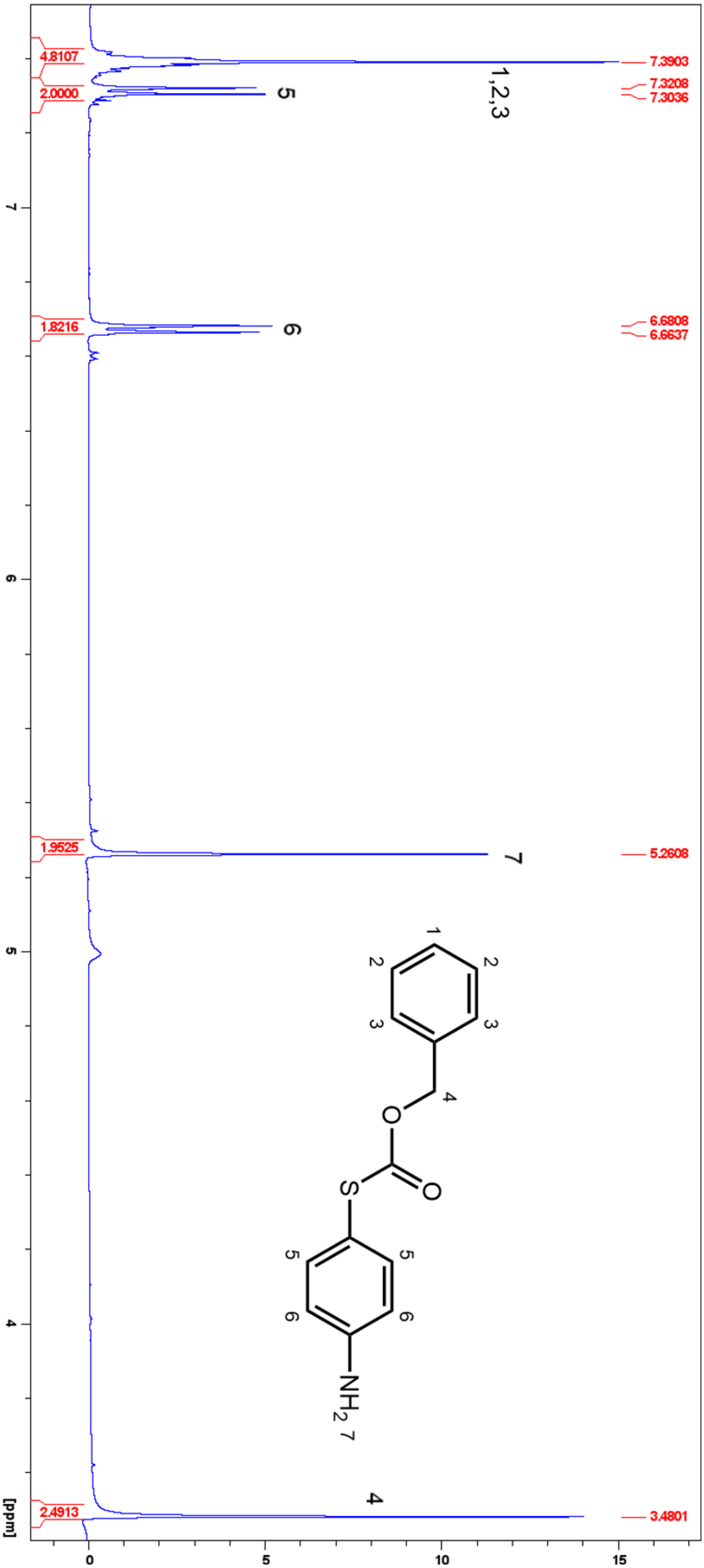


Figure A3.6: ^{1}H NMR of $\mathbf{6}^A$.

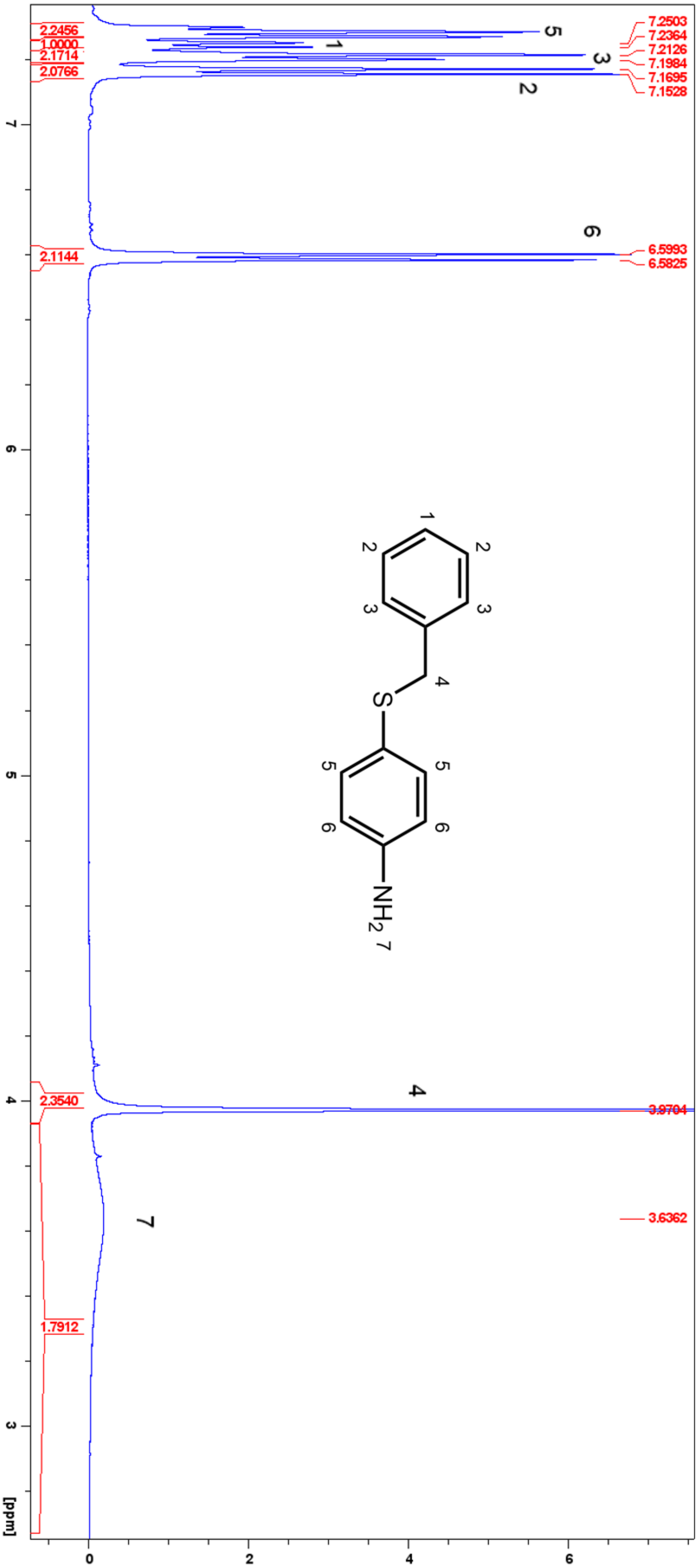


Figure A3.7: ^{1}H NMR of **6**$^{\mathbf{B}}$.

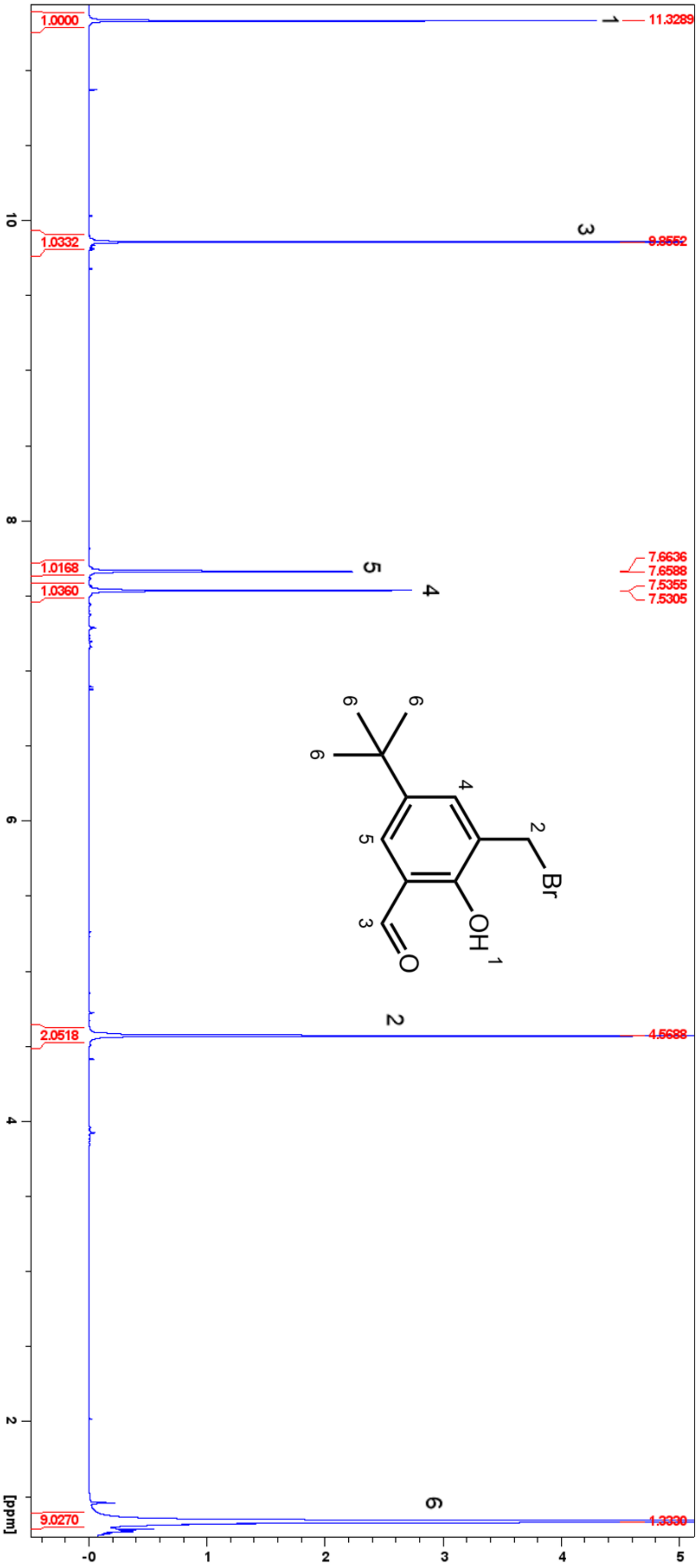


Figure A3.8: ^{1}H NMR of **7**.

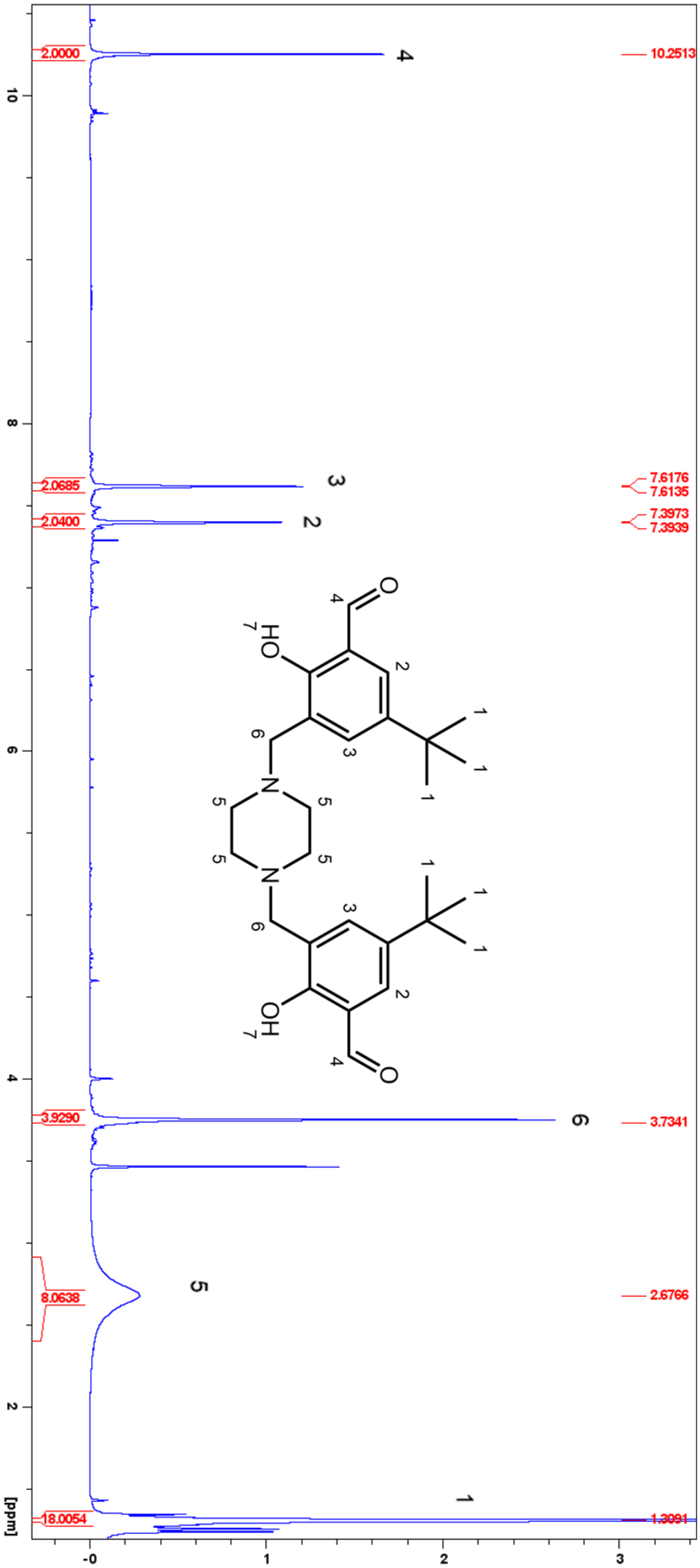


Figure A3.9: ^{1}H NMR of **8**.

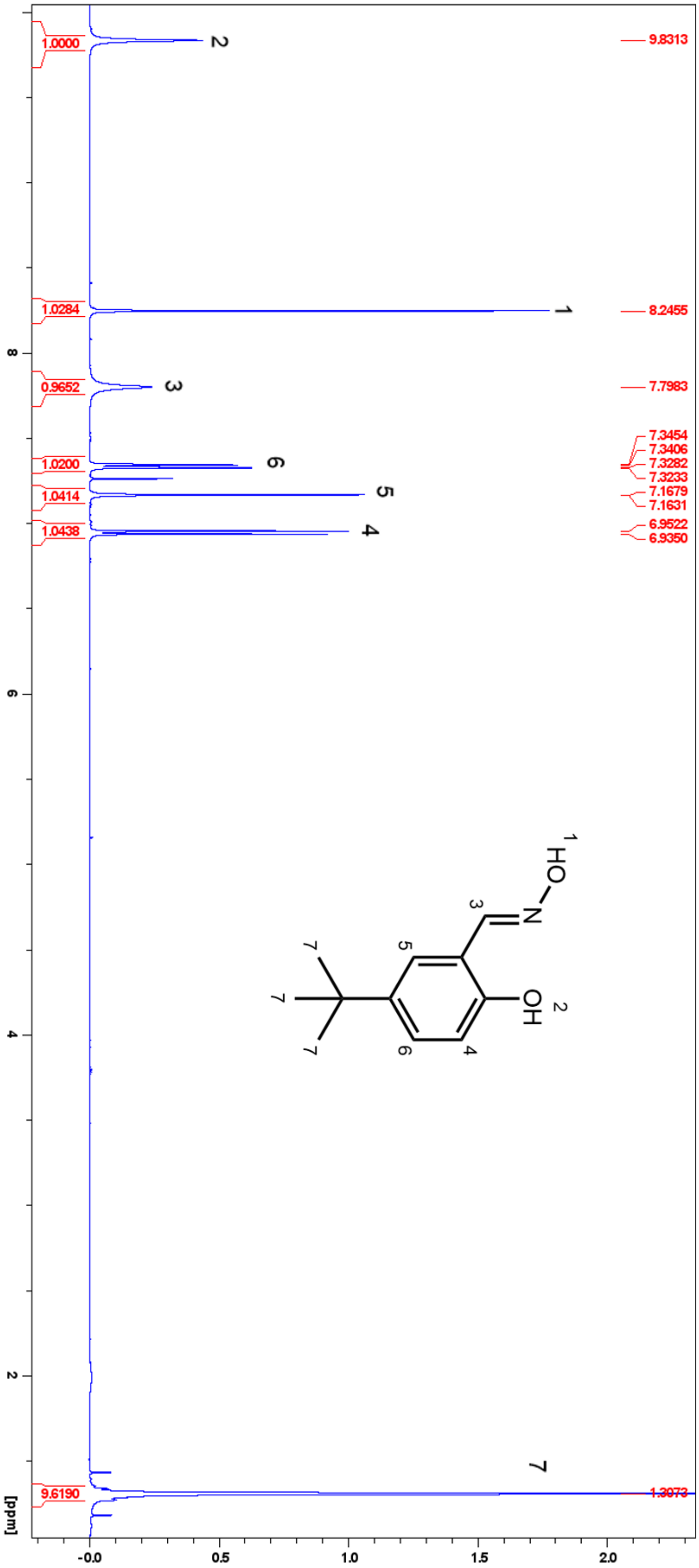


Figure A3.10: ^{1}H NMR of **9**.

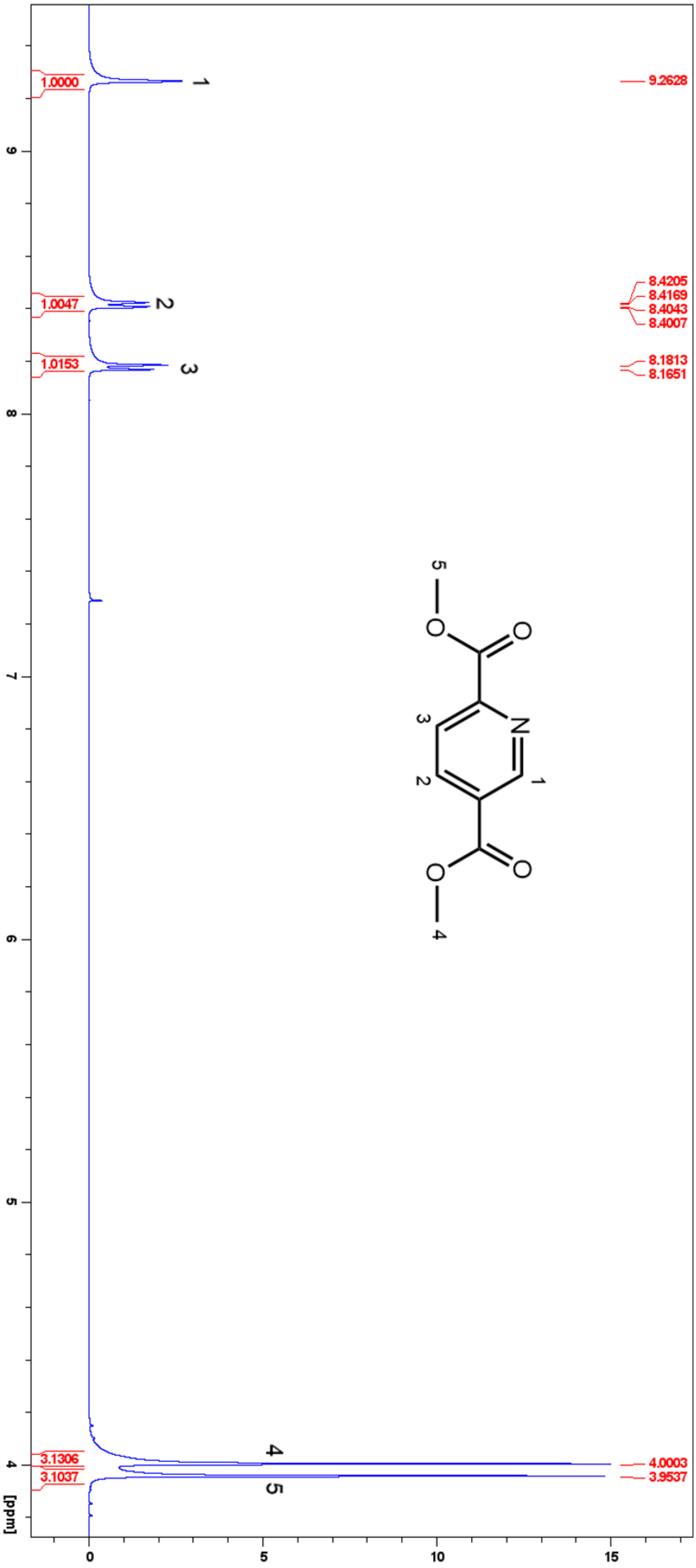


Figure A3.11: ^{1}H NMR of **10**.

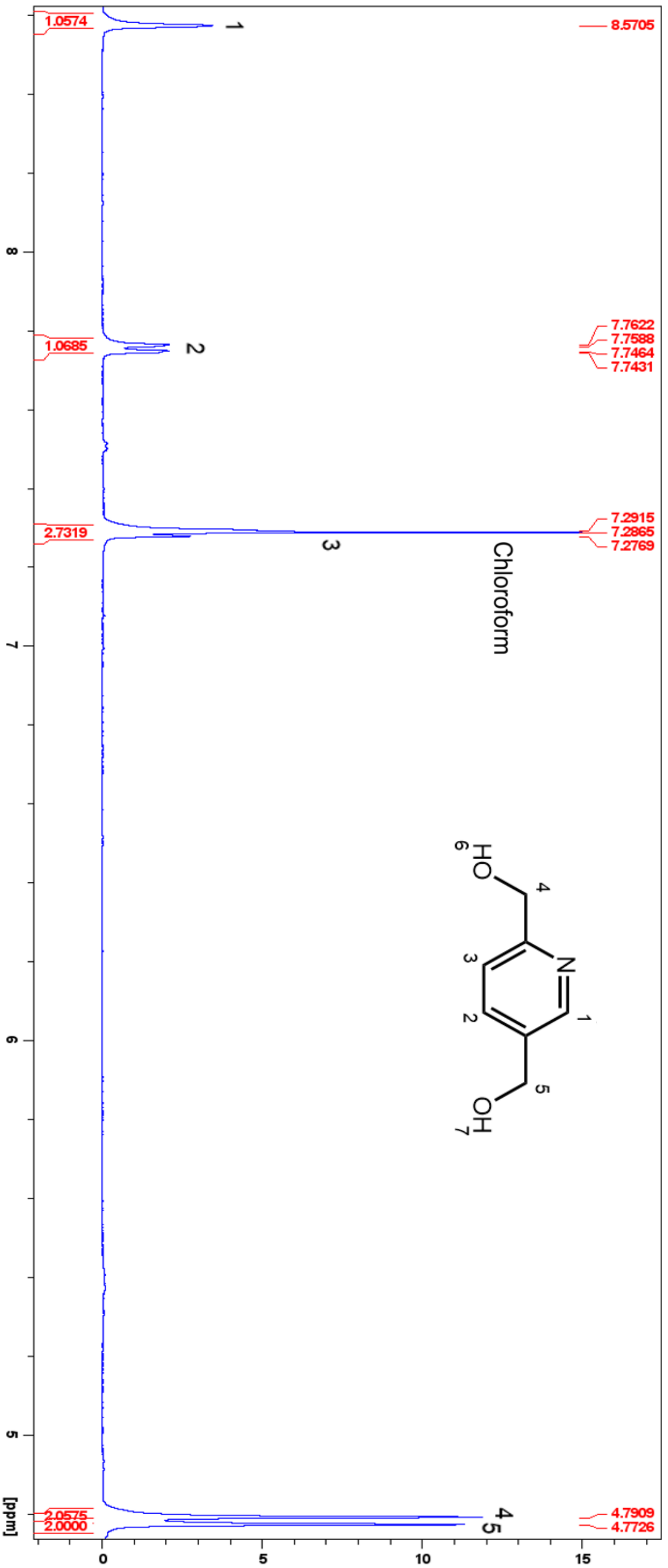


Figure A3.12: ^{1}H NMR of **11**.

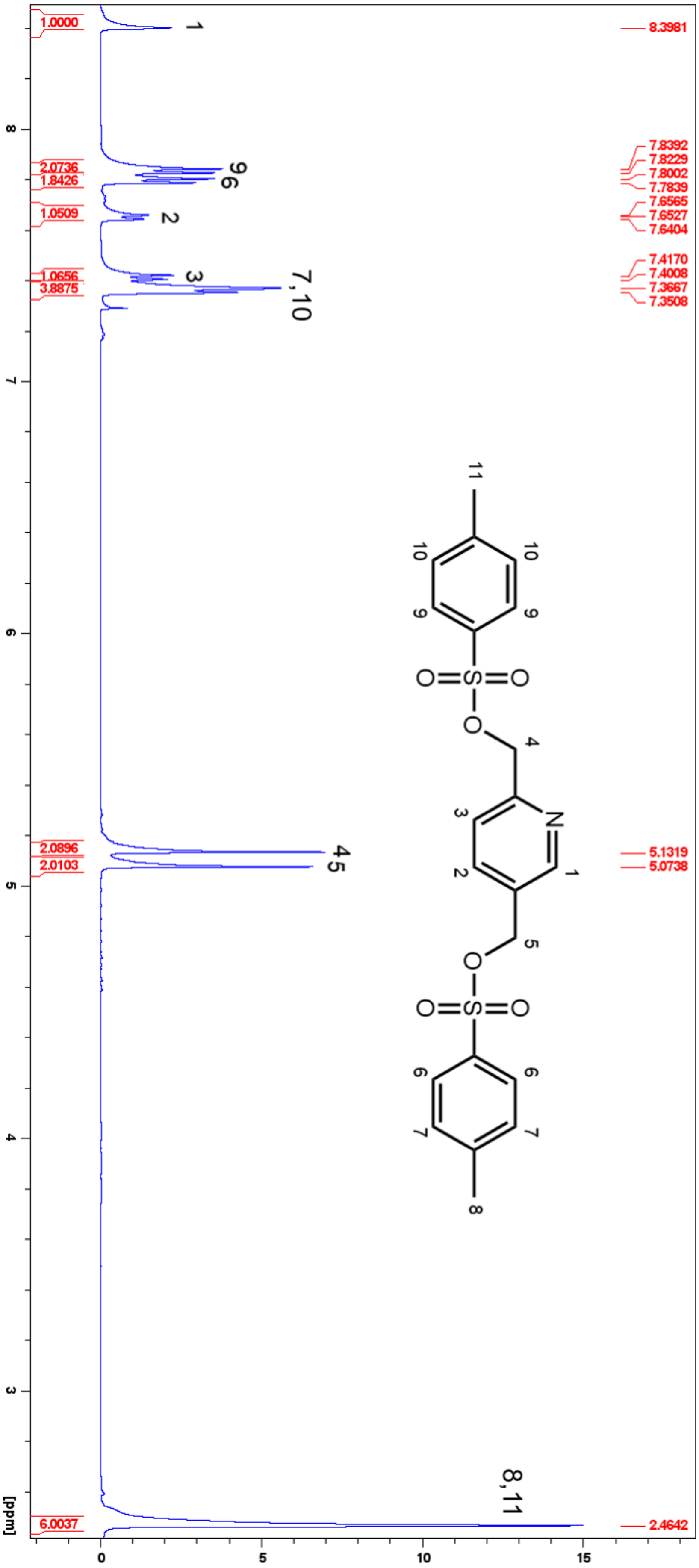


Figure A3.13: ^{1}H NMR of **12**.

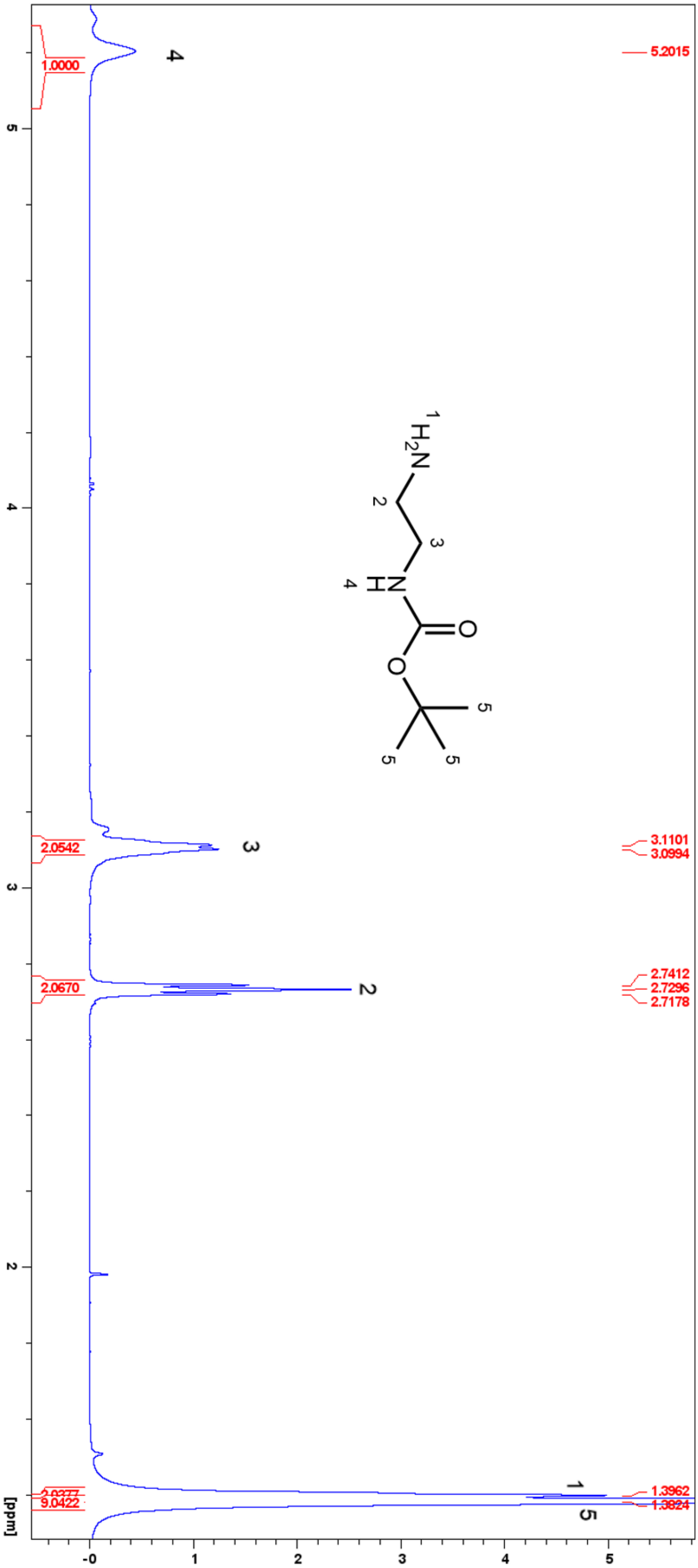


Figure A3.14: ^{1}H NMR of **14**.

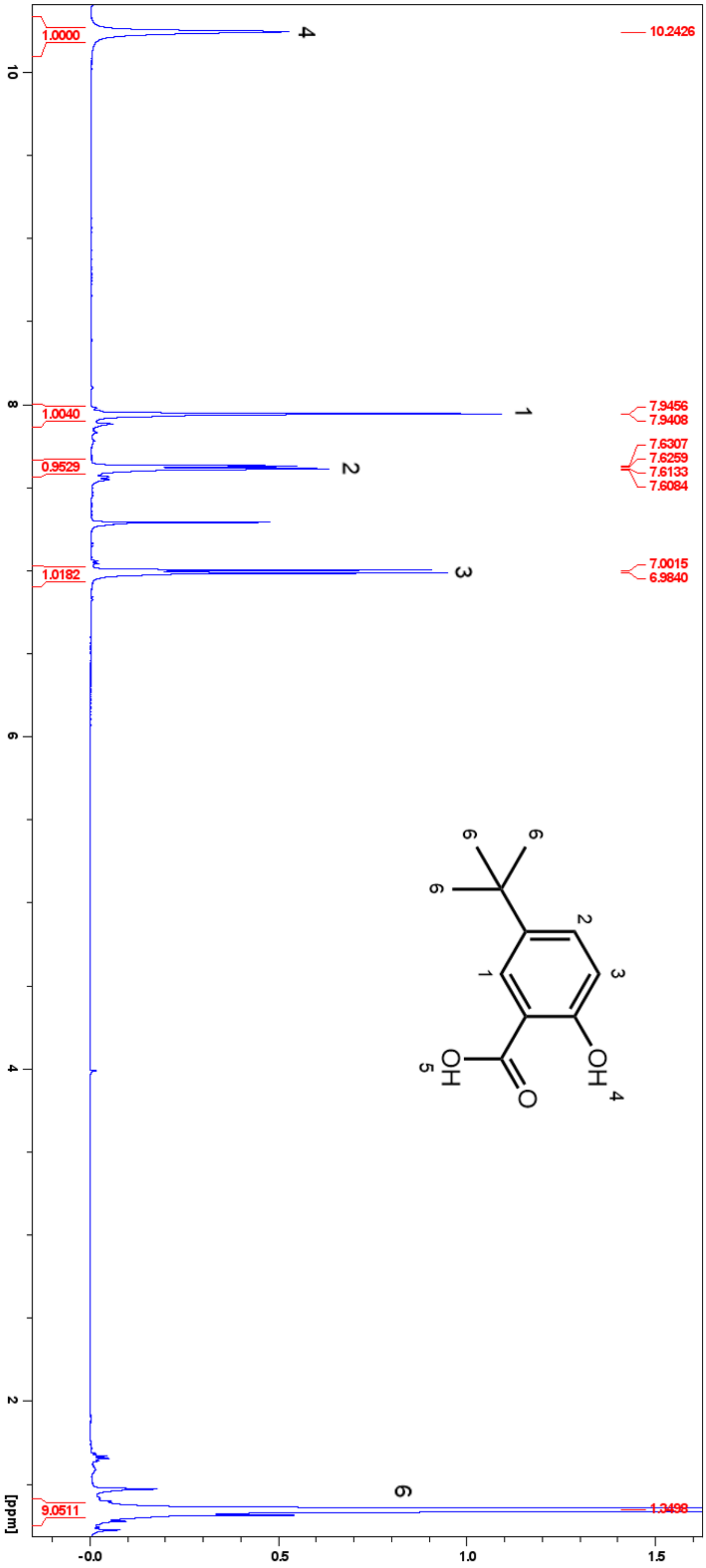


Figure A3.15: ^{1}H NMR of **15**.

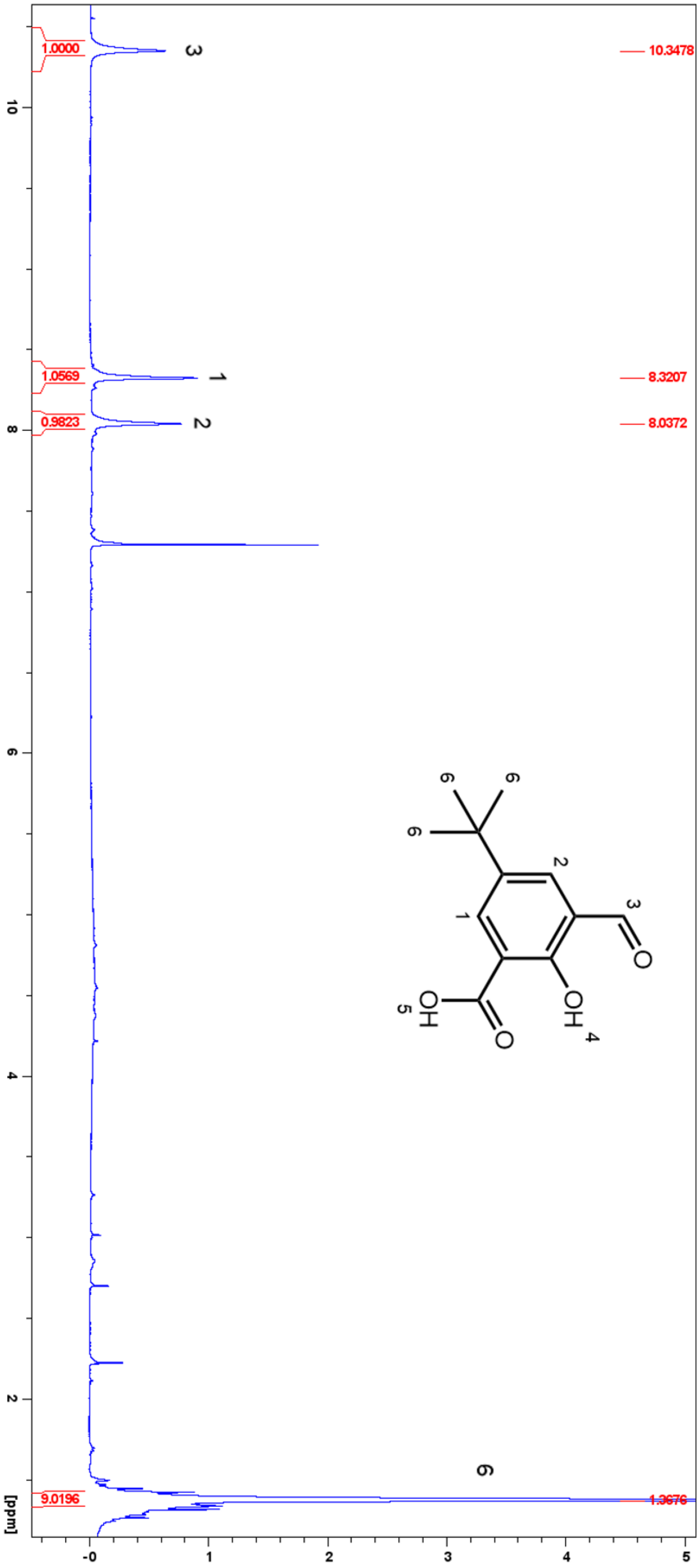


Figure A3.16: ^{1}H NMR of **16**.

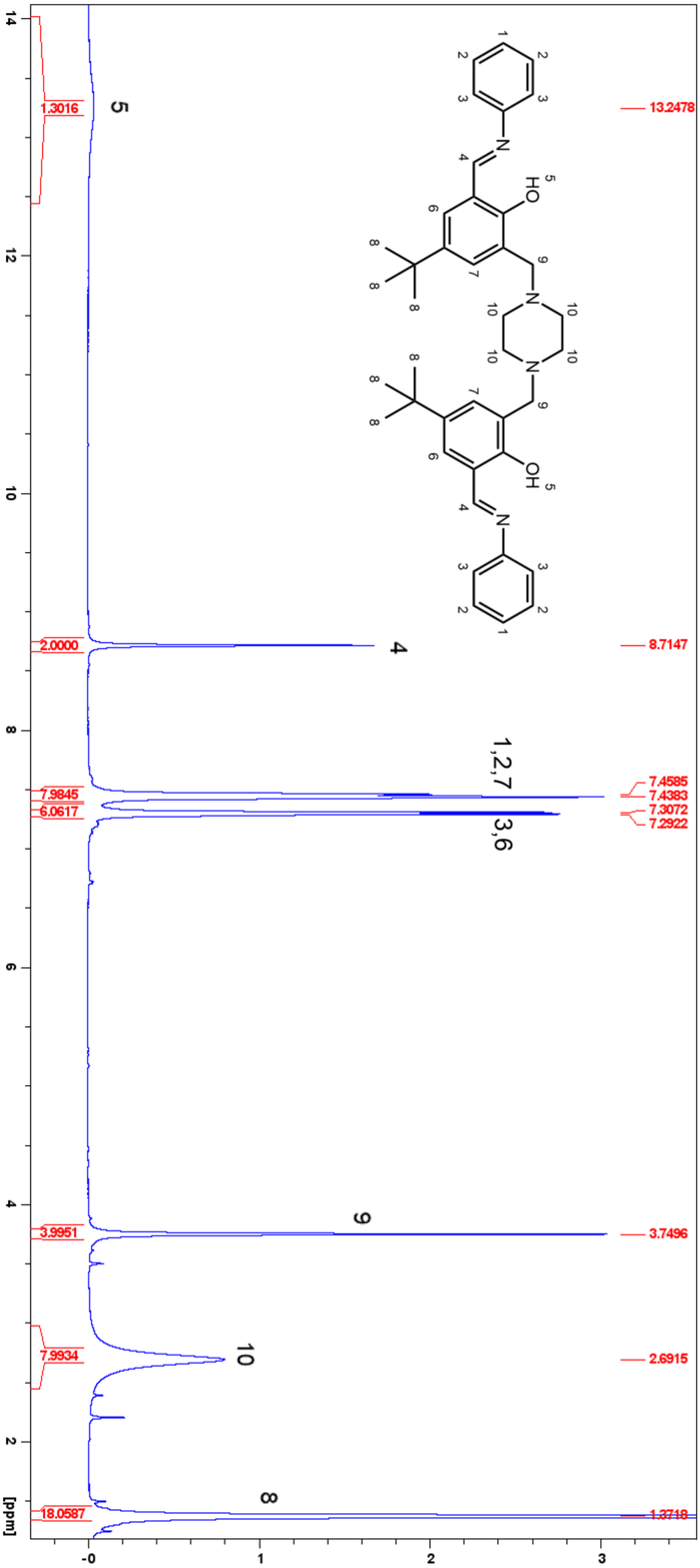


Figure A3.17: ^{1}H NMR of **39**.

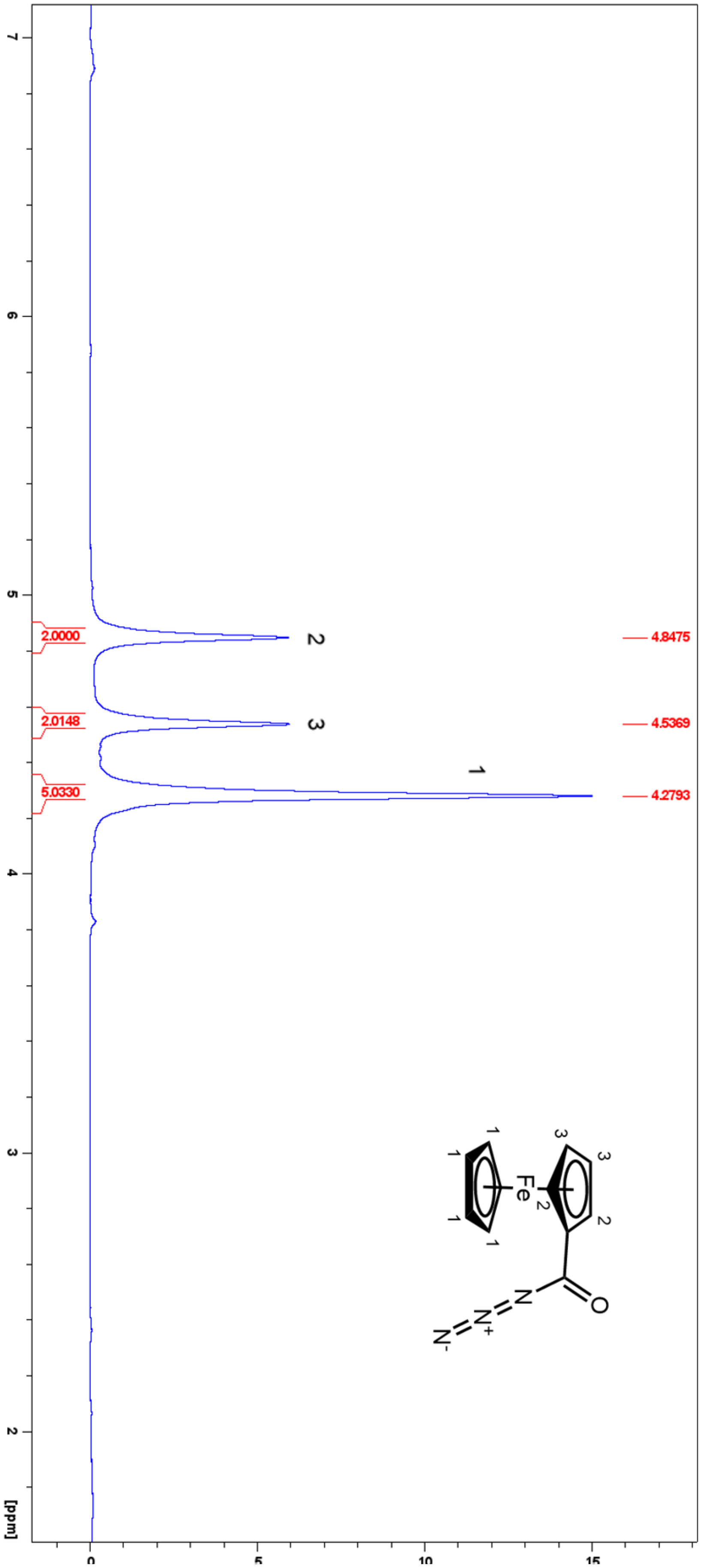


Figure A3.18: ^{1}H NMR of **40**.

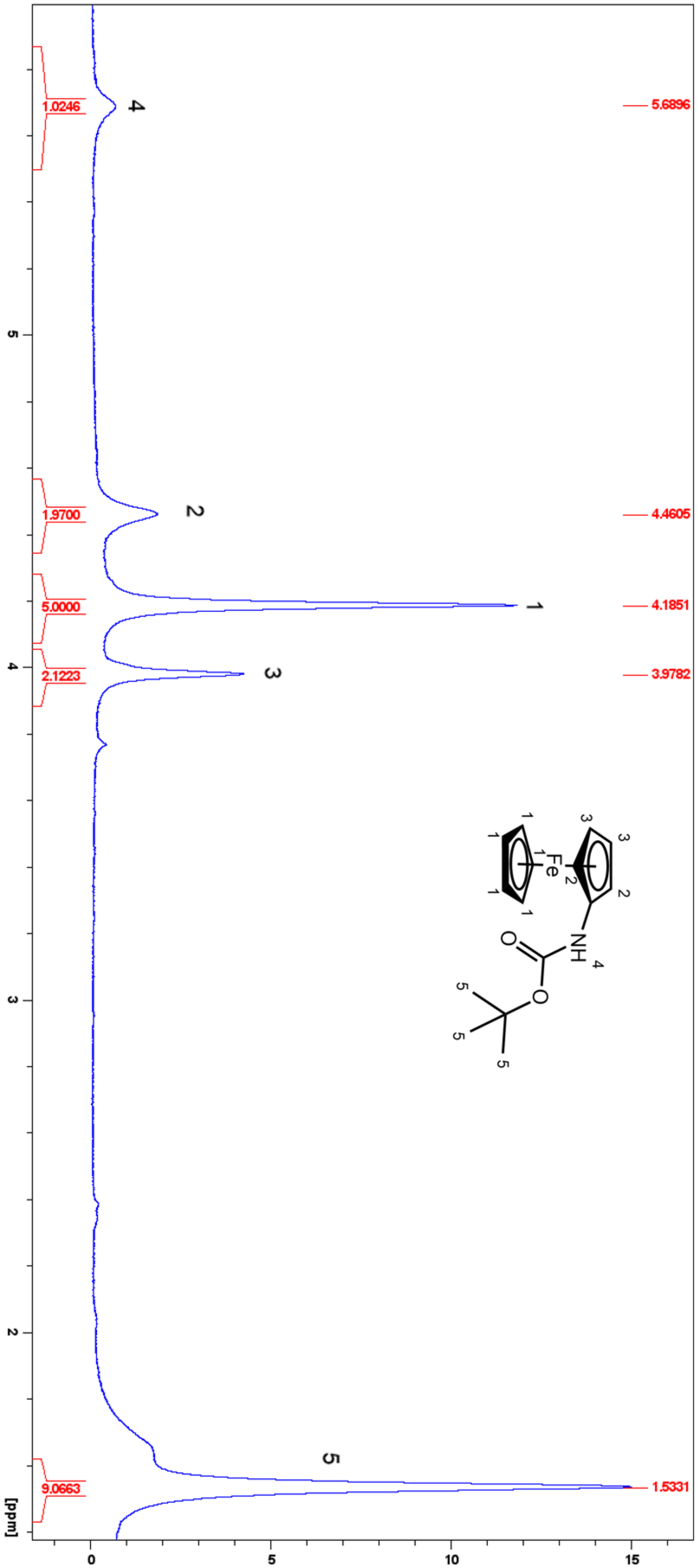


Figure A3.19: ^{1}H NMR of **41**.

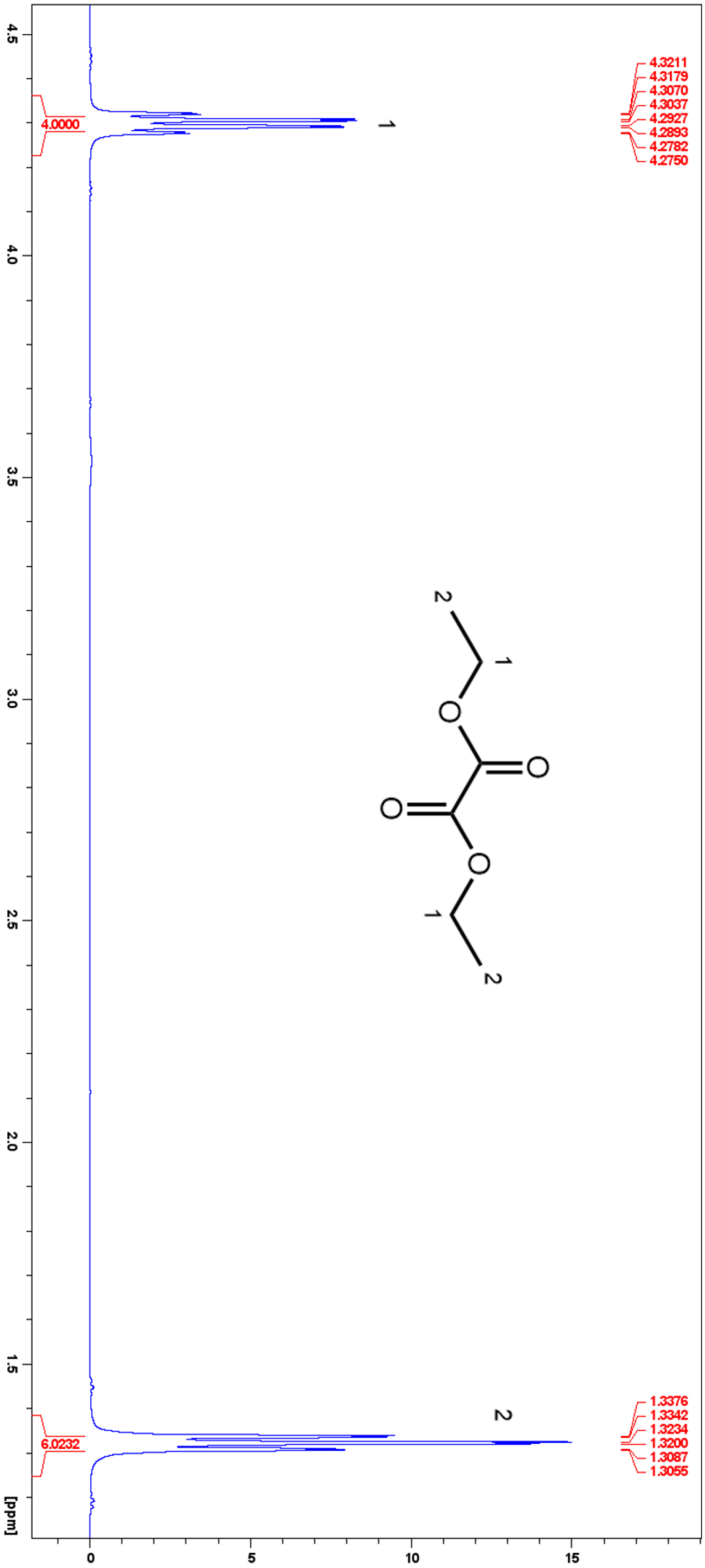


Figure A3.20: ^{1}H NMR of **45**.

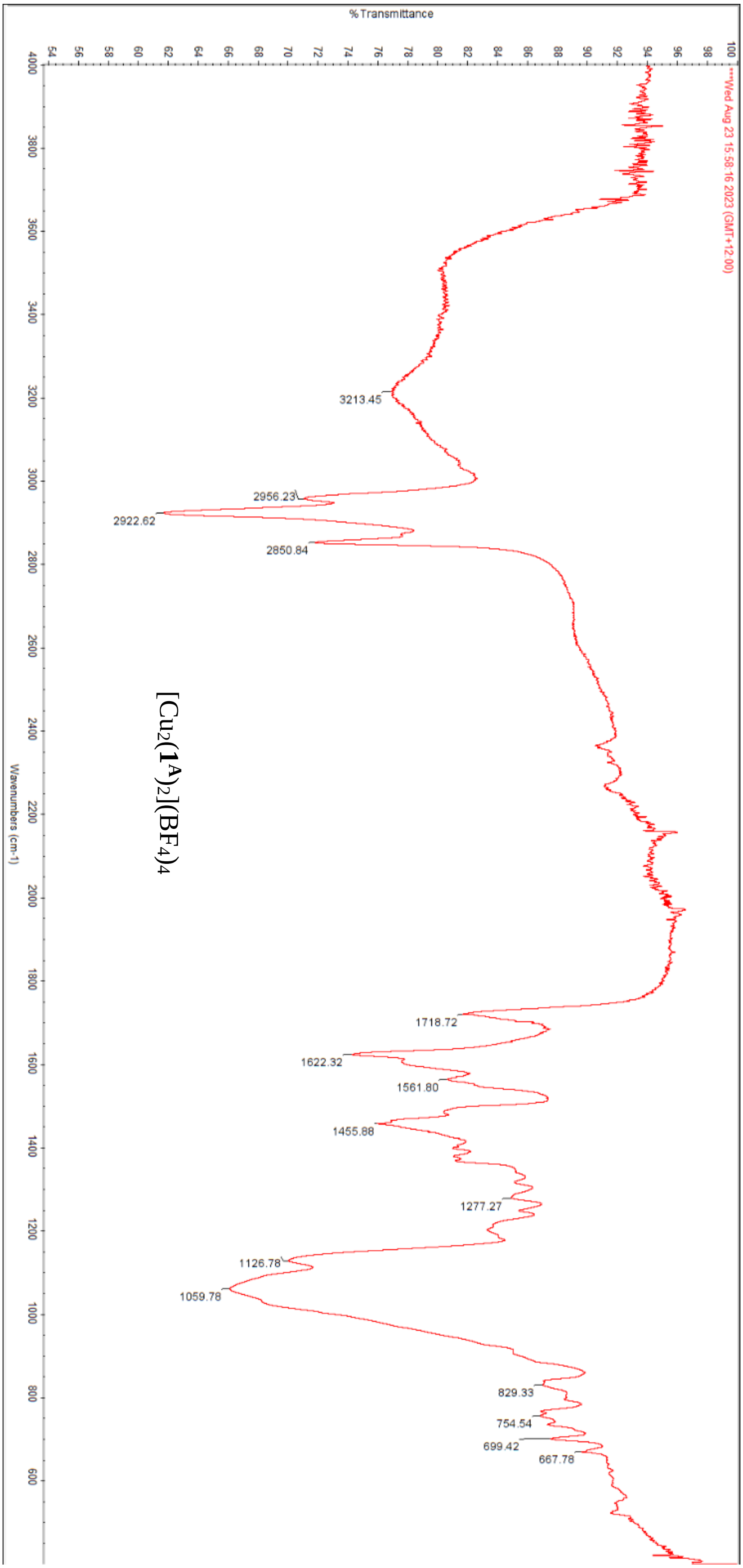


Figure A3.21: IR of $[Cu_2(\mathbf{1}^A)_2](BF_4)_4$.

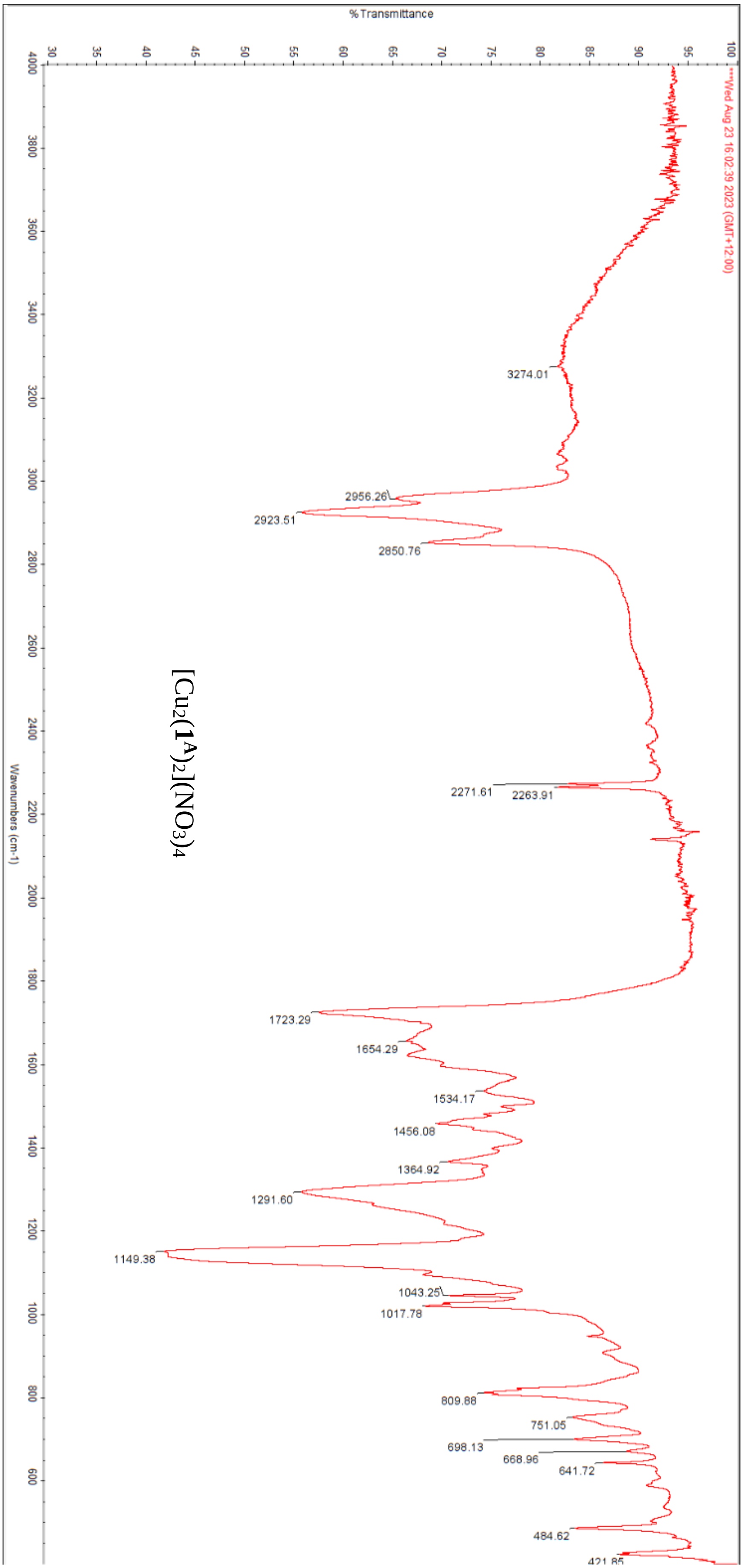


Figure A3.22: IR of [Cu$_2$(**1^A**)$_2$](NO$_3$)$_4$.

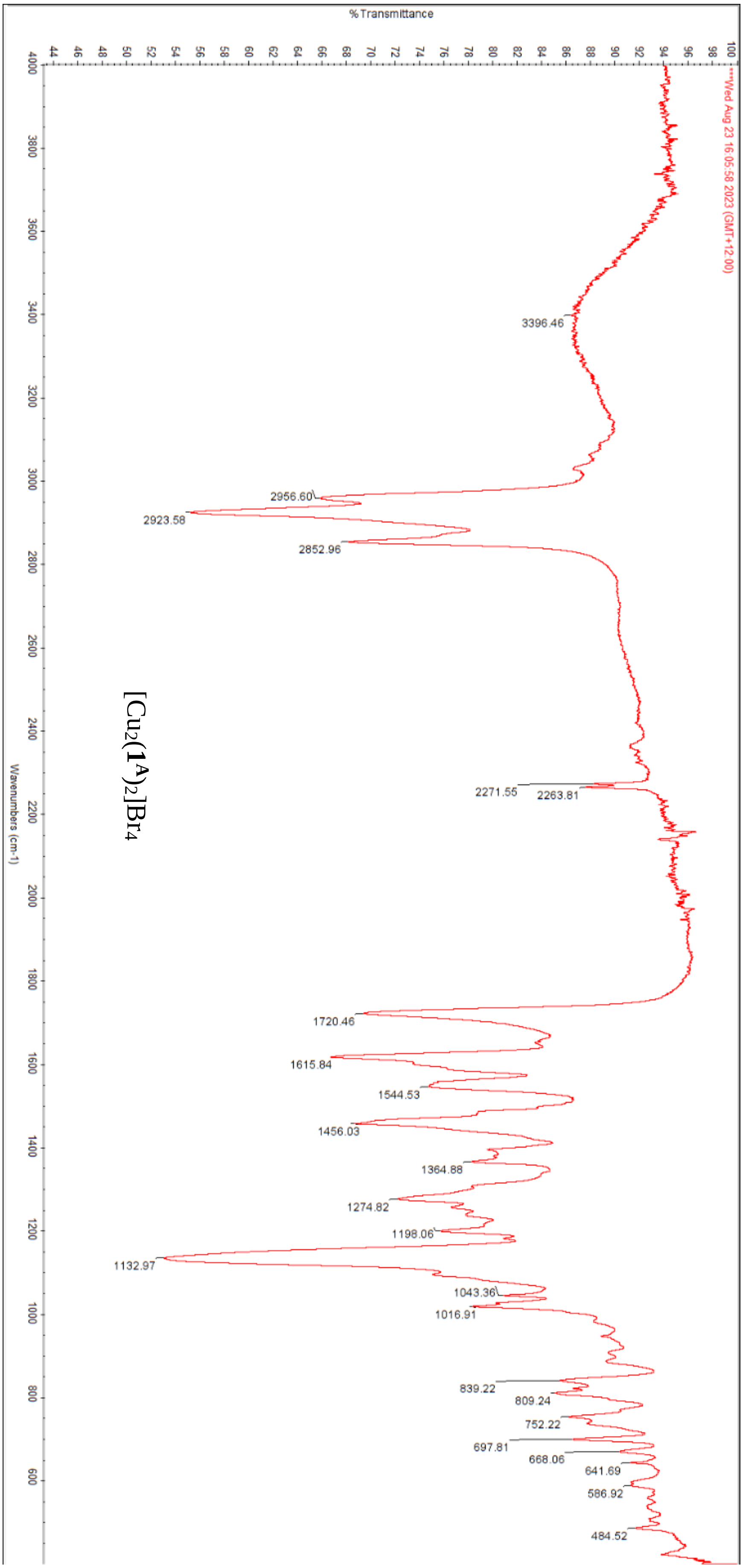


Figure A3.23: IR of $[Cu_2(\mathbf{1}^A)_2]Br_4$.

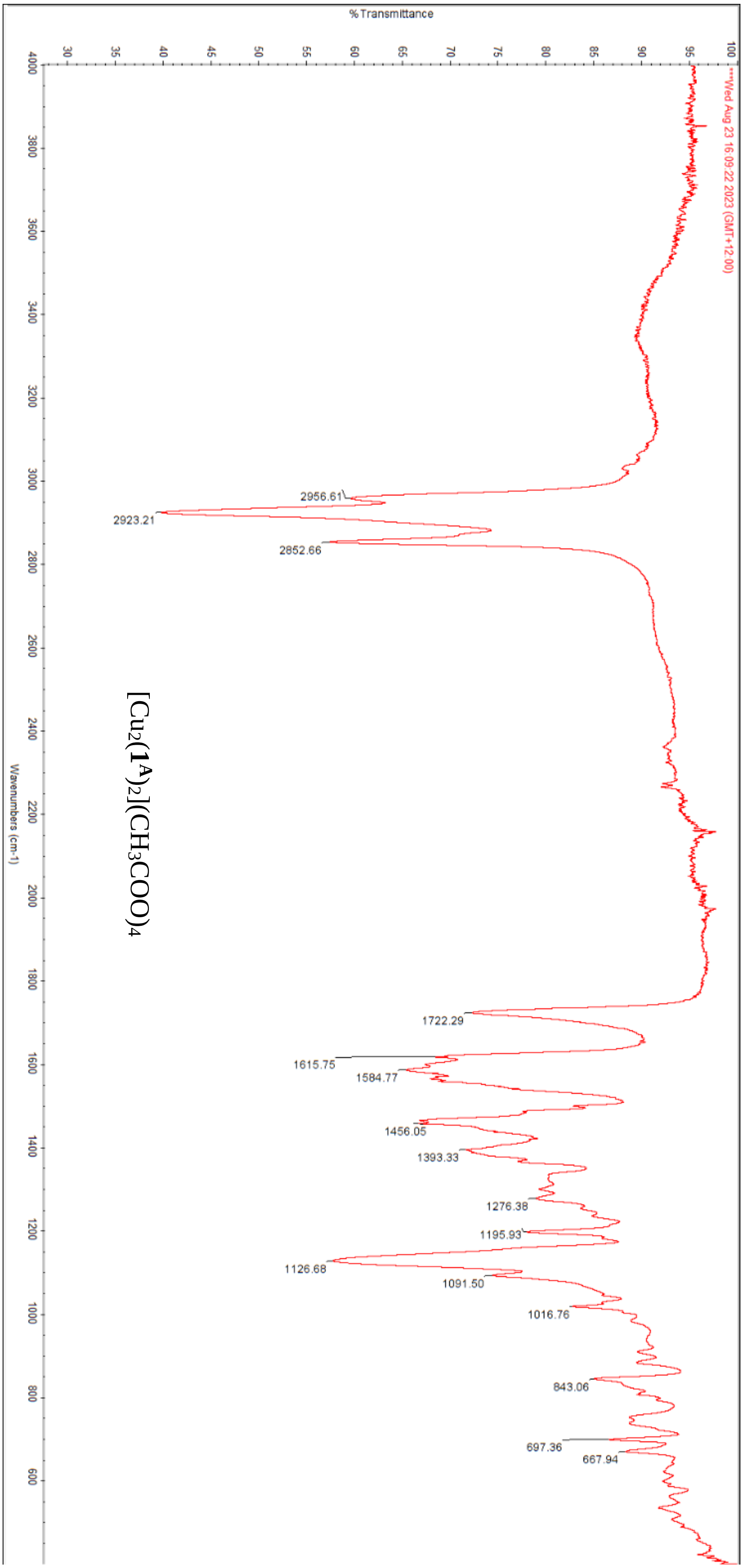


Figure A3.24: IR of [Cu$_2$(**1^A**)$_2$](CH$_3$COO)$_4$.

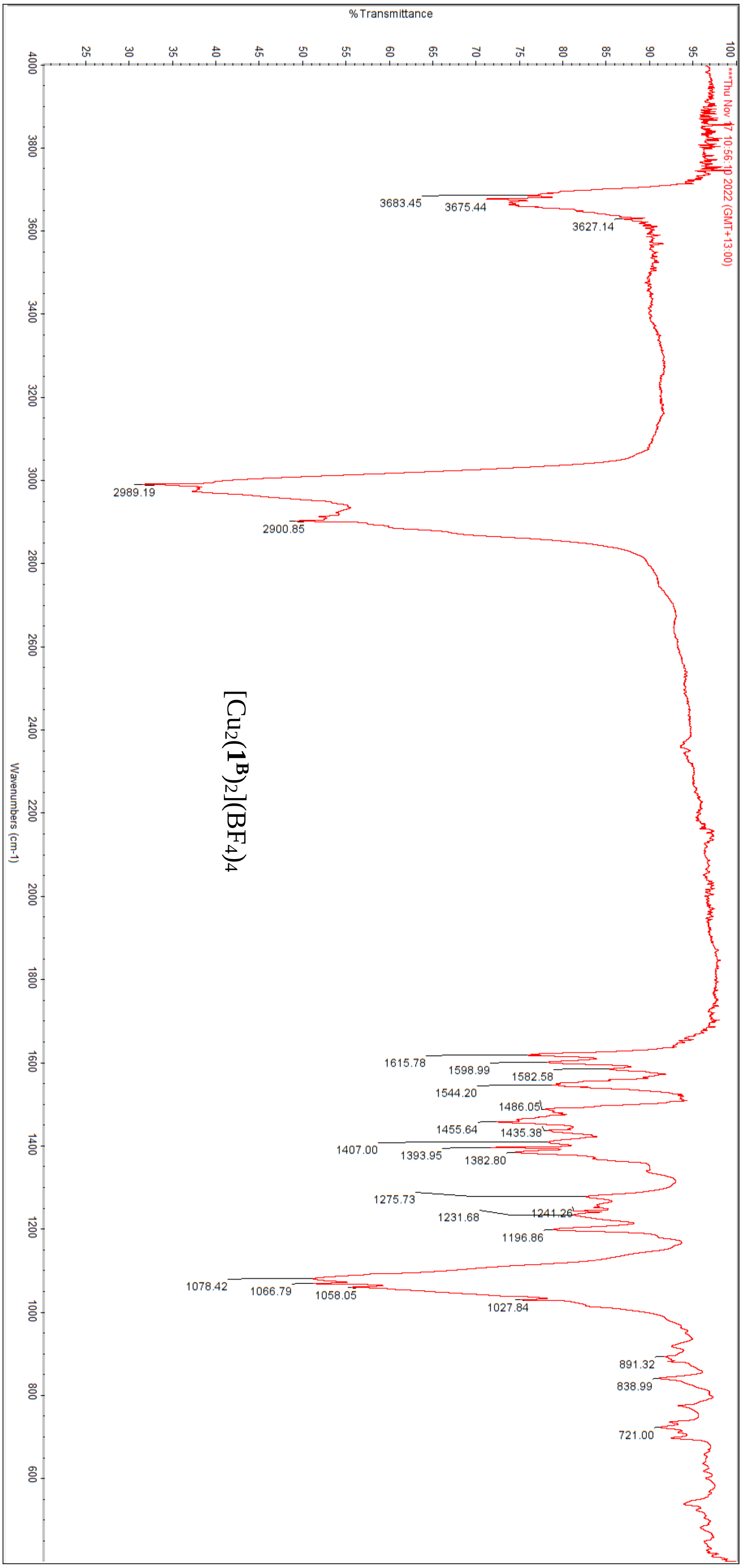


Figure A3.25: IR of [Cu$_2$($\mathbf{1^B}$)$_2$](BF$_4$)$_4$.

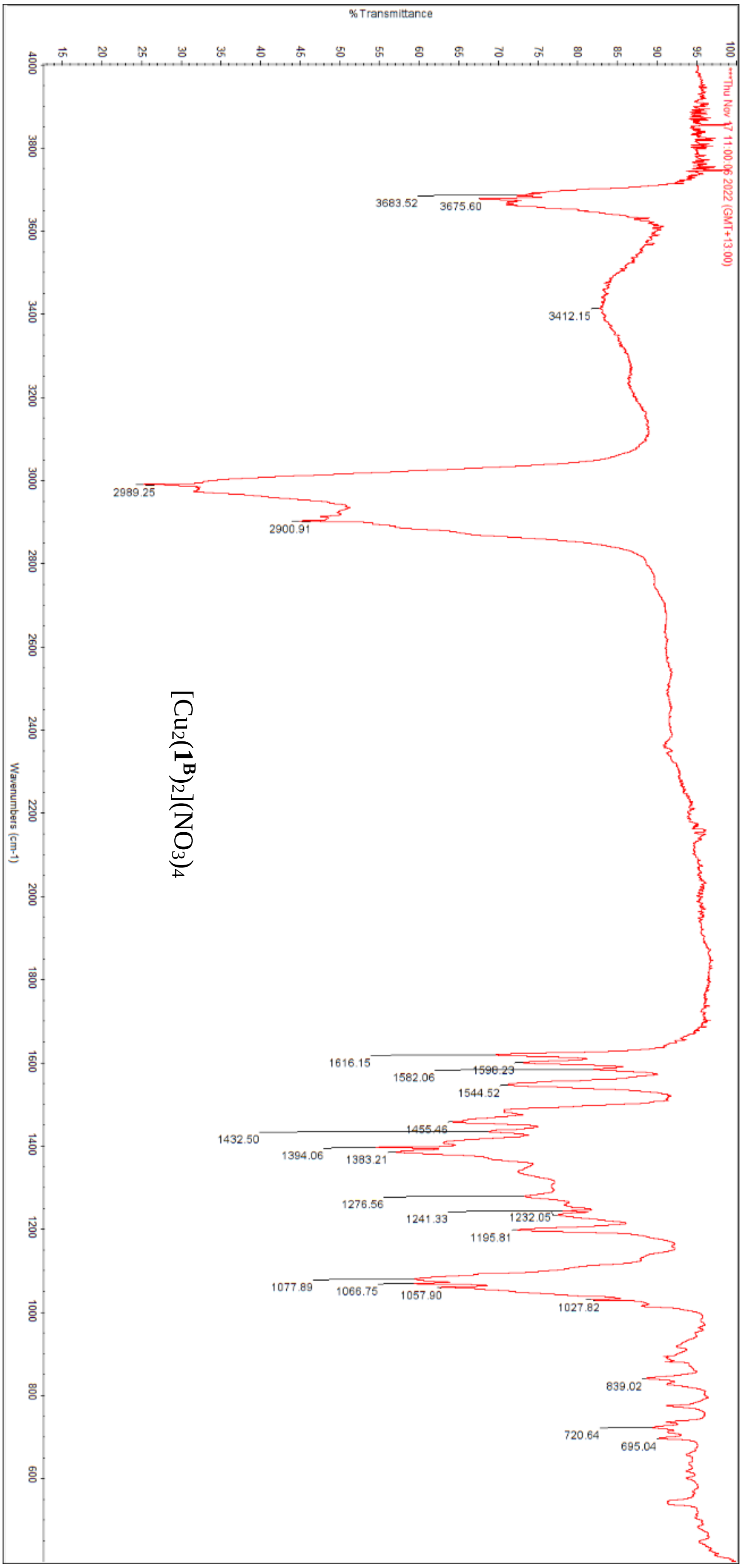


Figure A3.26: IR of $[Cu_2(\mathbf{1}^B)_2](NO_3)_4$.

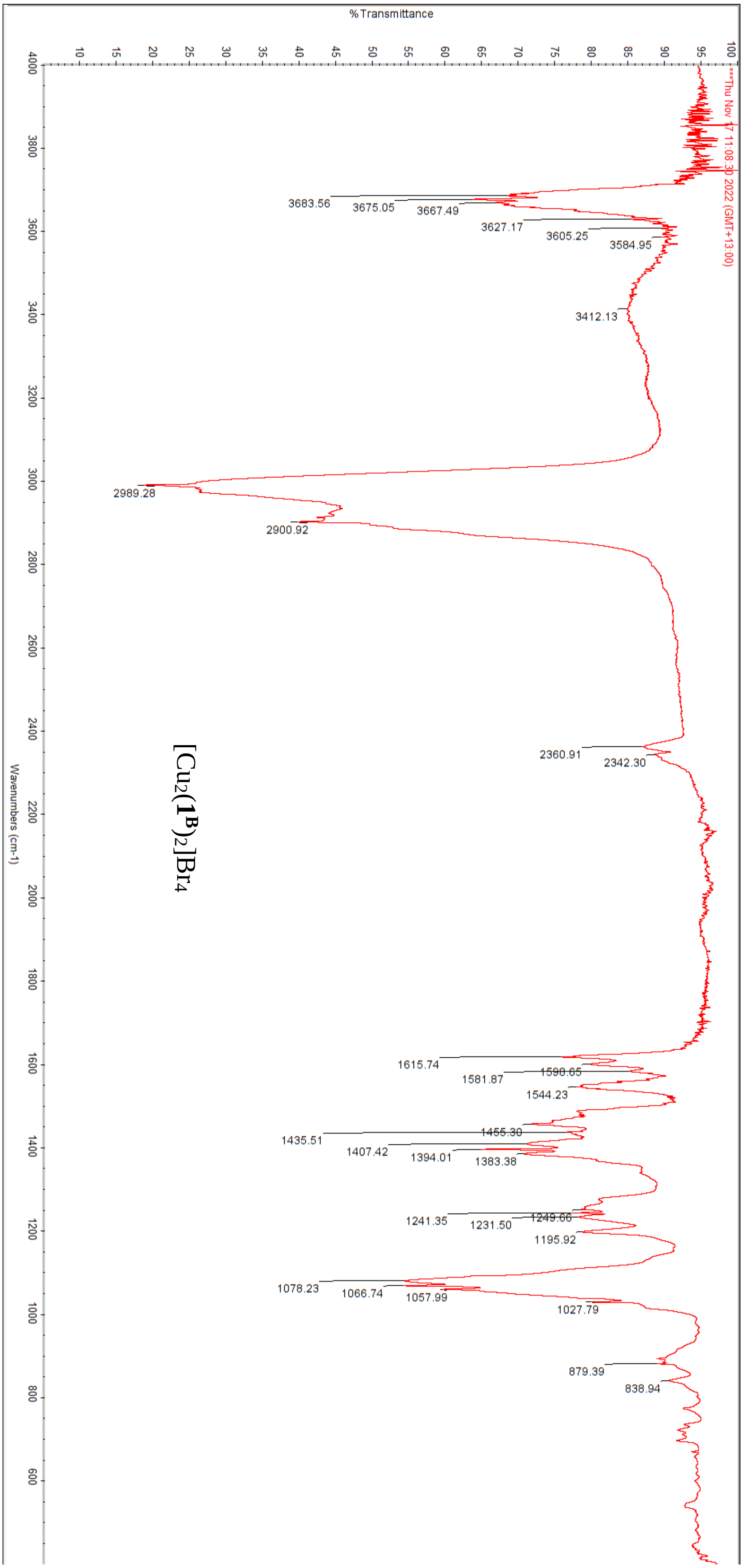


Figure A3.27: IR of $[Cu_2(\mathbf{1^B})_2]Br_4$.

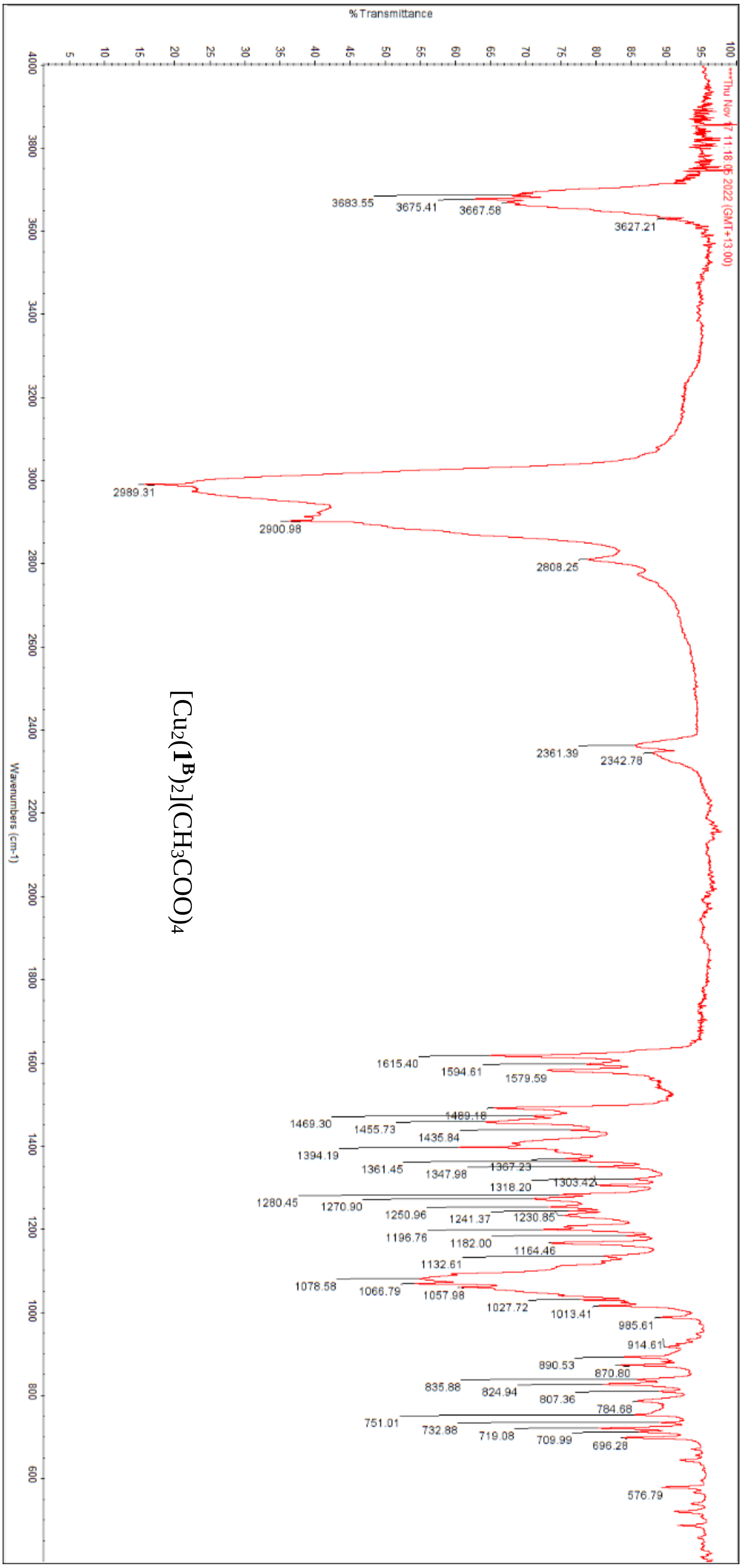


Figure A3.28: IR of $[Cu_2(\mathbf{1^B})_2](CH_3COO)_4$.

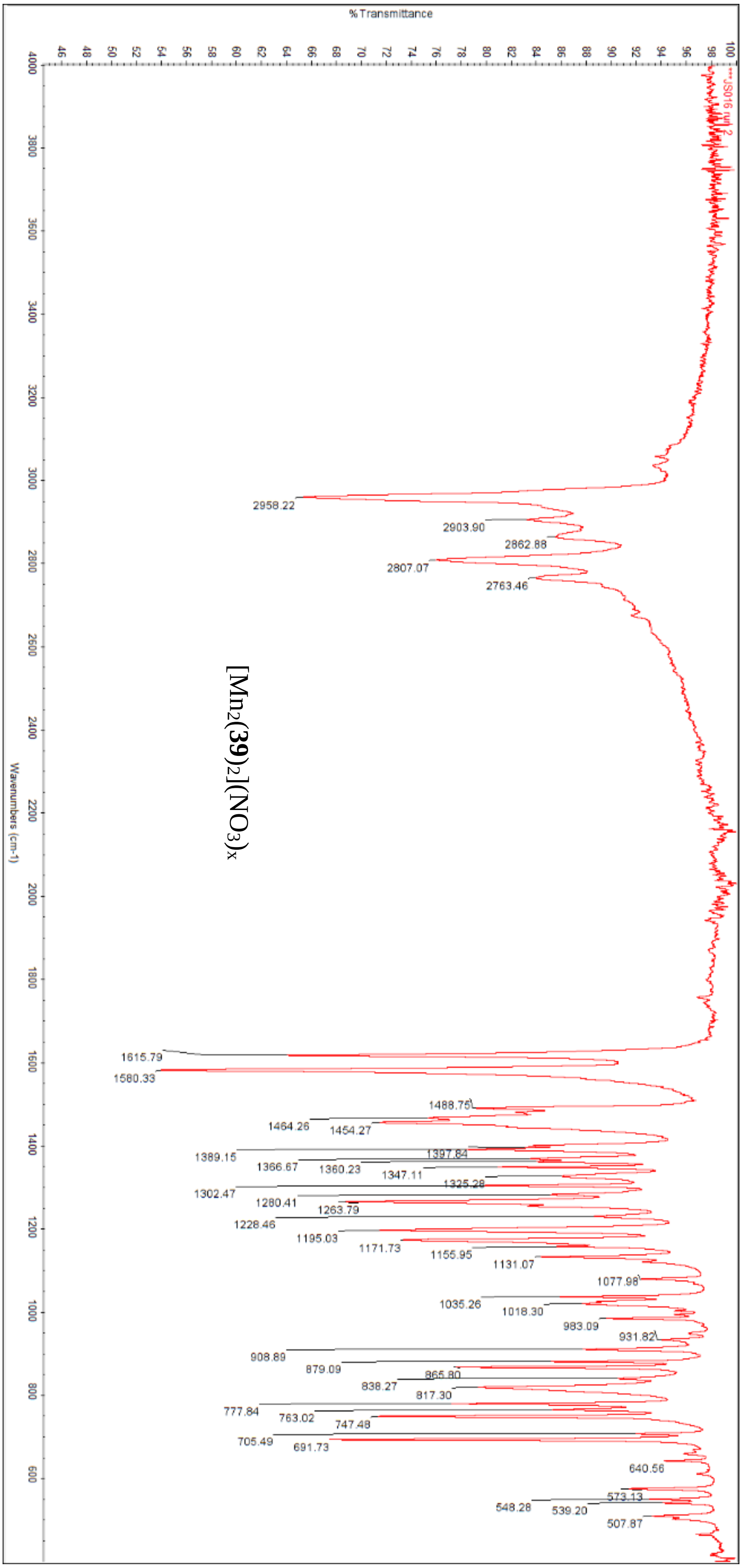


Figure A3.29: IR of $[Mn_2(\mathbf{39})_2](NO_3)_x$.

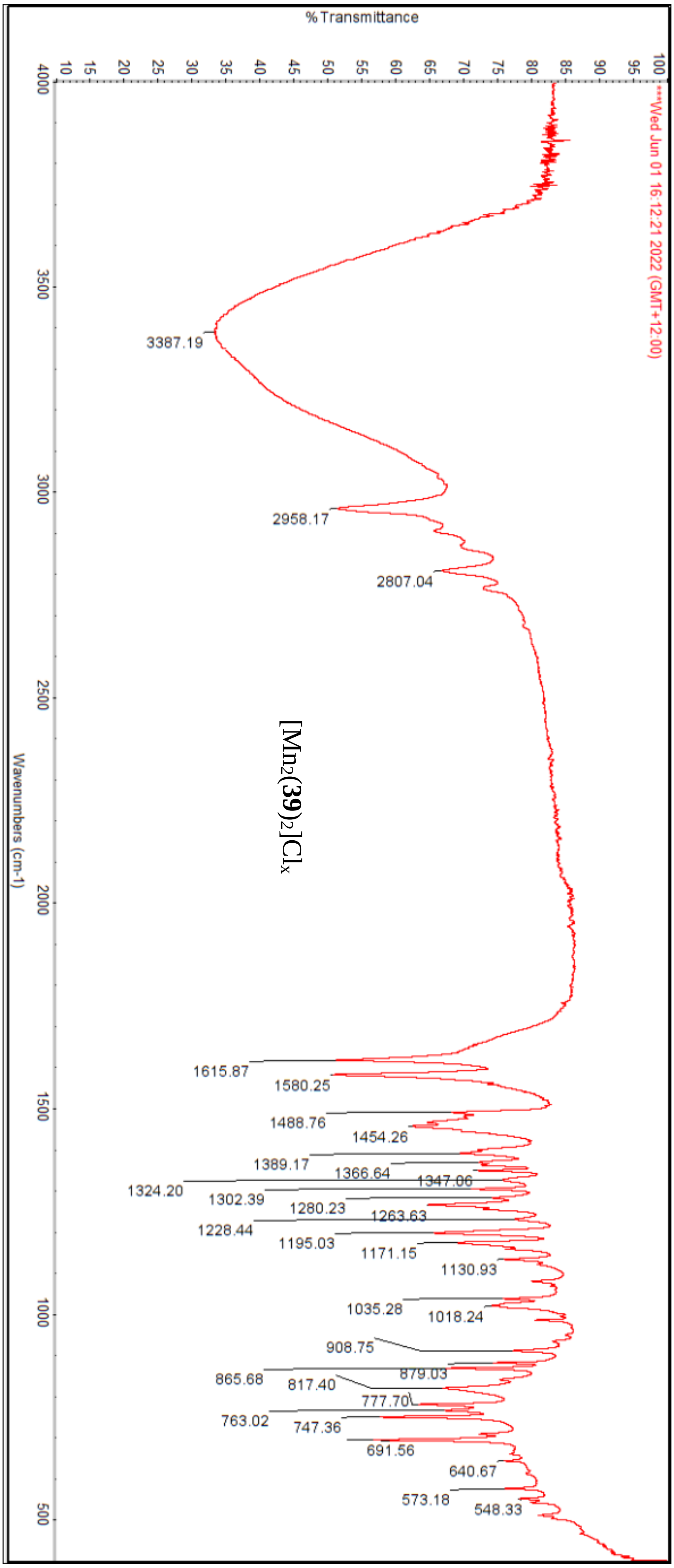


Figure A3.30: IR of $[Mn_2(\mathbf{39})_2]Cl_x$.

Appendix 4 – Raw Python Code

The following lines of code are a simple example of how to use the functions outlined in Chapter 2.

```
# For filtering spectrum, use %matplotlib inline otherwise use
%matplotlib widget

%matplotlib widget
import Data_analysis as da
directory_address = r'C:\Users\Desktop\Python example directory'
home = directory_address
data_dict = {}
folder_dict = {}


da.print_files(home_dir = home, technique = 'Raman', folder_dict =
folder_dict)


file_list_Raman = ['File 1','File 2','File 3','File 4','File 5','File 6']
da.process_data(file_list_Raman, technique = 'Raman', folder_dict =
folder_dict, data_dict = data_dict)


plot_list = [['File 1','File 2'],['File 3','File 4','File 5','File 6']]
da.quick_plot(plot_list = plot_list, plot_orientation = 'vert', technique
= 'Raman', data_dict = data_dict, label_peaks = False, auto_axis =
'both', save_figure = True, outputdir = home)
```

The following is the full code for the functions outlined in Chapter 2 that has been saved with the file name “Data_analysis”.

```
''' Useful functions to import form here are print_files (which prints
the files in folders and adds the addresses to a dictionary),
process_data (which extracts all the data from given files and adds it to
    a dictionary), peak_labels (puts the peak labels on the graph),
my_r2_score (can be used to calculate the R^2 value of a dataset) and
quick_plot (quickly plots a given dataset)'''
#   from Data_analysis import print_files, process_data, peak_labels,
my_r2_score, quick_plot

import os
import time, sys
import numpy as np
import pandas as pd
import scipy.sparse as sparse
import matplotlib.pyplot as plt
import scipy.sparse.linalg as splinalg
from scipy.signal import find_peaks
from sklearn.metrics import r2_score
from colorama import Fore, Back, Style
from IPython.display import clear_output
from sklearn.linear_model import LinearRegression
```

```
NMR_files = ['acqu', 'acqu2', 'acqu2s', 'acqus', 'nuslist', 'fid',
'prosol_History', 'pulseprogram', 'scon2', 'ser', 'shimvalues',
'specpar', 'topshimreport', 'vtc_pid_settings', 'waltz65', 'garp', 'EA',
'automatic', 'zg30', 'zgpg30', 'deptsp135', 'noesygpph', 'cosygpmfqf',
'hmbcetgpl3nd', 'cosylrqf', 'deptsp90', '1i', '1r', '2ii', '2ir', '2ri',
'2rr', 'assocs', 'clevels', 'curdat2', 'base_info', 'baslpnts', 'intrng',
'outd', 'peaklist', 'peakrng', 'peaks', 'proc', 'proc2', 'proc2s',
'procs', 'title']

NMR_folder_list = ['pdata', 'lists', 'form', 'pp', 'cpd', 'wave', 'gp',
'wave', 'vc', 'olex2']

colours = ['#000000', '#FF0000', '#00FF00', '#0000FF', '#B0B0B0',
'#FF00FF', '#A47C43', '#FF8300', '#3BA32F', '#931FDF', '#A02222',
'#F3D438', '#00FFFF', '#FFFF00']

colour_names = ['Black', 'Red', 'Green', 'Blue', 'Grey', 'Pink', 'Brown',
'Orange', 'Dark Green', 'Purple', 'Maroon', 'Gold', 'Light Blue',
'Yellow']

def comparing_lists(list_1, list_2):
    ''' Looks at the first list and if anything in the first list matches
anything in the second list, it returns true'''
    for z in range(len(list_1)):
        if all(list_1[z] != a for a in list_2) != True:
            return True

def check_file_type(file_list, file_list_address, string_to_find,
file_type, technique):
    ''' Looks into each file in the list and checks if the string is
contained in the first line of the file'''
    for z in range(len(file_list)):
        if file_list[z].startswith('.') != True:
            if file_type == '.txt': # For Single Fluorescence Files and
UV Files
                if file_list[z].endswith(file_type):
                    if open(os.path.join(file_list_address,
file_list[z])).read(1) == string_to_find:
                        return True
            if file_type == '.xlsx': # For Fluorescence Scan Files
                if file_list[z].endswith(file_type):
                    if (pd.read_excel(os.path.join(file_list_address,
file_list[z]))).columns.values[0] == string_to_find:
                        return True
            if technique == 'Raman':
                if file_list[z].endswith(file_type):
                    if open(os.path.join(file_list_address,
file_list[z])).read(1).isdigit():
                        return True

def investigate(folder, file_type, NMR_files, file_startswith,
technique):
    ''' Looks two folder steps in from the given folder and returns true
if there is a file with the file type in the name'''
    for x in range(len(os.listdir(folder))):
        if all(os.listdir(folder)[x] != a for a in NMR_files) == True:
            if os.path.isfile(os.path.join(folder,
os.listdir(folder)[x])) != True:
```

```
                if any (file_type in a for a in
(os.listdir(os.path.join(folder, os.listdir(folder)[x])))):
                    if file_type == '.txt':
                        for z in
range(len(os.listdir(os.path.join(folder, (os.listdir(folder)[x]))))):
                            if ((os.listdir(os.path.join(folder,
(os.listdir(folder)[x]))))[z]).endswith(file_type):
                                if open(os.path.join(folder,
(os.listdir(folder)[x]), (os.listdir(os.path.join(folder,
(os.listdir(folder)[x]))))[z])).read(1).startswith(file_startswith):
                                    return True
                    if file_type == '.xlsx':
                        for z in
range(len(os.listdir(os.path.join(folder, (os.listdir(folder)[x]))))):
                            if ((os.listdir(os.path.join(folder,
(os.listdir(folder)[x]))))[z]).endswith(file_type):
                                if pd.read_excel(os.path.join(folder,
(os.listdir(folder)[x]), (os.listdir(os.path.join(folder,
(os.listdir(folder)[x]))))[z])).columns.values[0].startswith(file_startsw
ith):
                                    return True
                    elif file_type != '.txt' and file_type != '.xlsx':
                        if technique == 'Raman':
                            for z in
range(len(os.listdir(os.path.join(folder, (os.listdir(folder)[x]))))):
                                if ((os.listdir(os.path.join(folder,
(os.listdir(folder)[x]))))[z]).endswith(file_type):
                                    if (open(os.path.join(folder,
(os.listdir(folder)[x]), (os.listdir(os.path.join(folder,
(os.listdir(folder)[x]))))[z])).read(1)).isdigit():
                                        return True
                        else:
                            return True

def update_progress(progress):
    bar_length = 80
    if isinstance(progress, int):
        progress = float(progress)
    if not isinstance(progress, float):
        progress = 0
    if progress < 0:
        progress = 0
    if progress >= 1:
        progress = 1

    block = int(round(bar_length * progress))
    clear_output(wait = True)
    text = f"{'Progress: |' + ('█' * block) + (' ' * (bar_length -
block)) + '| ' + str(progress * 100) + '%'}"
    print(text)

def print_files(home_dir = 'a', technique = 'Raman', folder_dict = 'a',
NMR_files = NMR_files):
    ''' Technique can be Raman (.csv), IR (.CSV), UV (.txt), Fluorescence
(.xlsx or .txt).
    This returns a list of files that fit the technique type within all
the folders in a given directory'''
    file_startswith = 'Nothing unless file is .txt'
    folder_list = []
```

```
    printing_list = []
    for address, dirs, files in os.walk(home_dir):
        if (os.path.split(address)[1] in NMR_folder_list) != True:
            if os.path.split(address)[1].isdigit() != True:
                folder_list.append(os.path.split(address)[1])
    if technique == 'Raman':
        file_type = '.csv'
    if technique == 'IR':
        file_type = '.CSV'
    if technique == 'UV':
        file_type = '.txt'
        file_startswith = '"'
    if technique == 'Fluorescence':
        type_of_file = input('"Scan" or "Single" spectra?')
        if type_of_file == 'Scan' or type_of_file == 'scan':
            type_of_file = 'Scan'
            file_type = '.xlsx'
            file_startswith = 'WL [nm]'
        elif type_of_file == 'Single' or type_of_file == 'single':
            type_of_file = 'Single'
            file_type = '.txt'
            file_startswith = 'W'
    for address, dirs, files in os.walk(home_dir):
        folder_of_file = os.path.split(address)
        folder_of_folder = os.path.split(folder_of_file[0])
        if (folder_of_file[1] in NMR_folder_list) != True:
            if folder_of_file[1].isdigit() != True:

update_progress((folder_list.index(folder_of_file[1])+1)/(len(folder_list
)))
        if investigate(address, file_type, NMR_files, file_startswith,
technique) == True:
            printing_list.append(folder_of_file[1])
        if any (file_type in z for z in files):
            if file_type == '.txt' or file_type == '.xlsx' or technique
== 'Raman':
                if check_file_type(files, address, file_startswith,
file_type, technique) == True:
                    printing_list.append('  ' + folder_of_file[1])
            if file_type != '.txt' and file_type != '.xlsx' and technique
!= 'Raman':
                printing_list.append('  ' + folder_of_file[1])
        for a in range(len(files)):
            if file_type in files[a]:
                if file_type == '.txt':
                    if open(os.path.join(address,
files[a])).read(1).startswith(file_startswith):
                        folder_dict[files[a][:-4]] = address
                        printing_list.append('    ' + files[a][:-4])
                if file_type == '.xlsx':
                    if pd.read_excel(os.path.join(address,
files[a])).columns.values[0].startswith(file_startswith):
                        folder_dict[files[a][:-5]] = address
                        printing_list.append('    ' + files[a][:-5])
                if technique == 'Raman':
                    if open(os.path.join(address,
files[a])).read(1).isdigit():
                        folder_dict[files[a][:-4]] = address
                        printing_list.append('    ' + files[a][:-4])
```

```
                if file_type != '.txt' and file_type != '.xlsx' and
technique != 'Raman':
                    folder_dict[files[a][:-4]] = address
                    printing_list.append('    ' + files[a][:-4])
        if (folder_of_file[1] in NMR_folder_list) != True:
            if folder_of_file[1].isdigit() != True:
                for a in range(len(printing_list)):
                    if printing_list[a].startswith('    '):
                        print('    ' + Back.LIGHTWHITE_EX +
printing_list[a][4:] + Style.RESET_ALL)
                    else:
                        print(printing_list[a])

def xlsx_analysis(file_list, data_dict = 'a', folder_dict = 'a'):
    ''' Assumes the first column is the absorbance wavelength and the
column headers are the excitation wavelengths. Will return an input to
the data_dict as a dictionary containing a dataframe of the data
(Dataframe), excitation wavelengths ('Exc WL'), absorbance wavelengths
('Abs WL') and each individual data series for each exitation
wavelength'''
    for x in range(len(file_list)):
        update_progress((x + 1)/(len(file_list)))
        data_dict[file_list[x]] = {}
        file_data = pd.read_excel(os.path.join(folder_dict[file_list[x]],
file_list[x] + '.xlsx'), header = 0)
        key_list = ['Dataframe', 'Exc WL', 'Abs WL']
        for y in file_data.columns[1:]:
            key_list.append(y[4:-1])
        for z in range(len(key_list)-2):
            if z == 0:
                data_dict[file_list[x]][key_list[z]] = file_data
            if z == 1:
                data_dict[file_list[x]][key_list[z]] = [float(a[4:-1])
for a in file_data.columns[1:]]
            data_dict[file_list[x]][key_list[z+2]] =
pd.read_excel(os.path.join(folder_dict[file_list[x]], file_list[x] +
'.xlsx'), header = None, skiprows = [0], usecols = [z])

def process_data(file_list, folder_dict = 'a', data_dict = 'a', technique
= 'Raman'):
    ''' Technique can be Raman (.csv), IR (.CSV), UV (.txt), or
Fluorescence (.xlsx or .txt). Raman assumes that the csv file has the
separate frames in a single column. This processes the files and returns
the data in the data dictionary with the file name as the key. For
accurate spectrum filtering of Raman data, %matplotlib widget can't be
used'''
    file_startswith = 'Nothing unless file is .txt'
    separator = ','
    if technique == 'Raman':
        file_type = '.csv'
        header = input("Header? If no header don't input anything")
        if header == '':
            header = None
    if technique == 'IR':
        file_type = '.CSV'
        header = None
    if technique == 'UV':
        file_type = '.txt'
        header = 1
```

```
    if technique == 'Fluorescence':
        type_of_file = input('"Scan" or "Single" spectra?')
        if type_of_file == 'Scan' or type_of_file == 'scan':
            type_of_file = 'Scan'
            file_type = '.xlsx'
            header = 0
        elif type_of_file == 'Single' or type_of_file == 'single':
            type_of_file = 'Single'
            file_type = '.txt'
            header = 0
            separator = '\t'
    if technique == 'Raman':
        #assumes csv file contains separate frames, as blocks in a single
column
        N_pixels = input("Number of pixels? (if 1024 for the FERGIE don't
input anything)")
        if N_pixels == '':
            N_pixels = 1024
        elif N_pixels != '':
            N_pixels = int(N_pixels)
        print(Back.LIGHTWHITE_EX + "\nFor accurate spectrum filtering of
Raman data, %matplotlib widget can't be used\n%matplotlib inline works
well for filtering\n" + Style.RESET_ALL)
        filtering = input("Filter the spectrum? (if not don't input
anything)")
        for x in range(len(file_list)):
            update_progress((x)/(len(file_list)))
            df_WN = pd.read_csv(os.path.join(folder_dict[file_list[x]],
file_list[x] + file_type), header = header, usecols = [0], nrows =
N_pixels).values
            intensity =
(pd.read_csv(os.path.join(folder_dict[file_list[x]], file_list[x] +
file_type), header = header, usecols = [1])).values
            N_frames = int(intensity.shape[0]/N_pixels)
            intensity_array =
np.transpose(intensity.reshape(N_frames,N_pixels))
            data_dict[file_list[x] + ' WN'] = df_WN
            if filtering == '':
                baseline = baseline_als(np.mean(intensity_array, axis =
1), 100000, 0.01, niter = 10)
                data_dict[file_list[x]] = np.mean(intensity_array, axis =
1) - baseline
            elif filtering != '':
                plt.plot(df_WN, intensity_array)
                plt.minorticks_on()
                plt.grid(which = 'major', color = '0.55', linestyle = '-
')
                plt.grid(which = 'minor', color = '0.4', linestyle = ':')
                plt.show()
                x_focus = int(input('Wavelength to compare'))
                x_index = (np.abs(df_WN - x_focus)).argmin()
                x_lims = [np.min(intensity_array[x_index,:]),
np.max(intensity_array[x_index,:])]
                plt.close()
                plt.hist(intensity_array[x_index,:], bins =
np.linspace(x_lims[0], x_lims[1], 50))
                plt.show()
                print('Number of spectra before filtering:',
intensity_array.shape[1])
```

```
                cutoff = input("Filter above or below the threshold
(Don't input anything if above)")
                threshold = input('Enter the threshold intensity:')
                threshold = np.float(threshold)
                if cutoff == '': #above
                    bad = intensity_array[x_index,:] > threshold
                    bad_spec = np.where(bad)
                    print('Number of spectra above threshold
intensity:',bad_spec[0].shape[0])
                    print('Indices of deleted spectra:', bad_spec)
                if cutoff != '': #below
                    bad = intensity_array[x_index,:] < threshold
                    bad_spec = np.where(bad)
                    print('Number of spectra below threshold
intensity:',bad_spec[0].shape[0])
                    print('Indices of deleted spectra:', bad_spec)
                # use the indices of the bad frames to remove the
corresponding spectra from the data frame
                good = np.delete(intensity_array, bad_spec[0], axis = 1)
                print('Number of spectra after filtering:',
good.shape[1])
                baseline = baseline_als(np.mean(good, axis = 1), 100000,
0.01, niter = 10)
                data_dict[file_list[x]] = np.mean(good, axis = 1) -
baseline
                plt.pause(3)
        update_progress(1)

    elif technique == 'IR' or technique == 'UV' or (technique ==
'Fluorescence' and type_of_file == 'Single'):
        for x in range(len(file_list)):
            update_progress((x + 1)/(len(file_list)))
            intensity =
(pd.read_csv(os.path.join(folder_dict[file_list[x]], file_list[x] +
file_type), header = header, sep = separator, usecols = [1])).values
            df_WN = pd.read_csv(os.path.join(folder_dict[file_list[x]],
file_list[x] + file_type), header = header, sep = separator, usecols =
[0])
            if technique == 'IR':
                data_dict[file_list[x]] = (np.mean(intensity ,axis =
1)[:-1])
                data_dict[file_list[x] + ' WN'] = ((df_WN.values)[:-1])
            elif technique == 'UV' or (technique == 'Fluorescence' and
type_of_file == 'Single'):
                data_dict[file_list[x]] = np.mean(intensity ,axis = 1)
                data_dict[file_list[x] + ' WL'] = df_WN.values

    elif technique == 'Fluorescence' and type_of_file == 'Scan':
        xlsx_analysis(file_list = file_list, data_dict = data_dict,
folder_dict = folder_dict)

def baseline_als(y, lam, p, niter=10):
    L = len(y)
    D = sparse.csc_matrix(np.diff(np.eye(L), 2))
    w = np.ones(L)
    for i in range(niter):
        W = sparse.spdiags(w, 0, L, L)
        Z = W + lam * D.dot(D.transpose())
        z = splinalg.spsolve(Z, w*y)
```

```
        w = p * (y > z) + (1-p) * (y < z)
    return z

def peak_labels(x, y, plot, technique):
    if technique == 'IR':
        y = y*-1+100
    baseline = baseline_als(y, 100000, 0.01, niter=10)
    baselined = y - baseline
    if technique == 'Raman':
        peaks,_ = find_peaks(y, width = [1, 30], prominence = (100,
None), height = (baseline*1.1, None), threshold = (5, None), distance =
1)
    if technique == 'IR':
        peaks,_ = find_peaks(y, prominence = 1, height = 10, distance =
250)
    if technique == 'UV' or technique == 'Fluorescence':
        peaks,_ = find_peaks(y, height = baseline*1.15, distance = 10)
    for j in peaks:
        if technique != 'IR':
            plot.annotate('{0:.0f}'.format(float(x[j])), xy = (x[j],
y[j]), xycoords = 'data', xytext = (0, +15), textcoords = 'offset
points', fontsize = 12,
                          rotation = 90, clip_on = True, horizontalalignment
= 'center', verticalalignment = 'bottom', arrowprops = dict(arrowstyle="-
”, connectionstyle="arc3,rad=.0"))
        if technique == 'IR':
            plot.annotate('{0:.0f}'.format(float(x[j])), xy = (x[j],
(y[j]*-1+100)), xycoords = 'data', xytext = (0, -15), textcoords =
'offset points', fontsize = 12,
                          rotation = 90, clip_on = True, horizontalalignment
= 'center', verticalalignment = 'top', arrowprops = dict(arrowstyle="-”,
connectionstyle="arc3,rad=.0"))

def regr_line(x, y, ax = None, xlabel = 'I', ylabel = '[x]', color
='black', showlegend = False):
    '''Performs simple linear regression and creates a plot which can be
added to any existing axis via ax'''
    if ax is None:
        ax = plt.gca()
    model = LinearRegression()
    model.fit(x,y)
    b0 = model.intercept_
    b1 = model.coef_
    if np.min(x) > 0 and np.max(x) > 0:
        new_x = np.linspace(0.95 * np.min(x), 1.05 * np.max(x), 100)
    elif np.min(x) < 0 and np.max(x) > 0:
        new_x = np.linspace(1.05 * np.min(x), 1.05 * np.max(x), 100)
    else: # both min and max are negative
        new_x = np.linspace(1.05 * np.min(x), 0.95 * np.max(x), 100)
    regr_line = b0 + b1 * new_x
    line = ax.plot(new_x, regr_line, ls = '--', color = color)
    if showlegend:
        ax.legend(line, [f'{ylabel} = {b0:{5}.{3}} + {b1[0]:{5}.{3}} x
{xlabel}'])
    return

def my_r2_score(y_axis_values, x_axis_values, colours = 'a'):
    if colours == 'a':
        colours = '#000000' #Black
```

```
    line_regrs = LinearRegression()
# If a graph is plotted, this will fit a regression line
    x_axis_values = np.asarray(x_axis_values).reshape(-1, 1)
    regr_line(x_axis_values, y_axis_values, color = colours)
    line_regrs.fit(x_axis_values, y_axis_values)
    y_axis_pred = line_regrs.predict(x_axis_values)
    r2_value = r2_score(y_axis_values, y_axis_pred)
    return r2_value

def quick_plot(plot_list = 'list containing the lists of data to be
plotted', plot_orientation = 'vertical', technique = 'a', data_dict =
'a', save_figure = False, outputdir = 'only needed if saving a figure',
label_peaks = False, auto_axis = 'both'):
    # If just one graph to be plotted, have one list in plot_list list
i.e. [[dataset_1, dataset_2]]
    # If setting axis yourself, set auto_axis to 'x' (you want to set y
lims), 'y' (you want to set x lims) or 'none' (you want to set x and y
lims)
    # Can't handle fluorescent scans
    if 'vert' in plot_orientation:
        fig, ax = plt.subplots(len(plot_list), 1,
figsize=(6,4*len(plot_list)))
    if 'hor' in plot_orientation:
        fig, ax = plt.subplots(1, len(plot_list),
figsize=(6*len(plot_list),4))
    if technique == 'Raman' or technique == 'IR':
        wave_type = ' WN'
        xlabel = r'Wavenumber (cm$^{-1}$)'
    if technique == 'Fluorescence' or technique == 'UV':
        wave_type = ' WL'
        xlabel = 'Wavelength (nm)'

    if len(plot_list) == 1:
        new_y_lims = []
        for a in range(len(plot_list[0])):
            ylims = [np.min(data_dict[plot_list[0][a]]),
np.max(data_dict[plot_list[0][a]])]
            if ylims[0] > 0 and ylims[1] > 0:
                new_y_lims.append(0.9 * ylims[0])
                new_y_lims.append(1.1 * ylims[1])
            elif ylims[0] < 0 and ylims[1] > 0:
                new_y_lims.append(1.1 * ylims[0])
                new_y_lims.append(1.1 * ylims[1])
            elif ylims[0] < 0 and ylims[1] < 0:
                new_y_lims.append(1.1 * ylims[0])
                new_y_lims.append(0.9 * ylims[1])
            ax.plot(data_dict[(plot_list[0][a] + wave_type)],
data_dict[plot_list[0][a]], color = colours[a], linewidth = 1)
            ax.set_yticks([])
            ax.set_xlabel(xlabel, fontsize = 12)
            if auto_axis == 'both' or auto_axis == 'x':
                ax.autoscale('x', tight = True)
            if label_peaks == True:
                peak_labels(data_dict[(plot_list[0][a] + wave_type)],
data_dict[plot_list[0][a]], ax, technique)
                if technique != 'IR':
                    new_y_lims.append(1.1 * np.max(new_y_lims))
                if technique == 'IR':
                    new_y_lims.append(np.min(new_y_lims)/0.9-15)
```

```
        if technique != 'IR':
            if auto_axis == 'both' or auto_axis == 'y':
                ax.set_ylim([np.min(new_y_lims), np.max(new_y_lims)])
            for a in range(len(plot_list[0])):
                ax.annotate((plot_list[0][a] + ' (' + colour_names[a] +
')'), xy=(0.02, (0.93-a/12)), xycoords='axes fraction', color =
colours[a], bbox={'alpha': 0.1}, fontsize = 8)
        if technique == 'IR':
            ax.invert_xaxis()
            if auto_axis == 'both' or auto_axis == 'y':
                ax.set_ylim([np.min(new_y_lims), 100])
            for a in range(len(plot_list[0])):
                ax.annotate((plot_list[0][a] + ' (' + colour_names[a] +
')'), xy=(0.03, (0.05+a/10)), xycoords='axes fraction', color =
colours[a], bbox={'alpha': 0.1}, fontsize = 8)
        if auto_axis == 'x' or auto_axis == 'none':
            y_min = int(input('Input the lower y limit'))
            y_max = int(input('Input the upper y limit'))
            ax.set_ylim(y_min, y_max)
        if auto_axis == 'y' or auto_axis == 'none':
            x_min = int(input('Input the lower x limit'))
            x_max = int(input('Input the upper x limit'))
            ax.set_xlim(x_min, x_max)

    if len(plot_list) > 1:
        if auto_axis == 'x' or auto_axis == 'none':
            varying_y_lims = input("Do you want to set different y limits
for each graph? (If no, don't input anything)")
            if varying_y_lims == '':
                y_min = int(input('Input the lower y limit'))
                y_max = int(input('Input the upper y limit'))
        if auto_axis == 'y' or auto_axis == 'none':
            varying_x_lims = input("Do you want to set different x limits
for each graph? (If no, don't input anything)")
            if varying_x_lims == '':
                x_min = int(input('Input the lower x limit'))
                x_max = int(input('Input the upper x limit'))
        for x in range(len(plot_list)):
            new_y_lims = []
            for a in range(len(plot_list[x])):
                ylims = [np.min(data_dict[plot_list[x][a]]),
np.max(data_dict[plot_list[x][a]])]
                if ylims[0] > 0 and ylims[1] > 0:
                    new_y_lims.append(0.9 * ylims[0])
                    new_y_lims.append(1.1 * ylims[1])
                elif ylims[0] < 0 and ylims[1] > 0:
                    new_y_lims.append(1.1 * ylims[0])
                    new_y_lims.append(1.1 * ylims[1])
                elif ylims[0] < 0 and ylims[1] < 0:
                    new_y_lims.append(1.1 * ylims[0])
                    new_y_lims.append(0.9 * ylims[1])
                ax[x].plot(data_dict[plot_list[x][a] + wave_type],
data_dict[plot_list[x][a]], color = colours[a], linewidth = 1)
                ax[x].set_yticks([])
                ax[x].set_xlabel(xlabel, fontsize = 12)
                if auto_axis == 'both' or auto_axis == 'x':
                    ax[x].autoscale('x', tight = True)
                if label_peaks == True:
```

```
                    peak_labels(data_dict[(plot_list[x][a] + wave_type)],
data_dict[plot_list[x][a]], ax[x], technique)
                    if technique != 'IR':
                        new_y_lims.append(1.1 * np.max(new_y_lims))
                    if technique == 'IR':
                        new_y_lims.append(np.min(new_y_lims)/0.9-15)
            if technique != 'IR':
                if auto_axis == 'both' or auto_axis == 'y':
                    ax[x].set_ylim([np.min(new_y_lims),
np.max(new_y_lims)])
                for a in range(len(plot_list[x])):
                    ax[x].annotate((plot_list[x][a] + ' (' +
colour_names[a] + ')'), xy=(0.02, (0.93-a/12)), xycoords='axes fraction',
color = colours[a], bbox={'alpha': 0.1}, fontsize = 8)
            if technique == 'IR':
                ax[x].invert_xaxis()
                if auto_axis == 'both' or auto_axis == 'y':
                    ax[x].set_ylim([np.min(new_y_lims), 100])
                for a in range(len(plot_list[x])):
                    ax[x].annotate((plot_list[x][a] + ' (' +
colour_names[a] + ')'), xy=(0.03, (0.05+a/10)), xycoords='axes fraction',
color = colours[a], bbox={'alpha': 0.1}, fontsize = 8)
            if auto_axis == 'y' or auto_axis == 'none':
                if varying_y_lims == '':
                    ax[x].set_ylim(y_min, y_max)
                else:
                    y_min = int(input(f'Input the lower y limit for plot
{x}'))
                    y_max = int(input(f'Input the upper y limit for plot
{x}'))
                    ax[x].set_ylim(y_min, y_max)
            if auto_axis == 'y' or auto_axis == 'none':
                if varying_x_lims == '':
                    ax[x].set_xlim(x_min, x_max)
                else:
                    x_min = int(input(f'Input the lower x limit for plot
{x}'))
                    x_max = int(input(f'Input the upper x limit for plot
{x}'))
                    ax[x].set_xlim(x_min, x_max)

    if 'hor' in plot_orientation:
        plt.subplots_adjust(left = 0.02, right = 0.98, top = 0.98, bottom
= 0.125, wspace = 0.06, hspace = 0.15)
    if 'vert' in plot_orientation:
        plt.subplots_adjust(left = 0.05, right = 0.95, top = 0.98, bottom
= 0.07, wspace = 0.11, hspace = 0.2)
    if len(plot_list) == 1:
        plt.subplots_adjust(left = 0.05, right = 0.95, top = 0.98, bottom
= 0.15, wspace = 0.11, hspace = 0.2)
    if save_figure == True:
        figure_name = input('Figure name?')
        fig.savefig(os.path.join(outputdir, figure_name + '.png'), dpi =
600)
```