THE UNIVERSITY *of* EDINBURGH

Investigating the Biomedical Applications of Coordination Cages

Isis Amanda Middleton

Degree of Doctor of Philosophy

School of Chemistry

The University of Edinburgh

2023

For Mum and Dad

Abstract

Metallocages represent an exciting field of supramolecular chemistry concerned with the assembly of specific ligands and metals to form discrete structures. Coordination cages have applications in catalysis, molecular recognition and drug delivery. Accessing the biomedical applications of cages has received growing interest over the past two decades with, a more recent focus on utilising the cage cavity for the encapsulation of radioisotopes as a means of fast and efficient radiolabelling. This work follows on from previous work by the Lusby group, which demonstrated the synthesis of a kinetically locked and robust Co^{III} tetrahedral capable of hosting $[^{99m}Tc]TcO_4^-$ and altering the radioisotopes biodistribution.

The change in bioaccumulation from the stomach and thyroid to the liver indicated a potential interaction with proteins. Chapter two describes an investigation into the binding between a series of coordination complexes and Human Serum Albumin (HSA) using techniques including NMR, MS, dialysis and isothermal titration calorimetry. The study showed a strong interaction ($K_d \approx 2$ μM) between the Co^{III} tetrahedral from the previous study and HSA, compared to the Co^{III} mononuclear complex, which exhibited much weaker binding ($K_d \approx 200$ μM). Computational modelling indicated the interaction was likely the result of multiple electrostatic interactions, with potential applications in protein-mediated cage delivery.

Chapter three described the formation of three novel Co^{III} tetrahedra whereby the *in vivo* stability and host-guest chemistry could be changed by altering the external functionality of the cages. Using NMR and radiochemical TLC, the stability of the new systems under a range of conditions was determined, and the new Co^{III} ethanolamine functionalised cage exhibited a similar radiochemical encapsulation (EC_{95} = 4.4 μM) as the previously defined Co^{III} tetrahedral (EC_{95} = 1.6 μM). The new cage systems also serve as a scaffold for which further bioconjugation could occur through the binding of peptides to alter the delivery of the system, highlighting the potential of these kinetically inert Co^{III} tetrahedra as targeted delivery vessels.

Chapter four focused on determining the scope of cages applicable for biomedicine by synthesising a series of novel Pd_2L_4 systems. The cages were tested for their water solubility and stability in the presence of bio-prevalent species. NMR studies indicated that the underivatised Pd_2L_4 systems possessed a half-life of <10 minutes in the presence of NaCl, indicating a lack of biological stability. A series of more strongly coordinating ligands were synthesised, and attempts to assemble the cages resulted in low symmetry NMRs, presumably due to the increased strength of ligand coordination hindering the assembly equilibrium.

Overall a comprehensive investigation was completed into the activities of coordination cages *in vivo* and their potential applications. Whereby the external functionality of the cage is imperative for the bioactivity and stability of the system.

Lay Summary

Small molecular cages can be formed through the assembly of smaller complimentary parts. These molecular cages are hollow, and the internal cavity can be used to enclose compact species known as guests. This thesis investigates the applications of these cages in medicine and their potential to act as delivery vessels for imaging agents and therapeutics.

A drug delivery vessel works like a car picking up a passenger (guest) and delivering it to the required destination within the body. When considering the suitability of a potential drug, their interactions with species found within the body must be well understood. They must exhibit good stability under biological conditions, such as varying pH, slightly elevated temperatures and good water solubility. This thesis investigates the interactions between a series of small molecular cages and common biological species, notably proteins.

It was found that chemical alteration of the cages was required to ensure they were robust enough to survive within the body. It was also found that changing the chemical groups on the outside of the cage impacted the activity of the cage within the body. Strong interactions between the cages and proteins were determined to be responsible for where the cage travelled to within the body.

Table of Contents

Acknowledgements

Firstly, I would like to thank my supervisor Prof Paul Lusby for the opportunity to work on such an interesting project and your continued support and guidance over the last 3 and a half years. Especially for your ongoing assistance and presence throughout the many hurdles (including Covid-19) which I encountered during my studies, it is greatly appreciated.

I would also like to thank my secondary supervisor Prof Alison Hulme and Dr Christopher Coxon for the use of your laboratory equipment and the advice from yourselves and your students concerning all things biology. I would like the thank Prof Stephen Archibald and his research group, notably Dr Isaline Renard, for their ongoing collaborations in the world of radiochemistry.

Thank you to the entire Lusby group; past and present members. Thank you to Helen, Bec, Aaron, Will, Cora and Patrick for welcoming me into the group and keeping me involved throughout the various coronavirus lockdowns, even if that involved playing copious online games. Thank you to the current group members Laura, Ayesha, Ellie, Hang and Emer for always remembering everyone's birthdays with cake. Special thanks to Emer for joining the project and acting as a constant sounding board for all my ideas and analyses, some of which were better than others, I leave knowing this project is in the best hands. Thank you to all my project students for all your hard work, commitment and for teaching me how to be a better teacher.

Thank you to the staff over at the Edinburgh Protein Purification Facility, notably Dr Matthew Nowicki, for your ongoing patience teaching me all things proteins; from protein crystallography to isothermal titration calorimetry. I've learned a lot!

Thank you to all the technical staff in the School of Chemistry. Notably, Lorna Murray and Juraj Bella for never saying 'that's not possible' and for training me on almost every NMR technique under the sun. Thank you to Faye Cruicshank and Logan Mackay for braving the world of supramolecular mass spectrometry with me, it's been fun.

Thank you to all the friends I've made it Edinburgh, when I moved here 4 years ago I didn't know a single soul and miraculously I managed to build a whole new Scottish family. Special thanks to Becca, Rose, Phoebe, Reece and Blaine for keeping me going during some of the tougher days of lockdown, I wouldn't have finished this PhD without you. Thank you to my friends from the book club, which quickly became the wine club, for your ongoing willingness to help me practise for presentations and interviews, even when you don't understand what I'm talking about.

Lastly, I would like to thank my family, especially my Mum and Dad whose commitment to always answering the phone regardless of the time or reason for the call doesn't go unnoticed.

Declaration

I hereby declare that except where specific reference is made to other sources, the work contained in this thesis is the original work of the author. It has been composed by the candidate and has not been submitted, in whole or in part, for any other degree, diploma, or other qualification.

Isis Amanda Middleton

List of Abbreviations

AA	Ammonium Acetate
BArF	Tetrakis(3,5-bis(trifluoromethyl)phenyl]borate
d	doublet
DCE	Dichloroethane
DCM	Dichloromethane
DMF	N,N-dimethylformamide
DMSO	Dimethyl Sulfoxide
DOSY	Diffusion Ordered Spectroscopy
Eq	equivalents
ESI	Electrospray ionisation
Et_2O	Diethyl ether
Et_3N	Triethylamine
EtOH	Ethanol
h	hours
m	multiplet
MS	mass spectrometry
m/z	Mass to charge ratio
MeCN	Acetonitrile
MeOH	Methanol
min	minutes
nm	nanometre
NMR	Nuclear Magnetic Resonance
OTf	Trifluoromethanesulfonate
PET	Positron Emission Tomography
ppm	Parts per million
q	quartet
rt	room temperature
s	seconds
SCC	Supramolecular Coordination Complex
SPECT	Single photo emission computed tomography
t	triplet
THF	Tetrahydrofuran
UV-Vis	Ultraviolet–visible

Chapter 1 : Introduction to the Biomedical Applications of Coordination Cages

Chapter 1 : Introduction to the Biomedical Applications of Coordination Cages

1.1 Supramolecular Coordination Complexes

Supramolecular chemistry concerns the studies of complexes comprised of a discrete number of components held together by non-covalent forces.[1] These non-covalent forces such as hydrogen bonds, hydrophobic interactions and metal-ligand interactions, are critical to the structural integrity of fundamental complexes found in all walks of life, synthetic and natural. Over the last five decades, the understanding and applications of supramolecular chemistry have seen exponential growth from the discovery of crown-ether complexes to the synthesis of complex molecular machines.[2,3]

Supramolecular structures often comprise a larger host system that can complimentarily bind a guest through a range of non-covalent interactions. Perhaps one of the most prominent examples of this type of bonding relationship is that between a protein and its ligand or more specifically an enzyme and its substrate. From catalysis to drug design and biomedical imaging to molecular machines, understanding and making advancements in supramolecular chemistry is at the forefront of modern science.[1]

Supramolecular chemistry has shown promise with regard to biological applications for decades, including imitating enzyme catalysis with well-defined cages.[4] Enzymes are often comprised of hundreds of amino acids, which can be tedious to sequence and subsequently the protein can become complex to synthesise and emulate. However, if the active site mechanism can be identified with an understanding of the interactions between the enzyme and the substrate, it can be simulated by a smaller, simpler synthetic complex. Figure 1.1 shows an early example of this approach, exemplified by Breslow and colleagues whereby they synthesised mimics of ribonucleases capable of completing the acid-base catalysis.[5] Such complexes are much easier to synthesise than biologically occurring proteins and can offer a cheap, efficient alternative.

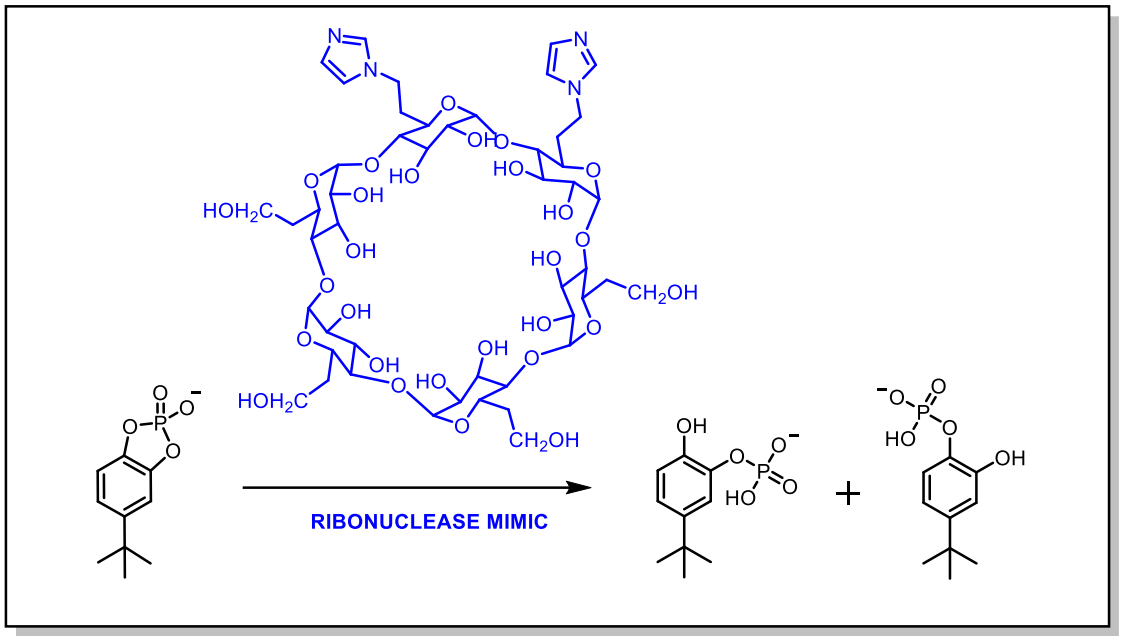


Figure 1.1. Acid-Base catalysis of catechol cyclic phosphate by β-cyclodextrin bis-imidazole acting as ribonuclease mimic, exemplified by Breslow and Schmuck.[5]

More recently, the links between supramolecular chemistry and biology are expanding beyond imitating enzymes to using coordination complexes in drug design, delivery and bioimaging.[6] Host-guest chemistry can be specifically designed for not only complimentary binding and catalysis but also for the encapsulation and transportation of species. Since Pedersen discovered crown-ethers in the 1960s the ability to selectively encapsulate a guest within a host has been fine-tuned, allowing controlled uptake and release.[2] This work describes the biomedical applications of coordination cages at present, with a focus on their ability to encapsulate and deliver guests of medical interest.

1.1.1 Self-Assembly Principles

Supramolecular coordination complexes (SCCs) are defined as discrete complexes comprised of more than one metal centre coordinated to multidentate ligands.[7] In the development of metallocages, the metals and ligands chosen can be used to design specific cage configurations. Self-assembled structures are advantageous as they are typically the thermodynamically favoured product. They adhere to the "maximum site occupancy" principle, which indicates that each possible metal binding site is occupied and the geometry is fulfilled to its maximum.[8] This implies that in a stable self-assembled complex, the principle of maximum site occupancy has been obeyed and the process is entropically driven. Therefore, by combining multidentate rigid ligands with metal centres which have intrinsic geometry they will self-assemble to finely tuned structures. The assembly process is reliant on the metals being able to reversibly coordinate ligands until the most stable geometry is achieved,

there is extensive research focused solely on measuring the equilibrium between the different self-assemblies.[9]

Metallosupramolecular self-assembling cages can pose some advantages over organic self-assembly systems due to the varying geometries of metals, allowing the specific and selective design of a cage tailored towards its role.[7] Transition metals can be favourable due to the different coordination geometries that are exhibited and whilst initial research highlighted octahedral transition metal complexes, extensive work over the last century has demonstrated the occurrence of a variety of geometries.[10] The electronic influence on geometry stems mainly from the principles of crystal field theory, which details the breaking in the degeneracy of orbitals upon their interactions with a static electric field; caused by neighbouring anions.[11]

When designing SCCs they can be either edge-directed or face-directed. As shown in Figure 1.2A, when a complex is edge-directed the ligands make up the edges of the polyhedra resulting in a porous complex.[12] Alternatively, Figure 1.2B exemplifies a face-directed polyhedra where the ligands assemble to form the faces of the complex preventing the creation of large pores.

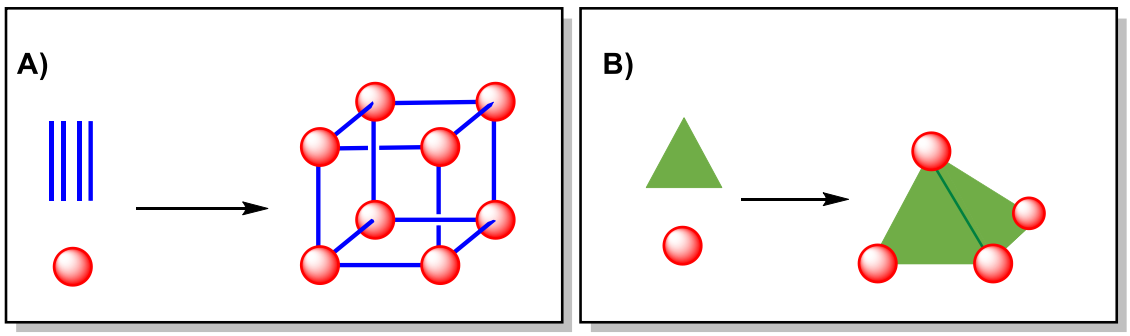


Figure 1.2. Illustrations of cage assembly A) Edge-directed assembly B) Face -directed assembly. The red spheres represent metal centres, the blue lines are edge directed ligands and the green triangles are face directed ligands.

The majority of the coordination cages being optimised for drug delivery are porous in nature where the host encapsulation is reliant on non-covalent interactions, and the host-guest complex exists in equilibrium with the free guest and host. For example, Cheng-Yong Su and co-workers published a study detailing the slow release of the encapsulated guest (5-FU) from their designed cage in vivo, whereby the equilibrium presents a lack of control over the release of the guest.[13] The porous nature of many of these cages, combined with the tendency for host-guest chemistry to be altered in biological mediums, leads to uncontrolled guest release. Figure 1.3 highlights a possible way of avoiding this equilibrium between the assembled host-guest complex and the free host and guest by utilising face-directed assembly to sterically prevent guest dissociation.

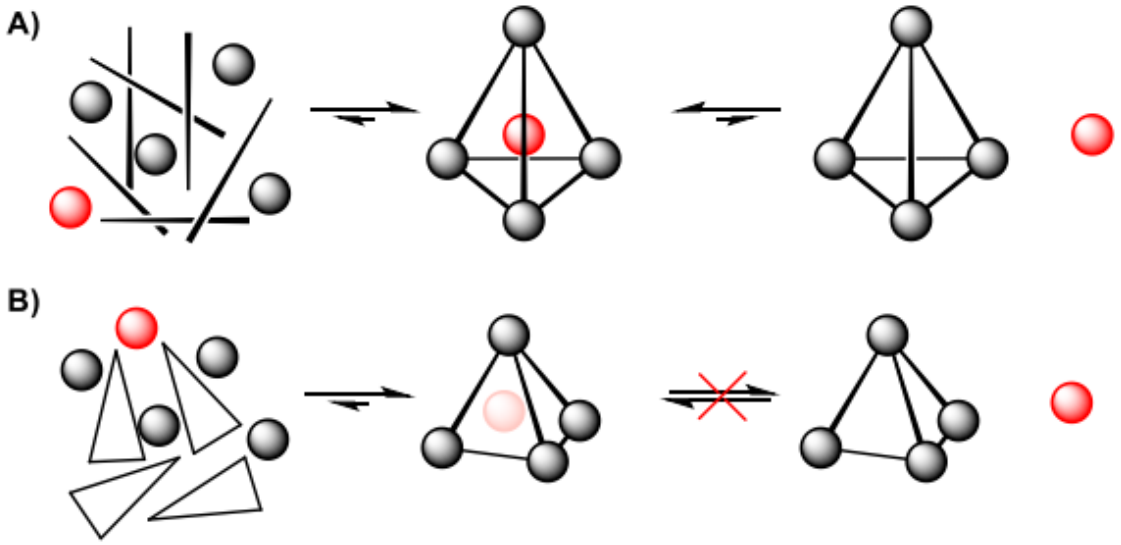


Figure 1.3. Equilibria of self-assembled coordination cages (black) and their guests (red). A) Edge-directed assembly of a kinetically locked porous cage with guest association. B) Face-directed assembly of a windowless cage with guest thermodynamically encapsulated.

There are two mechanisms for guest release, a simple in and out exchange as shown in Figure 1.3A or dissociation of the guest upon cage disassembly. By assembling a face-directed cage with no pores, the only mode of guest release is through disassembly of the cage, which can be mediated by creating robust, stable cages.

The Ward group produced a study in 2002 detailing a Co^{II} tetrahedron which template assembled around an anion guest, BF_4^-, with near complete encapsulation, seen in Figure 1.4.[14] The study included a series of NMR experiments at variable temperatures whereby there appeared to be no guest exchange in or out of the cage. Their cage utilised intertwining ligands that prevented pores large enough for guest disassociation while maintaining a kinetically stable host-guest complex.

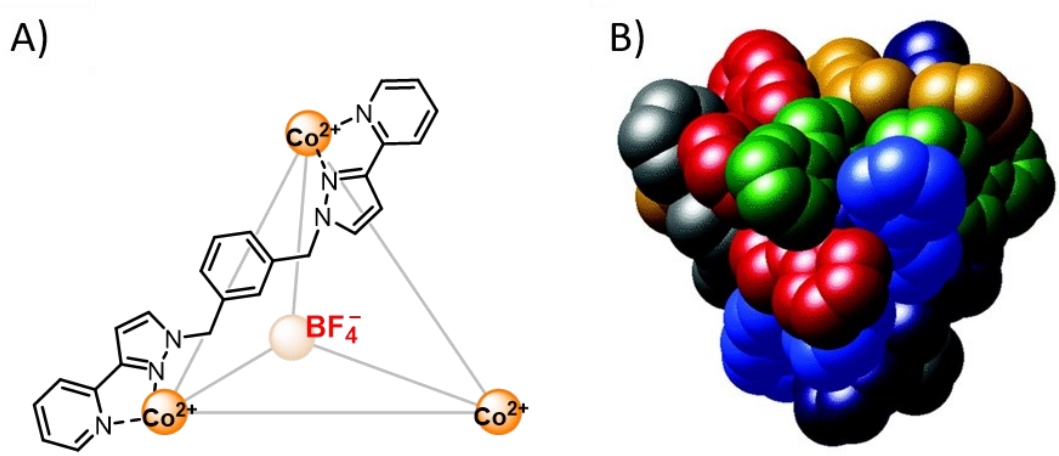


Figure 1.4. A Co^{II} tetrahedron with complete encapsulation of anionic guest A) Simplified structure representation of Co^{II} tetrahedron and B) X-ray crystallography structure of cage with encapsulated guest, adapted from reference. [14]

When considering the biomedical potential of coordination cages, fine-tuning their intrinsic structure is imperative; it determines which guests can be encapsulated and the strength of the interactions. As understanding of the physical chemistry behind the interactions between compounds advances, the ability to design and control hosts escalates. This is exemplified by the employment of computational modelling techniques such as Density Functional Theory to predict complexations.[15] Jelfs and Tarzia published a detailed review in 2016, discussing the capability of modern computational modelling techniques to predict the assembly geometry of metal-organic frameworks and forecast the properties of the resultant structures.[16]

1.2 Supramolecular Coordination Complexes (SCCs) as Biomedical Delivery Vessels

Coordination complexes have long been used as therapeutics, from the development of cisplatin in the late 20th century to ruthenium-based complexes and more recently advanced SCCs.[17] For example, in 2013 Vajpayee *et al.* designed a series of M_4L_4 arene-Ru self-assembled rectangular complexes capable of inhibiting cell cycles.[18] The study showed the complexes halted the cell cycle at the G0 phase and were effective in both in human lung and liver cancer cell lines which had exhibited cisplatin resistance. As the field expanded, therapeutic complexes were designed whereby their bioactivity was no longer solely reliant on metal toxicity. Coordination cages containing first and second-row transition metals can comprise a less toxic and cheaper alternative, whereby the resultant complexes can act solely as delivery vessels. Discrete cages could theoretically be ideal vessels for biologically relevant species given their ability to alter the properties of their encapsulated guest;

however, at the time of writing there are no coordination cages in clinical trials as targeted delivery vessels.

Yoshizawa and co-workers synthesised cytotoxic Pd^{II} and Pt^{II} M_2L_4 cages shown in Figure 1.5, capable of hosting small molecules, pyrene and caffeine.[19] Their study suggested that upon encapsulation of the guests, the complexes were more stable and were not reduced by glutathione to their active cytotoxic counterparts. While this was an example of a cytotoxic biomedical cage encapsulating a guest and reducing its bioactivity, similar work in this field attempted the opposite effect. With the aim of synthesising biomedical cages that act as delivery vessels to cytotoxic guests, which would increase the bioactivity of the complex as a whole. Advantages of utilising the cage as a delivery vessel include being able to add properties such as specific delivery and solubility to the guest-drug without having to modify the drug itself. This method was nicknamed 'the Trojan horse strategy' by Therrein *et al.* in 2008, whereby a highly cytotoxic drug could be carried directly to the required tumour cell and only released and active once within the cell.[20]

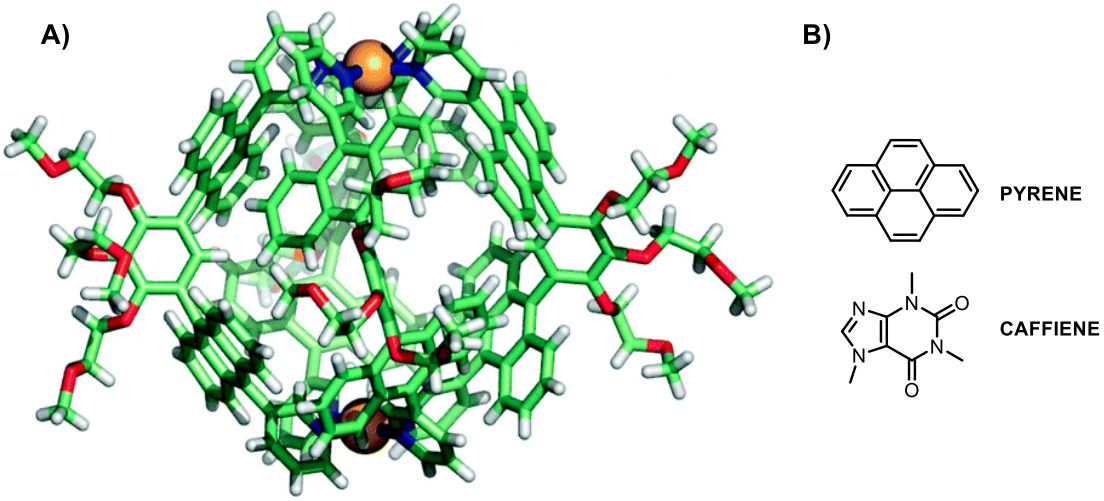


Figure 1.5. Structure of Pd^{II} M_2L_4 cage and its corresponding guests designed by Yoshizawa. A) X-Ray Crystallographic image of the cage, adapted from reference.[19] B) Guests of the cage, pyrene shown top and caffeine shown bottom.

The concept of utilising a coordination cage as a delivery vessel was shown in a later study by Therrien, which used fluorescent probes encapsulated by a hexaruthenium metallocage to monitor guest dispersion.[21] An increased level of fluorescence was observed within the cell when the fluorescent probe was administered within the cage, implying the capabilities of SCCs to enhance the cellular uptake of guests. The same group has since used the same ruthenium cage to encapsulate photodynamic therapeutic porphyrin and subsequently deliver it into cancerous cells, as shown in Figure 1.6.[22] They proposed that by administering the photosensitiser within the cage, undesired extracellular photosensitization is prevented as the therapeutic is only active once released from the

complex within the cell, evidencing the ability of coordination cages to alter the properties of encapsulated drugs.

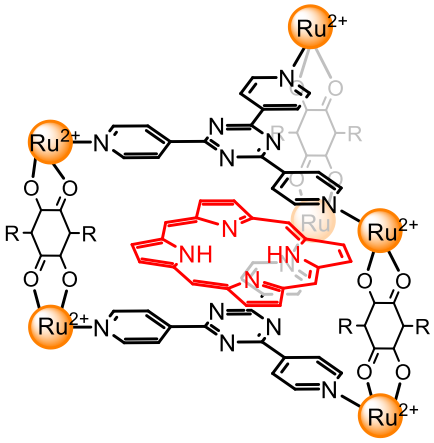


Figure 1.6. Structure of hexaruthenium metallocage hosting a photodynamic therapeutic; porphyrin. Adapted from work by Therrien et al. [22]

Similar drug limitations were overcome by Stang and colleagues in 2018, who designed an organoplatinum(II) metallacycle capable of encapsulating curcumin, a naturally occurring potential anti-cancer therapeutic.[23] Curcumin has incredibly low water solubility and the study showed an increased intracellular effect as a result of administration within the cage which provided water solubility to the hydrophobic guest. While many coordination cages are charge dense and exhibit good water solubility they can simultaneously possess a hydrophobic core, capable of encapsulating non-polar guests and altering their solubility. Supramolecular structures can also be utilised to increase the biological half-life of encapsulated species, as evidenced in a report by Wang *et al.,* which showed the formation of metal-organic frameworks (MOFs) capable of hosting anti-cancer drug 5-Fluorouracil.[24] Typically, 5-fluorouracil has a terminal half-life of around 8-20 min before being catabolised by the liver, however due to the complex structure of the host MOF designed by Wang and co-workers, they showed the drug was released over at least a 7.5 h period. Whilst this indicates a good improvement of the drug's *in vivo* half-life the study hypothesised that the release was due to complex instability.[24]

Due to the solubility issues associated with MOFs, Cheng-Yong Su *et al.* developed smaller zinc-based M_4L_4 tetrahedral cages capable of hosting 5-Fluorouracil (Figure 1.7).[13] They evidenced the controlled release of the substance over time in a biological human simulation, where the release of drug molecules lasted up until 108 h. They hypothesised that a series of strong hydrogen bond interactions between the tetrahedral cage and the fluorouracil, along with good size complementarity, stabilised the host-guest complex and increased the half-life of the guest. This work was further progressed by

the same group who showed the cage was also able to metabolise the anti-cancer drug Carmofur (HCFU) into another active form 5-FU.[25]

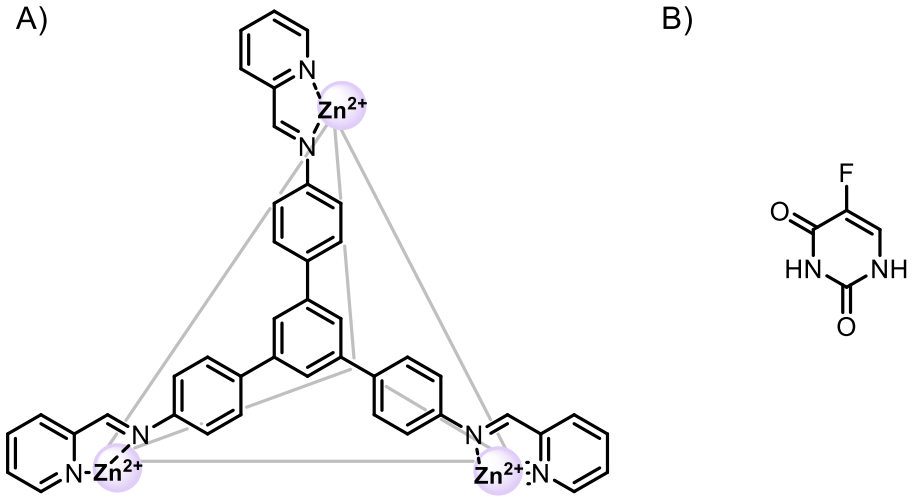


Figure 1.7. Zn_4L_4 tetrahedral cage capable of binding 5-Fluorouracil as evidenced by the Chen-Yong Su and co-workers. A) Zn_4L_4 tetrahedral cage and B) 5-Fluorouracil. [13]

Structural cage instability within biological media appears to be a reoccurring limitation when considering coordination cages a delivery vessels, however some works have attempted to change this limitation into an advantage. For example, Dastidar *et al.* created a Cu^{II}-metal organic polyhedral capable of hosting anti-cancer drug, doxorubicin with a pH-dependent release.[26] The complex self-assembles from tris-pyridyl ligands encapsulating doxorubicin within a single layered sphere. Given the low pH of cancerous cells, they investigated the cage's stability in a range of environments and showed that upon acidification, the cage disassembled, releasing the drug.

Wing-Wah Yam *et al.,* also completed a study highlighting the potential use of pH control with concerning mediating the encapsulation of anti-cancer therapeutics.[27] They completed a proof-of-principle study, successfully designing a series of alkynyl-platinum^{II} terpyridine molecular rectangles (Figure 1.8A) capable of hosting various metal-based anti-cancer therapeutics. Following a decrease in pH, the pyridyl groups of the Pd-based complex become protonated, altering the internal cavity and disfavouring guest encapsulation. Similarly, in 2015, Ward developed a Co^{II} polyhedral coordination cage (Figure 1.8B) which hosted a series of small therapeutics, including Parkinson's drug 'Amantadine', finding that upon reduction of the pH the binding association was drastically reduced.[28]

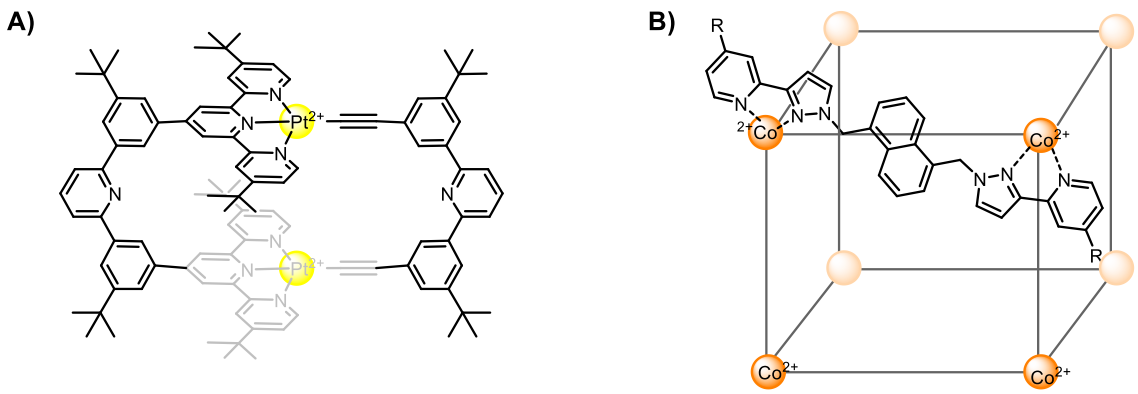


Figure 1.8. Structure of biomedical cages with pH control A) Pt^{II} rectangle complex by Wing-Wah Yam et al. [27] *and B) Co^{II} cubic complex by Ward* et al.[28]

Whilst much of this work attempts to utilise the host-guest instabilities at different acidities, it highlights a limitation of these complexes as reliable drug delivery vessels. The variety of pH levels and the array of competitive anions *in vivo* can cause cage disassembly and dissociation of encapsulated guests.

Similar structural limitations were observed by the Crowley group after they successfully designed a Pd_2L_4 cage capable of effectively hosting the anticancer drug cisplatin, shown in Figure 1.9.[29] The Pd_2L_4 cage disassembled in the presence of biologically abundant species including Cl^- and bioreductants, preventing Crowley and co-workers from further developing the cage for biomedical applications. However, more recent work completed by Casini *et al*. showed a similar Pd_2L_4 cage containing cisplatin decorated with a fluorescent label to image its cellular uptake.[30] They used proton NMR to evidence the cage encapsulation of two cisplatin molecules, parallel to the research completed by Crowley *et al.* They then showed the cage containing the encapsulated cisplatin had a substantially lower IC_{50} value than either the lone cisplatin or cage. However, although the cellular uptake of the fluorescently labelled ligand could be monitored, the structural integrity of the cage could not, leading to questions regarding the stability of the host-guest complex and host *in vivo*.

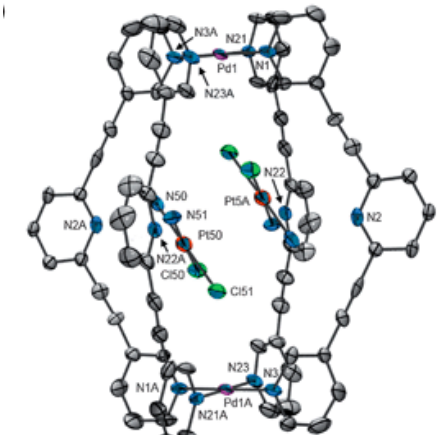


Figure 1.9. X-ray Crystallography structure of Pd_2L_4 cage capable of binding two molecules of cisplatin as determined by Crowley and co-workers. Figure adapted from reference.[29]

1.2.1 Targeted Delivery – Achieving Tumour Specificity

To advance the applications of coordination cages as drug delivery vessels, they need to be designed with a level of disease specificity. Within cancer research, many limitations with current therapeutics are related to a lack of tumour specificity and resultant toxicity to healthy cells.[31] A previous SPECT imaging study completed by the Lusby group carried out on naïve mice showed that the biodistribution of $[^{99m}Tc]TcO_4^-$ was altered from the thyroid and the stomach to the liver upon administration when encapsulated within a Co^{III} tetrahedral cage.[32] Metallocages as drug delivery vessels can be made tumour specific in a number of ways including; conjugating a directing group, hypoxic control and employing the enhanced permeability and retention (EPR) effect. Figure 1.10 visualises functionalising a directing group to the cage exterior to control where the complex travels within the organism and thus where the guest accumulates.

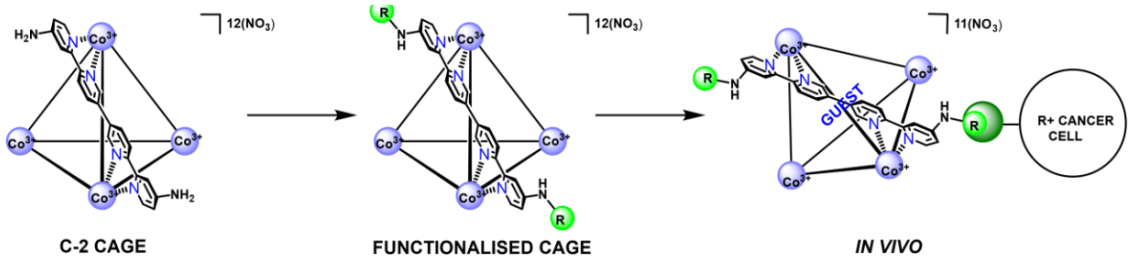


Figure 1.10. Proposed scheme for functionalisation of the Co^{III} tetrahedral cage designed by the Lusby group for targeted biomedical applications

Directing groups in cancer research refer to species such as peptides or monoclonal antibodies which are targeted towards either a proliferative pathway or species.[31] Cancerous cells result from excessive,

uncontrolled growth and division, and for this to occur, specific proteins and receptors are over-expressed, promoting proliferative activity.[31] Monoclonal antibodies are proteins produced from a single cell-line and capable of binding to tumorigenic proteins and inhibiting their activity. There has been substantial research into synthesising small synthetic peptides which bind tumour-specific proteins and can also work as directing groups.[33] Conjugating directing peptides and cytotoxins to a delivery body not only ensures a high payload but can also increases the half-life of the peptide within the organism.[34]

The method of achieving tumour specificity with metallocages by functionalising a bioactive group, capable of targeted delivery, to the exterior of the cage was probed by Casini *et al.* in 2017, where they reported conjugating a linear peptide to the ligand in a Pd_2L_4 cage via an amide bond.[35] Figure 1.11 shows the model linear peptide Ac-NLEFK-Am (the name is derived from the acetylated and amidated constituent amino acids) bioconjugated to their metallocage, highlighting the potential for the cage as a targeted delivery vessel. There were two reported strategies; one attempted to conjugate the peptide to the cage after it had been assembled and the second functionalised the ligands before assembling the cage. Their analysis of the cages formed was completed using ESI-MS coupled with HPLC and they evidenced that the latter approach was more successful. This followed on from some cage functionalisation work Casini had completed in 2016, where they conjugated fluorophores to the outside of the same cage in an attempt to fluorescently image the cage.[36] In both works, they successfully employed standard peptide coupling techniques to exofunctionalise the species onto the cage.

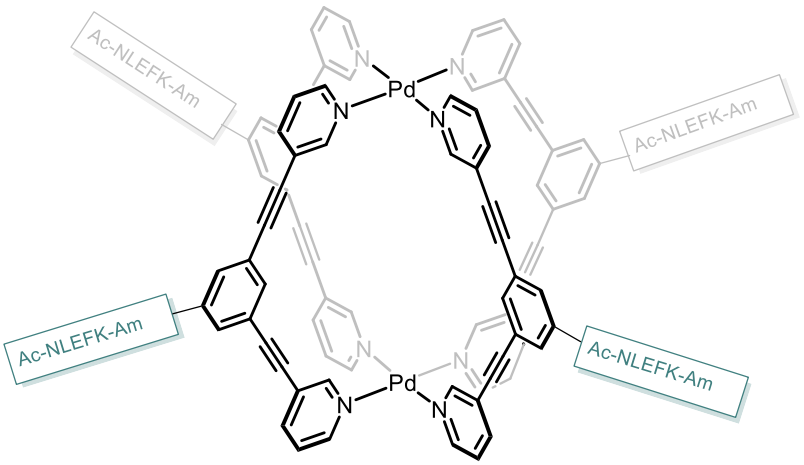


Figure 1.11. Structure of Pd_2L_4 cage with model peptide "Ac-NLEFK-Am" exo-functionalised, as developed by Casini and co-workers. [35]

Since these initial findings the Casini group have made significant progress. In 2018 they used the same amide bond formation previously developed to tether directing peptides to their cage.[37] By attaching peptides targeted to αvβ3 integrins over-expressed solely in melanoma A_{375} cells to the cage, they demonstrated a two-fold increase in the cytotoxicity of the encapsulated cisplatin. The same study also included *ex vivo* studies demonstrating the reduced toxicity towards healthy tissue for the encapsulated cisplatin in comparison to the free cisplatin. In 2021, the Casini group went on to show that by bioconjugating PepH3, a penetrative peptide fragment, to the exterior of their Pd_2L_4 cage it was able to transpose the highly selective blood-brain barrier whilst encapsulating SPECT imaging agent $[^{99m}Tc]TcO_4^-$.[38] Figure 1.12 shows a comparison between the basic Pd_2L_4 system which is unable to penetrate the blood-brain barrier versus the PepH3 exofunctionalised cage with encapsulated $[^{99m}Tc]TcO_4^-$ as exemplified by Casini and colleagues.

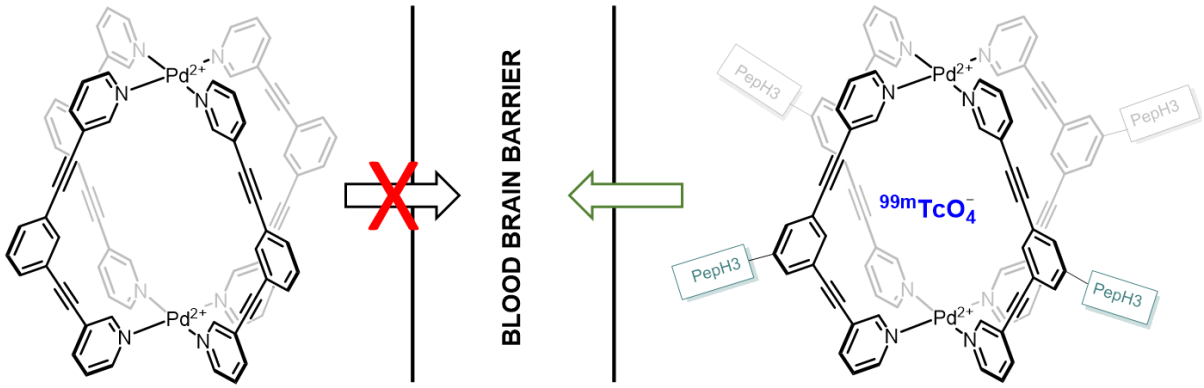


Figure 1.12. Pd_2L_4 cage exofunctionalised with PepH3 by Casini et al. *enabling it to penetrate blood-brain barrier with encapsulated $[^{99m}Tc]TcO_4^-$ in comparison to basic Pd_2L_4 cage. [38]*

However, these cages still face structural limitations regarding water solubility and stability within biological mediums. In 2019 Casini and co-workers published a study featuring BODIPY functionalised cages where they described instability in PBS after incubating for 4 h and precipitation out of water over a 24 h period.[39] This indicated that whilst they have optimised the functionalisation capabilities of their system there are still limitations regarding its delivery reliability *in vivo* due to instability.

In 2019, Stang and colleagues produced a multi-purpose metallocage capable of not only targeted delivery, achieved through the bioconjugation of a directing ligand, but also dual anti-cancer activity.[40] This was accomplished through the creation of a large sandwich complex, whereby platinum prodrugs constituted the cage edges, a large multidentate ligand comprised the cage faces and a photosensitizer was contained within, the metallocage was then assembled into nanoparticles using amphiphilic copolymers. As can be seen by the copolymer structure in Figure 1.13, the nanoparticles could be externally functionalised via the azide groups, allowing for the bioconjugation of targeting

ligand, cRGDfK, onto the surface of the structure using click-chemistry. Whilst the complex was shown to have good activity against cancer (in mice), which was overwise drug resistant, the paper quoted a half-life of 1.87 ± 0.24 h for the complex implying a level of *in vivo* instability.

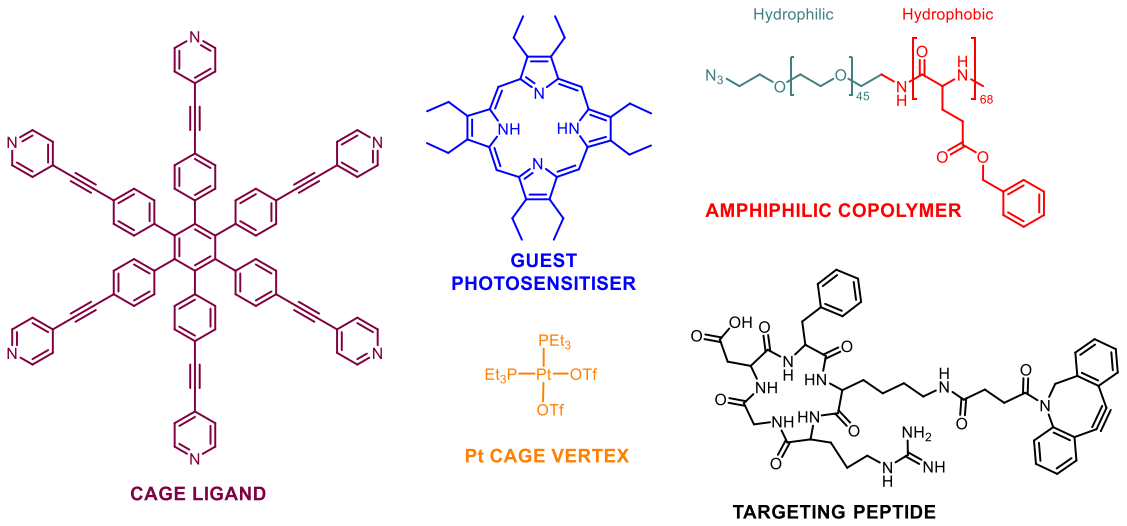


Figure 1.13. Structural components of metallocage containing nanoparticles capable of targeted photodynamic therapy as designed by the Stang group.[40]

The Stang group has published a series of papers utilising amphiphilic copolymers, which encapsulate metallocages into larger nanoparticles, increasing the biological half-life of the complexes as well as improving solubility and stability.[41,42] The large nanoparticles are not only specific towards cancerous cells via the bioconjugation of targeting peptides but the large structures also satisfy the molecular weight threshold for the EPR effect.

The EPR factor is a theory that accounts for the increased uptake and retention of larger molecules in tumours. It was first described by Maeda and Matsumura separately in the late 1980s, whereby Maeda detailed the excessive accumulation of Evans blue dye in tumour tissues compared to in the blood.[43] Studies have identified a threshold of about 40 kDa by which if the compound surpasses this molecular weight it is likely to be disproportionally accumulated in solid tumours compared to the rest of the organism.[43] This is considered to be a result of the rapidly growing tumour tissue having an altered vasculature, aimed at increasing the availability of oxygen and nutrients to the malnourished cells. Consequently, the cells receive an increased uptake of particles, including those of larger sizes, however their lymphatic system remains ill-equipped to deal with the removal of larger nanoparticles; thus leading to their accumulation.[44] This retention effect favours the uptake and accumulation of larger drug delivery vessels and is a form of passive targeting. However, most metallocages designed as drug delivery vessels are smaller than the proposed threshold required to achieve the EPR effect. Figure 1.14 illustrates the method employed by Crowley and co-workers in 2017, where they designed

a coordination cage which contained multiple discrete cavities capable of hosting multiple guests simultaneously, managing to increase the size without compromising the host-guest chemistry achieved by smaller cavities.[45]

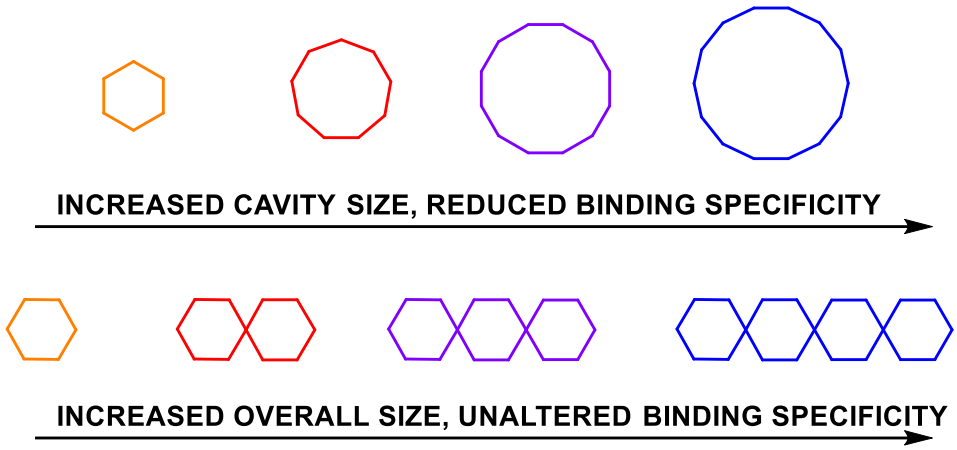


Figure 1.14. Illustration of the effect of cavity size on binding specificity and the utilisation of multiple cavities to increase complex size whilst maintaining guest specificity. [45]

Zheng *et al.* built on this work by assembling nanoparticles of ~ 80 nm from small metallocages containing Pt-drugs, where they showed the complex was capable of encapsulating therapeutics with similar IC_{50} to cisplatin, without compromising binding capabilities.[46] However, these examples represent a minority of the research as most cages being currently investigated have a molecular weight of approximately 2-3 kDa. Even with the employment of larger cages the sizes are still unlikely to surpass the 40 kDa threshold estimated necessary for the EPR effect.

Another way of achieving tumour specificity is by utilising the occurrence of hypoxia. Whilst the vasculature around the tumour will attempt to supply the rapid growth of cells with oxygen there is often still an oxygen deficit observed in tumour cells.[47] This leads to tumour hypoxia in which the lack of oxygen present causes the cells to exhibit a reductive microenvironment, a result of the overexpression of oxidoreductases. There has been substantial research into creating bioreductive prodrugs that are non-cytotoxic in physoxic (normal) tissue but are reduced in hypoxic tissue leading to fragmentation and release of the cytotoxic element. [47] For example, Figure 1.15 shows PR-104, a pre-prodrug which continued to Phase II clinical trials in 2010 and is rapidly converted from its phosphate ester form to its cognate alcohol *in vivo*.[48] The pro-drug, PR-104A, exploits tumour hypoxia by its radical reduction to DNA crosslinking metabolites in the absence of oxygen.

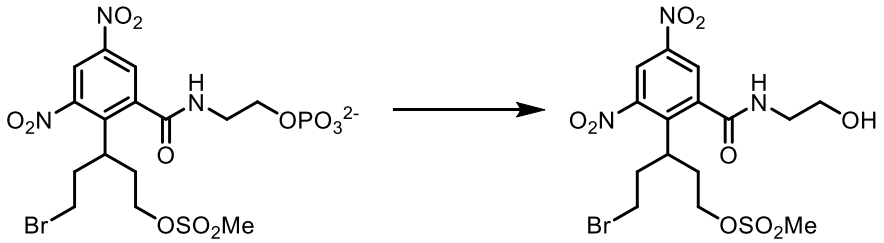


Figure 1.15. The conversion of pre-pro-drug PR-104 to its active form pro-drug PR-104A. [48]

There are also a range of transition metal complex prodrugs that were designed to utilise the reducing nature of hypoxic tumours to achieve specificity. The first examples of these to make it to clinical trials were utilising heavy metals; Pt^{IV} and Ru^{III}, where the prodrugs exhibited the metals in their oxidised states but upon reduction the compounds' *in vivo* activity is activated.[49]

A key example of where this technique has been effectively applied is seen in the evolution of cisplatin. Cisplatin contains Pt^{II}, which is reactive and toxic, studies have highlighted its substitution for Pt^{IV} enabling it to be safely delivered to the cell before being reduced to its more reactive counterpart.[50]

Theoretically, an intact cage containing a cytotoxic drug could have controlled disassembly only in the reducing microenvironment exhibited in hypoxic tumour tissues, releasing a cytotoxic guest at the desired destination. This was shown by Lippard in 2014, who described a cage capable of hosting two Pt (IV) prodrugs, which upon reduction within biological conditions are released in their active forms from the cage.[51]

1.2.2 Host-Guest Chemistry and Radiolabelling

Biomedical delivery cages can host a variety of small species including but not limited to cytotoxins. Much of the research discussed has concerned metallocages hosting small therapeutics such as cisplatin, as documented by Casini and Crowley, or photosensitizers, as described by Therrien and Stang.[29,30,22,18] The tendency of coordination cages to host anions has long been documented but has only recently been applied to biomedical imaging and diagnostics. [32,38,52] In 2018 the Lusby group synthesised a cage capable of hosting radiotracer anion; $[^{99m}Tc]TcO_4^-$.[32] Figure 1.16 shows the results of a series of SPECT imaging studies *in vivo* for both the lone radiotracer and the cage containing the radiotracer guest as carried out in naïve mice.[32]

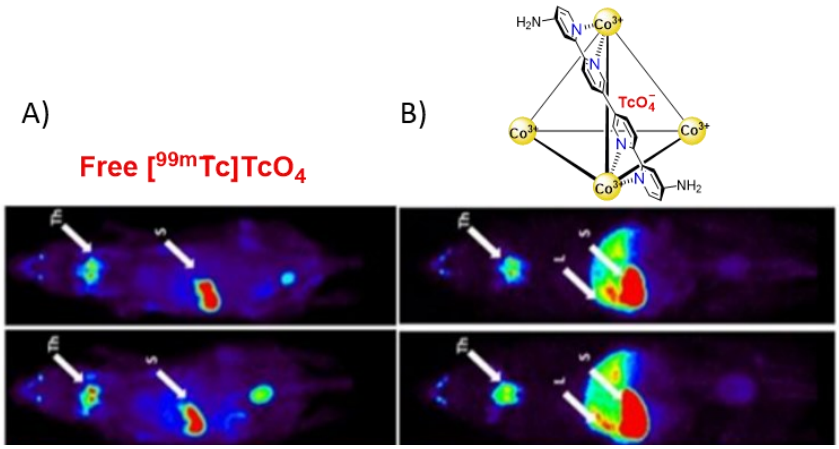


Figure 1.16. SPECT images of $[^{99m}Tc]TcO_4^-$ in Co^{III} tetrahedral adapted from Grantham et al. *2018 paper.. A) SPECT images from free pertechnate showing high thyroid and stomach uptake. B) SPECT images from $[^{99m}Tc]TcO_4^-$ in Co^{III} tetrahedral showing increased liver uptake*

SPECT is a form of bio-imaging that creates 3D images mapping the destination of radioactive tracers which are administered to the organism being monitored.[53] More specifically, the instrument has gamma detectors which can register the emission of radiation by the tracers creating an image of the organism and where the tracers are located; this technique is commonly used for cancer imaging. However, this imaging technique has limitations regarding relatively low quality images and slow scan time resulting in a singular end picture.[53] As a result of this limitation, in 2018, Lusby *et al.* were only able to identify where the cage had accumulated within the mice and not the path it was taking throughout the organism. Investigating alternative guests could potentially lead to the employment of different imaging techniques and increased biomedical applications.

Since the Lusby group published the 2018 study, evidencing the altered biodistribution of $[^{99m}Tc]TcO_4^-$ upon administration within their Co^{III} tetrahedral cage, there has been a couple more studies investigating the encapsulation of anionic guests with radio-isotopic analogues within coordination cages.[38] Notably, Mooibroek and Reek published a paper in 2022 detailing the encapsulation of ReO_4^- and $PtCl_4^-$, cold analogues of $[^{99m}Tc]TcO_4^-$ and $[^{195m}Pt]PtCl_4^{2-}$ respectively.[54] Figure 1.17 shows molecular models of the Pd_2L_4 system encapsulating the two guests as determined by Reek and co-workers. NMR titrations were subsequently employed to determine the affinity between the anionic guests and the Pd_2L_4 host, remarkably $PtCl_4^-$ had a $K_a = 10^4\ M^{-1}$.

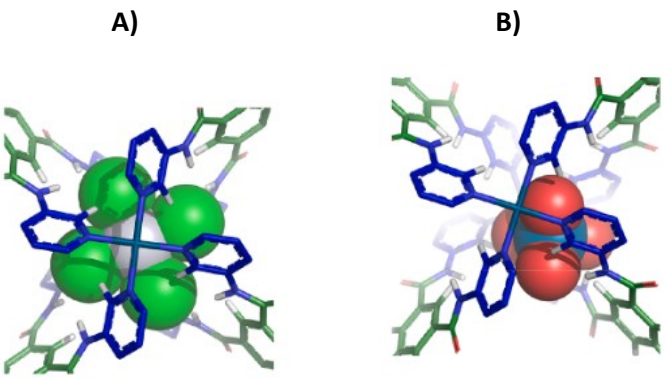


Figure 1.17. Molecular models of a Pd_2L_4 cage encapsulating A) $PtCl_4^-$ and B) ReO_4^- as developed by Mooibroek and Reek using DFT calculations.[54]

Whilst there are few examples of coordination complexes encapsulating radioisotopic anions, a 2022 study by Holland and d'Orchymont described the incorporation of ^{68}Ga and ^{89}Zr into a large supramolecular rotaxane.[55] Where they completed a series of PET imaging studies with the positron-emitting isotopes, demonstrating the ability of rotaxanes to radiolabel proteins. However, these systems are complex to synthesise and exhibit unknown stability *in vivo* limiting their applications.

1.3 Interactions of SCCs with biomolecules

The activities of many drugs are reliant on their interactions with biomacromolecules such as nucleotides and proteins. It is hypothesised that the altered bioaccumulation of $[^{99m}Tc]TcO_4^-$ when encapsulated with the Co^{III} tetrahedral, as shown by Lusby and co-workers, could be a result of protein interactions with the cage occurring in the liver.[32] Therefore, it is integral that there are robust methods in place to not only accurately categorize the type of interaction taking place but also to numerically quantify said interactions. Understanding molecular associations allows for the manipulation, understanding and prediction of complex molecular pathways that lie at the heart of health and disease. Therefore it remains imperative that all metallo-complexes with applications within biomedicine have either minimal interactions with biomolecules or those interactions are well characterised and utilised.

Perhaps one of the earliest and most prominent examples of an organometallic complex binding to a biomolecule being utilised as a therapeutic is the discovery of cisplatin. Following the synthesis of cisplatin in 1965 and its application as a potent chemotherapy agent, there has been growing interest in metal-based drugs and their potential.[56] However, clinically approved metal based drugs are still relatively rare and this is primarily attributed to metal toxicity and the low solubility of neutral metal complexes.

The limitations of metal based drugs are evenly balanced by the huge potential they possess. Metal coordination centres with intrinsic geometry can willingly bind to both the bases in nucleotides and the amino acids making up proteins: two major drug targets. When looking at metallodrugs their activity can come as either a result of the metal, ligand or both. The aforementioned coordination geometry of the metal centre acts as a three dimensional scaffold to build upon allowing the active targeting of a range of biomolecules.

1.3.1 The Interactions between Coordination Complexes and Nucleic Acids

There has been significant work investigating the binding between metallodrugs and DNA, beyond the preliminary findings of cisplatin. For example, in 2009, Yang and co-workers evidenced the binding of a Co^{III} based complex to a DNA-duplex containing mismatched base pairs.[57] Figure 1.18 shows the structure of the sterically bulky aryl-based cobalt complex designed to fit into DNA strands where the bases are mismatched, they also utilised the stereochemistry of the octahedral metal centre, showing that the lambda and delta forms binding to separate parts of the decanucleotide structure.

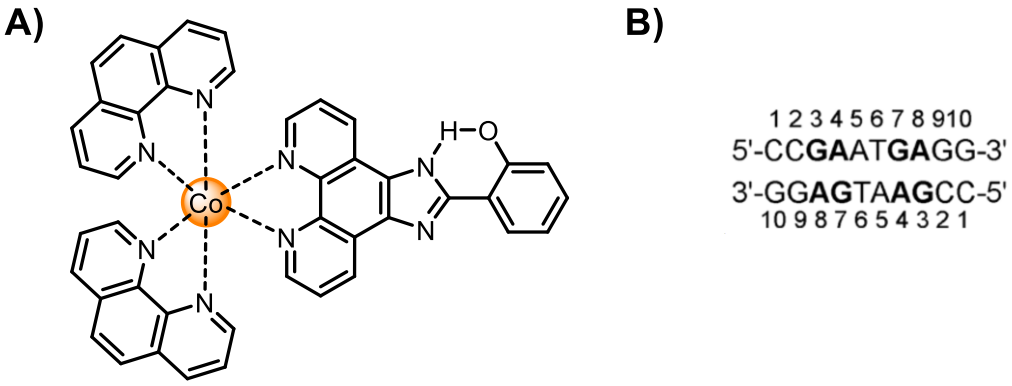


Figure 1.18. A) The structure of $[Co(phen)_2(HPIP)]^{3+}$ developed by Yang and co-workers and B) the DNA sequence used that it was evidenced to bind to.[57]

In 2013, Scott and his group synthesised an iron-based helical complex, shown in Figure 1.19, capable of triggering cell apoptosis with nanomolar potency. The cages were three times more potent against cancer (p53 negative) cell lines than the healthy controls indicating a good level of specificity and highlighting the diverse activity of biomedical coordination cages.[58] The Scott group have recently further developed these helical structures that can imitate cationic antimicrobial peptides. The helical complexes selectively bind bacterial DNA and key bacterial enzymes, displaying good antimicrobial action against strains that were otherwise exhibiting resistance.[59]

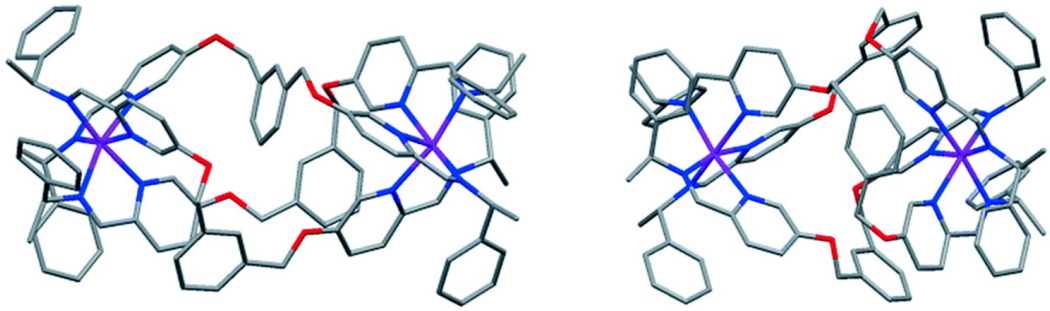


Figure 1.19. X-ray Crystallography structures of Fe^{II} based helical coordination complexes capable of triggering cell apoptosis as designed by Scott and co-workers. [58]

Hannon and co-workers have also published multiple papers investigating the binding of metal complexes to DNA, including seminal work evidencing and characterising a series of three-way DNA junction (3WJ) helicate binders. Figure 1.20 shows the X-ray crystal structure of a supramolecular helicate $[M_2L_3]^{4+}$ with trigonal-antiprismatic geometry capable of perfectly slotting into the cavity of a 3WJ.[60] Three-way junctions occur in both RNA as a part of splicing and translation as well as in DNA as a stage of DNA replication, making them a viable drug target.[60] By crystallising the supramolecular helicate into 3WJ, Hannon and colleagues were able to interpret the nature of the binding, which appeared to be a product of multiple interactions. The electrostatic interaction between the cationic metal centres and the anionic phosphate backbone of the DNA along with the extensive π-stacking between the phenyl system within the ligand the nucleotide bases present at the junction.

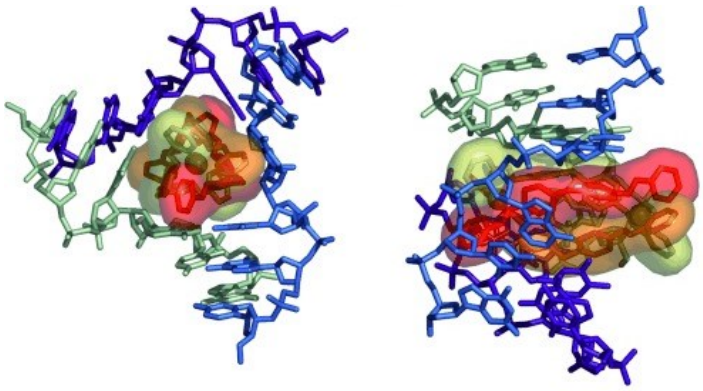


Figure 1.20. X-Ray Crystallography structure of Helicate Binding 3WJ as proposed by Hannon and co-workers. Figure adapted from reference.[60]

Over the last 15 years, Hannon has built on this initial work, including a study released in 2020 whereby they form a pseudo-rotaxane from a helicate. By wrapping a curcubit group around the helical structure and trapping it in place with additional branching structures at either end, they show the importance of structure, charge and symmetry in binding 3WJs.[61] The complexes designed appear to be highly versatile and capable of binding multiple DNA and RNA structures, most recently shown by their interactions with the RNA of SARS-CoV-2. The helicates were shown not only to bind to dynamic sections of the 5' UTR region but also to subsequently suppress the replication of SARS-CoV-2. [62]

Following these initial studies by Hannon, targeting DNA 3WJ's with supramolecular helicates has grown with popularity. Vazquez and colleagues recently combined traditional supramolecular design with peptide chemistry to synthesise a series of chiral helicates, where Figure 1.21 shows the structure. They evidenced DNA binding of the foldamers using fluorescence spectroscopy and when labelled with rhodamine could be used to label DNA replication sites in viable cells.[63]

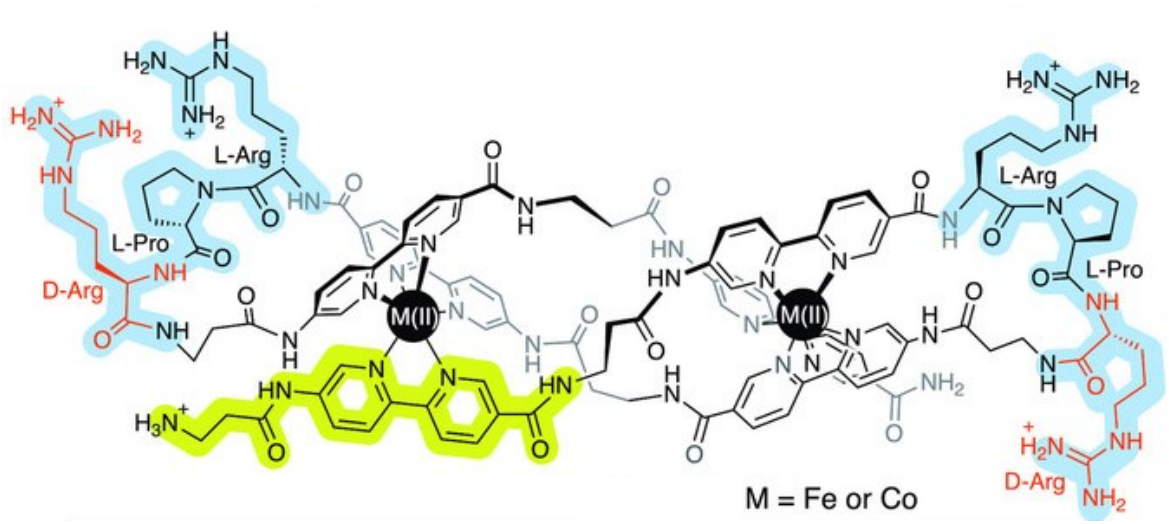


Figure 1.21. Structure of M^{II} based peptide helicate containing two loops consisting of the heterochiral l-Arg-l-Pro-d-Arg sequence. Figure adapted from study by Vazquez et al. [63]

The targeting of DNA using supramolecular complexes is not only limited to helicates binding 3WJ's as described, the anionic phosphate backbone of DNA along with the bases π-systems, generates potential binding sites for a variety of structures. Nitschke *et al.* have employed discrete tetrahedral $[M_4L_4]^{+8}$ systems, which are capable of binding DNA 3WJs and subsequently quenching fluorescence, as seen in Figure 1.22.[64] Given the tendency of their system to bind unpaired DNA bases, they also used the quenching of fluorescence as a detection method for unpaired DNA bases in double-stranded DNA. They later evidenced the ability of their tetrahedral structure to bind DNA G-quadruplexes forming an aggregate, preventing known binders such as S1 nuclease from binding and degrading the DNA.[65]

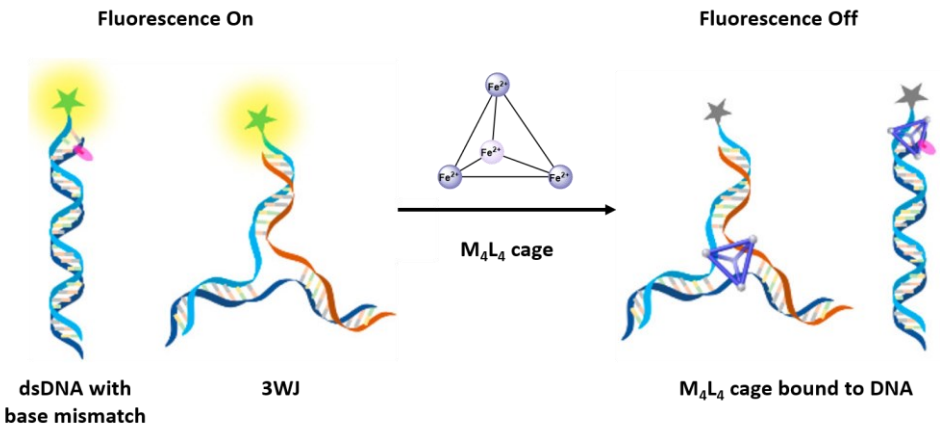


Figure 1.22. Schematic of M_4L_4 tetrahedral cage binding to dsDNA and 3WJ DNA. Figure adapted from work by Nitschke and co-workers. [64]

The versatility of these systems was further demonstrated by Huang and co-workers who used the Fe_4L_4 bound to the 3WJ as a probe for two distinctive miRNA's; a family of non-coding RNA structures thought to be integral in cell-cell communication.[66] Detection of specific miRNA's can be complex but has great potential in the detection of diseases; the usage of coordination cages in this area has applications in the early diagnosis and treatment of cancer.

1.3.2 Coordination Complexes and Their Interactions with Proteins

When designing therapeutics, administration must be considered and traditionally this is done either intravenously or orally, both leading to the drug circulating in the blood stream. So whilst complexes may have strong and specific binding to DNA, their activity within the blood stream, consisting of proteins and ions must be carefully considered. Similarly to polynucleotides, proteins are intrinsically charged, a result of anionic and cationic amino acid residues. These areas of electrostatic dipole along with the π-systems contained within amino acid residues provide proteins with similar ability to bind coordination complexes. This is furthered by the availability of free reactive amino acids, often positioned in a complimentary nature to form the active site of the protein.[67] This binding between proteins and metal complexes can be utilised in the form of molecular recognition, whereby there has been extensive research into utilising the intrinsic geometry of metal complexes to bind proteins. [68] Coordination complexes can be designed and their 3D structure fine-tuned to fit within the active site of an enzyme thus, they are increasingly being used as inhibitors. There are three modes of action for coordination complexes whereby it can originate from either the metal centre, the ligand, or the combined complex.

One of the first examples of using metal complexes as inhibitors came from Meggers and co-workers who utilised a central 'inert' ruthenium to act as a scaffold on which to attach biologically active

ligands which would complimentarily occupy the active site of protein kinases.[69] The ruthenium-based complex was designed to imitate the natural product, Staurosporine, which acts as a potent ATP-competitive protein kinase inhibitor, structures shown in Figure 1.23. This study exhibited the first example of a ruthenium complex serving as a protein kinase inhibitor, as well as exemplifying the capabilities of metal complexes to bind to proteins.

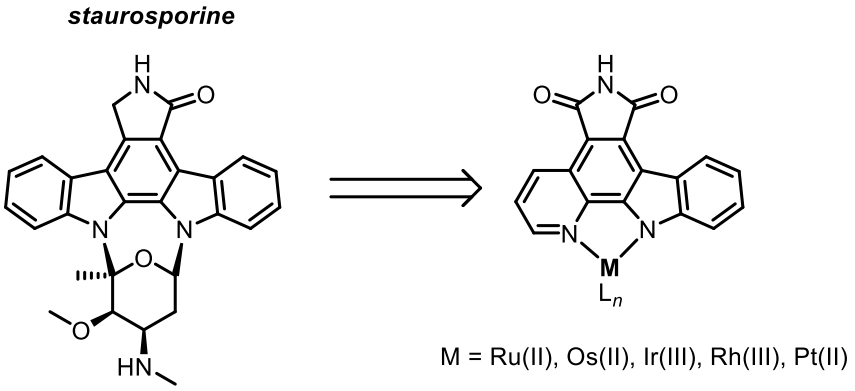


Figure 1.23. Structure of natural product staurosporine versus the Ruthenium-based complex designed by Meggers to exhibit similar activity.

Whilst some research focused on using the metal centre purely as a geometry to scaffold ligands around to compliment the active site, others have found the intrinsic activity of the metal binding to biomolecules more useful. Merlino and co-workers investigated the metalation of proteins by ruthenium complexes and evidenced when ruthenium complexes are administered intravenously when recovered most of the ruthenium is found complexed to plasma proteins.[70] Whilst the nature of the coordination is not fully known it appears that the ligands of the parent compound are substituted for the protein forming pseudo-octahedral complexes. There is strong evidence that the resulting anti-metastatic activity arises from the complex forming adducts with human serum albumin. Figure 1.24 exhibits a crystal structure of an anti-cancerous ruthenium compound forming an adduct with lysozyme protein isolated from a hen egg white published by Merlino and co-workers.[71] This later study utilised the solid state information to accurately characterise the adduct, illustrating the specific amino acid residues involved.

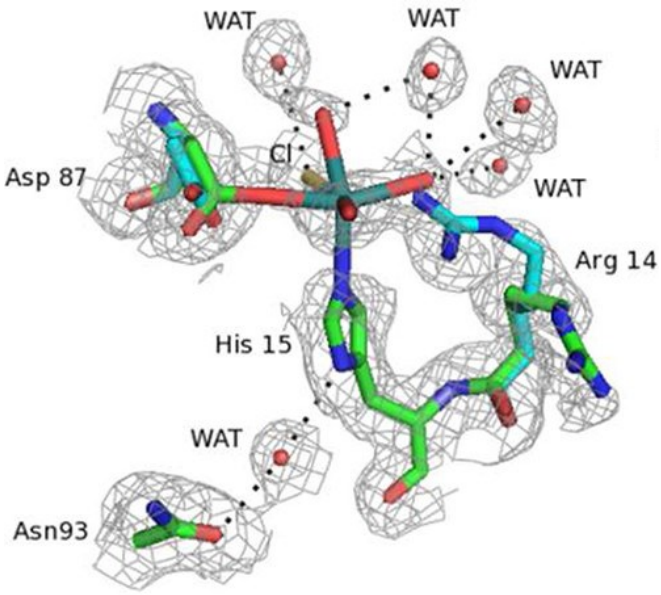


Figure 1.24. X-ray crystallography structure of Ru^{III} complex forming a protein adduct with a lysozyme protein and specific amino acid interactions, figure adapted with work by Merlino and colleagues.[71]

Given the growing interest in discrete coordination complexes as both therapeutics and delivery vessels it is essential to consider the activity of the metal centre, the ligand and the complex as a whole. Compared to the relatively extensive research on the interactions of metal complexes with proteins there is considerably less investigation into binding between coordination cages and proteins.

In 2014, Chi and colleagues evidenced the binding of a novel ruthenium-based complex with green fluorescent protein, which disrupts the tripeptide chromophore of the protein upon binding. Figure 1.25 shows the x-ray crystallography structure of the Ru^{II} based complex bound to the protein and the involvement of the arginine residue in binding.[72] The disruption has a quenching effect, which is tracked using UV-Vis absorption and emission experiments followed by gel electrophoresis, circular dichroism and atomic microscopy indicating that the alteration of the protein structure upon binding the complex subsequently caused aggregation. By monitoring the spectral changes at 512 nm they carried out binding titrations, determining a K_a of 7.4 x 10^8 M^{-1}. The strong interactions provides precedent that discrete complexes, which are able to diffuse through the cell membrane, could be used in targeting proteins.

Figure 1.25. Chi and co-workers demonstrated arene Ru-based supramolecular coordination complex for efficient binding and selective sensing of green fluorescent protein, figure adapted with reference.[72]

Figure 1.26 exemplifies work completed by Stang and colleagues a couple of years later, where a discrete organoplatinum metallocycle was mixed with tobacco mosaic virus (TMV) binds forming an aggregate.[73] TMV has complex physicochemical properties resulting from an RNA core coated in proteins and measuring 300 nm in length. Under neutral conditions the protein coating possesses a negative charge and thus when introduced to the cation metallocycle the structures electrostatically interact. As the complex and TMV species aggregate together they produce an aggregation-induced emission (AIE), which is spectroscopically monitored and used to quantify binding. They hope to pave the path for a new group of biohybrid species composed of metal-organic complexes and

anisotropically shaped bio-nanoparticles. Therefore they evidence the power of electrostatic attraction to build functional complexes with new properties.

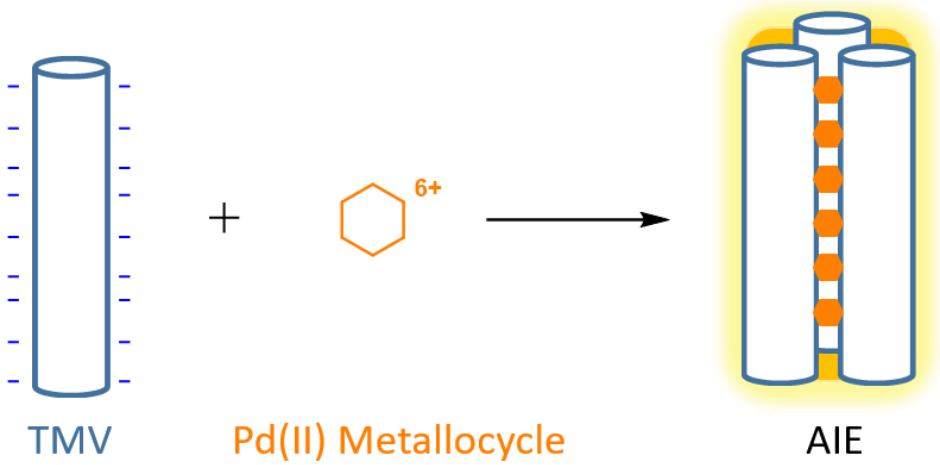


Figure 1.26. Organoplatinum metallocyle developed by Stang and co-workers capable of binding tobacco mosaic virus (TMV) and forming an aggregate displaying Aggregate Induced Emission (AIE).

Both Chi and Stang found evidence of protein aggregation, a phenomena plaguing the pharmaceutical industry, whereby upon structural changes the proteins stick together and can subsequently precipitate from solution. If a species preferentially binds the denatured version of the protein this can shift the equilibrium and further favour protein aggregation.[74] The ion pairing that occurs with electrostatic interactions can also lead to aggregations as shown by Fujita and co-workers who developed a series of sugar coated metallo-spheres capable of binding proteins.[75] The palladium-based $M_{12}L_{24}$ spheres were covalently coated with a series of saccharides, whereby the external functionalisation of the metallocages led to the binding of lectins and subsequently the formation of aggregates.

One of the only other examples of an investigation into the interactions between coordination complexes and proteins was published by Therrien and co-workers in 2015.[76] They identified three ruthenium-based metalloprisms, which they had previously found interacted with a series of isolated amino acids, and studied the interactions between the complexes and a series of plasma proteins. Figure 1.27 highlights the electrostatic nature of proteins, as exemplified by transferrin, whereby the small, charged metallocage can form complimentary interactions on the protein's surface. ^{1}H NMR titrations were used to monitor the protein in the presence of the metalloprisms, and whilst no significant shifts were observed in the spectra an overall decrease in protein and metallocycle concentration implied the formation of insoluble aggregates. This was further validated by a series of

experiments monitoring the structure of the proteins via circular dichroism prior to and following the addition of the coordination complexes, whereby significant structural changes were observed. Unfortunately, they had no success with either x-ray crystallography, in the generation of a co-crystal or with mass spectrometry. Similarly to the aforementioned studies, Therrien and co-workers attributed the interactions between the proteins and the coordination cages to be a result of electrostatic interactions.

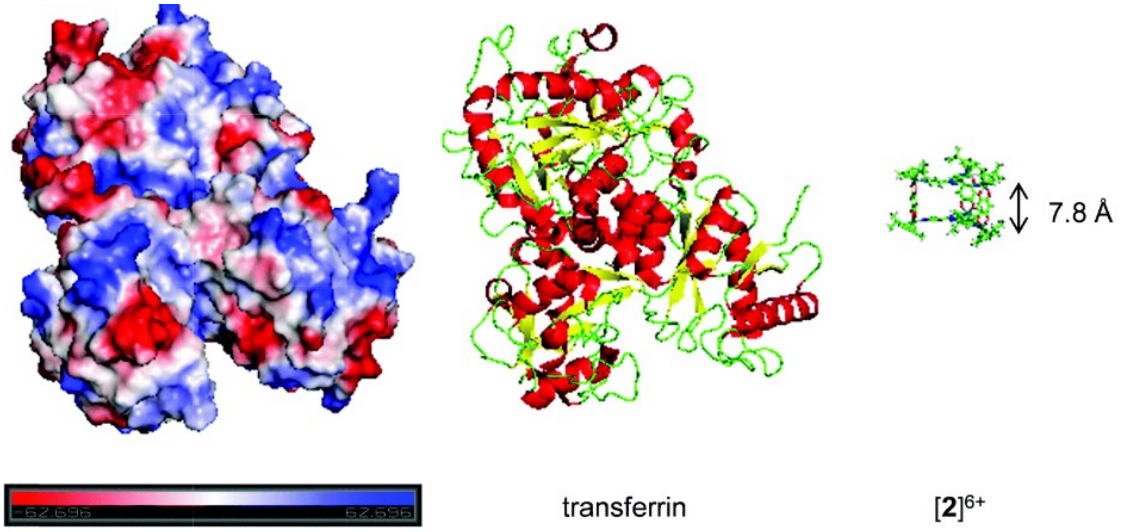


Figure 1.27. Comparison in size and charge density between blood plasma protein; transferrin and a Ru^{II} based metallocage developed by Therrien and co-workers. Figure adapted from work by the Therrien group.

Understanding the interactions between coordination cages and proteins is not only imperative to accessing the activity of the complexes *in vivo*, it also highlights the possibility of using proteins as delivery species.

Interactions of SCCs with Other Biomolecules

The blood plasma is predominantly composed of water, accounting for 91%, proteins (8%) and salts, sugars, fats and hormones account for the final 1%. There has been extensive research into the tolerance of coordination cages to metal salts with particular emphasis on NaCl, which exists at a blood plasma concentration of approximately 150 mM.[77] There are well defined mechanisms for the disassembly of coordination cages in the presence of salt given the propensity of Cl^- to bind to cationic metal centres. Figure 1.28 shows the mechanism proposed by Crowley and co-workers, whereby upon the addition of NaCl to their Pd_2L_4 cage the ligands were displaced by Cl^- and $PdCl_4^{2-}$ formed.[78] The tolerance of coordination cages to salt appears to directly correlate to the strength of the coordination bonds and stability of the cage, in which the 'kinetically locked' Co^{III} tetrahedral developed by the Lusby group remain unaffected in the presence of a high concentration of salt.[32]

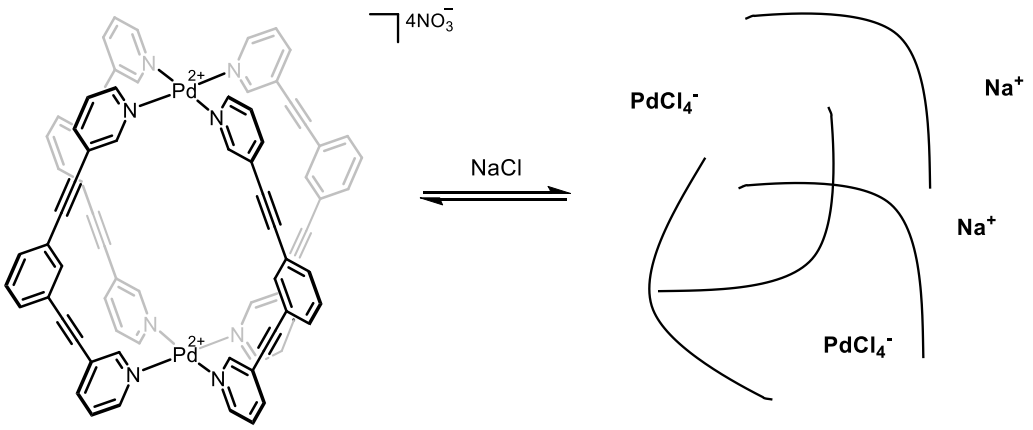


Figure 1.28. Pd_2L_4 cage disassembly in the presence of NaCl, evidenced by Crowley and colleagues.[78]

Regarding the potential interactions between coordination complexes and sugars, there has been limited research investigating the possible binding, with negligible focus on the blood plasma sugars specifically.

Yang and co-workers published a 2015 study, exhibiting the hierarchical self-assembly of platinum based metallocycle with polysaccharide, heparin. Figure 1.29A shows the structure of the cationic metallocycle, which was electrostatically attracted to the anionic charges on the glycosaminoglycan polymer shown in Figure 1.29B. It is important to monitor the concentration and function of heparin in the blood and thus they developed a coordination complex which upon binding heparin produces aggregation-induced emission which can be monitored and detected.[79] Although their work further demonstrated the ability to cationic coordination complexes to form strong electrostatic interactions, it failed to probe whether the complex could bind neutral saccharides, such as those found in blood plasma.

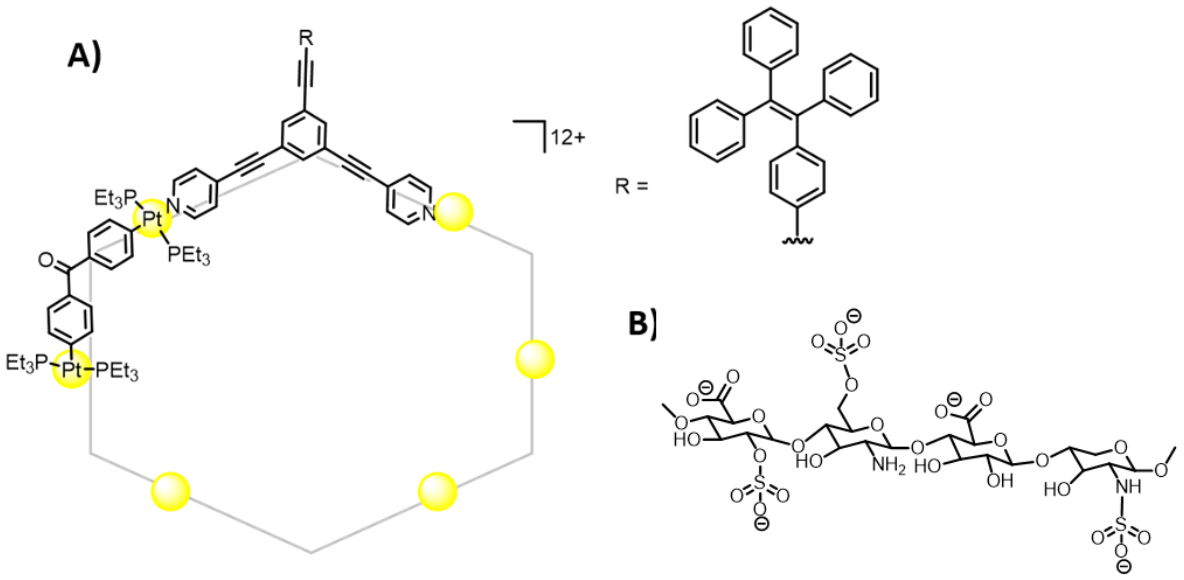


Figure 1.29. Pt^{II} based metallocycle capable of binding polysaccharide Heparin as developed by Yang and co-workers. A) Structure of the Pt^{II} metallocycle and B) Structure of Heparin

1.4 Coordination Cage Requirements for Biomedical Applications

When considering the potential biomedical applications of a coordination cage, water solubility is imperative. Low water solubility can lead to *in vivo* precipitation as well as decrease the effective concentration of the cage within the organism. Traditionally coordination complexes are assembled from rigid, aromatic ligands which coordinate cationic metal centres forming SCCs with hydrophobic cores. Solubility in polar mediums is attributed to electrostatic interactions between the charged cage and the solvent, as well as interactions between the counter anions and the solvent.

When targeting coordination cage dissolution in water, strongly hydrating counter anions such as NO_3^- are commonly utilised, whereby the interactions between the shell of NO_3^- anions and water help solubilise the complex. However, reliance on counter anions for coordination cage water dissolution can limit the host-guest capability, where guest-encapsulation can cause anion metathesis and upon decreasing the number of hydrophilic counter anions the solubility of the cage decreases.

Intrinsically, self-assembled cages exist in an effective equilibrium between their assembled and disassembled constituent parts. Under assembly conditions the equilibrium is likely biased to favour the formed cage and when transferring the coordination complex into biological conditions the equilibrium position can change. Movement into biological conditions can represent a change in pH, an abundance of bioreductants, proteins and salts alongside a sharp decrease in concentration.

Therefore it is imperative that coordination cages being considered for biomedical applications exhibit good stability under biological conditions, preventing premature disassembly or guest leakage.

1.5 Summary

Supramolecular coordination complexes are the focus of significant research, including investigations into their biomedical applications. Whilst the majority of preliminary research utilises the cytotoxic nature of metal complexes in the direct treatment of aggressive diseases such as cancer, more recent advances have sought to use coordination cages as delivery vessels which are biologically inert and capable of hosting biomedically relevant guests.[30]

The advancements in computational modelling have enabled the specific design and synthesis of cages which are capable of binding a variety of guests. Coordination cages have been synthesised capable of hosting: photodynamic therapeutics, small chemotherapy agents and bioimaging radioisotopes.[21,32,80] Research in this field focuses on the external functionalisation of coordination cages to target the delivery of the encapsulated guest and maximise the therapeutic and diagnostic potential.

However, the bioactivity of coordination cages *in vivo* and more specifically in the blood plasma is still somewhat unknown. Whilst there are many examples of supramolecular complexes binding DNA there are significantly fewer studies investigating the interactions between coordination cages and other plasma species, notably proteins.[58] Coordination complexes can also disassemble when transferred into harsher conditions and show instability *in vivo* in the presence of salts and bioreductants. Advances by Lusby *et al.* included synthesising a kinetically robust Co^{III} coordination cage capable of encapsulating imaging agent $[^{99m}Tc]TcO_4^-$ and following a series of SPECT experiments evidenced that the host-guest complex remained intact.[32] However, this remains one of the only examples of a coordination cage capable of retaining the encapsulated guest and remaining stable *in vivo*. This growing field represents an exciting application of coordination cages where the properties of the overall host-guest complex can alter the properties of the encapsulated guest.

1.6 References

1 I. V. Kolesnichenko and E. V. Anslyn, *Chem. Soc. Rev.*, 2017, **46**, 2385–2390.

2 C. J. Pedersen, *J. Am. Chem. Soc.*, 1967, **89**, 2495–2496.

3 J.-P. Sauvage, *Angew. Chemie Int. Ed.*, 2017, **56**, 11080–11093.

4 V. Mouarrawis, R. Plessius, J. I. van der Vlugt and J. N. H. Reek, *Front. Chem.*, 2018, **6**, 623.

5 R. Breslow and C. Schmuck, *J. Am. Chem. Soc.*, 1996, **118**, 6601–6605.

6 A. Casini, B. Woods and M. Wenzel, *Inorg. Chem.*, 2017, **56**, 14715–14729.

7 T. R. Cook, Y. R. Zheng and P. J. Stang, *Chem. Rev.*, 2013, **113**, 734–777.

8 J. F. Lemonnier, L. Guénée, G. Bernardinelli, J. F. Vigier, B. Bocquet and C. Piguet, *Inorg. Chem.*, 2010, **49**, 1252–1265.

9 T. J. Burchell and R. J. Puddephatt, *Inorg. Chem.*, 2005, **44**, 3718–3730.

10 K. Bowman-James, *Acc. Chem. Res.*, 2005, **38**, 671–678.

11 J. S. Griffith and L. E. Orgel, *Q. Rev. Chem. Soc.*, 1957, **11**, 381–393.

12 C. Zhu, M. Pan and C. Su, *Isr. J. Chem.*, 2019, **59**, 209–219.

13 W. Q. Xu, Y. Z. Fan, H. P. Wang, J. Teng, Y. H. Li, C. X. Chen, D. Fenske, J. J. Jiang and C. Y. Su, *Chem. - A Eur. J.*, 2017, **23**, 3542–3547.

14 R. L. Paul, Z. R. Bell, J. C. Jeffery, J. A. McCleverty and M. D. Ward, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 4883–4888.

15 C. J. Cramer and D. G. Truhlar, *Phys. Chem. Chem. Phys.*, 2009, **11**, 10757–10816.

16 A. Tarzia and K. E. Jelfs, *Chem. Commun.*, 2022, **58**, 3717–3730.

17 C. G. Hartinger, M. A. Jakupec, S. Zorbas-Seifried, M. Groessl, A. Egger, W. Berger, H. Zorbas, P. J. Dyson and B. K. Keppler, *Chem. Biodivers.*, 2008, **5**, 2140–2155.

18 V. Vajpayee, S. Lee, S. H. Kim, S. C. Kang, T. R. Cook, H. Kim, D. W. Kim, S. Verma, M. S. Lah, I. S. Kim, M. Wang, P. J. Stang and K. W. Chi, *Dalt. Trans.*, 2013, **42**, 466–475.

19 A. Ahmedova, R. Mihaylova, D. Momekova, P. Shestakova, S. Stoykova, J. Zaharieva, M. Yamashina, G. Momekov, M. Akita and M. Yoshizawa, *Dalt. Trans.*, 2016, **45**, 13214–13221.

20 B. Therrien, G. Süss-Fink, P. Govindaswamy, A. K. Renfrew and P. J. Dyson, *Angew. Chemie*, 2008, **120**, 3833–3836.

21 O. Zava, J. Mattsson, B. Therrien and P. J. Dyson, *Chem. - A Eur. J.*, 2010, **16**, 1428–1431.

22 A. Garci, J. P. Mbakidi, V. Chaleix, V. Sol, E. Orhan and B. Therrien, *Organometallics*, 2015, **34**, 4138–4146.

23 S. Datta, S. K. Misra, M. L. Saha, N. Lahiri, J. Louie, D. Pan and P. J. Stang, *Proc. Natl. Acad. Sci. U. S. A.*, 2018, **115**, 8087–8092.

24 H. N. Wang, X. Meng, G. S. Yang, X. L. Wang, K. Z. Shao, Z. M. Su and C. G. Wang, *Chem. Commun.*, 2011, **47**, 7128–7130.

25 W. Xu, Y. Qiu, S. He, S. Peng, B. Xie, M. Zhong, T. Jiang, X. Liu, W. Yin and J. Jiang, *Inorg. Chem. Commun.*, 2018, **87**, 24–26.

26 K. Sarkar, M. Paul and P. Dastidar, *Chem. Commun.*, 2016, **52**, 13124–13127.

27 A. K. W. Chan, W. H. Lam, Y. Tanaka, K. M. C. Wong and V. W. W. Yam, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 690–695.

28 W. Cullen, S. Turega, C. A. Hunter and M. D. Ward, *Chem. Sci.*, 2015, **6**, 625–631.

29 J. E. M. Lewis, E. L. Gavey, S. A. Cameron and J. D. Crowley, *Chem. Sci.*, 2012, **3**, 778–784.

30 A. Schmidt, V. Molano, M. Hollering, A. Pöthig, A. Casini and F. E. Kühn, *Chem. - A Eur. J.*, 2016, **22**, 2253–2256.

31 N. A. Seebacher, A. E. Stacy, G. M. Porter and A. M. Merlot, *J. Exp. Clin. Cancer Res.*, 2019, **38**, 1–39.

32 B. P. Burke, W. Grantham, M. J. Burke, G. S. Nichol, D. Roberts, I. Renard, R. Hargreaves, C. Cawthorne, S. J. Archibald and P. J. Lusby, *J. Am. Chem. Soc.*, 2018, **140**, 16877–16881.

33 S. Marqus, E. Pirogova and T. J. Piva, *J. Biomed. Sci.*, 2017, **24**, 1–15.

34 W. jin Jeong, J. Bu, L. J. Kubiatowicz, S. S. Chen, Y. S. Kim and S. Hong, *Nano Converg.*, 2018, **5**, 1–18.

35 J. Han, A. Schmidt, T. Zhang, H. Permentier, G. M. M. Groothuis, R. Bischoff, F. E. Kühn, P. Horvatovich and A. Casini, *Chem. Commun.*, 2017, **53**, 1405–1408.

36 A. Schmidt, M. Hollering, M. Drees, A. Casini and F. E. Kühn, *Dalt. Trans.*, 2016, **45**, 8556–8565.

37 J. Han, A. F. B. Räder, F. Reichart, B. Aikman, M. N. Wenzel, B. Woods, M. Weinmüller, B. S. Ludwig, S. Stürup, G. M. M. Groothuis, H. P. Permentier, R. Bischoff, H. Kessler, P. Horvatovich and A. Casini, *Bioconjug. Chem.*, 2018, **29**, 3856–3865.

38 B. Woods, R. D. M. Silva, C. Schmidt, D. Wragg, M. Cavaco, V. Neves, V. F. C. Ferreira, L. Gano, T. S. Morais, F. Mendes, J. D. G. Correia and A. Casini, *Bioconjug. Chem.*, 2021, **32**, 1399–1408.

39 B. Woods, D. Döllerer, B. Aikman, M. N. Wenzel, E. J. Sayers, F. E. Kühn, A. T. Jones and A. Casini, *J. Inorg. Biochem.*, 2019, **199**, 110781.

40 G. Yu, B. Zhu, L. Shao, J. Zhou, M. L. Saha, B. Shi, Z. Zhang, T. Hong, S. Li, X. Chen and P. J. Stang, *Proc. Natl. Acad. Sci. U. S. A.*, 2019, **116**, 6618–6623.

41 G. Yu, S. Yu, M. L. Saha, J. Zhou, T. R. Cook, B. C. Yung, J. Chen, Z. Mao, F. Zhang, Z. Zhou, Y. Liu, L. Shao, S. Wang, C. Gao, F. Huang, P. J. Stang and & X. Chen, *Nat. Commun.*, 2018, **9**, 1–18.

42 X. Lin, F. Chen, X. Yu, H. Wang, H. Qiu, Y. Li, S. Yin and P. J. Stang, *Proc. Natl. Acad. Sci. U. S. A.*, 2022, **119**, 1–8.

43 H. Maeda, J. Wu, T. Sawa, Y. Matsumura and K. Hori, *J. Control. Release*, 2000, **65**, 271–284.

44 R. Bazak, M. Houri, S. El Achy, W. Hussien and T. Refaat, *Mol. Clin. Oncol.*, 2014, **2**, 904–908.

45 D. Preston, J. E. M. Lewis and J. D. Crowley, *J. Am. Chem. Soc.*, 2017, **139**, 2379–2386.

46 Z. Yue, H. Wang, D. J. Bowers, M. Gao, M. Stilgenbauer, F. Nielsen, J. T. Shelley and Y. R. Zheng, *Dalt. Trans.*, 2018, **47**, 670–674.

47 W. R. Wilson and M. P. Hay, *Nat. Rev. Cancer*, 2011, **11**, 393–410.

48 C. P. Guise, M. R. Abbattista, R. S. Singleton, S. D. Holford, J. Connolly, G. U. Dachs, S. B. Fox, R. Pollock, J. Harvey, P. Guilford, F. Doñate, W. R. Wilson and A. V. Patterson, *Cancer Res.*, 2010, **70**, 1573–1584.

49 E. Reisner, V. B. Arion, B. K. Keppler and A. J. L. Pombeiro, *Inorganica Chim. Acta*, 2008, **361**, 1569–1583.

50 M. D. Hall, C. T. Dillon, M. Zhang, P. Beale, Z. Cai, B. Lai, A. P. J. Stampfl and T. W. Hambley, *J. Biol. Inorg. Chem.*, 2003, **8**, 726–732.

51 Y. R. Zheng, K. Suntharalingam, T. C. Johnstone and S. J. Lippard, *Chem. Sci.*, 2015, **6**, 1189–1193.

52 R. Custelcean, *Chem. Soc. Rev.*, 2014, **43**, 1813–1824.

53 M. M. Khalil, J. L. Tremoleda, T. B. Bayomy and W. Gsell, *Int. J. Mol. Imaging*, 2011, **2011**, 15.

54 X. Schaapkens, J. N. H. Reek and T. J. Mooibroek, *Inorg. Chem. Commun.*, 2022, **139**, 109284.

55 F. d'Orchymont and J. P. Holland, *Angew. Chem. Int. Ed.*, 2022, **61**, 1–11.

56 N. Cutillas, G. S. Yellol, C. De Haro, C. Vicente, V. Rodríguez and J. Ruiz, *Coord. Chem. Rev.*, 2013, **257**, 2784–2797.

57 H. Chen, C. Dou, Y. Wu, H. Li, X. Xi and P. Yang, *J. Inorg. Biochem.*, 2009, **103**, 827–832.

58 R. A. Kaner, S. J. Allison, A. D. Faulkner, R. M. Phillips, D. I. Roper, S. L. Shepherd, D. H. Simpson, N. R. Waterfield and P. Scott, *Chem. Sci.*, 2016, **7**, 951–958.

59 D. H. Simpson, A. Hapeshi, N. J. Rogers, V. Brabec, G. J. Clarkson, D. J. Fox, O. Hrabina, G. L. Kay, A. K. King, J. Malina, A. D. Millard, J. Moat, D. I. Roper, H. Song, N. R. Waterfield and P. Scott, *Chem. Sci.*, 2019, **10**, 9708–9720.

60 A. Oleksi, A. G. Blanco, R. Boer, I. Usón, J. Aymamí, A. Rodger, M. J. Hannon, M. Coll, P. M. J Hannon, A. Oleksi, A. G. Blanco, R. Boer, P. M. Coll, I. Usón, J. Aymamí and P. A. Rodger, *Angew. Chem. Int. Ed.*, 2006, **45**, 1227–1231.

61 C. A. J. Hooper, L. Cardo, J. S. Craig, L. Melidis, A. Garai, R. T. Egan, V. Sadovnikova, F. Burkert, L. Male, N. J. Hodges, D. F. Browning, R. Rosas, F. Liu, F. V. Rocha, M. A. Lima, S. Liu, D. Bardelang and M. J. Hannon, *J. Am. Chem. Soc.*, 2020, **142**, 20651–20660.

62 L. Melidis, H. J. Hill, N. J. Coltman, S. P. Davies, K. Winczura, T. Chauhan, J. S. Craig, A. Garai, C. A. J. Hooper, R. T. Egan, J. A. McKeating, N. J. Hodges, Z. Stamataki, P. Grzechnik and M. J. Hannon, *Angew. Chem. Int. Ed. Engl.*, 2021, **60**, 18144–18151.

63 J. Gómez-González, Y. Pérez, G. Sciortino, L. Roldan-Martín, J. Martínez-Costas, J. D. Maréchal, I. Alfonso, M. Vázquez López and M. E. Vázquez, *Angew. Chem. Int. Ed.*, 2021, **60**, 8859–8866.

64 J. Zhu, C. J. E. Haynes, M. Kieffer, J. L. Greenfield, R. D. Greenhalgh, J. R. Nitschke and U. F. Keyser, *J. Am. Chem. Soc.*, 2019, **141**, 11358–11362.

65 J. Zhu, Z. Yan, F. Bošković, C. J. E. Haynes, M. Kieffer, J. L. Greenfield, J. Wang, J. R. Nitschke and U. F. Keyser, *Chem. Sci.*, 2021, **12**, 14564–14569.

66 C.-Y. Qu, L.-S. Zhou, L.-H. Shu and Q. Huang, *ACS Omega*, 2021, **6**, 3330–3335.

67 K.-I. Okazaki, T. Sato and M. Takano, *J. Am. Chem. Soc.*, 2012, **134**, 8918–8925.

68 K. J. Kilpin and P. J. Dyson, *Chem. Sci*, 2013, **4**, 1410–1419.

69 L. Zhang, P. Carroll and E. Meggers, *Angew. Chem., Int. Ed*, 1999, **99**, 521–523.

70 A. Vergara, I. R. Krauss, D. Montesarchio, L. Paduano and A. Merlino, *Inorg. Chem*, 2013, **52**, 10714–10716.

71 A. Vergara, G. D'errico, D. Montesarchio, G. Mangiapia, L. Paduano and A. Merlino, *Inorg. Chem*, 2013, **52**, 4159.

72 A. Mishra, S. Ravikumar, Y. Ho Song, N. Saravanan Prabhu, H. Kim, S. Ho Hong, S. Cheon, J. Noh and K.-W. Chi, *Dalt. Trans.*, 2014, **43**, 6032.

73 Y. Tian, X. Yan, S. L. Manik, Z. Niu and P. J. Stang, *J. Am. Chem. Soc*, 2016, **138**, 12033–12036.

74 E. Y. Chi, S. Krishnan, T. W. Randolph and J. F. Carpenter, *Pharm. Res.*, 2003, **20**, 1325–1336.

75 N. Kamiya, M. Tominaga, S. Sato and M. Fujita, *J. Am. Chem. Soc.*, 2007, **129**, 3816–3817.

76 L. E. H. Paul, B. Therrien and J. Furrer, *Org. Biomol. Chem*, 2015, **13**, 953.

77 N. Nezafati, F. Moztarzadeh and S. Hesaraki, *Biotechnol. Bioprocess Eng.*, 2012, **17**, 746–754.

78 D. Preston, A. Fox-Charles, W. K. C. Lo and J. D. Crowley, *Chem. Commun.*, 2015, **51**, 9042–9045.

79 L.-J. Chen, Y.-Y. Ren, N.-W. Wu, B. Sun, J.-Q. Ma, L. Zhang, H. Tan, M. Liu, X. Li and H.-B. Yang, *J. Am. Chem. Soc.*, 2015, **137**, 11725–11735.

80 R. Cosialls, C. Simo, S. Borros, V. Gçomez-Vallejo, C. Schmidt, J. Llop, A. Belen Cuenca and A. Casini, *Chem. – A Eur. J.*, 2022, **29**, 1–7.

Chapter 2 : Investigating the Interactions between Coordination Complexes and Human Serum Albumin

Chapter 2 : Investigating the Interactions of Coordination Complexes with Human Serum Albumin

2.1 Introduction

Previous research by the Lusby group indicated that by encapsulating SPECT radioisotopic anion pertechnetate (TcO_4^-) within a Co^{III} tetrahedral cage (**C1**) the bioaccumulation could be switched from primarily the thyroid and the stomach to the liver (see Figure 2.1).[1] Given that the study proved the cage stayed intact and the pertechnetate remained encapsulated, it was concluded that the change in radioisotope distribution was a direct result of cage bioaccumulation. It was subsequently considered that the bioaccumulation of the coordination cage could result from interactions with biomolecules found in the liver, possibly caused by electrostatic interactions with the highly charged assembly. The liver exists in mice, similarly to humans, as the largest internal organ and is responsible for the production of hormones, bile, vitamins and proteins.[2]

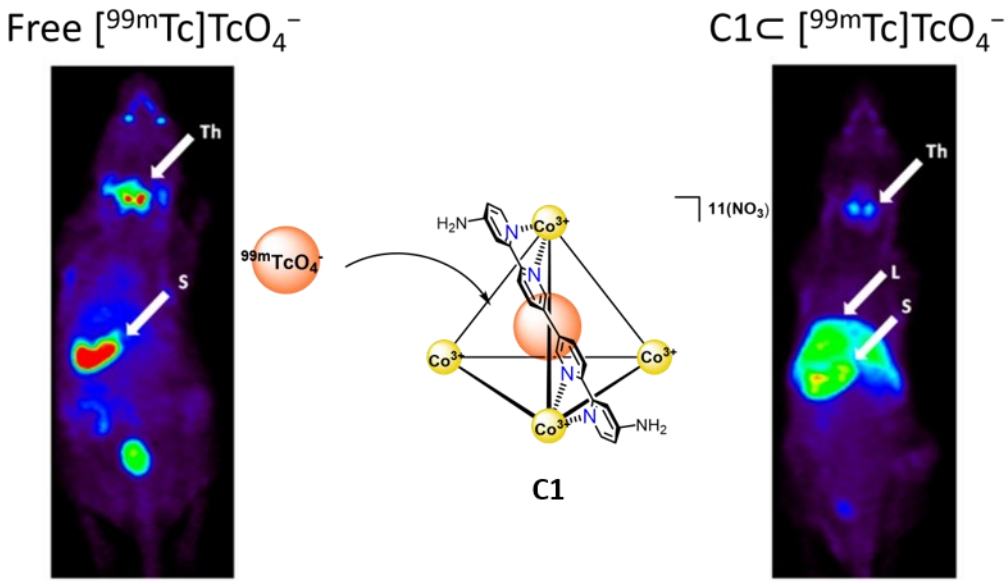


*Figure 2.1. Comparison of [99mTc]TcO4– uptake in naïve mice compared to [^{99m}Tc][TcO_4⊂**C1**]$^{11+}$ showing a decreased thyroid and stomach bioaccumulation and an increased liver accumulation upon encapsulation. Figure adapted from Lusby* et al.[1]

The most abundant protein produced in the liver is albumin, found at a blood plasma concentration of 40 mg/mL.[3] Albumin is a known binder of a variety of species, including common hydrophobic and hydrophilic drugs, providing it with applications in pharmaceutical delivery.[4] It was proposed that although the cage itself was not decorated with any specific targeting groups, it might still be capable of binding albumin and this was leading to bioaccumulation in the liver.

Studying the interactions of small molecules with proteins is a well-established field given the proficiency of proteins as modern drug targets. However, whilst using coordination cages for medical applications is emerging as a popular area, there is a deficit of research investigating the interactions of these species with bio-macromolecules. As discussed in Chapter 1 there are a few examples of coordination complexes binding DNA junctions but very little is known about the interactions between coordination complexes and proteins.[5,6] At the time of writing, only one example was found investigating the interactions of coordination cages with plasma proteins and how these interactions might affect the application of SCCs as biomedical delivery vessels.[7]

2.1.1 Aims of this project

Given the growing attention to using coordination cages as delivery vessels for radioisotopes and small drug molecules, understanding their interactions with prevalent bio-macromolecules is essential. This study outlines a series of versatile techniques determining the binding between coordination cages and human serum albumin (HSA), one of the most prevalent macromolecules in blood plasma.

The aims included employing a fragment-based drug design outlook whereby a range of coordination complexes are proposed based on the **C1** system used in the previous SPECT imaging study.[1] Allowing a structure-activity relationship insight into which aspects of **C1** are essential for interactions with the protein.

Herein, the suitability of traditional supramolecular binding techniques such as 1D NMR titrations, DOSY and MS are compared against more biological indirect methods, including Rapid-Equilibrium Dialysis and Gel Electrophoresis. The study aimed not only to investigate and characterise the specific case of the **C1** system described here, but also to provide a methodology for future research into the interactions between proteins and supramolecular coordination complexes more generally.

2.2 Results and Discussion

2.2.1 Synthesis of Coordination Complexes

Given the unknown nature of the potential interaction between **C1** and HSA, a library of complexes for the study were proposed. This includes **C1** which when encapsulating pertechnetate altered its bio-accumulation along with complexes; **C2**, **C3** and **C4** (Figure 2.2). **C2**, synthesised previously by the Lusby group, is included to probe whether the peripheral amine groups on the parent cage are interacting with the protein, via hydrogen bonding or electrostatic interactions, possibly upon protonation.[1] **C3** allows an investigation into the effects of both shape and charge density, existing in a similar size but with elongated shape and half the charge. Previous work has shown how symmetrical helical structures can perfectly compliment the triangular gap in a DNA 3WJ since proteins are inherently asymmetrical and it is hypothesised that this binding method will not be emulated with proteins.[6] **C3** is inherently more strained than its tetrahedral counterparts and thus the effect of intrinsic molecular strain on subsequent binding can be investigated.

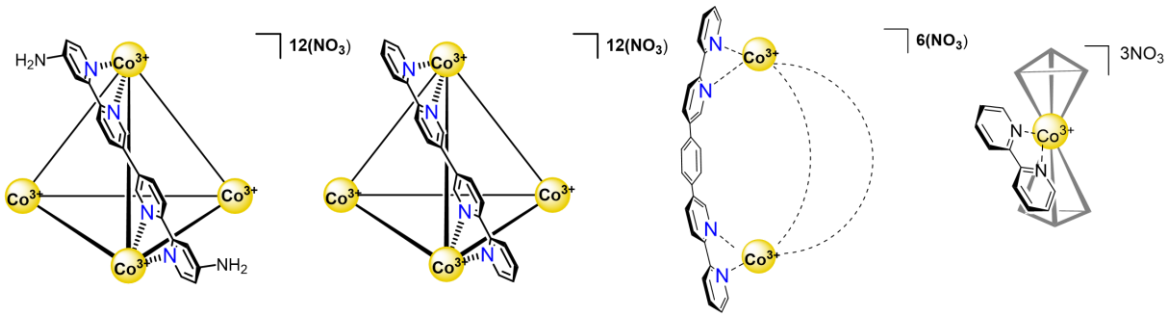


*Figure 2.2. Structure of the four coordination complexes investigated in this work. From left to right; **C1**, **C2**, **C3** and **C4**.*

C4 provides a fragment-based approach to the binding, whereby the complex mimics a single vertex of either the cage or the helicate, thus will provide insight into the importance of multiple charged sites. Thus, if similar binding was observed between HSA and the four complexes, it would imply that they are likely electrostatically 'sticking' to the protein via a single metal vertex. Conversely, stronger association of the tetrahedra compared to the dinuclear or mononuclear complexes would suggest that protein interaction involves polyvalent binding through multiple charged vertices.

C2, **C3** and **C4** complexes were synthesised using the previously described self-assembly followed by metal oxidation to kinetically lock the structures, as shown in Figure 2.3.[8] Whilst **C4** is structurally limited to form one complex product, the three other ligands can form multiple structures, as has been previously detailed by the Lusby group.[8] However, with short rigid ligands such as those in Figure

2.3, M_4L_6 tetrahedra are by far the most common outcome, with sometimes the helicate observed as a kinetic product.

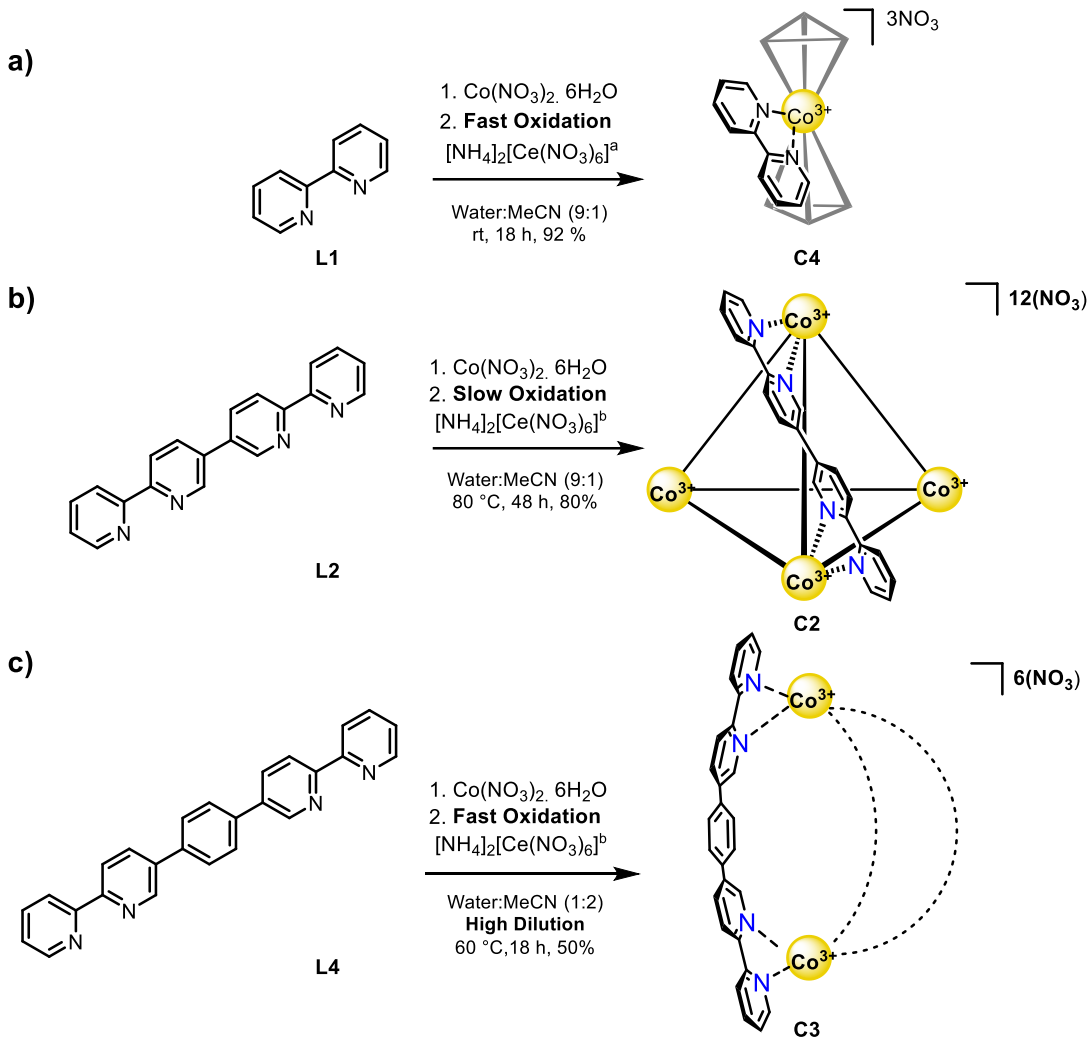


Figure 2.3. Schematics for the self-assembly reactions of (a) ***C4*** *(b)* ***C2*** *(c)* ***C3***

As **C4** is the only possible product with 2,2'-bipyridine (**L1**), it is unsurprising that this complex could be isolated in good yield by heating with $Co(NO_3)_2{\cdot}6H_2O$ for 1 h at 90 °C in a water : MeCN (9:1) mix followed by rapid addition of cerium ammonium nitrate. In contrast, assembly of quaterpyridine (**L2)** into **C2** utilised slow addition of the oxidant, which favours the energetically less-strained tetrahedral structure, as has previously been described by the Lusby group.[1]

Meanwhile, **C3** assembly from **L4** was slightly more challenging, whereby strong dilution and heat were required to bias the pre-oxidation equilibrium between the helicate and the tetrahedra. It was also observed that upon heating *in vacuo* to remove the solvents prior to precipitation of the complex,

some conversion from the helicate to the tetrahedral geometry occurred. It is considered that this is due to the intrinsic strain of the helicate and therefore applying heat kinetically favours the less strained. This was resolved by repeating the self-assembly at lower concentration and reducing the heat whilst removing the solvent *in vacuo*.[8]

The assembly of **C1** was further optimised from the previously reported procedure, improving the yield from 26% to 94%. This was achieved by utilising a 'trans-ligation' method as opposed to a direct assembly, which uses the more labile **C2** as a template onto which more strongly coordinated ligands are exchanged. This method has been utilised by other groups, including Crowley and co-workers who evidenced that by mixing sub-stoichiometric amounts of a more strongly binding donor ligand to a cage system, an asymmetric heteroleptic structure could be formed, as seen in Figure 2.4.[9] To ensure the complete conversion from the template cage to the product cage an excess of the stronger donor ligand was employed, with 2 Eq of **L3** per every equivalent of **L2** in **C2**.

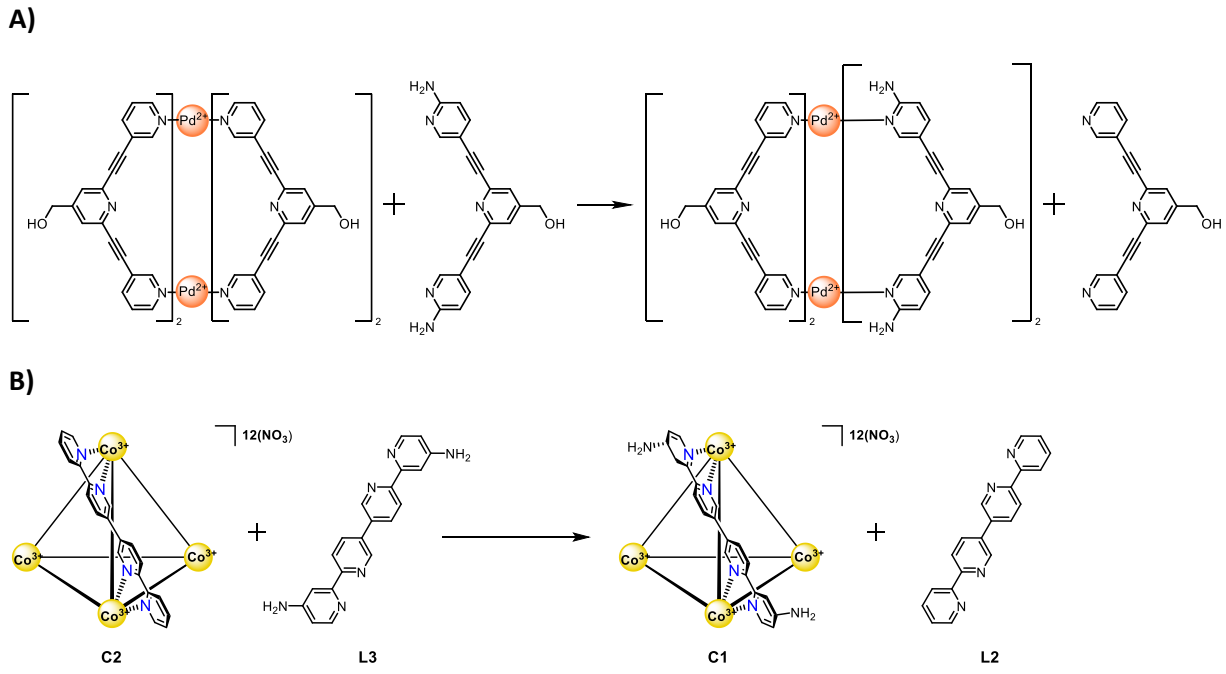


*Figure 2.4. Transligation of labile cages to form more stable cages (a) Formation of asymmetric heteroleptic cage by Crowley and co-workers[9] (b) Transligation of **C2** to form **C1** this work.*

Provisional attempts at the transligation involved simply adding an excess of **L3** to **C2** in 9:1 D_2O: CD_3CN. This experiment showed limited success with a mixture of species appearing in the NMR, which was attributed to the strength of the bipyridyl-Co^{III} bond in **C2** preventing ligand exchange preventing ligand exchange (Figure 2.5B). Trans-ligation was aided by the addition of a sub-stoichiometric amount (0.05 mol%) of Co^{II}, which was hypothesised to improve the rate of substitution through an electron-

transfer mechanism, wherein **C2** is partially reduced to give a more-dynamic system that is more amenable to ligand substitution. Further improvements were also made by optimising the solvent system, considering the low solubility of the **L3**. It was found that adding 1% DMF promoted significantly faster transligation, leading to a higher yield. The templating ligand was then recovered via filtration and selectively eluted providing indirect evidence of a successful transligation.

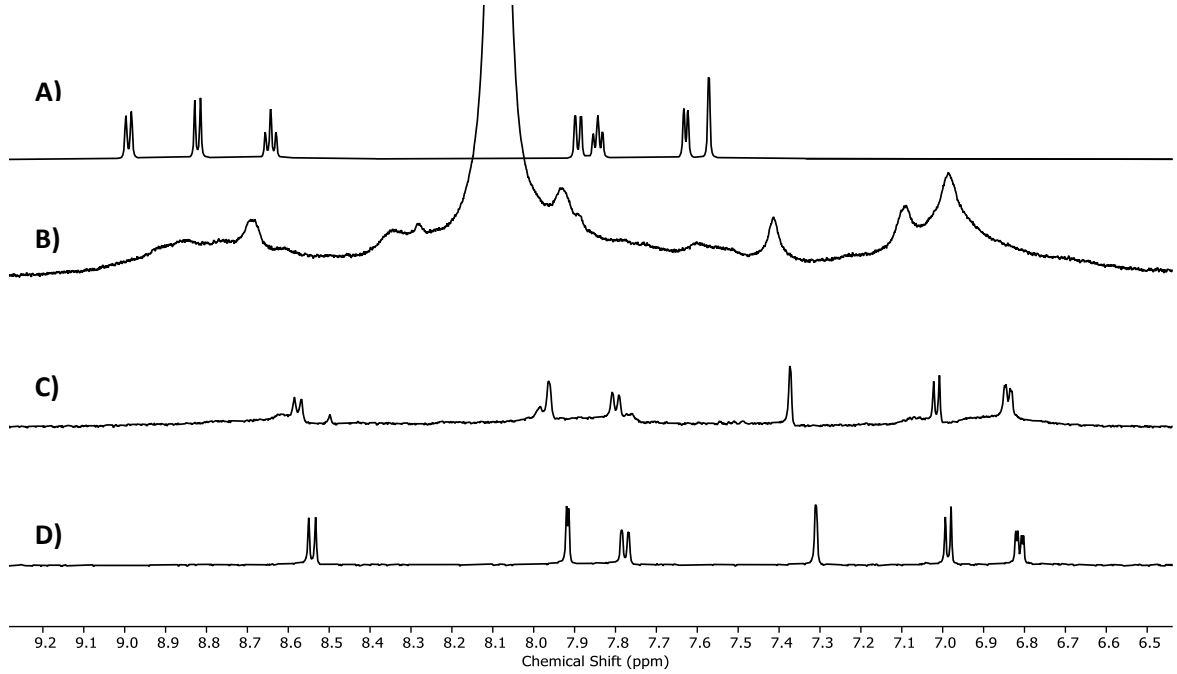


*Figure 2.5. Stack of ^{1}H NMR (500 MHz, D_2O) spectra showing the transligation optimisation for the formation of **C1** from **C2**. A) Template **C2** B) Reaction mixture containing an excess of **L3** with **C2** in 1% DMF in 3:1 D_2O and CD_3CN C) Transligation product with the addition of 0.05 Eq Co^{II} in CD_3CN and D_2O D) Pure **C1** transligated from **C2** with 0.05 Eq Co^{II} and 1% DMF in 3:1 D_2O and CD_3CN.*

2.2.2 Investigating Protein Binding

2.2.3 Spectroscopic Methods

Techniques for investigating the binding between the protein and the complexes were split into spectroscopic and indirect methods. The former involved using advanced analytical tools for identifying the cage-protein complex either through direct observation or spectroscopic changes, and will be discussed in the following section.

Mass Spectrometry

Mass spectrometry (MS) was the first technique used to investigate a possible interaction between **C1** and HSA since it is capable of direct detection of metal coordination complexes that have bound proteins via a range of interactions. This technique has been developed and optimised over the past decade or so, to allow not only the determination of protein mass but also the detection of ligands bound to folded proteins via non-covalent interactions. This is achieved through "soft" desolvation and ionisation of the protein-ligand complex, otherwise known as native MS.[10]

Robinson and co-workers have been pioneering this field, moving away from traditional MS conditions that can lead to the dissociation of weakly associated biomolecules and utilising more mild conditions that allow the visualisation of intact non-covalent complexes.[11] They defined a new era for the gentle top-down study of ligand interactions with proteins as 'nativeomics'; a combination of native MS with the 'omics', notably proteomics.[12] They describe not only identifying the protein with ligands bound in their native form but then follow this with multi-stage mass spectrometry (MS^n), a technique whereby the product ions formed in the first stage are subject to re-fragmentation providing information on the ligands bound.

In an attempt to study the binding between the HSA and **C1**, an excess of the cage was added to protein and incubated before ultrafiltration centrifugal concentration (MWCO 10 kDa) and buffer exchange into ammonium acetate solution (AA) at pH 7.4, serving also to remove excess salt. Previous research has shown that AA is an ideal solution for native MS given its volatility and stabilizing abilities.[13] Before attempting MS on the protein and **C1**, it was first probed whether ammonium acetate interacted strongly with the outside of the cage, as this could potentially perturb cage-protein binding. This was out-ruled by monitoring the addition of an excess of AA into **C1** (1 mM) in D_2O, time points were also run to ensure the complex was stable in solution over time (Figure 2.6). The slight deshielding of the internal proton is attributed to an excess of ammonium acetate leading to some weak encapsulation within the cage.

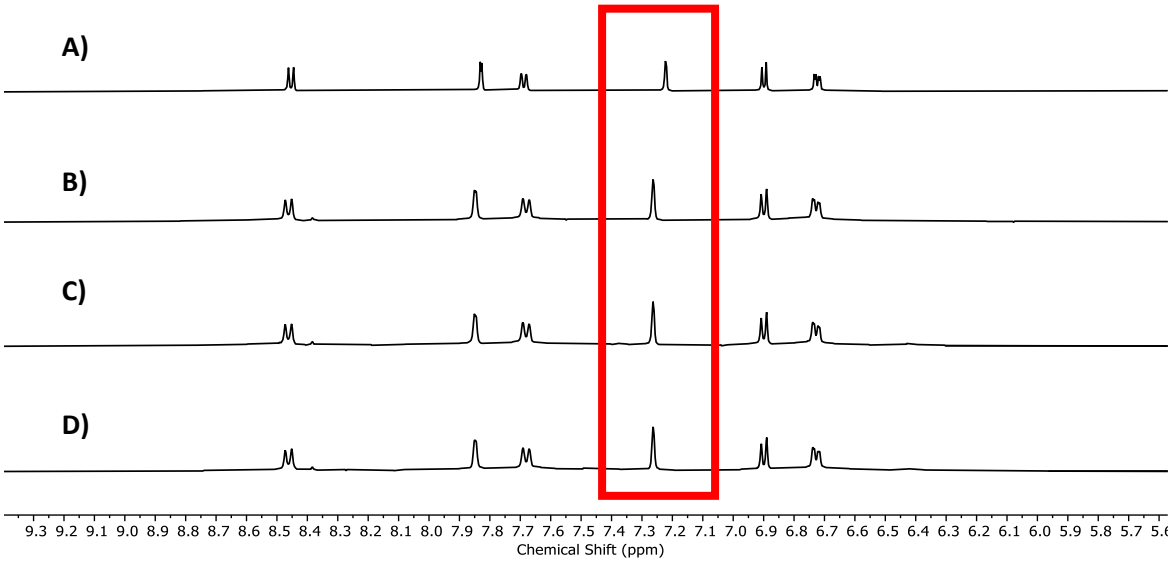


*Figure 2.6. ^{1}H NMR (500 MHz, D_2O) spectra evidencing **C1** stability in 100 mM AA. A) **C1** in D_2O B) **C1** in 100 mM AA in D_2O at t = 0, C) **C1** in 100 mM AA in D_2O at t = 6 h and D) **C1** in 100 mM AA in D_2O at t = 12 h. The internal proton of the cage is highlighted in red.*

The sample was then stored on ice before being analysed on a nanospray electrospray ionization (nESI) quadrupole/time-of-flight mass spectrometer. nESI and MALDI make up the most prominently utilised ion sources for protein visualisation, allowing a mild ionisation without disturbing the protein-ligand complex.[13] In this study, an Advion Nanomate was used, a chip-based electrospray ionisation instrument capable of generating a nano-electrospray from minimal material.

It has previously been found in the Lusby group that specific M_4L_6 cage systems can be challenging to analyse via MS. The NO_3^- counter-anions of **C1** make it water soluble, leading to a solution with high surface tension and vaporisation enthalpy that can hinder the performance in ESI-MS. In order to obtain a steady ion current of the aqueous **C1** solution on the spectrometer, a low gas pressure (0.6 psi) and a higher voltage (1.5 V) are required. The spectrometer parameters were then optimised by changing the voltages at the sampling and extraction cones, whereby too low voltages led to excessive spectral noise whereas too high a voltage caused cage fragmentation and only the ligand could be seen. The effects of sampling cone voltage on the MS of **C1** are shown in Figure 2.7.

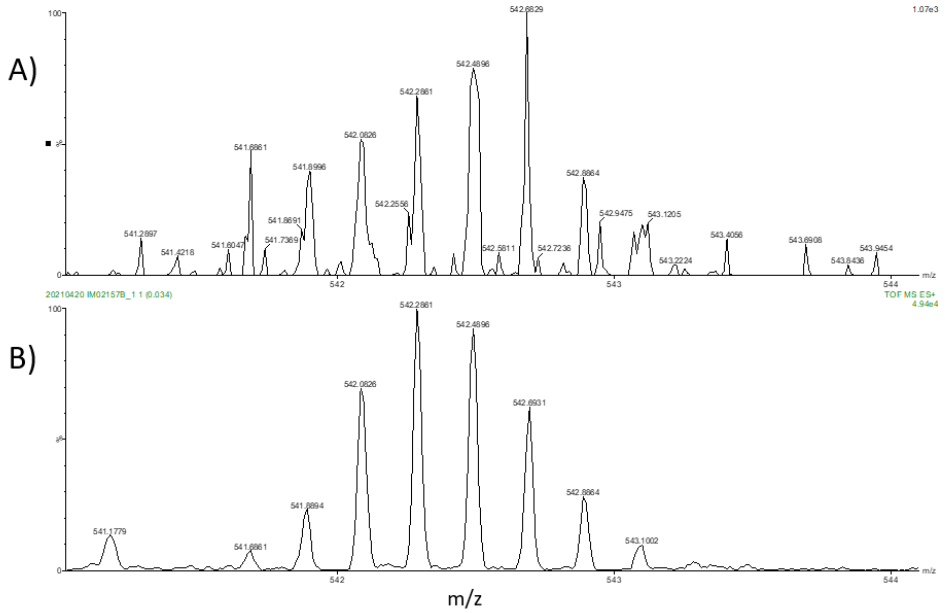


*Figure 2.7. Comparison of **C1** ESI-MS spectrum (5+ charge species) at different spectrometer voltages. A) Low voltage B) Higher Voltage*

It has been observed that by replacing the NO_3^- counter anions of the cage with PF_6^-, the solubility of the cage changes and it becomes easier to analyse via MS. However, for protein binding the cage needs to be water soluble so that it can exist in the same buffer system as the protein. Additionally, this study exists to probe the interactions of the **C1** system used in the SPECT imaging thus it needed to be the NO_3^- analogue of the cage that was investigated.

Initially, the protein-cage solution (in 100 mM AA) was run under ionisation conditions that were optimised for the cage only; however the conditions for ionisation appeared to be too gentle for the protein, as can be seen in Figure 2.8A. The gentle ionisation conditions led to a lack of resolution of the native protein signals, potentially due to residual salt, which is known to adversely affect the signal/noise (SN) ratio (this phenomenon is referred to as salt interference).[14] Whilst it is normally cations such as Na^+ that cause interference, a common protein contaminant in the form of NaCl, it is hypothesised that the counter NO_3^- anions could also be perturbing the SN ratio. Although it had been hoped that the majority of salt would be removed via the previously mentioned centrifugal filtration it appeared some had been retained.

As previous research has indicated that an increased sampling cone voltage can lead to salt dissociation, this was also attempted. [15] This appeared to help resolve the protein signals to a stage where the mass and charge state became clearer. However, the cost of this higher voltage appeared to be fragmentation of the cage, which subsequently could only be seen in the form of the free ligand

(Figure 2.8B). Even after further manipulation of the gas pressure and voltages on both the NanoMate and the mass spectrometer, there was no pronounced signals that could be confidently assigned to a cage:protein adduct.

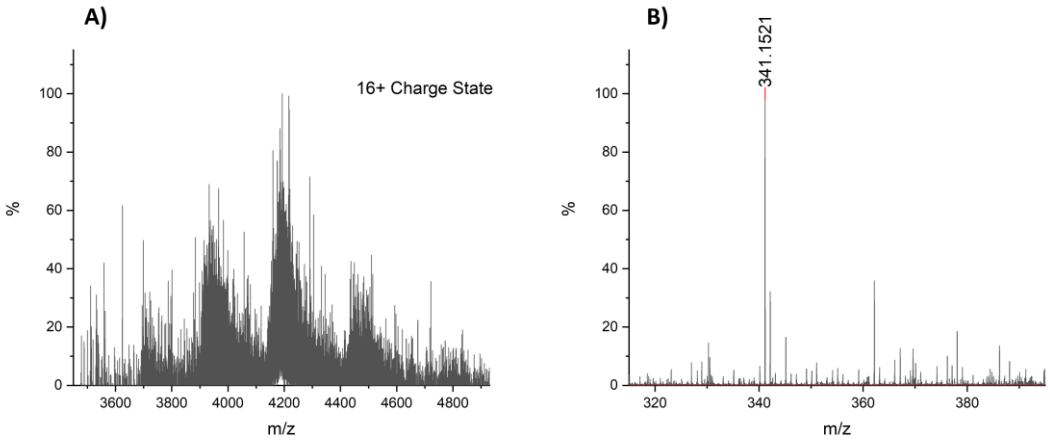


*Figure 2.8. ESI-MS of the Protein:Cage binding mixture at different voltages. A) Low voltage causing lack of resolution of 16+ charge state of HSA and B) High voltage causing cage fragmentation, loss of multiply charged **C1** species and only visualisation of the labelled ligand m/z signal.*

It was therefore hypothesised that the drastically different conditions required for analysis of the cage and protein independently could be hindering MS analysis of cage-protein interaction. Perhaps with utilisation of a tribrid ESI mass spectrometer capable of MS^n, as detailed by the Robinson group in a paper first outlining 'nativeomics', it might be possible to further study this apparently complicated system using MS.[12]. Moreover, given that there was still cage present in the protein sample following centrifugal concentration (as a method for desalting and buffer exchange) and the MWCO of the membrane was larger than the cage, retention of the cage is evidence of binding. This concept is further probed in the section "Centrifugal Ultrafiltration" below.

Nuclear Magnetic Resonance

Following the first investigations using MS, which relied on destructive ionisation of the protein-cage complex, NMR was attempted. Given that binding of the protein to the cage might be evidenced by the change in the magnetic environment of the protons (or relevant nuclei) located near the interaction, NMR could serve as a good technique for monitoring and characterising binding.

Traditionally this is done by either simply titrating the ligand into a solution of protein and monitoring the change in chemical shift of the protein signals, whereby shielding or de-shielding effects result from a change in the chemical environment caused by the binding of the ligand.[16] Alternatively, the protein is titrated into a solution of ligand and chemical shift perturbations of the ligand elude to

binding of the protein. Due to the significant overlap of signals in the 1D NMR of a protein, these "simple" experiments come with limitations but can stand as a good starting point. This is highlighted by Figure 2.9 whereby a change in the protein spectrum is evidenced upon the addition of cage but the changes are relatively minimal and difficult to quantify or numerically track. It is also important to consider the concentration limitations of the protein, whereby HSA remains folded and in solution only at lower concentrations than are typically optimal for NMR.

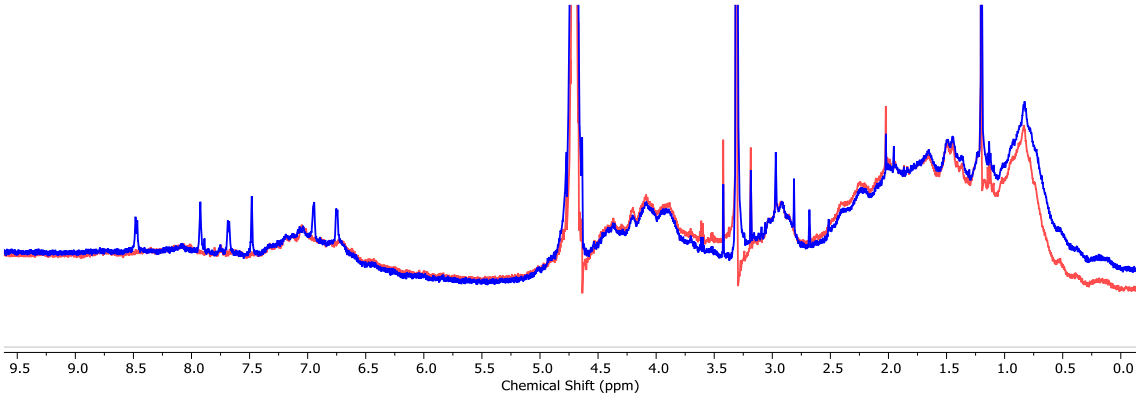


*Figure 2.9. Superimposed ^{1}H NMR (601 MHz, D_2O) Spectra of free protein in 0.1 M PBS in D_2O (red) and a protein with 5 Eq **C1** Cage in d-PBS (blue).*

In the preliminary NMR experiments, HSA protein was visualised in deuterated AA (100 mM) at 0.1 mM concentration; towards the upper boundary of the proteins' solubility. An 800 MHz spectrometer was utilised to gain the best resolution possible of the low concentration (Figure 2.10). The dispersal of sharp peaks across a wide chemical shift range is compliant with folded protein, whereby a narrow range of broader protein signals can be evidence of denaturing.[17] Whilst similar methods have been employed by others they are always used in conjunction with other, more reliable data, for example Williamson and Eaton, who showed the multi-site binding of epigallocatechin gallate (EGCG) to HSA using a combination of 1D-NMR titrations, multi-dimensional NMR experiments and isothermal titration calorimetry.[18] Given the inaccuracy of the K_d values obtained by other utilising simple one-dimensional titration data it was decided not to attempt to quantify the data obtained.

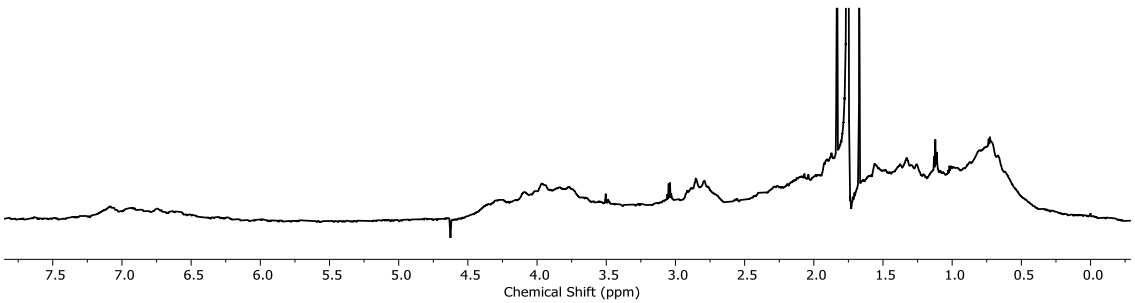


Figure 2.10. ^{1}H NMR (800 MHz, D_2O) spectrum of HSA Protein in 100 mM AA.

Slightly more complicated NMR experiments can be utilised to further understand the nature of the

binding, including saturation transfer difference (STD) and WaterLOGSY.[19] Both experiments rely on the Nuclear Overhauser effect, the transfer of spin polarization from one spin-active nucleus to another. It is important to consider whether changes in the protein or the ligand are being monitored; ligand based analysis can be preferential since less protein is required.

STD NMR allows the detection of weak binding events between ligands and proteins, typically for the calculation of K_d values between 10^{-8} M to 10^{-3} M. This technique works by monitoring spectral changes in the ligand spectrum when in the presence of protein, however whilst the traditional method above requires the protein concentration to be in the mM, for STD NMR the protein concentration can be as low as nM. The experiment works by having the ligand in a huge excess over the protein, and the collection of two spectra; off-resonance and on-resonance. The off-resonance spectrum will essentially just show the standard proton spectrum for the solution, whilst the on-resonance spectrum will selectively saturate the protein. If the protein is bound to the ligand the saturation of the protein will be transferred from the protein body to the bound ligand via the nuclear overhauser effect. Therefore any change in the intensity of the ligand peaks upon protein saturation is evidence of binding. By subtracting the on-resonance spectrum from the off-resonance spectrum the difference spectrum is obtained; the larger the peak in the difference spectrum the stronger the interaction of the proton with the protein.

Control experiments were completed to ensure that selective saturation of the protein could be achieved. A solution of protein in AA (100 mM in D_2O) was prepared, an STD experiment was then run and the difference spectrum was compared to the original spectrum (Figure 2.11). Given that there were no signals present in the on-resonance spectrum in just the presence of the protein, it was substantiated that there were no impurities that would interfere. Moreover, the on-resonance spectrum clearly showed the selective saturation of the protein. Given that the experimental conditions seemed optimal, the spectra were acquired using the same parameters in the presence of excess cage.

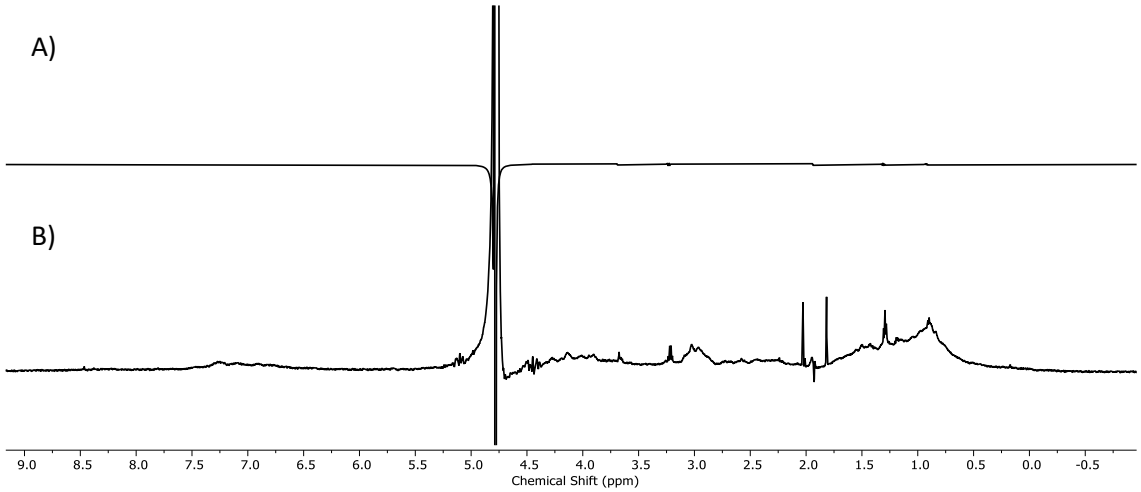


Figure 2.11. ^{1}H NMR (601 MHz, D_2O with 100 mM AA) Control STD NMR Experiment on Free Protein. A) On-resonance with saturation at -4 ppm, B) Reference spectrum of free HSA.

Spectra were obtained from an incubated solution of **C1** (2mM) and HSA protein (100 µM) in AA (100 mM in D_2O) and a difference spectrum generated (Figure 2.12). The difference spectrum is shown in blue against the reference spectra shown in red. If there is a strong interaction between the protein and the coordination complex there will be a transfer of saturation when the protein is selectively irradiated and the cage on-resonance signals will decrease intensity and thus the difference spectrum will have more significant intensity. When comparing the difference spectrum to the reference spectrum, it appears that protons H_B and H_A are binding strongest to the protein; these signals correspond to the outer most protons of the cage (see assignment Figure 2.12). They are also closest to the cationic metal centre, which could be hypothesised to have the strongest electrostatic interaction with negatively charged anionic amino-acid residues.

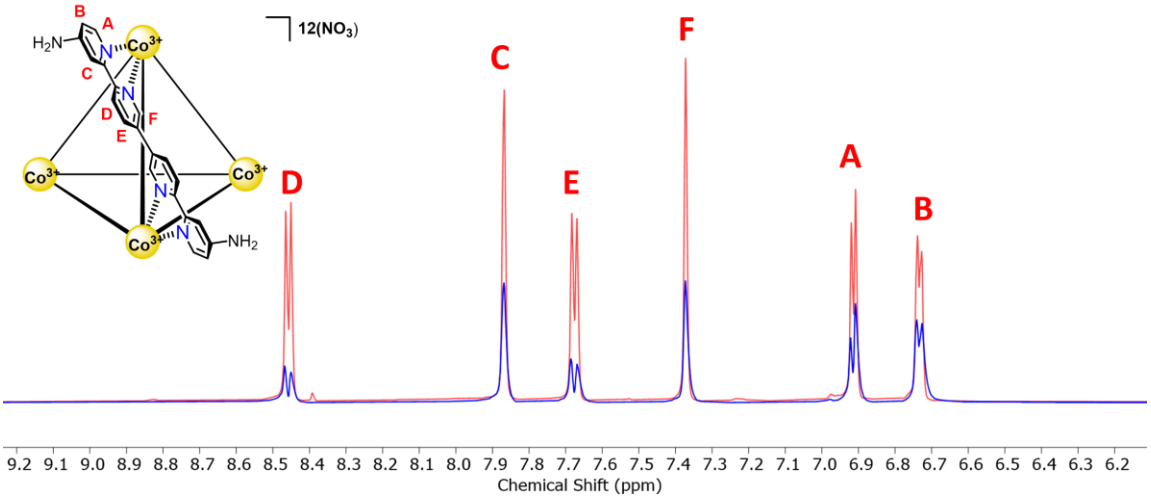


*Figure 2.12. ^{1}H NMR (601 MHz, D_2O with 100 mM AA) STD NMR Spectra of **C1** with HSA Protein. Reference spectrum showed in red and difference spectrum shown in blue.*

This binding can be quantified using amplification factors, determined via the equation underneath Table 2.1. At a given saturation time of 3.9 s, the amplification factor is defined as 'the average number of ligand molecules saturated per molecule of receptor'.[19] It uses the relative STD effect (I_{STD}/I_0), the ratio of the difference and reference integrals, along with the concentrations of ligand and protein. In short, a larger amplification factor is evidence of a stronger interaction. It is important to note the A_{STD} is also highly affected by saturation time along with ligand concentration. A longer saturation time allows more ligands to become saturated, decay of ligand saturation is slow and thus increasing saturation time will increase A_{STD}. A high enough ligand concentration is required to prevent already saturated ligand from returning to the binding site and the amplification factor will reach a plateau once the receptor binding site is saturated (*i.e.*, the ligand concentration is greater than the K_d).

To form a SAR comparison, the procedure was repeated for the three remaining complexes; **C2**, **C3** and **C4**. The spectral parameters, namely the saturation time (3.9 s), were maintained throughout allowing the generation of comparable amplification factors as displayed in Table 2.1. As can be seen **C1** has by far the largest amplification factor 1.8 a.u implying the most interactions with the protein compared to **C4,** which is much lower at 0.23 a.u. This is consistent with the hypothesis that multiple points of electrostatic interaction occur with the **C1** tetrahedral system compared to the **C4**. Perhaps slightly surprising is the low amplification factor of **C2** compared to **C1**, given that these cages are approximately the same size and shape. This would imply that the terminal amine groups of **C1** are required to enhance the interaction with HSA. The similarity between the amplification factors of C2 and C4 indicates that the interaction between C2 and the protein is occurring at a single vertex, analogous to C4. It is also noted that **C3**-protein sample showed some precipitate at the bottom of the NMR tube. Therefore, the low amplification factor of **C3** (0.11 a.u) may be due to a lower concentration rather than a weaker interaction.

Table 2.1. Amplification Factors calculated for each coordination complex from the STD NMR. Saturation time was 3.9 s, complex concentration 2 mM and HSA concentration 100 µM

Complex	**A_{STD}***
C1	1.80
C2	0.29
C3	0.11*
C4	0.23

Ultraviolet-Visible Spectroscopy

Another popular method employed for visualising the binding between two species is UV-Vis spectroscopy. Absorption traces of species can be distinctive and subject to change upon alteration of the structure of the chromophore. There are many examples of this technique being used to evidence binding including Kanakis and colleagues who showed flavonoids bind to HSA using UV-Vis spectroscopy. In this example, they observed changes in the absorption spectra of the flavanoids upon binding protein, allowing a simple titration assuming a one-one binding method.[20]

Given the electron configuration of the Co^{III} centres, the four complexes studied here are innately coloured and absorb light between 200-400 nm. This results from charge transfer between the pyridine based ligands and the cationic metal centres. Therefore it was hypothesised that the cages would have distinctive absorption spectra which could be used to monitor the interactions of the complexes with proteins. The UV/Vis trace for each complex is shown in Figure 2.13.

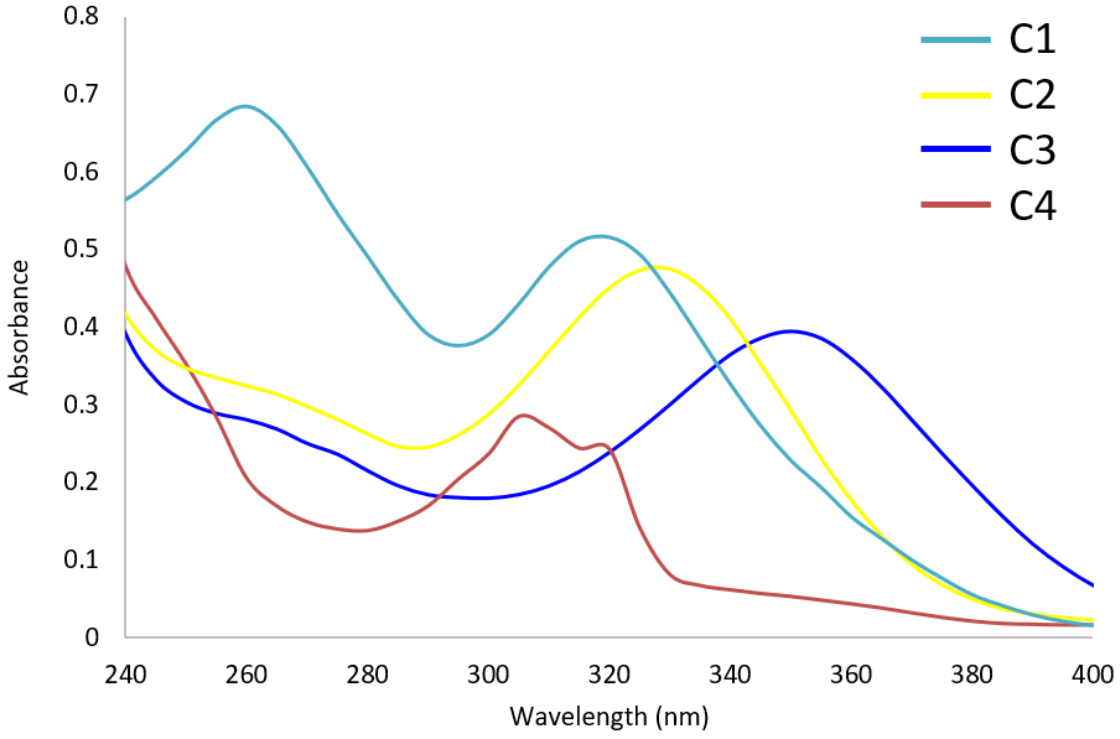


Figure 2.13. UV-Vis absorption spectra of the four complexes

Calibration curves were completed to determine the extinction coefficients of the four complexes, utilising the Beer-Lambert law to calculate the molar extinction coefficient at a series of known concentrations (Table 2.2). There is an almost direct relationship between the magnitude of the extinction coefficient and the number of Co^{III} centres in the complex. It appears each Co^{III} centre bound to three ligands contributes approximately 40,000 $M^{-1}cm^{-1}$ to the extinction coefficient, validating the

role of the transition metal in the absorption spectrum of the complex. The change in absorption maxima further highlights the different electronic effects of the ligand, perhaps best illustrated by comparing the QPTY cage and **C1** whereby they are differentiated simply by an ortho-amine on the ligand and yet this donating group changes the Co^{III}-pyridine bond sufficiently to alter the absorption spectrum of the complex.

Table 2.2. Extinction coefficients ($M^{-1}cm^{-1}$) of the four coordination complexes at stated wavelengths (nm). Calculated using the Beer-Lambert law on UV-Vis calibration data.

Complex	E ($M^{-1}cm^{-1}$)	Max Wavelength (nm)
C1	150,000	320
C2	150,000	335
C3	88,000	350
C4	30,000	305

A calibration curve was also generated for HSA, to both verify the calibration method and the protein being used, an extinction coefficient of 37,000 cm^{-1} M^{-1} was obtained which was in agreeance with literature.[20] The absorption maxima of the protein occurs at 280 nm whilst the maximal absorption resulting from the Co^{III} centres occurs at a slightly longer wavelength of over 300 nm. Two initial binding titration assays were carried out; the addition of **C1** into the protein (see Figure 2.14 below) and another where the protein was titrated into **C1**. As can be seen in Figure 2.14 below, there is purely an additive effect upon combining the cage and protein together, with no shift in the absorption maxima of the cage.

Similarly, there is no change in the absorption maxima wavelength of the protein. Whilst this doesn't disprove that the protein and cage are interacting it does indicate that any interactions taking place between the protein and the cage are not effecting the structure of the chromophore within the protein or the Co^{III}-pyridine bonds. It also further evidences that the cage and protein remain in solution upon interacting and that the cage remains intact, as seen by the absorption trace.

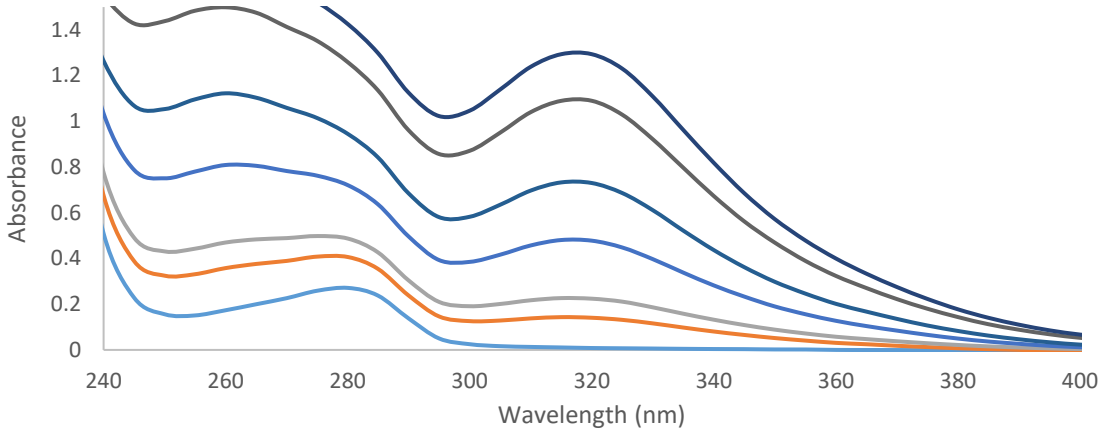


*Figure 2.14. UV-Vis Absorption Spectra of **C1** titration into a solution of HSA Protein (shown as the bottom light blue data set)*

Whilst UV-Vis cannot be directly used to determine any interactions between the complexes and the protein, the technique has been used in several of the indirect methods to quantify concentration following a binding experiment.

2.2.4 Indirect Assessment of Binding

The experiments described thus far have involved monitoring spectroscopic changes upon combining the protein and cage. Given the somewhat different analytical requirements for probing the protein and cage, for example, the contrasting preferable concentrations for NMR or the ionisation conditions for MS, it was considered that indirect, non-spectroscopic methods may provide more information.

Centrifugal Ultrafiltration

HSA has been shown to bind metal complexes for over 20 years with some of the first work performed by Shaw who determined the binding of Auranofin, an imitative natural product used to treat rheumatoid arthritis.[21] This was initially investigated using Penefsky spin columns, a technique developed by Harvey Penefsky in the 1970s, whereby a solution mixture is centrifuged through Sephadex gel leading to a filtration process whereby the larger components will elute first and the smaller components later.[22] Both Shaw and Penefsky showed that if spun independently, the drug of interest would be initially retained within the gel whereas when introduced and bound to the protein it would elute faster due to complexation to the larger protein. Shaw showed that a binding constant could be obtained by pulling the serum-Auranofin reaction mixture through Pelensky spun columns, whereby they were able to monitor the increasing ratio of Aurofin bound to protein at interval time points. It is worth noting that this method of binding constant determination only gives an approximation is not at true equilibrium.

The concept of using size exclusion to analyse the binding of free ligand and protein bound has been hugely developed since the initial studies by Pefensky. This includes but is not limited to techniques such as equilibrium dialysis, ultrafiltration, ultracentrifugation, gel filtration, electrophoresis and the partition method.[23] These techniques must be further subdivided to account for whether they allow the study of the system whilst in equilibrium. The two techniques focussed on here are equilibrium dialysis and ultrafiltration.

Both techniques function by loading both ligand and protein into one fraction and quantifying the physical separation of bound protein:ligand complex and free ligand across a permeable membrane. As implied by the name, equilibrium dialysis allows the ligand to pass freely between the fractions across the membrane thereby reaching an equilibrium. Ultrafiltration also consists of two compartments separated by a membrane, similarly the permeable membrane pertains to a molecular weight cut-off which allows the passing of the small ligand but retains the protein. Although in equilibrium dialysis, the ligand passively passes through the membrane over enough time to reach an equilibrium point, in ultrafiltration, centrifugal force is required to pull the ligand and some solvent through the membrane. Preliminarily ultrafiltration was used to get an indication of the binding capabilities of the four complexes with HSA. Vivaspin® centrifugal concentrators are found in most protein labs and provide a fast and simple method for gaining an initial insight into protein-ligand binding, as visualised in Figure 2.15.

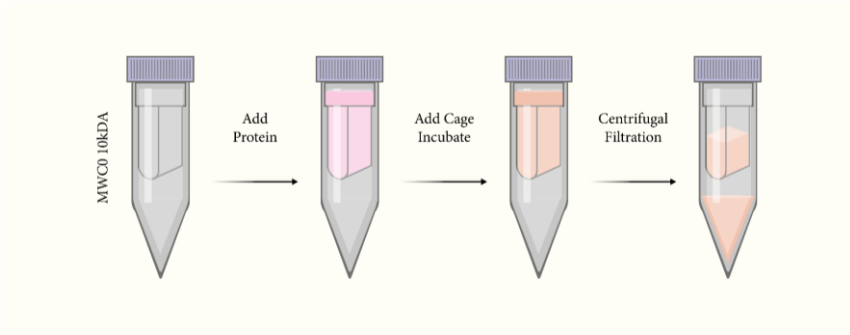


Figure 2.15. Visualisation of centrifugal ultrafiltration technique for preliminary protein binding investigation.

The permeable membrane of centrifugal concentrators is traditionally composed of polyethersulfones (PES), a family of thermoplastics with high hydrophilicity, temperature stability and controllable size permeability. They contain a subunit design of aryl-SO_2-aryl and subsequently can non-specifically bind larger lipophilic groups. To determine whether there was an interaction between any of the complexes and the PES membrane, control experiments were performed where solutions of the complexes in buffer were centrifuged through the membrane and the concentration of the retentate compared against the concentration of the filtrate. A molecular weight cut-off of 10 kDa was employed, which is

less than half of the protein of interest (HSA ≈ 66 kDa) and over double that of the largest complex (**C1** ≈ 3kDa). The hypothesis was made that if there was no interaction between the complex and the PES membrane then the entire concentration of complex loaded onto the centrifugal concentrator should be located in the filtrate after centrifugation, having passed cleanly through the membrane. As seen in Figure 2.16, the percentage in retentate for each complex for the control was under 10%, evidencing minimal to no interactions between the complex and the membrane. This was further verified by measuring the concentration in the filtrate.

Each complex was then incubated with a stoichiometric (1:1) concentration of protein for 1 h before being centrifuged As seen in Figure 2.16, this led to a large increase in the percentage remaining in the retentate for all four complexes, most notably the two tetrahedral cages. Given that the complexes pass through the membrane in the absence of the protein, it was concluded that they are retained in the initial compartment due to their interactions with the protein. For the two tetrahedral complexes, 70% of the complexes become bound to the protein at equimolar concentration. Again, the data indicates that the more highly charged tetrahedra interact most with the protein compared to the smaller, lower charged **C3** and **C4**. Around 40% of **C4** interacts with the protein, preventing it from passing through to the retentate. This seems like a relatively strong interaction and further evidences the potential for small charged species to non-specifically form electrostatic interactions with "sticky" proteins like HSA. The experiments were completed in triplicates and the data shown is the mean average.

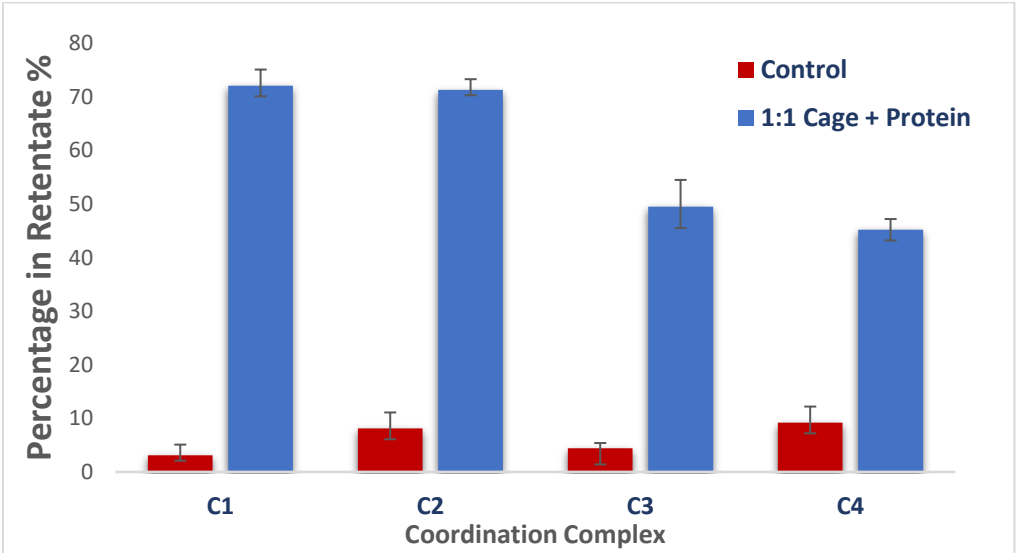


Figure 2.16. Centrifugal Ultrafiltration Results showing Percentage of Complex Retained in the Retentate for each Coordination complex in the absence (control red) and presence (1:1 stoichiometry in blue) of protein.

Rapid Equilibrium Dialysis

As mentioned above ultrafiltration cannot provide quantitative information on the system as movement through the membrane is irreversible and therefore the system is not at equilibrium. In contrast, equilibrium dialysis is based on the reversible passage of small molecules through a permeable membrane with a pre-determined molecular weight cut-off, whereby the receptor of interest is contained within one compartment. Equilibrium dialysis has been used for over 20 years to measure and quantify protein binding, more specifically it has huge applications in the field of plasma protein binding (PPB). It is quoted as avoiding the limitations of non-specific binding and large plasma volumes that can hinder the ultrafiltration and ultracentrifugation. Equilibrium dialysis has seen a fair amount of development over the last couple of decades, culminating in Thermofischers development of a 96-well Teflon tray with single use rapid equilibrium dialysis (RED) inserts, a schematic of which is shown in Figure 2.17.[24]

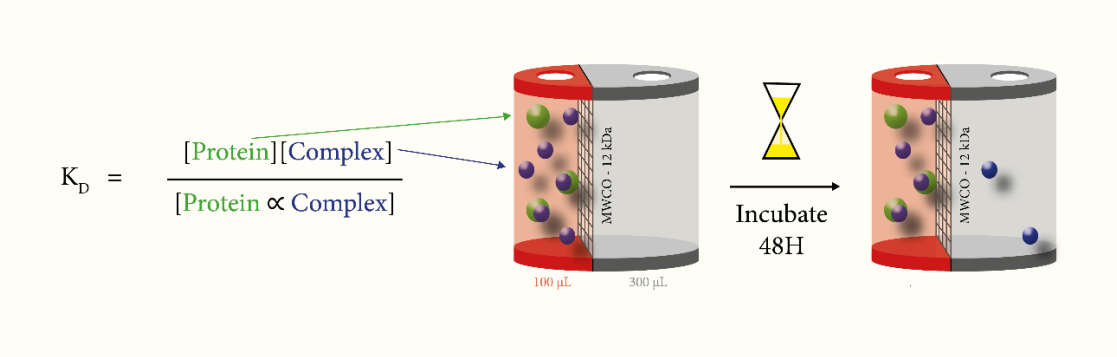


Figure 2.17. Visualization of Rapid Equillibrium Dialysis Inserts and methodology.

Before any binding experiments were undertaken, it was ensured that the complexes did not interact with the permeable membrane, composed of a cellulose matrix with varying molecular weight cut-offs. Any interaction between the complexes and the membrane would lead to a deviation from the passive equilibration process and a potential distribution bias. A solution of **C1** in buffer was loaded into the smaller red compartment of the RED insert, whilst a solution of the same buffer was loaded into the other, larger white compartment. The RED inserts used had a molecular weight cut-off of 8 kDa. The tray was then incubated (37 °C) on a rocker (200 rpm) for 24 h, before cooling and recording the resultant concentration of the cage in each compartment.

After 24 h, it was found that there remained a bias towards the red compartment in which the cage had been loaded. It was therefore hypothesised that the molecular weight cut-off selected might have been slightly too small, leading to steric hindrance between the membrane the cage diffusing through. The experiment was repeated with a slightly more permeable membrane, constituting a molecular weight cut-off of 12 kDa, with the other parameters maintained. There still remained a slight bias

within the compartment that the cage was loaded into and therefore the length of the incubation time was also increased. The optimisation conditions of the control experiment are exemplified in Table 2.3 below, whereby the optimal conditions for equilibrium were achieved with the 12 kDa MWCO membrane and 48 h incubation time, leading to a cage concentration ratio of 1.0 between the two compartments (*i.e.*, complete equilibration).

Table 2.3. Tabulation of the Optimisation Conditions for the Control Rapid Equilibrium Dialysis

MWCO of Membrane (kDa)	**Incubation Time (h)**	**Concentration Ratio (red/white)**
8	24	3.0
12	24	1.3
12	30	1.2
12	48	1.0

The relatively long time required to reach equilibrium may be a result of either a steric hindrance for the cage upon passing through the membrane or a potential reversible interaction between the cage and the membrane. There is extensive research into the comparison on permeable equilibrium dialysis and ultrafiltration membranes which essentially suggest that the hydrophilic cellulose and polyethersulfone are the two gold standards.[25] However, these studies investigate the applications of these membranes with respect to proteins and small drug molecules, thus it is noted here that these are structurally distinct to the metallo-organic cages investigated here. The control experiment was repeated with the remaining three complexes, to determine whether all three equilibrate under the same parameters.

Binding experiments were then carried out by introducing the complexes to HSA, then allowing the system to dialyse under the conditions obtained from the control experiments. The concentration parameters and calculation methods were then adapted from a literature study using the same dialysis kit.[26] Utilising a protein concentration of 80 μM in AA ensured that the protein remained in its folded and monomeric form as previously confirmed by high resolution NMR, native MS and size exclusion HPLC. The complexes were loaded such that in the absence of protein/protein binding they would equilibrate with a resulting concentration of 50 μM. All complex concentrations were pre-calculated and then verified using UV-Vis spectroscopy, as described in the UV-Vis Spectroscopy section above.

Calculations were performed using the Excel Workbook designed by Shave and co-workers, which uses the input of P_C (the ratio of ligand in each compartment in the absence of protein), P_t (the ratio of ligand in each compartment in the presence of protein), l0 (the ligand concentration used in the experiment, assuming equal distribution across the compartments) and t0 (the concentration of

protein loaded into the red compartment) to calculate the K_d for the system (Table 2.4).[26] Shave and co-workers completed a full derivation of the system by defining all the terms and using Wolfram Mathematica to solve for the required variable and editing the resultant formula into an accessible workbook. The resultant concentration ratios from the binding experiments and the derived binding constants are exemplified below as a mean average of the values calculated from the repeated experiments. It is noted in each experiment the concentration of the complexes were recorded before and after the experiment to ensure mass balance was being maintained.

Table 2.4. Rapid Equilibrium Dialysis results for the four complexes. The ratio indicates the complex concentration in the red compartment divided by the white compartment following incubation. Mass balance describes whether the final recorded complex concentration matched the initial concentration. K_d *and* K_a *are calculated using the method described by Shave et al.* [26]

Complex	**Ratio**		**Mass Balance**	K_d **(μM)**	K_a **(M^{-1})**
	No Protein	**With Protein**			
C1	1.00	1.61	YES	85	12000
C2	0.99	2.17	YES	31	33000
C4	1.01	1.31	YES	223	4500
C3	1.60	2.85	NO	N/A	N/A

As can be seen from Table 2.4, **C1** and **C2** show a significant bias towards the protein containing chamber, indicating an interaction. Given the initial hypothesis that any interactions found would likely be a result of the cationic complexes interacting with the anionic residues of the protein, and thus this electrostatic interaction would be stronger for the complexes with the higher charge, it appears consistent that the tetrahedral complexes show the strongest binding. The similarity in binding constants between the two cages indicates that the functional groups on the outside of the cage are probably not pivotal in the binding capabilities. Provisionally, STD NMR experiments had indicated that **C1** interacts most strongly, which was hypothesised could be caused by the cage amine groups hydrogen bonding with amino acid residues. However, given the equivalent binding constants with **C2** which exhibits not external functionality it is considered that the interactions are largely electrostatic.

Following this structure-activity relationship approach, the fragment based **C4** complex probes whether the size of the complex is important. The multiple vertices of the tetrahedral cages provide multiple electrostatic interaction points with available amino acids, whilst **C4** only pertains one. The

significantly larger dissociation constant of **C4** evidences this theory whereby the interaction is 10-fold weaker than the tetrahedra.[26]

Whilst both the tetrahedral cages, **C1** and **C2**, and the mononuclear complex **C4** exhibit a clear non-biased equilibration in the absence of protein, this was not the case for **C3**. Not only was there a pronounced concentration bias (1.6) in the red compartment in the absence of protein, which was much higher than the required 1.1, but there also appeared to be issues with mass balance. The amount of complex loaded into the insert at the beginning of the experiment was significantly higher than the compartments combined at the end of the experiment. It was hypothesised that this could be caused by a couple of factors; instability leading to disassembly into insoluble ligand, instability leading to formation of insoluble oligomers; instability leading to formation of soluble but larger oligomers or strong interactions with the permeable membrane.

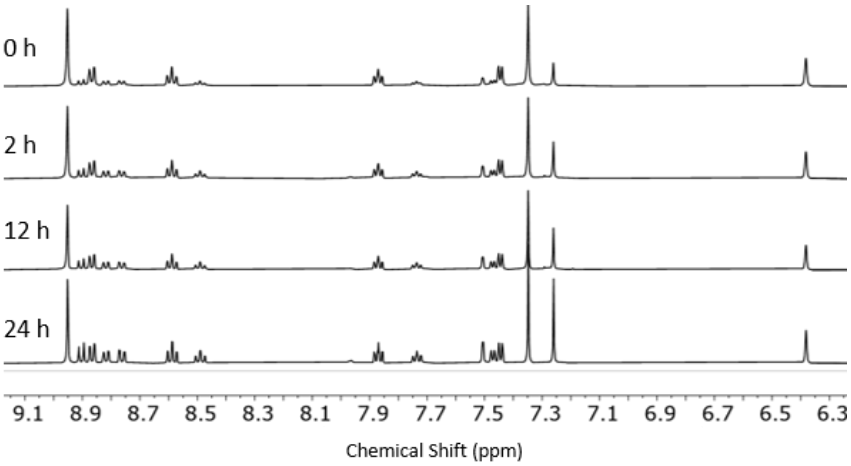


*Figure 2.18. ^{1}H NMR (500 MHz, D_2O) Spectra showing a solution of 2 mM **C3** over time.*

Each compartment was analysed using UV-Vis absorption after the completion of the experiment to calculate the concentration of complex in each compartment. As previously described, the varying Co^{III} assemblies possess distinct absorption traces and thus it was concluded that the formation of a larger oligomer unable to cross the MWCO membrane is unlikely to have caused the lack of equilibration. This is due to the absorption trace of each fraction clearly showing only the presence of **C3**; and in the red chamber some additional HSA protein. There is a relatively high level of strain possessed by **C3** and thus a series of control experiments were run measuring the concentration of **C3** in the dialysis solution over time. Thus investigating whether the heat combined with intrinsic strain would lead to disassembly.

It was found that heating **C3** (2 mM) in deuterated water at 90 °C over 24 h led to slow conversion from helicate to the tetrahedral species (Figure 2.18). This was contrary to what was expected given

the optimisation of conditions for the formation tetrahedral cage as exemplified in literature. [8] Moreover, given the decrease in absorption of **C3** after the dialysis this did not align with the conversion to the tetrahedral over time, which shows a similar UV/Vis spectrum. Essentially it appears likely that **C3** is decomposing into either something that doesn't absorb in the 200-400 nm wavelength range or it is insoluble. Precipitation may also be an explanation for the lack of equilibration in the control experiment with only **C3**, where insoluble impurities may block the pores preventing passage from one compartment to the other.

Next, imitating the conditions of the equilibrium dialysis, a solution of **C3** (0.1 mM) in AA (100 mM in D_2O) was made up and incubated at 37 °C, 200 rpm for 48 h and the NMR and UV-Vis recorded before and after. This showed a decrease in the concentration of **C3** indicating that the **C3** dissembles into insoluble free ligand over time. It is considered that this process may well be solvent and temperature dependent, whereby at higher concentrations and in non-buffered solvents, the complex may form the more stable tetrahedral from helicate as opposed to breaking down. In buffer, the presence of extra coordinating anions (*i.e.*, acetate) may facilitate the disassembly of **C3** into its constituent parts. As such, the apparent relatively strong binding of this compound was regarded as an artifact of the disassembly process and was thus discarded.

It is also noted that the dissociation model used to calculate K_d assumes 1:1 binding. Although the nature of the binding is unknown, and HSA possesses a large surface area that may possess multiple receptor sites, a 1:1 binding model is used for ease of calculations and to gain a provisional insight into the strength of binding. Thus the binding constants here should not be taken as definitive values and instead as a starting point to start to evaluate the differences in binding strength between the complexes.

Control experiments were also completed, showing that the binding experiments work both when loading the Co^{III} complexes into the red compartment with the protein, as described in the literature, but also when loaded into the protein-free white compartment (see experimental section). This indicates that the system is reaching equilibrium and further confirms that the Co^{III} complexes are able to reversible cross the semi-permeable membrane. Moreover, the red compartments ability to retain the protein under the optimised parameters was investigated. Preliminary literature research indicated a propensity for protein leakage over time, which could be mediated through the control of rpm and incubation length.[27] However, after incubating the protein in the red compartment at 37 °C, 200 rpm, minimal leakage, 3 %, was found in the white compartment.

Selective Precipitation Binding Methods

A further indirect method for determining protein binding was adapted from the work of the Coxon group, whereby they measured peptides binding to HSA by selectively precipitating protein and protein bound species.[28] There are two well-used methods for precipitating HSA from a solution, trichloroacetic acid (TCA) or organic solvent such as ethanol.[29] Given the known low organic solvent solubility of the four cobalt complexes investigated here, precipitation using TCA was investigated. Preliminarily this required stability tests to measure the tolerance of the four complexes to the acid, both their solubility and their ability to remain assembled over time. Any complex precipitation caused by insolubility or disassembly to free ligand upon addition of the acid would prevent selective protein precipitation and thus lead to unusable data. NMR spectra of the four complexes in D_2O (1 mM, 0.5 mL) were recorded before and after the addition of 50 µL of TCA (15 % w/v, D_2O). Whilst **C2**, **C4** (Figure 2.19) and **C3** showed unchanged spectra upon addition of the acid, this was not the case for **C1**, which showed a set of new peaks.

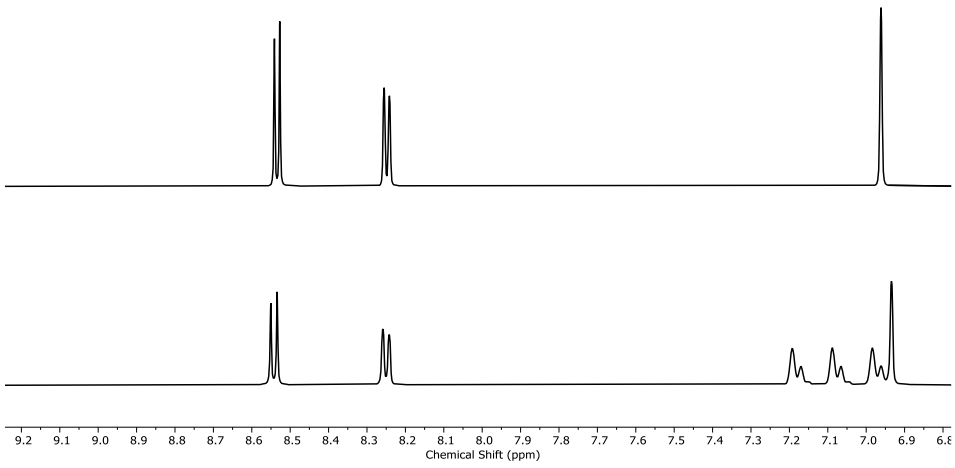


*Figure 2.19. ^1H NMR Spectra of **C4** in D_2O (top) and in D_2O with 10% TCA (15% solution) (bottom)*

Upon further investigation with acetic acid, **C1** also exhibits new peaks with the addition of acid (Figure 2.20). The possibility of the second set of species in spectrum A belonging to free ligand can be outlawed. There appear to be two possible causes; the pyridyl-amine is becoming protonated leading to the formation of a new species or the acid is a slow-exchange guest within the cage. It seems unlikely that the amine is being protonated, this is due to the inactivity of the amine group. Once coordinated to the cobalt the electron-density from the lone pair on the nitrogen are likely donating into the pyridyl ring system which in turn is coordinating to the cationic cobalt, inactivating the amine. This phenomenon is discussed further in the experimental section of Chapter 3, considering the various methods for functionalising **L3** via the amine. Given the similarities of the two tetrahedral cavities it seems highly unlikely that the acid would be a strong guest within **C1** and not **C2**. It is

hypothesised that the amine groups are playing a role in the reaction between the cage and the acid but no further conclusions are drawn here and **C1** is ruled out for the precipitation experiments.

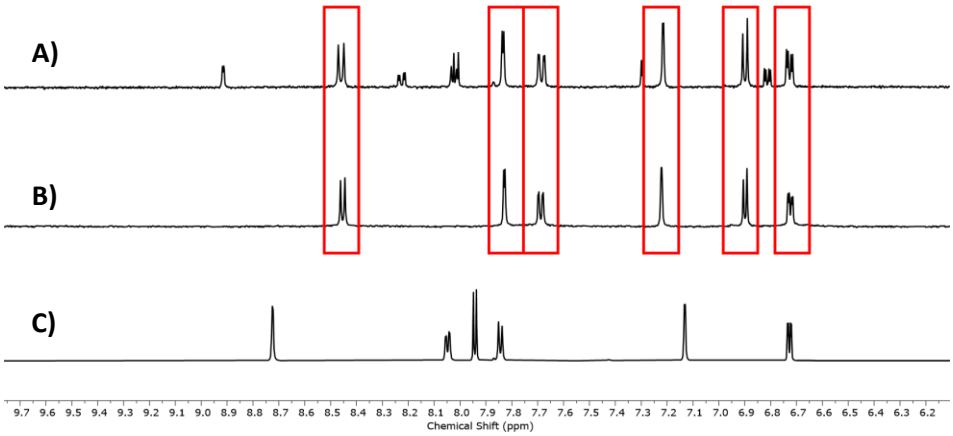


*Figure 2.20. ^{1}H NMR (500 MHz, D_2O) Spectra showing effect of acid on **C1**. A) **C1** in 0.1% acetic acid B) **C1** and C) **L3** in 0.1% Acetic Acid. **C1** signals are highlighted within the red boxes.*

Continuing with **C2**, **C3** and **C4** complexes, the precipitation experiments were analysed using UV-Vis instead of NMR, analogous to the equilibrium dialysis experiments. For the control experiments, the complexes were dissolved in AA (100 mM in H_2O) at a known concentration, before TCA (15 % w/v) was added to each followed by vortexing, centrifugation and re-measuring the concentration of the supernatant. The experiments were repeated in the presence of protein, allowing for a 24 h incubation period prior to the addition of the acid and recording the resulting concentration of complex in supernatant solution. The concentration of each complex in the supernatant was compared against the corresponding control and a percentage bound calculated as shown in Figure 2.21. The experiments were repeated over a variety of protein concentrations. The data obtained from these experiments are shown in Figure 2.21, alongside an estimate of their K_d values (Table 2.5).

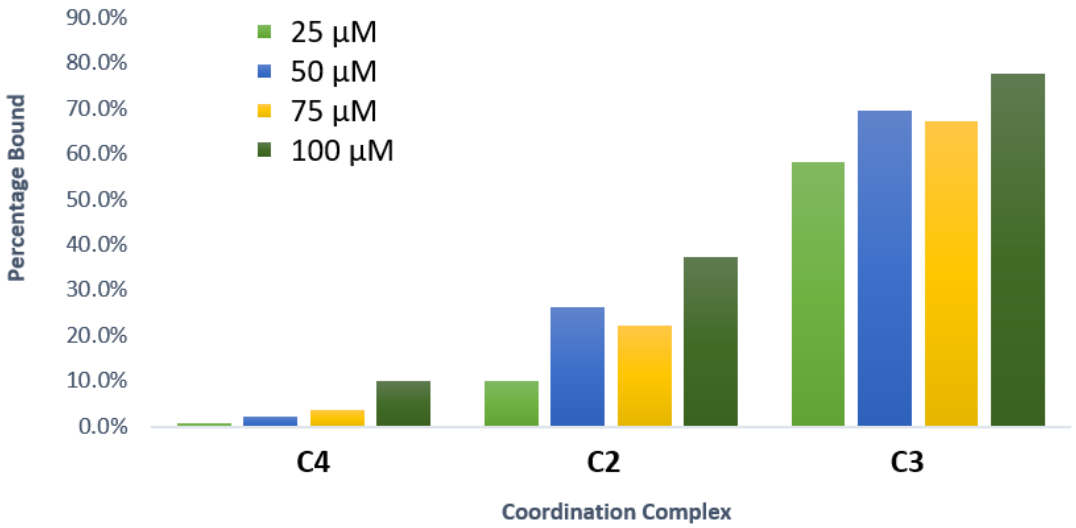


*Figure 2.21. Precipitation Data for **C2**, **C3** and **C4** showing the percentage of complex bound at 25 µM (green), 50 µM (blue), 75 µM (yellow) and 100 µM (dark green) of HSA. Complex concentration is at 50 µM.*

*Table 2.5. Dissociation Constants calculated from precipitation experiments for **C4**, **C2** and **C3**.*

Complex	K_d (µM)
C4	500
C2	150
C3	50

Given the range of the data for both **C4** and **C2** an extrapolation was required to calculate the concentration at which 50% of the complex was bound, therefore slightly decreasing the reliability of the method. Moreover this method precipitates the bound protein and complex out of solution and thus disturbs the equilibrium. Subsequently the dissociation constants are used to formulate a pattern between the complexes' binding strength rather than as absolute values. The same pattern is exemplified here as with the previous experiments whereby the larger species with the higher charge densities seems to more strongly interact with the protein. Whereby the mononuclear C4 exhibits the weakest binding and the larger and more highly charged C2 and C3 complexes correlate to lower K_d values.

Whilst the acid controls indicate **C3** complex is stable under the reaction conditions, given the outcome of other binding experiments described earlier in this chapter, the strong binding data here is treated with caution. Given the propensity of **C3** to disassemble into free insoluble ligand this could easily cause a reduction in complex concentration, not as a result of selective protein precipitation.

2.2.5 Gel Electrophoresis

One of the few examples within literature investigating the interactions of coordination cages with biomolecules, is the 2021 work by Nitschke and co-workers which evidences the binding of a similar Fe^{II} based tetrahedron to DNA G-Quadruplexes. One of the methods used to show the binding was native polyacrylamide gel electrophoresis (PAGE) whereby it was shown that upon the addition of cage, large aggregates form bands along the top of the gel (Figure 2.22). [30]

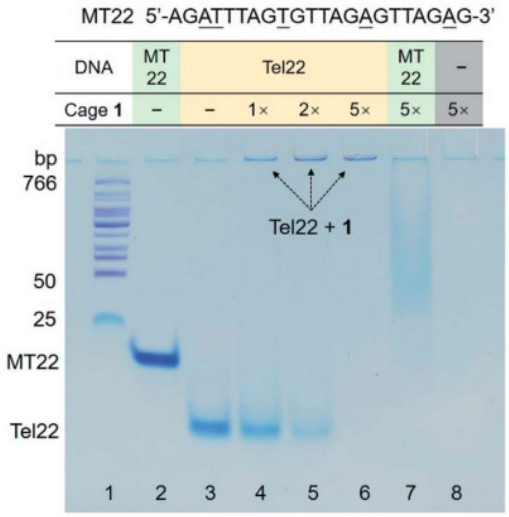


Figure 2.22. Interaction between DNA G4 (Tel22) and Fe^{II} tetrahedral cage studied by polyacrylamide gel electrophoresis (PAGE), figure adapted from work by Nitschke et al.[30]

A similar experiment was carried out with **C1**, whereby four HSA solutions were loaded into a pre-cast polyacrylamide gel and run against a protein ladder. As shown in Figure 2.23, lane A contains the protein ladder and lane B contains free protein; these are reference points for the binding experiments. For lanes C and D, the protein was incubated (37 °C, 24 h) with 1 Eq and 5 Eq of **C1**, respectively, and lane E contained a diluted solution of protein. As can be seen by the thick bands at the top of lanes B to E it is difficult to determine whether there is any change in the protein band. Given the inaccuracy of this method for measuring molecular weight and the already very large size of the protein it is hypothesised that it would be difficult to see the addition of a single cage onto the protein band.

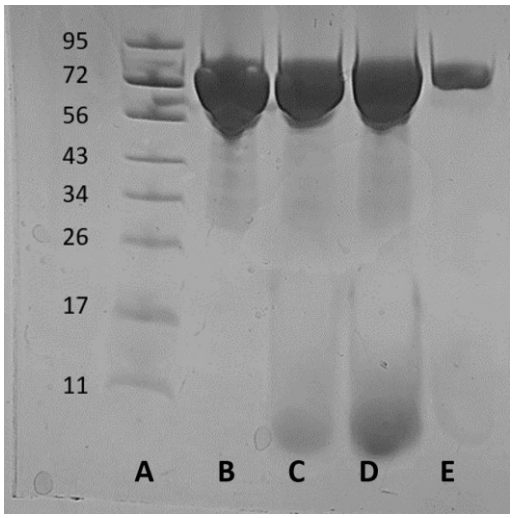


*Figure 2.23 . Gel Electrophoresis of the HSA in the presence of **C1**. A) Protein ladder B) HSA Protein C) HSA Protein with 1 Eq **C1** D) HSA Protein with 5 Eq **C1** E) HSA Protein at 5X dilution*

It is worth noting that the cage appears to be visible under the gel and staining conditions, which appears to show there is free cage at the bottom of the gel. This may be of use for future studies, for example, investigating the interactions between cages and lower molecular weight biomolecules such as DNA fragments (similar to the Nitschke work) or smaller peptides. No further conclusions are drawn from the gel electrophoresis work in this study.

2.2.6 Isothermal Calorimetry Titration

As technologies develop, there is a shift towards high-throughput methods which allow the rapid generation of binding data for a series of receptors and substrates. Typically the most popular methods for studying the thermodynamics of equilibrium systems include isothermal titration calorimetry and surface plasmon resonance.[31]

Isothermal titration calorimetry (ITC) was pioneered in the late 20th century whereby it was described as a technique to 'Determine the *K* in Minutes'.[32] It works by monitoring the heat change associated with a ligand binding to a receptor. These heat changes are then compared against a reference cell and binding constants can be obtained from the observed change in enthalpy of the system.[33]

Alternatively, surface plasmon resonance (SPR) is an optical technique that immobilises the receptor onto a surface, typically a thin gold film via amine-coupling bonds, and a mobile substrate is allowed to flow over the film. When binding occurs between the mobile substrate and the immobilised receptor there is a perceived change in the refractive index observed off of the gold film. The technique works by monitoring the change in refraction upon binding to obtain numerical binding data for the dynamic system, similarly to ITC.[31]

Within medicinal chemistry, these two techniques are used to validate the binding between drugs and their target proteins/receptors. However, the techniques can be expensive, optimisation can be lengthy and often the equipment is not available in all labs. Given the requirement of the generation of the gold film layer with bound protein for surface plasmon resonance, it was concluded that ITC would be a more straightforward method of analysis. This is furthered by the breadth of thermodynamic information that can be gathered from just a few experiments, including enthalpy and stoichiometry of binding. Until this point a 1:1 binding method had been assumed but not evidenced.

All ITC experiments were carried out at 25 °C on a MicroCal Auto iTC200 (GE Healthcare) instrument using water within the reference cell. Initially, controls were run that involved adding a solution **C1** in AA (100 mM in H_2O) into a blank solution of AA, to ensure there was no background solvation effects. However, preliminary experiments showed an endothermic isotherm implying that as the cage became more diluted in AA heat was being consumed. It was concluded that the heat changes upon cage dilution into AA were significant and thus the conditions could not be used as a control. Alternative, phosphate buffered saline (PBS) was employed), which subsequently showed only small exothermic heats of dilution. These conditions were used for all subsequent experiments.

Titrating **C1** into HSA protein yielded the isotherm shown in Figure 2.24. This excellent fit was obtained by processing the data using the MicroCal PEAQ-ITC Analysis software. Carrying out an ITC measurement using **C4** (see Figure 2.25) showed, as expected, a much weaker interaction. The thermodynamic parameters for these experiments are summarised in Table 2.6. Errors for the thermodynamic parameters occurring from the 1:1 and 1:5 binding model fits are included, whereby small errors are indicative of a good fit to the model.

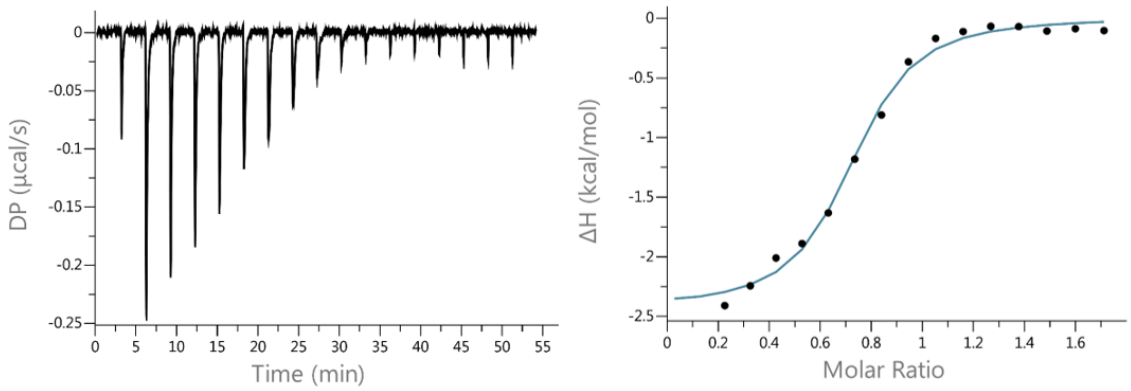


*Figure 2.24. Isothermal Titration Calorimetry Isotherm for **C1** titrated into HSA Protein.*

The thermodynamic data shows that the strength of binding between **C1** and the protein is significant, with a K_d of 1.66 µM. The magnitude of this association constant is comparable to several

"conventional" ligands that are considered strong binders.[34] The stoichiometry of binding suggests there is 0.650 Eq of **C1** per protein, a sub-stoichiometric *N* value can result from not all of the protein existing in its folded active form. Typically, when working with protein a portion will have misfolded and remains inactive, which is especially true when working at higher concentration as oligomers can form deactivating the protein. Given that these experiments were performed towards the upper boundary of HSA solubility using a 200 µM stock solution, it is possible that the sub-stoichiometry is a result of some misfolding leading to a lower concentration of active protein. This would lead to an underestimation of binding stoichiometry, thus inferring a 1:1 model seems reasonable. All the data was modelled across a series of binding isotherms to determine the model with the best fit.

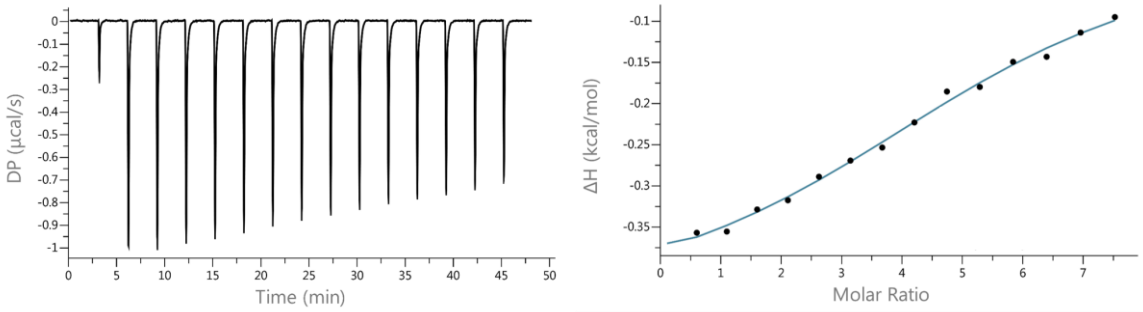


*Figure 2.25. Isothermal Titration Calorimetry Isotherm for **C4** titrated into HSA Protein.*

With a 1:1 binding model and a strong K_d it is inferred that the association must involve multiple metal vertices of the cage and these are likely simultaneously interacting with multiple anionic amino-acid residues. Another explanation would be the favourable displacement of more solvent due to the relatively large size of **C1**. Both of these interpretations are supported by the 100-fold weaker binding of the mononuclear **C4** complex (K_d =147 µM). The much smaller size may also explain the 5:1; C4:HSA, binding model, implying multiple non-specific interactions with single anionic sites across the proteins surface.

*Table 2.6. Kinetic and Thermodynamic Parameters obtained from ITC experiments on **C1** and **C4***

Parameter	**C1**	**C4**
Stoichiometry	0.650 ±0.025	5.24 ±1.11
K_d (µM)	1.66 ±0.67	147 ±107
ΔH (kcal/mol)	-2.71 ±0.21	-0.480 ±0.157
ΔG (kcal/mol)	-7.90	-5.24
-TΔS (kcal/mol)	-5.19	-4.77

The higher error for the K_d of **C4** are noted as a result of the weak binding lying at the upper boundary

of the instrument's capabilities. This is evidenced by the isotherm shown in

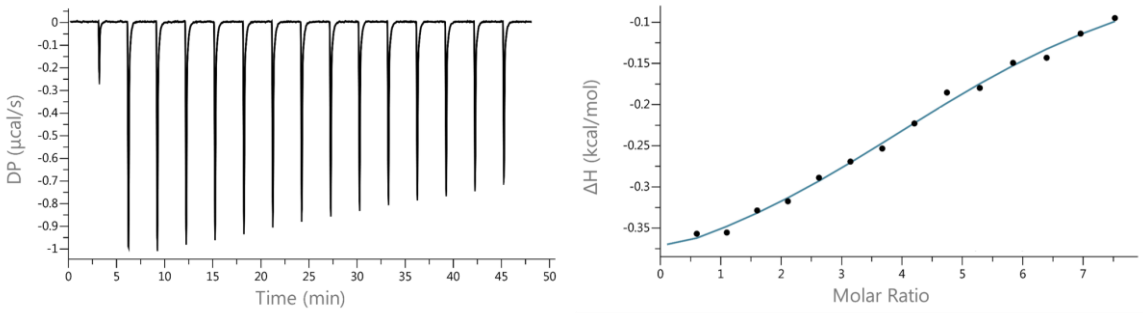


Figure 2.25 whereby there minimal change between the heat released at the beginning of the experiment and the end, suggesting the system barely reached a saturation point. This lack of saturation is more obvious in the processed data, which shows that Δ*H* has not reached a plateau. Given the quality of the data and comparisons made against the control, it was concluded that the heat changes were still significant and thus the parameters could still be determined. The **C4** K_d value serves as a comparison against **C1** signifying the weaker binding and the higher ratio of complexes per protein.

2.2.7 Computational Modelling of Binding

Computational modelling has become an essential tool for screening potential ligands for known receptors in drug discovery. More specifically, molecular docking is used to predict the non-covalent interactions between a defined small molecule and a receptor, proposing a preferred orientation and a docking score which indicates the predicted strength of binding. Molecular docking is often used in conjunction with molecular dynamics; a computer simulation dealing with the energies and force fields of atoms within a system over a defined period of time. Used in conjunction the two techniques can provide an insight into which ligands will bind a receptor, the conformation of the ligand and the binding site on the receptor.[35]

Over the last couple of decades an array of virtual screening software has been developed, including the freely available AutoDock and more recently upgraded AutoDock Vina. AutoDock Vina was developed by Olson and co-workers at the Scripps Institute and has been used extensively, boasting over 12,000 citations for their original paper.[36] The software utilises the Lamarckian Genetic Algorithm in a two tier method for first locating the active site of the receptor and docking the ligand into said site, and then finding the optimal conformation and orientation.[35] The various outcomes of different binding sites and conformations are then scored based on the strength of the predicted interactions. A crystal structure of **C1** was obtained via slow water evaporation (Figure 2.26) and the cif file

converted for usage in the docking studies. For further discussion of the crystal structure see Chapter 3.

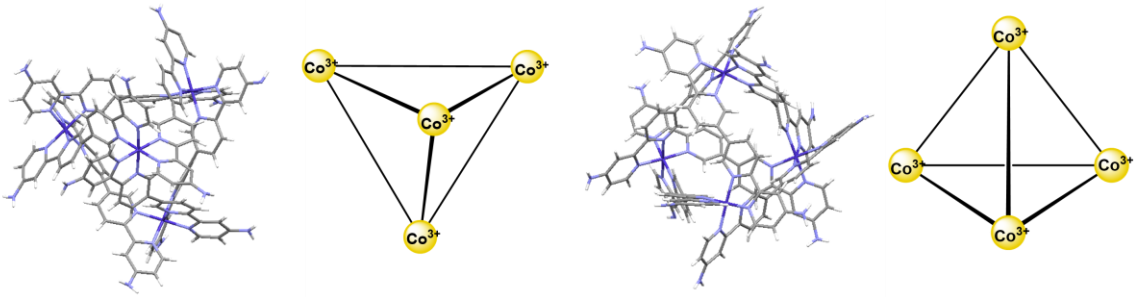


*Figure 2.26. X-Ray Crystallography Structure of **C1** shown from two perspectives.*

The cage file was prepared for docking using MGLTools, also developed by the Scripps Institute, allowing the control of torsion within the cage. Given that MGLTools was designed for the application of small drug molecules, unless implicitly stated by the user simple molecular orbital theory is applied and most bonds are viewed as rotational. It is assumed here that **C1** retains its' distinct and rigid tetrahedral shape when interacting with the protein.

This assumption is evidenced and justified by the experiments described previously which all indicate that the cage remains intact upon binding to the protein. Preliminary computational docking experiments suggested the cage became highly distorted and fragmented upon associating with the protein. Given the experimental evidence implying the cage remains intact, the number of torsions within the cage were limited along with confining the degrees of freedom, allowing the programme to model and dock the intact cage. Since none of the experimental analysis indicated where on the protein binding was occurring, the docking simulation grid box was set to cover the entire protein, thus exploring all potential binding sites. Two potential binding sites were located, with the first showcasing the highest docking score and the second with the cage in a series of orientations (Table 2.7).

Table 2.7. Summary of the two binding sites as calculated from AutoDock Vina. For binding site 2 there were 8 orientations and the average distance relative to binding site 1 determined.

Binding Site	Affinity (kcal/mol)	Distance from rmsd l.b	Distance from rmsd u.b	Orientations
1	-13.1	0.00	0.00	1
2	-10.7	32.7	40.1	8

The low number of proposed binding sites is consistent with the 1:1 binding stoichiometry determined from the ITC experiments. Interestingly, when overlaying the proposed cage binding sites and orientation with the isoelectronic map of the protein, it doesn't appear to be binding at an anionic site of the protein (Figure 2.27). Whilst binding site 1 does seem to be located on an anionic area of the protein, binding site B seems to be located between anionic and cationic surface areas. This could imply there is direct interactions with specific anionic amino-acid residues as opposed to a general electrostatic attracted to an anionic region.

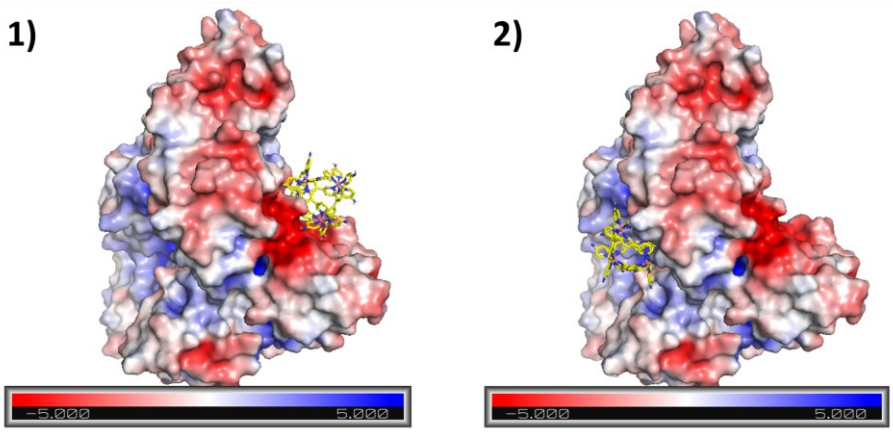


Figure 2.27. Visualisations of the binding sites 1 and 2 determined using AutoDock Vina overlaid the electrostatic map of HSA generated using PyMOL (APBS Electrostatics)

Given the unexpectedly cationic nature of binding site 2, an investigation of the amino-acids involved in the interactions with the cage was carried out using PyMOL (Figure 2.28). Where the amino-acids with the closest proximity to the cage were identified and visualised. At one cage vertex, shown at the top of Figure 2.28, two glutamate residues can be seen pointing towards the metal centre. This interaction may either be an electrostatic interaction between the anionic residue and the cationic metal centre or it may be a result of hydrogen bonding between the amine on **C1** and the carboxylate on the residue. In fact, Glu 276 is a mere 2 Å away from the amine group on the vertex of **C1**, well within a typical hydrogen bond distance range.

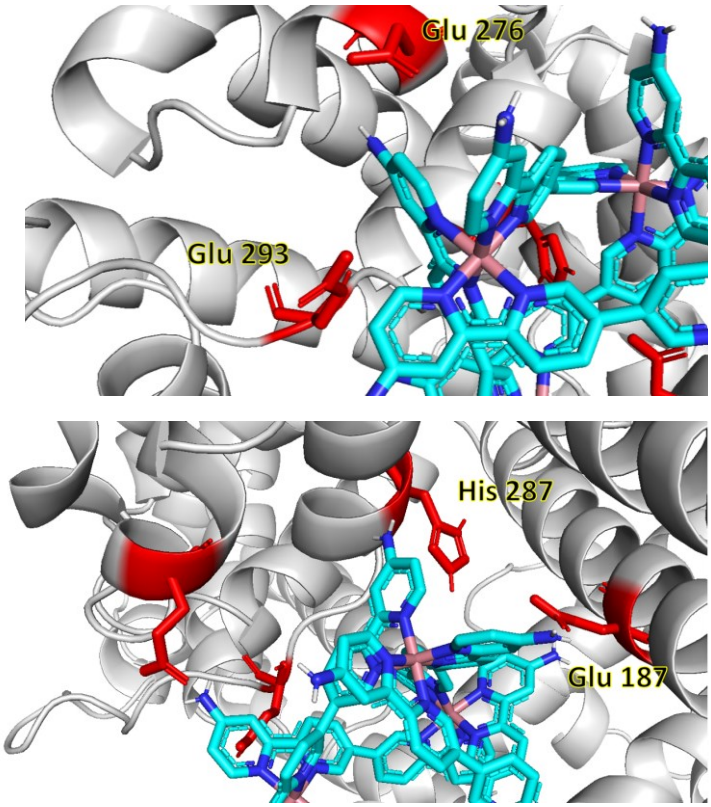


*Figure 2.28. Map of binding site 2 showing amino-acids (red) present and predicted to involved in binding **C1** (Blue), visualised using PyMOL.*

Interestingly, located at the vertex pointing into the protein (shown at the bottom of Figure 2.28) is another glutamic acid along with a histidine residue. Whilst histidine residues are known for their ability to coordinate cationic metal centres, the metal centre coordination here is fully saturated therefore coordination is highly unlikely. There may be a hydrogen bond interaction between the amine group on the cage and the protonated histidine. Whilst this mapping of the binding site can gave an insight into the type of interactions and the amino-acids involved it is not definitive. AutoDock Vina is typically designed for the screening of libraries of small drug molecules and not rigid, highly charged architectures as described here. This is corroborated by the restrictions and torsion placed on the cage system to allow the program to run effectively. These results are used in conjunction with the previous techniques to help visualise what the binding site might look like.

2.2.8 Protein Crystallisation

Given the evidence of a relatively strong and stoichiometric interaction between **C1** and the protein alongside the preliminary computational models suggesting a specific binding site, an attempt at co-crystallisation was made. HSA protein has been crystalized over a thousand times both independently and bound to other species, initial crystallisation conditions were identified by searching the RCSB Protein Data Bank and locating the work of Ghuman and co-workers.[4] Initial attempts were made to crystallise the protein independently and then soak the crystals with the **C1** complex in attempt to crystallise them together. However, attempts to gain singular HSA crystals were unsuccessful and although crystals were obtained these were assumed to be phosphate salt crystals. Alternativery, co-crystals were attempted by loading both HSA and **C1** (100 mg/mL of each) into the same well and allowing them to bind before crystallation. This resulted in the generation of large orange crystals which upon inspection appeared to be protein crystals (Figure 2.29). Given the orange colour attributed normally to **C1** it was hypothesised that these could be crystals of the cage bound to the protein, subsequently the crystals were frozen and sent to the national synchrotron. At time of writing we are awaiting the results.

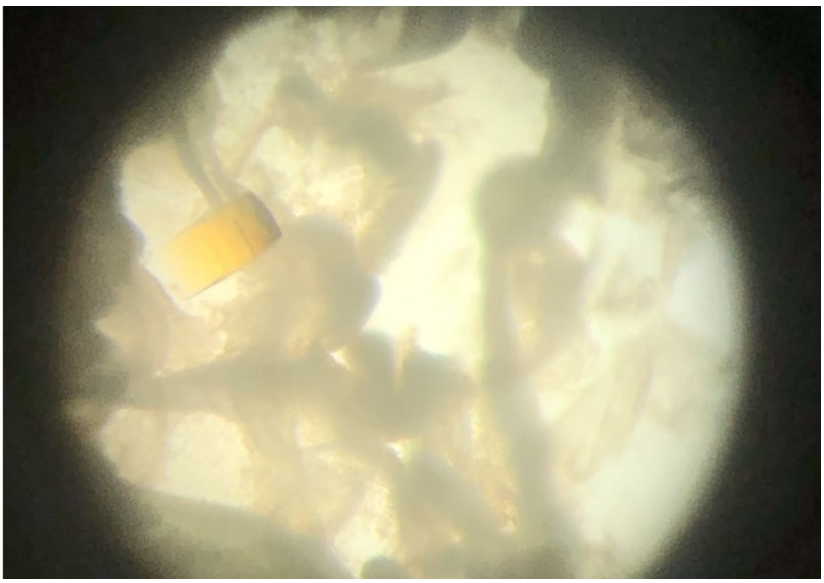


*Figure 2.29. Photograph of **C1** and HSA crystallisation experiment. Grey cloud like structures are protein precipitate and orange rectangular crystal is suspected **C1**:HSA co-crystal.*

2.3 Conclusions and Future Work

In summary a series of experiments have been performed on both the parent **C1** and three related complexes. Whilst methods such as using MS, UV-Vis spectroscopy and gel electrophoresis were unsuccessful for decisively showing any interactions between **C1** and HSA, further experiments indicated favourable coordination complex:protein binding.

STD NMR experiments indicated all of the complexes interacted with HSA but by comparison of the amplification factors the binding was found to be significantly stronger for **C1** than for **C4.** This was further built upon by centrifugal ultrafiltration whereby **C1** showed 70% of the complex remained above the permeable membrane when in the presence of stoichiometric protein, compared to **C4** which indicated a 40% retention of complex in the retentate.

Binding constants were then calculated using equilibrium dialysis, selective protein precipitation and isothermal titration calorimetry. Whilst there was variation between the three methods, the same trend was seen in each case, whereby the higher charged tetrahedral species exhibited a stronger binding constant that the smaller, lower charged species. This validates the hypothesis that the binding is due to electrostatic interactions and implies that the tetrahedral cage is likely interacting via multiple cationic metal centres. Using ITC and computational docking, **C1** was evidenced to be 1:1 binding aided by interactions between both glutamic acid and histidine residues with the cage vertices. ITC showed that there were likely five **C4** binding each monomer of protein.

As a strong binder of HSA, this increases the number of potential biomedical applications of **C1**. It is hypothesised that if the original SPECT imaging studies were repeated by co-loading **C1** with HSA the resultant bioaccumulation of the encapsulated [^{99m}Tc][TcO_4^-] might be further altered.

For example, a variety of cancerous cell lines over produce glycoprotein90 (gp60), a cell surface protein capable of binding HSA. This binding has been utilised by Abraxane, a commercially available targeted drug delivery system composed of HSA and paclitaxel, a chemotherapy agent used against a variety of cancers. Abraxane is FDA approved for use against metastatic breast cancer and metastatic lung cancer and pancreatic cancer as HSA facilitates tumour and cellular uptake.[37] For a drug to have applications based on its ability to bind to proteins the association potency must be K_d << 10 μM which is the case for **C1** and HSA. [16]

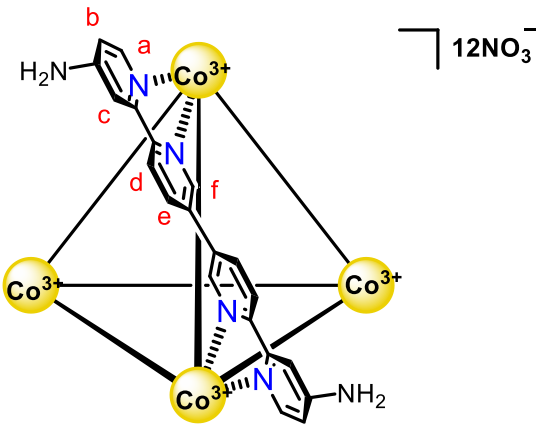
2.4 Experimental

2.4.1 General

All reagents and solvents were purchased from Alfa Aesar, VWR, Fluorochem or Merck and used without further purification. Phosphate buffered saline (PBS), was produced using buffer tablets purchased from fischer scientific and ultra-pure deionised water or D_2O to a resultant concentration of 0.1 M. All buffer was filtered through Millex-HP Syringe Filter Units (0.45 µm, polyethersulfone) prior to use. HSA was purchased as a lyophilised, desalted, non-denatured white powder (>98% purity) from Merck and protein solutions were stored at 4 °C. For Rapid Equillibrium Dialysis (RED) the RED Device Inserts (8 kDa and 12 kDa cut off) and disposable plate base were purchased from Thermo Scientific.

All ^{1}H, ^{13}C, DOSY and STD NMR spectra were recorded on either a 600 MHz Bruker AV III equipped with a TCI cryo-probe (Ava600), 500 MHz Bruker AV III equipped with a DCH cryo-probe (Ava500), a 500 MHz Bruker AV IIIHD equipped with a Prodigy cryoprobe (Pro500) or a 400 MHz Bruker AV III equipped with BBFO+ probe (Ava400) at a constant temperature of 300 K. Chemical shifts are reported in parts per million. Coupling constants (*J*) are reported in hertz (Hz). Standard abbreviations indicating multiplicity were used as follows: m = multiplet, q = quartet, t = triplet, d = doublet, s = singlet, bs = broad singlet. All analysis was performed with MestReNova, Version 12.0.3. MS recorded using a Synapt Waters ESI Mass Spectrometer fitted with Nanomate® and analysed in MassLynx. All UV-Vis spectrometry recorded using Varian Cary® 50 UV-Vis Spectrophotometer.

2.4.2 Synthesis of Coordination Complexes



C1

Optimised synthesis in comparison to previous literature.[1]

L3 (44 mg, 0.13 mmol) was suspended in solution of **C2** (30 mg, 0.01 mmol) in H_2O (4 mL), MeCN (500 µL) and DMF (50 µL). A solution of aqueous $Co(NO_3)_2.6H_2O$ (2 mM, 275 µL) was added before heating at 90 °C for 48 h. The mixture was cooled before diluting with H_2O (10 mL) and filtering through celite. The filtrate was freeze dried and further purified by size exclusion chromatography (sephadex LH-20) yielding an orange solid. (32.1 mg, 0.01 mmol, 96%)

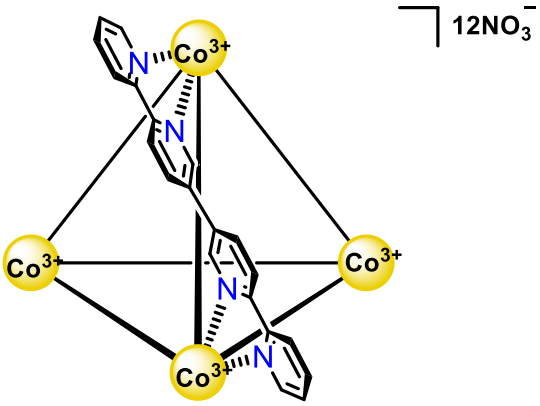
^{1}H NMR (600 MHz, D_2O) δ 8.55 (d, *J* = 8.4 Hz, 12H, H_d), 7.92 (d, *J* = 2.7 Hz, 12H, H_c), 7.78 (dd, *J* = 8.4, 1.8 Hz, 12H, H_e), 7.31 (d, *J* = 1.8 Hz, 12H, H_f), 6.99 (d, *J* = 7.0 Hz, 12H, H_a) 6.81 (dd, *J* = 7.0, 2.7 Hz, 12H, H_b). ^{13}C NMR (126 MHz, D_2O) δ 158.4, 156.5, 153.4, 148.5, 143.2, 137.3, 124.6, 114.7, 112.0, 110.0..

Extinction Coefficient

The spectrophotometer was zeroed with de-ionised water, and a solution of **C1** titrated into water, the linear data was fitted (Figure S1 7) and the Beer-Lambert Law applied to calculate the extinction coefficient.

ε = 150,000 M^{-1} cm^{-1} at 320 nm

ε = 140,000 M^{-1} cm^{-1} at 280 nm



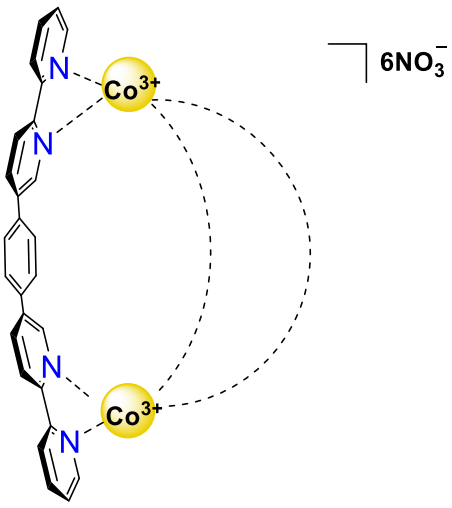
C2

Synthesised according to previous literature.[1]

Extinction Coefficient work

The spectrophotometer was zeroed with de-ionised water, and a solution of **C2** titrated into water, the linear data was fitted (Figure S1 6) and the Beer-Lambert Law applied to calculate the extinction coefficient.

ε = 160,000 M^{-1} cm^{-1} at 335 nm



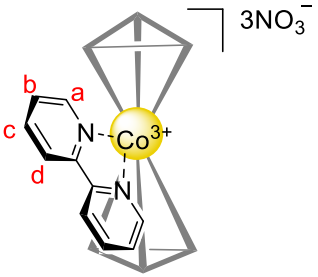
C3

Synthesised according to previous literature. [8]

Extinction Coefficient

The spectrophotometer was zeroed with de-ionised water, and a solution of **C3** titrated into water, the linear data was fitted (Figure S1 5) and the Beer-Lambert Law applied to calculate the extinction coefficient.

ε = 88,000 M^{-1} cm^{-1} at 350 nm



C4

$Co(NO_3)_2.6H_2O$ (125 mg, 0.43 mmol) and **L1** (200mg, 1.28 mmol) were weighed into a 10-20 mL microwave vial and suspended in H_2O: MeCN (4:1, 10mL). The vial was sealed before bubbling with nitrogen for 5 min, the reaction mixture was then heated at 90 °C for 18 h. The mixture was then cooled to rt before adding a solution of cerium ammonium nitrate (350 mg, 0.64 mmol) in MeCN (10 mL) and stirring vigorously for a further 15 min. MeCN (40 mL) was then used to precipitate out the product which was isolated via filtration and washed with diethyl ether (20 mL). Yielded beige solid (282 mg, 92 %)

^{1}H NMR (500 MHz, D_2O) δ 8.73 (d, *J* = 8.1, 1H, H_d), 8.45 (t, *J* = 7.9 Hz, 1H, H_c), 7.70 (t, *J* = 7.6 Hz, 1H, H_b), 7.38 (d, *J* = 7.7 Hz, 1H, H_a). ^{13}C NMR (126 MHz, D_2O) δ 155.83, 150.82, 143.87, 131.06, 127.04.^{1}H DOSY NMR (600 MHz, D_2O): D = 3.15e^{-10} m^2 s^{-1} ; calculated hydrodynamic radius = 7.5 Å. ESI-MS (m/z): calculated for 1+ 651.127; found 651.115 (1+).

Extinction Coefficient

The spectrophotometer was zeroed with de-ionised water, and a solution of **C4** titrated into water, the linear data was fitted (Figure S1 4) and the Beer-Lambert Law applied to calculate the extinction coefficient.

ε = 38,000 M^{-1} cm^{-1} at 305 nM

2.4.3 Direct Analysis of Binding

Mass Spectrometry

C1 (100 µM) and HSA (20 µM) in AA buffer (100 mM, 1 mL) were desalted using Vivaspin 500 centrifugal concentrators before being incubated for 1 h (37 °C, 200 rpm). The solutions were then filtered through Millipore Express® PLUS polyethersulfone membrane before cooling to 0 °C in an ice bath and running on a Synapt Waters ESI Mass Spectrometer fitted with Nanomate®.

Started the NanoMate® with a gas pressure 0.6 psi, voltage 1.5 V and the spectrometer with 30 V and 5 V for the sampling cone and extraction cone respectively. Data was recorded for a minimum of 2 min in the range of 100-2000 m/z.

Ultraviolet-Visible Spectrophotometry

Titration Method 1

A stock solution of **C1** (0.5 mM in 100 mM AA buffer pH 7.4) was titrated into a solution of HSA protein (1 mL, 10 µM in 100 mM AA buffer pH 7.4) at 1 µL increments, and the absorption spectrum was recorded between 200 and 400 nm. Using the absorbance at 320 nm of **C1**, a Beer-lambert plot was used to verify the extinction coefficient of the complex, seen in Figure S 2.1.

ε = 130,000 M^{-1} cm^{-1} at 320 nM

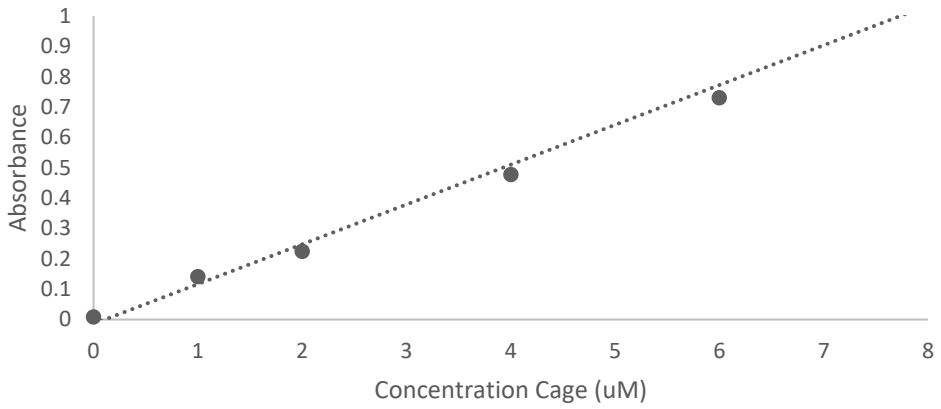


*Figure S 2.1. UV-Vis Spectrometry Titration Data showing **C1** titrated into solution of HSA Protein (10 µM) in ammonium acetate pH 7.4.*

Titration Method 2

A stock solution of HSA protein (0.5 mM in 100 mM AA buffer pH 7.4) was titrated into a solution of **C1** Cage (1 mL, 5 µM in 100 mM AA buffer pH 7.4) in 1 µL increments, and the absorption spectrum was recorded between 200 and 400 nm. Using the absorbance at 280 nm of the HSA protein, a Beer-lambert plot was used to verify the extinction coefficient of the protein.

ε = 30,000 M^{-1} cm^{-1} at 280 nM

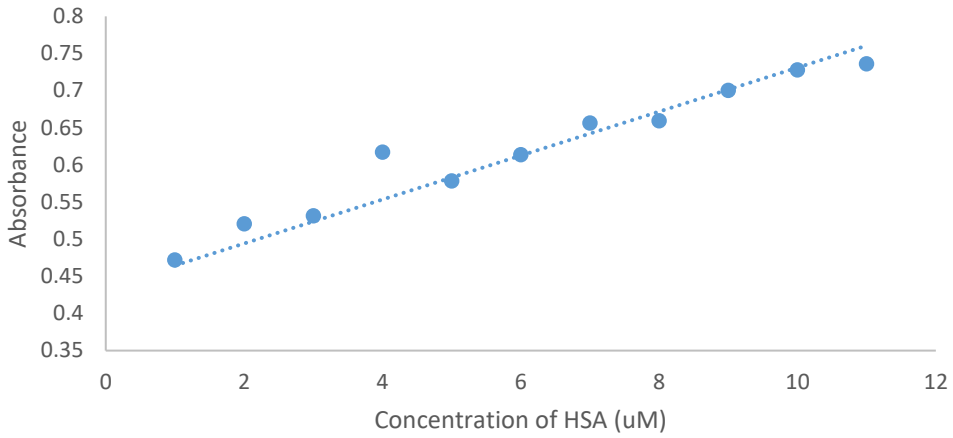


*Figure S 2.2. UV-Vis Spectrometry Titration Data showing HSA Protein titrated into solution of **C1** (5 µM) in ammonium acetate pH 7.4.*

Saturation Transfer Difference NMR

General Sample Preparation

A stock solution of HSA protein (1 Eq, 3 mL, 50 µM in 0.1 M PBS Buffer pH 7.4 90% D_2O; 10% DI H_2O) was prepared and stored at 4 °C. Each complex (30 Eq, 1.5 mM) was added to an aliquot of protein stock solution (0.5 mL) and incubated at 37 °C, 200 rpm for 1 h.

Samples prepared of all four complexes; the **C4**, **C3**, **C2** and **C1**.

Data Collection

STD spectra were collected by using the standard pseudo-2 D pulse sequence, STDDIFFESGP.3. STD NMR spectra were recorded on a Bruker Avance 600 MHz spectrometer equipped with a TCI cryoprobe and a relaxation delay of 4 s. Data processing was performed on a PC with Bruker Topspin v4.1.3 software. The total saturation time was set to 3.9 s. The proton carrier was placed at -4 and 40 ppm for on- and off-resonance saturation, respectively. The on- and off-resonance experiments were recorded in an interleaved fashion to avoid any experimental inconsistencies and minimize the effect of any radiofrequency induced by sample temperature changes.

Amplification factors were calculated from the off and on-resonance spectra by applying the Mayer and Meyer protocol.[38]

2.4.5 Indirect Analysis of Binding

Rapid Equillibrium Dialysis (RED)

Control Experiments

Determining Membrane and Incubation Time

A solution of **C1** (500 µM, 100 µL) was pipetted into the red chambers of two RED tubes and followed by the addition of dialysis buffer (100 mM AA Buffer, 300 µL) into the corresponding white chambers. The RED tubes were covered with adhesive tape, to prevent evaporation before sealing entire tray with three layers of parafilm. The tray was then incubated at 37 °C, 200 rpm for 24 h before being removed from incubator and allowed to cool before recording the concentration of **C1** in both the red and white chambers, using UV-Vis absorption. The process was repeated with longer incubation times until the system appeared to have equilibrated.

*Table S 2.1. Optimisation of **C1** equilibrium in RED dialysis insert in the absence of protein.*

MWCO of Membrane (kDa)	**Incubation Time (h)**	**Concentration Ratio (red/white)**
8	24	3.0
12	24	1.3
12	30	1.2
12	48	1.0

Initial Binding Experiments

Method 1

Ran in duplicates; loaded HSA (100 µL, 50 µM, 100 mM AA buffer pH 7.4) into red chamber and loaded **C1** (300 µL, 250 µM, 100 mM AA buffer pH 7.4) into white chamber. The RED tubes were covered with adhesive tape, to prevent evaporation before sealing entire tray with three layers of parafilm. The tray was then incubated at 37 °C, 200 rpm for 48 h.. The concentration of **C1** in each compartment was then recorded using UV-Vis spectroscopy and the ratio between the two compartments calculated.

Method 2

Ran in duplicates; loaded a solution of HSA (50 µM) and cage (100 µL, 750 µM , 100 mM AA buffer pH 7.4) into the red chamber and loaded buffer (300 µL, 100 mM AA Buffer pH 7.4) into the white chamber. The RED tubes were covered with adhesive tape, to prevent evaporation before sealing entire tray with three layers of parafilm. The tray was then incubated at 37 °C, 200 rpm for 48 h.The

concentration of **C1** in each compartment was then recorded using UV-Vis spectroscopy and the ratio between the two compartments calculated.

Table S 2.2. Concentration ratio between red and white chambers of RED dialysis inserts following control binding methods 1 and 2.

Method	Concentration Ratio (red/white)
1	1.12
2	1.13

Binding Experiments

Following the protocol proposed by Shave and collegues.[26]

Run in duplicates; loaded HSA (100 µL, 80 µM in 100 mM AA buffer at pH 7.4) into the red chamber and complex (300 µL, 66.7 µM in 100 mM AA buffer at pH 7.4) into the white chamber. Sealed tubes with adhesive tape and covered tray with three layers of parafilm, before incubating at 37 °C, 200 rpm for 48 h. The concentration of **C1** each compartment was then recorded using UV-Vis and the ratio between the two compartments calculated.

Using the excel workbook written by Shave and colleagues a K_d can be calculated for each system using the values detailed below.

Pc = ratio of the compound concentrations in the red and white chambers in the control experiment in the absence of the protein

Pt = is the ratio of complex between the compartments (red/white) in the presence of protein

t0 = concentration of protein loaded into red compartment

l0 = ligand concentration assuming equal distribution across the chambers

Red/White vol = total volume of each compartment

Table S 2.3. Parameters for determining dissociation constants via *RED dialysis for the four complexes*

Complex	Pc	Pt	t0 (µM)	l0 (µM)	Red vol (µL)	White vol (µL)
C1	1.00	1.61	50	80	100	300
C2	0.99	2.17	50	80	100	300
C3	1.60	2.85	50	80	100	300
C4	1.01	1.31	50	80	100	300

Precipitation Binding Method

Control Experiment – General

Complex (2.5 x 10^{-5} mmol, 50 µM) was dissolved in buffer (0.5 mL, 100 mM AA) and then incubated at 37 °C, 200 rpm for 24 h. The solution was then allowed to cool to rt before adding an aliquot of trichloroacetic acid (15%, 50 µL). The resulting mixture was vortexed and then centrifuged (12,000 rpm, 1 minute) and the concentration of the supernatant determined using UV-Vis spectroscopy.

Procedure completed for the four complexes: **C1**, **C2**, **C3** and **C4**.

Binding Experiment Protocol - General

Complex (2.5 x 10^{-5} mmol, 50 µM) and HSA (0-100 µM) were dissolved in buffer (0.5 mL, 100 mM AA) and then incubated at 37 °C, 200 rpm for 24 h. The solution was then allowed to cool to rt before adding an aliquot of trichloroacetic acid (15%, 50 µL). The resulting mixture was vortexed and then centrifuged (12,000 rpm, 1 minute) and the concentration of the supernatant determined using UV-Vis spectroscopy. This protocol was repeated at varying concentrations of HSA protein.

Procedure completed for the four complexes: **C1**, **C2**, **C3** and **C4**.

K_d was estimated by graphically deriving the amount of HSA at which 50% of the complex was bound. Where the data did not show evidence of reaching 50 % bound, extrapolation utilising a linear plot was performed.

Centrifugal Ultrafiltration – Binding Experiment

General Procedure for Centrifugal Ultrafiltration

Complex (2.5 x 10^{-5} mmol, 50 µM) and HSA (2.5 x 10^{-5} mmol, 50 µM) were dissolved in buffer (0.5 mL, 100 mM AA) and then incubated at 37 °C, 200 rpm for 24 h. An aliquot (100 µL) of the solution was loaded into a Vivaspin® 500 centrifugal concentrator with a MWCO of 10 kDa. Followed by centrifugation (12,000 rpm, 3 x 7 mins), a further 100 µL of AA (100 mM) was loaded into the centrifugal concentrator following each centrifugation. UV-Vis spectroscopy used to determine complex concentration.

Procedure completed for the four complexes: **C1**, **C2**, **C3** and **C4** and in the absence of protein to obtain control data.

Data Processing

Using the known initial concentrations, the percentage of complex retained above the MWCO membrane was calculated and compared against the control as tabulated below.

Table S 2.4. Centrifugal Ultrafiltration Data showing percentage of complex retained in Insert in the process and absence of protein

Complex	**Percentage Retained in Insert %**	
	Absence of Protein	Presence of Protein
C1	3.1	72.1
C2	8.1	71.3
C3	4.4	49.5
C4	9.2	45.2

2.4.6 Gel Electrophoresis

Gel Electrophoresis was carried out in Mini Gel Tank using precast NuPAGETM 4-12% Bis-Tris Gel (1 mm x 10 &15 well) (Invitrogen by ThermoFisher Scientific). NuPAGETM MES SDS (sodium dodecyl sulphate) Running Buffer (20X) was diluted in distilled water to 1 x MES SDS to be used as the running buffer.

Samples were prepared by mixing equal amounts of sample and 2 x SDS buffer. Before loading, all samples were placed on a thermoplate for 10 minutes >80 °C. A 10-100 kDa MW standard marker was used (Fisher BioReagentsTM EZ-RunTM Prestained Rec Protein Ladder) as received. 10 µL of marker and sample was loaded onto the gel and it ran for 30 minutes at 200 V (Bio-Rad PowerPacTM Basic).

Component	**Ingredients**	**Volume**
2 x SDS	Tris pH 6.8 (0.5 M)	2.5 mL
	glycerol	2 mL
	10% w/v SDS	2 mL
	β-mercaptoethanol	1 mL
	0.1% w/v bromophenol blue	0.5 mL

2.4.7 Isothermal Titration Calorimetry

Sample Preparation

Stock solutions (4 mM) of each complex in 0.1 M PBS buffer were filtered through 0.45 µM Millex syringe filters and concentration verified using UV-Vis absorption, before storing at room

temperature. A solution of HSA (10 mg/mL in 0.1 M PBS buffer pH 7.4) was made, syringe filtered and stored at 4 °C, to be diluted when required.

Experimental Set-up

All ITC experiments were carried out at 25 °C on a MicroCal Auto iTC200 in 0.1 M PBS buffer at pH 7.4. For each experiment the initial delay was set to 180 s, stirring speed 750 rpm, duration of injection 1 s, spacing between injections 180 s, initial injection volume 0.5 µL and filter period 5 s. For further experiment details specific to complexes see Table S 2.5 below. Control experiments were run for each complex; buffer into buffer, buffer into protein and complex into buffer. Minimal heats of dilution were observed here, validating the complex binding data. Data was processed using the MicroCal PEAQ-ITC Analysis software, relying on the Wiseman isotherm. **C1** data summarised in Table S 2.6 and Figure S 2.3. **C4** data summarised in Table S 2.7 and Figure S 2.4.

Table S 2.5. Isothermal Titration Calorimetry Experiment Parameters.

Cell	Cell Concentration (mM)	Syringe	Syringe Concentration (mM)	Reference Power (uCal/sec)	Total Number of Injections	Injection volume (uL)
HSA	0.05	**C1**	0.50	5	17	2
Buffer	1.00	**C1**	0.50	5	17	2
HSA	4.00	**C4**	4.00	3	15	2.5
Buffer	1.00	**C4**	4.00	3	15	2.5
Buffer	1.00	Buffer	1.00	3	15	2.5

*Table S 2.6. **C1** interacting with HAS. ITC Data calculated using MicroCal PEAQ-ITC Analysis software, relying on the Wiseman isotherm*

Experiment	[Syr] (M)	[Cell] (M)	N (sites)	N Error (sites)	K_d (M)	K_d Error (M)	ΔH (kcal/mol)	ΔH Error (kcal/mol)	ΔG (kcal/mol)
1	4.00E-03	1.00E-04	5.09	1.52	1.94E-04	1.88E-04	-0.564	0.262	-5.06
2	4.00E-03	1.00E-04	5.32	1.11	1.27E-04	8.21E-05	-0.42	0.124	-5.32
3	4.00E-03	1.00E-04	5.32	0.707	1.19E-04	4.96E-05	-0.455	0.0860	-5.35
Average	4.00E-03	1.00E-04	5.24	1.11	1.47E-04	1.07E-04	-0.480	0.1573	-5.24

*Table S 2.7. **C4** complex interacting with HAS. ITC Data calculated using MicroCal PEAQ-ITC Analysis software, relying on the Wiseman isotherm*

Experiment	[Syr] (M)	[Cell] (M)	N (sites)		N Error (sites)	K_d (M)	K_d Error (M)	ΔH (kcal/mol)	ΔH Error (kcal/mol)	ΔG (kcal/mol)	-TΔS (kcal/mol)
1	5.0E-04	5.0E-05	7.0E-01		1.4E-02	1.1E-06	2.9E-07	-2.4E+00	1.1E-01	-8.1E+00	-5.7E+00
2	5.0E-04	5.0E-05	6.5E-01		3.0E-02	1.8E-06	8.2E-07	-2.8E+00	2.5E-01	-7.8E+00	-5.0E+00
3	5.0E-04	5.0E-05	6.2E-01		3.0E-02	2.0E-06	9.2E-07	-2.9E+00	2.7E-01	-7.8E+00	-4.9E+00
AVERAGE	5.0E-04	5.0E-05	6.5E-01		2.5E-02	1.7E-06	6.7E-07	-2.7E+00	2.1E-01	-7.9E+00	-5.2E+00

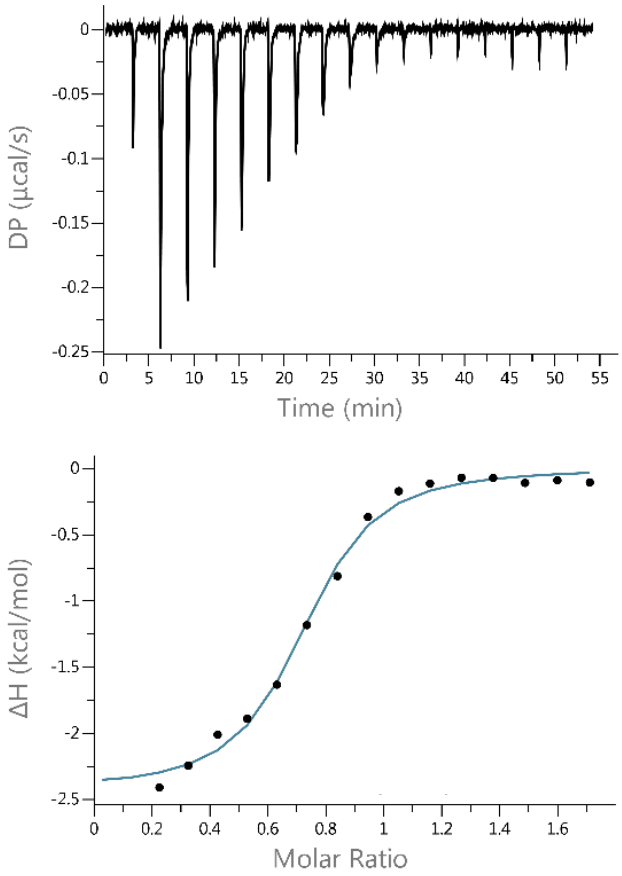


*Figure S 2.3. **C1** ITC Isotherm produced using MicroCal PEAQ-ITC Analysis software, relying on the Wiseman isotherm.*

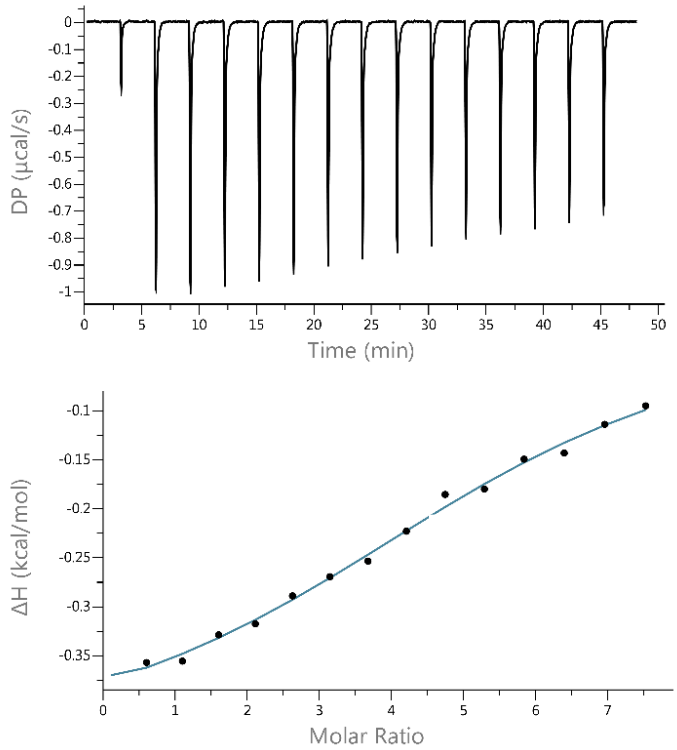


Figure S 2.4. ***C4*** *complex ITC Isotherm produced using MicroCal PEAQ-ITC Analysis software, relying on the Wiseman isotherm.*

2.4.8 Computational Modelling – AutoDock Vina

File Preparation

For all computational modelling purposes a crystal structure of HSA (PDB 1BM0) was obtained a 1999 study by Kobayashi *et al.* used in its monomeric form.[39] AutoDock Tools (downloaded as part of the MGLTools software package developed by the Sanner lab at the Scripps Research Institute) was used to load the protein file pdb.[36] Using AutoDock Tools residual water molecules were removed, polar hydrogens were added and Kollman charges were applied to the protein and the file saved as a pdbqt.

An obtained crystal structure of **C1** was utilised for all computational modelling purposes, see Chapter 3 'X-ray Crystallography'. This file was loaded into AutoDock Tools whereby any residual water molecules were removed before limiting the torsion of the cage prior to the docking experiment. Upon selecting the cage structure as the ligand the torsion tree was limited to 0 rotational bonds and thus maintaining the structure of the cage upon binding before saving as a pdbqt.

Configuration Details

Downloaded AutoDock Vina from the Scripps Research Institute and utilised the previously prepared receptor (HSA) file and ligand (cage) file along with the following parameters to run the program. Exporting the output data into both a text file and a structural pdbqt file.

Parameters:

Center Coordinates (x, y, z) = (29.535, 31.826, 23.500)

Grid box size (x, y, z) = 60, 60, 58

Energy range = 4

Exhaustiveness = 10

Output Table:

Mode	Affinity (kcal/mol)	Distance from rmsd l.b	Distance from rmsd u.b
1	-13.1	0	0
2	-11.0	32.033	39.073
3	-11.0	32.027	39.751
4	-10.9	31.575	38.999
5	-10.9	31.575	39.433
6	-10.6	33.538	41.458
7	-10.6	33.543	40.554
8	-10.1	33.516	41.168
9	-10.1	33.518	40.355

2.5 References

1 B. P. Burke, W. Grantham, M. J. Burke, G. S. Nichol, D. Roberts, I. Renard, R. Hargreaves, C. Cawthorne, S. J. Archibald and P. J. Lusby, *J. Am. Chem. Soc.*, 2018, **140**, 16877–16881.

2 A. Kalra, E. Yetiskul, C. J. Wehrle and F. Tuma, *Physiology*, StatPearls Publishing, 2022.

3 A. M. Merlot, D. S. Kalinowski and D. R. Richardson, *Front. Physiol.*, 2014, **5**, 298.

4 J. Ghuman, P. A. Zunszain, I. Petitpas, A. A. Bhattacharya, M. Otagiri and S. Curry, *J. Mol. Biol.*, 2005, **353**, 38–52.

5 J. Zhu, C. J. E. Haynes, M. Kieffer, J. L. Greenfield, R. D. Greenhalgh, J. R. Nitschke and U. F. Keyser, *J. Am. Chem. Soc.*, 2019, **141**, 11358–11362.

6 A. Oleksi, A. G. Blanco, R. Boer, I. Usón, J. Aymamí, A. Rodger, M. J. Hannon, M. Coll, P. M. J Hannon, A. Oleksi, A. G. Blanco, R. Boer, P. M. Coll, I. Usón, J. Aymamí and P. A. Rodger, *Angew. Chem. Int. Ed.*, 2006, **45**, 1227–1231.

7 L. E. H. Paul, B. Therrien and J. Furrer, *Org. Biomol. Chem*, 2015, **13**, 953.

8 M. J. Burke, G. S. Nichol and P. J. Lusby, *J. Am. Chem. Soc.*, 2016, **138**, 9308–9315.

9 D. Preston, J. E. Barnsley, K. C. Gordon and J. D. Crowley, *J. Am. Chem. Soc.*, 2016, **138**, 10578–10585.

10 K. Ishii, M. Noda and S. Uchiyama, *Biophys. Physicobiology*, 2016, **13**, 87.

11 H. Hernández and C. V. Robinson, *Nat. Protoc.*, 2007, **2**, 715–726.

12 J. Gault, I. Liko, M. Landreh, D. Shutin, J. Reddy Bolla, D. Jefferies, M. Agasid, H.-Y. Yen, M. J. G W Ladds, D. P. Lane, S. Khalid, C. Mullen, P. M. Remes, R. Huguet, G. McAlister, M. Goodwin, R. Viner, J. E. Syka and C. V Robinson, *Nat. Methods*, 2020, **17**, 505–508.

13 D. P. Donnelly, C. M. Rawlins, C. J. Dehart, L. Fornelli, L. F. Schachner, Z. Lin, J. L. Lippens, K. C. Aluri, R. Sarin, B. Chen, C. Lantz, W. Jung, K. R. Johnson, A. Koller, J. J. Wolff, I. D. G. Campuzano, J. R. Auclair, A. R. Ivanov, J. P. Whitelegge, L. Paša-Tolić, J. Chamot-Rooke, P. O. Danis, L. M. Smith, Y. O. Tsybin, J. A. Loo, Y. Ge, N. L. Kelleher and J. N. Agar, *Nat. Methods*, 2019, **16**, 587–594.

14 H. Metwally, R. G. Mcallister and L. Konermann, *Anal. Chem*, 2015, **87**, 2434–2442.

15 C. Hao, R. E. March, T. R. Croley, J. C. Smith and S. P. Rafferty, *J. Mass Spectrom.*, 2001, **36**, 79–96.

16 E. Barile and M. Pellecchia, *Chem. Rev.*, 2014, **114**, 4749.

17 R. Page, W. Peti, I. A. Wilson, R. C. Stevens and K. Wüthrich, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 1901–1905.

18 J. D. Eaton and M. P. Williamson, *Biosci. Rep.*, 2017, **37**, 1–10.

19 A. Viegas, J. Ao Manso, F. L. Nobrega and E. J. Cabrita, *J. Chem. Educ*, 2011, **88**, 990–994.

20 C. D. Kanakis, P. A. Tarantilis, M. G. Polissiou, S. Diamantoglou and H. A. Tajmir-Riahi, *J. Mol. Struct.*, 2006, **798**, 69–74.

21 J. R. Roberts, J. Xiao, B. Schliesman, D. J. Parsons and C. F. Shaw, *Inorg. Chem*, 1996, **35**, 424–433.

22 H. S. Penefsky, *Methods Enzymol.*, 1979, **56**, 527–530.

23 C.-M. Toma, S. Imre, C.-E. Vari, D.-L. Muntean and A. Tero-Vescan, *Processes*, 2021, **9**, 382.

24 N. J. Waters, R. Jones, G. Williams and B. Sohal, *J. Pharm. Sci.*, 2008, **97**, 4586–4595.

25 M. Palencia and B. L. Rivas, *J. Memb. Sci.*, 2011, **372**, 355–365.

26 S. Shave, N. T. Pham, C. B. Śmieja and M. Auer, *Methods Protoc.*, 2020, **3**, 1–20.

27 S. van Liempd, D. Morrison, L. Sysmans, P. Nelis and R. Mortishire-Smith, *J. Lab. Autom.*, 2011, **16**, 56–67.

28 V. D'Aloisio and C. Coxon, University of Edinburgh, 2022.

29 Y. Y. Chen, S. Y. Lin, Y. Y. Yeh, H. H. Hsiao, C. Y. Wu, S. T. Chen and A. H. J. Wang, *Electrophoresis*, 2005, **26**, 2117–2127.

30 J. Zhu, Z. Yan, F. Bošković, C. J. E. Haynes, M. Kieffer, J. L. Greenfield, J. Wang, J. R. Nitschke and U. F. Keyser, *Chem. Sci.*, 2021, **12**, 14564–14569.

31 G. K. Krishnamoorthy, P. Alluvada, S. Hameed Mohammed Sherieff, T. Kwa and J. Krishnamoorthy, *Biochem. Biophys. Reports*, 2020, **21**, 100712.

32 T. Wiseman, S. Williston, J. F. Brandts and L. N. Lin, *Anal. Biochem.*, 1989, **179**, 131–137.

33 M. W. Freyer and E. A. Lewis, *Methods Cell Biol.*, 2008, **84**, 79–113.

34 H. Olsen, A. Andersen, A. Nordbø, U. E. Kongsgaard and O. P. Børmer, *BMC Clin. Pharmacol.*, 2004, **4**, 1–9.

35 D. R. Houston and M. D. Walkinshaw, *J. Chem. Inf. Model.*, 2013, **53**, 384–390.

36 O. Trott and A. J. Olson, *J. Comput. Chem.*, 2009, **31**, 455–461.

37 N. Nateghian, N. Goodarzi, M. Amini, F. Atyabi, M. R. Khorramizadeh and R. Dinarvand, *Chem. Biol. Drug Des.*, 2016, **87**, 69–82.

38 M. Mayer and B. Meyer, *Angew. Chem. Int. Ed.*, 1999, **38**, 5907–5908.

39 S. Sugio, A. Kashima, S. Mochizuki, M. Noda and K. Kobayashi, *Protein Eng. Des. Sel.*, 1999, **12**, 439–446.

Chapter 3 : Towards the Targeted Delivery of M_4L_6 Cages

Chapter 3 : Towards the Targeted Delivery of M_4L_6 Cages

3.1 Introduction

Previous research by the Lusby group had identified a kinetically inert Co^{III} based M_4L_6 tetrahedral capable of strongly encapsulating the biomedical radioisotope $[^{99m}Tc]TcO_4^-$ and altering its bioaccumulation, shown in Figure 3.1. Upon encapsulation within **C1** the distribution of TcO_4^- changes from the thyroid and the stomach to the liver.[1] As discussed in Chapter 2 this was hypothesised to be a result of interactions between the cage and proteins present in the liver, preventing further circulation. The work evidenced the potential of the cage as a delivery vessel and its ability to alter the biodistribution of $[^{99m}Tc]TcO_4^-$, although no attempt was made to control and target the bioaccumulation.

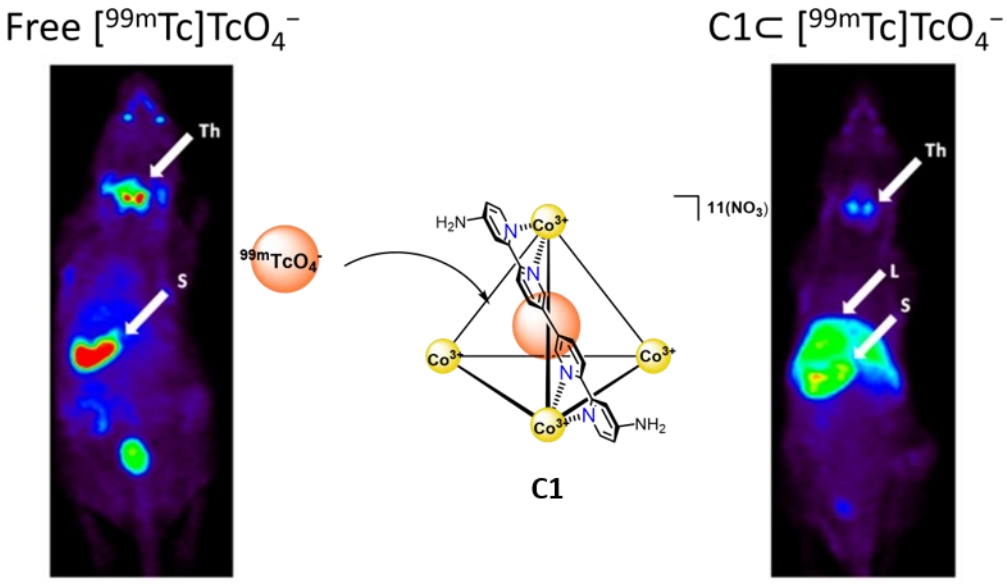


Figure 3.1. Comparison of $[^{99m}Tc]TcO_4^-$ uptake in naïve mice (left) vs $[^{99m}Tc][TcO_4\subset C\text{-}1]^{11+}$ (right). Free $[^{99m}Tc]TcO_4^-$ was injected into a naïve anaesthetised animal 40 min prior to a 20 min SPECT acquisition. Caged $[^{99m}Tc]TcO_4^-$ was injected into a naïve anaesthetised animal 20 min prior to a 100 min SPECT acquisition. Figure adapted from reference.[1]

Coordination cages as biomedical vessels is an emerging field yet there remains very few examples of successful targeted delivery of an encapsulated guest. One key example of targeted delivery of a metallocycle via bioconjugation was reported in 2018 by Stang and colleagues.[2] They assembled a multifunctional metallocage composed of porphyrin rings capable of photodynamic therapy and chemotherapeutic platinum metal centres (Figure 3.2A). In order to ensure the *in vivo* stability of the host–guest complex and enhance its solubility in aqueous media, the coordination complex was

embedded into polymeric nanoparticle by assembly with the RGD-PEG-b-PEBP ligand (Figure 3.2B). This formed large micelle-like nanoparticles which featured the externally positioned RGD peptides as active targeting species, because the peptide features a cyclo(Arg-Gly-Asp-D-Phe-Lys) (cRGDfK) moiety which specifically binds the αvβ3 integrin receptor, a key marker of tumor vasculature and thus targets the nanoparticle towards cancerous tissue. The Stang group have since built on this work by assembling a library of metallocycle containing nanoparticles using amphiphilic copolymers whereby they exemplify the potential of non-covalent bioconjugation. [3,4]

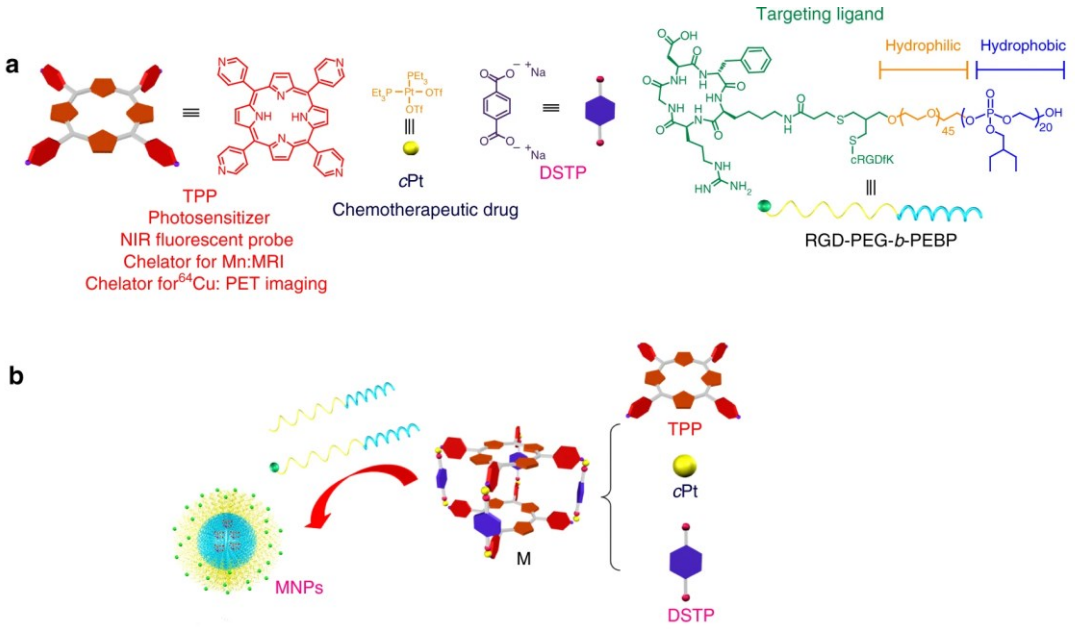


Figure 3.2. Schematic diagrams of the MNPs serving as a multifunctional theranostic platform as developed by Stang and co-workers, figure adapted with reference.[2]

Casini and co-workers have spent the last couple of years attempting to control the biodistribution of a series of Pd_2L_4 lantern cages by exo-functionalising with targeting peptides.[5] Most recently, they conjugated peptide PepH3 to their ligand before assembling the cage (Figure 3.3), the peptide represented a small fragment of Dengue virus type-2 protein (DEN2C) and exhibited the ability to translocate endothelial cells.[6] Subsequently they argued that they had assembled a coordination cage that was able to cross the blood brain barrier. Whilst there are limitations working with potentially labile Pd_2L_4 systems *in vivo*, which are discussed in detail in Chapter 4. This is nominally the first example of covalently bioconjugating a targeting group onto the outside of a coordination cage and altering the biodistribution of the cage and its encapsulated guest; in this case $[^{99m}Tc]TcO_4^-$.

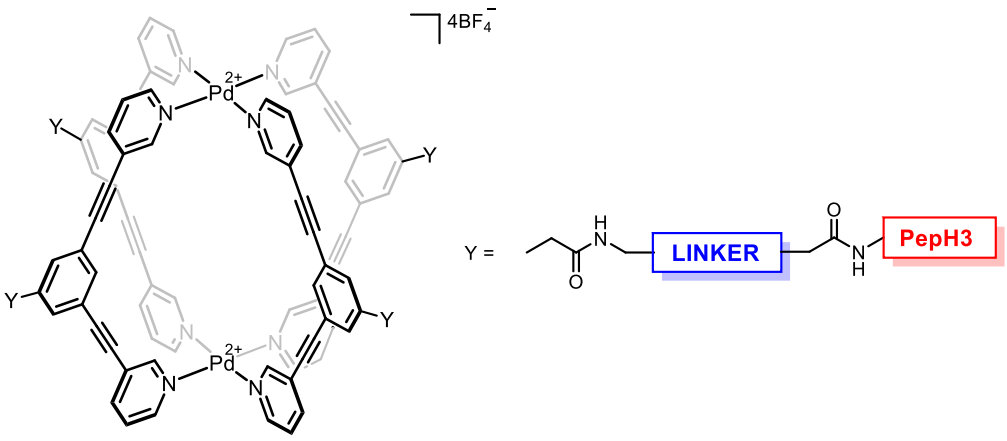


Figure 3.3. Bioconjugation of Pd_2L_4 cage with PepH3 to allow cage blood brain barrier penetration as described by Casini and co-workers.[6]

Targeted delivery is essential in both modern therapeutics and biomedical imaging/diagnostics, as controlling the biodistribution of species increases their efficacy, whilst reducing cytotoxicity against healthy cells. Stang and colleagues work served to highlight the potential of non-covalent bioconjugation via supramolecular assembly to associate targeting species with the coordination cages in comparison to the more traditional covalent route achieved by the Casini group.

3.1.1 Radiolabelling in Theranostics

Nuclear medicine concerns the use of radioactive materials *in vivo* for either diagnostics or therapeutics. Single photon emission computed tomography (SPECT) and positron emission tomography (PET) are the two most popular imaging techniques used to track the bio-distribution of radioisotopes.[7] When considering the suitability of a radioisotopes for medicinal purposes, both the half-life and ease of labelling must be considered. For example, ^{99m}Tc is the most widely used radioisotope in medicine and exhibits a half-life of approximately 6 hours providing adequate time for the radiolabelling process.[8] The synthetic incorporation of the radioisotope into a specific ligand (radiolabelling) must be relatively simple to minimise exposure for the handler along with wastage of half-life time.

^{99m}Tc is produced from a Mo generator as a solution of $[^{99m}Tc]TcO_4^-$ in saline, if used directly the free anion accumulates in the thyroid and the stomach.[8] Figure 3 shows some of the most common radiopharmaceuticals of which $[^{99m}Tc]TcO_4^-$ is incorporated, including ^{99m}TcMAG3, which serves to monitor renal function (Figure 3.4A).[8] Technetium exametazime, sold under the name Ceretec (Figure 3.4B), is formed by radiolabelling ^{99m}Tc with the isomeric forms of exametazime. It tracks cerebral blood flows to aid in the diagnosis of strokes and vascular malfunctions.[9] Lastly, Figure 3.4C shows Myoview commonly used in myocardial perfusions to investigate whether there has been damage to

the blood flow around the heart.[10] Whilst the synthetic process for the ligand might be longer, the formulation that involves incorporation of ^{99m}Tc into the radiopharmaceutical typically takes no more than a couple of hours.[8]

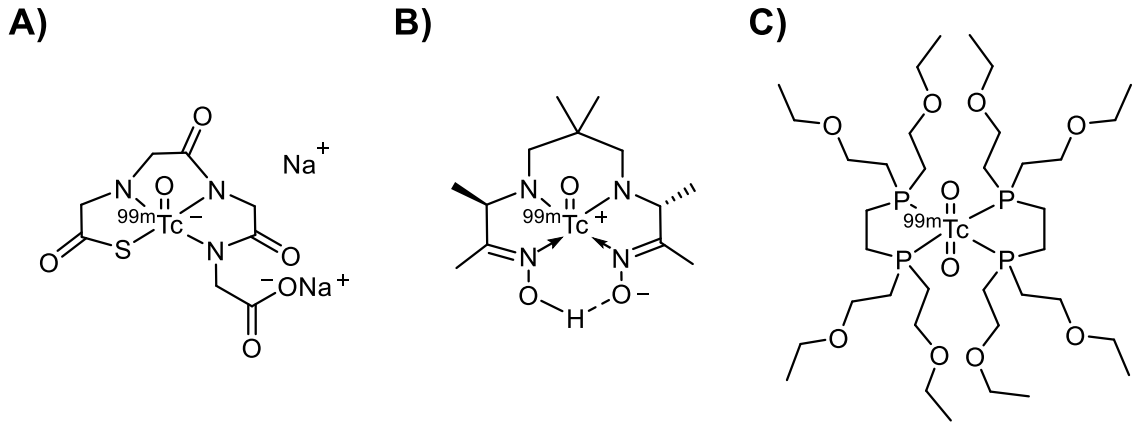


Figure 3.4. Examples of some FDA approved ^{99m}Tc containing radiopharmaceuticals. A) ^{99m}TcMAG$_3$ for renal function imaging, B) ^{99m}Tc-d,l-HM-PAO (Ceretec) for imaging of cerebral blood flow and C) ^{99m}Tc-tetrofosmin (Myoview) for imaging blood flow around the heart.[8]

The encapsulation of [^{99m}Tc]TcO_4^- within **C1** evidenced the potential of coordination cages as a fast and simple radiolabelling method.[1] Formulation of the ^{99m}Tc radiolabelled **C1** was achieved by adding a solution of the cage into a fresh saline solution of [^{99m}Tc]TcO_4^- direct from the generator and incubating for 5 min. This process is considerably faster and less complex than for the described FDA approved radiopharmaceuticals, such as Ceretec and Myoview, currently used in clinic.

3.1.2 Aims

This work aims to alter and control the biodistribution of the M_4L_6 cages previously studied by the Lusby group by altering their exterior properties. External functionalisation of the cobalt cage with a variety of targeting ligands or antibodies would generate a library of cages capable of being instantaneously radio-labelled and controlling the biodistribution of biomedically relevant guests such as; ^{99m}Tc.

The focus will first be directed towards covalently altering the ligand and investigating the moieties of the ligand which are essential for the assembly of a stable cage. The work then aims to make a comparison between pre- and post-assembly modifications when considering routes to the successful bioconjugation of coordination cages. The last section concerns the synthesis of M_4L_6 cages towards targeted delivery using the possibility of non-covalent bioconjugation routes in comparison to the more traditional covalent route.

It is imperative that the solubility and stability of the novel externally functionalised cages are analogous to or better than the previously synthesised **C1**. Therefore this work aims to probe the suitability of the resultant cages by testing their tolerance to biological conditions. Looking at both the stability of the empty cages in the presence of bioabundant species such as salts and reductants as well as the stability of the host-guest complexes.

The work will also probe the host-guest scope of the cages to determine their delivery capabilities within theranostics. Investigating both commonly used radioisotopic anions such as $[^{99m}Tc]TcO_4^-$ for SPECT imaging alongside novel guests such as $[^{124}I]IO_4^-$. The work also aims to investigate methods into preventing guest leakage from the cage systems following encapsulation, with the idea of preventing premature release *in vivo*.

3.2 Results and Discussion

3.2.1 Overall Strategy and Design

The targeted delivery of therapeutics is the method by which a drug moiety is designed to accumulate specifically in the area in which it is required. For example, when cancerous cells replicate and create micro environments that display distinctive characteristics in comparison to healthy tissue, such as over expressing certain receptors and proteins.[11] Subsequently, therapeutics can be designed that containing moieties which complement the over expressed, cancer associated receptor and proteins, increasing the specificity of the drug.

When considering the bioconjugation of a targeting group to the therapeutic of interest, this can be achieved by either covalently tethering the species or non-covalently binding it. The Thordarson group recently published a review comparing the different methods of conjugating peptides to therapeutics and the various benefits that attached peptides can provide.[12] This includes a discussion on the recent abundance of research into using transformable peptide monomers (TPMs) which self-assemble into nanoparticles, improving the stability and biological half-life of the individual monomers.[13] This is analogous to the aforementioned research by Stang and co-workers, where they incorporated a metallocage into a self-assembled nanoparticle structure using a peptide containing ligand.[2]

In comparison to this non-covalent assembly approach there has been significant research within the area of targeted drug delivery into covalently tethering peptides onto the surfaces of nanoparticles; namely gold nanoparticles.[14] The bioconjugation of gold nanoparticles is traditionally achieved via thiol bonds formed on the surface of the structure, however this can be unpredictable and lead to varying levels of conjugation between nanoparticles.[14]

The smaller and more defined nature of coordination cages in comparison to nanoparticles allows for well-defined binding of the targeting groups at controlled sites. Following the previous research in the Lusby group with **C1**, this work aims to focus on altering the external functionality of the cage in order to see how it effects the biodistribution. In Figure 3.5, two bio-conjugation strategies are outlined, firstly covalent bonding the biomedically relevant moiety (R shown in purple) to the outside of the M_4L_6 cage, and secondly non-covalent bioconjugation via anion metathesis whereby the biomedically relevant group (R^- shown in pink) is bound to the cage via electrostatic interactions.

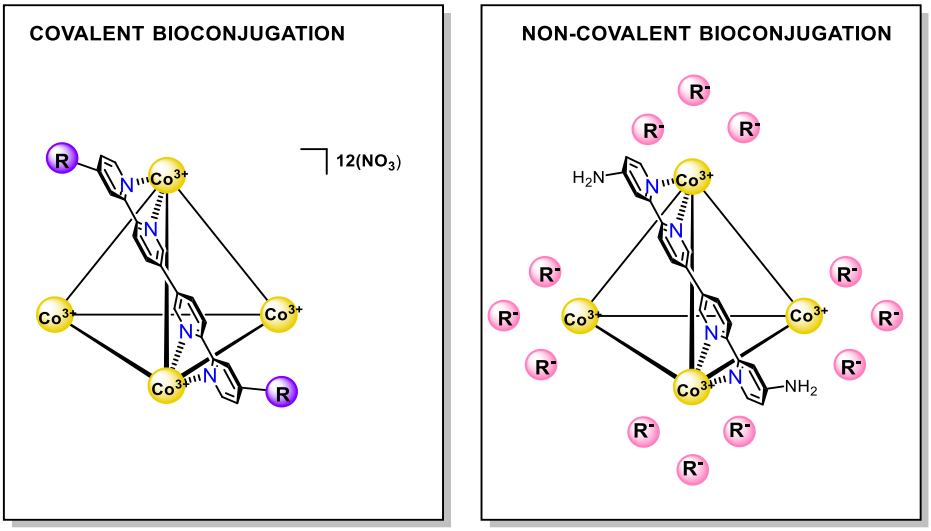


Figure 3.5. Bio-conjugation strategies: Covalent bioconjugation, covalently bonding biomedically relevant species (R shown in purple) to the cage. Non-covalent bioconjugation, non-covalent interactions associate the biomedically relevant species in its anionic form (R^- shown in pink) with the cage.

When considering covalent bioconjugation there are two approaches, outlined in work by Casini and colleagues shown in Figure 3.6.[15] In approach 1, the targeting group, shown as a peptide, is directly bound onto the outside of the pre-assembled cage, defined as a post-assembly modification. In approach 2, the targeting group is first attached to the ligand before then self-assembling the functionalised ligand into a cage. Whilst, the research of Casini and colleagues opted for approach 2, both methods are attempted in this study in an effort to optimise the synthetic process.

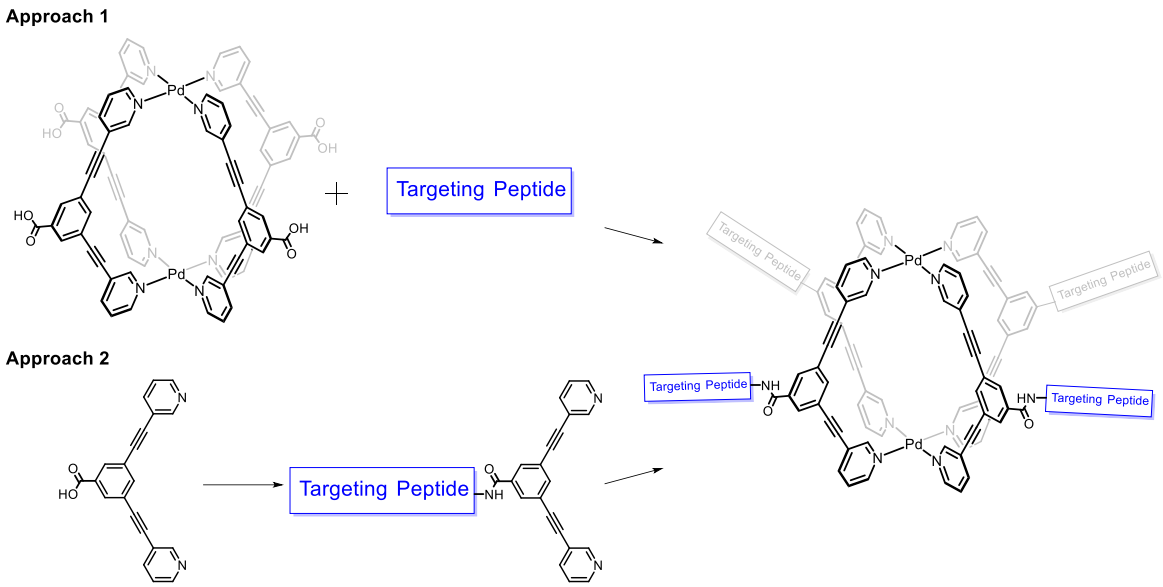


Figure 3.6. Scheme of the two different bioconjugation approaches as highlighted by Casini: (i) direct tethering of the metallacage to the peptide (ii) initial anchoring of the ligand to the peptide, followed by metallacage self-assembly. [15]

Invariably, both approach 1 and approach 2 require an intermediate ligand with a functional group the targeting moiety can be attached to easily. Popular bioconjugation reactions include esterification,[16] amide bond formation[15] and click-chemistry[17]; notably click-chemistry has been gaining traction recently due to both the synthetic ease of bond formation and stability of the resultant triazole.[18] Comparatively, ester and amide bonds are both more susceptible to cleavage by prevalent enzymes than triazoles, which could lead to premature release of the targeting group.

In developing functionalised cages, it was noted that the basic structure of **C1** should be maintained as it has shown good solubility and stability under biological conditions. Notably, an electron donating group, such as an amine, is required in the external 4-pyridyl position to strengthen the Co^{III}-pyridine bond and thus making the cage stable in biologically relevant conditions.

Figure 3.7 identifies three target ligands sharing the same quaterpyridine backbone but with various different functional handles for bioconjugation. Figure 3.7A shows a methoxy featuring ligand (**L5**), in comparison to the traditional **L3** the amine is replaced with an ether group, with the aim to investigate whether the lone pair on the oxygen would sufficiently strengthen the ligand-metal coordination bond so that the cage is stable under biological conditions.

Both Figure 3.7B and Figure 3.7C show ligands (**L6, L7** and **L8**) which utilise the *para*-positioned amine group to increase strength of the resulting cobalt coordination bond. Functionalisation of the ligand with ethanolamine and ethylenediamine, enables the bioconjugation of the ligand via either ester or amide bond formation, whilst **L8** features an alkyne group capable of reacting with azides via click-

chemistry to form a triazole bond. The various target ligands once synthesised and assembled into cages, will not only provide an insight into the role of the *para*-substituted group on cage stability but also the suitability of different bonds for bioconjugation.

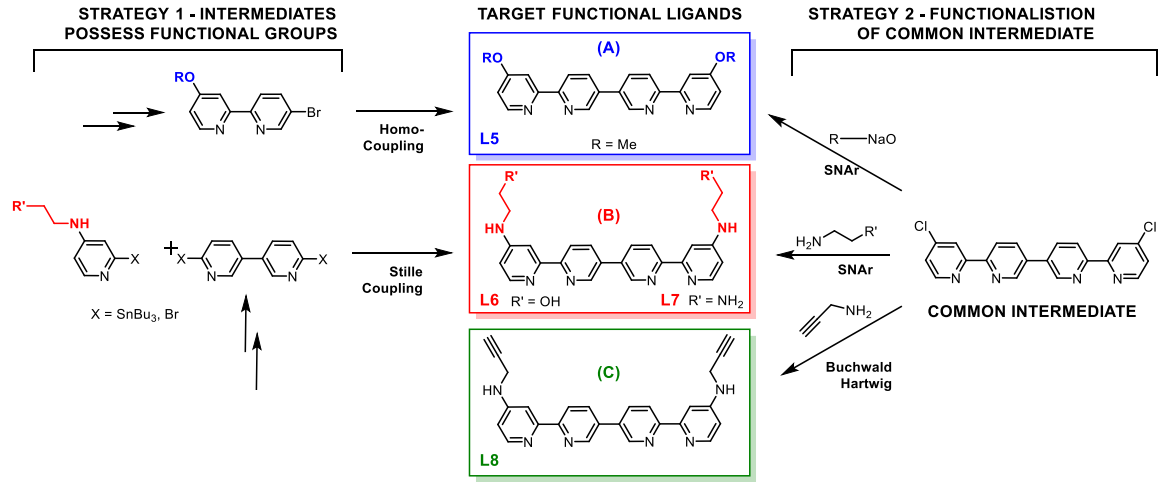
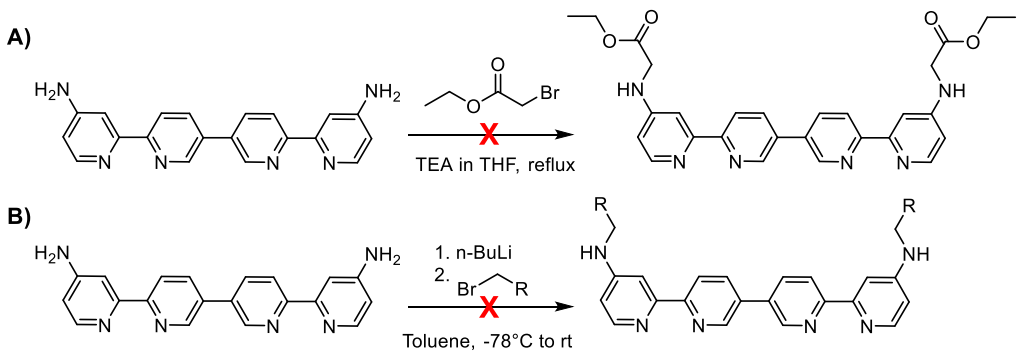


Figure 3.7. Visualisation of synthetic routes to target functional ligands (A) ***L5*** *(B)* ***L6*** *and* ***L7*** *Ligand and (C)* ***L8****. Strategy 1 focusses on installation of the functional groups prior to the formation of the ligand backbone whilst Strategy 2 utilises a common ligand intermediate where the functionalisation occurs in the final step.*

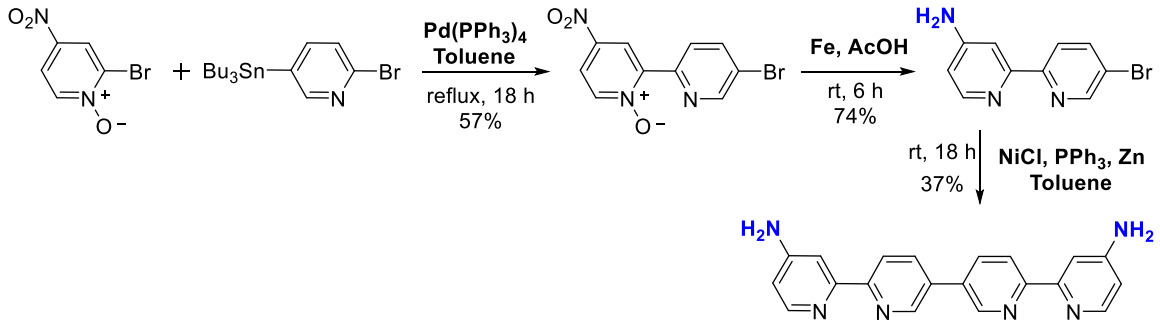
It is noted that preliminary research following the previous synthesis of **C1**, attempted to functionalise the **L3** by directly reacting the amine. However, it transpired that the amine group of the **L3** was incredibly unreactive likely due to its propensity to donate its electrons into the delocalised pyridine ring.[19] The two attempted approaches for direct functionalisation of the ligand are highlighted in Scheme 3.1 where R represents an anchor to functionalise onto, either a free amine or alkyne.[19]

Whilst, superficially the amine group would appear to be a good anchor, its positioning makes it remarkably inert and so direct formation of an amide bond is not possible here.[20] In fact, the nitrogen has such as lack of nucleophilicity that even prior deprotonation with n-butyllithiuim proved ineffective (Scheme 3.1B).



*Scheme 3.1. Approaches to functionalising **L3** directly. R indicates the group is being conjugated onto the ligand. Attempted by Dr William Grantham.*

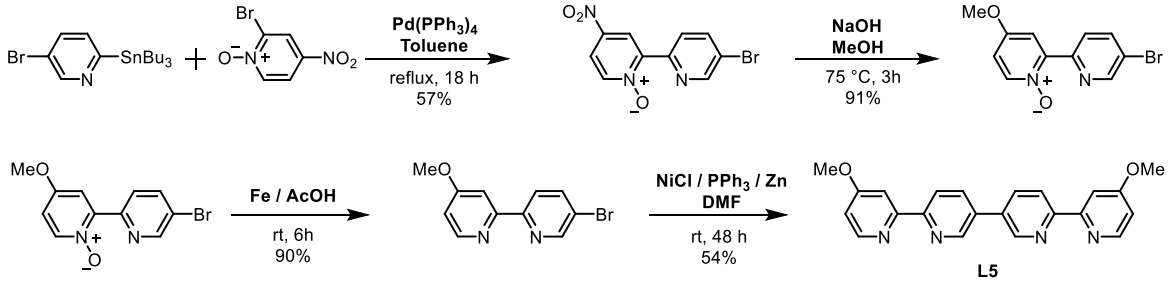
Due to the low reactivity of **L3**, none of the proposed synthetic routes to the target ligands identified in Figure 3.7 utilised the previously synthesised **L3** as an intermediate. Alternatively, initial research focused on synthesising the target ligands via a similar route to **L3** (Scheme 3.2), where the functionality of the ligand is incorporated into the intermediates prior to the formation of the quaterpyridine backbone, identified as Strategy 1 in Figure 3.7. °



*Scheme 3.2. Reaction scheme for the synthesis of **L3** with amine-functional groups highlighted in blue for clarity.*

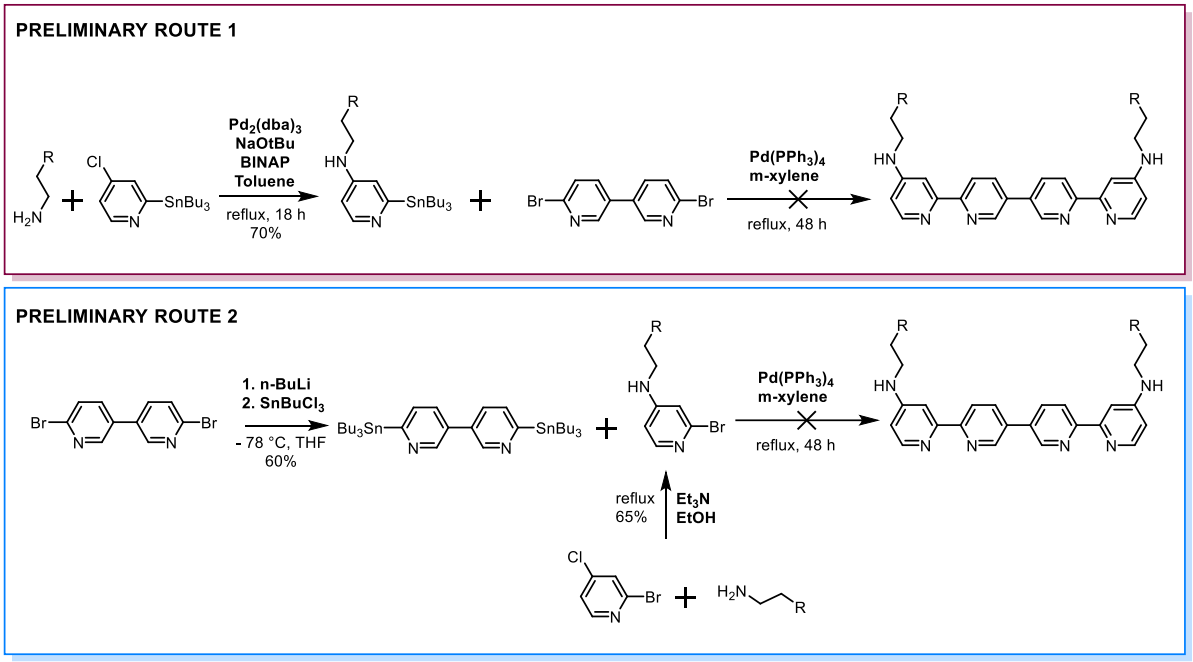
Previous attempts had been made by the Lusby group to obtain **L5** via Strategy 1 utilising the final homo-coupling step as shown in Figure 3.7, with full synthetic scheme shown in Scheme 3.3. However, the synthesis was found to be lengthy with low overall yield (3%) and the intermediates were unrequired for other ligands within the lab.[19]

PREVIOUS SYNTHETIC ROUTE



*Scheme 3.3. Previously optimised synthetic route to **L5** as identified by Lusby and co-workers.*[19]

Similar synthetic routes were designed for target ligands, **L6** and **L7**, where intermediates were functionalised prior to the formation of the quaterpyrdine ligand backbone as defined by Strategy 1 (Scheme 3.4). For both routes in Scheme 3.4, the reaction scheme failed at the final Stille cross-coupling step. Upon considering the literature it seemed likely that the increased electron-density in the pyridine ring provided by the *para*-substituted amine group served to deactivate the substrate by stabilising the Ar-X bond and preventing oxidative addition and/or trans-metalation.[21] Interestingly, this occurred both when the electron-donating amine group was on the stannylated substrate (preliminary route 1) and when it was the aryl-halide (preliminary route 2). The ability of electron-rich groups to deactivate palladium cross coupling reactions is discussed further in the synthetic section of Chapter 4.

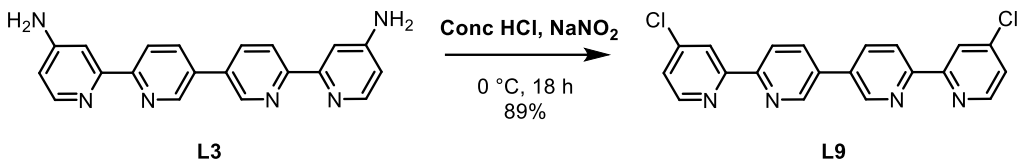


*Scheme 3.4. Preliminary proposed synthetic routes to **L6** and **L7**.*

3.2.2 Synthesis of Common Intermediate

Following the limited successes of Strategy 1 to obtain the targeted ligands by functionalisation of intermediates prior to formation of the ligand backbone, Strategy 2 identified a common ligand intermediate (**L9**). This compound would then be used through a single step functionalisation following the formation of the quaterpyridine ligand backbone, shown on the right of Figure 3.7. This method proved much more successful due to the usage of a common intermediate minimising the synthetic processes to access the library of target ligands.

Whilst direct attempts at functionalisation of **L3** were unsuccessful, an attempt was made to access the common intermediate, **L9** from the previously synthesised **L3**. The Sandmeyer reaction was identified as a way of converting the amine functional group into the halide (Scheme 3.5).Going directly from **L3** to **L9** through a Sandmeyer reaction proved to be effective (89% yield) only on a small scale. The insolubility of **L3** required 40% hydrochloric acid and heat for dissolution and still exhibited saturation at about 0.5 mg/mL. Given the need for temperature control and concentrated acid neutralisation this reaction proved impractical for large scales.



*Scheme 3.5. Schematic of the Sandmeyer reaction for the conversion of **L3** to **L9***

Direct synthesis of **L9** (*i.e.*, not via **L3**) was decided upon with two possible routes highlighted in Figure 3.8 as Strategy 1 and Strategy 2. The first method involved synthesising 5,-bromo-4-chloro-2,2'-bipyridine (**1**) via a Stille coupling of chloro-stannylpyridine (**2**) with 2-iodo-5-bromopyridine and a subsequent homocoupling. The second route involved the formation of 5,5'-dibromo-2,2'-bipyridine (**3**) via a Stille cross coupling, followed by another Stille cross coupling with **2** to yield **L9** in a four-step synthesis.

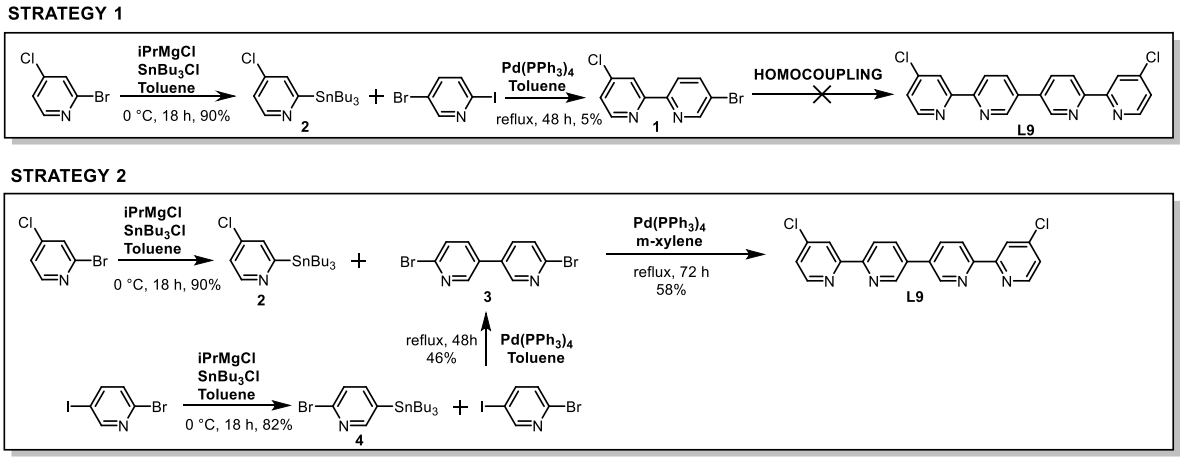
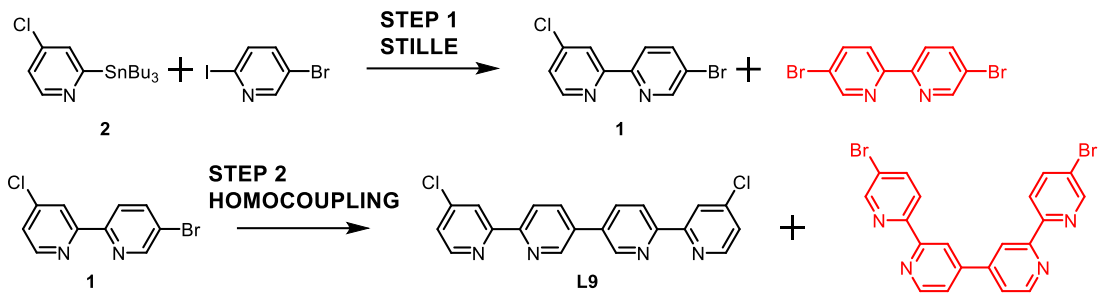


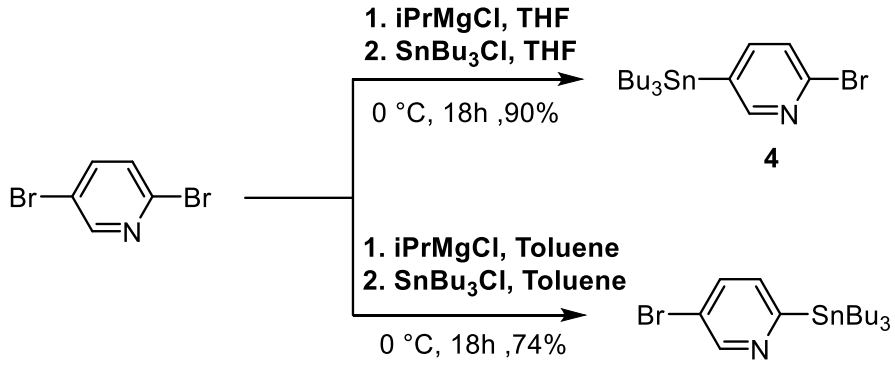
*Figure 3.8. Two reaction schematics for making **L9**. Strategy 1) Stille cross coupling followed by Negishi cross coupling Strategy 2) Combination of two successive Stille cross couplings.*

Both routes utilised the intermediate **2**, which was obtained in a one-step synthesis from chloro-bromopyridine using iPrMgCl following by $SnBu_3Cl$, and had a yield of 90%. Strategy 1 involved coupling **2** with 2-iodo-5-bromopyridine, however, this reaction resulted in partial homo-coupling generating 5,5'-bromo-2,2'-bipyridine (as shown in Scheme 3.6) as well as the intended product. The overall yield to **1** was about 5%, this was then followed by a low yielding homo-coupling reaction. It is hypothesised that **1** could react with itself both via the bromine and the chlorine substituents, leading to the formation of different oligomers, visualised in Scheme 3.6 in red. Both the reactions involved in Strategy 1 were found to be low yielding leading to the pursuit of Strategy 2.



*Scheme 3.6. Proposed Strategy 1 for the synthesis of **L9** with the possible by-products shown in red.*

Preliminarily difficulties were faced obtaining the selective stannylation at the 5-position of the 5-iodo-2-bromopyridine, the first step of Strategy 2. It was hoped that since the iodine is a better leaving group than bromine and subsequently more reactive the placement of the halogens would favour the 5-site for stannylation, however when the reaction was initially attempted in toluene this was not the case. Previous work by Getmanenko and Twieg reported that changing the solvent from toluene to the more coordinating solvent, THF could favour the 5-tributylstannyl-2-bromopyridine (**4**) isomer over 5-bromo-2- tributylstannylpyridine, exemplified in Scheme 3.7.[22] This occurs as a result of the THF coordinating to the pyridine centre and sterically crowding the 2-bromo position preventing access to the Grignard reagent and thus favouring the reaction at the less crowded 5-position.



Scheme 3.7. Selective stannylation of dibromopyridine via *solvent control.*[22]

Upon optimisation, the two resulting stannylation reactions were both high yielding (90% and 82%, shown in Figure 3.8). The following two Stille reactions of Strategy 2 worked relatively well with yields of 46% and 58%, with the loss in yield likely a result of the poor solubility of both **3** intermediate and **L9**. Low solubility of these multi-pyridyl systems is common and was observed with **L3** which only dissolved significantly in DMSO.

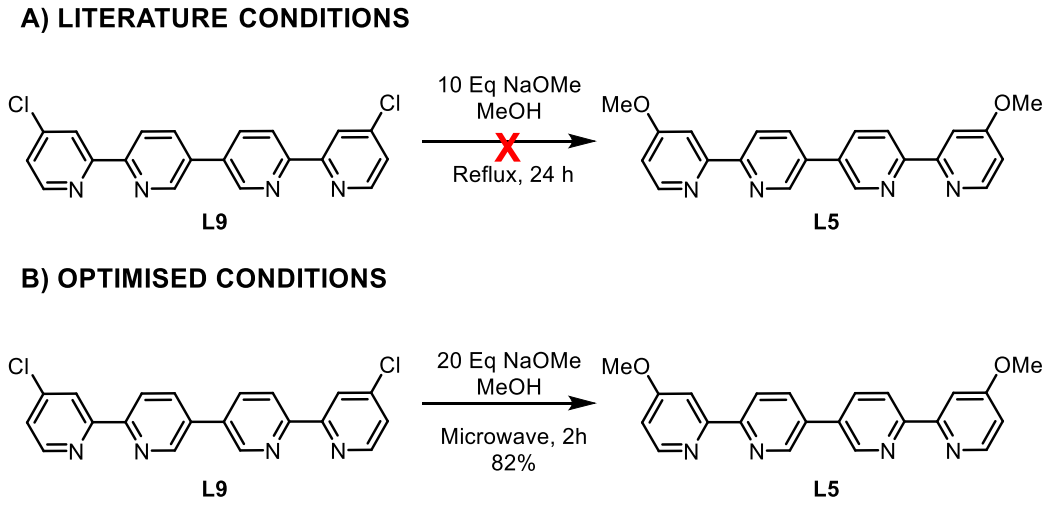
3.2.3 Functionalisation via S_NAr Reaction

Synthesis of Ether Ligand

The initial study that identified **C1** as being capable of altering the biodistribution of $[^{99m}Tc]TcO_4^-$ evidenced the importance of having an electron-donating group in the *para* position of the pyridine ring.[1] When investigating the same M_4L_6 system where the ligand was a simple quaterpyridine with no functionalisation the resultant cage was unstable in biological serum and in the presence of bio-reductases such as glutathione.[1] The ability of the *para*-amine to donate electron-density into the ring system and subsequently stabilise the resultant Co^{III}-ligand bond is imperative for cage integrity in biological conditions. It was hypothesised that the electron donating capability of the amine group might be emulated by an ether group, where the lone pair of electrons on the oxygen might similarly stabilise the Co^{III}-ligand bond.

As previously mentioned, attempts to synthesise **L5** in a similar fashion to **L3** were low yielding and time consuming. Consequently, **L9** was identified as an intermediate from which to obtain the ether ligand. The conversion of 2 or 4-halopyridines to the equivalent methoxy compound is well known and can be achieved using a simple S_NAr reaction.[23] The standard literature conditions stated 5 Eq of sodium methoxide per halogen at reflux in MeOH over 24 h. These conditions gave minimal conversion (Scheme 3.8A), with the low solubility of **L9** in MeOH likely a significant factor.

To improve dissolution, higher temperatures were employed, with the reaction repeated in a high pressure vial with 140 °C microwave irradiation for 2 hours. Also, an increase in the equivalents of sodium methoxide to 10 Eq per chloride was used. This led to the complete conversion of **L9** to **L5** and a resultant yield of 82% (Scheme 3.8B).

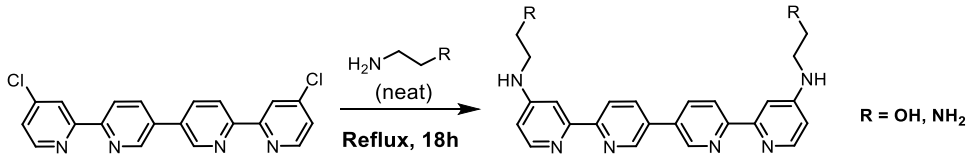


*Scheme 3.8. Formation of **L5** from **L9**. Comparison of the literature conditions versus the optimised conditions.*

Synthesis of Amine Containing Ligands

Target ligands **L6** and **L7** are amine-based ligands analogous to **L3**. They feature a secondary amine with short alkyl spacer to the bio-conjugation handle in the form of either a free alcohol (**L6**) or a free amine (**L7**). Following the successful synthesis of **L5**, it was hypothesised that a S_NAr reaction could also be used to obtain **L6** and **L7**. Literature conditions indicated the use of a slight excess of amine (ethanolamine and ethylenediamine), Et_3N as a base and to reflux in an alcohol solvent such as ethanol.[24]

However, there were issues with a lack of ligand solubility and the reaction mixture showed only starting material after 48 h. Given the basic nature of both ethanolamine and ethylenediamine, and their relatively high boiling points, the S_NAr reactions were re-attempted by refluxing the ligand in the neat amines, thereby it is the substrate, the base and the solvent.[25] Refluxing in neat amine proved successful with a yield of 67% and 81% for the **L6** and **L7** respectively (Scheme 3.9).



*Scheme 3.9. Reaction scheme for the successful S_NAr of **L9** with ethanolamine for **L6** (67%) and for **L7** (81%).*

The S_NAr reaction significantly changes the magnetic proton environments in the pyridine ring systems. This can be seen in Figure 3.9, the blue box which highlights the shielding of the *meta*-proton (shown in blue) from 7.4 ppm (**L9**) to 6.6 ppm (ethanolamine and ethylenediamine). The two S_NAr products, are quite similar but distinguishable by the positioning of the alkyl peaks, where there is more deshielding in **L6** at 3.2 and 3.6 ppm in comparison to **L7**, as a result of the more electronegative alcohol group.

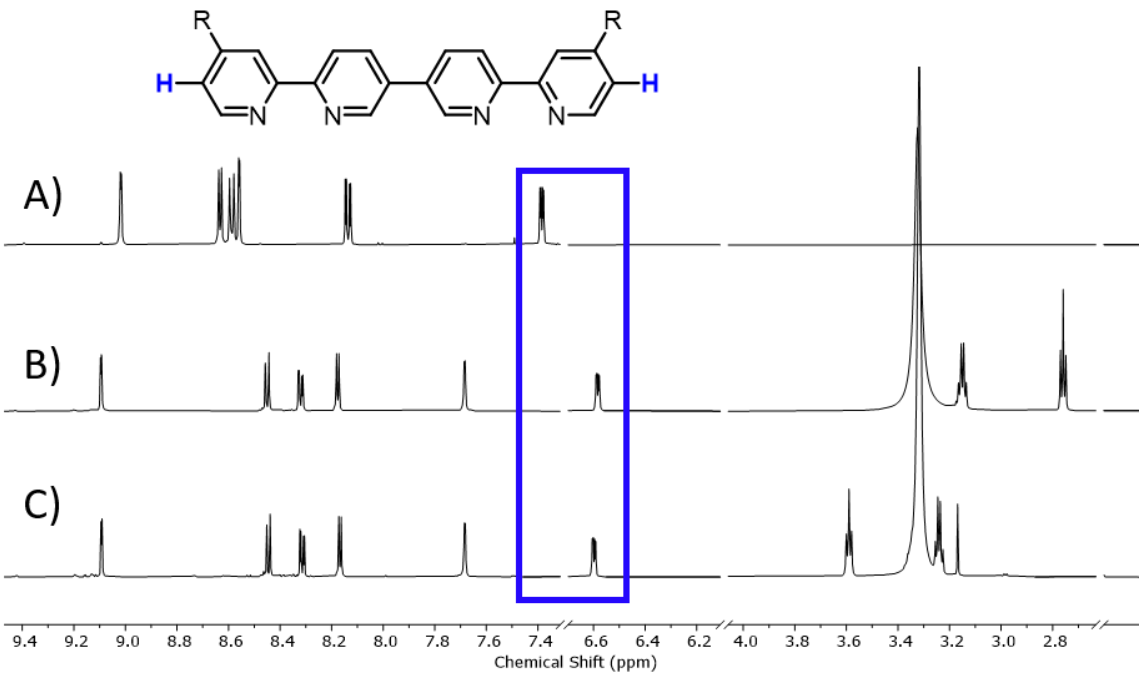


*Figure 3.9. Partial ^{1}H NMR (500 MHz, d_6-DMSO) Spectra showing successful S_NAr reaction with **L9** A) **L9** (in $CDCl_3$) B) **L7** and C) **L6**. R = either Cl for **L9**, $NH_2CH_2CH_2NH_2$ for **L7** or $NH_2CH_2CH_2OH$ for **L6**.*

It is noted here that the S_NAr method was attempted on a range of alternative amines alongside ethanolamine and ethylenediamine, however, it was not universally successful (Figure 3.10). When attempted with alkyne based amines, which were of interest for subsequent "click" chemistry, the S_NAr reaction instead formed unidentifiable mixtures from which the products could not be extracted. No further attempts to explore these reactions were made.

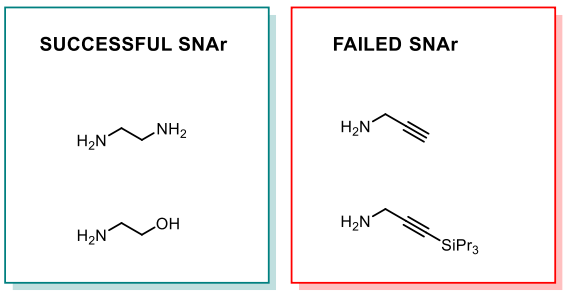


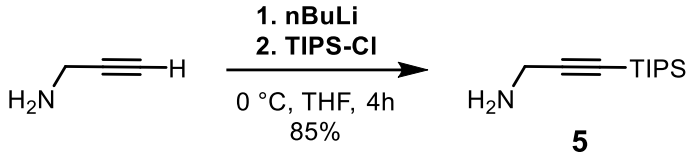
*Figure 3.10. S_NAr reaction substrate scope for suitability with **L9.***

3.2.4 Functionalisation via Buchwald-Hartwig reaction

*Synthesis of **L8** Featuring Alkyne*

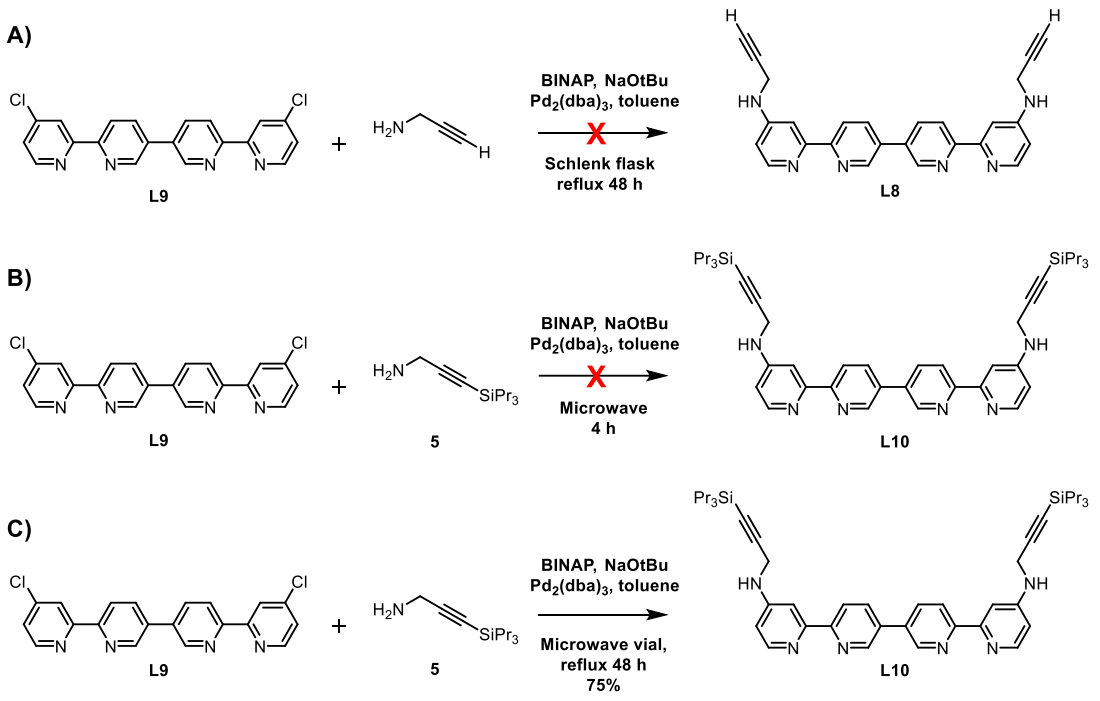
As the S_NAr route was unsuccessful in the installation of an alkyne group, a Buchwald-Hartwig coupling reaction using **L9** was envisaged. The Buchwald-Hartwig amination exists as a reaction for forming carbon-nitrogen bonds from aryl halides and was first reported in 1983 by Stephen Buchwald and John

Hartwig.[26] Initially propargylamine was used in the reaction (Scheme 3.11A), however the reaction failed and the free alkyne was considered to be the problem potentially due to its ability to react with aryl-halides via a Sonogashira coupling. In order to prevent reaction at the alkyne terminus, the protected version of propargylamine (**5**) was synthesised using nBuLi and triisopropylsilane, which proceeded well with an 85% yield (Scheme 3.10). [27]



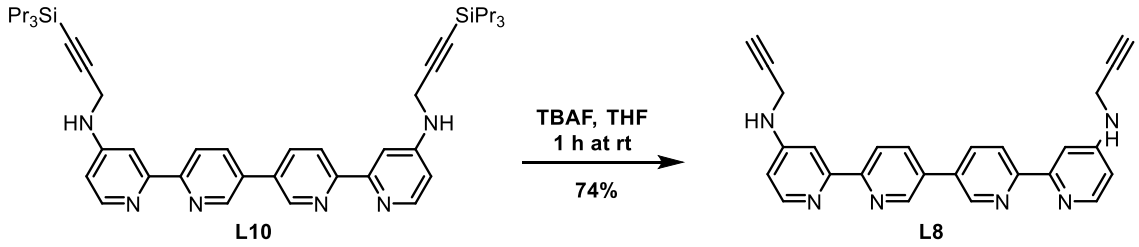
Scheme 3.10. Protection of free alkyne in propargylamine using nBuLi and triisopropylsilane in dry THF with conditions adapted from literature.[27].

The Buchwald-Hartwig coupling reaction was repeated with **5** and progressed well under reflux in dry toluene in an inert atmosphere within a microwave vial. Interestingly the reaction proved pressure sensitive and only progressed in good yield when carried out in a sealed microwave vial at high temperatures (Scheme 3.11C). Endeavours to scale the reaction in a large Schlenk flask, whilst maintaining all other conditions, proved unsuccessful presumably due to the reliance on increased pressure. Unfortunately the reaction did not progress at all when microwaved (Scheme 3.11B). The optimised Buchwald-Hartwig gave a yield of 75% but was unable to be scaled above 100 mg per batch.



*Scheme 3.11. Buchwald-Hartwig cross coupling reaction with **L9** to obtain alkyne functionalised ligand.*

Following the coupling reaction, **L10** was then deprotected using standard conditions of 2.5 Eq of TBAF per protecting group in THF with a yield of 74%, as shown in Scheme 3.12.[28] The stability of the **L8** was unknown and it was consequently de-protected directly prior to further requirement.



*Scheme 3.12. Reaction conditions for the deprotection of **L10** to form **L8***

Click chemistry on L8

Obtaining a ligand with an external alkyne handle is preferential for bioconjugation due to the reliability of the click reaction and the stability of the resultant bond. This technique was developed at the end of the 20th century by Sharpless and colleagues, and involves a 1,3-dipolar cycloaddition between an azide and an alkyne forming a triazole.[17] These reactions traditionally require a copper(I) catalyst with the mechanism shown in Figure 3.11, but have been optimised to proceed effectively under a variety of conditions.[18,29,30]

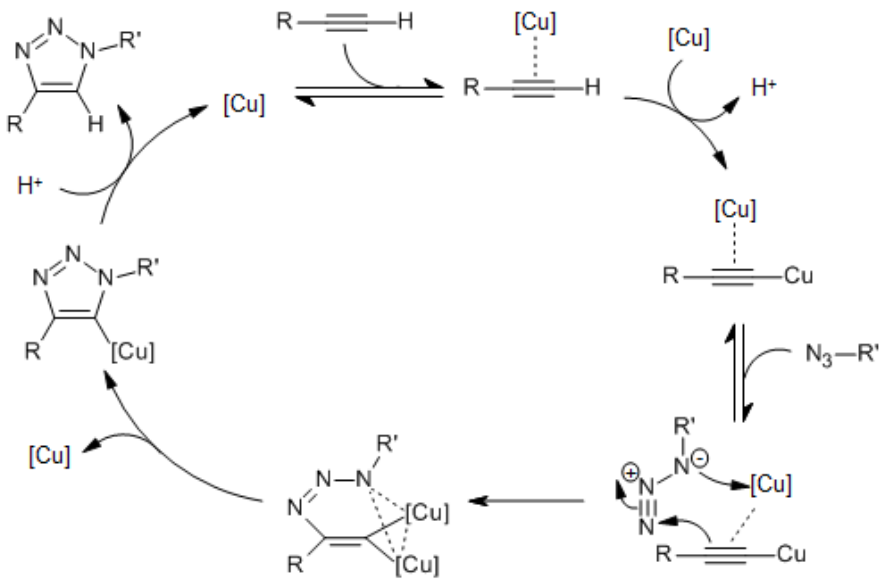


Figure 3.11. Mechanism for the copper catalysed click-chemistry reaction between an alkyne and an azide, as proposed by Sharpless and co-workers.[17]

Prior to attempting the click reaction with any targeting groups, the reaction was attempted with a simple polyethylene glycol (PEG) azide (**6**). As eluded in Chapter 2, it is hypothesised that the *in vivo* bioaccumulation of **C1** was related to its interactions with proteins in the liver and therefore by conjugating PEG groups to the external of the cage the binding with proteins might be reduced and the *in vivo* distribution altered.[1]

Due to the multidentate nature of **L8**, it was considered that the catalytic copper salt might form unwanted interactions with the ligand. Lipshutz and co-workers published a study in 2006 detailing the generation of a heterogeneous Cu/C catalyst capable of catalysing click reactions.[29] It was hypothesised that if immobilised onto a heterogeneous carbon surface the copper might be less likely to interact with the coordination sites of the ligands and the subsequent recovery of the catalyst would be straightforward. The literature procedure for the generation of the catalyst, involved loading $Cu(NO_3)_2$ onto Darko Carbon, to a resulting catalyst loading of 1.01 mmol/g. A 2015 study by the Lusby group evidenced that the Cu/C catalyst, designed by Lipshutz and co-workers, was capable of catalysing click chemistry between dialkynyl-pyridine units and PEG azides, Figure 3.12.[31] This previous reaction reported a good yield of 86% and it was hypothesised the same conditions which utilised Et_3N and DCE as solvents could be employed here.

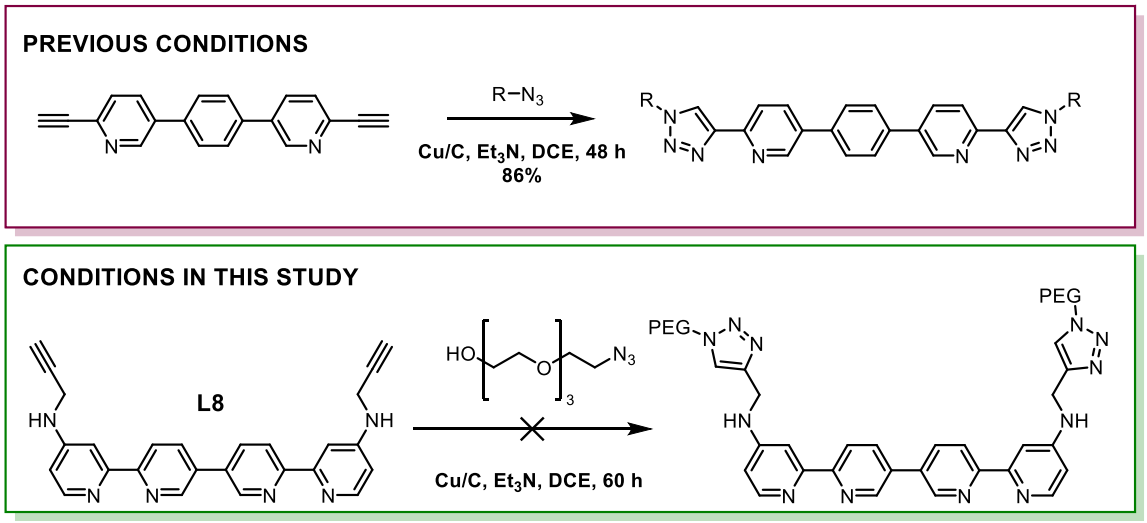


Figure 3.12. Visualisation of heterogenous catalysis conditions for click-chemistry of ligands. Previous Conditions: Successful click reaction using Cu/C catalyst (Cu/C, Et_3N, DCE, 48 h, 65°C, 75%) to form ligand of interest for the Lusby group.[31] Conditions unsuccessful in this study (Cu/C, Et_3N, DCE, 60 h, 65°C)

Unfortunately, when these conditions were attempted on **L8**, the reaction was unsuccessful. The reaction was sampled over 60 h and monitored via ^{1}H NMR (Figure 3.13B), however, only **6** could be identified. Moreover, after filtering the reaction mixture to remove the heterogeneous catalyst, the filtrate (Figure 3.13C) seemed to contain no triazole product or **L8** starting material. The lack of ligand present in both the reaction mixture and the product could either be a result of minimal dissolution in the solvent system or even though the copper was immobilised onto the carbon it was still interacting with the ligand and thus it remained in the retentate.

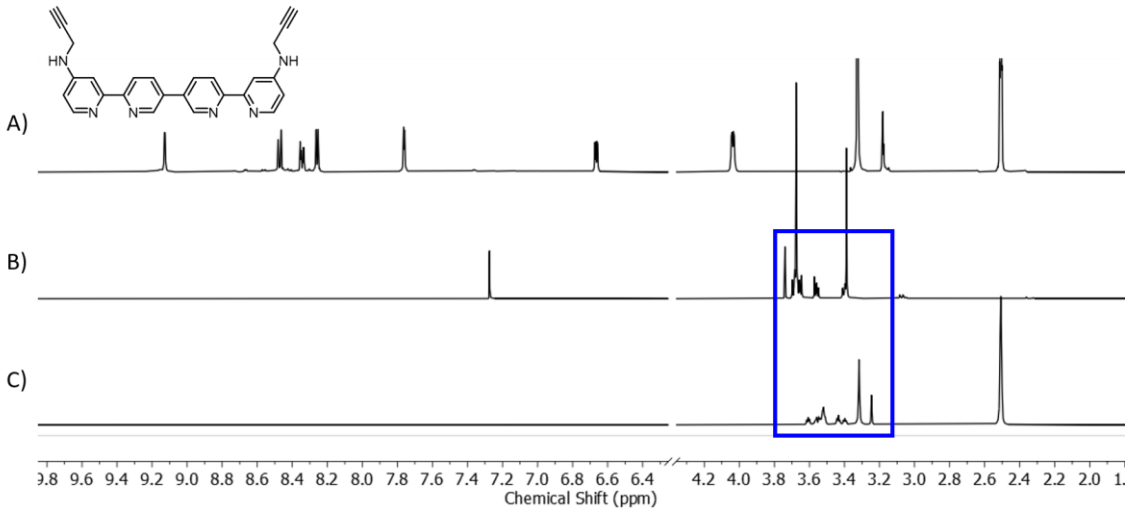


*Figure 3.13. ^{1}H NMR (500 MHz, d_6-DMSO) Attempted click chemistry of **6** with **L8** (Cu/C, Et_3N, DCE, 60 h, 65°C) A) **L8** B) reaction mixture ($CDCl_3$) and C) isolated product following filtration of reaction mixture.*

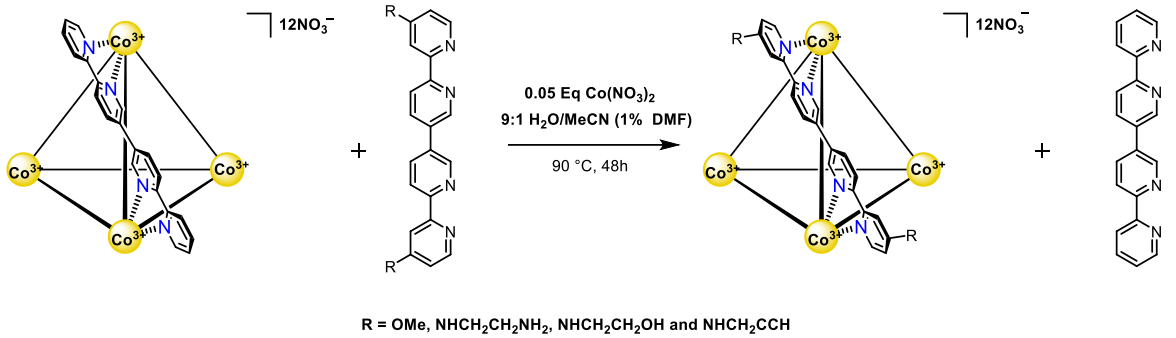
It was considered that the copper catalysed click-chemistry on the uncoordinated ligand is hindered by the ability of copper to coordinate to the free bipyridine units. In comparison, the previous click-

chemistry study by the Lusby group, shown in Figure 3.12, had utilised intermediates with only single pyridine units in comparison to the quaterpyridine backbone of **L8**. Whilst the product might exhibit a chelate effect, any coordination of the pyridyl-triazole group to the catalyst would still be significantly weaker in comparison to the *para*-amine substituted ligand, **L8**.

In the next section, performing the click reaction on the assembled cage is discussed. As the ligands are coordinated to the Co^{III} centers it was expected that interactions with the click catalyst would be reduced.

3.2.5 Cage Assembly

The previous method for **C1** assembly involved direct assembly of the ligands with Co^{II} followed by oxidation of the metal centres with cerium ammonium nitrate (see Chapter 2).[1] However, the yield of only 26% was deemed too low for the final stage of the multistep synthesis. Alternatively, a transligation method was adopted which utilised the more labile **C2** as a template whereby the less strongly coordinating **L2** would be displaced by **L5**, **L6**, **L7** or **L8** (Scheme 3.13). The transligation conditions were optimised for **C1**, as discussed in Chapter 2, the addition of a sub-stoichiometric amount (0.05 mol%) of Co^{II} was required to improve the rate of substitution presumably through an electron-transfer mechanism.



*Scheme 3.13. General conditions for the formation of stable functionalised cages from **C2**. Where R represents the functionalised groups: OMe, $NHCH_2CH_2NH_2$, $NHCH_2CH_2OH$ and $NHCH_2CCH$.*

Subsequently the same conditions were used when forming the functionalised cages, based on the newly synthesised ligands **L5**, **L6**, **L7** and **L8**. This was first attempted with **L5**, where an excess of ligand was added to **C2** and the reaction mixture was allowed to equilibrate at elevated temperatures over 48 h. The excess ligand was removed via celite filtration and the cage eluted in H_2O. The ^1H NMR spectra of the templating cage (Figure 3.14A), the eluted product (Figure 3.14B) and **L5** (Figure 3.14C) are shown below.

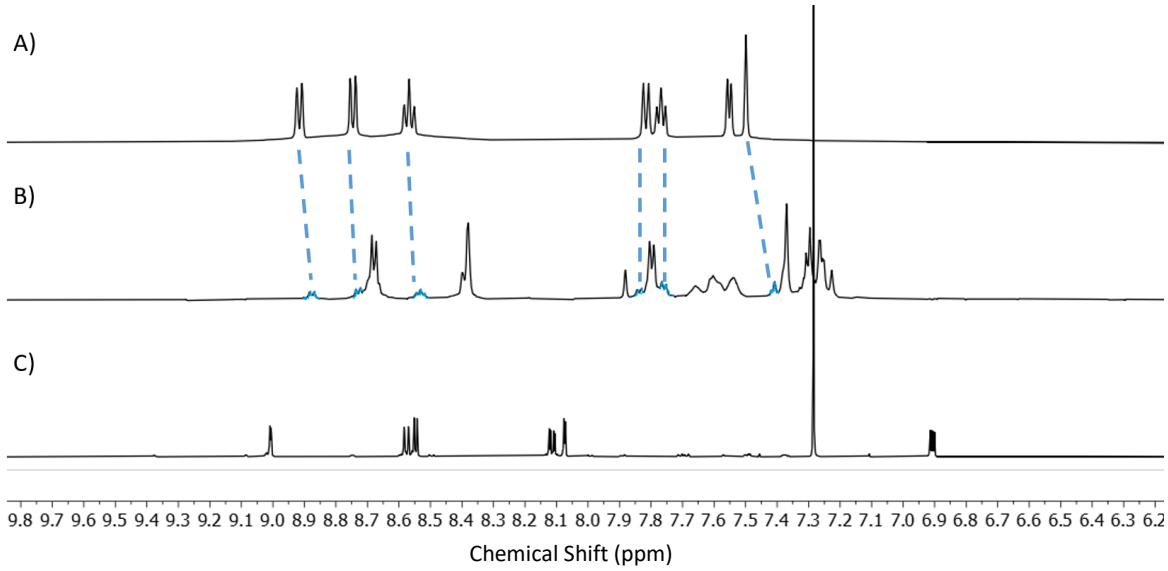


*Figure 3.14. Partial ^{1}H NMR (500 MHz, D_2O) showing **C5** assembly* via *transligation of QPTY cage. Achieved by the addition of an excess of **L5** to **C2** in D_2O, 10% CD_3CN and 1% DMF with 0.05 Eq $Co(NO_3)_2$ A) **C2** template cage B) **C5** C) **L5** (in $CDCl_3$)*

Whilst the majority of the obtained product was found to be **C5** there appeared to be approximately 10% **C2** impurity remaining (Figure 3.14B). This was hypothesised as a result of the system existing in an equilibrium between the two cages (Figure 3.15). In an attempt to bias the suspected equilibrium towards **C5**, the trans-ligation experiment was repeated using the isolated 90% **C5** product with a further excess of **L5**. Interestingly, this had no effect on the cage mixture composition.

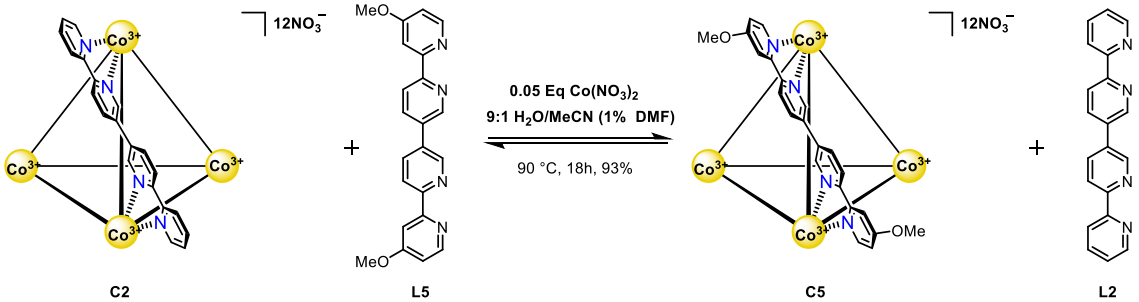
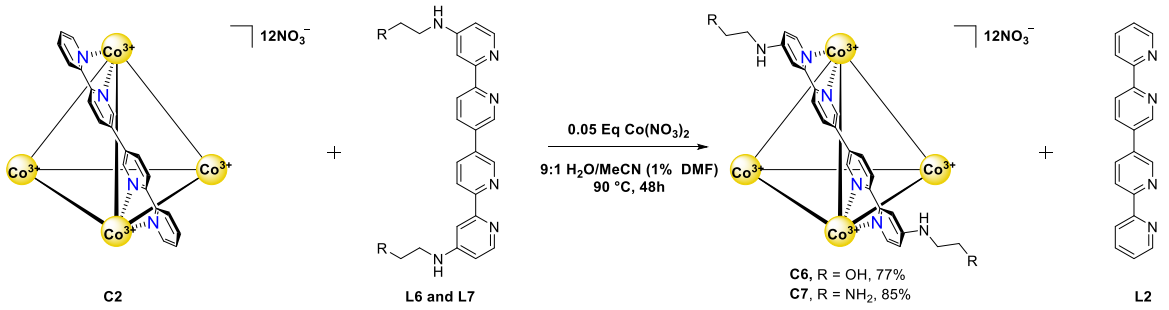


*Figure 3.15. Reaction scheme showing equilibrium present during trans-ligation of **C2** to form the **C5**.*

Size-exclusion chromatography was also attempted on the **C5** product, however, somewhat predictably the **C5** and **C2** impurity eluted simultaneously, presumably due to their near identical sizes. The existence of an equilibrium between **C5** and the more labile **C2** indicates that the Co^{III}-**L5** bond is reversible and potentially more labile than that of its Co^{III}-**L3** counterpart. Given the cage exhibited 90% purity and the only impurity was the well characterised **C2**, further studies of the cage were conducted on its impure form. The complex was further characterised using DOSY, ^{13}C NMR and MS (shown in Figures SI 33-35). The DOSY spectrum indicated a singular discrete species with a

hydrodynamic radius of 12.5 Å corresponding to the tetrahedral complex. The mass spectrum showed the presence of multiple, multiply charged species as a result of counter anion displacement, including the 3+ species found at 1005.1403 and the 4+ species found at 738.3683.

The transligation was then attempted using the two amino ligands, **L6** and **L7** (Scheme 3.14).



*Scheme 3.14. Reaction scheme for the transligation to form **C6** and **C7**. Conditions included 9:1 H2O/MeCN (1% DMF) with 0.05 Eq $Co(NO_3)_2$ and yields 77% - 85%.*

A comparison of ^{1}H NMR spectra for the free **L6** versus the respective assembled cage is shown in Figure 3.16. As expected, a shielding effect is observed for the aromatic signals (6.6 - 9.0 ppm) upon complexation, likely a result of π- π stacking as the ligands become sterically crowded around the coordination centre. Alternatively, when looking at the ethanolamine signals (3.2 – 3.8 ppm) there is a slight deshielding effects upon ligand assembly, presumably as a result of amine involvement in the Co^{III}-ligand bond removing electron density away from the alkyl region.

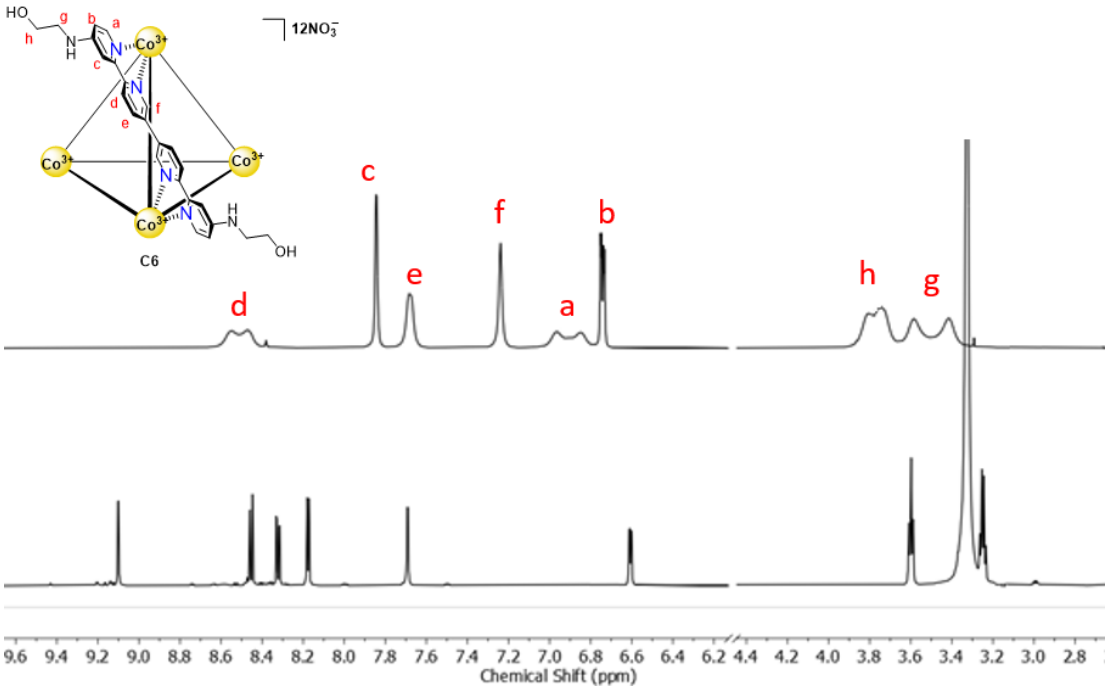


*Figure 3.16. Partial ^{1}H NMR (500 MHz, D_2O) showing **C6** assembly. Achieved by the addition of an excess of **L6** to **C2** in D_2O, 10% CD_3CN and 1% DMF with 0.05 Eq $Co(NO_3)_2$. Top spectrum) **C6**, bottom spectrum) **L6** (d_6-DMSO)*

The ^{1}H NMR of the cage shows broadening of several signals, notably at 8.5 and 6.9 ppm in the aromatic region and the two ethanolamine peaks in the alkyl region (3.2 – 3.8 ppm). It was hypothesised that this could be the result of the multiple conformations of the ethanolamine substituents. This could be caused by partial double bond character between the amine and pyridyl ring, leading to restricted rotation and thus multiple conformations. Variable temperature ^{1}H NMR experiments (Figure 3.17), showed that at higher temperature there is a significant sharpening of the signals. This process is reversible wherein when the sample was cooled to 27 °C the broadening returns. This experiment also shows the cage is stable at high temperature. **C6** was further characterised using DOSY, ^{13}C NMR and MS.

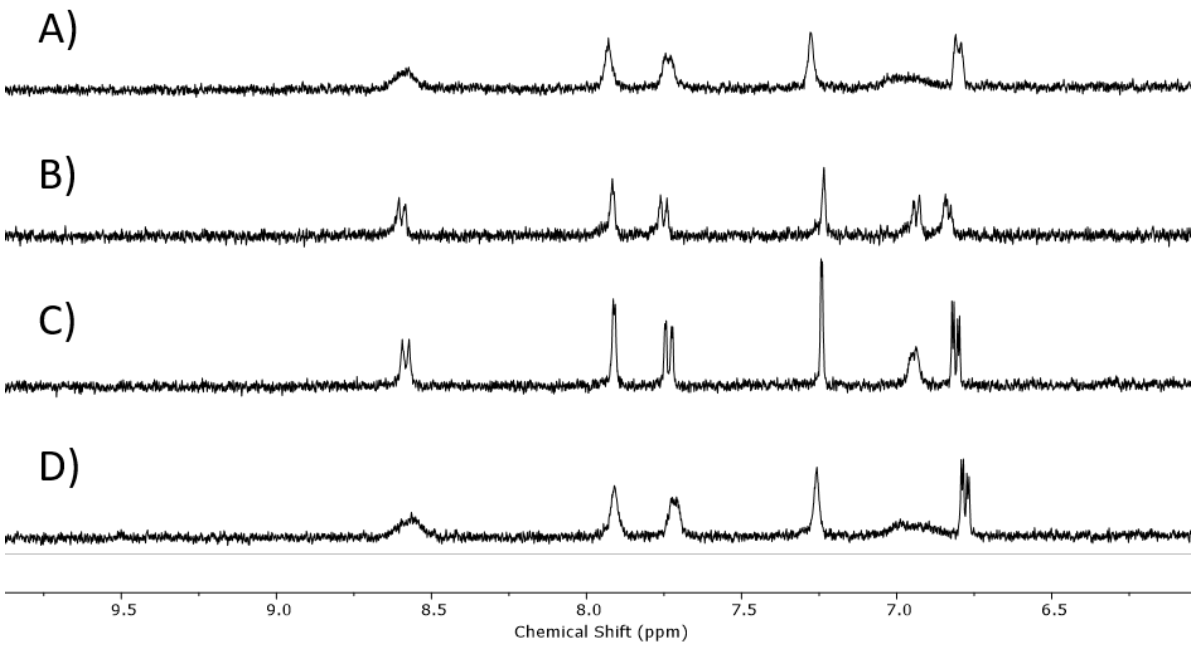


*Figure 3.17. Variable temperature ^{1}H NMR (400 MHz, D_2O) on a solution of **C6** (1 mM, 0.5 mL) A) 303 K B) 363 K C) 333 K D) 300 K. Carried out by Dr Émer Foyle.*

C7 was formed under the same conditions, shown in Figure 3.18, where the same shielding of the pyridine ring signals and deshielding of the alkyl signals was observed. Intriguingly the signals were even broader than in **C6**, and upon repeating the variable temperature experiments, the spectra remained broad. This indicated that there was likely a dynamic process occurring which was not resolved at the temperatures tested. **C7** was further characterised using DOSY and ^{13}C NMR.

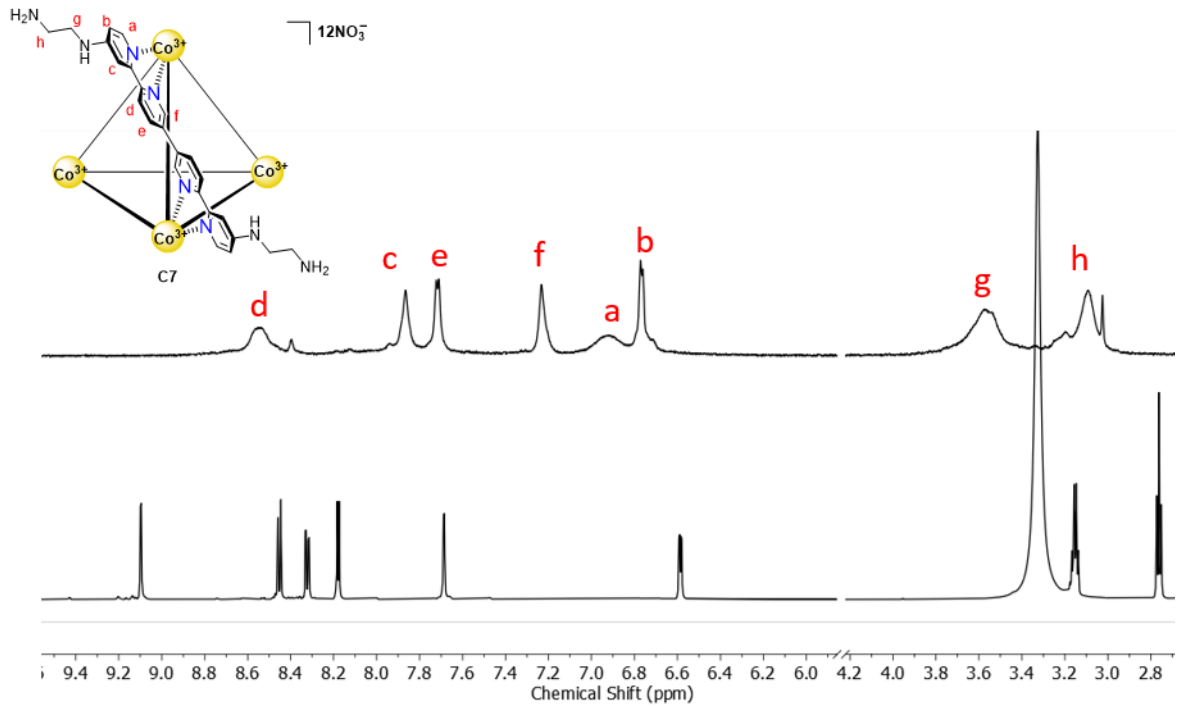


*Figure 3.18. Partial ^{1}H NMR (500 MHz, D_2O) showing **C7** assembly. Achieved by the addition of an excess of **L7** to **C2** in D_2O, 10% CD_3CN and 1% DMF with 0.05 Eq $Co(NO_3)_2$. Top spectrum) **C7**, bottom spectrum) **L7** (d_6-DMSO).*

Attempts were made to acquire MS of the **C7** but were unsuccessful. Initially the same conditions on the NanoMate® Synapt were used which proved successful for the other cages, however there appeared to be issues ionising the cage. Even when altering the applied voltage to the sample no multiply charged species were present. An anion metathesis was performed using KPF_6 with the aim of changing the cage solubility to a more volatile solvent, however the PF_6^- analogue of the cage in MeCN appeared to exhibit the same issues. It was concluded that the external primary amines were easily protonated and thus interfere with the traditional ionisation route via the removal of counter anions.

Lastly, a trans-ligation was attempted using **C8** with the same conditions as described earlier (Scheme 3.13). Similarly, the NMR of this cage showed significant broadening (Figure 3.19A). This can be seen in both the aromatic peaks between 6.8 and 9.0 ppm and the alkynyl peaks between 2.0 and 4.0 ppm. It is noted that although the conditions preliminarily proved successful for the formation of **C8**, the process proved difficult to repeat. As a result characterisation of the cage was limited to a ^{1}H NMR spectrum. Attempts at nanospray ESI MS was unsuccessful similar to **C7**.

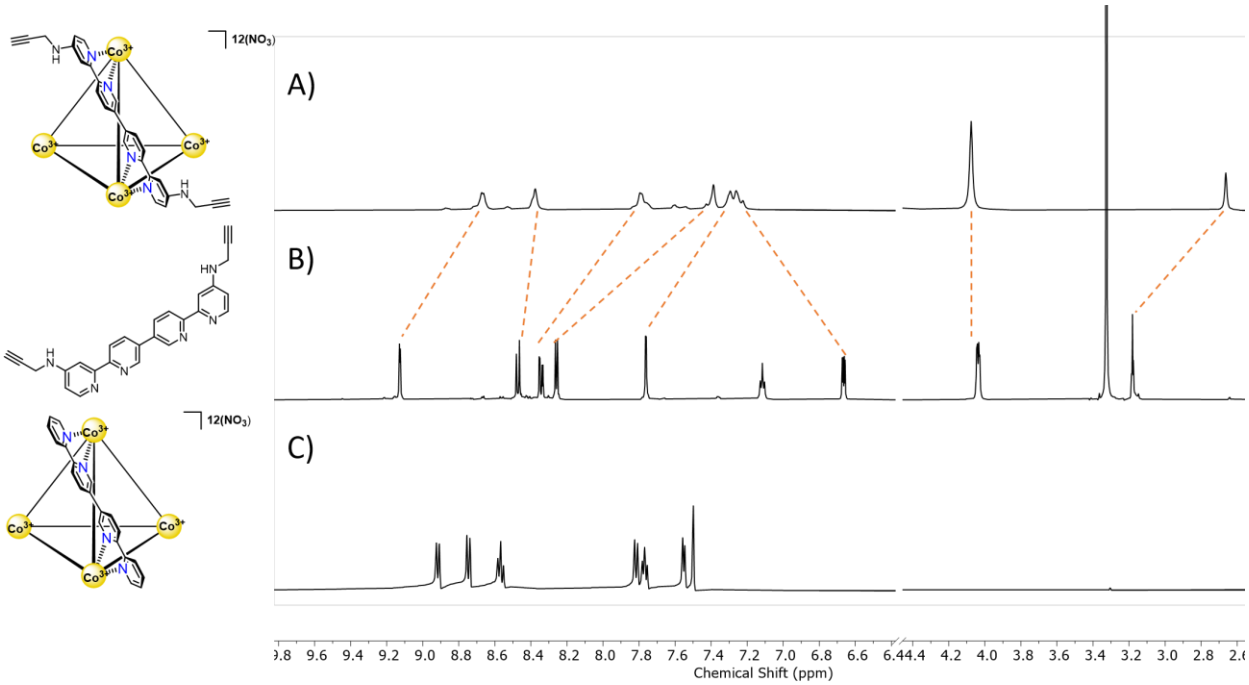


*Figure 3.19. ^{1}H NMR (500 MHz, D_2O) showing **C8** assembly via transligation achieved by the addition of an excess of **L8** to **C2** in D_2O, 10% CD_3CN and 1% DMF with 0.05 Eq $Co(NO_3)_2$. A) **C8** B) **L8** ($CDCl_3$) C) templating **C2**.*

It is also noted that an attempt was made to assemble **C9** prior to alkyne deprotection as visualised in, Figure 3.20. However, pure **C2** was isolated from the attempted transligation. This is considered to be a result of the lack of **L10** solubility, where the bulky hydrophobic nature of the triisopropylsilane protected ligand prevents its dissolution within the aqueous cage transligation solvent system.

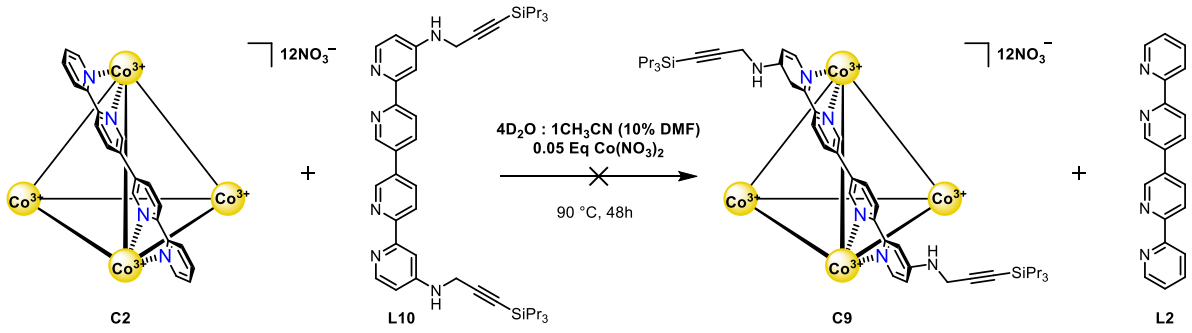


*Figure 3.20. Attempted transligation conditions of **L10** using template **C2** to form **C9.***

Click Chemistry on Cage

There are two methods for achieving external bioconjugation of a coordination cage, as described by Casini and co-workers and visualised in Figure 3.6, either post-assembly modification or pre-assembly.[15] Whilst previous studies by Casini and colleagues successfully assembled Pd_2L_4 cages from

ligands bioconjugated to peptides, attempts to perform click-chemistry on the free ligand were in this case unsuccessful.[15] Given the propensity of copper to coordinate to the free ligand therefore hindering its capabilities as a catalyst, post-assembly modification of the cage was attempted. It was hypothesised that since the ligands exist in a fully coordinated, saturated state within the Co_4L_6 cage they are less likely to interact with the metal catalyst and prevent the click-chemistry reaction of the external alkynes.

Two sets of reaction conditions were attempted (Figure 3.21), where 'Conditions A' utilise the same heterogenous Cu/C catalyst which was attempted with the free ligand. Alternatively, 'Conditions B' feature the more commonly used $CuSO_4$ with sodium ascorbate, which due to the presence of the free metal salt had been deemed unsuitable for the previous reaction with the ligand only. Both reactions employed an excess of PEG-azide (**6**).

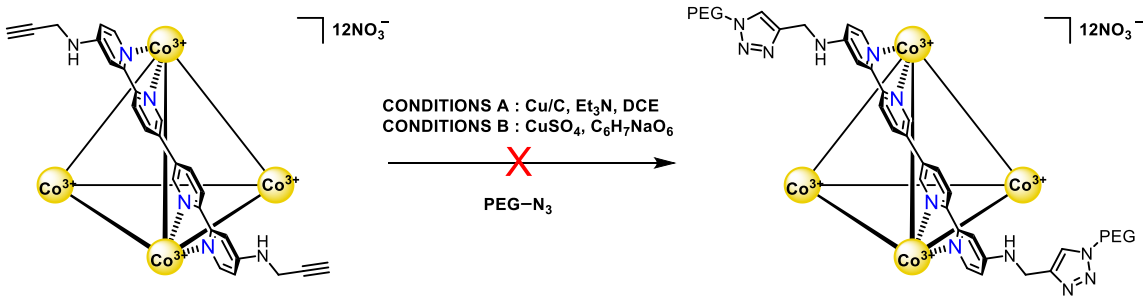


*Figure 3.21. Attempted conditions for click chemistry on **C8** with PEG-N_3. Conditions A : Cu/C, Et_3N, DCE and Conditions B : $CuSO_4$, Sodium Ascorbate.*

The ^{1}H NMR spectra of the reaction mixtures for both conditions after 48 h (Figure 3.22) show mixed results. For 'Conditions A', similarly to when attempted with the free ligand, the spectrum shows a lack of cage in solution, with only very weak signals in the aromatic region. The spectrum mostly highlights the presence of the PEG-azide, indicated by the red box, which although in excess is not the only species present in the reaction mixture. It is hypothesised that either **C8** has bound to the heterogeneous catalyst, removing it from solution and thus limiting its visibility by NMR or the reaction has occurred and the resulting product is water insoluble and thus not visible in the NMR.

On the other hand, the reaction mixture formed under 'Conditions B' gave more signals in the ^{1}H NMR spectrum (Figure 3.22B). Whilst there is still an excess of **6** present, there are other signals which could be indicative of a successful click-reaction. A second set of signals between 6.5 ppm and 10 ppm, highlighted in yellow, could be indicative on the 'clicked-cage', however the reaction stalled after the 48 h and the second species remained low in concentration. The reaction mixture also features an abundance of peaks in the alkyl region between 3.5 ppm and 5.0 ppm, highlighted in blue, which could be the **6** peaks becoming deshielded upon the formation of the triazole.

Given the small scales (2mg **L8**) both reactions were attempted on, further characterisation via either ^{1}H DOSY or HSQC was not possible.

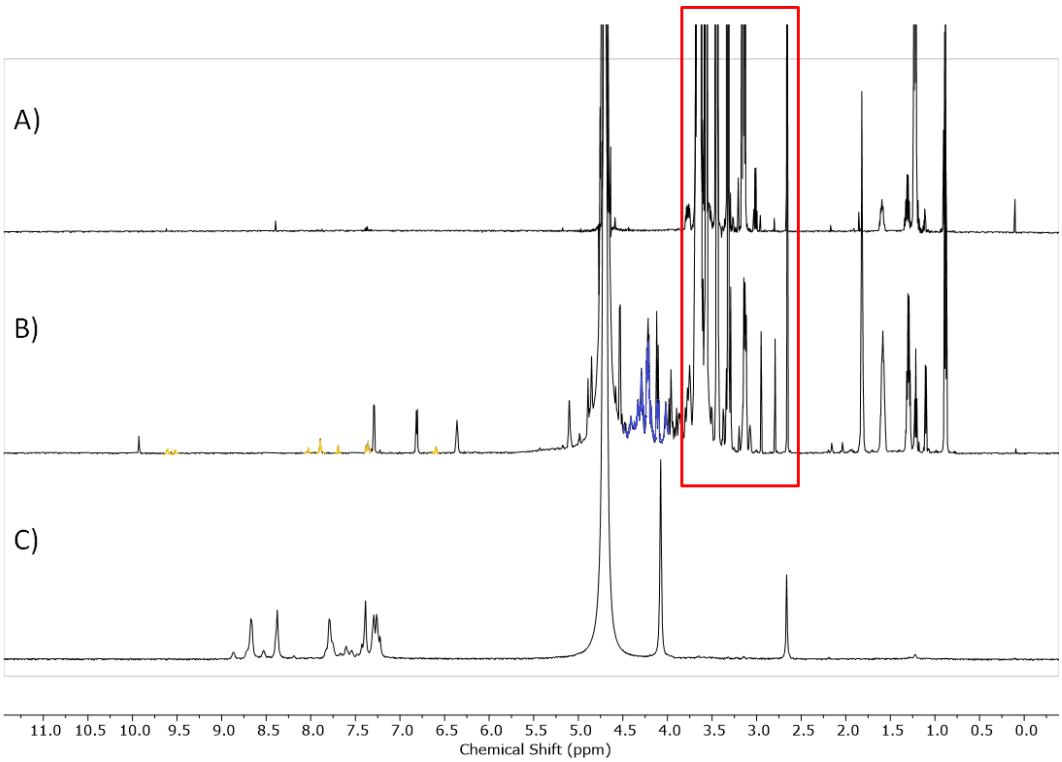


*Figure 3.22. ^{1}H NMR (500 MHz, D_2O) spectra show attempted click reaction on **C8** with PEG-N_3 following the two reactions conditions A) reaction mixture with conditions Cu/C, Et_3N, DCE, B) reaction mixture with conditions $CuSO_4$, $C_6H_7NaO_6$ and C) free **C8**. Red box highlights excess of PEG-azide.*

It is thought that the factors hindering reaction progression are likely either related to unwanted coordination of the ligand and the catalyst, or lack of solubility. There are limited examples of post-assembly coordination cage click chemistry; with two key examples provided by Zhou et al., in 2010 and by Chang and co-workers in 2022.[32,33] Zhou and co-workers faced similar issues with the traditional coupling conditions, $CuSO_4$ and sodium ascorbate, where they discuss the catalysts ability to reduce their labile metallo-structure. Alternatively they utilised $[Cu(CH_3CN)_4]PF_6$ which they found more promising although failed to definitively characterise their resultant structure with NMR and instead relying on gel permeation chromatography to determine size. Interestingly, Chang and colleagues utilised CuBr/PDMETA; a catalyst combination normally associated with atom transfer radical polymerization.[32]

In conclusion, although the click reaction between **C8** and **6** may have partially worked under 'Conditions B' given unreliability of the transligation reaction to form **C8** this route was not further explored.

X-Ray Crystallography

The Lusby group has previously published the crystal structure for **C1** containing ReO_4^-, the cold analogue of nuclear medicine radioisotope [^{99m}Tc]TcO_4^-(Figure 3.23).[1] This previous crystal structure was obtained by the slow diffusion of acetone into a solution of **C1** in MeCN which had been saturated with [NMe_4][ReO_4]. Upon addition of a [NMe_4][ReO_4] to the cage, counter anion metathesis begins to occur where the associated NO_3^- are replaced with ReO_4^- which consequently changes the cage's solubility. Therefore, the conditions for crystallisation of **C1** with ReO_4^- were achieved using vapour diffusion of acetone and MeCN, these conditions would be unsuitable for the empty NO_3^- **C1**.

The crystal structure of the empty cage was obtained via slow diffusion of MeCN into a solution of cage in water, forming single orange prism crystals. To better understand the effect of guest encapsulation, a comparison is made between the cage's solid state structure in the presence and absence of the guest.

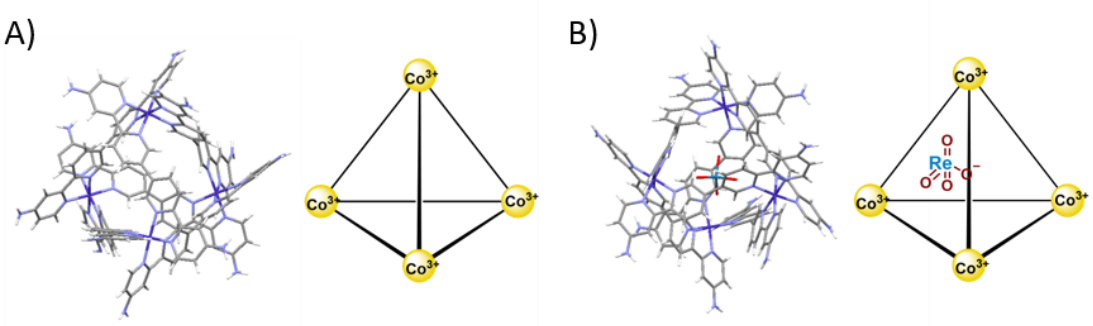


Figure 3.23. X-ray Crystallography Structures. A) Empty cage crystal structure with cartoon visualisation for clarity B) ***C1*** *with encapsulated ReO_4^- crystal structure (obtained by Dr William Grantham) with cartoon visualisation for clarity cage with guest.*

Upon ReO_4^- encapsulation the Co-Co distances range from 9.2 Å to 9.4 Å, whilst in the empty cage the Co-Co distances range from 9.3 Å to 9.4 Å. The distances imply the cage exhibits similar structure when empty and when existing as a host-guest complex. Although the slightly increased range in distances for the occupied cage may indicate the presence of a slight distortion upon encapsulation, the difference is considered to be within error of the technique and not significant.

Figure 3.24 shows a comparison of the dihedral angles between the bipyridyl rings of the two cages, A) in the presence of a guest and B) in the free cage. The average dihedral angle in the occupied cage was 65 ° whereas it was only 55 ° in the empty cage, this further indicates a slightly increased level of cage contortion upon encapsulation of the guest. It is not uncommon for a host to adapt in conformation when forming the host-guest complex and can be indicative of strong binding interactions enabling cage distortion.[34]

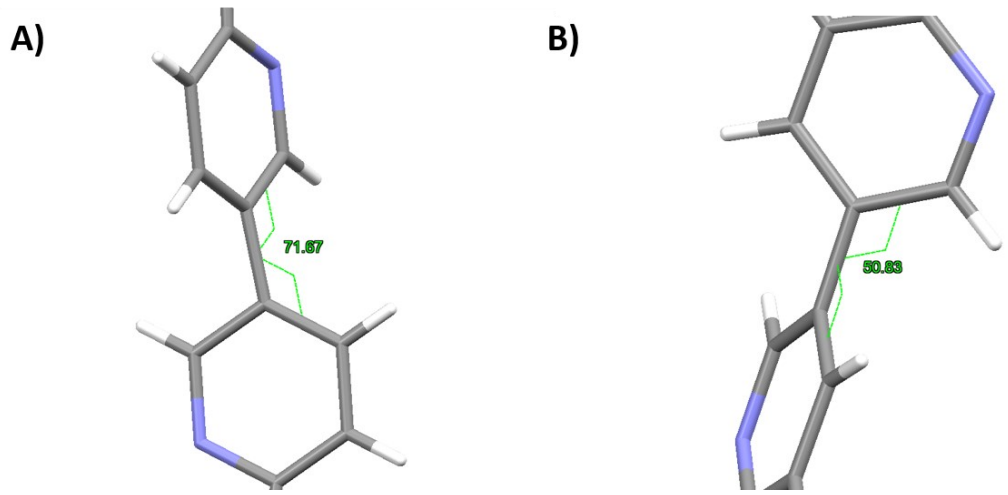


Figure 3.24. Visualisation the dihedral angles within of the two crystal structures as measured in Mercury 2021.3 (CDCC) A) filled cage B) empty cage.

3.2.6 Non-covalent Cage Functionalisation

The bioconjugation strategies until this point have focussed on covalently derivatising the outside of the cage. However, targeting groups can be associated to the cage via non-covalent interactions such as is seen in the self-assembly of nanoparticles using transformable peptide monomers (TPMs).[2] Given the highly cationic nature of the cage, it was thought that electrostatic interactions could be used to associate an anionic targeting group with the cage. This would be achieved via an anion metathesis where the NO_3^- counter anions are replaced by R^- moieties, where R is representative of a targeting groups such as a peptide fragment connected to an anionic terminus (Figure 3.25).

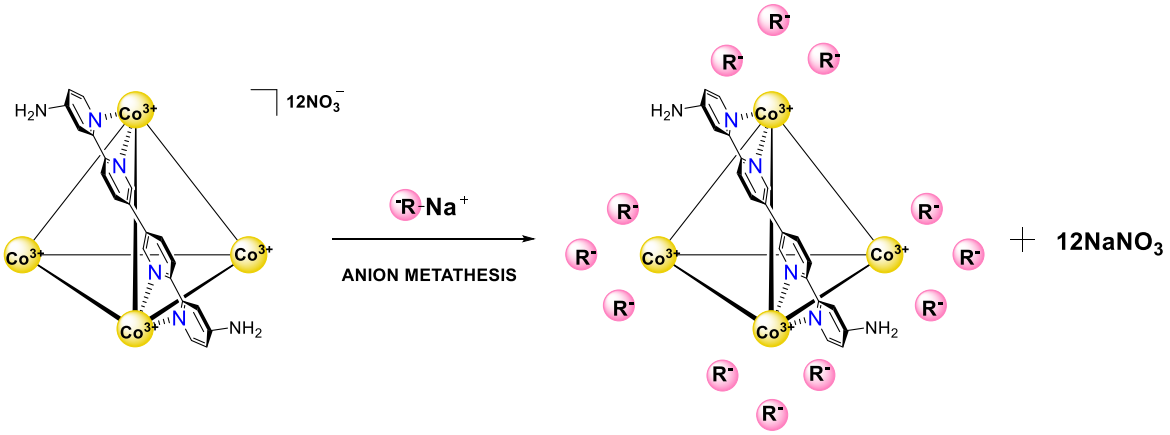


Figure 3.25. Visualisation of anion metathesis as a route to non-covalent bioconjugation, where R^- is representative of a biologically relevant group such as a peptide.

In an effort to draw a direct comparison between non-covalent and covalent bioconjugation, a synthesis was designed that utilised the same PEG-azide intermediates that had been previously used for covalent functionalisation (Figure 3.26). There was literature precedent for similar synthetic routes and thus this was chosen as a concise method to obtain the anionic R group, whilst utilising the intermediates already possessed.[35] The route progressed well with a high overall yield.

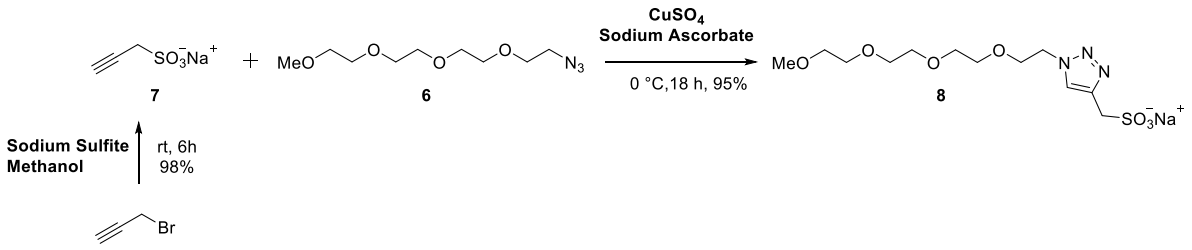


Figure 3.26. Reaction Scheme for the synthesis of anionic R groups.

Following the synthesis of **8**, a simple anion metathesis was attempted (Figure 3.27), where an excess of **8** was reacted with a solution of **C1** in D_2O aided by 20 min of sonication followed by heating at 60 °C for 18 h. A dark precipitate formed, which was isolated onto a glass filter and eluted into MeOH.

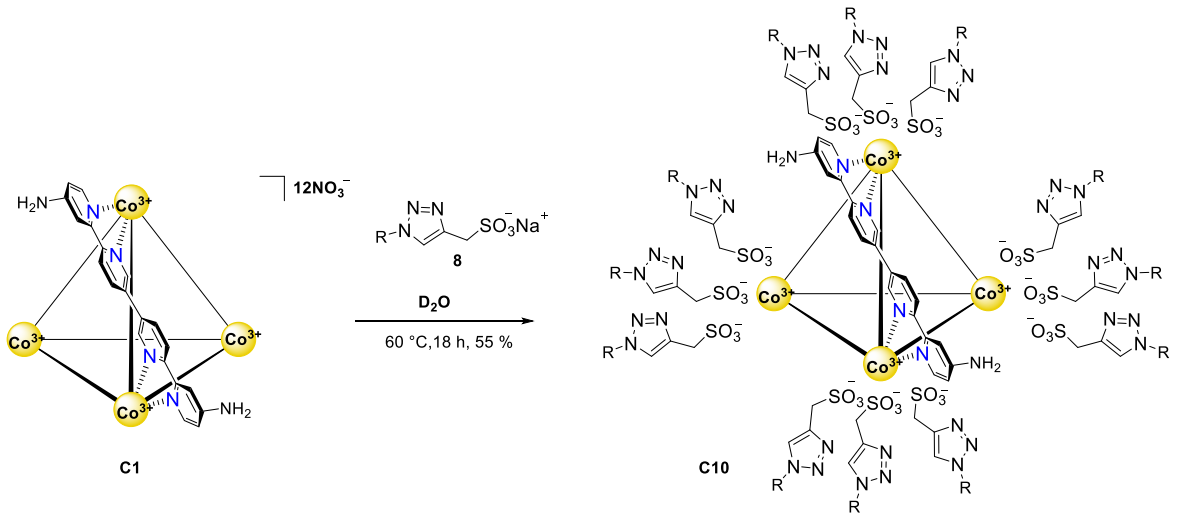


Figure 3.27. Reaction scheme for anion metathesis conditions.

Figure 3.28 shows a comparison of the ^{1}H NMR spectra following anion metathesis of **8** and the free **C1** to form **C10**. The cage signals have moved significantly downfield upon metathesis, and the distinctive triazole proton resonance can be seen as a yellow singlet at 8.05 ppm. Pleasingly, the triazole:cage signals integrate 1:1, which strongly infer that all twelve NO_3^- counter anions have been replaced with **8**. Moreover, the CH_2 of the sulfonate can be seen in blue at 4.25 ppm and integrates at a ratio of 2:1 in relation to the cage peaks, further evidencing the successful metathesis.

Unfortunately **C10** appeared to have limited solubility in MeOD, and dissolution was insufficient for further ^{13}C and DOSY NMR characterisation. Multiple attempts were made to gain the mass spectrum of the cage using ESI MS, however issues were encountered with ionisation and no multiply charged species could be observed.

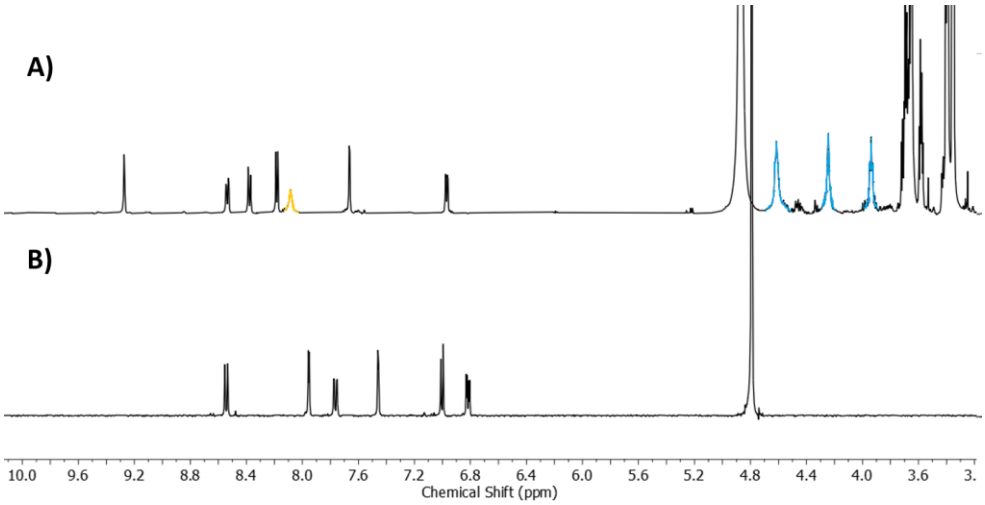


*Figure 3.28. ^{1}H NMR (500 MHz, D_2O) spectra exemplifying PEG-triazole sulfonate metathesis with **C1**. A) **C10** PEG-triazole sulfonate complex (MeOD) B) **C1**.*

Relying on the ^{1}H NMR spectrum of **C10**, the complex was mixed with ReO_4^- to explore whether the cage would still be capable of $[^{99m}Tc]TcO_4^-$ for imaging purposes. Figure 3.29 shows a titration of $[NBu_4][ReO_4]$ into **C10**. It is noted that unlike **C1**, **C10** is insoluble in water and thus the titration was carried out in MeOD, whereby the change in solvent could alter the binding by reducing the hydrophobic effect.

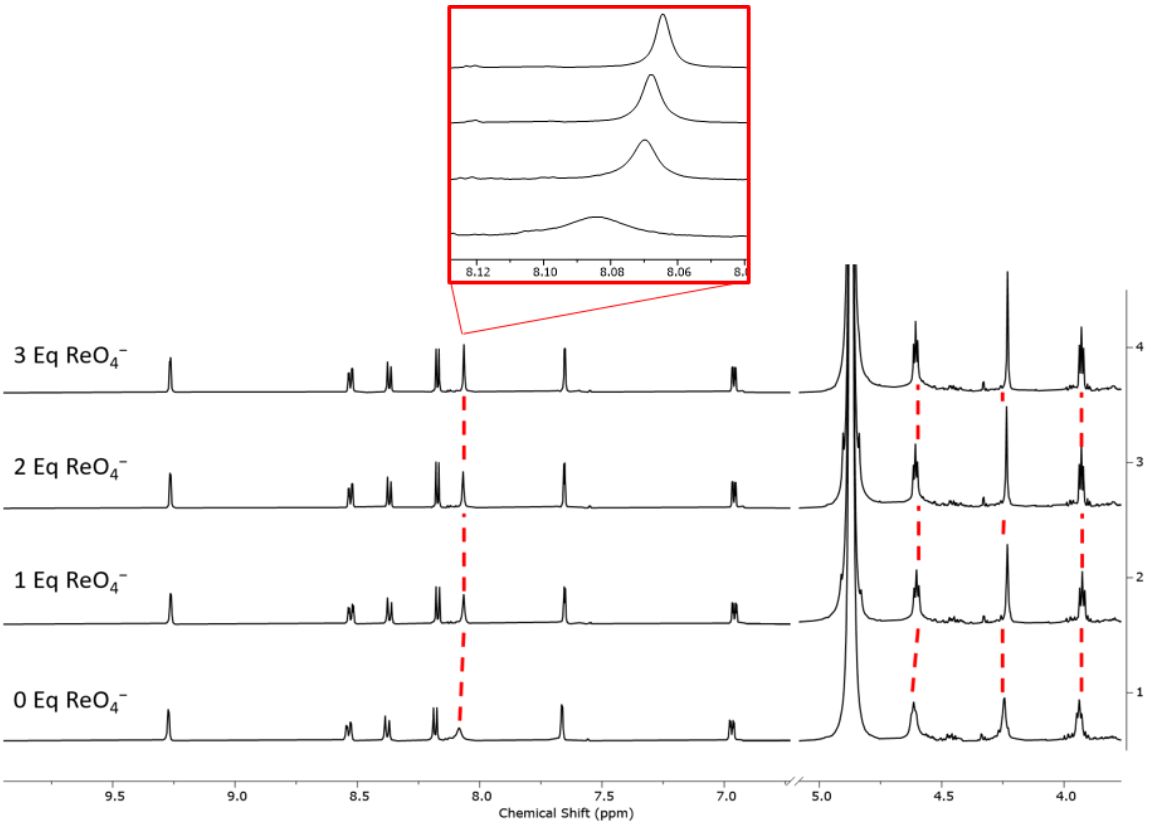


Figure 3.29. ^{1}H NMR (500 MHz, MeOD) titration of ReO_4^-[NBu_4ReO_4] into PEG anion cage.

Interestingly, there appears to be minimal perturbations in the cage signals upon addition of [NBu_4][ReO_4] implying it is not a strong guest in the cage under the conditions of the titration experiment. However, the signals associated with PEG-triazole sulfonate counter anions appear to narrow and sharpen, which can be followed by the dashed red lines in Figure 3.29. Specifically, the broad singlet at 8.1 ppm in empty **C10**, corresponding to the triazole-proton, appears to considerably narrow and shift downfield slightly upon the addition of just 1 equivalent of [NBu_4][ReO_4]. This could be indicative of the ReO_4^- associating to the outside of the cage partially displacing the external **8** counter anions.

3.2.7 Investigating Cage Stability

The design of **C1** was the result of the cage needing to possess two key features; water solubility and stability in biological conditions. Whilst the **C5**, **C6** and **C7** cage are all water soluble, their tolerance to bio-abundant species such as salts and reductants must be determined. Previous work by Lusby and co-workers evidenced the requirement of the *para*-positioned amine group for reducing cage lability.[1] In comparison, **C2** although stable in the presence of NaCl, was found to quickly disassemble in the presence of the bioreductant, glutathione.[19]

NMR Stability Studies

Most obviously, the cages are required to possess good stability in the presence of salts. Na^+ and Cl^- are the most prevalent ions with concentrations of ~140 mM and ~105 mM, followed by HCO_3^- at 27 mM and HPO_4^- at 2 mM *in vivo*.[36] Consequently, the stability of the cages in phosphate buffered saline (PBS), made up in D_2O, was assessed using 1H NMR spectroscopy (Figure 3.30). As can be seen, both **C5, C6** and **C7** show good stability over a period of 12 hours.

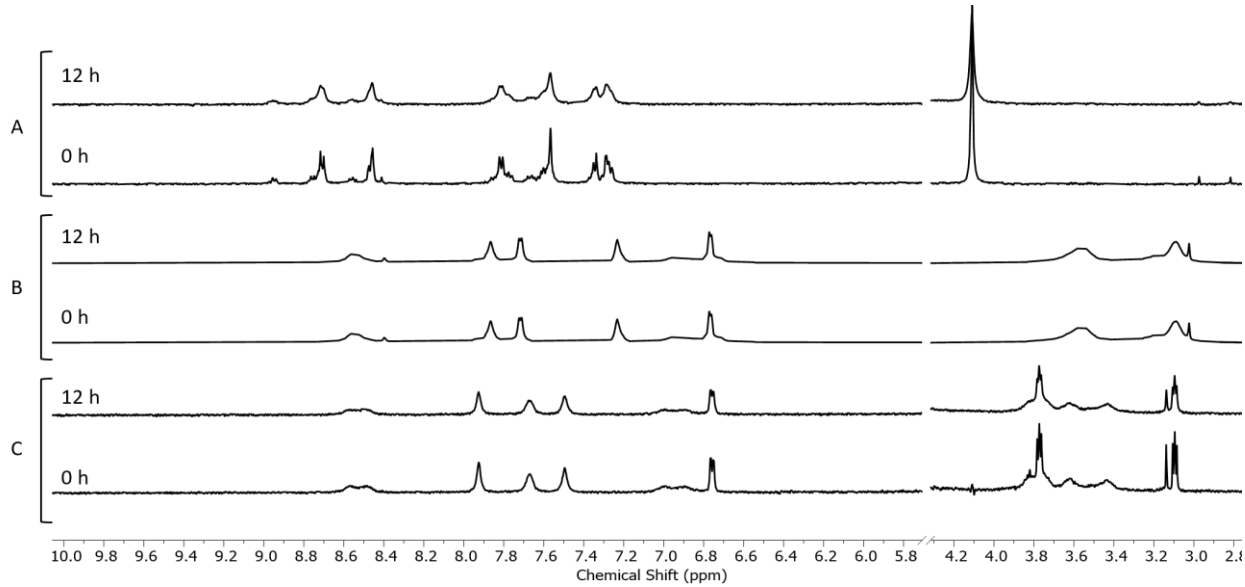


Figure 3.30. Partial ^1H NMR (500 MHz, D_2O, 0.1 M PBS) exhibiting stability of cages in d-PBS over time. A) ***C5****, B)* ***C7*** *C)* ***C6***

Following the successful stability tests in the presence of the salt, the cages were subsequently tested for their tolerance to glutathione. Glutathione is a bioreductant with a known mechanism for reducing and breaking down therapeutics, such as platinum based chemotherapy agent oxaliplatin.[37] When the **C5** was treated with excess glutathione, it was observed that just 30 min after the addition of glutathione the concentration of the **C5** in solution has significantly decreased and after 6 hours, only 10% of the initial concentration remains in solution (Figure 3.31). This indicates that although **C5** is slightly more stable than **C2** in the presence of glutathione, which disassembled just one hour after the addition, it is not as stable as **C1**. Again, it was hypothesised that the Co^{III} cage is being reduced by the glutathione,[38] evidenced by the formation of glutathione disulphide. It is likely that the more labile Co^{II} cage then disassembles.

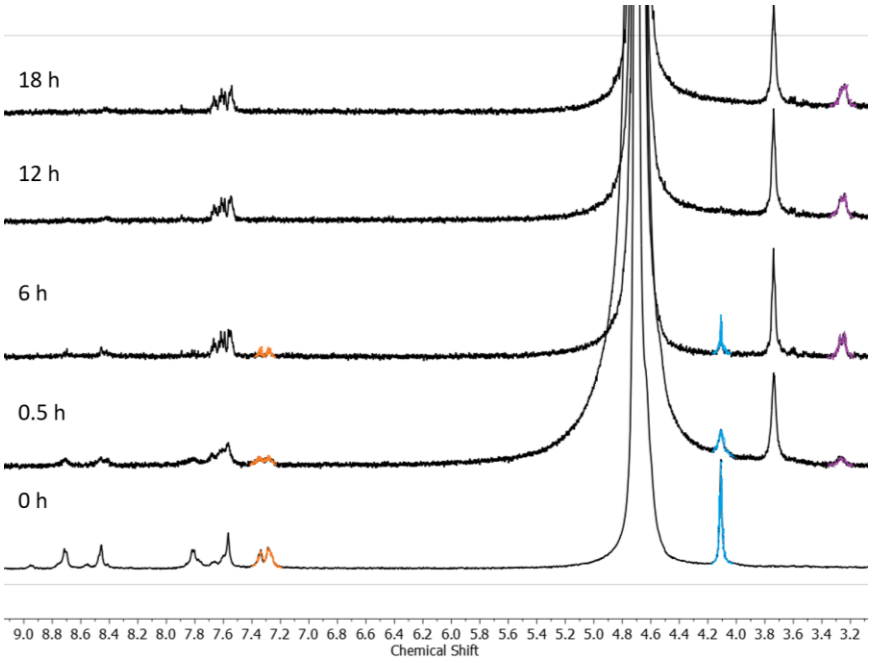


*Figure 3.31. Partial ^{1}H NMR (600 MHz, D_2O) of **C5** (1 mM) in presence of 10 Eq of glutathione over time. Aromatic cage signals shown in orange, methyl cage signal shown in blue and glutathione disulphide shown in purple.*

Similarly to **C1**, both the **C6** and the **C7** appeared stable in the presence of glutathione (Figure 3.32). The spectra were standardised against the internal standard and the cage resonances remained constant throughout the course of the experiment. This implies the two cages were unsusceptible to the reduction mechanism which had cause **C5** disassembly, presumably as a result of the increased strength of the Co^{III}-Ligand bond.

Observing the **C7** aromatic signals upon the addition of glutathione a distinct shift in chemical shift for all the signals is observed. The resultant signals are not analogous to the free ligand, which is insoluble in water, but instead imply an interaction between the cage and either the glutathione or another species present. Although glutathione is too large to entirely fit within the cage it is hypothesised that a strong external interaction could lead to a significant perturbation of the signals such as is observed in Figure 3.32A.

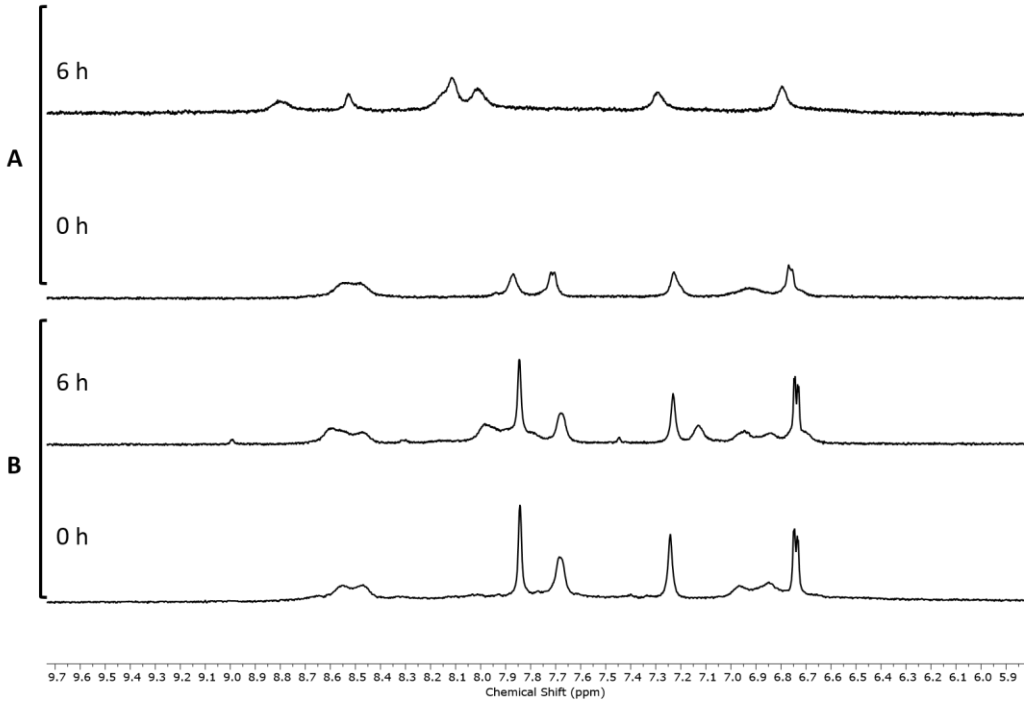


*Figure 3.32. Partial ^{1}H NMR (500 MHz, D_2O) of **C6** and **C7** in presence of 10 Eq of glutathione over time. A) **C7** and B) **C6**.*

Radiochemical Stability Studies

While the previous NMR studies provide an indication of the free cages ability to withstand various bio-abundant species, namely salt and glutathione, they fail to describe how the different conditions would affect the host-guest complex. Following the previous methodology developed by our collaborators at the University of Hull, the percentage of $[^{99m}Tc]TcO_4^-$ encapsulated over a series of different conditions was measured, including varying solvents, salt concentrations and transferring the radiolabelled cage into mouse serum.

Figure 3.33 shows the percentage of $[^{99m}Tc]TcO_4^-$ still encapsulated within the four cages following the transfer of a small aliquot (10 µL) of host-guest complex into mouse serum (100 µL) over a range of cage concentrations. The amount of $[^{99m}Tc]TcO_4^-$ remaining inside the cage was then monitored over the course of an hour using thin-layer chromatography.

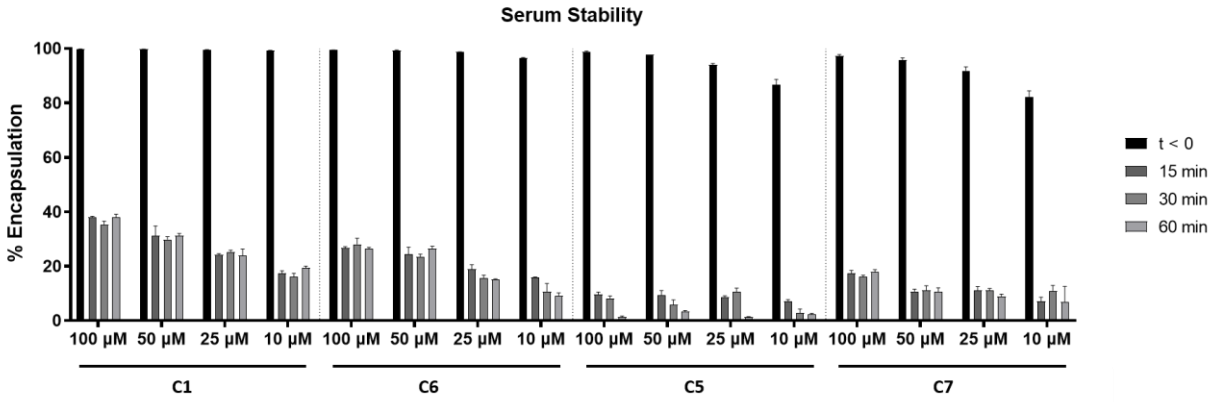


Figure 3.33. Stability of cage systems in mouse serum monitored by evaluation of the percentage of intact host-guest complex over time (% [^{99m}Tc]TcO_4^- encapsulated).

It is observed across the four data sets, that a sharp decrease in the percentage of encapsulated [^{99m}Tc]TcO_4^- occurs upon the initial transfer of the cage into the mouse serum. Given that the concentration of the host-guest complex considerably decreases as it is diluted upon addition to the serum, a series of controls were carried out to determine whether the perceived change in percentage encapsulation was a simple dilution effect. Figure 3.34 shows the percentage of [^{99m}Tc]TcO_4^- still encapsulated within **C1** following transfer of 10 µL of host-guest complex in a solution of water (100 µL), analogous to the methodology for the serum stability studies. Interestingly, 100% retention is observed throughout the course of the experiment and across the four investigated concentrations, indicating that the observed decrease in percentage encapsulated [^{99m}Tc]TcO_4^- in serum is a result of the species present in the serum.

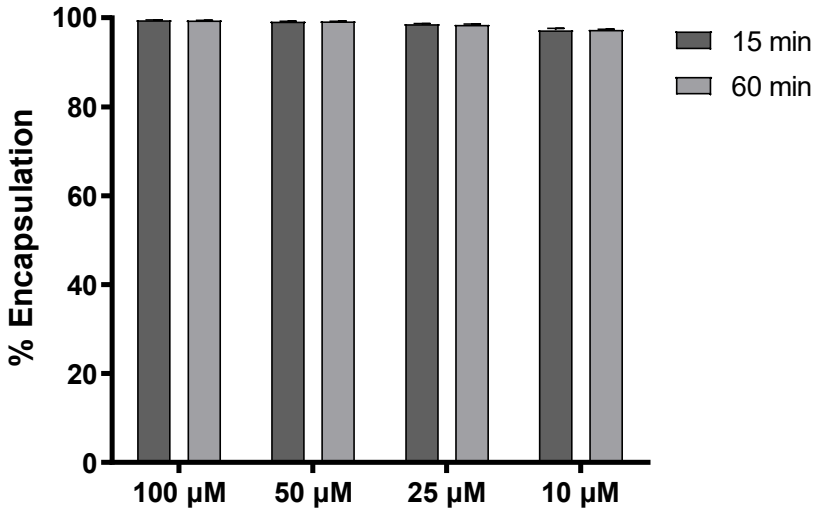


*Figure 3.34. Effect of water dilution on **C1:** [^{99m}Tc]TcO_4^- host-guest complex measured over the course of 1 hour, (% [^{99m}Tc]TcO_4^- encapsulated).*

Observing the data in Figure 3.33, the parent **C1** retains the highest percentage of bound [^{99m}Tc]TcO_4^-, with **C5** shows the lowest radioisotopic labelling when transferred into mouse serum. Given the proven instability of **C5** in the presence of just the bioreductant glutathione, it is unsurprising that the cage would readily disassemble over the course of the experiment losing the encapsulated guest. It was previously hypothesised that since **C5** remained partially intact upon the addition of glutathione for up to 6 hours, it may prove suitable for shorter imaging studies. However, the serum stability studies showed that after 1 hour **C5** had retained only a small percentage of encapsulated [^{99m}Tc]TcO_4^- and thus it was deemed too labile under biological conditions for imaging studies.

Initially it was hypothesised that due to the presence of a *para*-positioned amine group in both **L6** and **L7**, both the **C6** and **C7** would exhibit similar stability to the previously synthesised **C1**. The retention of encapsulated [^{99m}Tc]TcO_4^- of the **C6** cage was confirmed upon transfer into mouse serum, which showed 30% of the initial bound [^{99m}Tc]TcO_4^- compared to the 35% by **C1**. However, **C7** exhibits lower stability in serum, with only 20% encapsulated [^{99m}Tc]TcO_4^- remaining after 1 hour in serum at 100 µM. Assuming that the host-guest chemistry is very similar for these three cages, it would appear that the pendent groups are potentially facilitating guest release, possibly via a cage breakdown mechanism.

In an attempt to further investigate the stability of the host guest complexes, a series of competitive experiments were run, shown in Figure 3.35. The experiments were run identically to the previous serum stability studies, except the concentration of the host-guest complex was 2[EC_{95}] ensuring maximum encapsulation for each individual cage. The t=0 values are not included here for clarity but are defined as 100% prior to the transfer into the competitive solution.

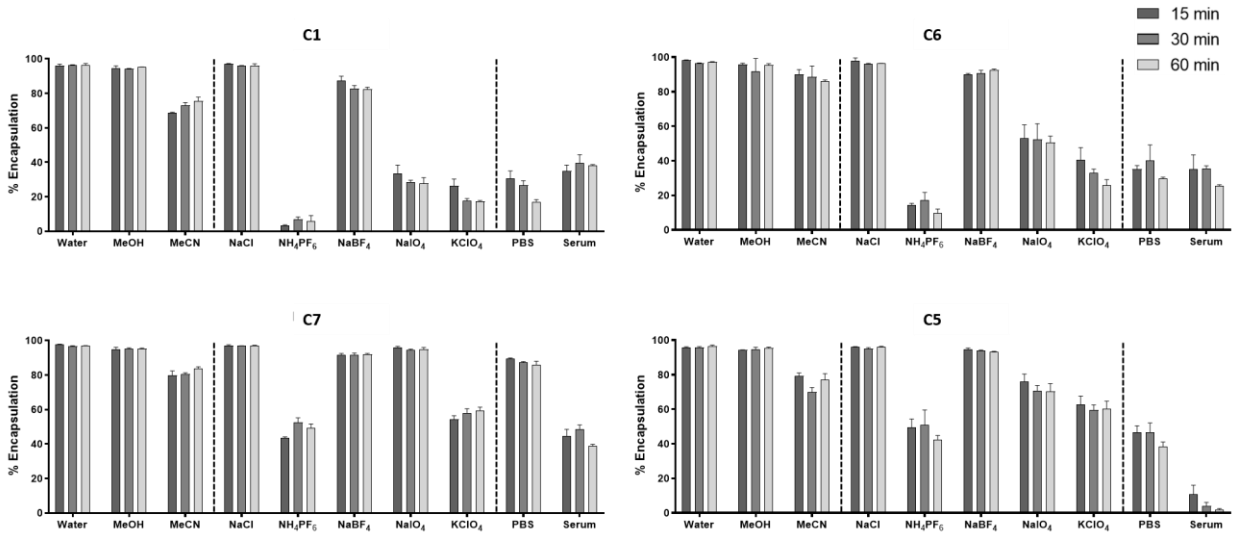


Figure 3.35. Monitoring the percentage of encapsulated [^{99m}Tc]TcO_4^- within the four cages in the presence of competitor species, over the course of an hour.

Pleasingly, the four host-guest complexes showed good stability in the presence of NaCl which supports the previous NMR studies. However, when transferred into solutions containing anions that may bind the cage better a more significant decrease in the percentage of [^{99m}Tc]TcO_4^- encapsulated was observed. Presumably, as the host-guest complex is transferred into a solution containing a high concentration of competitive guests and a low concentration of the original host-guest complex the equilibrium is biased towards the binding of the competitor. This hypothesis is further validated by the retention of [^{99m}Tc]TcO_4^- within the four cages when exposed to $NaBF_4$ where BF_4^- is known to be a weak guest.

The apparent increased retention of guests within the **C7** and **C5** is likely a result of them being investigated at higher concentrations due to their respective EC_{95} values (Figure 3.46).

3.2.8 Guest Anion Scope Study

The shape and structure of the cationic **C1** provides it with complementarity to tetrahedral anions. As mentioned, [^{99m}Tc]TcO_4^- is used most commonly with nuclear medicine and subsequently was the focus of the previous host-guest research and SPECT study. However, as nuclear medicines develop, more radioisotopes become available providing access to a range of different imaging techniques including PET. Within the context of radiolabelling coordination cages, the suitable radio-isotopic guests must be easily synthesised from the form in which they are eluted from the generator as well as having a proficient half-life and good interactions with the cage. This section focusses on investigating a series of potential "cold" guests to identify which would be suitable for bio-imaging purposes.

Guests inside ***C1***

A summary of the binding constants for a series of anions within both **C2** and **C1** is shown in Table 3.1. Previously, SO_4^{2-} had been investigated for its host-guest chemistry with **C2** due to its prevalence *in vivo* and therefore if they were found to be too strongly binding it might competitively block the intended guest from the cavity.[19] Fortunately, SO_4^{2-} was found to only be an incredibly weak guest within the cage presumably due to its large free energy of solvation. ClO_4^- was investigated to probe the nature of the cavity, due to its reduced hydrophilicity, in comparison to SO_4^-, it was found to be a relatively good guest within the cage.[19] The remaining investigated anions contain elements that have isotopes which are of interest for nuclear medicine such as fluorine, rhenium and iodine.

Table 3.1. Binding constants of various anions and C2 and C1s in D_2O. Errors are estimated to be less than 10%.

Guest	K_A (M^{-1})	
	C2	**C1**
SO_4^{2-}	100*	
BF_4^-	500*	1600
ClO_4^-	7100*	21000‡
ReO_4^-	61000*	46000‡
PF_6^-	91000*	
SO_3F^-		8700
IO_4^-		4000

** data collected by Michael Burke ‡ data collected by William Grantham*

PET imaging studies allow the generation of a series of images over the course of the experiment, in comparison to SPECT which traditionally only generates a single image at the end of the experiment. Therefore PET imaging would allow for tracking of the biodistribution of the host-guest complex throughout the course of the experiment. Most commonly, ^{18}F is used for PET imaging, due to its positron emitting property and its medically favourable half-life of 109.8 min, as any alternative radioisotopes to fluorine exhibit drastically shorter half-lives.[39]

The synthesis of $[^{18}F]BF_4^-$ has been developed and optimised over the last decade with initial methodology focussing on isotopic exchange, leading to suboptimal radiochemical yield, and more recent synthetic processes including the direct radio-fluorination of BF_3.[39] Given the accessibility of $[^{18}F]BF_4^-$ the host-guest chemistry within **C1** was investigated, however analogously to **C2** it was determined to be a relatively weak guest, with a K_a of just 1600 M^{-1}. Figure 3.36 and Figure 3.37 shows the NMR titration data and fitted 1:1 binding isotherm for this experiment.

Given the ability of competitor guests to occupy the cavity in place of the intended guest, it is concluded that a potential guest must have a binding constant comparable to ReO_4^- (cold version of

[^{99m}Tc]TcO_4^-) in **C1**. Consequently BF_4^- was determined to be too weak of a guest. This is further supported by the previous displacement assays, which showed that the percentage of encapsulation [^{99m}Tc]TcO_4^- remained high in the presence of excess BF_4^-.

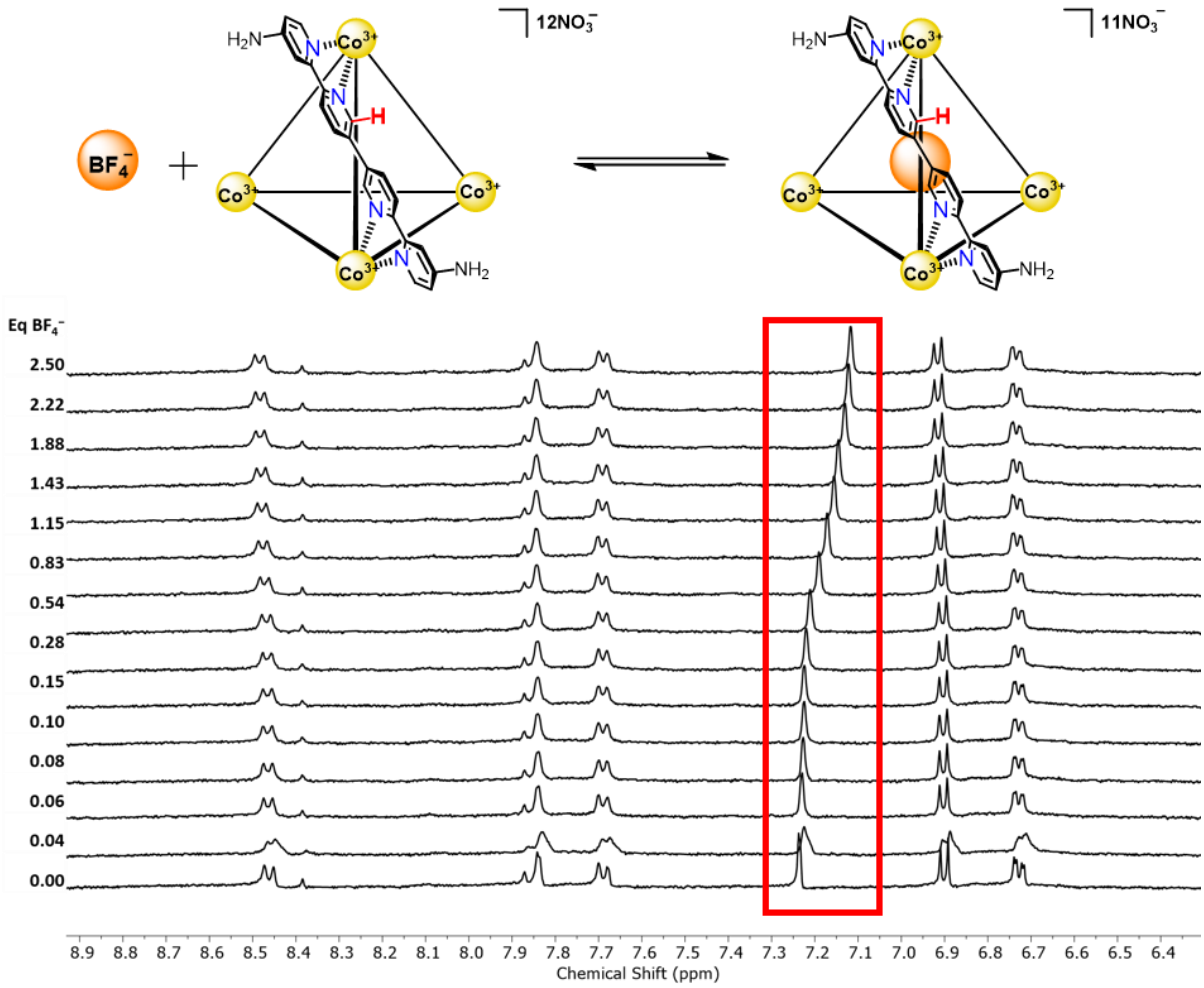


*Figure 3.36. ^{1}H NMR (400 MHz, D_2O) spectroscopic data for titration of KBF_4 into **C1**. Movement of internal proton highlighted in red.*

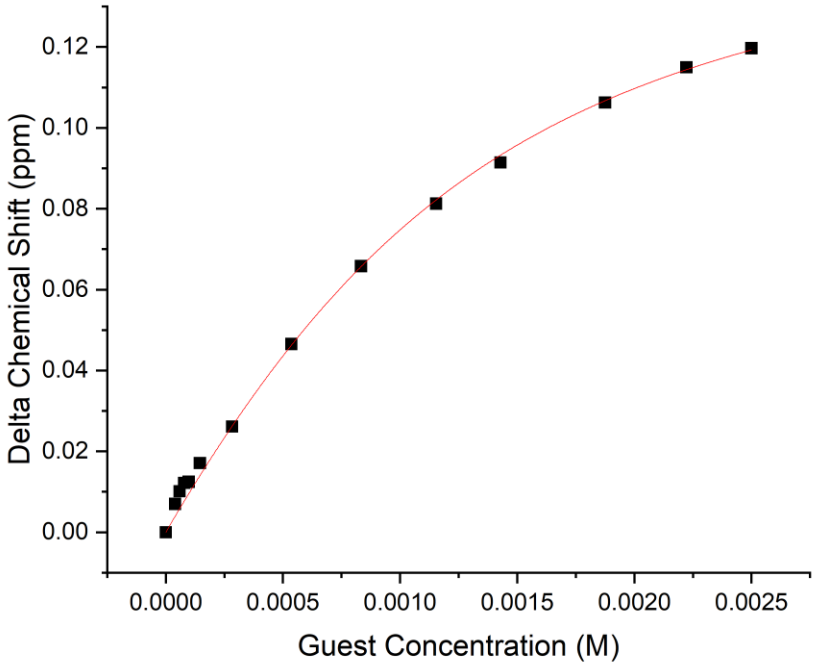


*Figure 3.37. Fitted curve for guest concentration versus change in chemical shift (ppm) for the internal proton of **C1** upon titration with BF_4^-.*

Disappointingly, NMR titration data indicated that SO_3F^- was also not interacting sufficiently with **C1** for radiolabelling purposes, exhibiting a K_a of 8700 M^{-1}. Presumably the hydrophilic nature of these anions leads to a higher desolvation cost offsetting the hydrophobic effect within the cage.

Alternatively, PF_6^- is much less hydrophilic favouring binding within the hydrophobic cavity, exhibiting a binding constant of 91,000 M^{-1} within **C2**. It is hypothesised that this increased K_a could also be a result of the larger anion size, providing better complementing the cage cavity. Whilst there has been some interest into $[^{18}F]PF_6^-$ in nuclear medicine as a blocker of uptake via the human sodium-ion symporter, the research has been largely hindered by the comparatively complex radiosynthesis of the anion.[40]

The anion IO_4^- is investigated due to positron emitting ^{124}I isotope, which is gaining traction as a potential PET tracer.[40] Alternatively to the majority of guests discussed until this point, IO_4^- was found to bind within **C1** via slow as opposed to fast exchange, as shown by the titration data in Figure 3.38. This slow exchange isn't caused by very high affinity, as the binding constant here was relatively low (4,000 M^{-1}) and similar to other anions that are in fast exchange (e.g., SO_3F^-). Rather, the slow exchanges indicates some form of energy barrier between the free and bound state that is greater than merely the breakage of the non-covalent interactions. When considering host-guest interactions for biomedical applications, large association constants (K_a > 40,000 M^{-1}) are required in order to

counter balance guest leakage under biological conditions. The number of equivalents of guest added during the course of the titration were limited due to precipitation of stock solutions.

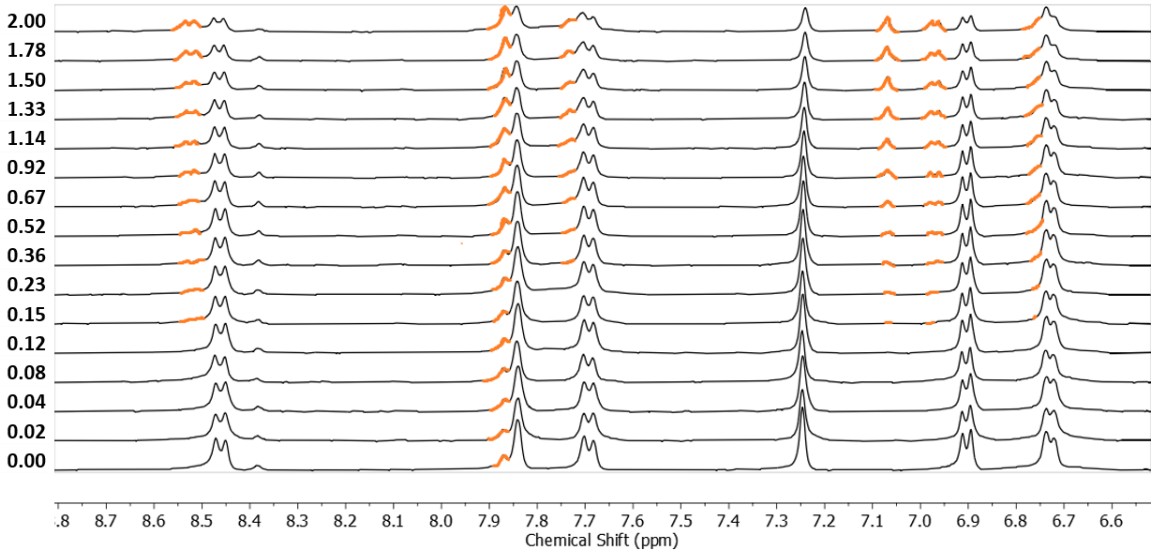


*Figure 3.38. ^{1}H NMR (400 MHz, D_2O) spectroscopic data for titration of IO_4^- ($NaIO_4$) into **C1**, with host-guest complex in slow-exchange shown in orange.*

Further ^{1}H NMR experiments monitoring the interactions between **C1** and IO_4^- over time indicated a level of instability of the host-guest complex. Figure 3.39 shows the increase of the slow-exchange host-guest complex (shown in light blue) as the equivalents of guest goes up. However, upon re-recording the same spectrum after 12 hours the slow-exchange species is lost and only "free" cage can be observed.

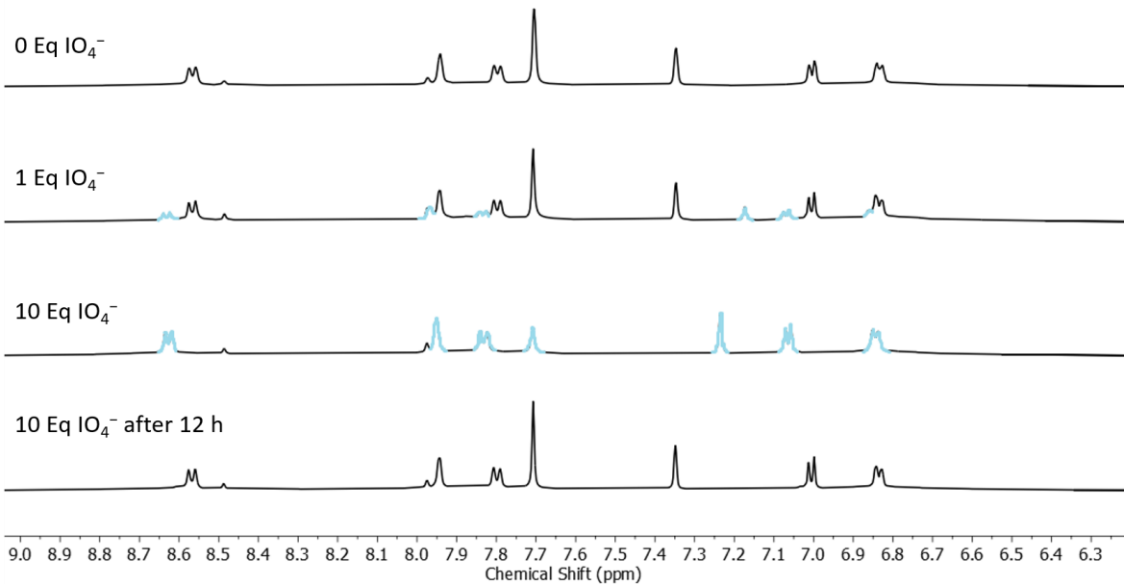
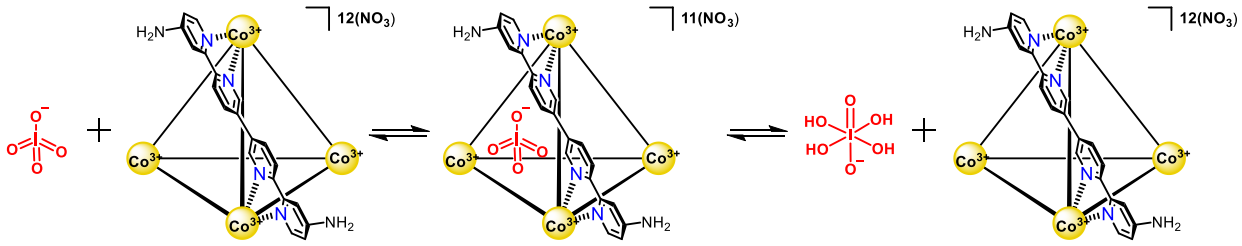


*Figure 3.39. ^{1}H NMR (400 MHz, D_2O) spectroscopic data for the generation of a slow-exchange species following addition of IO_4^- into **C1** and its stability over time.*

It is thought that over time metaperiodate (IO_4^-) can undergo hydrolysis to *ortho*-periodate ($H_4IO_6^-$) where the larger anion is no longer a guest within the cage.[41] The hydrolysis mechanism is not normally favoured under ambient conditions and the cage could be creating a microenvironment to catalyse the reaction. This could account for the relatively low binding constant within the cage where the equilibrium not only exists between the free cage and the host-guest complex but also with the hydrolysis product, as shown in Scheme 3.15**Error! Reference source not found.**.



*Scheme 3.15. Proposed equilibrium for metaperiodate encapsulation by **C1** and hydrolysis to non-encapsulated* ortho-*periodate.*

Functionalised Cage Binding

The previous section highlighted three anionic guests of **C1** which have potential in biomedical radiolabelling. Firstly, ReO_4^- was used as a cold analogue of [^{99m}Tc]TcO_4^- which has previously been encapsulated within **C1** and used in SPECT imaging studies. Secondly, PF_6^- which has shown to be a very strong guest within **C2** but investigations were limited due to the complex radiosynthesis of

[^{18}F]PF_6^-. Finally, IO_4^- which incorporated the interesting ^{124}I isotope and exhibited a slow-exchange binding interaction with **C1**.

Presumably, given the similarity in cavity size, shape and hydrophobicity between **C1** and the three novel cages (**C5**, **C6** and **C7**), they should exhibit analogous host-guest chemistry. Table 3.2 shows a summary of the binding constants for the **C6** and **C7**, as determined by ^{1}H NMR titrations of solutions of guest into cage, in unbuffered D_2O.

*Table 3.2. Summary of **C6** and **C7** binding constants*

Guest	K_a	
	C6	**C7**
ReO_4^-	2200	1000
PF_6^-	12000	N/A
IO_4^-	140,000	N/A

Interestingly, IO_4^- exhibited a slow-exchange interaction with **C6** similar to **C1** as shown by the generation of a second series of cage signals upon addition of the guest, visualised by the NMR data in Figure 3.40 by red peaks. The broadness and overlap of the signals makes the identification of the occupied host less clear, consequently the internal proton is monitored as highlighted by the red box and the concentration determined thereof. Following the addition of 0.5 Eq of IO_4^-[$NaIO_4$] the spectrum shows the presence of two sets of signals, implying a 1:1 ratio of host-guest complex to free cage. This is seen most clearly by the two internal proton signals (one black and one red) highlighted in the blue box, following the addition of half an equivalent of guest. By monitoring the increase in concentration of the host-guest complex throughout the titration a binding association 140,000 M^{-1} was estimated, which is more than sufficient for biomedical radiolabelling purposes. Due to the broadness of the signals the calculated K_a is used only as an indicator as to whether IO_4^- would be a good guest.

Intriguingly, the regeneration of free cage over time that had been observed for the $IO_4^-\subset$**C1** complex was not observed for **C6**. This was hypothesised to be a result of the different microenvironments surrounding the cages caused by their alternative external functionality.

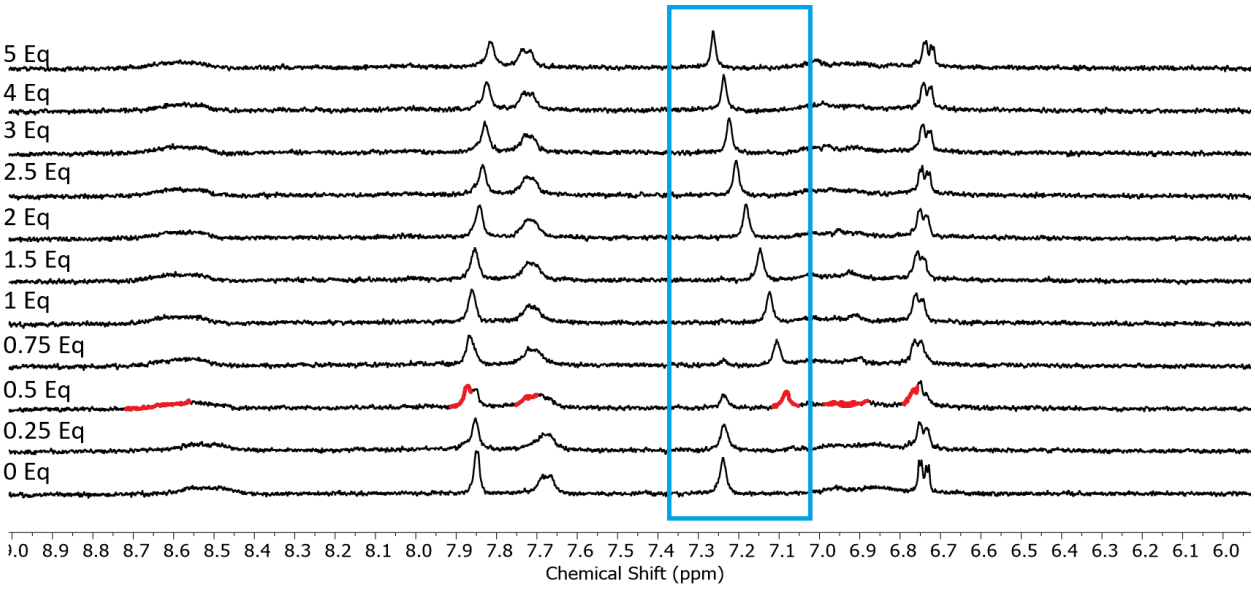


*Figure 3.40. ^{1}H NMR (400 MHz, D_2O) spectroscopic data for titration of IO_4^- into **C6**, with host-guest complex in slow-exchange shown in red and the internal proton signal highlighted in the blue box.*

Unfortunately, the proton signals for **C7** were too broad for the accurate determination of a binding constant with IO_4^-, as can be seen by the ^{1}H NMR titration data in Figure 3.41. Prior to the addition of any guest, the spectrum is broad but the cage is clearly visible, however upon the addition of just 1 Eq of $NaIO_4$ the cage signals have broadened and are considerably harder to isolate from the noise and baseline of the spectrum. While the concentration of free cage is observed to be decreasing upon the addition of the guest, it is impossible to accurately determine whether a new series of signals corresponding to the host-guest complex are appearing. It is thought that the guest could be interacting dynamically with the cage causing an expansion of the already broad signals leading to a complex NMR spectrum.

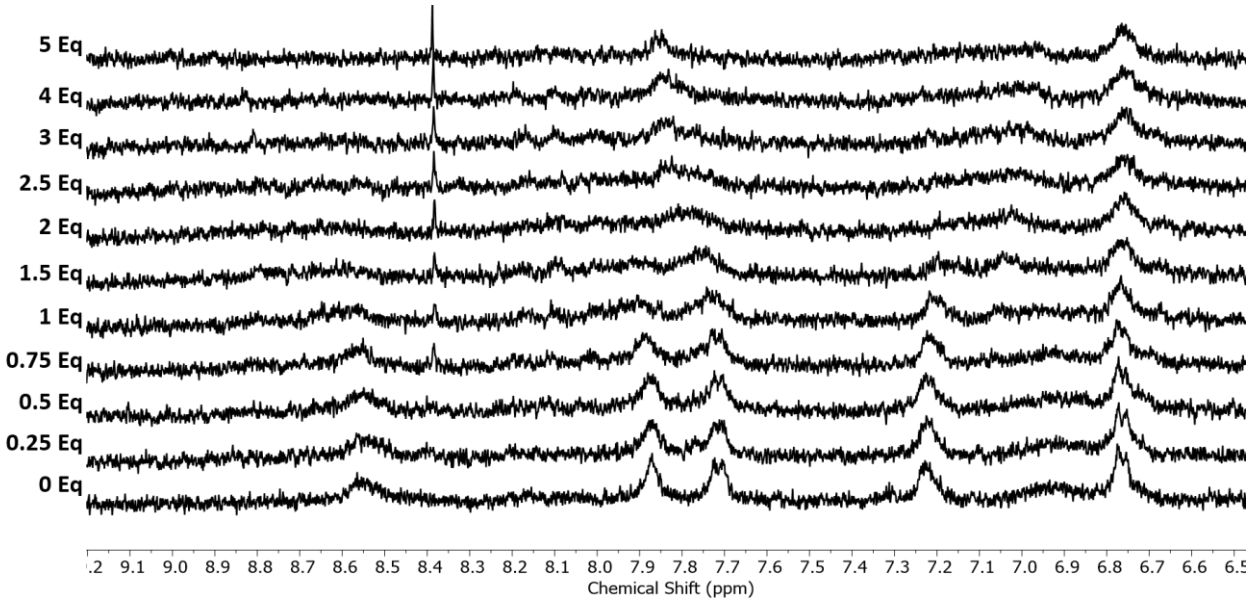


*Figure 3.41. ^{1}H NMR (400 MHz, D_2O) spectroscopic data for titration of IO_4^- [$NaIO_4$] into **C7**, the peaks are too broad to accurate depict or to calculate a binding constant.*

Figure 3.42 and Figure 3.43 show the ^{1}H NMR titration data between PF_6^- with **C6** and **C7** respectively. Interestingly, the binding constant for **C6** is considerably lower at only 12,000 M^{-1} compared to **C2** which is 91,000 M^{-1}. Potentially, the externally functionalised ethanolamine groups are creating a hydrophilic microenvironment where the alcohol groups are hydrogen bonding to the solvent molecules which are blocking the hydrophobic, PF_6^- anion.

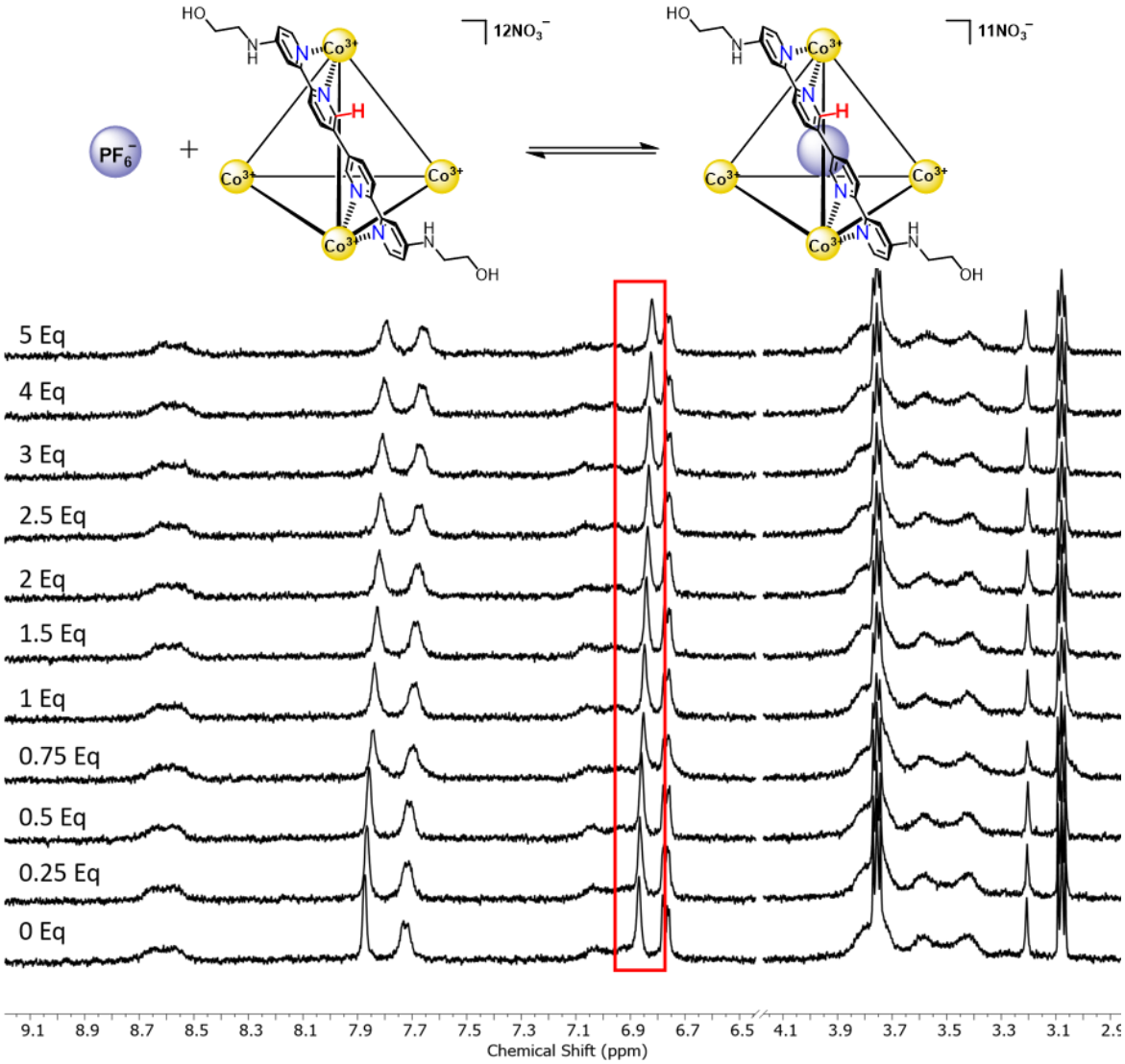


*Figure 3.42. ^{1}H NMR (400 MHz, D_2O) spectroscopic data for titration of PF_6^-[KPF_6] into **C6**, with host-guest complex in fast-exchange as highlighted by the shifts of the internal proton, shown in the red box.*

Fortunately, the **C7** titration data with PF_6^- was slightly less broad than with IO_4^- and thus the binding could be more easily visualised. As shown in Figure 3.43. the internal proton, highlighted in orange, clearly shifts upon the addition of PF_6^- however it appears to become almost saturated after the addition of only 0.25 Eq of guest. When observing the other peaks there are less significant perturbations although it is noted that the signal at 7.7 ppm shifts first downfield upon the addition of 0.25 Eq guest and then proceed to shift upfield throughout the remainder of the titration. This could be an indication of multiple binding methods, for example if the cage becomes internally saturated and subsequently the remaining interactions are observed on the surface of the cage. An attempt was made to fit the movements of the internal (orange) proton signal, however the 1:1 binding method could not be applied and consequently no binding constant was determined. It is considered that inaccuracies in cage or guest concentration could hinder the ability to reliably fit the data.

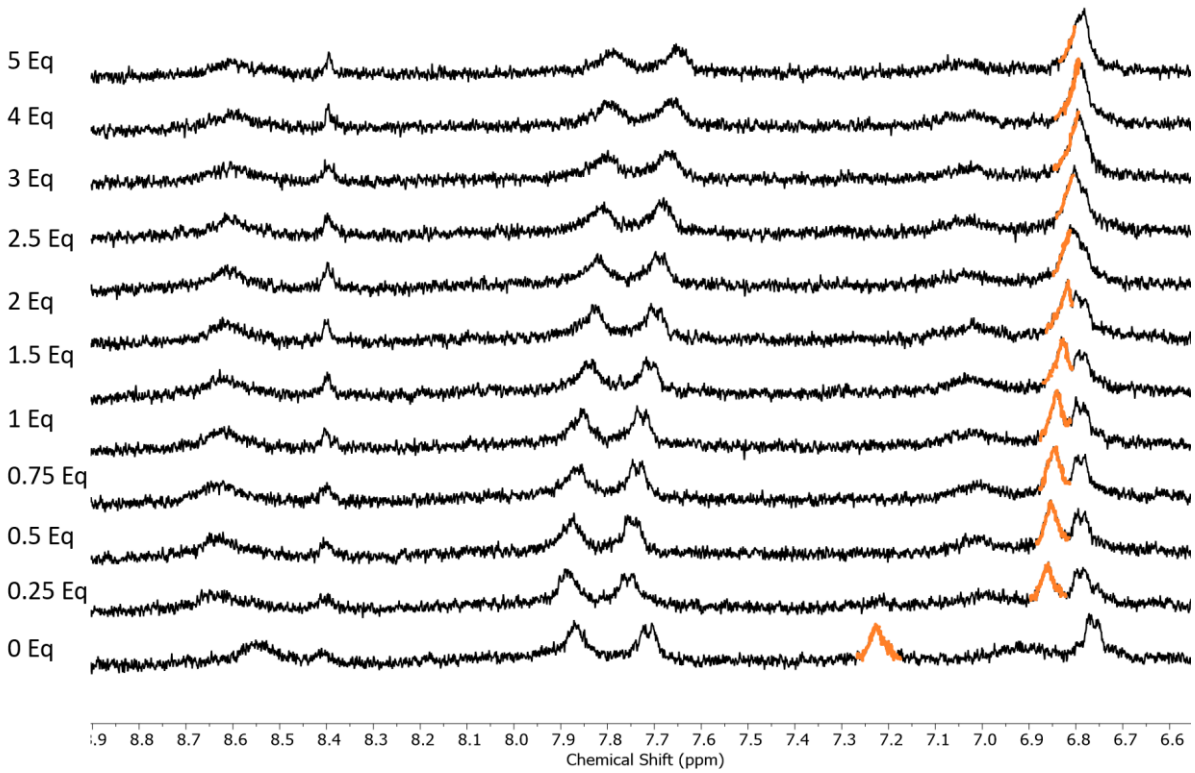


*Figure 3.43. ^{1}H NMR (400 MHz, D_2O) spectroscopic data for titration of PF_6^- [KPF_6] into **C7**, with host-guest complex in fast-exchange shown in orange although peaks are too broad to accurately depict.*

The phenomena of two different binding modes could also account for the shift movements observed for the titrations of [NBu_4ReO_4] with **C6** and **C7** as shown in Figure 3.44 and Figure 3.45 respectively, where upon the first addition of guest the cage signals shift significantly and then subsequently trend in the opposite direction throughout the course of the rest of the titration. The binding constants shown in Table 3.2 for ReO_4^- in the two cages are determined by removing the initial titration point and monitoring from the first addition, and are likely indicative of the guest interacting on the outside of the cage.

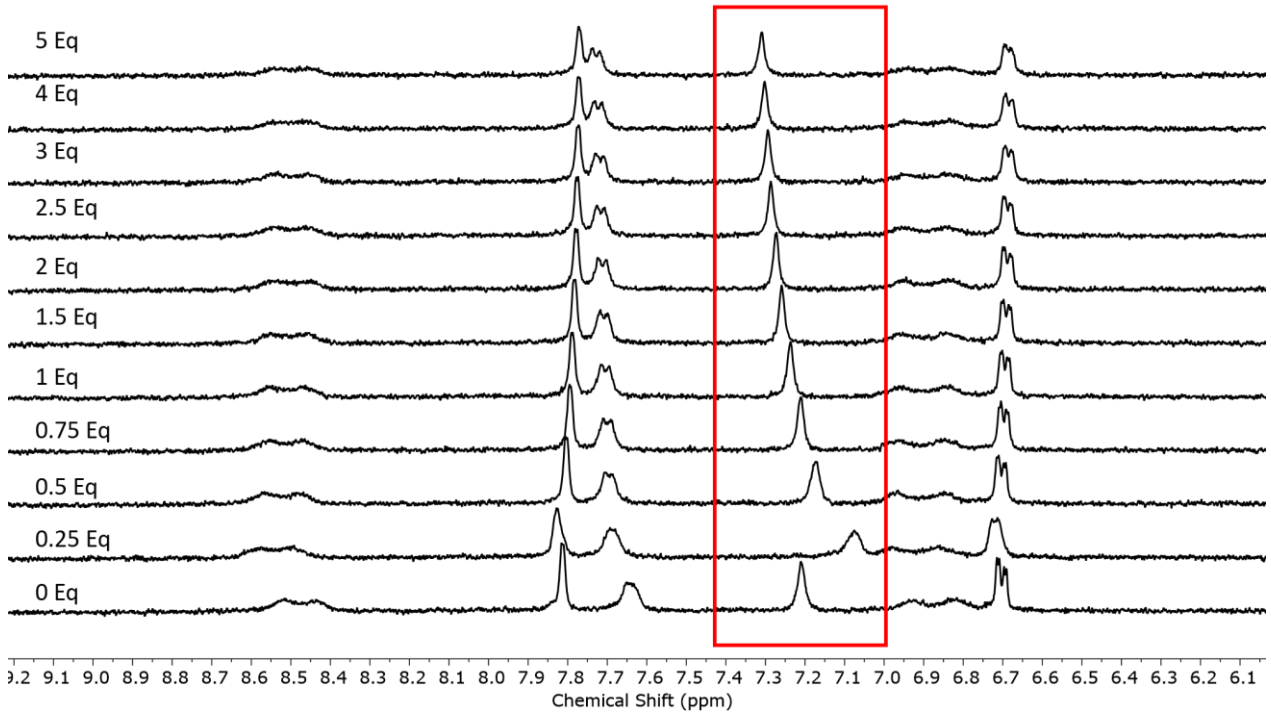


*Figure 3.44. ^{1}H NMR (400 MHz, D_2O) spectroscopic data for titration of ReO_4^- (NBu_4ReO_4) into **C6**, with host-guest complex in slow-exchange shown in red, which highlights the internal proton.*

Interestingly, upon the addition of 0.25 Eq ReO_4^- to **C7**, shown in Figure 3.45, it is possible that two sets of cage signals are observed as shown in the green box which contains two signals corresponding to the internal proton. The formation of a slow exchange host-guest complex which subsequently shifts throughout the titration is the same observation that is made for the binding for IO_4^- in **C6**. [NBu_4ReO_4] was used as the metal salt containing ReO_4^-, however it is known to have relatively low solubility in D_2O and subsequently the guest solution was vigorously shaken prior to each addition. It is thought that this method could lead to higher additions of guest than intended and could account for the early internal saturation of the cage. This hypothesis could also account for the strange trend observed when binding the ReO_4^- within **C6**.

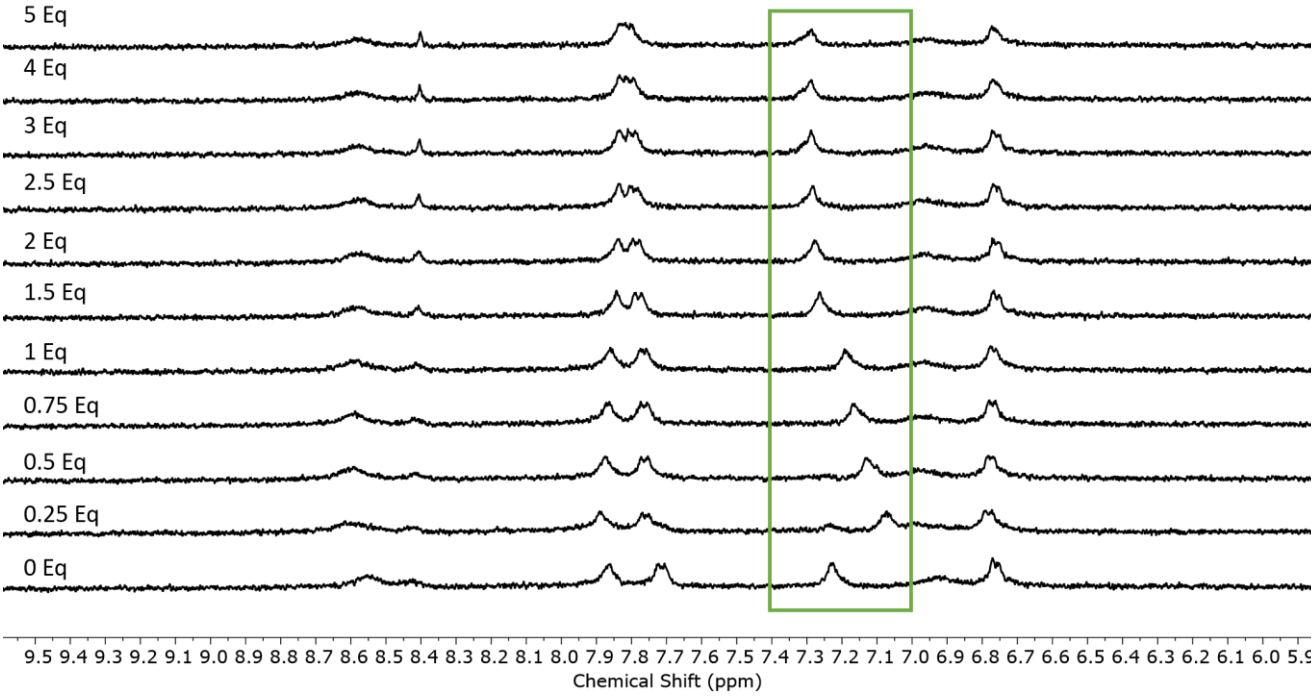


*Figure 3.45. ^{1}H NMR (400 MHz, D_2O) spectroscopic data for titration of ReO_4^- [NBu_4ReO_4] into **C7**, with host-guest complex in slow-exchange shown in green, which highlights the internal proton.*

Given the broadness of the titration spectra and the unpredictable trends observed, the calculated association constants (Table 3.2) are purely used as a preliminary indicator of which anions might make good guests within **C6** and **C7**. Further optimisation of the titration procedure is required to overcome the issues associated with signal breadth and host-guest complex solubility.

Radiochemical Guest Encapsulation

Following the ^{1}H NMR titrations of the guests within the cages, our collaborators in the Archibald group investigated the encapsulation of [^{99m}Tc]TcO_4^- using the method previously described. The concentration of each cage required to bind both half and 95% the activity of 1 MBq of [^{99m}Tc]TcO_4^- is shown in Figure 3.46, as EC_{50} and EC_{95} values. Interestingly, **C1** showed the best encapsulation of the guest with an EC_{95} of just 1.6 µM and closely followed by **C6** with an EC_{95} of 4.4 µM. Given the other similarities between **C5** and the previously synthesised **C2**, it is perhaps unsurprising that the **C5** also exhibits a relatively high EC_{95} of 12.2 µM.

Unexpectedly, **C7** exhibited the lowest levels of guest encapsulation with an EC95 of 33.8 µM. This potentially evidences the differences between the cold analogue ReO_4^- and the radioisotopic anion [^{99m}Tc]TcO_4^-, given the ^{1}H NMR titration indicated that ReO_4^- was a good guest within the cage. Considering the similarity in cavity between the four cages and the fact that they differ only in their

external functionalisation, it is inferred that the protruding groups of the cages could be interacting with the potential guests and affecting the internal binding.

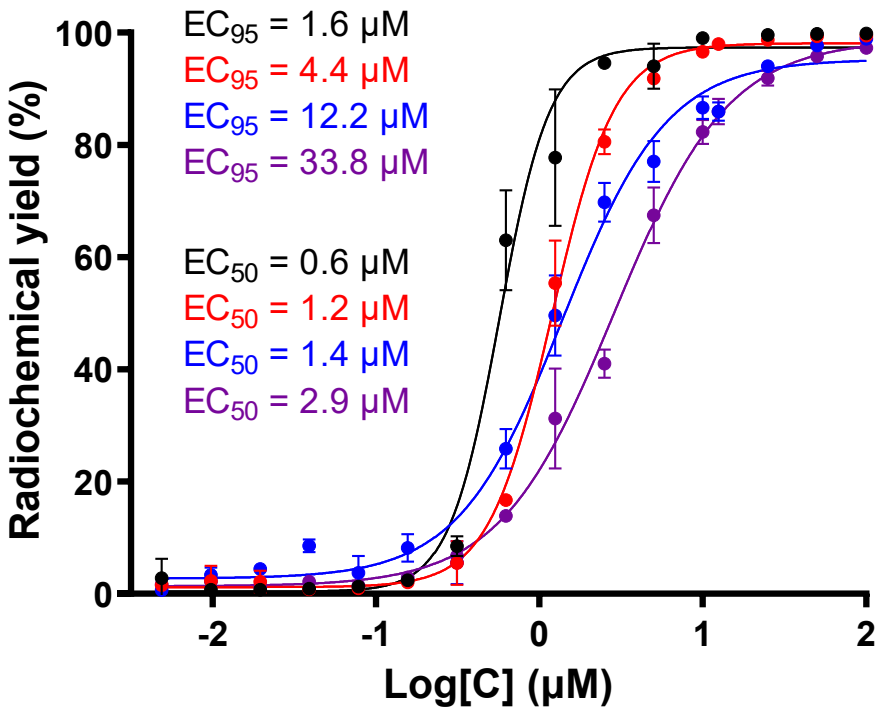


*Figure 3.46. Comparison of [^{99m}Tc]TcO_4^- encapsulation in **C1** (black), **C5** (blue), **C6** (red) and **C7** (purple) experiments performed by Dr Isaline Renard.*

3.2.9 Anion Metathesis to Control Guest Leakage

As evidenced by the loss of encapsulated [^{99m}Tc]TcO_4^- following the dissolution of the host-guest complexes into competitive medium, shown in Figure 3.35, there is a tendency for the guests to leak out of the cages. Even coordination cages and guests with strong binding constants exist in an equilibrium between host-guest complex and free host with disassociated guest. When considering host-guest chemistry *in vivo* the equilibrium between the assembled host-guest complex and the free parts is biased both because of dilution and competing analytes.

There are a couple of approaches for addressing the issue of guest leakage that can be broadly divided into two categories; assembling a 'closed' cage around the guest and modifying the cage post-assembly to prevent leakage. Fujita and co-workers reported the association of a series of capping ligands to the external of a coordination cage which covered the pores of the cage through which the guest typically travelled in and out (Figure 3.47).[42]

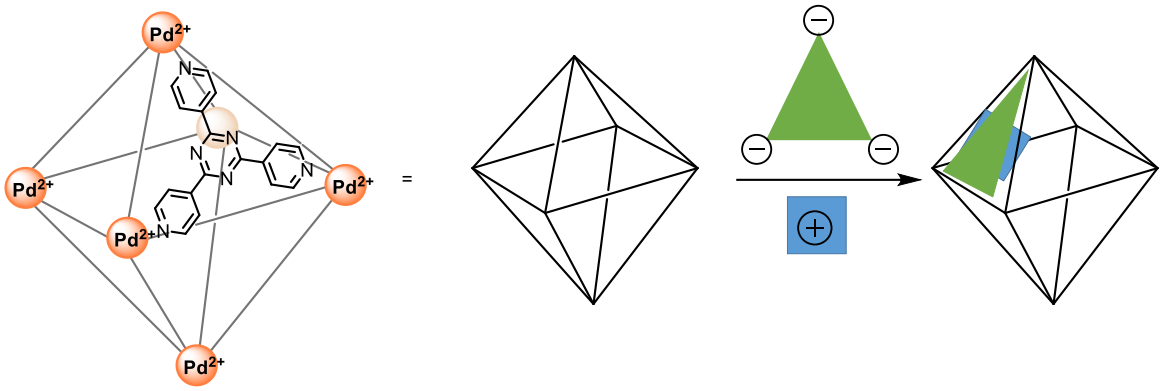


Figure 3.47. The portal capping of Fujita and co-workers Pd_6L_4 metallacage by four stoichiometric equivalents of a tripodal anion (green triangle), allowing for the encapsulation of four equivalents of a cationic guest (blue square).[42]

The technique of portal capping was largely employed by the Fujita group to enable the encapsulation of positively charged guests, it also served to illuminate the possibility of kinetically trapping a guest inside the cage using non-covalent external functionalisation. Therefore portal capping was attempted on **C1** to access non-equilibrium encapsulation of guests and prevent guest leakage; an imperative aim when regarding the biomedical applications of these cages for radiolabelling. Figure 3.48 shows a computer simulated molecular model of the free **C1** following by the capping of its portals by anionic trimesate **(L11)**.

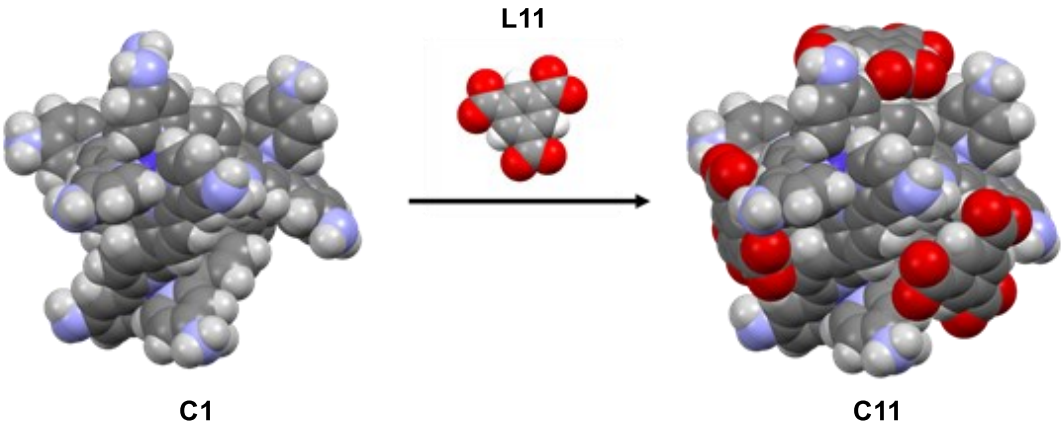


*Figure 3.48. SPARTAN simulated molecular model showing the capping of **C1** by four tripodal anionic ligands.*[43]

Sodium trimesate was identified as a capping ligand due to its good size complementarity with **C1**, as seen in Figure 3.48, as well as being commercially available. The capped cage (**C11**) was formed through the addition of 4 Eq of sodium trimesate to **C1** and characterised using ^{1}H NMR (shown in Figure 3.49) and DOSY NMR (Figure 3.50).

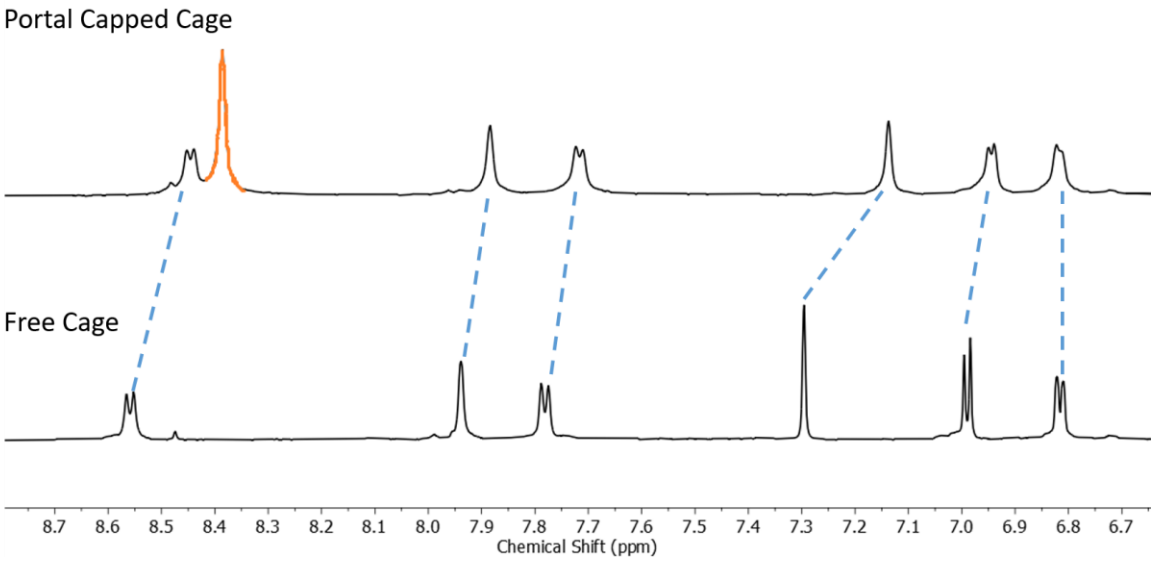


*Figure 3.49. ^{1}H NMR (400 MHz, D_2O) spectroscopic data of the formation of **C11**, formed by the addition of 4 Eq of sodium trimesate to **C1**. **L11** is highlighted in orange.*

C11 was found to have a diffusion coefficient of $9.73e^{-11}$ $m^2.sec^{-1}$ with a calculated hydrodynamic radius of 14.25 Å, which is slightly larger than the free **C1** (11.5 Å) and corresponds well with the molecular model. The concentration of **L11** appears in a slight excess of the cage, as shown by the orange signal in Figure 3.49, this is observed due to the relatively low water solubility of **C11**.

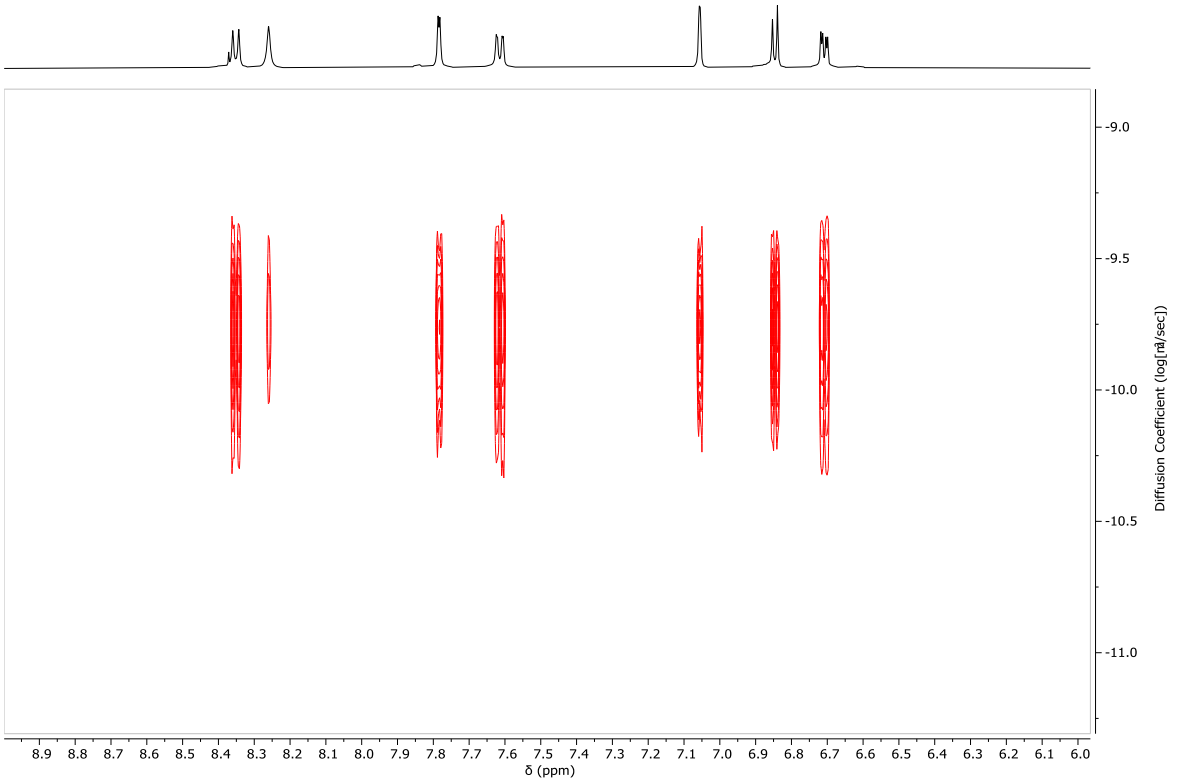


*Figure 3.50. ^{1}H DOSY NMR (500 MHz, MeOD) of **C11.***

To investigate the effect of the capping ligands on guest encapsulation ^{1}H NMR titrations were performed. A known guest was encapsulated within the cage and then **L11** was titrated into the solution and the cage signals monitored throughout. The anionic guests, PF_6^- and ReO_4^- were identified due to the usage of ^{18}F in PET imaging and [^{99m}Tc]TcO_4^-, which is structurally similar to ReO_4^- in SPECT imaging. Figure 3.51 shows the NMR spectra for first the encapsulation of ReO_4^- followed by the titration of **L11** into the solution to form **C11**.

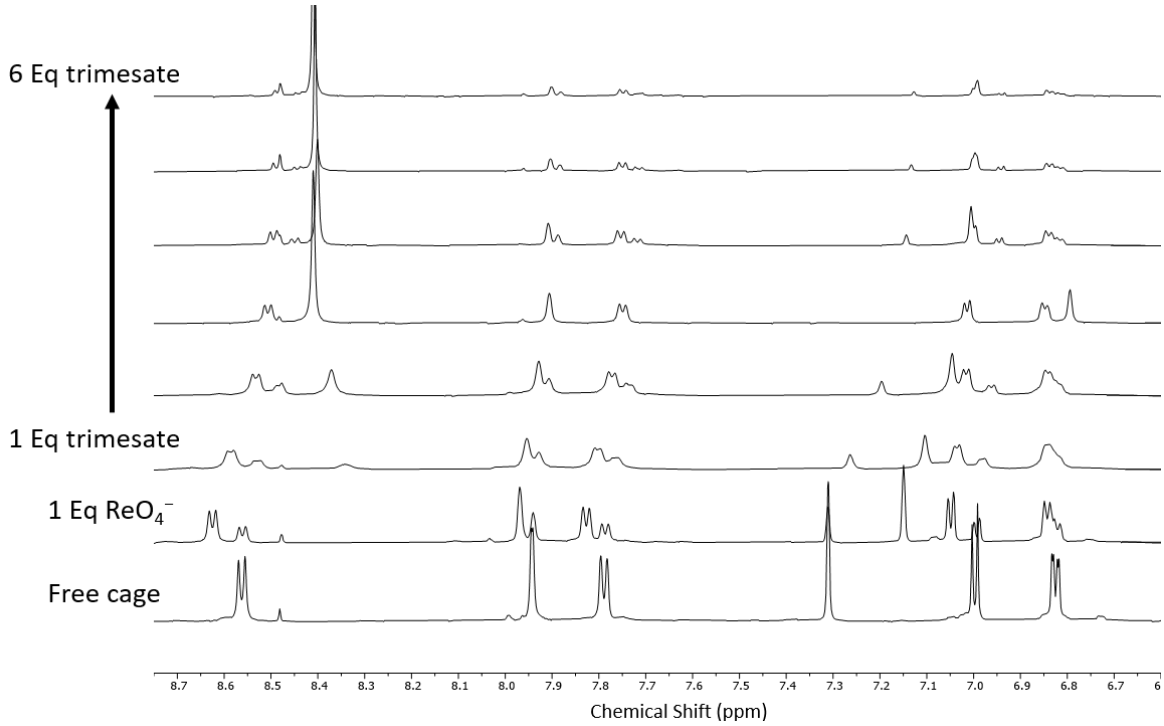


*Figure 3.51. ^{1}H NMR (400 MHz, D_2O) spectroscopic data for titration of **L11** into **C1** with encapsulated ReO_4^- to form **C11** with guest.*

Upon the addition of ReO_4^- two sets of signals can be observed corresponding to the host-guest complex and the free cage existing in slow-exchange with each other. The following titration of trimesate into solution caused a shielding effect for both sets of signals, indicating the capping ligand is interacting dynamically with both the occupied host and the free cage. It is also observed that throughout the course of the titration there is a considerable decrease in concentration of cage in solution, this is considered to be a result of the ion-pairing effect where the resulting complex is charge neutral leading to lower water solubility.

The same experiment was repeated using PF_6^- as the guest, as shown in Figure 3.52, starting with the generation of the host-guest complex followed by the titration of the capping ligand. Interestingly after the addition of 3 Eq of **L11** a second set of signals appears, indicating the formation of a slow-exchange species. Presumably, as **C11** is being formed it begins to hinder the movement of the guest in and out of the cage, slowing them down relative to the NMR time scale and thus generating a slow-exchange species.

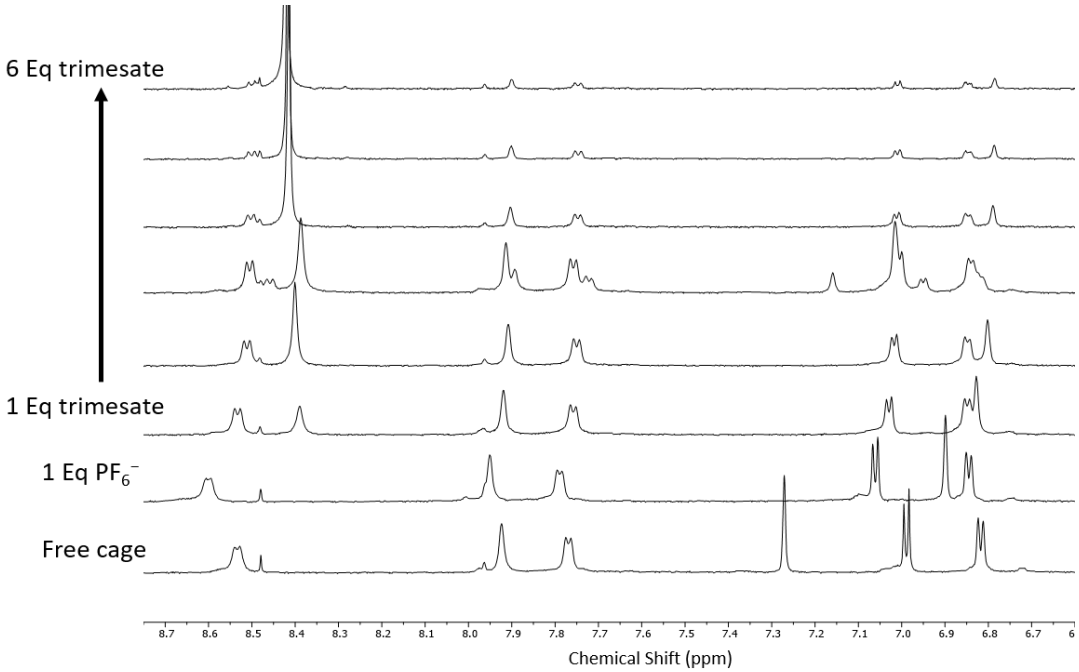


*Figure 3.52. ^1H NMR (400 MHz, D_2O) spectroscopic data for titration of **L11** into **C1** with encapsulated PF_6^- to form portal capped cage with guest.*

3.3 Conclusions

The aim of this research was to build on the previous research done by the Lusby and Archibald groups which evidenced that **C1** was capable of encapsulating $[^{99m}Tc]TcO_4^-$ and altering its biodistribution.[1] This work proposed that by modifying the external functionality of the Co^{III} tetrahedral cage the bioaccumulation could be changed and consequently controlled.

A series of ligands (**L5**, **L6**, **L7** and **L8**) were synthesised via a common intermediate ligand, **L9**, with particular versatility shown by a S_NAr reaction. Unfortunately, the synthetic process for **L8** was only optimised to achieve small amounts of the ligand, which slightly hindered further efforts to perform click chemistry on both **L8** and **C8**. Although, it was concluded that since the quaterpyridine based ligand was so highly coordinating to Cu, an essential catalyst for click-chemistry, post-assembly modifications was preferable as the ligands were fully saturated in the cage and therefore less interactive. Alternatively, the **C6** and **C7**s provided scaffolds for covalent bioconjugation via ester and amide bonds respectively.

Anion metathesis of the cage with PEG-triazole sulfonates was highlighted as a potential method for post-assembly non-covalent bioconjugation. NMR studies evidenced the new complex (**C10**) was readily formed and also exhibited different host-guest properties to the NO_3^- counter anion cages. The host-guest chemistry of the system was also altered upon portal capping of **C1**, as when the pores of the cage were blocked with **L11**, the guest anion PF_6^- changed from fast-exchange to slow exchange. Implying that PF_6^- is becoming kinetically trapped within capped cage and highlighting how external post-assembly modification could serve to hinder guest leakage.

The generation of an X-ray crystal structure of the empty **C1** visualised the level of strain and distortion the cage undergoes when encapsulating ReO_4^-. Following a series of host-guest titrations in the cages (**C1**, **C6** and **C7**), three radioisotopic anions were identified for further *in vivo* imaging studies. For SPECT imaging ReO_4^- (cold analogue of $[^{99m}Tc]TcO_4^-$) appeared to be a good guest in both the **C6** and **C7**s, although broadness of peaks and guest solubility slightly hindered accurate determination of binding constants. Moreover, PF_6^- (cold analogue of $[^{18}F]PF_6^-$) and IO_4^- (cold analogue of $[^{124}I]IO_4^-$) were identified as potential guests with PET imaging applications, which would allow for the tracking of the host-guest complex throughout the course of the imaging experiment.

Unfortunately **C5** was found to have insufficient stability under biological conditions, as evidenced by its disassembly in the presence of the bioreductant, glutathione and the loss of encapsulated $[^{99m}Tc]TcO_4^-$ when transferred to mouse serum. **C6** and **C7** will be used in SPECT and PET imaging studies to investigate whether the external functionalisation alters the bioaccumulation in comparison to **C1**. The two cages could also be externally functionalised using ester and amide bond formations

to bioconjugate targeting groups such as peptides onto the outside of the cage, to further control the bioaccumulation of the cages *in vivo*.

3.4 Experimental

3.4.1 General

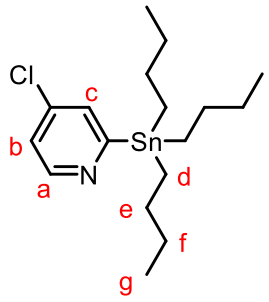
All ^{1}H, ^{13}C and ^{19}F NMR spectra were recorded on either a 500 MHz Bruker AV III equipped with a DCH cryo-probe (Ava500), a 500 MHz Bruker AV IIIHD equipped with a Prodigy cryoprobe (Pro500) or a 400 MHz Bruker AV III equipped with BBFO+ probe (Ava400) at a constant temperature of 300 K. Chemical shifts are reported in parts per million and are referenced to residual solvent.[44] Coupling constants (*J*) are reported in hertz (Hz). All analysis was performed with MestReNova, Version 12.0.3. Mass spectra were recorded using either a high resolution Bruker ToF instrument (DataAnalysis) for organic compounds or Waters ESI Synapt spectrometer fitted with a Nanomate® and analysed using MassLynx V4.1 for metal complexes.

3.4.2 Solvents, Reactants and Reagents

All reagents and solvents were purchased from VWR, Alfa Aesar, Fluorochem or Merck and used without further purification. Anhydrous solvents were used where drying was achieved with solvent purification system manufactured by Glass Contour. Column chromatography was carried out using either Geduran Si60 (40-63 µm) or Aluminum oxide activated (basic, Brockmann) as the stationary phase and TLC was performed on precoated Kieselgel 60 plates (0.20 mm thick, Merck) and observed under UV light at 254 nm or 365 nm. Size exclusion columns were packed used Sephadex-LH20 soaked in water. All reactions were carried out under N_2 inert atmosphere, unless stated otherwise. Phosphate buffered saline, was produced using buffer tablets purchased from fischer scientific and ultra-pure deionised water to a resultant concentration of 0.1 M. All reactions featuring organotin reagents were handled with ultra-nitrile gloves and contaminated glassware bleached following usage.

3.4.3 Ligand Synthesis

4-Chloro-2-(tributylstannyl)pyridine (2)

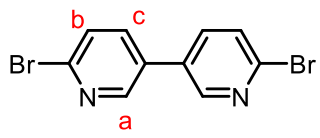


2

To a Schlenk flask under an inert atmosphere 2-bromo-4-chloropyridine (7.18 g, 37.3 mmol) was added and suspended in dry toluene (150 mL). The solution was then cooled to 0 °C and whilst stirring iPrMgCl (2 M in THF, 39.2 mmol, 19.6 mL) was added dropwise. The reaction mixture was left to stir for 1 hour at 0 °C before the addition of $SnBu_3Cl$ (39.4 mmol, 10.6 mL). After 18 hours and slowly being warmed to room temperature the reaction was quenched with H_2O (40 mL) and an aqueous extraction performed with hexane (3 x 40mL). The organic extract was dried over $MgSO_4$ and concentrated *in vacuo* before purifying via flash chromatography (basic alumina – 5% diethyl ether in hexane) to yield a clear and colourless oil. (13.5 g, 90%, 33.5 mmol)

^{1}H NMR (500 MHz, $CDCl_3$) δ 8.63 (dd, *J* = 5.4, 0.7 Hz, 1H, H_a), 7.45 – 7.40 (m, 1H, H_c), 7.15 (dd, *J* = 5.4, 2.3 Hz, 1H, H_b), 1.62 – 1.52 (m, 6H, H_d), 1.37 – 1.33 (m, 6H, H_e), 1.21 – 1.11 (m, 6H, H_f), 0.91 (t, *J* = 7.4 Hz, 9H, H_g) ppm. ^{13}C NMR (126 MHz, $CDCl_3$) δ 176.25, 151.11, 142.52, 131.88, 122.32, 28.99, 27.30, 13.65, 9.98 ppm. ESI MS: Calculated mass of $C_{17}H_{31}ClNSn$ = 404.1162 Found mass = 404.1158.

6-6'-Dibromo-3,3'-bipyridine (3)



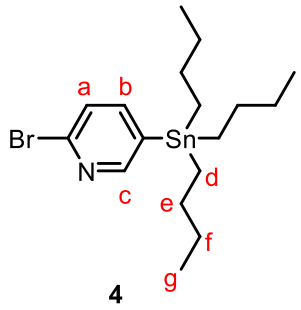
3

To a solution of 2-Bromo-5-iodopyridine (1.40 g, 4.92 mmol) in dry toluene (30.0 mL) was added **4** (2.00 g, 4.47 mmol) and the mixture sparged with N_2 for 20 min. Followed by the addition of $Pd(PPh_3)_4$ (0.260 g, 0.220 mmol) under an inert atmosphere, and then allowed to react at 120 °C for 48 hours. The reaction mixture was cooled to room temperature before removing the solvent *in vacuo* and

washing the residue with MeOH followed by celite filtration. The crude brown solid was purified via flash chromatography (5% Et_3N in DCM) yielding white solid. (0.650 g, 2.07 mmol, 46%)

^{1}H NMR (500 MHz, $CDCl_3$) δ 8.59 (dd, *J* = 2.6, 0.8 Hz, 1H, H_a), 7.74 (dd, *J* = 8.3, 2.6 Hz, 1H, H_c), 7.64 (dd, *J* = 8.3, 0.8 Hz, 1H, H_b) ppm. ^{13}C NMR (126 MHz, $CDCl_3$) δ 148.19, 142.43, 136.69, 131.65, 128.53 ppm. ESI MS: Calculated mass of $C_{10}H_7Br_2N_2$ = 312.89705 Found mass = 312.8964

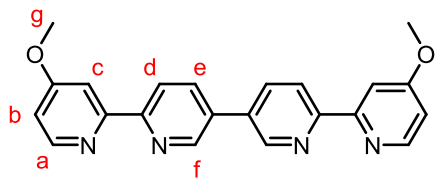
2-Bromo-5-(tributylstannyl)pyridine (4)



4

To a Schlenk flask under an inert atmosphere 2-Bromo-5-iodopyridine (6.40 g, 22.4 mmol) was added and suspended in dry THF (150 mL). The solution was then cooled to 0 °C and whilst stirring iPrMgCl (2 M in THF, 22.8 mmol, 11.4 mL) was added dropwise. The reaction mixture was left to stir for 1 hour at 0 °C before the addition of $SnBu_3Cl$ (23.5 mmol, 5.30 mL). After 18 hours and slowly being warmed to room temperature the reaction was quenched with H_2O (40mL) and an aqueous extraction performed with hexane (3 x 40mL). The organic extract was dried over $MgSO_4$ and concentrated *in vacuo* before purifying via flash chromatography (basic alumina – 5% diethyl ether in hexane) to yield a clear and colourless oil. (8.20 g, 82%, 18.3 mmol)

^{1}H NMR (500 MHz, $CDCl_3$) δ 8.38 – 8.31 (m, 1H, H_c), 7.64 – 7.55 (m, 1H, H_a), 7.47 – 7.41 (m, 1H, H_b), 1.64 – 1.44 (m, 6H, H_d), 1.34 (dq, J = 14.5, 7.3 Hz, 6H, H_e), 1.20 – 1.02 (m, 6H, H_f), 0.90 (t, J = 7.4 Hz, 9H, H_g) ppm. ^{13}C NMR (126 MHz, $CDCl_3$) δ 156.18, 146.17, 142.69, 135.78, 128.17, 28.95, 27.26, 13.62, 9.73 ppm. ESI MS: Calculated mass of $C_{17}H_{31}BrNSn$ = 448.0656 Found mass = 448.



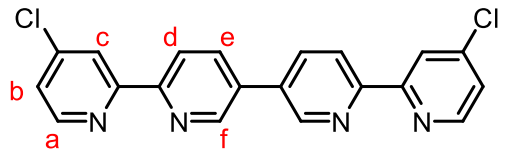
L5

L9 (0.02 g, 0.05 mmol), sodium methoxide (0.06 g, 1.06 mmol) and MeOH (1.5 mL) were combined under an inert atmosphere in a microwave vial (0.5 – 2 mL). The reaction was then microwaved (130

°C, absorption normal) for 2 hours before diluting with H_2O (20 mL) and extracting with chloroform (30 mL). Solvents were removed *in vacuo* to yield white solid. (16 mg, 0.04 mmol, 82%)

^{1}H NMR (601 MHz, $CDCl_3$) δ 9.01 (dd, *J* = 2.4, 0.8 Hz, 2H, H_f), 8.59 – 8.57 (m, 2H, H_a), 8.56 – 8.54 (m, 2H, H_d), 8.13 – 8.10 (m, 2H, H_e), 8.07 (d, *J* = 2.5 Hz, 2H, H_c), 6.91 (dd, *J* = 5.6, 2.5 Hz, 2H, H_b), 4.01 (s, 6H, H_g) ppm. ESI MS: Calculated mass of $C_{22}H_{19}N_4O_2$ = 371.15025 Found mass = 371.1487

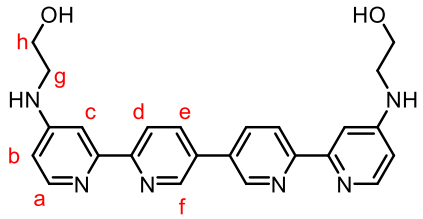
Note: Due to solubility issues the ^{13}C NMR was not obtained.



L9

To a solution of **3** (0.50 g, 1.59 mmol) in dry m-xylene (25 mL) was added **2** (1.60 g, 3.98 mmol) and the mixture degassed with N_2 for 20 min. This was followed by the addition of $Pd(PPh_3)_4$ (0.360 g, 0.320 mmol) under inert atmosphere, and the mixture allowed to react at 160 °C for 72 hours. The reaction mixture was cooled to room temperature before removing the solvent *in vacuo.* The solid product was dissolved in chloroform and filtered through celite. The filtrate was concentrated *in vacuo* before washing the residue with MeCN yielding a white solid. (0.35 g, 0.92 mmol, 58%)

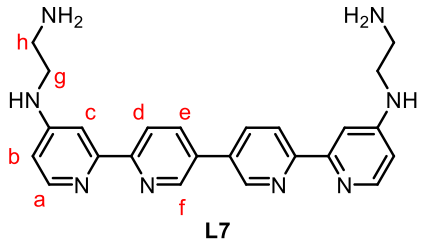
^{1}H NMR (500 MHz, $CDCl_3$) δ 9.01 (dd, *J* = 2.4, 0.8 Hz, 2H, H_f), 8.62 (dd, *J* = 5.2, 0.6 Hz, 2H, H_a), 8.57 (dd, *J* = 8.2, 0.8 Hz, 2H, H_d), 8.55 (dd, *J* = 2.0, 0.6 Hz, 2H, H_c), 8.13 (dd, *J* = 8.2, 2.4 Hz, 2H, H_e), 7.38 (dd, *J* = 5.2, 2.0 Hz, 2H, H_b) ppm. ^{13}C NMR (126 MHz, $CDCl_3$) δ 157.06, 154.68, 150.13, 147.51, 145.37, 135.29, 133.53, 124.13, 121.64, 121.53 ppm. ESI MS: Calculated mass of $C_{20}H_{13}Cl_2N_4$ = 379.0512 Found mass = 379.0508



L6

L9 (0.04 g, 0.11 mmol) and ethanolamine (1.01 g, 16.5 mmol, 1 mL) were combined in a microwave vial (0.5 – 2 mL) equipped with magnetic stirrer bar and heated at 170 ° C for 18 h. The product was precipitated by the addition of DCM (20 mL) and filtered to yield a beige solid (32.0 mg, 0.070 mmol, 67%).

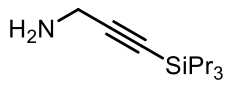
^{1}H NMR (601 MHz, d_6-DMSO) δ 9.09 (dd, J = 2.4, 0.9 Hz, 2H, H_f), 8.45 (dd, J = 8.2, 0.9 Hz, 2H, H_d), 8.31 (dd, J = 8.2, 2.4 Hz, 2H, H_e), 8.17 (d, J = 5.6 Hz, 2H, H_a), 7.68 (d, J = 2.5 Hz, 2H, H_c), 6.60 (dd, J = 5.6, 2.5 Hz, 2H, H_b), 3.59 (t, J = 6.0 Hz, 4H, H_h), 3.24 (q, J = 6.0 Hz, 4H, H_g) ppm. ^{13}C NMR (126 MHz, d_6-DMSO) δ 156.16, 155.22, 155.17, 149.82, 147.49, 135.37, 132.57, 120.94, 107.89, 104.32, 45.99, 41.18 ppm. ESI MS: Calculated mass of $C_{24}H_{25}N_6O_2$ = 429.20335 Found mass = 429.2038



L7

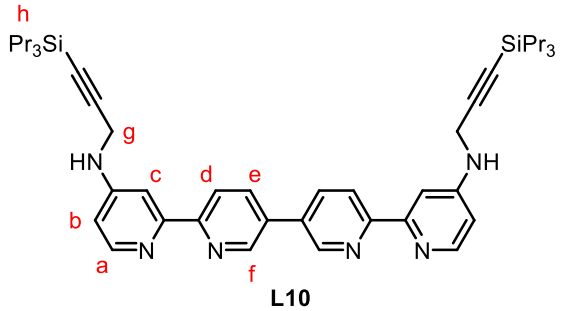
L9 (0.04 g, 0.11 mmol) and ethylenediamine (0.90 g, 15.0 mmol, 1 mL) were combined in a microwave vial (0.5 – 2 mL) equipped with magnetic stirrer bar and heated at 170 ° C for 18 h. The product was then precipitated by the addition of DCM (20 mL) and filtered to yield beige solid (38 mg, 0.09 mmol, 81%).

^{1}H NMR (601 MHz, d_6-DMSO) δ 9.09 (dd, J = 2.4, 0.8 Hz, 2H, H_f), 8.45 (dd, J = 8.2, 0.8 Hz, 2H, H_d), 8.32 (dd, J = 8.2, 2.4 Hz, 2H, H_e), 8.18 (d, J = 5.6 Hz, 2H, H_a), 7.68 (d, J = 2.4 Hz, 2H, H_c), 6.58 (dd, J = 5.6, 2.4 Hz, 2H, H_b), 3.15 (q, J = 6.4 Hz, 4H, H_g), 2.76 (t, J = 6.4 Hz, 4H, H_h) ppm. ^{13}C NMR (126 MHz, d_6-DMSO) δ 156.17, 155.27, 155.20, 149.76, 147.48, 135.35, 132.57, 120.96, 107.99, 104.22, 59.82, 45.03 ppm. Calculated mass of $C_{24}H_{27}N_8$ = 427.23532 Found mass = 427.2335



5

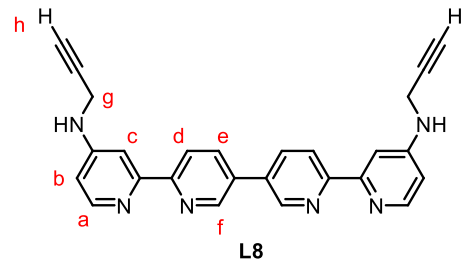
Synthesised according to literature.[27]



L10

Under inert atmosphere suspended **L9** (0.10 g, 0.26 mmol), BINAP (0.035 g, 0.053 mmol), NaOtBu (0.06 g, 0.66 mmol), $Pd_2(dba)_3$ (0.024 g, 0.026 mmol) and **5** (0.66 mmol, 162 µL) in dry toluene (10 mL) in a microwave vial (10 – 20 mL). Reaction heated to 115 °C for 96 h. Cooled to rt, diluted with H_2O (30 mL) and extracted with chloroform (30 mL) before drying over $MgSO_4$. Removed solvent *in vacuo* and purified via flash chromatography (0-10% MeOH, 5% Et_3N in chloroform) yielding beige solid (145 mg, 0.19 mmol, 75%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 8.95 (d, *J* = 2.3 Hz, 2H, H_f), 8.53 (d, *J* = 8.3 Hz, 2H, H_d), 8.38 (d, *J* = 5.6 Hz, 2H, H_a), 8.08 (dd, *J* = 8.3, 2.3 Hz, 2H, H_e), 7.81 (d, *J* = 2.5 Hz, 2H, H_c), 6.66 (dd, *J* = 5.6, 2.5 Hz, 2H, H_b), 4.15 (d, *J* = 5.4 Hz, 4H, H_g), 1.09 – 1.01 (m, 42H, H_h) ppm. ^{13}C NMR (126 MHz, $CDCl_3$) δ 155.43, 153.95, 150.27, 148.93, 147.18, 135.09, 133.10, 121.36, 108.20, 106.03, 102.58, 85.66, 33.85, 18.52, 18.47 ppm. ESI MS: Calculated mass of $C_{44}H_{61}N_6Si_2$ = 729.44908 Found mass = 729.4493



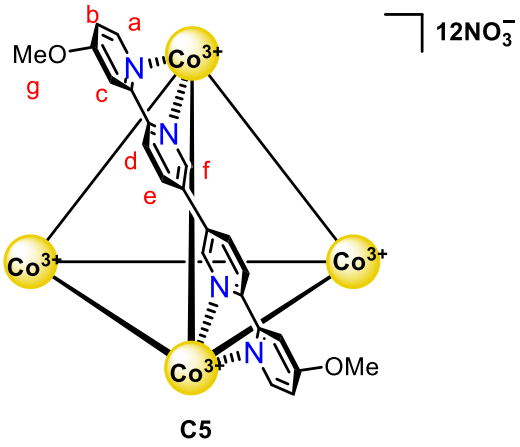
L8

To **L10** (0.10 g, 0.14 mmol) in dry THF (6 mL) under an inert atmosphere was dropwise added TBAF (0.549 mmol, 1M in THF, 0.55 mL). After addition was complete, the solution was left to react at room temperature for 1 hour. The solution was diluted with H_2O (20 mL) before extracting with chloroform (20 mL). The solvent was removed *in vacuo* and the solid product washed with diethyl ether, yielding white solid. (42 mg, 0.10 mmol, 74%)

^{1}H NMR (500 MHz, d_6-DMSO) δ 9.14 (d, *J* = 2.4 Hz, 2H, H_f), 8.47 (d, *J* = 8.4 Hz, 2H, H_d), 8.36 (dd, *J* = 8.4, 2.4 Hz, 2H, H_e), 8.26 (d, *J* = 5.7 Hz, 2H, H_a), 7.76 (d, *J* = 2.4 Hz, 2H, H_c), 6.68 (dd, *J* = 5.7, 2.4 Hz, 2H, H_b), 4.05 (dd, *J* = 5.8, 2.5 Hz, 4H, H_g), 3.22 – 3.13 (m, 2H, H_h) ppm. Calculated mass of $C_{26}H_{21}N_6$ = 417.18222 Found mass = 417.1810

Note: Due to solubility issues the ^{13}C NMR was not obtained.

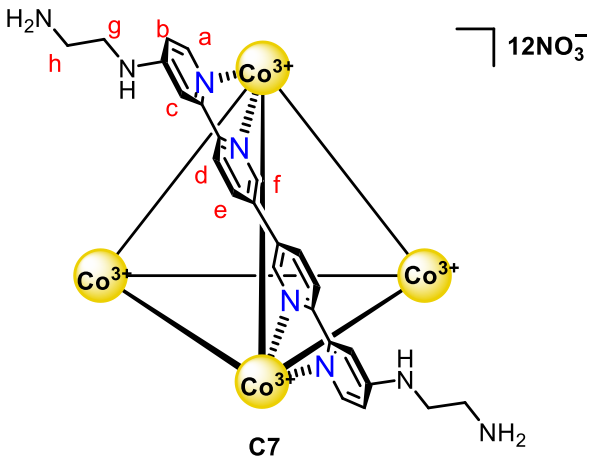
3.4.4 Cage Assembly



C5

L5 (12 mg, 16.2 µmol) was suspended in solution of **C2** (8 mg, 2.7 µmol) in H_2O (1.6 mL), MeCN (200 µL) and DMF (20 µL). A solution of aqueous $Co(NO_3)_2.6H_2O$ (2 mM, 60 µL) was added before heating at 90 °C for 48 h. The mixture was cooled before diluting with H_2O (3 mL) and filtering through celite. The filtrate was freeze dried and further purified by size exclusion chromatography (sephadex LH-20) yielding a yellow solid. (8.5 mg, 2.6 µmol, 93%)

^{1}H NMR (601 MHz, D_2O) δ 8.67 (d, *J* = 8.2 Hz, 12H, H_d), 8.38 (d, *J* = 3.0 Hz, 12H, H_c), 7.79 (d, *J* = 8.2 Hz, 12H, H_e), 7.37 (s, 12H, H_f), 7.34 – 7.28 (m, 12H, H_a), 7.28 – 7.22 (m, 12H, H_b), 4.08 (s, 36H, H_g) ppm. ^{13}C NMR (126 MHz, D_2O) δ 171.4, 155.8, 155.3, 151.6, 149.0, 144.1, 138.0, 125.9, 117.1, 115.1, 57.7 ppm. ^{1}H DOSY NMR (500 MHz, D_2O): D = 1.91 x 10^{-10} m^2 s^{-1} ; calculated hydrodynamic radius = 12.5 Å. ESI MS (m/z): calculated for 3+ 1005.1577; found 1005.1403 (3+), calculated for 4+ 738.3683; found 738.3683 (4+), calculated for 5+ 578.3043; found 578.2963 (5+), calculated for 6+ 471.5889 ; found 471.5800 (6+), calculated for 7+ 395.3603 ; found 395.3618 (7+).

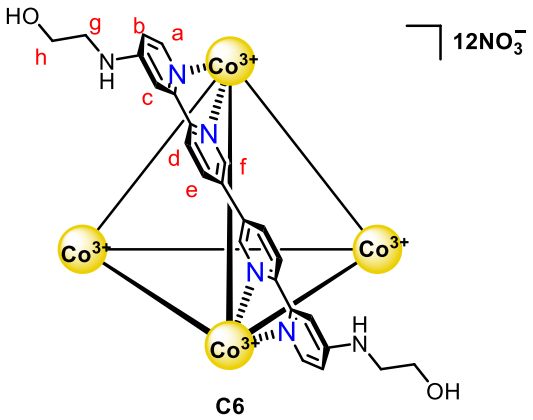


C7

L7 (54 mg, 127 µmol) was suspended in solution of **C2** (30 mg, 10.6 µmol) in H_2O (3.6 mL), CH_3CN (0.4 mL) and DMF (40 µL). A solution of aqueous $Co(NO_3)_2.6H_2O$ (2 mM, 250 µL) was added and the solution degassed for 10 min before heating at 90 °C for 48 h. The mixture was cooled before diluting with H_2O (8 mL) and filtering through celite. The filtrate was freeze dried and further purified by size exclusion chromatography with water (sephadex LH-20) yielding a yellow solid. (32 mg, 9.0 µmol, 85%)

^{1}H NMR (500 MHz, D_2O) δ 8.49 (m, 12H, H_d), 7.86 (s, 12H, H_c), 7.71 (d, J = 8.4 Hz, 12H, H_e), 7.23 (s, 12H, H_f), 6.92 (m, 12H, H_a), 6.76 (d, J = 7.4 Hz, 12H, H_b), 3.61 – 3.48 (m, 24H, H_h), 3.10 – 3.07 (m, 24H, H_g) ppm. ^{1}H DOSY NMR (500 MHz, D_2O): D = 1.93 e^{-10} m^2 s^{-1} ; calculated hydrodynamic radius = 12.7 Å.

Note: Due to solubility issues the ^{13}C NMR was not obtained.

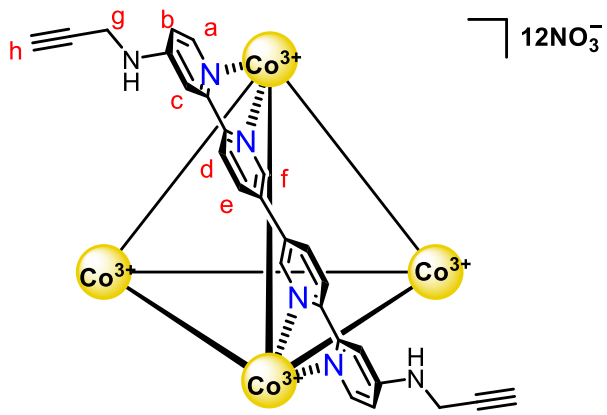


C6

L6 (54 mg, 127 µmol) was suspended in solution of **C2** (30 mg, 10.6 µmol) in H_2O (3.6 mL), CH_3CN (0.4 mL) and DMF (40 µL). A solution of aqueous $Co(NO_3)_2.6H_2O$ (2 mM, 250 µL) was added and the solution degassed for 10 min before heating at 90 °C for 48 h. The mixture was cooled before diluting with H_2O

(8 mL) and filtering through celite. The filtrate was freeze dried and further purified by size exclusion chromatography with water (sephadex LH-20) yielding an orange solid. (29 mg, 8.1 µmol, 77%)

^{1}H NMR (500 MHz, D_2O) δ 8.52 (m, 12H, H_d), 7.86 (s, 12H, H_c), 7.68 (s, 12H, H_e), 7.24 (s, 12H, H_f), 6.91 (m, 12H, H_a), 6.74 (d, *J* = 6.3 Hz, 12H, H_b), 3.76 (m, 24H, H_h), 3.50 (m, 24H, H_g) ppm. ^{13}C NMR (126 MHz, D_2O) δ 157.56, 156.59, 148.52, 143.12, 137.29, 124.63, 115.85, 113.41, 110.54, 107.48, 44.32, 41.24.^{1}H DOSY NMR (500 MHz, D_2O): D = 1.83 e^{-10} m^2 s^{-1} ; calculated hydrodynamic radius = 12.5 A. ESI-MS (m/z): calculated for 4+ 825.4529; found 825.4343 (4+), calculated for 5+ 647.7685; found 647.9666 (5+), calculated for 6+ 529.6420; found 529.6381 (6+), calculated 7+ 445.1201; found 445.1176 (7+), calculated 8+ 381.7325; found 381.7324 (8+), calculated 9+ 332.4268; found 332.4268 (9+).



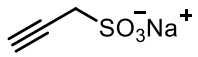
C8

L8 (30 mg, 72 µmol) was suspended in solution of **C2** (17 mg, 6 µmol) in H_2O (1.6 mL), CH_3CN (0.2 mL) and DMF (20 µL). A solution of aqueous $Co(NO_3)_2.6H_2O$ (2 mM, 128 µL) was added and the solution degassed for 10 min before heating at 90 °C for 48 hours. The mixture was cooled before diluting with H_2O (8 mL) and filtering through celite. The filtrate was freeze dried yielding a pink solid. (8 mg, 2.3 µmol, 38%)

^{1}H NMR (601 MHz, D_2O) δ 8.69 – 8.64 (m, 12H, H_d), 8.38 (s, 12H, H_c), 7.79 (s, 12H, H_e), 7.39 (s, 12H, H_f), 7.30 (s, 12H, H_a), 7.24 (d, *J* = 22.2 Hz, 12, H_b), 4.08 (s, 24H, H_g), 2.66 (s, 12H, H_h) ppm.

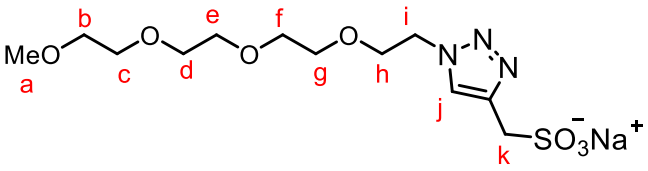
As discussed in "Chapter 3 – Cage Assembly" limited mass of this complex prevented further characterisation.

3.4.5 Post Assembly Modification



7

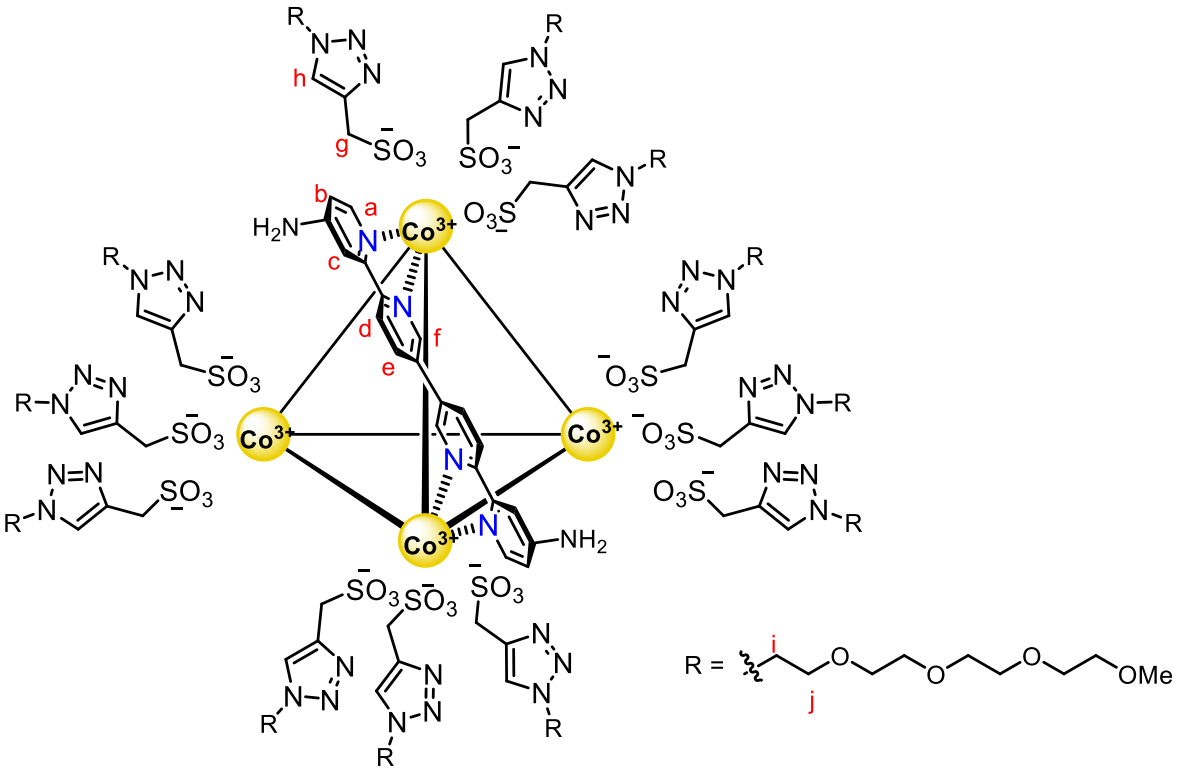
Synthesised according to literature.[35]



8

mPEG4-Azide (70 mg, 0.25 mmol, 75 µL) and sulfate alkyne intermediate (37 mg, 0.26 mmol) were added to a stirred solution of $CuSO_4.5H_2O$ (25 mg, 0.10 mmol), sodium ascorbate (100 mg, 0.50 mmol) in THF/EtOH/H_2O (1:2:2, 2.5 mL) and the solution degassed for 5 min. The reaction was heated for 18 hours at 65 °C before being cooled and diluted with acetone (5 mL). The mixture was filtered and the filtrate concentrated *in vacuo*, product dissolved in MeOH and filtered. The solvent removed from filtrate *in vacuo* yielding product as maroon oil. (90 mg, 0.24 mmol, 95%) Stored as a 1M stock solution in MeOH at 4 °C.

^{1}H NMR (400 MHz, MeOD) δ 8.05 (s, 1H, H_j), 4.60 (t, *J* = 4.7 Hz, 2H, H_i), 4.21 (s, 2H, H_k), 3.92 (t, *J* = 4.7 Hz, 2H, H_h), 3.64 – 3.52 (m, 12H, $H_{b,c,d,e,f,g}$), 3.35 (s, 3H, H_a) ppm. ^{13}C NMR (126 MHz, MeOD) δ 124.84, 71.55, 70.21, 70.13, 70.09, 70.06, 69.94, 69.86, 69.71, 69.03, 57.68, 50.40. High resolution ESI MS: Calculated mass of $C_{12}H_{23}N_3O_7SNa$ = 376.11489 Found mass = 376.1146 Calculated mass of $C_{12}H_{22}N_3O_7SNa_2$ = 398.09684 Found mass = 398.0971



C10

A suspension of **C1** (20 mg, 6.6 µmol) and mPEG-triazole (160 mg, 0.40 mmol) in MeOH (5 mL) was sonicated for 20 min. The mixture was then heated for 18 hours at 60 °C before being cooled and diluted with MeOH (10 mL). The mixture was filtered through celite and the filtrate concentrated *in vacuo*. The product was precipitated by the addition of H_2O (10 mL) and filtered and weighed. The isolated precipitate was stored as maroon stock solution in MeOD. (25 mg, 3.6 µmol, 55%)

^{1}H NMR (500 MHz, MeOD) δ 9.25 (d, *J* = 2.3 Hz, 12H, H_f), 8.51 (dd, *J* = 8.4, 2.3 Hz, 12H, H_e), 8.35 (d, *J* = 8.4 Hz, 12H, H_d), 8.16 (d, *J* = 6.9 Hz, 12H, H_a), 8.05 (s, 12H, H_h), 7.64 (d, *J* = 2.5 Hz, 12H, H_c), 6.94 (dd, *J* = 6.9, 2.5 Hz, 12H, H_b), 4.59 (t, *J* = 5.1 Hz, 24H, H_i), 4.22 (s, 24H, H_g), 3.91 (t, *J* = 5.1 Hz, 24H, H_j), 3.67 (s, 36H), 3.65 – 3.57 (m, 144) ppm.

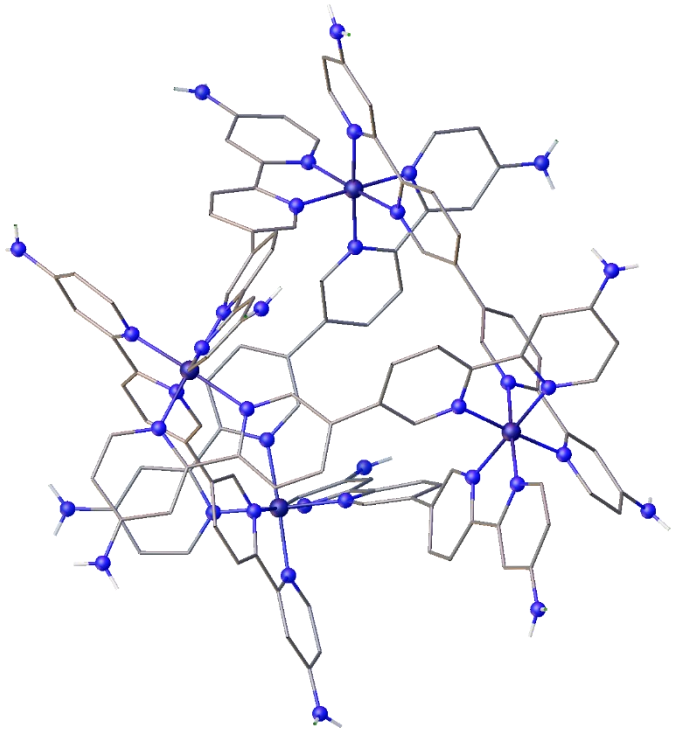
As discussed in "Chapter 3 – Non covalent functionalisation" limited solubility of this complex prevented further characterisation.

3.4.6 X-Ray Crystallography

Data Collection by: **Dr David Cordes – University of St Andrews**

Solved by: **Gary S Nichol – University of Edinburgh**

***R_1*=7.94%**



Experimental. Single orange prism crystals of **C1** recrystallised from a mixture of MeCN and water by vapour diffusion. A suitable crystal with dimensions 0.21 × 0.15 × 0.04 mm^3 was selected and mounted on a MITIGEN holder in Paratone oil on a Rigaku XtaLAB P200K diffractometer. The crystal was kept at a steady T = 173.15 K during data collection. The structure was solved with the ShelXT solution program using dual methods and by using Olex2 as the graphical interface.[45,46] The model was refined with ShelXL 2018/3 using full matrix least squares minimisation on F^2.[45]

Crystal Data. $C_{168}Co_4H_{168}N_{72}O_{36}$, M_r = 4007.45, monoclinic, $C2/c$ (No. 15), a = 24.4438(4) Å, b = 27.2173(5) Å, c = 26.7674(5) Å, β = 91.932(2)°, $\alpha = \gamma$ = 90°, V = 17798.1(6) $Å^3$, T = 173.15 K, Z = 4, Z' = 0.5, μ(Mo K_α) = 0.464, 108662 reflections measured, 16895 unique (R_{int} = 0.0510) which were used in all calculations. The final wR_2 was 0.2681 (all data) and R_1 was 0.0794 (I≥2 σ(I)).

Compound	C1
Formula	$C_{168}Co_4H_{168}N_{72}O_{36}$
$D_{calc.}$/ g cm^{-3}	1.496
μ/mm^{-1}	0.464
Formula Weight	4007.45
Colour	orange
Shape	prism
Size/mm^3	0.21×0.15×0.04
T/K	173.15
Crystal System	monoclinic
Space Group	C2/c
a/Å	24.4438(4)
b/Å	27.2173(5)
c/Å	26.7674(5)
α/°	90
β/°	91.932(2)
γ /°	90
V/Å^3	17798.1(6)
Z	4

Compound	C1
Z'	0.5
Wavelength/Å	0.71073
Radiation type	Mo K_α
Θ_{min}/°	1.667
Θ_{max}/°	25.682
Measured Refl's.	108662
Indep't Refl's	16895
Refl's I≥2 σ(I)	13129
R_{int}	0.0510
Parameters	1057
Restraints	770
Largest Peak	1.144
Deepest Hole	-1.077
GooF	1.058
wR_2 (all data)	0.2681
wR_2	0.2528
R_1 (all data)	0.0942
R_1	0.0794

3.4.7 Stability Studies

NMR Studies

Phosphate Buffered Saline Stability Studies

Procedure for Phosphate Buffered Saline Stability Studies

A solution of Co^{III} Cage (1 mM, 0.5 µmol, 0.5 mL) was prepared in 0.1 M PBS in D_2O (with t-BuOH internal standard) and the solution sonicated for 30 seconds before recording ^1H NMR spectra at interval time points. Time point ^1H NMR spectra stacked and normalised against the internal standard. Divide the integrals over the time points to work out cage consumed.

Glutathione Stability Studies

Procedure for Glutathione Stability Studies

To a solution of Co^{III} Cage (1 mM, 0.5 µmol, 0.5 mL) in D_2O (with t-BuOH internal standard) was added reduced glutathione (1.5 mg, 5 µmol) and the mixture sonicated before recording ^1H NMR spectra at interval time points. Time point ^1H NMR spectra stacked and normalised against the internal standard. Divide the integrals over the time points to work out cage consumed.

Radiochemical Yield Studies

The following experiments were completed by our collaborators in the Archibald group at the University of Hull, notably Dr Isaline Renard.

Concentration of $^{99}Mo/^{99m}Tc$ generator elution

10 mL of $[^{99m}Tc]TcO_4^-$ were eluted from a 99Mo/99mTc generator (Ultra Technekow™ FM, Curium) in saline. The eluate was cannulated through a Dionex™ OnGuard™ II Ag cartridge (ThermoFisher Scientific, catalogue #: 057089), pre-conditioned with 5 mL water, and a strong anion-exchange Bond Elut SAX cartridge (Agilent Technologies, catalogue #: 12102017), pre-conditioned with 1 mL MeOH, followed by 5 mL water. The activity trapped on the SAX cartridge was eluted with portions of saline solution of 500 µL, 1 mL and 3 x 500 µL. The bulk of the activity was found in the 1 mL elution (used for subsequent experiments).

Encapsulation of $[^{99m}Tc]TcO_4^-$ in cage

Stocks of the cage at the desired concentration were prepared in metal-free water. 1 µL of $[^{99m}Tc]TcO_4^-$ was added to 100 µL cage solution. The resulting mixture was stirred at 25°C in a shaker (600 rpm) for 15 min. Encapsulation was determined by iTLC (silica gel / H2O), where encapsulated $[^{99m}Tc]TcO_4^-$ remains on the baseline and free $[^{99m}Tc]TcO_4$- elutes with solvent front. The proportion of bound and free $[^{99m}Tc]TcO_4$- was determined by cutting the eluted TLC in half and analyzing each part using a Wallac 1480 Wizard 3" Automated Gamma Counter (Perkin Elmer, USA).

Serum stability

To evaluate serum stability of the encapsulated [^{99m}Tc]TcO_4- in conditions that mimic *in vivo* administration in mice, 10 µL of encapsulated species were added to 100 µL of pre-warmed mouse serum. The resulting mixture was incubated at 37°C and stirred at 600 rpm. Encapsulation was determined by iTLC (silica gel / H_2O). The proportion of bound and free [^{99m}Tc]TcO_4- was determined by cutting the eluted TLC in half and analyzing each part using a Wallac 1480 Wizard 3" Automated Gamma Counter (Perkin Elmer, USA).

Stability of encapsulated [^{99m}Tc]TcO_4^- in presence of competitor

[^{99m}Tc]TcO_4^- was encapsulated in the different cages as described before at 2X EC_{95}. 50 µL of encapsulated species were then added to 50 µL of competitor. In case of NaCl, NH_4PF_6, $NaBF_4$, $NaIO_4$, $KClO_4$, 100 µM solutions were used to afford a final concentration of 50 µM. The resulting mixture was stirred at 25°C in a shaker (600 rpm). Samples were taken at 15, 30 and 60 min. Encapsulation was determined by iTLC (silica gel / H_2O). The proportion of bound and free [^{99m}Tc]TcO_4^- was determined by cutting the eluted TLC in half and analyzing each part using a Wallac 1480 Wizard 3" Automated Gamma Counter (Perkin Elmer, USA).

3.4.8 Host-Guest Chemistry

K_a for cage and guest combinations determined through ^{1}H NMR titration in unbuffered D_2O. Two methods were used depending on the amount of cage possessed.

For the first method, a solution of cage with a guest was titrated into a solution of cage, thereby maintaining a constant concentration of cage. With anionic guests, addition of an excess of guest to cage lead to precipitation, presumably via anion metathesis to less coordinating anions and thus the titrations were limited.

For the second method a solution of guest was titrated into a solution of cage, causing a slight decrease in cage concentration throughout the course of the titration.

For both methods, the cage peaks were plotted over the course of the NMR titration. For the guests exhibiting binding in fast exchange the change in chemical shift of the internal proton was monitored throughout the titration. The guests in slow-exchange with the cage, the concentration of the host-guest complex was monitored throughout the titration.

Fast Exchange Method

For the addition of both guest and host into host:

$$\Delta\delta = \frac{\Delta\delta_{Max}}{2}\left(\frac{C_{Guest}}{C_{Host}} + \frac{1}{C_{Host}k_{Ass}} + 1 - \sqrt{\left(\frac{C_{Guest}}{C_{Host}} + \frac{1}{C_{Host}k_{Ass}} + 1\right)^2 - \frac{4C_{Guest}}{C_{Host}}}\right)$$

$\Delta\delta_{\mathrm{Max}} =$ maximal change in chemical shift

$\mathrm{C}_{\mathrm{Guest}} =$ concentration of guest

$\mathrm{C}_{\mathrm{Host}} =$ concentration of host

$\mathrm{k}_{\mathrm{Ass}} =$ association constant

Origin Function: y=(Dmax/2)*(G/H0+1/H0/Ka+1-sqrt((G/H0+1/H0/Ka+1)^2-4*G/H0))

For the addition of guest into host:

$$C_{Guest} = \frac{G_0\Delta v}{v_0 + \Delta v}$$

$$C_{Host} = \frac{H_0 v_0}{v_0 + \Delta v}$$

$$\Delta\delta = \frac{\Delta\delta_{Max}}{2}\left(\frac{G_0\Delta v}{H_0 v_0} + \frac{v_0 + \Delta v}{H_0 v_0 k_{Ass}} + 1 - \sqrt{\left(\frac{G_0\Delta v}{H_0 v_0} + \frac{v_0 + \Delta v}{H_0 v_0 k_{Ass}} + 1\right)^2 - \frac{4G_0\Delta v}{H_0 v_0}}\right)$$

$\mathrm{G}_0 =$ initial concentration of guest

$\mathrm{H}_0 =$ initial concentration of host

$\mathrm{v}_0 =$ initial volume

$\Delta\delta_{\mathrm{Max}} =$ maximal change in chemical shift

$\mathrm{k}_{\mathrm{Ass}} =$ association constant

Origin Function: y=(Dmax/2)*(G0*x/H0/V0+(V0+x)/H0/V0/Ka+1-sqrt((G0*x/H0/V0+(V0+x)/H0/V0/Ka+1)^2-4*G0*x/H0/V0))

3.4.9 Slow Exchange Method

$$y = \frac{\left(x + P + \left(\frac{1}{K_a}\right)\right) + \sqrt{\left(x + P + \left(\frac{1}{K_a}\right)\right)^2 - (4Px)}}{2}$$

$\mathrm{x} = \text{concentration of guest}$

$\mathrm{P} = \text{concentration of host}$

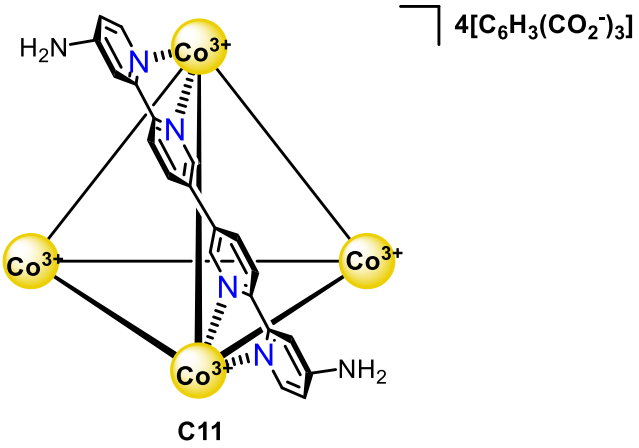
$\mathrm{y} = \text{concentration of host-guest complex}$

$\mathrm{K_a} = \text{association constant}$

Origin function = ((x+h+(1/K))-((x+h+(1/K))^2-(4*h*x))^(1/2))/2

3.4.10 Anion Metathesis

Experiment carried out by Samuel Leith with supervision.



C11

Stock solution of sodium trimesate (0.071 M) was prepared in D_2O. A solution of ^{t}BuOH (0.4 M, 100 µL) internal standard was also prepared in D_2O. A solution of **C1** (1.0 mM, 500 µL each) was prepared in D_2O, and ^{1}H NMR spectra were obtained followed by the addition of four stoichiometric equivalents of sodium trimesate (26.0 µL of stock solution). The solution was sealed, shaken and the ^{1}H NMR and ^{1}H DOSY spectra obtained.

^{1}H NMR (601 MHz, D_2O) δ 8.46 (d, *J* = 8.4 Hz, 1H), 8.37 (s, 1H), 7.89 (s, 1H), 7.72 (d, *J* = 8.4 Hz, 1H), 7.16 (s, 1H), 6.95 (d, *J* = 7.1 Hz, 1H), 6.83 (s, 1H) ppm. ^{1}H DOSY NMR (600 MHz, d_6-DMSO): D = -9.73 log(m^2.sec^1) ; calculated hydrodynamic radius = 14.25 Å.

^{1}H NMR Pseudo-Titrations in the Presence of a Guest

Stock solutions of KPF_6 (0.049 M, 100 µL) and TBA-ReO_4 (0.041 M, 100 µL) were prepared in unbuffered D_2O.

A solution of **C1** (1.0 mM, 500 µL) were prepared in D_2O , and ^{1}H NMR spectrum obtained for each solution followed by the addition of half a stoichiometric equivalent of guest. The solutions were sealed, shaken and their ^{1}H NMR spectra were acquired. Followed by a titration one stoichiometric equivalent of trimesate stock solution (7 µL) to the cage solution before ^{1}H NMR spectra were gathered, for a total of six stoichiometric equivalents.

3.5 References

1 B. P. Burke, W. Grantham, M. J. Burke, G. S. Nichol, D. Roberts, I. Renard, R. Hargreaves, C. Cawthorne, S. J. Archibald and P. J. Lusby, *J. Am. Chem. Soc.*, 2018, **140**, 16877–16881.

2 G. Yu, S. Yu, M. L. Saha, J. Zhou, T. R. Cook, B. C. Yung, J. Chen, Z. Mao, F. Zhang, Z. Zhou, Y. Liu, L. Shao, S. Wang, C. Gao, F. Huang, P. J. Stang and & X. Chen, *Nat. Commun.*, 2018, **9**, 1–18.

3 X. Lin, F. Chen, X. Yu, H. Wang, H. Qiu, Y. Li, S. Yin and P. J. Stang, *Proc. Natl. Acad. Sci. U. S. A.*, 2022, **119**, 1–8.

4 G. Yu, B. Zhu, L. Shao, J. Zhou, M. L. Saha, B. Shi, Z. Zhang, T. Hong, S. Li, X. Chen and P. J. Stang, *Proc. Natl. Acad. Sci. U. S. A.*, 2019, **116**, 6618–6623.

5 J. Han, A. Schmidt, T. Zhang, H. Permentier, G. M. M. Groothuis, R. Bischoff, F. E. Kühn, P. Horvatovich and A. Casini, *Chem. Commun.*, 2017, **53**, 1405–1408.

6 B. Woods, R. D. M. Silva, C. Schmidt, D. Wragg, M. Cavaco, V. Neves, V. F. C. Ferreira, L. Gano, T. S. Morais, F. Mendes, J. D. G. Correia and A. Casini, *Bioconjug. Chem.*, 2021, **32**, 1399–1408.

7 M. M. Khalil, J. L. Tremoleda, T. B. Bayomy and W. Gsell, *Int. J. Mol. Imaging*, 2011, **2011**, 15.

8 K. Shahzad, A. S. A. Majid, M. Khan, M. A. Iqbal and A. Ali, *Rev. Inorg. Chem.*, 2021, **41**, 151–198.

9 J. Pijarowska-Kruszyna, U. Karczmarczyk, A. Jaroñ, E. Laszuk, M. Radzik, P. Garnuszek and R. Mikolajczak, *Nucl. Med. Rev.*, 2017, **20**, 88–94.

10 T. A. Holly and R. O. Bonow, *Clinical Nuclear Cardiology*, Elsevier, 2010.

11 N. A. Seebacher, A. E. Stacy, G. M. Porter and A. M. Merlot, *J. Exp. Clin. Cancer Res.*, 2019, **38**, 1–39.

12 C. R. Forest, C. A. C. Silva and P. Thordarson, *Pept. Sci.*, 2020, **113**, 24205.

13 Y. Zhao, T. Ji, Y. Zhao, Y. Ding, J. Wang, R. Zhao, J. Lang, H. Qin, X. Liu, J. Shi, N. Tao, Z. Qin and G. Nie, *Angew. Chem. Int. Ed. Engl.*, 2016, **55**, 1055.

14 J. Zong, S. L. Cobb and N. R. Cameron, *Biomater. Sci.*, 2017, **5**, 872–886.

15 J. Han, A. Schmidt, T. Zhang, H. Permentier, G. M. M. Groothuis, R. Bischoff, F. E. Kühn, P. Horvatovich and A. Casini, *Chem. Commun.*, 2017, **53**, 1405–1408.

16 R. He, B. Finan, J. P. Mayer and R. D. DiMarchi, *Molecules*, 2019, **24**, 1–34.

17 H. C. Kolb, M. G. Finn and B. Sharpless, *Angew. Chem. Int. Ed*, 2001, 2004–2021.

18 H. C. Kolb and K. B. Sharpless, *Drug Discov. Today*, 2003, 8, 1128–1137.

19 W. Grantham and P. J. Lusby, University of Edinburgh, 2019.

20 J. Han, A. F. B. Räder, F. Reichart, B. Aikman, M. N. Wenzel, B. Woods, M. Weinmüller, B. S. Ludwig, S. Stürup, G. M. M. Groothuis, H. P. Permentier, R. Bischoff, H. Kessler, P. Horvatovich and A. Casini, *Bioconjug. Chem.*, 2018, **29**, 3856–3865.

21 P. Ruiz-Castillo and S. L. Buchwald, *Chem. Rev.*, 2016, **116**, 12564–12649.

22 Y. A. Getmanenko and R. J. Twieg, *J. Org. Chem.*, 2008, **73**, 830–839.

23 M. Zurro, S. Asmus, S. Beckendorf, C. Mück-Lichtenfeld and O. G. Mancheño, *J. Am. Chem. Soc.*, 2014, **136**, 13999–14002.

24 P. Li, J. Yu, F. Hao, H. He, X. Shi, J. Hu, L. Wang, C. Du, X. Zhang, Y. Sun, F. Lin, Z. Gu, D. Xu, X. Chen, L. Shen, G. Hu, J. Li, S. Chen, W. Xiao, Z. Wang, Q. Guo, X. Chang, X. Tian and T. Lin, *ACS Med. Chem. Lett.*, 2017, **8**, 841–846.

25 S. Ranganathan, Y. B. R. D. Rajesh and I. Karle, *Synlett*, 2007, **8**, 1215–1218.

26 P. A. Forero-Cortéscortés and A. M. Haydl, *Org. Process Res. Dev.*, 2019, **23**, 1487–1483.

27 V. Vaněk, J. Pícha, B. Fabre, M. Buděšínský, M. Lepšík and J. Jiráček, *Eur. J. Org. Chem.*, 2015, **2015**, 3689–3701.

28 A. K. H. Hirsch, F. Diederich, M. Antonietti and H. G. Börner, *Soft Matter*, 2009, **6**, 88–91.

29 B. H. Lipshutz and B. R. Taft, *Angew. Chem Int. Ed.*, 2006, **45**, 8235–8238.

30 D. Zhao, S. Tan, D. Yuan, W. Lu, Y. H. Rezenom, H. Jiang, L. Q. Wang and H. C. Zhou, *Adv. Mater.*, 2011, **23**, 90–93.

31 P. R. Symmers, M. J. Burke, D. P. August, P. I. T. Thomson, G. S. Nichol, M. R. Warren, C. J. Campbell and P. J. Lusby, *Chem. Sci.*, 2015, **6**, 756–760.

32 D. Zhao, S. Tan, D. Yuan, W. Lu, Y. H. Rezenom, H. Jiang, L.-Q. Wang, H.-C. Zhou, D. Zhao, D. Yuan, W. Lu, Y. H. Rezenom, H. Zhou, S. Tan, H. Jiang and L. Wang, *Adv. Mater.*, 2011, **23**, 90–93.

33 L. An, P. De La Torre, P. T. Smith, M. R. Narouz and C. J. Chang, *Angew. Chemie Int. Ed.*, 2023, **62**, 1–7.

34 T. K. Ronson, J. P. Carpenter and J. R. Nitschke, *Chem*, 2022, **8**, 557–568.

35 L. R. Staben, J. Chen, J. dela Cruz-Chuh, G. Del Rosario, M. A. Go, J. Guo, S. C. Khojasteh, K. R. Kozak, G. Li, C. Ng, G. D. Lewis Phillips, T. H. Pillow, R. K. Rowntree, J. Wai, B. Q. Wei, K. Xu, Z. Xu, S. F. Yu, D. Zhang and P. S. Dragovich, *J. Med. Chem.*, 2020, **63**, 9603–9622.

36 N. Nezafati, F. Moztarzadeh and S. Hesaraki, *Biotechnol. Bioprocess Eng.*, 2012, **17**, 746–754.

37 H. Ehrsson, I. Wallin and J. Yachnin, *Med. Oncol.*, 2002, **19**, 261–265.

38 S. Nayak, K. V. Reddy and A. C. Dash, *Transit. Met Chem*, 2014, **39**, 177–187.

39 H. Jiang and T. R. DeGrado, *Theranostics*, 2018, **8**, 3918–3931.

40 A. Khoshnevisan, K. Chuamsaamarkkee, M. Boudjemeline, A. Jackson, G. E. Smith, A. D. Gee, G. O. Fruhwirth and P. J. Blower, *J. Nucl. Med.*, 2017, **58**, 156–161.

41 S. Arndt, P. J. Kohlpaintner, K. Donsbach and S. R. Waldvogel, *Org. Process Res. Dev.*, 2022, **26**, 2564–2613.

42 H. Takezawa, R. Tabuchi, H. Sunohara and M. Fujita, *J. Am. Chem. Soc.*, 2020, **142**, 17919–17922.

43 Y. Shao, L. F. Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S. T. Brown, A. T. B. Gilbert, L. V. Slipchenko, S. V. Levchenko, D. P. O'Neill, R. A. DiStasio, R. C. Lochan, T. Wang, G. J. O. Beran, N. A. Besley, J. M. Herbert, C. Yeh Lin, T. Van Voorhis, S. Hung Chien, A. Sodt, R. P. Steele, V. A. Rassolov, P. E. Maslen, P. P. Korambath, R. D. Adamson, B. Austin, J. Baker, E. F. C. Byrd, H. Dachsel, R. J. Doerksen, A. Dreuw, B. D. Dunietz, A. D. Dutoi, T. R. Furlani, S. R. Gwaltney, A. Heyden, S. Hirata, C. P. Hsu, G. Kedziora, R. Z. Khalliulin, P. Klunzinger, A. M. Lee, M. S. Lee,

W. Liang, I. Lotan, N. Nair, B. Peters, E. I. Proynov, P. A. Pieniazek, Y. Min Rhee, J. Ritchie, E. Rosta, C. David Sherrill, A. C. Simmonett, J. E. Subotnik, H. Lee Woodcock, W. Zhang, A. T. Bell, A. K. Chakraborty, D. M. Chipman, F. J. Keil, A. Warshel, W. J. Hehre, H. F. Schaefer, J. Kong, A. I. Krylov, P. M. W. Gill and M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2006, **8**, 3172–3191.

44 G. R. Fulmer, A. J. M Miller, N. H. Sherden, H. E. Gottlieb, A. Nudelman, B. M. Stoltz, J. E. Bercaw, K. I. Goldberg and M. Beckman, *Organometallics*, 2010, 2176–2179.

45 G. M. Sheldrick, *Struct. Chem.*, 2015, **71**, 3–8.

46 O. V. Dolomanov, L. J. Bourhis, R. J. Gildea, J. A. K. Howard and H. Puschmann*, *Appl. Crystallogr.*, 2008, **42**, 339–341.

Chapter 4 : Investigating the Biomedical Potential of More Labile M_2L_4 Complexes

Chapter 4 : Investigating the Biomedical Potential of more Labile M_2L_4 Complexes

4.1 Introduction

4.1.1 Context

The stability, solubility and innate bioactivity are all key factors in developing coordination cages for biomedical applications. Whilst the research discussed up until this point has utilised the Co^{III} kinetically locked tetrahedra, it may be possible to utilise alternative cages as long as they maintain the same levels of structural integrity and bio-compatibility *in vivo*.

There has been significant research investigating the host-guest capabilities of water soluble cages, with key examples including Pd_6L_4 octahedra reported by the Fujita group and a Fe_4L_6 tetrahedron developed by Nitschke and co-workers (Figure 4.1).[1,2] Both systems possess rigid apolar ligands that form apolar cavities thus favouring the binding of guests via the hydrophobic effect. Nitschke and co-workers showed that a library of anionic guests could bind to the Fe_4L_6 coordination cage with K_a values ranging from 1 x 10^3 M^{-1} for PF_6^- to 1.1 x 10^6 M^{-1} for ClO_4^-.[1] The strong binding interactions are a product of electrostatic interactions, non-traditional hydrogen bonding and shape complementarity. Meanwhile, Fujita and co-workers described the encapsulation of neutral guests whereby the Pd^{II} octahedra encapsulated smaller adamantine derivatives in a 1:4 host:guest binding stoichiometry (Figure 4.1B) and larger neutral guests such as tetrabenzylsilane with a 1:1 model.[2]

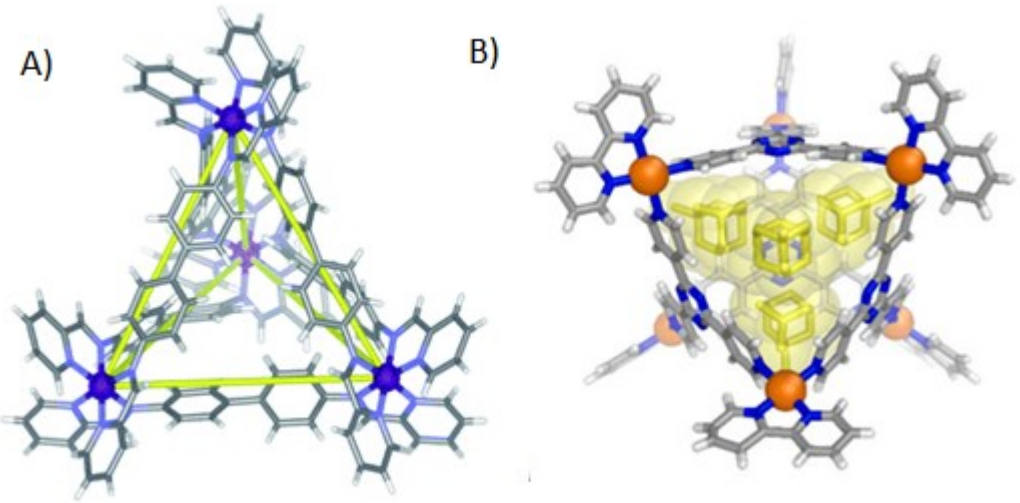


Figure 4.1. Water soluble cage examples. A) Fe based tetrahedron by the Nitschke group B) Pd based octahedral by the Fujita group with adamantine guest. Figures adapted from references.[1,2]

Cage water solubility and strong guest binding are both important factors when considering the biomedical applications of a potential coordination complex. Water accounts for 90% of the blood plasma composition, therefore a lack of water solubility would lead to a low effective concentration

of the coordination cage *in vivo*.[3] Whilst there is a vast body of research identifying water soluble cages and their host-guest capabilities, there is a substantially less work investigating whether those same cages would be suitable for biomedical applications.

M_2L_4 systems are becoming an increasingly popular class of cage since their inception by Steel and co-workers in 1998.[4] Preliminary studies by the Crowley group showcased a Pd_2L_4 cage capable of binding the chemotherapy agent cisplatin (Figure 4.2).[5]

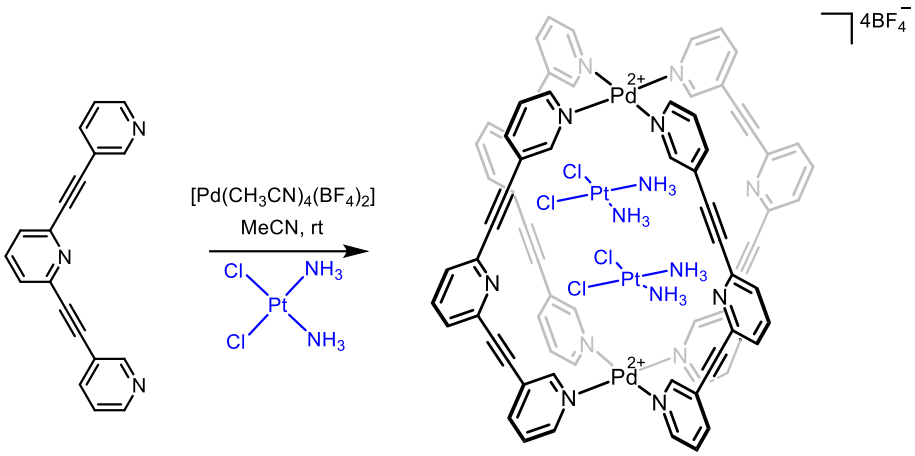


Figure 4.2. Pd_2L_4 cage capable of binding cisplatin adapted from work by the Crowley group.[5]

This phenomena has subsequently been utilised by the Casini group, who have exo-functionalised a Pd_2L_4 cage to control the bio-distribution of the cage and the encapsulated cisplatin.[6,5,7] Casini and co-workers subsequently published work involving the binding of the radioisotopic [^{99m}Tc]TcO_4^- to the Pd_2L_4 cage systems, and attempting to image the biodistribution of their cage in naïve mice through SPECT.[8] This included the bioconjugation of peptide PepH3, capable of blood brain barrier penetration, to the ligand via a solubilising polyethylene glycol linker, which was then assembled into a M_2L_4 cage using $Pd(BF_4)_2$ (Figure 4.3). The external functionalisation of the Pd_2L_4 cage apparently enabled the system to transgress the blood brain barrier in a way the un-functionalised cage was unable to. The encapsulation of [^{99m}Tc]TcO_4^- by similar Pd_2L_4 systems has been further validated by recent research by Reek and colleagues, who probed the biomedically relevant guest scope of the systems.[9]

When monitoring the biodistribution of the cage *in vivo* following the SPECT imaging studies, Casini and co-workers used ICP-MS to quantify the palladium content in the different tissues of the injected mice.[8] Subsequently, the palladium distribution was used to directly infer cage biodistribution, however the stability of these cages *in vivo* was not thoroughly investigated. Cage disassembly in the presence of bioreductants and salts must be considered over the course of the imaging study, Pd^{II} distribution may not equate to cage distribution if the system is unstable. In comparison to the kinetically locked Co^{III} cages, in which the bipyridine units form a strong chelating interaction to each

metal ion, the Pd_2L_4 systems are constructed via 8 monodentate Pd-pyridine bonds and therefore each ligand is more prone to displacement.

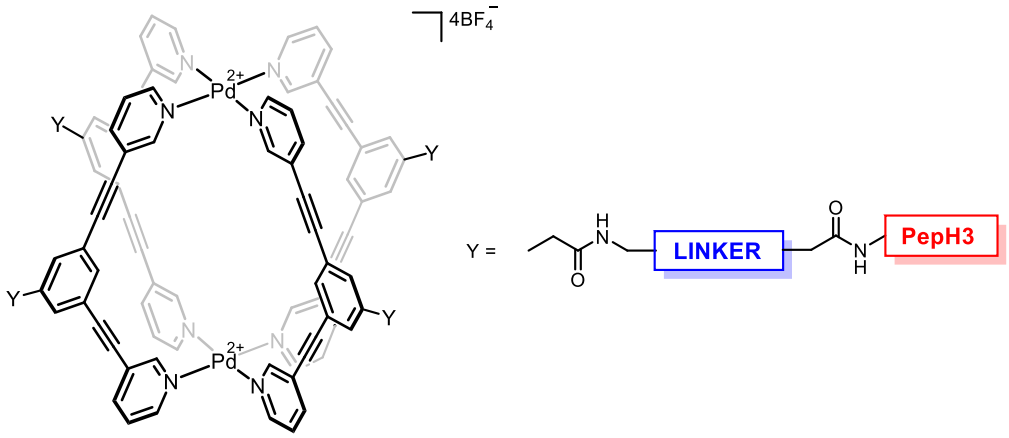


Figure 4.3. Bioconjugation of Pd_2L_4 cage with PepH3 to allow cage blood brain barrier penetration as described by Casini and co-workers.[8]

Following the inception of these M_2L_4 cages, there has been an effort to determine their stability against biologically relevant nucleophiles. Crowley *et al.* has shown that their cisplatin binding cage rapidly disassembles in the presence of chloride, with 8 Eq of the halide ions resulting in close to instantaneous breakdown, with a half-life ($t_{1/2}$) of less than 1 minute.[10] Subsequently, a mechanism of Pd_2L_4 cage disassembly in the presence of chloride ions was proposed including the displacement of ligands by the chloride ions and the formation of $[Pd_2L_2Cl_4]$ metallo-macrocycle.[10,11] However, there is limited understanding of how these systems behave *in vivo*.

4.1.2 Aims

The simple ligand synthesis and assembly of Pd_2L_4 cages makes them attractive candidates not only for biomedical applications but also for applications within catalysis and molecular recognition.[12,13] The aforementioned kinetically locked Co^{III} tetrahedra possess proven stability *in vivo* and strong association to bio-medically relevant guests, however their synthetic accessibility is not trivial. Thus, the aim of the research presented in this chapter was to probe the biomedical suitability of Pd_2L_4 cages that are typically used in the Lusby group for catalytic applications.[12,14,15] Additionally new ligand designs are proposed and investigated, with a focus on creating cage systems that possess the necessary attributes for biological applications; notably good stability and water solubility. Lastly, the host-guest chemistry of these cage systems will be discussed, including work with both anionic and neutral species, thereby determining whether the interactions are sufficient for delivery purposes *in vivo*.

4.2 Results and Discussion

4.2.1 Overall Strategy of Ligand and Cage Design

The ligands described in this chapter all possess the same two core structures as were first described separately by Hooley and Crowley.[16,17] Figure 4.4 visualises how the two underivatized cages are isostructural and only differ in the central aromatic spacer, wherein **C12** possess a pyridine ring whereas **C13** possess' a benzene ring. The various applications of **C12** have since been investigated, including recent research by the Lusby group evidencing the cages' ability to catalyse Michael addition reactions in the absence of a base. [12] Similarly **C12** has been awarded significant attention for its favourable host:guest and catalytic properties since it was first synthesised by the Crowley group.[16] However, whilst the guest scope of both the cage systems has been probed, their behaviour *in vivo* remains unknown. The aforementioned research completed by Casini and co-workers focuses only on the *in vivo* studies of functionalised cages.[7,8] To date there are no *in vivo* studies on **C12** and **C13**.

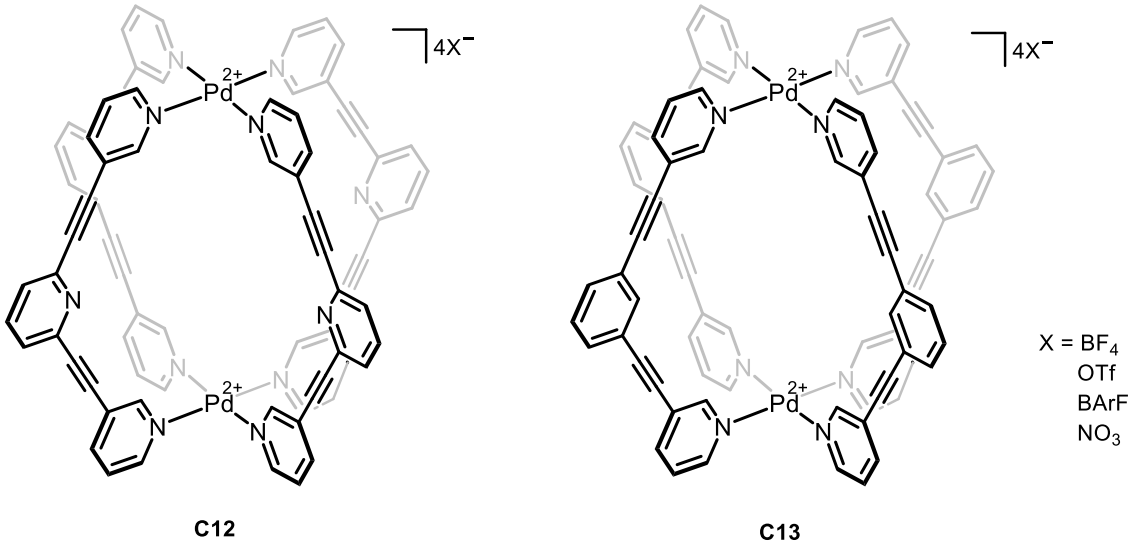


*Figure 4.4. Cage structures **C12** and **C13.***

The goals of this research were to increase the cages' water solubility and synthesise more strongly coordinating ligands (Figure 4.5). Extended aromatic ring systems are known for their hydrophobic nature and subsequent low water solubility, therefore the majority of research on Pd_2L_4 systems has been carried out in organic solvents.[5,12] Whilst water solubility can be improved via anion metathesis to more strongly hydrating anions such as NO_3^- (c.f., BF_4^- or $BArF^-$), the solubility of most P_2L_4 systems is still limited to MeCN, MeOH, dimethyl sulfoxide, or mixture of these solvents with water.[18] Improving the aqueous solubility of the cages by variation of the ligand removes the reliance on the counter anions, which is advantageous given the abundance of biological anions and their tendency to metathesise.

Citing a lack of water solubility, a recent study by Severin and co-workers showed that functionalisation of a Pd_2L_4 cage with a short PEG chain provided good water disolution.[13] Similarly, the functionalisation of a tri-aryl ligand with a short peptide improved the solubility of the resultant Pd_2L_4 system designed by Casini and co-workers.[8] Herein, an attempt to improve the water solubility of the cages by the addition of primary alcohols to the ligands was attempted; both a monosubstituted and disubstituted hydroxyl derivative were proposed (**L14** and **L15,** Figure 4.5). It is noted that by featuring the alcohol group on the central ring this may alter the movement of guests in out and of the cage, due to steric bulk. Moreover, the positioning of the methylene alcohol unit on the top and bottom of **L15** may both effect the Pd-Pyridyl bond and sterically hinder the interactions of both guests and biological species with the Pd^{II} centres.

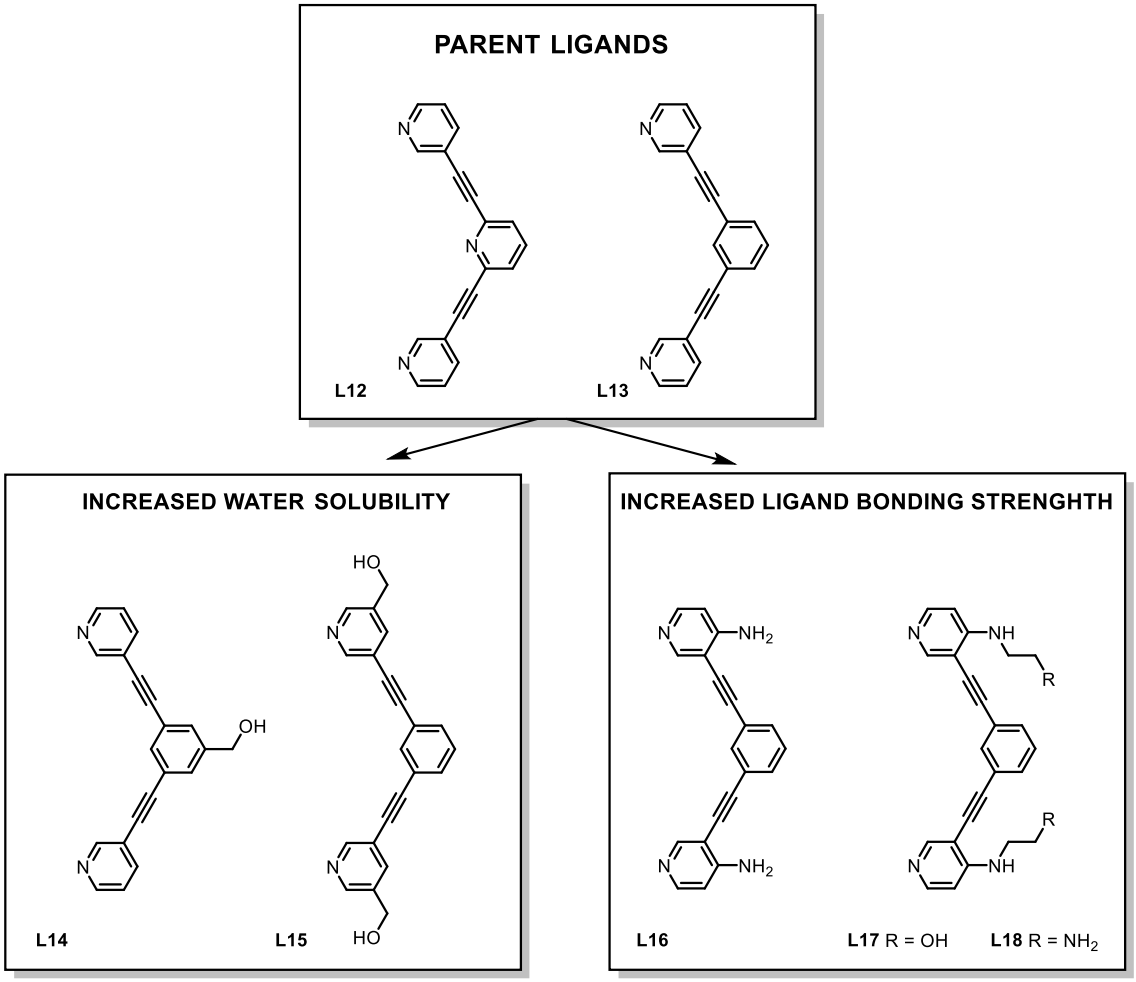


Figure 4.5. Proposed ligands.

The synthetic method proposed to make the ligands more strongly coordinating comes directly from the work with the Co^{III} cages, wherein attaching either an NH_2 or an N-alkyl group at the *para* position makes the coordinating pyridyl lone pair a stronger sigma donor.[19] A similar technique has been

exploited by the Crowley group, who found that external *ortho*-substituting the ligand with an amine improved the stability towards histidine and cysteine (Figure 4.6).[18,19] In this instance, the increased stability may be due to either (a) the increased donor strength of the pyridine ligand or (b) the steric bulk shielding the Pd^{II} from attack by the incoming nucleophile, or both. Although *ortho*-substitution lead to increased cage stability in the presence of biological nucleophiles and reductants, it also caused distortion of the cage which may affect the host-guest chemistry and lead to a decreased stability with respect to an equivalent non-distorted cage. The reported cage distortion is presumably due to steric repulsion of the crowded, substituted amine groups resulting in a twisting of the assembled ligands. In this work, we envisaged that the addition of amines in the *para*-position of the ligand would increase the strength of the coordination bond to the Pd^{II} centre via a stronger sigma donor, without sacrificing any stability due to twisting.

When considering the effect of external functional groups on the ligand and resultant cage solubility, **C13** was chosen as the comparison point. Therefore, the proposed ligands as shown in Figure 4.5 feature the central benzene ring as opposed to the pyridine ring system. This was decided to limit the number of ligands that were required.

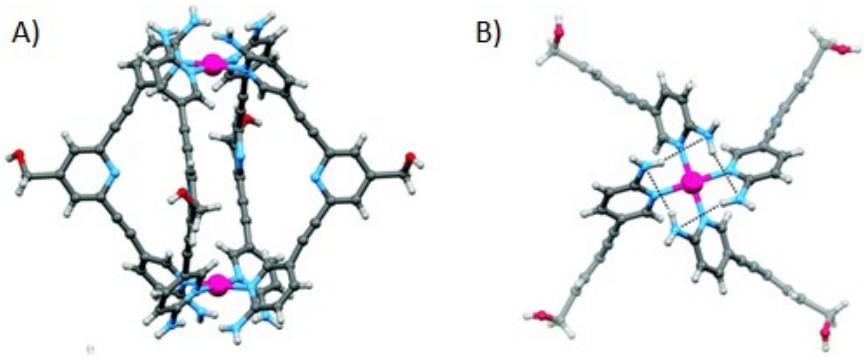


Figure 4.6. X-ray crystallography of Pd_2L_4 cage system designed by the Crowley group investigating bond strength when substituting an amine at the ortho-pyridyl position.[18]

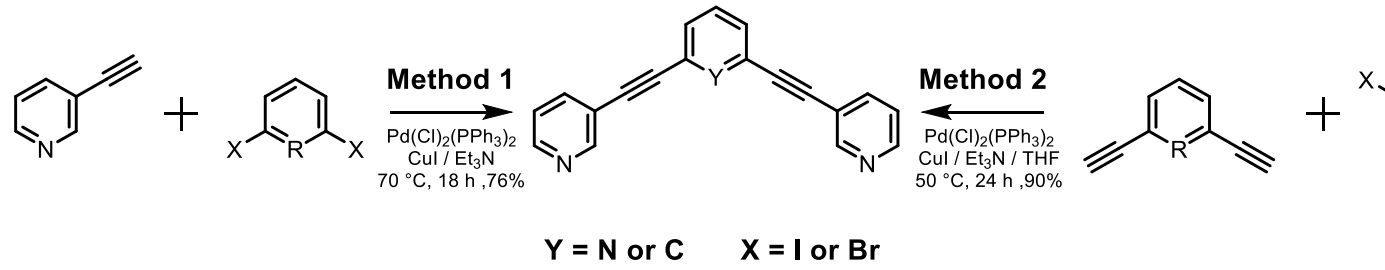
4.2.2 Synthesis

The synthetic section of this study in divided into two sub-sections. Firstly, the synthetic design of the ligands including altering of functional groups to control cage properties. Secondly, the cage assembly routes including the selection of Pd salts, conditions and determination of resultant cages. The properties of the ligands effect the nature of the coordination bond between the palladium and the ligand and thus ligand design directly affects cage assembly.

Ligand Synthesis

Synthesis of Unfunctionalised L12 and L13

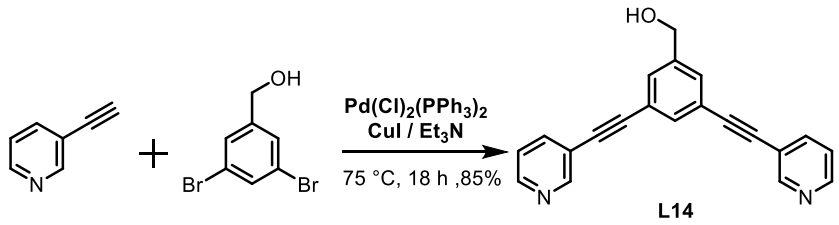
The synthesis of the non-derivatised ligands have been reported separately by the groups of Crowley and Hooley. These two groups use different methods (Scheme 4.1), whereby they differ only in the location of alkyne disconnection. Method 1, herein used to synthesis the two non-derivatised ligands was designed by the Hooley group, whilst method 2 was developed by the Crowley group.[16,17] The accessibility of these tri-aryl ligands makes them good candidates for pharmaceutical applications because the synthesis is high yielding and easily scalable. The synthesis also avoids more toxic compounds such as the organo-tin reagents required in Chapter 3.



*Scheme 4.1. Reaction Scheme showing the two synthetic routes to the non-derivatised ligands (**L12** and **L13**). Method 1 as developed by Hooley and Method 2 as developed by Crowley.* [16,17]

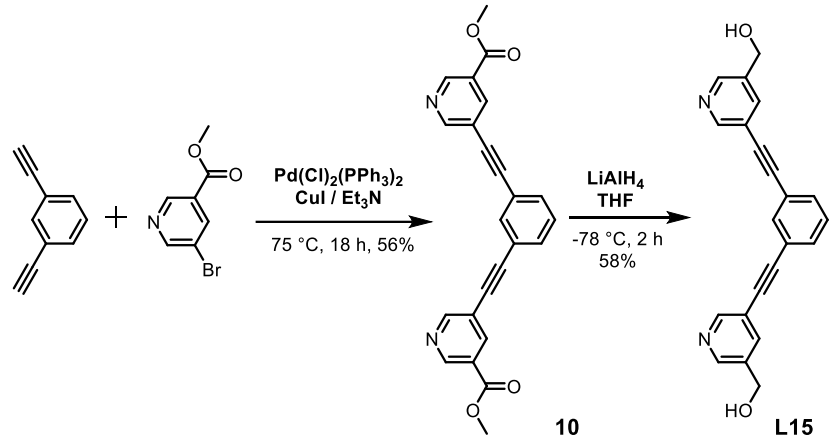
Synthesis of Polar Ligands for Water Soluble Cages

The synthesis of the more polar ligands were completed analogously to the two non-derivatised ligands, using 'Method 1'. Scheme 4.2 shows the successful synthesis of **L14** with a yield of 75% via a Sonogashira style cross-coupling reaction reported by Bakkari and Vincent.[20] **L14** solubility appeared to be similar to the two non-derivatised ligands with good dissolution in most common organic solvents including DCM, chloroform and THF, alongside increased solubility in more polar organic solvents, namely MeOH.



*Scheme 4.2. Synthesis of **L14** between ethynyl-pyridine and dibromobenzene* via *Sonogashira reaction with a yield of 85%.*

The synthetic route to **10** is outlined in Scheme 4.3 using 'Method 2' Sonogashira cross-coupling conditions, followed by a $LiAlH_4$ ester reduction of pendent ester groups to yield **L15**.



*Scheme 4.3. Synthesis of **L15**.*

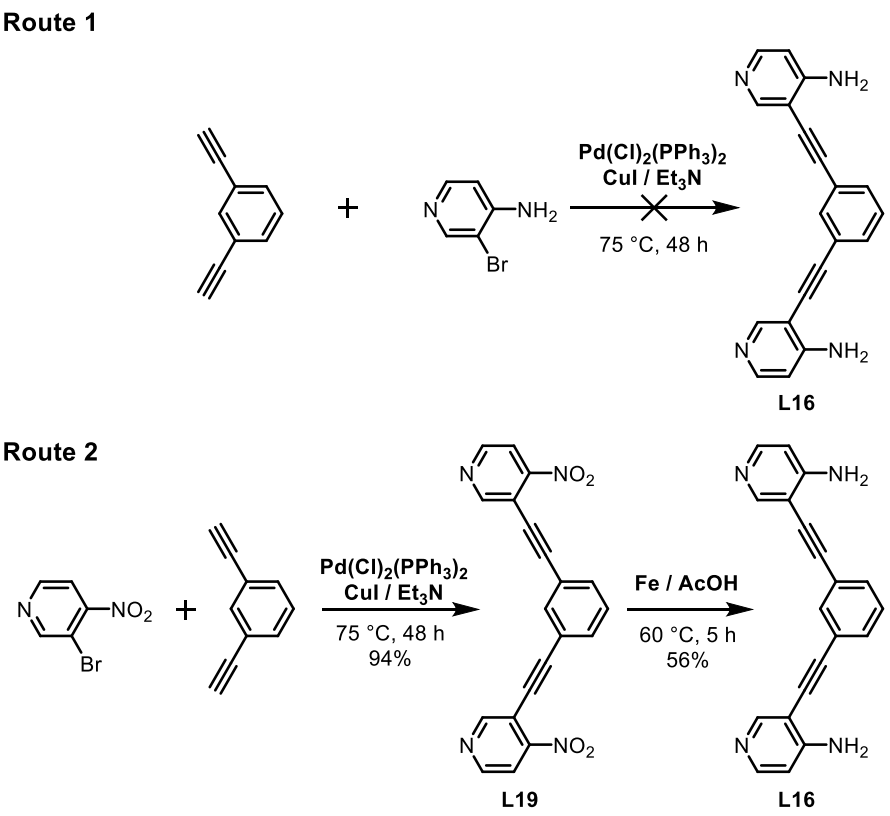
L15 possessed similar solubility to monosubstituted **L14** with increased solubility in more polar organic solvents such as MeOH and ethanol in comparison to **L12** and **L13**. Whilst both the methyl-alcohol mono and di-substituted ligands showed increased polar solvent solubility, neither were soluble in neat water. Polar solvent dissolution was attempted with both 100% deionised water and mixed aqueous solvent systems with up to 50% MeOH and 50% MeCN added. Given that both ligands possess extended aromatic structures, it is perhaps not surprising that the ligands themselves are not water soluble. However, it was hoped that the addition of the alcohol groups may still serve to water solubilise the assembled cages.

Synthesis of Ligands with Increased Pd Binding Strength

The synthesis of more strongly coordinating ligands with the aim to create more stable Pd_2L_4 cages was attempted using similar Sonogashira conditions. The novel "amine-ligand" (**L16** in Figure 4.5) was designed with the aim of investigating the effect of a *para*-positioned donating group on the strength of the Pd-pyridine bond.

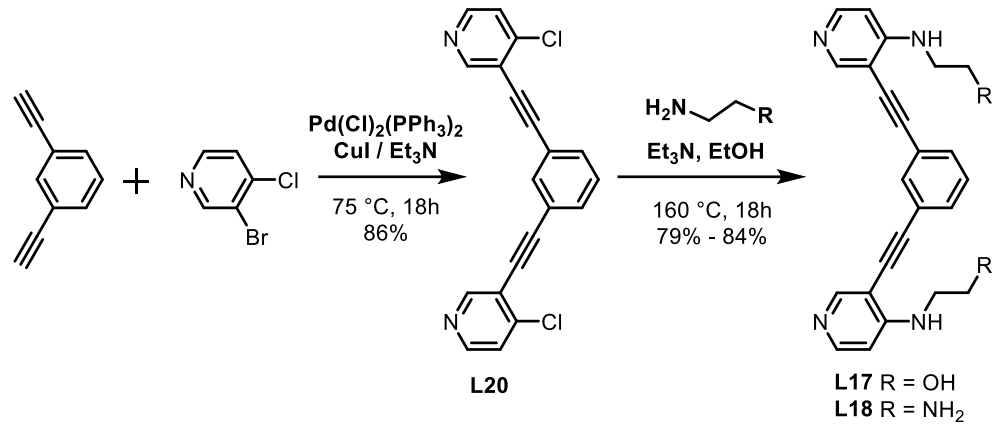
The same Sonogashira conditions that had been successful in preparing the other ditopic ligands proved ineffective for **L16** (Scheme 4.4). Upon heating the two reagents in the presence of $Pd(Cl)_2(PPh_3)_2$ and CuI in Et_3N for 18 h, reaction analysis indicated only the presence of starting materials. The concentration of diethynyl-benzene was decreasing over time but this was attributed to the instability of the starting material and not to formation of the required product. Upon studying the literature and considering the mechanism of the palladium cross coupling, it was considered that the electron donating properties of the amine might be deactivating the pyridyl-halide ring towards oxidative addition because the amine group increases the electron density such that the carbon-bromine bond is strengthened.[18,21]

To overcome the issue of an inactive pyridyl halide, route 2 was attempted whereby an electron withdrawing nitro group is used in place of the deactivated amine group (Scheme 4.4). **L19** was successfully obtained in high yield using the same Sonogashira conditions as previously described and purified yielding bright yellow crystals (94%). The iron reduction method discussed by Carlsson et al. was applied here and features a variation on the Bechamp reaction using iron filings in acetic acid.[21] Although when tracking the reaction using ^1H NMR it appeared to have run to completion, the resultant yield was relatively low (56%). It was hypothesised that the iron could be coordinating to the ligand and subsequently hindering ligand recovery.



*Scheme 4.4. Reaction schemes for the attempted and successful syntheses of **L16.***

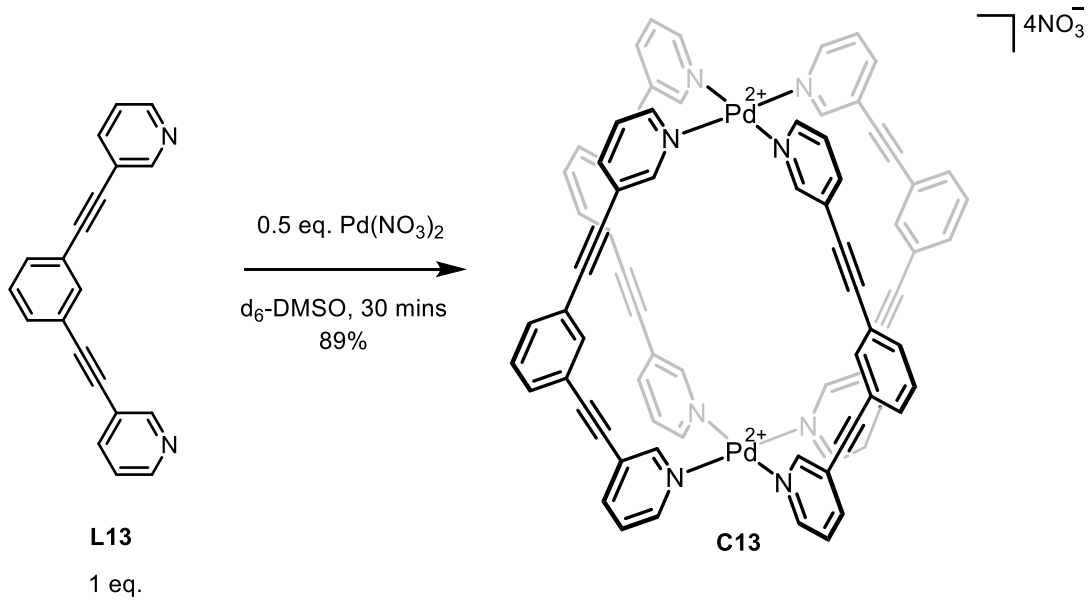
Following the success of ligand functionalisation via S_NAr in Chapter 3, it was employed here to synthesise **L17** and **L18** described in Scheme 4.5. **L20** was obtained in high yield (86%), further evidencing the increased activity of pyridyl halides substituted with electron withdrawing groups in Sonogashira coupling reactions. The S_NAr reaction was then completed on the purified **L20** using ethanolamine for **L17** and ethylenediamine for **L18**. Traditional S_NAr conditions utilising Et_3N as a base and ethanol as a solvent proved successful for the synthesis of ligands **L17** and **L18**, in comparison to the S_NAr conditions required in Chapter 3 whereby the reaction proceeded in neat ethanolamine (or equivalent amine). **L17** displayed the best polar solvent dissolution and whilst it was still not soluble in 100% deionised water, it was partially soluble in a 1:1 mixture of water and MeCN. Both **L16** and **L18** showed good dissolution in MeOH but lower polar solubility in comparison to **L17.**



*Scheme 4.5. Synthetic Route to **L17** (R = OH) and **L18** (R = NH_2).*

Cage Assembly

Different research groups have detailed various assembly methods for the synthesis of Pd_2L_4 systems.[2,22] The preliminary work by the Hooley group described mixing stoichiometric quantities of **L13** and palladium nitrate in dimethyl sulfoxide-d_6 for 0.5 h to obtain **C13** (Scheme 4.6).[22] These conditions were repeated with palladium triflate and palladium tetraphenylborate salts producing pure cage in high yields. However, the high boiling point of dimethyl sulfoxide (189 °C) makes this synthetic route less suitable for subsequent isolation of the cage samples.



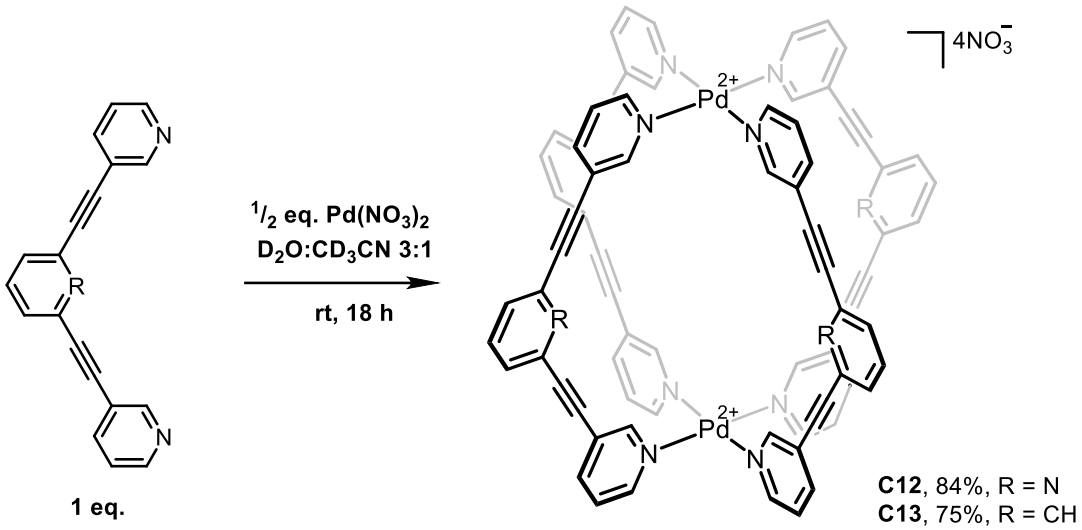
Scheme 4.6. Pd_2L_4 Cage Assembly Conditions Optimised by Hooley and co-workers.[22]

Meanwhile, the Crowley group employed similar assembly conditions with a 1:2 ratio between the palladium salt and the ligand in dimethyl sulfoxide, which were left to assemble up for up to 6 h and the resultant cage precipitated using either ethyl acetate or diethyl ether.[18] The yields obtained from these cage assembly reactions were relatively low (57%) due to material lost during precipitation. Alternatively, when assembling a large Pd^{II} octahedron Fujita and co-workers described heating a mixture of the ligands with the palladium nitrate salt in water to first allow the cage to assemble and then obtaining the solid cage via the slow evaporation of water.[2]

Cage assemblies are generally evidenced using ^{1}H NMR whereby upon complexation of the metal centres there is a change in the magnetic environment of the ligand protons. More specifically, the protons closest to the coordination bond typically show a shift downfield as they become more deshielded by the 2+ charge on the Pd.

Direct Assembly using Pd Nitrate Salts

Direct assembly was performed on **L12** and **L13** as shown in Scheme 4.7, using nitrate counter anions to increase the water solubility of the resulting cage. Stoichiometric amounts of the ligands were stirred with $Pd(NO_3)_2$ in a mixture of 1:3 CD_3CN and D_2O. This solvent system was chosen because the cage was predicted to be soluble whilst the ligands were sparingly soluble thus any remaining free ligand could be removed via filtration. Cage dissolution in the reaction solvent mixture shifts the equilibrium towards the product as well as making the purification process easier. Moreover, MeCN and water are both relatively volatile, in comparison to dimethyl sulfoxide and thus allow access to the pure solid cage.



*Scheme 4.7. The formation of Pd_2L_4 cages **C12** and **C13**.*

The assembly of **C13** was attempted first on an NMR scale and ^1H spectra were recorded at interval time points to track the rate of cage assembly (Figure 4.7). The ligand and Pd salt were mixed together prior to recording the first time point t = 0 h with an approximate dead time of 15 min. Cage self-assembly appears to be almost complete by the first time point, indicating that the assembly reaction takes place instantaneously. When observing the two **L13** resonances at 8.5 ppm and 8.7 ppm it is noted that they are still present at t = 0 h and only seem to completely diminish after 48 h, this is attributed to the presence of a slight excess of ligand. This preliminary NMR scale assembly can also be used to evidence the stability of the cage over time, whereby for the 48 h period the cages stays assembled and in solution. Good solution stability is required within biomedicine as a long half-life is required to administer the drug and for it to disperse throughout the organism.

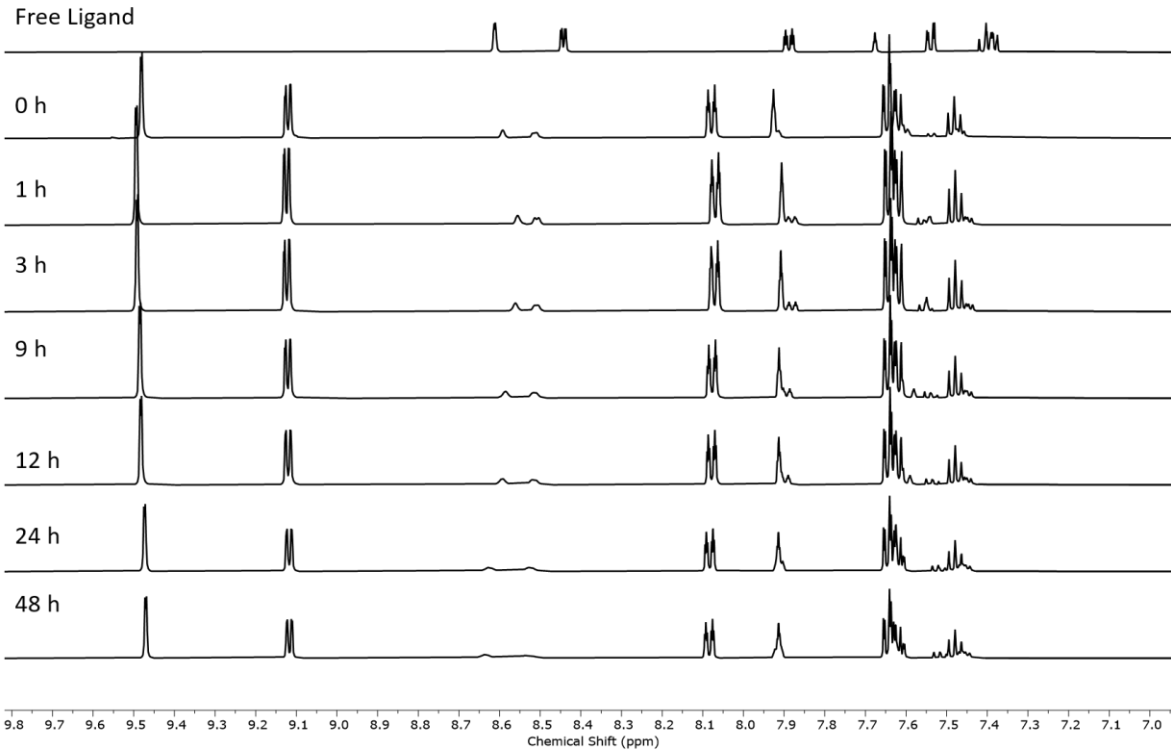


*Figure 4.7. ^{1}H NMR (500 MHz, 3 D_2O : 1 CD_3CN) spectroscopic data showing **C13** assembly mixture of **L13** and $Pd(NO_3)_2$ over time.*

Following the success of the preliminary NMR-scale studies, the same conditions were then repeated with for **C12**; **L12** and $Pd(NO_3)_2$ (2:1) were sonicated in water and MeCN before stirring for 18 h. The ^{1}H NMR showed the formation of a high symmetry species with chemical shifts indicative of cage formation, whereby all proton signals become deshielded upon cage assembly with the two protons adjacent to the coordinating pyridine most affected (Figure 4.8). The two non-derivatised cage systems were successfully assembled and isolated in good yields of 79% and 84% for **C13** and **C12** respectively. Both cages were further characterised by ^{13}C NMR, DOSY and MS.

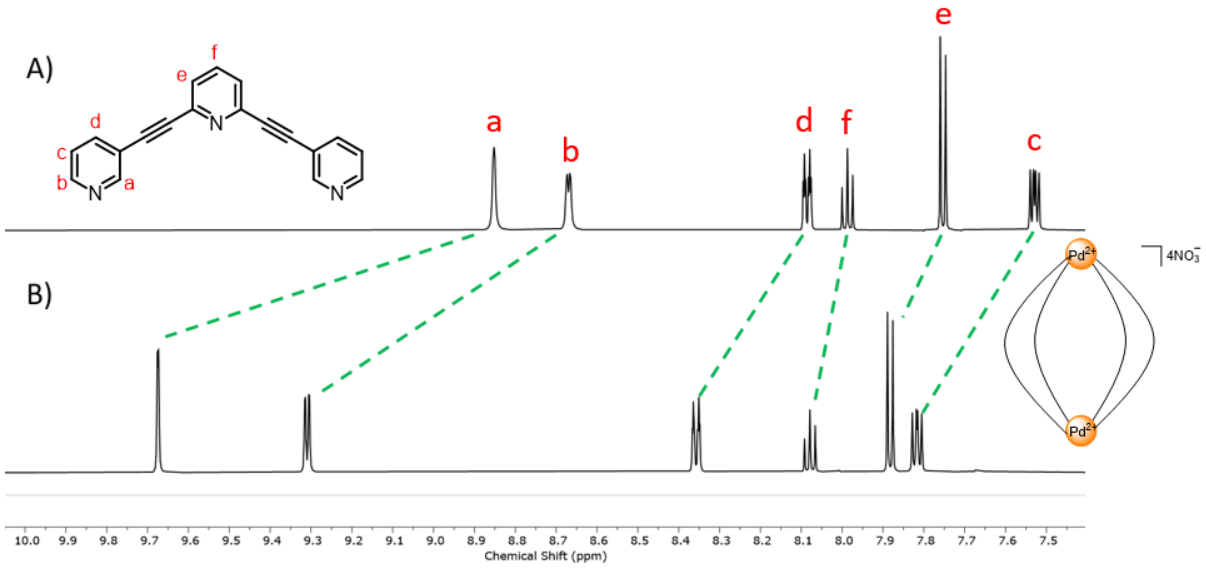
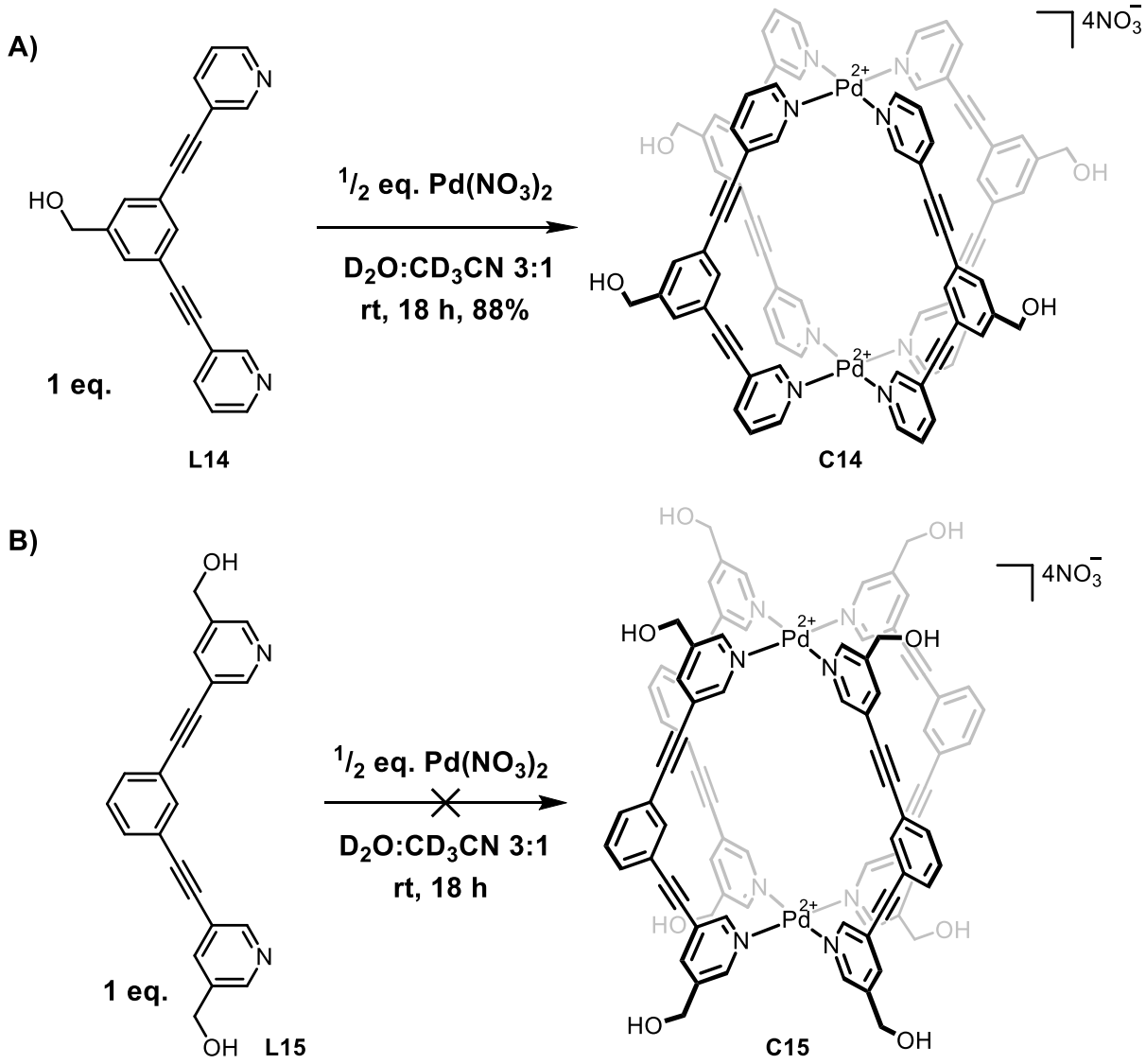


*Figure 4.8. ^{1}H NMR (500 MHz, 3 D_2O : 1 CD_3CN) **C12** assembly spectra A) **L12** in CD_3CN and B) **C12**. The green dashed lines connect the protons from their chemical shift in the free ligand to the chemical shift of the assembled cage.*

Unfortunately neither of the two non-derivatised cage systems were subsequently water soluble, analogous to the previous research conducted on these cage systems by Hooley and co-workers who indicated the Pd_2L_4 cage was only soluble in DMSO.[17,22] However, upon experimenting with the solvent compatibility of **C12** and **C13**, the cages were found to be soluble in mixed aqueous solvents with either 50% MeOH or MeCN. Whilst this solubility is perhaps better than expected, it is not acceptable for biomedical applications, which tolerate only very small amounts of DMSO or MeOH. In an attempt to access more water soluble cages the same assembly conditions were then attempted with **L14** and **L15** (Scheme 4.8).



*Scheme 4.8. Reaction schemes for assembly of water soluble cage systems **C14** and **C15.***

The reaction occurred successfully with **L14** generating **C14** in yield of 88% (Scheme 4.8), with the corresponding ^{1}H NMR showing the characteristic deshielding of the ligand protons upon assembly. Interestingly, the free ligand showed broad signals for the two *ortho*-pyridyl protons (Figure 4.9A)

which upon complexation resolved into sharp peaks. Presumably due to the halting of a dynamic ligand interaction upon cage assembly.

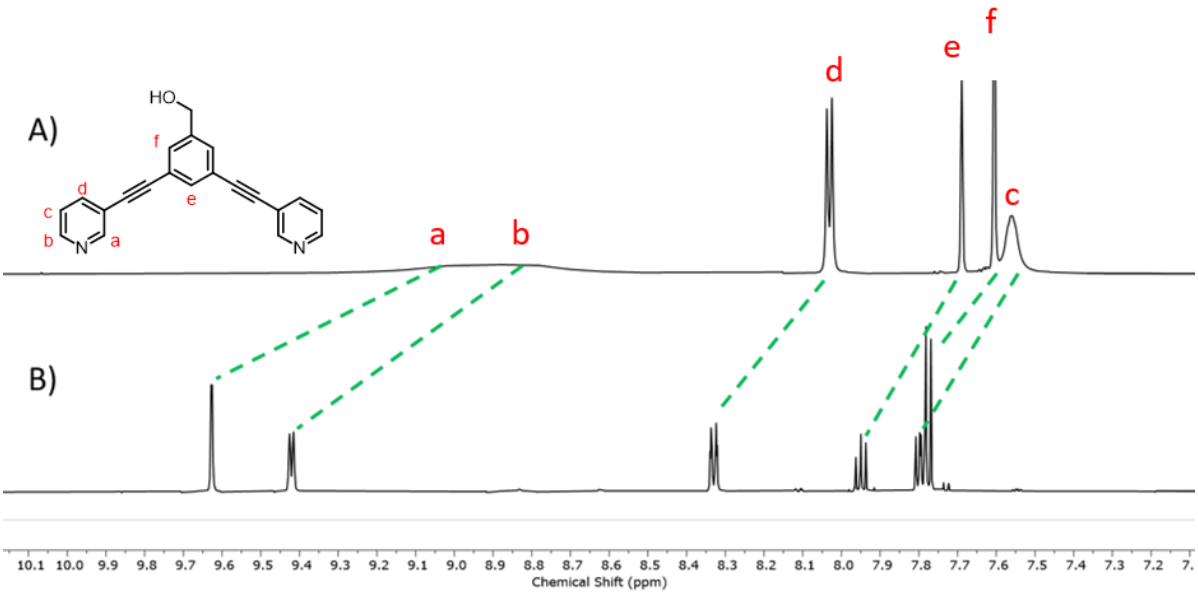


*Figure 4.9. ^{1}H NMR (500 MHz, 3 D_2O: 1 CD_3CN) **C14** assembly spectra A) **L14** in CD_3CN and B) **C14** in 3 D_2O : 1 CD_3CN. The green dashed lines connect the protons from their chemical shift in the free ligand to the chemical shift of the assembled cage.*

Unfortunately, the assembly reaction with **L15** did not yield cage (or at least a cage that could be isolated). Treatment of **L15** with $Pd(NO_3)_2$ (2:1) in water and MeCN (3:1) led to a heterogeneous mixture, where the ^{1}H NMR showed neither cage nor ligand (Figure 4.10B). In an attempt to solubilise the mixture d_6-DMSO was added, however, no spectral changes were observed (Figure 4.10C). One possible explanation is the formation of large insoluble oligomers, whereby instead of forming small discrete structures the ligand coordinates to the palladium and forms longer polymeric species. Typically larger oligomers can have low solvent solubility and thus this could explain the heterogeneous mixture. It was considered that since **L15** and **L14** have similar solubility, **C15** need not be pursued. Moreover, a complicated cage assembly did not align with the synthetic convenience that made these systems attractive.

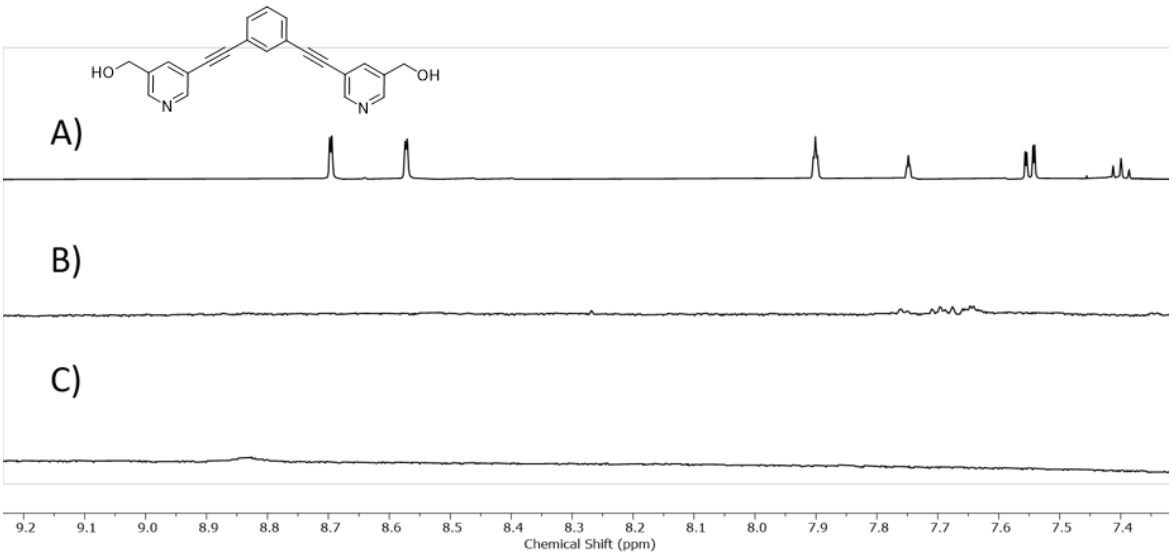
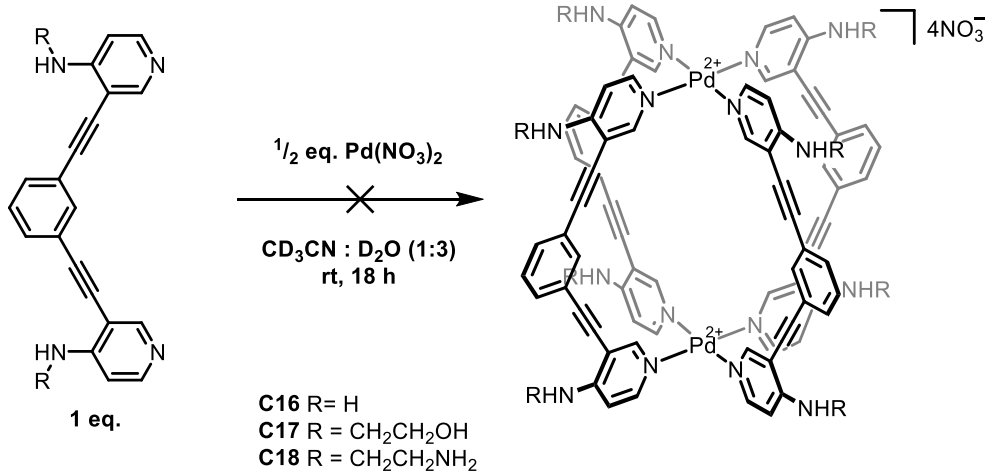


*Figure 4.10. ^{1}H NMR (500 MHz, 3 D_2O : 1 CD_3CN) Attempted assembly of **C15** A) **L15** in CD_3CN B) $Pd(NO_3)_2$ and **L15** in 3 D_2O : 1 CD_3CN C) $Pd(NO_3)_2$ and **L15** in 3 D_2O : 1 CD_3CN with 10% d_6-DMSO*

Next, direct cage assembly was attempted using the more strongly coordinating ligands; **L16**, **L17** and **L18**. Initially the same conditions that worked for the previous cages were used, however these were unsuccessful (Scheme 4.9).



Scheme 4.9. Assembly conditions for the formation of electron rich cages using $Pd(NO_3)_2$

There appeared to be minimal formation of **C16** when **L16** was treated with $Pd(NO_3)_2$ in D_2O and CD_3CN for 18 h (Figure 4.11). **L16** was minimally soluble in the solvent mixture used for the cage assembly so the free ligand was recorded in deuterated chloroform. Figure 4.11B shows the assembly reaction mixture, which highlighted by the orange dashed lines shows the ligand amongst a mixture of other species. The broadening in the baseline of (Figure 4.11B) is presumed to be a result of a multitude of

species as opposed to a discrete Pd_2L_4 assembly. It was hypothesised that the Pd-pyridine bond may be less reversible as a result of the stronger sigma bond donation of the ligand, subsequently leading to the formation of more insoluble oligomers as the system is unable to rearrange to from the thermodynamic product as it would with the non-derivatised systems. In an attempt to overcome this, the system was heated for 24 h (Figure 4.11C), however this lead to an even broader spectrum with no obvious increase in sharp signals that could be attributed to the Pd_2L_4 product. It is noted that the broad humps appearing in the range of 9.4 ppm to 10 ppm are analogous with the chemical shifts normally seen upon Pd complexation of these systems (Figure 4.11C). However, given the broadness of the spectrum indicating a mixture of species, it was concluded that the cage was not cleanly assembling.

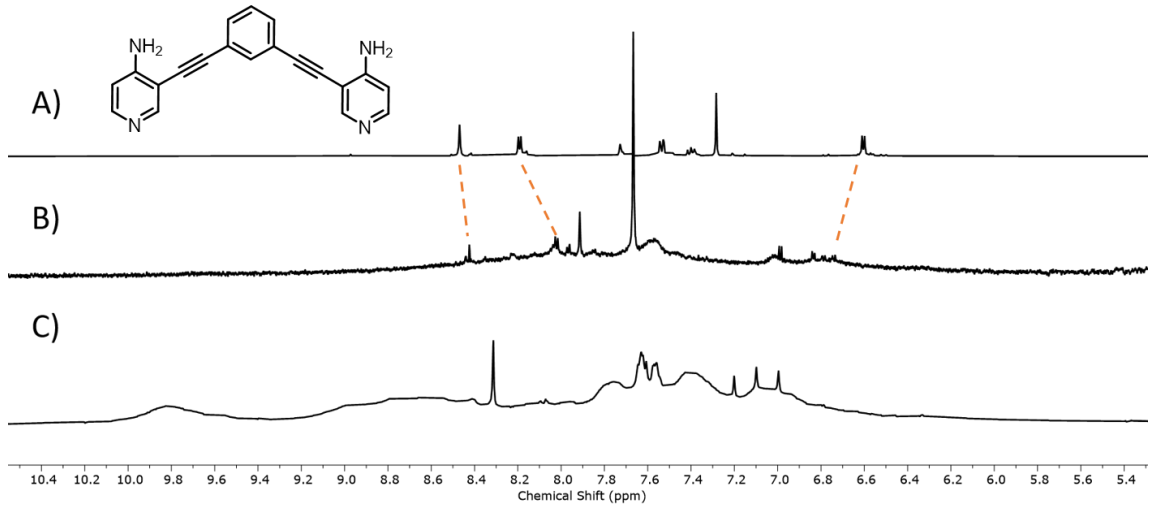


Figure 4.11. ^{1}H NMR (500 MHz, A: $CDCl_3$, B/C: $3D_2O$ $1CD_3CN$) Attempt at ***C16***$\cdot 4NO_3^-$ *assembly A)* ***L16*** *($CDCl_3$). B) Assembly mixture after the addition of 0.5 Eq $Pd(NO_3)_2$ to* ***L16*** *C) (B) after 24 h heating at 60 °C*

A similar phenomena occurred with **L18**; when attempting to assemble a Pd_2L_4 cage with $Pd(NO_3)_2$, only the free ligand could be seen in solution. Previous research by the Lusby group suggested that this could be a result of the primary amine interacting with the $Pd(NO_3)_2$ and preventing it from coordinating with the pyridine. In future research, this could be investigated by protecting the amine, such as with a tert-butyloxycarbonyl group.

An attempt was made to assemble **L17** using the same conditions as the two previous ligand assembly experiments. The ligand was partially insoluble in the aqueous reaction mixture and thus the free ligand is shown in chloroform, as shown in Figure 4.12A. In contrast to the previous experiments with **C16**, this reaction shows a single symmetrical species with the same number of aromatic proton environments as the free ligand. Interestingly, the ligand assembly spectrum showed a mixture of signals that are deshielded and shielded in relation to the free ligand. Given the propensity of the two

protons adjacent to the pyridine to significantly de-shield upon Pd^{II} complexation, it is difficult to conclude simply from the proton NMR spectra whether assembly occurred.

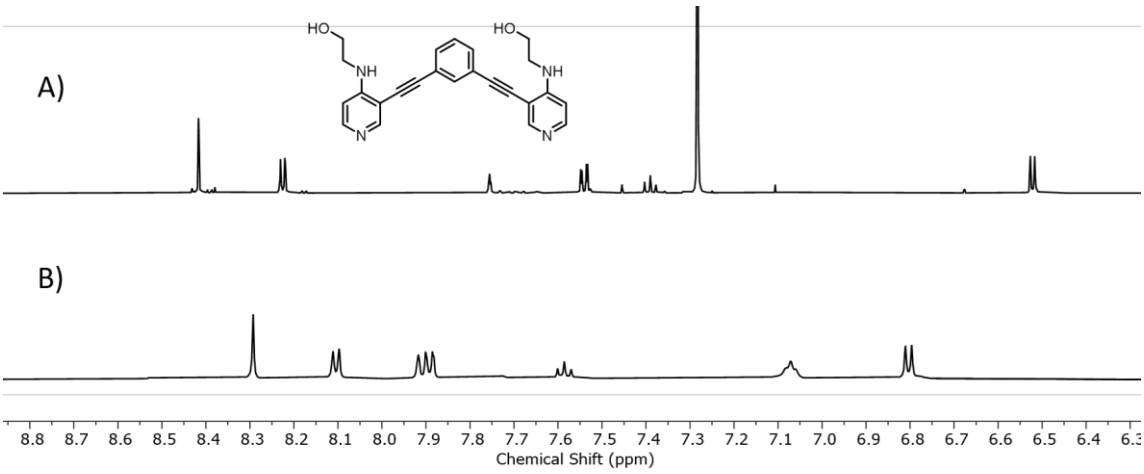


*Figure 4.12. ^{1}H NMR (500 MHz, A: $CDCl_3$, B: $3D_2O$ $1CD_3CN$) Attempt at **C17** assembly A) **L17** and B) Assembly mixture after the addition of 0.5 Eq $Pd(NO_3)_2$ to **L17.***

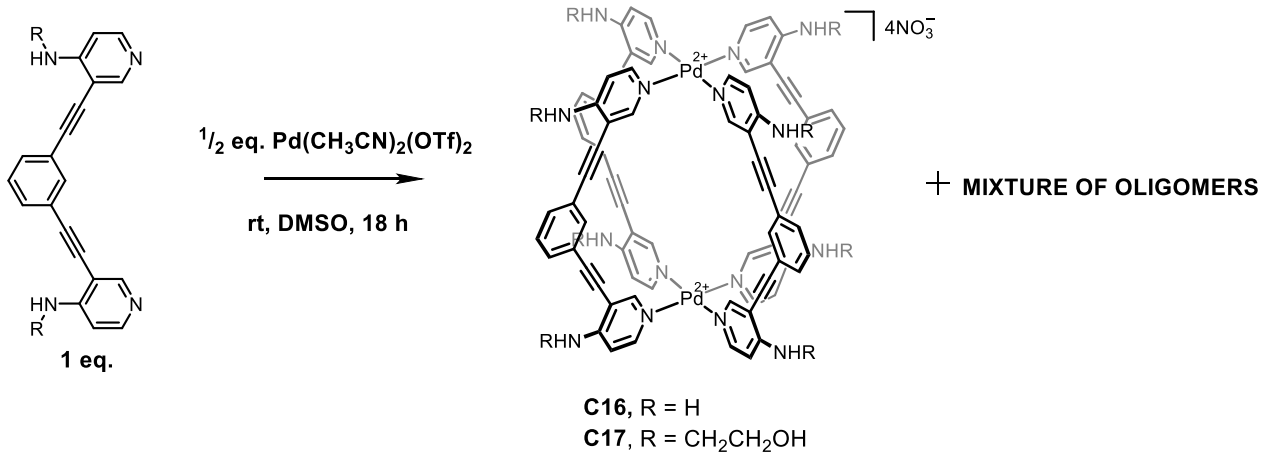
Initially nanospray MS was utilised to help determine the composition of the **L17** assembly, however the spectrum showed only singularly charged free ligand and no multiply charged complexation species. Whilst this indicated the amine ligand assembly may have been unsuccessful, the lack of highly charge species in the spectrum could also be a result of the formation of a difficult to ionise species. Subsequently ^{1}H DOSY NMR was attempted and a diffusion coefficient extracted of $1.15e^{-10}$ m^2sec^{-1} indicating a molecular weight of approximately 2260 g/mol and an atomic radii of 11.0 Å. Given the atomic radii of **C13** was 10.5 Å, this indicated that the assembly may have been successful. However, due to the lack of supplementary 1D NMR and MS data, it is not possible to conclude with full certainty the assembly of **L17** into **C17**, under the conditions attempted.

In summary, the assembly reactions for the two non-derivatised ligands (**L12** and **L13**) and the hydroxyl-monosubstituted **L14** were successful. However, the attempted assembly of the four ligands (**L15**, **L16**, **L17** and **L18**) with altered functionality on the coordinating rings proved more complex. Presumably the electron donating nature of the *para*-position amine group, present in **L16**, **L17** and **L18**, increased the strength of the pyridine sigma donor bond sufficiently to hinder the reversibility required to assemble the Pd_2L_4 cages. This does not account for **L15** assembly whereby the *meta*-methylhydroxyl is not expected to substantially influence the resultant pyridine sigma bond strength. Alternatively, it is hypothesised that the positioning of the substituted group in closer proximity to the pyridine, in comparison to the amine functionalised ligands, could kinetically hinder the Pd^{II} coordination.

Direct Assembly using Alternative Pd Salts

Given the inability to assemble **L16**, **L17** and **L18** into cages using $Pd_2(NO_3)_2$, assembly reactions using different palladium salts were attempted. Different counter-anions can have a huge effect on the cage assembly and resultant cage properties including stability and solubility.[23] It was considered that alternative palladium salts might serve to better template the formation of the cage and a subsequent anion metathesis of the cage could be carried out to obtain the nitrate cages.

Previous research by the Lusby group has utilised $Pd(OTf)_2$ for cage assembly conditions.[24] A similar method was used here whereby a 2:1 mixture of ligand and $Pd(OTf)_2$ were mixed in d_6-DMSO and the ^{1}H NMR was recorded, Scheme 4.10. This was attempted for both **L17** (Figure 4.13) and **L16** (Figure 4.14).



*Scheme 4.10. Assembly conditions for the formation of **C16** and **C17** using $Pd[MeCN]_2(OTf)_2$.*

For assembly of **L17**, the ^{1}H NMR showed the presence of many signals in the aromatic range (Figure 4.13), which could be caused by multiple species and/or low symmetry species. Interestingly, the same spectrum shows distinctive deshielded and sharp peaks between 10.5 ppm and 10.9 ppm, which could be a result of the assembled Pd_2L_4 system (highlighted by the orange dashed line). Potentially, the presence of the triflate anion either acts as a template for cage formation, or facilitates the reversible binding of the pyridine ligands. Nonetheless, the multitude of signals indicates that the reaction did still not reach thermodynamic equilibrium.

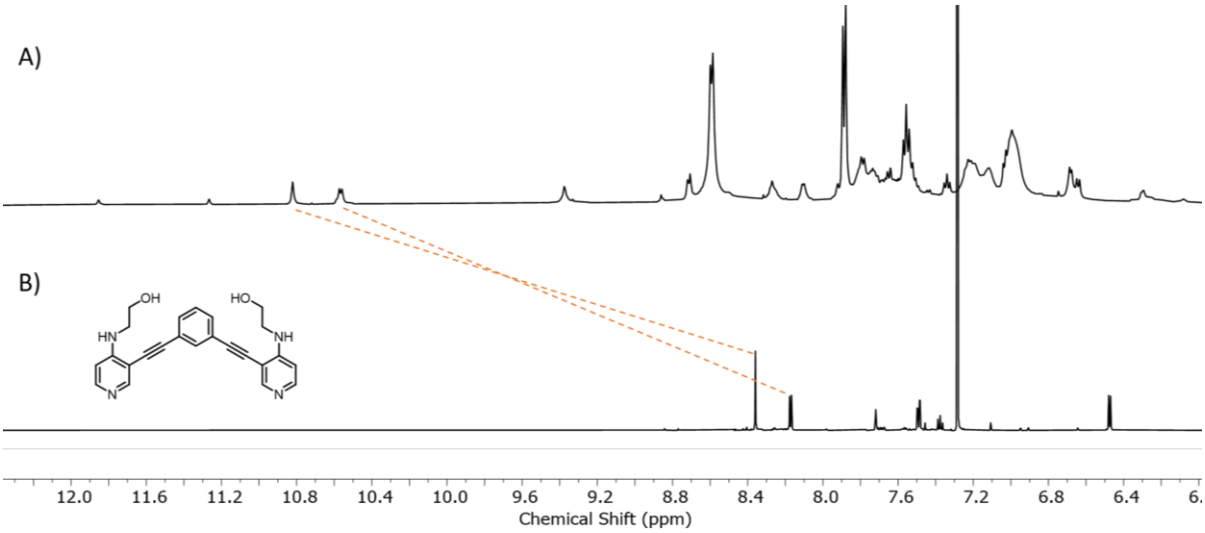


*Figure 4.13. ^{1}H NMR (500 MHz, A: d_6-DMSO, B: $CDCl_3$) Attempt at **C17**·4OTf$^-$ assembly A) Assembly mixture after the addition of 0.5 Eq Pd[MeCN]$_2$(OTf)$_2$ to **L17** and B) **L17.***

In the assembly of **L16** with Pd(OTf)$_2$[MeCN]$_2$, the ^{1}H NMR spectrum showed the reaction mixture was significantly cleaner (Figure 4.14). However, the spectrum shows double the number of proton environments in the aromatic region than would be expected for a Pd_2L_4 cage assembly. This increased number of signals is likely either the result of a mixture of two species, or the formation of a low symmetry discrete assembly. The integrals of the aromatic peaks were determined as approximately equal indicating either a 1:1 mixture of two symmetrical species or one discrete assembly with two magnetically distinct ligand environments.

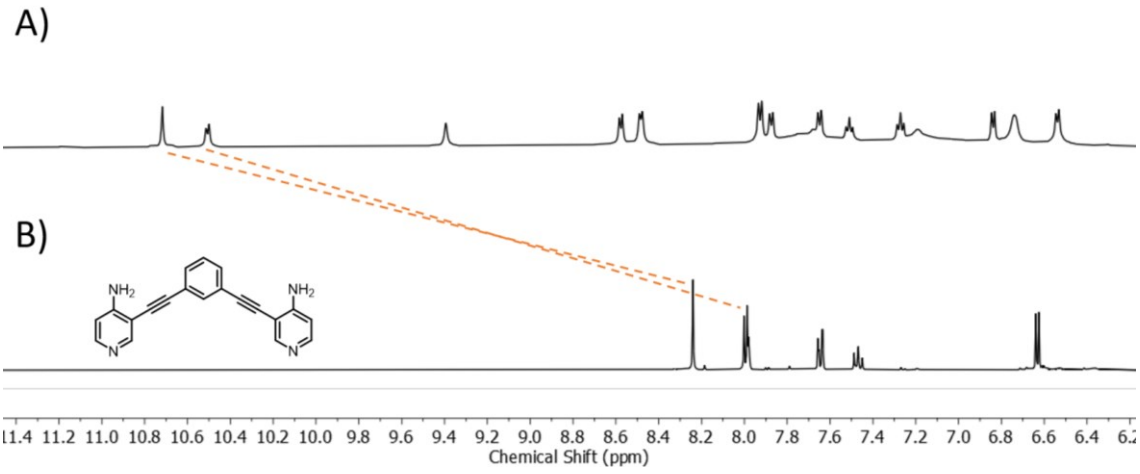


*Figure 4.14. ^{1}H NMR (500 MHz, A: d_6-DMSO, B: $CDCl_3$) Attempt at **C16**·4OTf cage assembly A) Assembly mixture after the addition of 0.5 Eq Pd[MeCN]$_2$(OTf)$_2$ to **L16** and B) **L16***

Hiraoka and co-workers recently published research documenting how different anionic guests could lead to the formation of various interlocked Pd_xL_{2x} assemblies, as shown in Figure 4.15.[25] In their study they describe a tetranuclear interlocked palladium cage composed of two binuclear cages, quadruply interlocked together. They detail that not only does the interlocked cage system lose its' intrinsic symmetry but the deshielding effect of coordination is increased. Similar interlocked species have been obtained by the Clever group whereby they observed that the gradual conversion of a Pd_2L_4 assembly into an interlocked Pd_4L_8 geometry when the reaction was heated over time.[26] It therefore seems plausible that the formation of an apparent low symmetry species from **L16** could be explained by an interlocked structure.

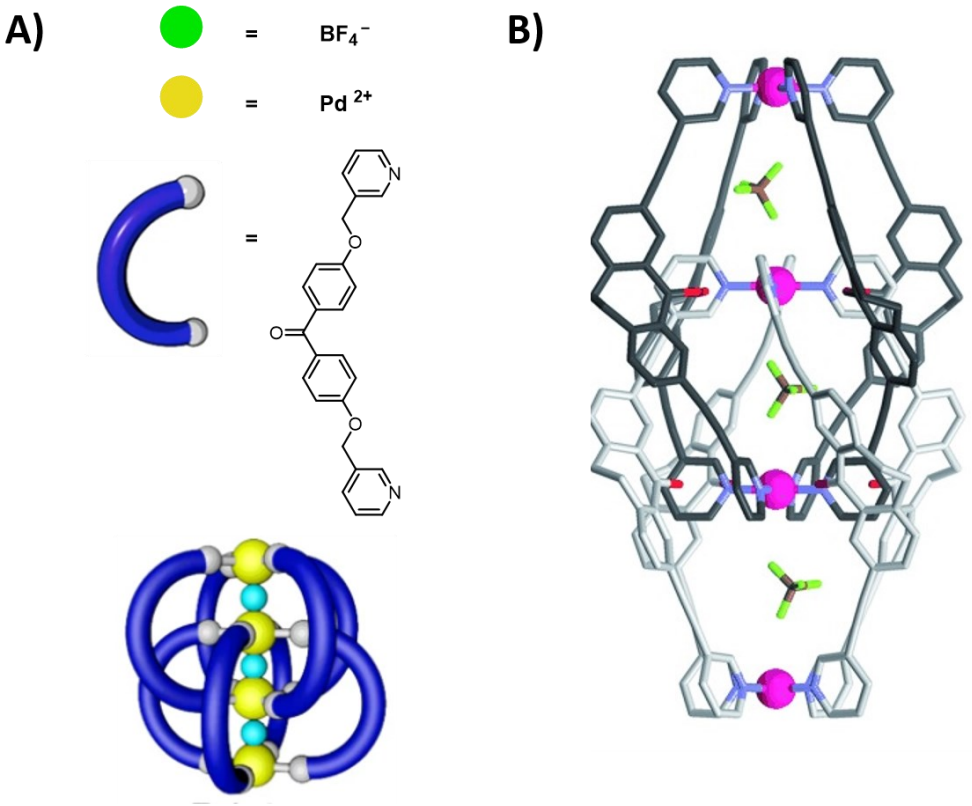


Figure 4.15. Interlocked Pd_4L_8 cages as described by A) Hiraoka and co-workers and B) Clever and co-workers. Figures adapted from a 2019 study by Hiraoka and colleagues and a 2020 study produced by Clever and colleagues.[25, 26]

In an attempt to better characterise the assembly product of both **L16** and **L17**, DOSY experiments were carried out in d_6-DMSO. The ^{1}H DOSY NMR of **L16** assembly mixture (Figure 4.16) shows that the aromatic signals may be a result of multiple species. This is highlighted by the two horizontal sets of signals (red and blue dashed lines, Figure 4.16) with diffusion coefficients of $8.9e^{-11}$ $m^2\,sec^{-1}$ and $7.3e^{-11}$ $m^2\,sec^{-1}$ corresponding to atomic radii of 12.3 Å and 15.0 Å respectively. When considering the error in the DOSY experiment, and the relatively small size difference, it was concluded that this data was not clear enough to disprove the existence of a single species.

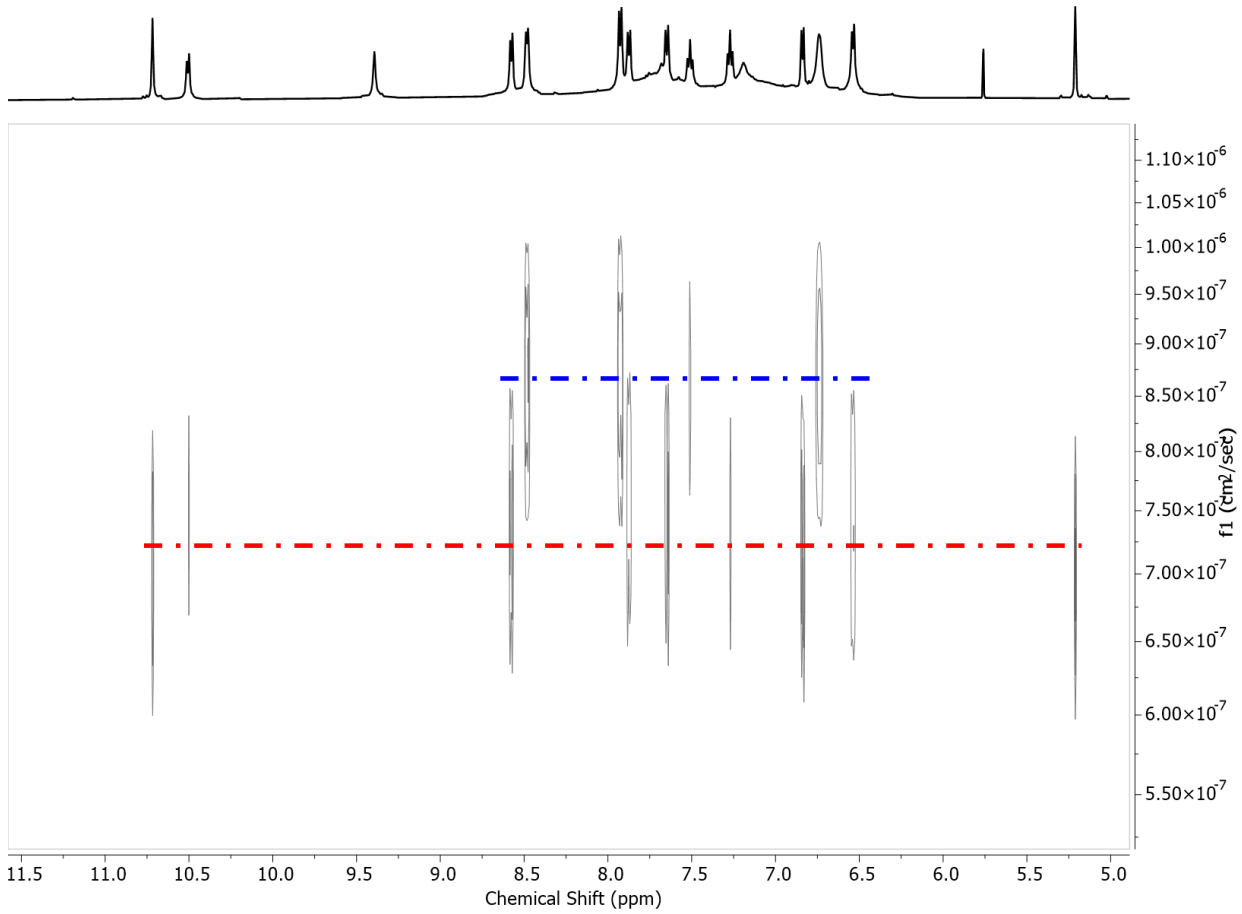


*Figure 4.16. ^{1}H NMR (500 MHz, d_6-DMSO) Assembly mixture of 0.5 Eq $Pd(OTf)_2[MeCN]_2$ and **L16**. Dotted lines indicate mean point for the two apparent species.*

Similarly, the ^{1}H DOSY for the **L17** assembly mixture did not fully define whether the assembly reaction was a mixture of species or a more complex coordination geometry. Given the broad nature of the spectrum it was difficult to distinguish single points (Figure 4.17), and a large band can be seen in the DOSY spectrum corresponding to an atomic radii of around 16 Å. This is larger than that of the simple **C12** and **C13** cages (approximately 11 Å). Independently, this data is not sufficient to determine what structure(s) had been obtained; the increased size could be a result of either the formation of the Pd_2L_4 geometries with increased external functionalisation or the generation of an alternative, larger structures (*i.e.*, interlocked Pd_4L_8 structures).

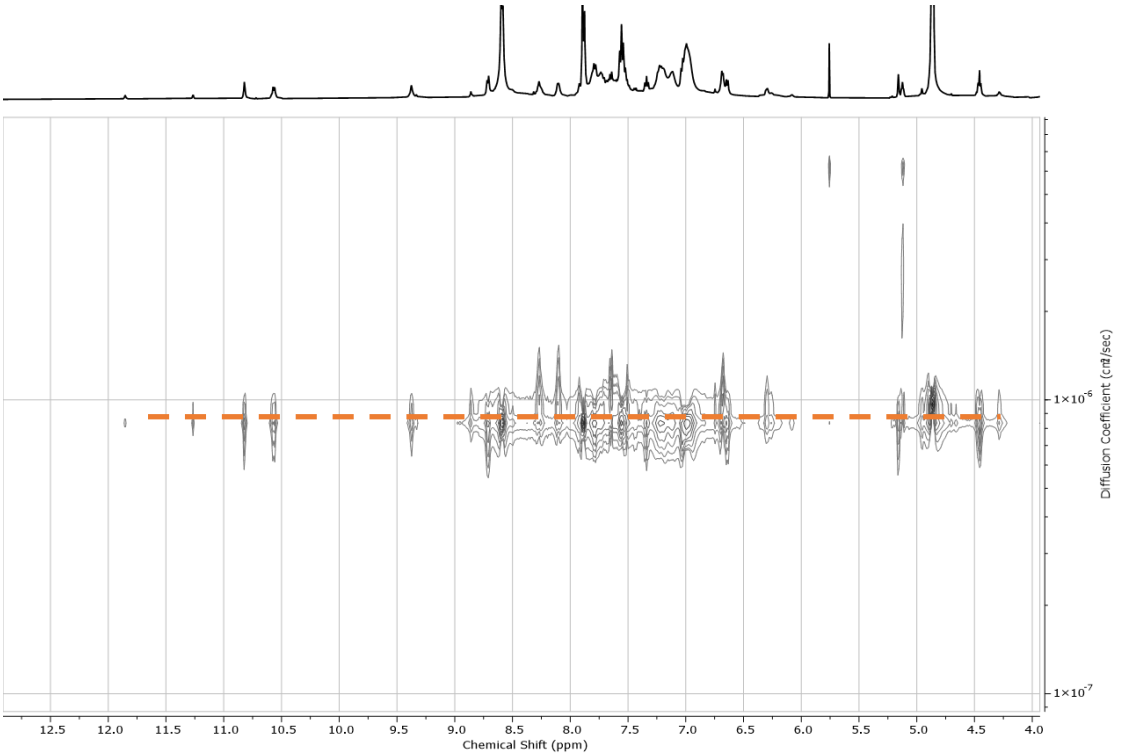


*Figure 4.17. ^{1}H NMR (500 MHz, d_6-DMSO) Assembly mixture of 0.5 Eq $Pd(OTf)_2[MeCN]_2$ and **L17**. Dotted lines indicate mean point for the two apparent species. Dotted line indicates mean point for apparent species.*

Nanospray ESI mass spectra of the **L17** and **L16** assembly reactions were also obtained (Figure 4.18 and Figure 4.19, respectively). Whilst both mass spectra show a multitude of signals the major species for each assembly mixture aligns with the 2+, 3+ and 4+ charge states for the two Pd_2L_4 cages. The full break down of each charge state can be found in experimental and figures in the Appendix.

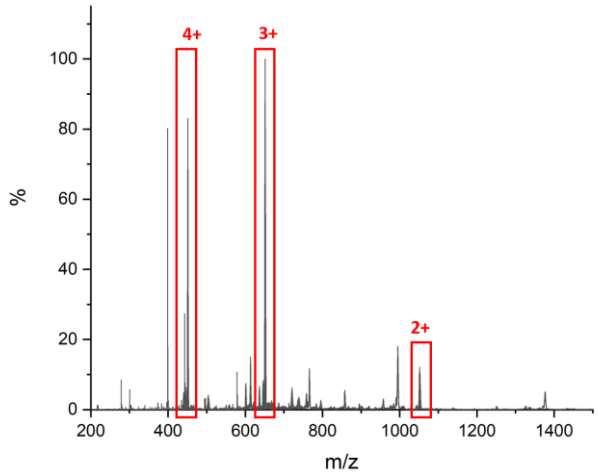


*Figure 4.18. Full mass spectrum of **C17**·4OTf assembly mixture containing 0.5 Eq $Pd(OTf)_2[MeCN]_2$ and **L17.***

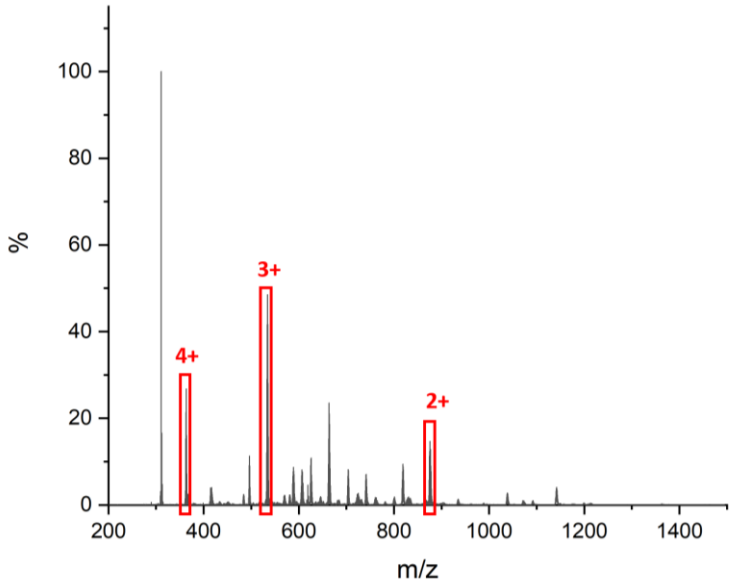


*Figure 4.19. Mass spectrum of the **C16**·4OTf cage assembly mixture (0.5 Eq $Pd(OTf)_2[MeCN]_2$ and **L16**.*

Although, MS provides evidence that the Pd_2L_4 cages have been formed, it must be considered that multiple different cage assemblies can give rise to identical m/z values. For example, ionisation of Pd_4L_8 through the loss of four OTf^- counter anions would have the same mass over charge signal as Pd_2L_4 losing two counter anions. The difference in these species would only manifest in the isotopic distribution of the signal. This is explored in Figure 4.20, whereby a comparison is made between the predicted isotopic distribution for the 2+ Pd_2L_4 cage and the 4+ Pd_4L_8 cage charge states. The narrow 0.25 m/z isotopic separation of a 4+ Pd_4L_8 signal is not observed in the experimental spectrum, Figure 4.20C, indicating the sole presence of the Pd_2L_4 structure. MS is not a quantitative technique and high intensity signals only indicates that the species ionised easily and not that it is the most prominent species in solution. Therefore, there may be other coordination geometries in solution that are not ionising (or are breaking down) under the same conditions and therefore are not detectable.

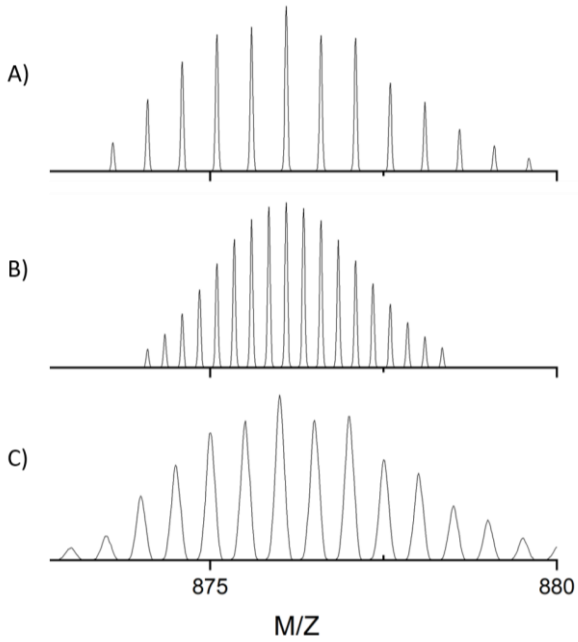


*Figure 4.20. ESI-MS m/z signals for A) Predicted signal for 2+ Pd_2L_4 of **C16**·4OTf B) Predicted signal for 4+ Pd_4L_8 cage and C) the **C16**·4OTf cage assembly mixture (0.5 Eq $Pd(OTf)_2[MeCN]_2$ and **L16)** indicating 2+ charge state for Pd_2L_4*

In combination, the NMR experiments and MS point to the formation of the desired **C16** and **C17** Pd_2L_4 systems, however, it is likely that these only exists as part of a complex mixture of geometries. While there was not sufficient time to further explore these reactions, it could be possible that size exclusion chromatography could be used to separate and purify the different complexation structure. Alternatively, guest templation could be used to bias the reaction mixtures towards a single species. As an example, Fujita and co-workers attempted to assemble a platinum octahedron and instead obtained a kinetically trapped mixture of oligomers.[27] They subsequently showed that the mixture resolved to give a single species upon the addition of sodium adamantine carboxylate, which acts as a strong binding guest. They also showed that once assembled, the guest could be removed with affecting the cage structure. Presumably a similar method could be used here to resolve the apparent mixture of species, potentially by adding a known guest of these cage systems such as benzoquinone.

4.2.3 Stability Studies of Cages

When investigating complexes for their biomedical applications the biological half-life ($t_{1/2}$) must be considered. This is defined as the time required for the concentration of a biologically active substance to decrease from the maximum plasma concentration (c_{max}) by half. This can vary significantly and is often difficult to calculate accurately. For example, the platinum based chemotherapy agent Oxaliplatin which is used in the treatment of colorectal cancers has a half-life of 14 min.[28] This short half-life was determined by Ehrsson *et al.* using liquid chromatography and photometric detection and contradicted previous research which used ICP-MS to track the concentration of Pt in the blood leading to a $t_{1/2}$ = 189 h.[28] They concluded that Oxaliplatin only remained intact as the pre-administered complex for a short amount of time but that the free Pt remained in the blood for significantly longer leading to the drastically different half-times. The breakdown of Oxaliplatin was due to reactions with small biological species such as glutathione which acts as a bioreductant and larger bio-abundant species like albumin.[28] In the case of Oxaliplatin, the activity of the drug is dependent on platinum coordinating to DNA and thus activity is not affected by the short half-life of the intact drug. When considering the use of coordination cages for encapsulation and delivery of bio-medically relevant guests, the stability of the three dimensional structure within plasma is imperative. This section focuses on stability of the cages in the presence of two major components of blood plasma; salts and bioreductants.

Salt Stability

Blood is composed of a mixture of blood cells (45%) and plasma (55%), with plasma an aqueous mixture of ions, proteins and nutrients.[3] Na^+ and Cl^- are most prevalent with concentrations of ~140 mM and ~105 mM, this concentration of NaCl equates to approximatly 0.9% w/v; the same concentration found in saline.[29] The prevalence of various salts in blood plasma is shown in

Table 4.1. When considering radiolabelling, the process is often completed within saline solutions. Traditionally generators will produce a solution of [$^{99m}TcO_4^-$]TcO_4^- in saline and radiolabelling will occur in the presence of an excess of salt. This section investigates the stability of the assembled cages in solutions of NaCl with concentrations up to 0.9% as would be encountered upon guest encapsulation and intravenous *in vivo* administration.

Table 4.1. Various Ion Concentrations found in human blood plasma

Ion	Plasma Concentration mmol/L
Na^+	142.0
K^+	5.0
Mg^{2+}	1.5
Ca^{2+}	2.5
Cl^-	103.0
HCO_3^-	27.0
HPO_4^{2-}	1.0
SO_4^{2-}	0.5

Salt Stability of C13 and C14

Provisionally an attempt was made to dissolve **C13** in deuterated phosphate buffer saline (0.1 M) however, this lead to a heterogeneous mixture. This could be explained by either the low intrinsic solubility of the cage, or cage disassembly, which would give the insoluble ligand.

To determine the tolerance of **C13** to NaCl under homogeneous conditions, saline experiments were carried out in a mixture of d_6-DMSO and water, and followed by ^{1}H NMR spectroscopy (Figure 4.21). For comparison purposes, the first two spectra show free ligand and assembled cage in d_6-DMSO. Initially, **C13** was first dissolved in d_6-DMSO before slowly adding deuterated saline (0.9% m/v), reaching a final ration of d_6-DMSO:saline of 4:96. Upon addition of the saline to the solution of cage in DMSO, a precipitate formed spontaneously and the ^{1}H NMR spectrum of the solution showed no **C13** was present (Figure 4.21C). The precipitate was isolated and dissolved in DMSO (Figure 4.21D), which by comparison to the authentic sample was determined to be the free ligand.

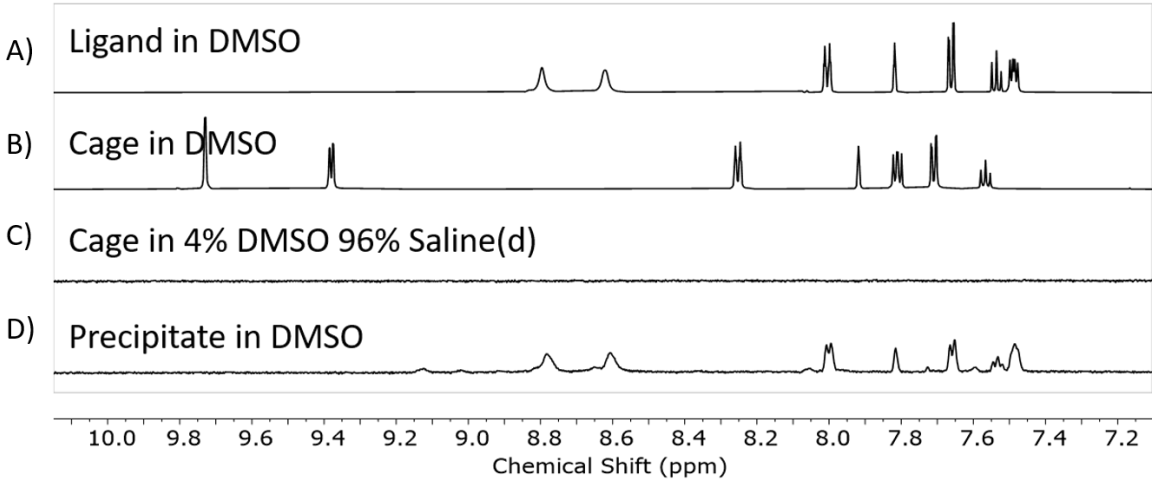


*Figure 4.21. ^{1}H NMR (500 MHz, d_6-DMSO) Investigating effect of NaCl on **C13**. A) Free **L13** B) **C13** C) **C13** in with 96% deuterated saline D) Isolated precipitate from (C)*

The most likely explanation for the cage disassembly is displacement of the ligands by the Cl^- anions, generating tetrachloroplatinate, as had previously been suggested by Crowley (Figure 4.22). [5,11] They also showed that by adding an excess of $AgSbF_6$ to the reaction mixture the Cl^- anions are scavenged, releasing the Pd^{II} cations and regenerating the cage.

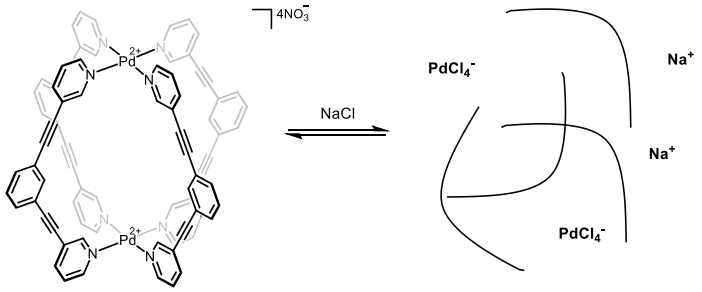


Figure 4.22. Proposed mechanism for cage disassembly in the presence of NaCl.

However, given that **C13** was insoluble in 96% saline solution the experiment was repeated using lower concentrations of saline such that a homogenous solution was maintained throughout. An aliquot (50 µL) of deuterated saline was added to a solution of **C13** in d_6-DMSO (450 µL) (Figure 4.23A). As can be seen in Figure 4.23B the cage completely disassembles into free ligand with just 0.09% saline, a tenth of the concentration found in plasma. The spectrum was recorded with a dead time of 10 min implying that the disassembly of the cage was almost instantaneous when exposed to NaCl.

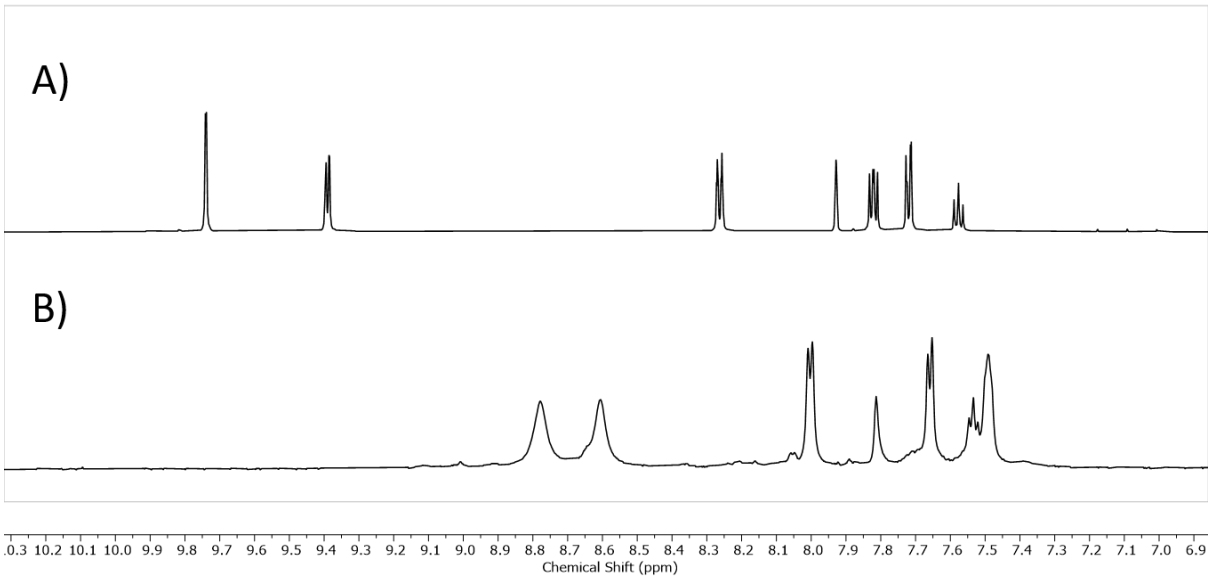


*Figure 4.23. ^{1}H NMR (500 MHz, d_6-DMSO) Investigating effect of NaCl on **C13**. A) **C13** B) **C13** and 0.09% (w/v) saline*

The addition of deuterated saline to a solution of cage in DMSO was repeated with **C14** to investigate whether solubility of the cage or external functionality would affect the cages interaction with salt. A series of time-point experiments were recorded to track cage disassembly however the first spectrum recorded 15 min after the addition of deuterated saline (90% d_6-DMSO and 10% saline) showed complete disassembly to **L14**, which was maintained over the following 12 h (Figure 4.24).

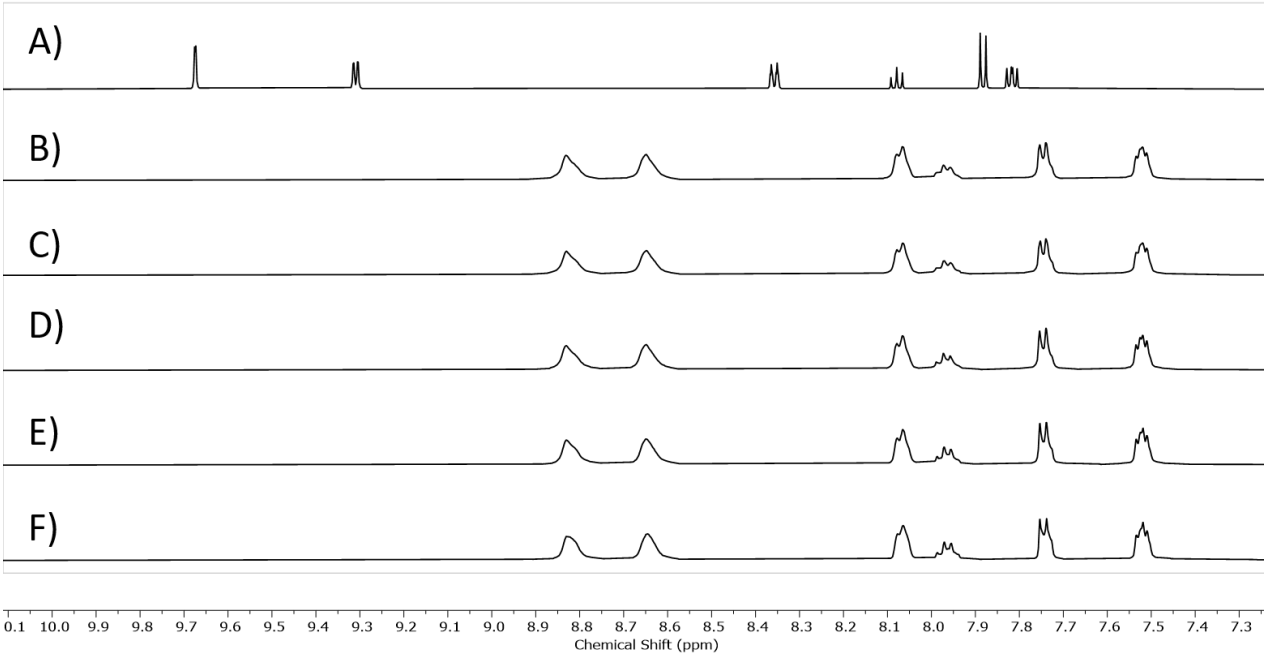


*Figure 4.24. ^{1}H NMR (500 MHz, d_6-DMSO) Investigating effect of NaCl on C**14**. A) **C14** B-D) **C14** after the addition of 10% saline at time points B) t = 0.25 h C) t = 1 h D) t = 2 h E) t = 6 h F) t = 12 h.*

Both **C13** and **C14** indicated a saline half-life of less than 10 min given that complete breakdown occurred before the first ^{1}H NMR time point.

Salt Stability on **C12**

The same experiments were repeated for **C12**, to investigate whether an altered interior cavity would affect the NaCl stability. Under conditions with a majority DMSO solvent (90% d_6-DMSO and 10% saline (0.9% (w/v)), the initial ^{1}H NMR spectrum showed only partial decomposition, with 50% intact cage alongside free ligand (Figure 4.25). Over time, the proportion of ligand increased however, this process was relatively slow. Overall, the kinetic profile appears to follow an initial rapid breakdown in the first 15 min followed by more gradual disassembly. Whilst the system does show instability in the presence of salt, it appears remarkably more stable than both **C13** and **C14**. It is not obvious why such a small difference in structure would lead to such different behaviour.

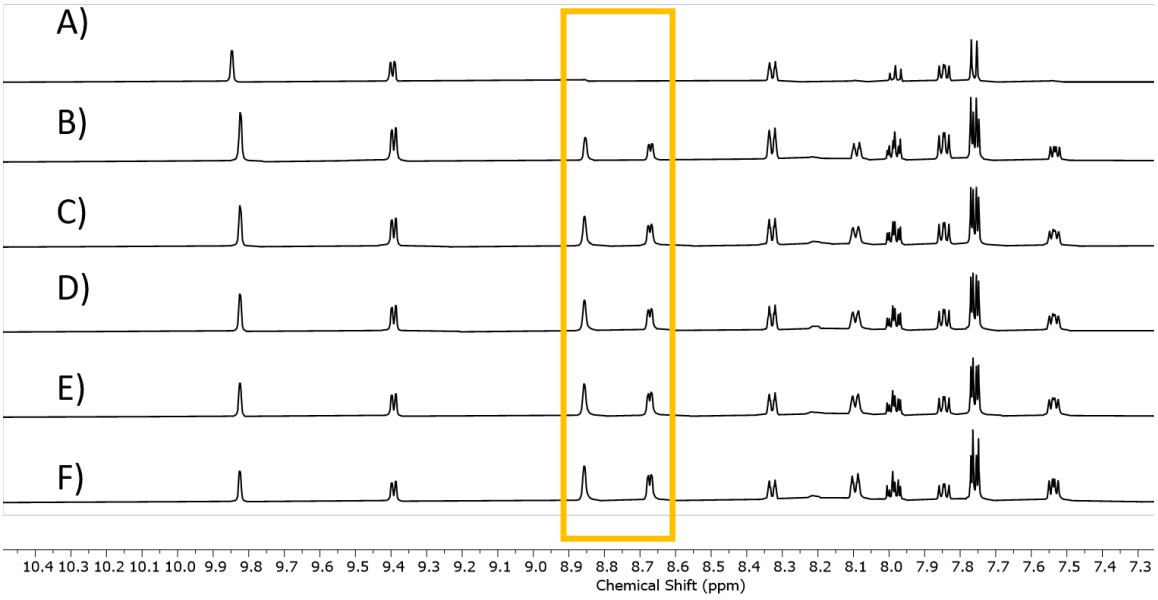


*Figure 4.25 . ^{1}H NMR (500 MHz, d_6-DMSO) Investigating effect of NaCl on **C12**. A) **C12** in B-D) **C12** after the addition of 10% saline at time points B) t = 0.25 h C) t = 1 h D) t = 2 h E) t = 6 h F) t = 12 h. Exemplary ligand signals highlighted in orange box.*

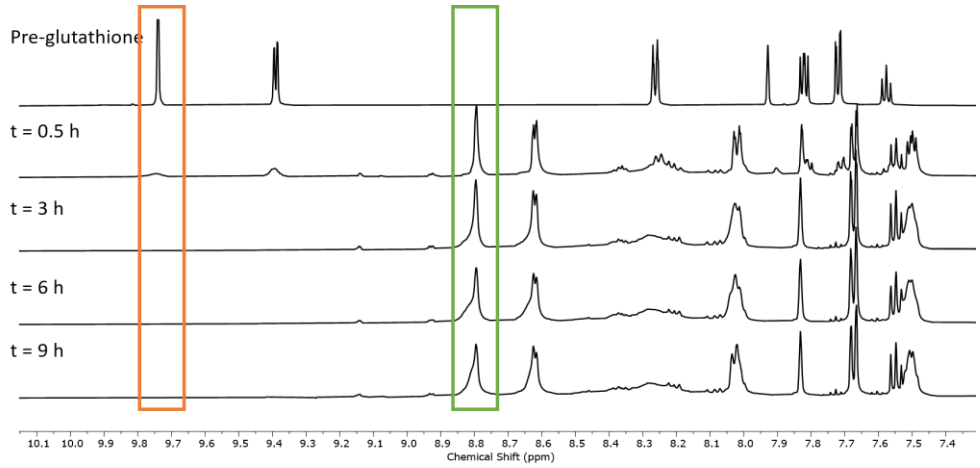
Glutathione Stability

Plasma is largely composed of ions, proteins and nutrients, including glutathione an abundant organic molecule functioning as an antioxidant and preventing damage to important cellular components. As a bioreductant, glutathione has a known mechanism for reducing and breaking down therapeutics, including chemotherapy agent Oxaliplatin mentioned above.[28]

The biological half-life of metallocages is therefore dependent on their tolerance to glutathione, which exists at a cellular concentration of between 1 mM and 15 mM depending on cell type.[30] Herein a series of ^{1}H NMR spectra were recorded at interval time points over 9 h after the addition of

glutathione to a solution of **C13** in d_6-DMSO and the generation of ligand shown in the green box (Figure 4.26A). After just 30 min only around 10% of the initial cage concentration remained, highlighted in the orange box (Figure 4.26A), indicating low stability in the presence of glutathione. However, the disassembly is not instantaneous thus implying the mechanism by which glutathione breaks down the cage works considerably slower than that of NaCl.

A)



B)

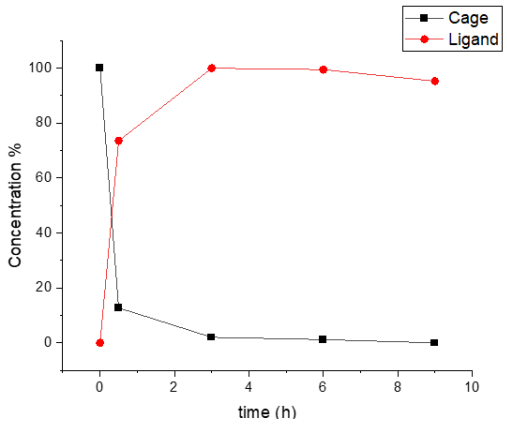


*Figure 4.26. Stability of **C13** in the presence of glutathione. A) Partial ^{1}H NMR Spectra (500 MHz, d_6-DMSO) of 1 mM **C13** and time point intervals after the addition of Glutathione (**C13** and **L13** resonances highlighted in orange and green respectively) B) Percentage consumption of cage by glutathione and percentage generation of ligand over time.*

It is concluded that neither the non-derivatised cage systems (**C12** and **C13**) nor the more soluble **C14** possess sufficient stability for *in vivo* applications. This apparent lack of stability is likely from the lability of the Pd-pyridine bond and the propensity of the Pd to interact with other species. On-going research within the group aims to synthesise a hybrid of **L16** and **L17**, whereby both the *ortho* and the *para* positions host a highly electron donating amine group, utilising both kinetic and thermodynamic stability. The aim is to synthesise less labile Pd_2L_4 cages which may possess better tolerance against bioabundant species including salt and bio-reductants.

4.2.4 Host-Guest Chemistry

In addition to remaining stable in the presence of biologically abundant species, the cages must also encapsulate guests with a strong binding affinity in order to avoid premature release of potentially toxic species during the targeted delivery of imaging agents or therapeutics.

Previous research by the Lusby group involved encapsulating $[^{99m}TcO_4^-]TcO_4^-$ within a Co^{III} tetrahedral cage, performing SPECT imaging studies following administration to a mouse.[19] Not only was the biodistribution of the $[^{99m}TcO_4^-]TcO_4^-$ altered but upon analysing the tissue post mortem the host-guest complex was found intact with $[^{99m}TcO_4^-]TcO_4^-$ still encapsulated within the cage.[19] A cold equivalent anion of $[^{99m}Tc]TcO_4^-$ was used to estimate the association constant within **C1**, whereby ReO_4^- binds with a K_a of 46,000 M^{-1} in D_2O. Given the abundance of competitive anions *in vivo*, the Pd_2L_4 systems must encapsulate the intended biomedical guest with an association constant of a similar magnitude.

Anionic Salt Guests

Often radioisotopes are generated in their anionic form. For example, ^{99m}Tc is eluted from a ^{99}Mo generator as a saline solution of $[^{99m}Tc]TcO_4^-$, or ^{18}F is formed as an aqueous solution of $[^{18}F]$ fluoride ions from the proton irradiation of $^{18}H_2O$.[31] Therefore, if cages are capable of encapsulating radioisotopes in their generated anion forms it removes the need for complicated syntheses that consume radioisotope half-life time.

To further develop the previous SPECT imaging research carried out in the Lusby group, an attempt was made to encapsulate ReO_4^- within a Pd_2L_4 system due to the analogous size and geometry with $[^{99m}Tc]TcO_4^-$.[19] Recently, Casini and co-workers developed a Pd_2L_4 cage system capable of binding $[^{99m}Tc]TcO_4^-$ which was similarly determined using ReO_4^-.[8] Casini's work showed that upon the addition of 2 eq ReO_4^- there was some shielding of cage protons. However, the change in chemical shift was minimal implying weak binding, which was not further quantified into a binding association. Furthermore, their experiments were completed in d_6-DMSO a highly polar competitive solvent and although necessary to dissolve the components it is plausibly hindering binding of the guest. Moreover, although the binding experiments can provide an indication as to whether the ReO_4^- is a guest, the solvent system used is not representative of conditions *in vivo*.

In an attempt to more closely mirror the conditions *in vivo,* a series of ^1H NMR pseudo titrations were carried out with **C13** (1 mM) in a water/MeCN solvent system (3 D_2O: 1 CD_3CN). The addition of MeCN was required to solubilise **C13**, which does not dissolve in 100% D_2O. In addition to NBu_4ReO_4, pseudo-titration titrations were also completed using KSO_3F and $NaIO_4$ (Figure 4.27). $[^{18}F]SO_3F^-$ is considered as a potential guest with applications as a PET imaging agent. Whilst initially ^{124}I was seen as an

unattractive choice for nuclear medicine due to a complex radioactive decay, in recent years advancements in PET technology has overcome this and led to increased interest in ^{124}I. Consequently, IO_4^- was investigated here as a novel guest for radiolabelling. In all experiments, ^1H NMR spectra were recorded after the addition of 0.5 Eq and 1 Eq of the selected anionic guest.

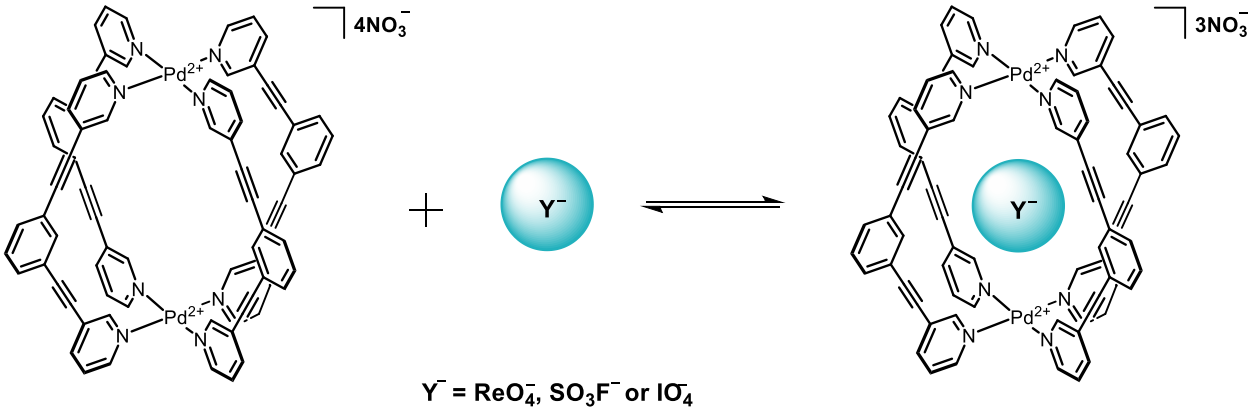


Figure 4.27. Host-guest interaction equilibrium between anionic guest and cationic cage.

Over the course of the ^1H NMR pseudo-titration of **C13** with NBu_4ReO_4 the concentration of cage in solution decreases (Figure 4.28). Following the addition of 0.5 Eq of NBu_4ReO_4 less than 20% of cage remained in solution which subsequently fell to negligible levels at 1 Eq It was proposed that the reduction in concentration could either be (a) a result of ion pairing between the cationic cage and the anionic guests, displacing the strongly hydrated NO_3^- counter anions and forming a less soluble system or (b) the cage disassembling. Additionally, there are minimal shift perturbations in the cage that does remain in solution following the addition of 0.5 Eq ReO_4^- implying that the cage is at best binding ReO_4^- only very weakly in this solvent system.

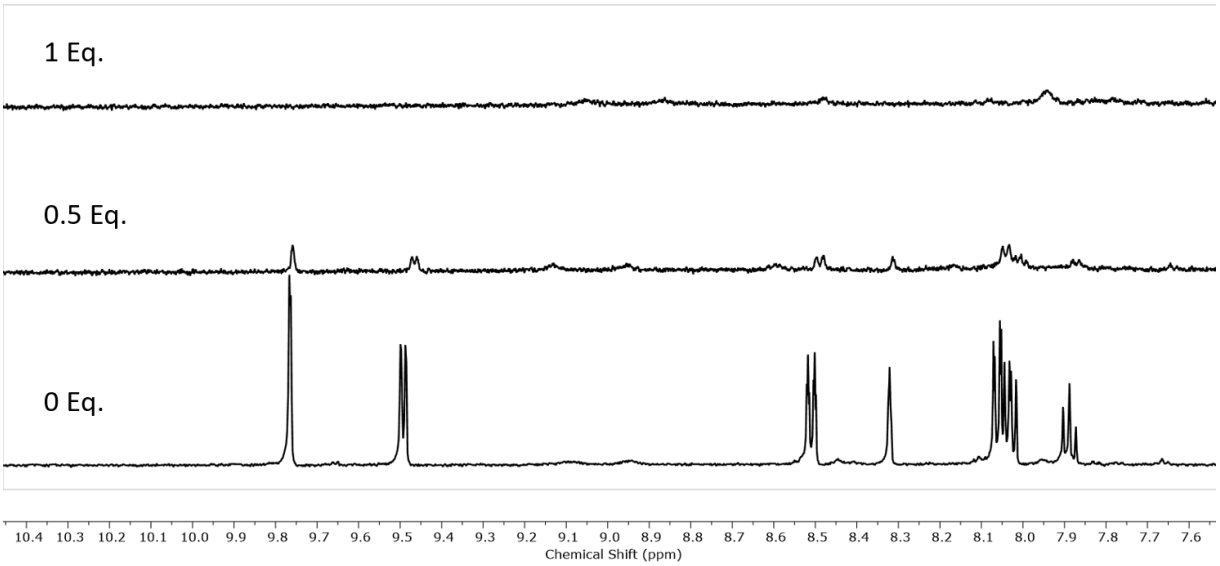


*Figure 4.28. ^{1}H NMR (500 MHz, 3 D_2O: 1 CD_3CN) Pseudo titration of NBu_4ReO_4 into 1 mM **C13***

Similarly as with the addition of NBu_4ReO_4, both pseudo-titrations using KSO_3F and $NaIO_4$ showed the concentration of **C13** in solution decreasing upon addition of the anionic guest (Figure 4.29 and Figure 4.30). Whilst the concentration of **C13** only decreases by 50% in the presence of 0.5 eq KSO_3F and $NaIO_4$ compared with the addition of 0.5 eq of NBu_4ReO_4, both show an absence of cage in solution with 1 eq of metal salt. This again appears to suggest that the aqueous solubility of the cage is highly reliant on association with the NO_3^- counter anions. The 50% reduction in cage concentration upon the addition of 0.5 eq of both KSO_3F and $NaIO_4$ could indicate that a 1:1 host guest complex has formed and this species is insoluble.

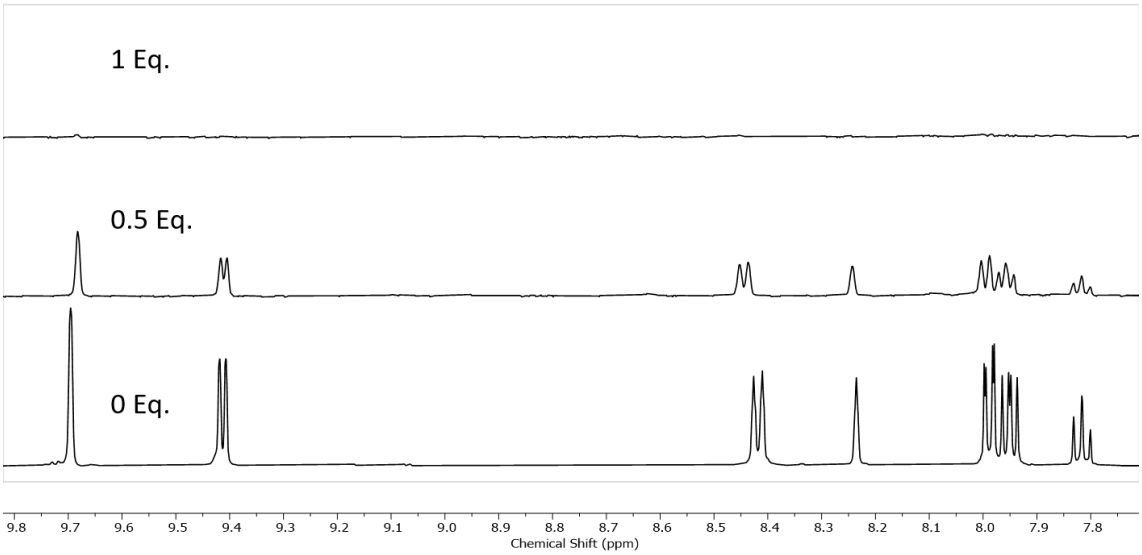


*Figure 4.29. ^{1}H NMR (500 MHz, 3 D_2O: 1 CD_3CN) Pseudo titration of KSO_3F into 1 mM **C13.***

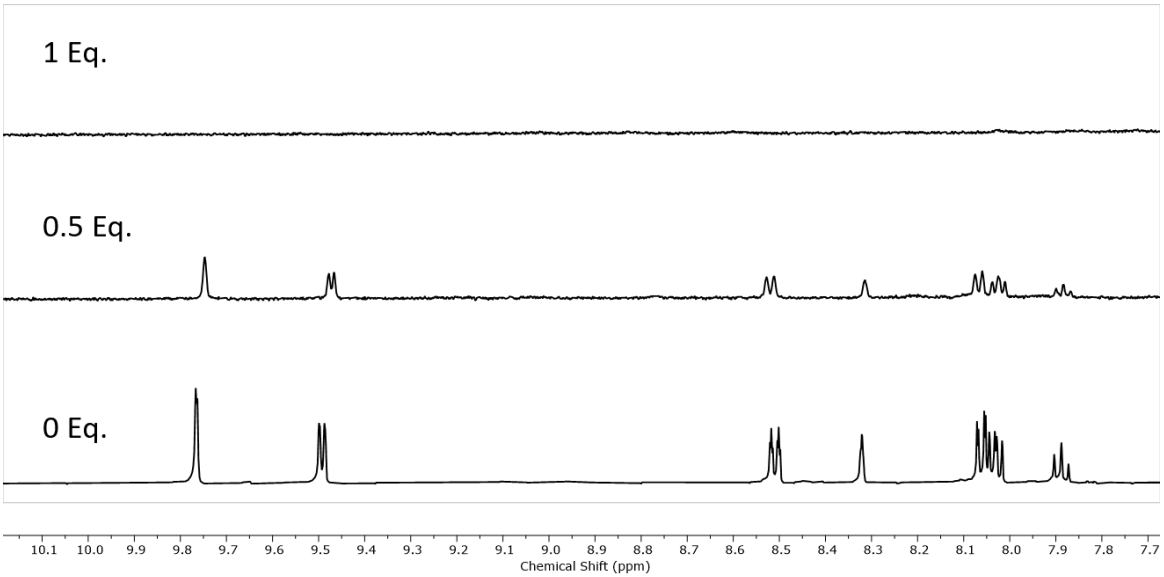


*Figure 4.30. ^{1}H NMR (500 MHz, 3 D_2O: 1 CD_3CN) Pseudo titration of $NaIO_4$ into 1 mM **C13***

In an attempt to maintain cage solubility over the course of the pseudo-titration the experiments were repeated in MeOD. It is noted that the host-guest chemistry of the systems can change when altering solvent, for example the solvaphobic effects will decrease moving from water into MeOH. KPF_6 was herein used as a cold ^{18}F source due to its preferable MeOH dissolution, moreover previous research by the Lusby group had identified it as a strong guest within Co^{III} tetrahedral cages.

Preliminary guest binding in MeOD between **C13** and the two guests, KPF_6 and NBu_4ReO_4, shows that adding the salts causes a decrease in cage concentration (Figure 4.31) analogous to the experiments carried out in water. Interestingly, upon the addition of ReO_4^-, a second species also appears, most

obviously identified by the signal at 9.55 ppm (highlighted in red). This was postulated to be the host-guest complex existing in slow exchange with the free host. The switch from fast-slow exchange often occurs in the K_a range of 10^4-10^6 M^{-1}, thus indicating that in this solvent, ReO_4^- could be a reasonably strong guest.[32]

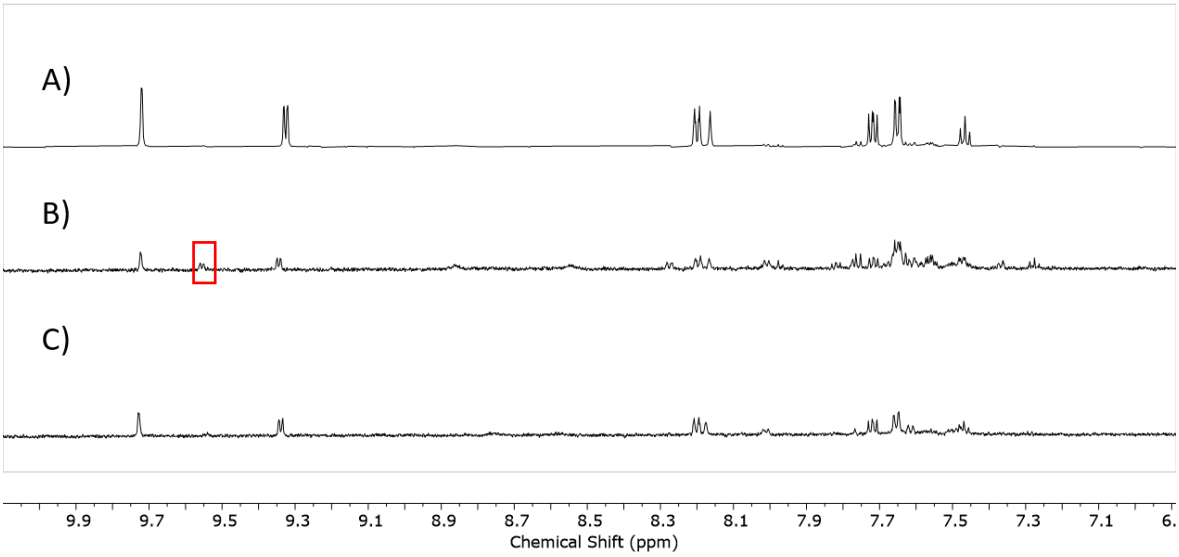


*Figure 4.31. ^{1}H NMR (500 MHz, MeOD) A) 1 mM **C13** B) (A) with 1 Eq NBu_4ReO_4 C) (A) with 1 Eq KPF_6. Red box highlights the generations of a second species.*

Upon the addition of PF_6^- into a solution of **C13**, there appears to be minimal shifts in the signals. Whilst it is possible that the anion is bound strongly within the cage and the host-guest complex precipitates out of solution, it is also possible that the PF_6^- anion is merely replacing the NO_3^- counter anion associated externally to the cage and subsequently reducing system solubility since PF_6^- is a much more hydrophobic anion and will thus interact more weakly with water. All attempts to solubilise and subsequently identify the precipitate in organic solvent were unsuccessful.

The host-guest experiments were repeated using **C12**. Upon the addition of 1 Eq NBu_4ReO_4 to a solution of **C12** in MeOD, a second set of signals appeared which could be due to the formation of a slow exchange host-guest complex (Figure 4.32B). However, it should be noted that the concentration of both free cage and presumed host-guest complex following the addition of NBu_4ReO_4 are significantly lower than the initial concentration of free cage. The ^{1}H NMR spectrum of **C12** in MeOD following the addition of 1 Eq KPF_6 indicates only a decrease in cage concentration yet with no shift perturbations of the peaks, implying that the anion is unlikely a good guest within the cage.

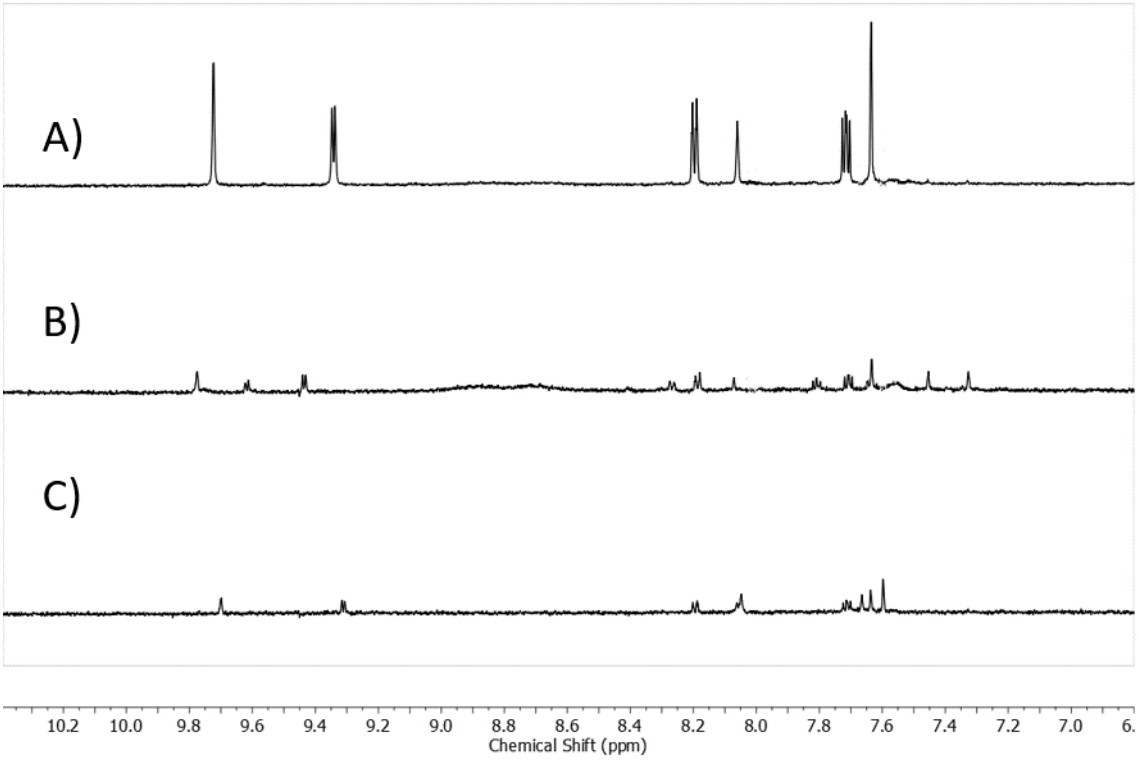


*Figure 4.32. ^{1}H NMR (500 MHz, MeOD) A) 1 mM **C12** B) (A) with 1 Eq NBu_4ReO_4 C) (A) with 1 Eq KPF_6*

These host-guest binding studies highlight one major feature; lack of inherent water solubility required the addition of the guest anion. The propensity of the cages to crash out of solution upon addition of anionic guests implies that the solubility of the cage system is highly reliant on the nitrate counter anions to such an extent that replacing just one strongly hydrating NO_3^- counter anions with a more hydrophobic anion (e.g., ReO_4^-, PF_6^- etc) leads to a drastically reduced solubility. It is concluded that since a second set of signals appears upon the addition of NBu_4ReO_4 to both **C12** and **C13**, ReO_4^- is likely a guest within both cages, which stands to validate the previous research by Casini and co-workers on similar systems.

Alternative Neutral Guests

The cationic nature of these cages allows them to attract and encapsulate complimentary anions through electrostatic interactions. However, the aromatic backbone of the ligands also enables the resultant coordination cages to encapsulate neutral guests through the hydrophobic effect. Binding of neutral guests is both enthalpically and entropically favourable in water based solvent systems. Hydrophobic binding has been widely explored when researching water soluble coordination cages, notably, by Fujita and co-workers who have used a series of Pd-octahedra capable to bind series neutral, aploar guests from adamantine derivatives to tetrabenzylsilane.[2] However, it should be noted that the structure of these octahedra are quite different to the Pd_2L_4 cages described here, as they

possess large aromatic surfaces which surround the cavity and thus facilitate apolar binding. In contrast, the perpendicular alignment of ligands in the Pd_2L_4 cages disfavours the formation of interactions between the guests and the cages' aromatic surfaces. The Lusby group has also repeatedly shown that the inward facing *ortho*-pyridyl protons can act as moderate H-bond donors, and thus interact with guests via H-bond interactions (Figure 4.33).[12,15,33]

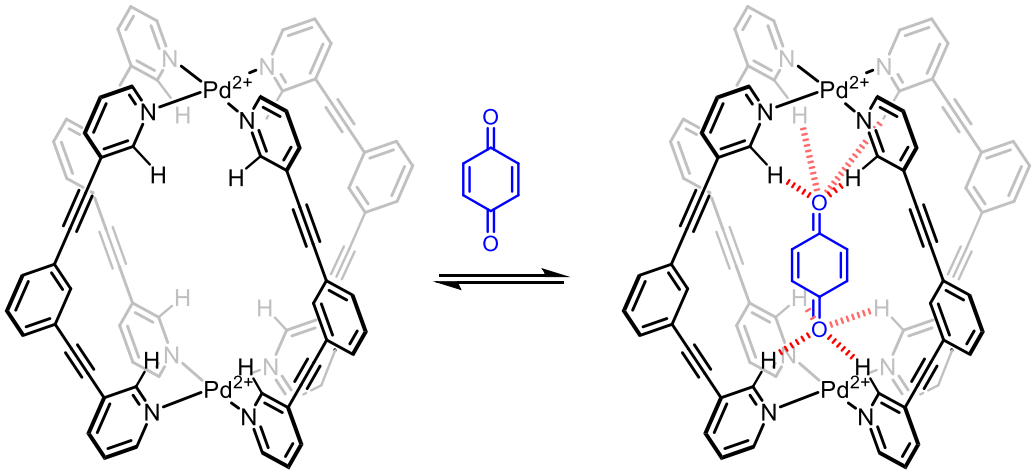


*Figure 4.33. Encapsulation of benzoquinone within **C12**·4BArF, facilitated by hydrogen bonds illustrated by dashed lines.*

The Lusby group have found that *p*-benzoquinones are ideal guests for the non-derivatised **C12** and **C13**, as they can simultaneously hydrogen bond to both the "top" and "bottom" of the cage (Figure 33). It has been found that the highest affinity is observed in DCM using BArF counter anions, that are too large to fit in the cavity. Under these conditions, unsubstituted benzoquinone has a binding association constant of 8000 M^{-1} in **C13** and 1000 M^{-1} in **C12**.[34] Halo-substituted benzoquinones are capable of binding within the Pd_2L_4 cavities, but give significantly weaker binding association constants. The host-guest interaction is reliant on the electron rich carbonyl, hydrogen bonding with the electron poor internal *ortho*-protons of the pyridine rings, whereby by substituting the *p*-benzoquinone protons for electron withdrawing halogens, reduces the strength of the carbonyl to proton interaction. Tetrafluorobenzoquinone (fluoranil) has a K_a of just 120 M^{-1} in **C13** (BArF counter anions) when measured in d_2-DCM and a K_a of just 80 M^{-1} in **C12** (BArF counter anions).[34]

Notably, the group have also discovered a hydration reaction whereby encapsulated fluoranil hydrolysed and the subsequent product enolises to give a strongly binding anion (Figure 4.34). This guest in particular was identified as a potential guest that might be of interest for ^{18}F PET imaging.

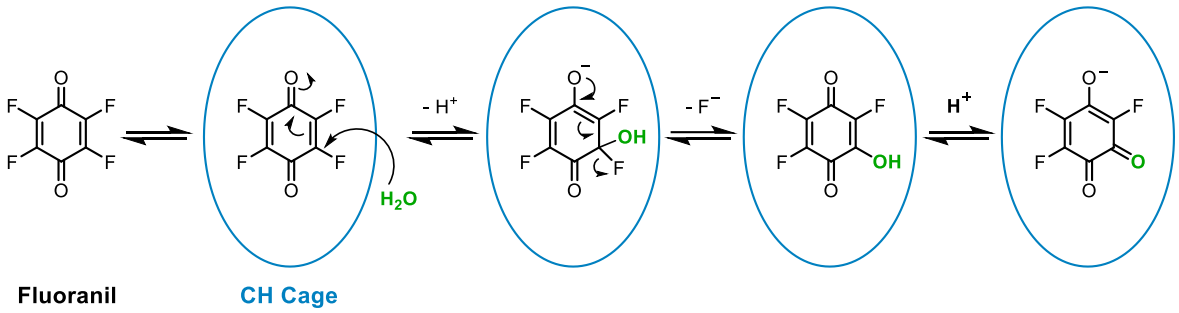


Figure 4.34. Fluoranil hydrolysis proposed mechanism adapted from previous unpublished research by the Lusby group.

In order to utilise the guest binding of benzoquinones for biomedical applications, it is necessary to investigate in a more polar solvent that is more representative of conditions *in vivo*. As such, experiments were attempted in MeOD and d_6-DMSO whereby the hydrophobic effect will be lower in comparison to water but are hypothesised will still aid binding. Upon addition of 2 Eq Fluorinil into a solution of **C13** in d_6-DMSO, no shifts in the cage signals (Figure 4.35) were observed indicating minimal interactions. Whilst the strength of the interaction could be further probed by adding a larger excess of fluorinil it is considered that the interaction is not sufficient for delivery purposes under these conditions.

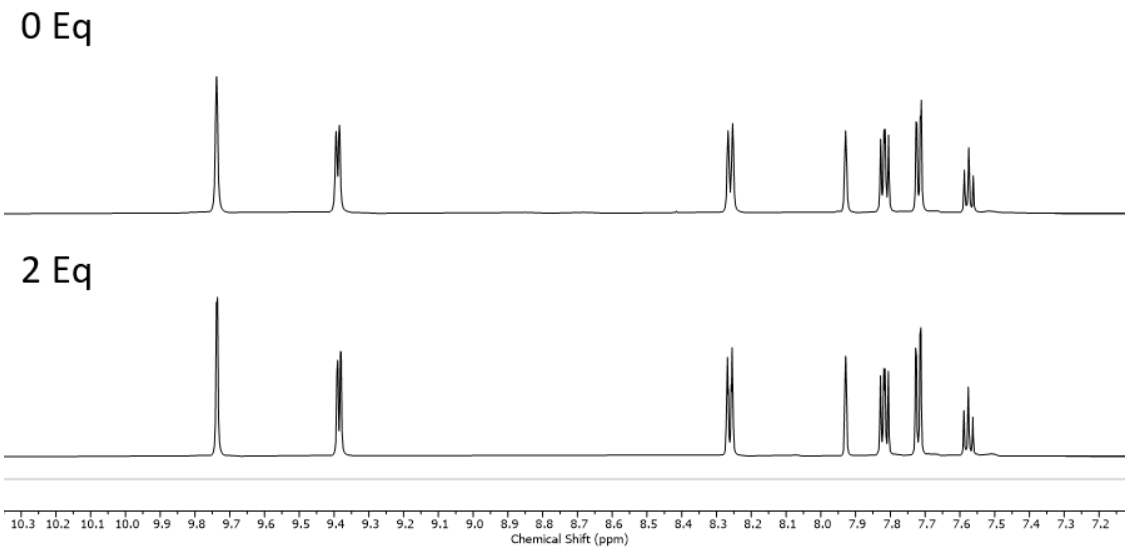


*Figure 4.35. ^{1}H NMR (500 MHz, d$_6$-DMSO) Preliminary fluoranil guest binding study in **C13**. Whereby X Eq indicates the number of equivalents of fluorinil added to **C13**.*

To improve dissolution a solution of Fluorinil was titrated into a solution of **C14** in MeOD to investigate guest interactions in an alternative solvent. Minimal shift perturbation was detected with no apparent saturation point, even following the addition of 3.5 Eq of fluorinil (Figure 4.36). It is worth noting that although binding appeared to be minimal the neutral nature of the guest led to no decrease in the

cage concentration over the course of the titration. It was therefore concluded that the interactions are too weak for drug delivery applications.

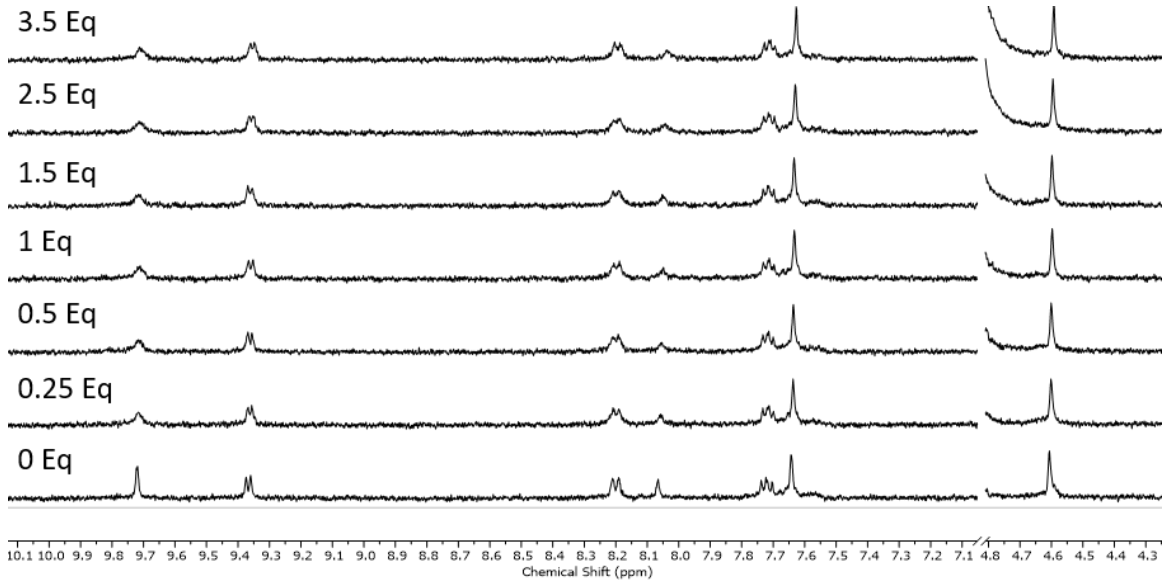


*Figure 4.36. ^{1}H NMR (400 MHz, MeOD) Titration of fluoranil into 1 mM **C14**, whereby X Eq indicates the number of equivalents of fluoranil added to the cage solution.*

Interestingly a pseudo titration of fluoranil with **C12** in MeOD showed the formation of a second set of signals, potentially from the generation of a host-guest complex that exists in slow exchange with the free host (Figure 4.37). Previous research by the Lusby group indicated that fluoranil had a weak guest encapsulation in **C12** (K_A = 80 M^{-1}) thus it is surprising to observe a species in slow exchange in this case. It is possible that the fluoranil is reacting with residual water and converting into its more strongly binding enolate form (such as described in Figure 4.34) within **C12** accounting for the generation of the second set of signals over time. Future research would include synthesising the enolate product and directly measuring the host-guest chemistry in these systems here under conditions that are relevant for use *in vivo*.

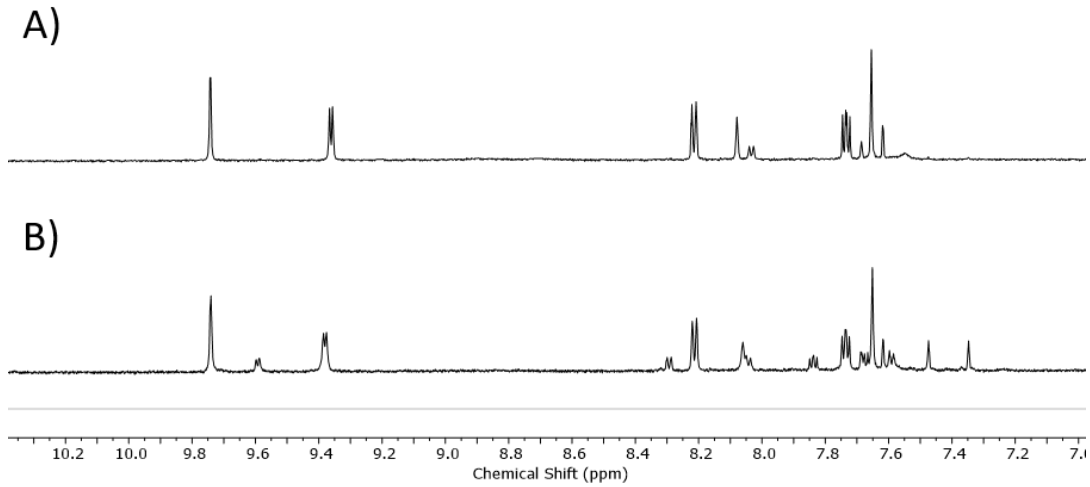


*Figure 4.37. ^{1}H NMR (400 MHz, MeOD) Preliminary fluorinil guest binding study in **C12**. A) free **C12** B) **C12** with 2 Eq fluoranil*

4.3 Conclusions and Future Work

The aims of the research included synthesising a library of ligands and subsequent Pd_2L_4 cages to probe their potential applications and limitations within biomedicine, and more specifically as delivery vessels. A library of five novel ligands were synthesised with various substituents aiming to alter the water solubility and stability of the resultant cages.

A series of cage assembly reactions were then completed and indicated that by altering the binding strength of the ligands the cage assembly process became hindered. Whilst the assembly of the non-derivatised ligands (**L12** and **L13**) and mono-hydroxyl substituted **L14** progressed with high yields (79-88%), the assembly of the ligands with increased sigma bond donation (**L16** and **L17**) lead to unrefined mixtures. Presumably as the increased strength of the coordination bond does not permit the error checking usually required to access the thermodynamically preferred Pd_2L_4 assembly.

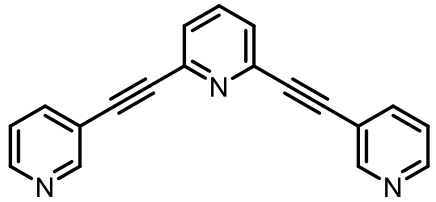
The host-guest chemistry and stability of the non-derivatised complexes was probed, investigating the effect of both salt and bioreductant glutathione on cage stability along with the suitability of various biomedically relevant species as guest molecules. The cage systems proved unstable in the presence of both NaCl, in which Pd^{II}-Cl coordination led to instantaneous cage disassembly, and glutathione which caused complete cage degradation within just 30 min. The low water solubility of the cage systems and their dependency on NO_3^- counter anions for solubility led to precipitation of the cage upon titration with anionic guest molecules. However, preliminary results indicate that ReO_4^- could be a strong guest within both **C12** and **C13**, with the host:guest complex existing as a slow-exchange on the NMR time scale.

It is concluded that without drastically increasing the ligand binding strength of these tri-aryl Pd_2L_4 cage systems the Pd-ligand coordination bond remains too labile to ensure stability within biological conditions. Moreover, high water solubility of the cage systems independent of the associated counter anions is imperative for strong anion guest encapsulation without cage precipitation. By improving both the stability and water solubility of these cage systems they could show potential as biomedical delivery vessels but the systems discussed here proved unsuitable. Future research would utilise alternatives to the direct assembly routes attempted here in order to access the more stable cages. Including trans-ligation, similar to the optimised method in Chapters 2 and 3, along with using a guest template to alter cage geometry.

4.4 Experimental

4.4.1 Synthesis of Ligands

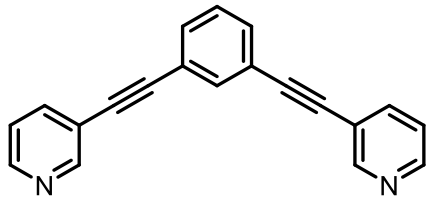
Non-derivatised Ligands



L12

Synthesised according to literature procedure. [33]

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.86 (d, *J* = 1.9 Hz, 2H), 8.63 (dd, *J* = 4.9, 1.9 Hz, 2H), 7.91 (dt, *J* = 7.9, 1.9 Hz, 2H), 7.76 (d, *J* = 7.8 Hz, 1H), 7.56 (d, *J* = 7.8 Hz, 2H), 7.34 (dd, *J* = 7.9, 4.9 Hz, 2H).

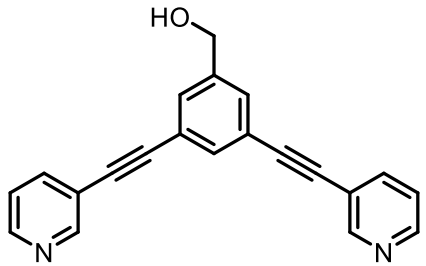


L13

Synthesised according to literature procedure.[22]

^{1}H NMR (500 MHz, CD_3CN+D_2O) δ 8.67 (dd, *J* = 2.1, 0.9 Hz, 2H), 8.50 (dd, *J* = 5.0, 1.7 Hz, 2H), 7.95 (dt, *J* = 7.9, 1.9 Hz, 2H), 7.74 (d, *J* = 1.3 Hz, 1H), 7.60 (dd, *J* = 7.7, 1.6 Hz, 2H), 7.48 – 7.44 (m, 3H).

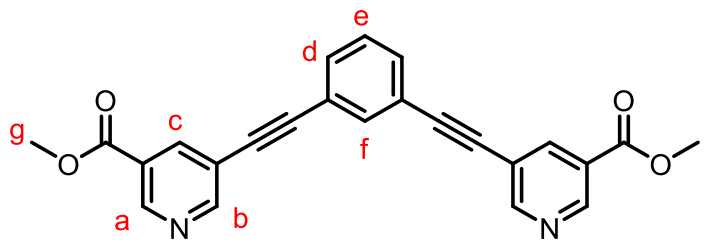
Water Soluble Ligands



L14

Synthesised according to literature procedure.[35]

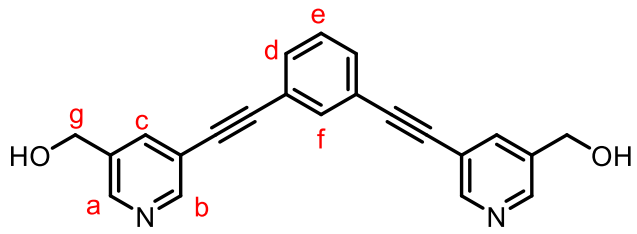
^{1}H NMR (500 MHz, $CDCl_3$) δ 8.83 (bs, 2H), 8.63 (bs, 2H), 7.84 (d, *J* = 7.9 Hz, 2H), 7.68 (t, *J* = 1.6 Hz, 1H), 7.58 (d, *J* = 1.6 Hz, 2H), 7.34 (bs, 2H), 4.76 (d, *J* = 4.3 Hz, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 152.25, 148.75, 141.94, 138.51, 133.71, 130.14, 123.75, 123.24, 91.56, 86.77, 64.21, 1.02.



10

Synthesised by Sam Oldknow according to literature.[17]

^{1}H NMR (500 MHz, $CDCl_3$) δ 9.20 (s, 2H, H_b), 8.95 (s, 2H, H_a), 8.45 (s, 2H, H_c), 7.79 (d, *J* = 1.6 Hz, 1H, H_f), 7.59 (dd, *J* = 7.8, 1.6 Hz, 2H, H_d), 7.44 (t, *J* = 7.8 Hz, 1H, H_e), 4.01 (s, 6H, H_g). ^{13}C NMR (126 MHz, $CDCl_3$) δ 165.14, 155.47, 149.64, 139.42, 134.90, 132.19, 128.85, 122.79, 92.51, 85.81, 77.22, 52.65, 1.02.



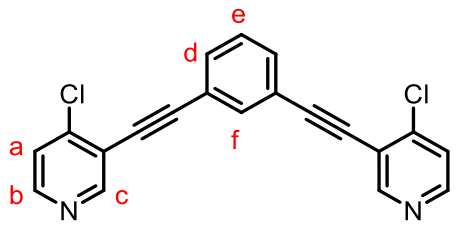
L15

Under an inert atmosphere, a solution of **10** (60.0 mg, 0.150 mmol) in dry THF was cooled to -78 °C. $LiAlH_4$ (0.360 mmol, 1 M, 0.360 mL) was added and the reaction was stirred for 2 h. The reaction was then warmed to rt and quenched with ice water (3 mL), and the solution was filtered. The precipitate was washed with THF (20 mL) and the combined organic solutions were dried over anhydrous $MgSO_4$ and concentrated *in vacuo.* The product was purifiied by column chromatography on Et_3N deactivated silica (1-10% MeOH in 5% Et_3N: DCM) to yield **L15** (30 mg, 0.0880 mmol, 58%) as a beige solid.

^{1}H NMR (601 MHz, $CDCl_3$) δ 8.70 (d, *J* = 2.0 Hz, 2H, H_b), 8.57 (d, *J* = 2.0 Hz, 2H, H_c), 7.90 (t, *J* = 2.0 Hz, 2H, H_a), 7.75 (td, *J* = 1.6 Hz, 1H, H_f), 7.55 (dd, *J* = 7.9, 1.6 Hz, 2H, H_d), 7.40 (t, *J* = 7.9 Hz, 1H, H_e), 4.08 (s, 4H, H_g). ^{13}C NMR (126 MHz, $CDCl_3$) δ 155.48, 149.65, 139.44, 134.90, 132.19, 125.70, 122.79, 120.17, 92.51, 85.79, 77.22, 52.65.

Species not located in ESI mass spectrum.

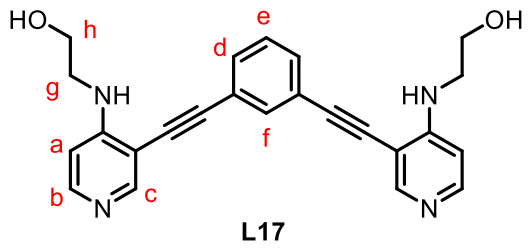
Ligands with Increased Binding Strength



L20

Under an N_2 atmosphere, 3-bromo-4-chloropyridine (189 mg, 0.990 mmol) and 1,3-diethynylbenzene (50.0 mg, 0.400 mmol) were suspended in dry degassed Et_3N (20.0 mL). $Pd(PPh_3)_2Cl_2$ (24.0 mg, 0.0360 mmol) and CuI (11.0 mg, 0.0590 mmol) were added and the reaction was heated to 80 °C and stirred for 18 h. The reaction mixture was filtered and the retentate was washed with chloroform (10 mL) leaving a crude yellow solid which was purified by column chromatography on Et_3N deactivated silica (1% Et_3N in DCM) yielding a white product (120 mg, 0.336 mmol, 86%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 8.78 (s, 2H, H_c), 8.48 (d, *J* = 5.4 Hz, 2H), 7.83 (t, *J* = 1.7 Hz, 1H, H_f), 7.63 (dd, *J* = 7.8, 1.7 Hz, 2H, H_d), 7.44 (t, *J* = 7.8 Hz, 1H, H_e), 7.42 (d, *J* = 5.4 Hz, 2H, H_a). ^{13}C NMR (126 MHz, $CDCl_3$) δ 153.41, 149.22, 145.24, 134.91, 132.35, 128.80, 124.10, 122.82, 120.52, 96.40, 83.75, 77.22. High resolution ESI MS: Calculated mass of $C_{20}H_{11}Cl_2N_2$ = 349.0294, Found mass = 349.0307.

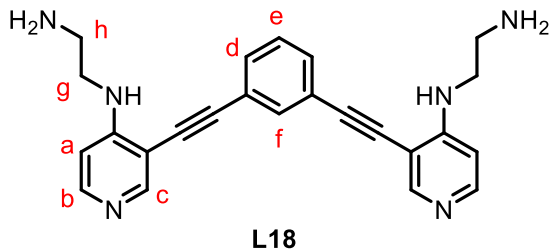


L17

L6 (50.0mg, 0.140 mmol), ethanolamine (180 mg, 2.90 mmol), Et_3N (51.0 mg, 0.572 mmol) and ethanol (2.00 mL) were stirred together at 160 °C in a sealed microwave vial for 18 h. The reaction was cooled to rt and the solvent was removed *in vacuo.* The product was precipitated with water (4.00 mL), yielding a beige solid (46.0 mg, 0.115 mmol, 79%).

^{1}H NMR (601 MHz, $CDCl_3$) δ 8.36 (s, 2H, H_c), 8.17 (dd, *J* = 5.9 Hz, 2H, H_b), 7.72 (t, *J* = 1.7 Hz, 1H, H_f), 7.49 (dd, *J* = 7.7, 1.7 Hz, 2H, H_d), 7.37 (t, *J* = 7.87 Hz, 1H, H_e), 6.47 (d, *J* = 5.9 Hz, 2H, H_a), 3.95 (t, *J* = 5.7 Hz, 4H, H_h), 3.43 (q, *J* = 5.7 Hz, 4H, H_g). ^{13}C NMR (126 MHz, d_6-DMSO) δ 153.91, 152.18, 149.76, 134.34,

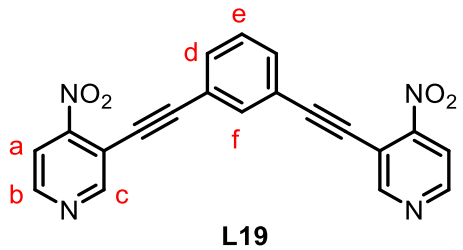
131.89, 129.52, 123.40, 105.11, 103.80, 96.07, 84.88, 57.87, 48.32. High resolution ESI MS: Calculated mass of $C_{24}H_{22}N_4O_2$ = 399.18155 Found mass = 399.1813.



L18

L6 (40.0 mg, 0.114 mmol), ethylenediamine (140 mg, 2.29 mmol), Et_3N (41.0 mg, 0.458 mmol) and ethanol (2.00 mL) were stirred together at 160 °C in a sealed microwave vial for 18 h. The reaction was cooled to rt and the solvent was removed *in vacuo.* The product was precipitated with water (4.00 mL), yielding a beige solid (38.0 mg, 0.0960 mmol, 84%).

^{1}H NMR (601 MHz, $CDCl_3$) δ 8.42 (s, 2H, H_c), 8.23 (d, *J* = 5.8 Hz, 2H, H_b), 7.75 (t, *J* = 1.7 Hz, 1H, H_f), 7.54 (dd, *J* = 7.8, 1.7 Hz, 2H, H_d), 7.40 (d, *J* = 7.8 Hz, 1H, H_e), 6.52 (d, *J* = 5.8 Hz, 2H, H_a), 3.34 (q, *J* = 11.8, 5.9 Hz, 4H, H_g), 3.07 (t, *J* = 5.9 Hz, 4H, H_h). ^{13}C NMR (126 MHz, $CDCl_3$) δ 153.37, 152.20, 149.82, 134.27, 131.40, 128.70, 124.09, 123.34, 104.21, 96.48, 83.75, 44.63, 40.63. High resolution ESI MS: Calculated mass of $C_{24}H_{25}N_6$ = 397.21352 Found mass = 397.2119.

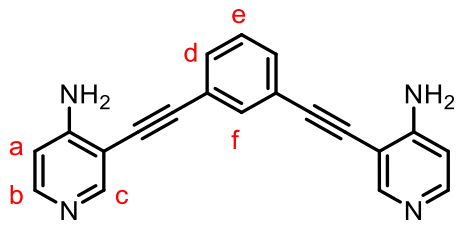


L19

A solution of 3-bromo-4-nitropyridine (200 mg, 0.990 mmol) and 1,3-diethynylbenzene (50.0 mg, 0.396 mmol) in dry degassed Et_3N (20 mL) was stirred for 20 min and followed by the addition of $Pd(PPh_3)_4Cl_2$ (24.0 mg, 0.0360 mmol) and CuI (11.0 mg, 0.0590 mmol). The reaction was heated to 80 °C under N_2 for 18 h and then filtered and the retentate washed with chloroform (10 mL). The resulting crude yellow solid was purified by column chromatography on Et_3N deactivated silica (1% Et_3N in DCM) yielding a bright yellow solid (138 mg, 0.372 mmol, 94%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 9.08 (s, 2H, H_c), 8.82 (d, *J* = 5.3 Hz, 2H, H_b), 7.94 (d, *J* = 5.3 Hz, 2H, , H_a), 7.95 (t, *J* = 1.6 Hz, 1H, H_f), 7.70 (dd, *J* = 7.8, 1.6 Hz, 2H, H_d), 7.49 (t, *J* = 7.8 Hz, 1H, H_e). ^{13}C NMR (126

MHz, $CDCl_3$) δ 155.82, 154.28, 150.08, 135.47, 133.27, 129.01, 122.42, 116.92, 113.50, 98.66, 82.61. High resolution ESI MS: Calculated mass of $C_{20}H_{11}N_4O_4$ = 371.07748 Found mass = 371.0764.



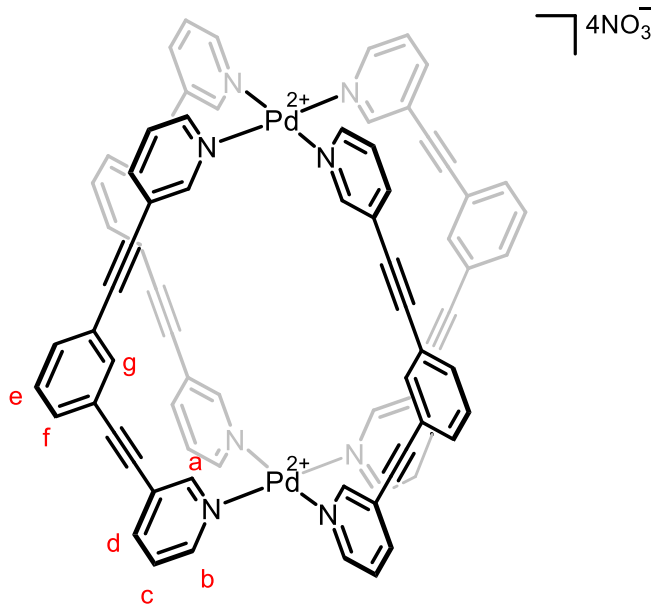
L16

Fe powder (109 mg, 1.94 mg) was added to a stirred suspension of **L19** (60 mg, 0.162 mmol) in glacial acetic acid (8 mL) under an inert atmosphere. The reaction mixture was then heated to 60 °C for 5 h before cooling to room temperature and basifying to pH 13 using NaOH (3M, 35 mL). The basic aqueous mixture was extracted with chloroform (3 x 75 mL), dried over $MgSO_4$ and solvents removed *in vacuo* to afford **L16** (28 mg, 0.0903 mmol, 56%) as a beige solid.

^{1}H NMR (500 MHz, $CDCl_3$) δ 8.47 (s, 2H, H_c), 8.20 (d, *J* = 5.7 Hz, 2H, H_b), 7.73 (t, *J* = 1.6 Hz, 1H, H_f), 7.54 (dd, *J* = 7.2, 1.6 Hz, 2H, H_d), 7.40 (d, *J* = 7.2 Hz, 1H, H_e), 6.60 (d, *J* = 5.7 Hz, 2H, H_a). ^{13}C NMR (126 MHz, d_6-DMSO) δ 155.03, 152.55, 149.18, 134.46, 131.58, 129.34, 123.56, 118.47, 108.73, 95.48, 85.10. High resolution ESI MS: Calculated mass of $C_{20}H_{15}N_4$ = 311.12912 Found mass = 311.1302.

4.4.2 Cage Assembly

Direct Assembly with Pd Nitrate Salts

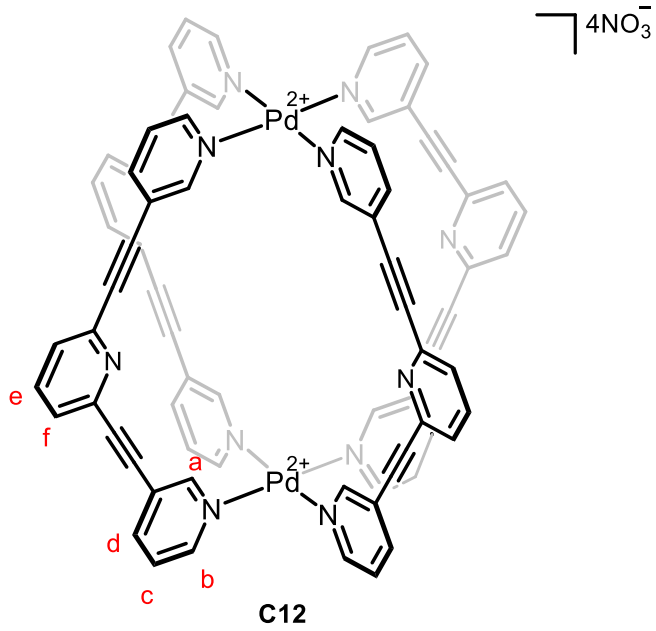


C13

L13 (50.0 mg, 0.178 mmol) was added to a solution of $Pd(NO_3)_2$ (24.0 mg, 0.0891 mmol) in MeCN (2.5 mL) and DI H_2O (2.5 mL) in a vial equipped with magnetic stirrer. The reaction mixture was sonicated for 20 min, before sealing and allow to react at room temperature for 18 h. The reaction mixture was then concentrated *in vacuo* and freeze dried to remove remaining DI H_2O, the resulting solid was then washed with MeCN which yielded a beige powder (56.0 mg, 0.0354 mmol, 79%).

^{1}H NMR (500 MHz, d_6-DMSO) δ 9.74 (d, *J* = 2.0 Hz, 8H, H_a), 9.39 (dd, *J* = 5.9, 2.0 Hz, 8H, H_b), 8.26 (dt, *J* = 8.0, 2.0 Hz, 8H, H_d), 7.93 (t, *J* = 1.7 Hz, 4H, H_g), 7.82 (dd, *J* = 8.0, 5.9 Hz, 8H, H_c), 7.72 (dd, *J* = 7.8, 1.7 Hz, 8H, H_f), 7.58 (t, *J* = 7.8 Hz, 4H, H_e). ^{13}C NMR (126 MHz, d_6-DMSO) δ 153.35, 151.03, 143.28, 134.48, 133.81, 130.50, 127.79, 122.76, 122.21, 94.20, 85.65. ^{1}H DOSY NMR (600 MHz, d_6-DMSO): D 22 = 1.05 x 10^{-10} m^2 s^{-1}; calculated hydrodynamic radius = 10.5 Å. ESI-MS (m/z): 729.1 (2+) 465.4 (3+) 333.5 (4+).

C12

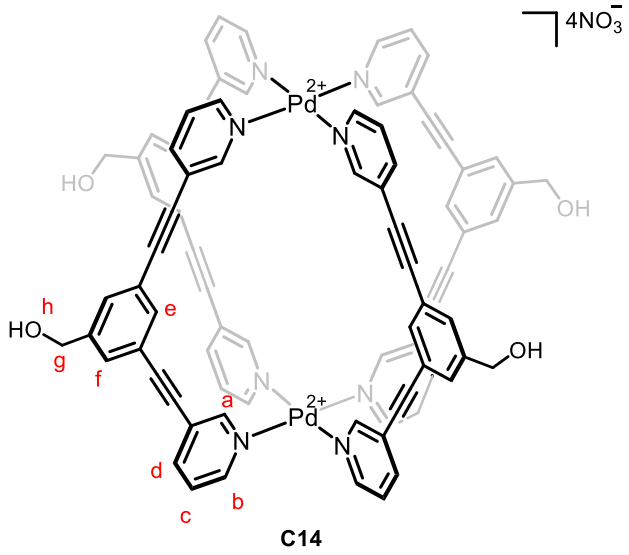


C12

L12 (21.0 mg, 0.075 mmol) was added to a solution of $Pd(NO_3)_2$ (10.0 mg, 0.0375 mmol) in MeCN (0.6 mL) and DI H_2O (1.8 mL) in a microwave vial (2-5 mL) equipped with magnetic stirrer. The reaction mixture was sonicated for 20 min, before sealing and allow to react at room temperature for 18 h. The reaction mixture was then concentrated *in vacuo* and freeze dried to remove remaining H_2O, the resulting solid was then washed with MeCN which yielded a beige powder (25.0 mg, 0.0158 mmol, 84%).

^{1}H NMR (500 MHz, d_6-DMSO) δ 9.85 (d, *J* = 1.6 Hz, 8H, H_a), 9.40 (dd, *J* = 5.9, 1.5 Hz, 8H, H_b), 8.33 (dt, *J* = 8.0, 1.5 Hz, 8H, H_d), 7.98 (t, *J* = 7.7 Hz, 4H, H_e), 7.85 (dd, *J* = 8.0, 5.9 Hz, 8H, H_c), 7.76 (d, *J* = 7.7 Hz, 8H,

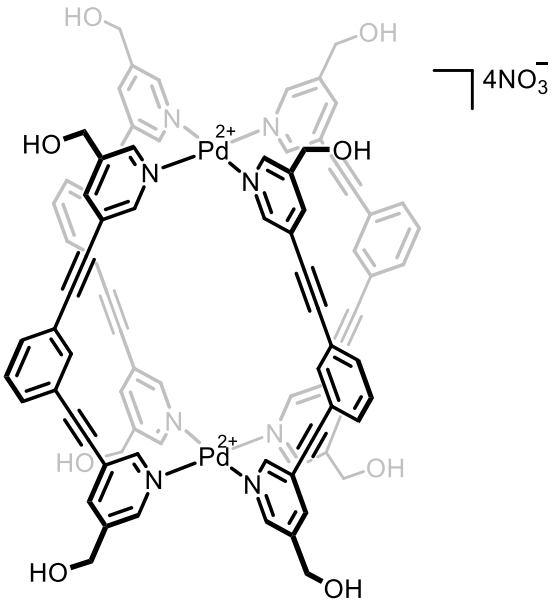
H_f). ^{13}C NMR (126 MHz, d_6-DMSO) δ 153.39, 150.95, 145.23, 143.24, 132.73, 131.32, 127.77, 122.81, 121.98, 94.36. ^{1}H DOSY NMR (600 MHz, d_6-DMSO): D 22 = 1.05 x 10^{-10} m^2 s^{-1} ; calculated hydrodynamic radius = 10.5 Å. ESI-MS (m/z): 731.1 (2+) 466.7 (3+) 334.5 (4+).



C14

L14 (23 mg, 0.075 mmol) was added to a solution of $Pd(NO_3)_2$ (10 mg, 0.0375 mmol) in MeCN (0.6 mL) and H_2O (1.8 mL) in a microwave vial (2-5 mL) equipped with magnetic stirrer. The reaction mixture was sonicated for 20 min, before sealing and allow to react at rt for 18 h. The reaction mixture was then concentrated *in vacuo* and freeze dried to remove remaining H_2O, the resulting solid was then washed with MeCN which yielded a beige powder (28 mg, 0.0164 mmol, 88%).

^{1}H NMR (601 MHz, d_6-DMSO) δ 9.67 (d, *J* = 1.6 Hz, 8H, H_a), 9.31 (dd, *J* = 6.0, 1.4 Hz, 8H, H_b), 8.36 (dt, *J* = 8.0, 1.6 Hz, 8H, H_d), 8.08 (t, *J* = 7.8 Hz, 4H, H_e), 7.88 (d, *J* = 7.8 Hz, 8H, H_f), 7.81 (dd, *J* = 8.0, 7.8 Hz, 8H, H_c), 5.45 (t, *J* = 5.6 Hz, 4H, H_h), 4.54 (d, *J* = 5.7 Hz, 8H, H_g). ^{13}C NMR (126 MHz, d_6-DMSO) δ 154.06, 152.42, 151.55, 143.84, 142.44, 139.67, 138.50, 129.07, 127.82, 122.08, 93.85, 83.81. ^{1}H DOSY NMR (600 MHz, d_6-DMSO): D 22 = 1.01 x 10^{-10} m^2 s^{-1}; calculated hydrodynamic radius = 10.5 Å. ESI-MS (m/z): 789.1 (2+) 505.40 (3+) 363.13 (4+).

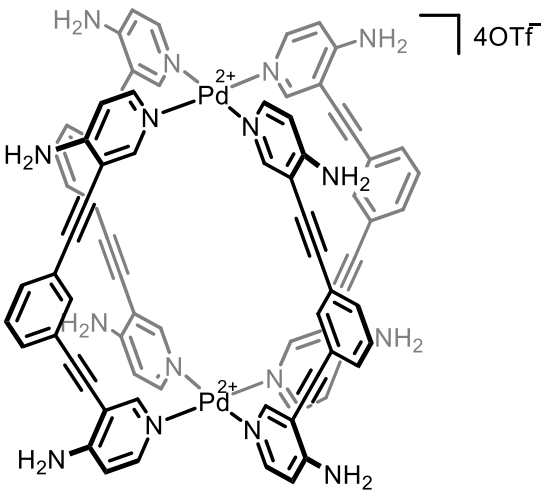


C15

A solution of $Pd(NO_3)_2$ (4 mg, 0.0146 mmol) and **L15** (10 mg, 0.029 mmol) in d_3-MeCN (0.6 mL) and D_2O (1.8 mL) was sonicated for 20 min, sealed and stirred for 18 h. NMR showed mixture of species.

Direct Assembly with Alternative Pd Salts

C16·4OTf

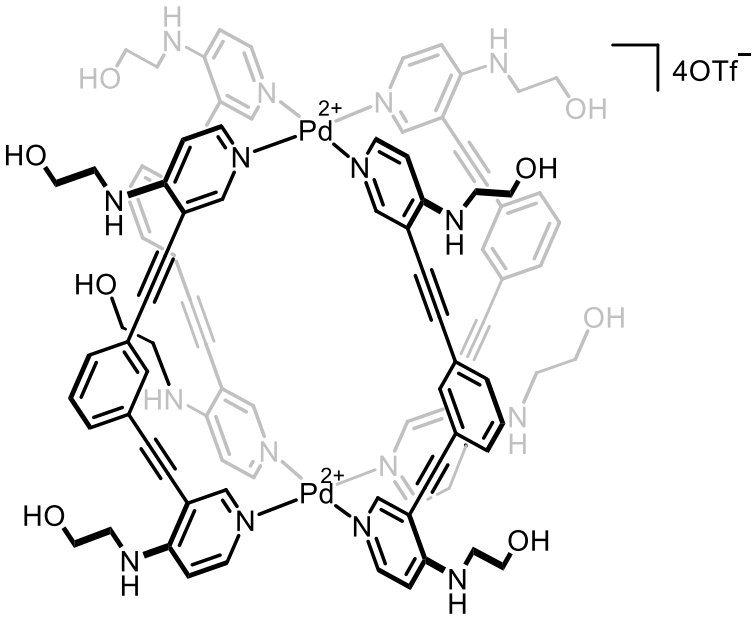


L16 (23 mg, 0.0750 mmol) and $Pd[MeCN]_2(OTf)_2$ (10 mg, 0.0180 mmol, 133 mg/mL, 75 µL) were sonicated together in d_3- MeCN (0.6 mL) for 2 h to give **C16** in quantitative yields (0.0375 mmol).

^{1}H DOSY NMR (600 MHz, d_6-DMSO): D 22 = 1.01 x 10^{-10} m^2 s^{-1}; calculated hydrodynamic radius = 10.5 Å. ESI-MS (m/z): 534.5 (2+) 876.2 (3+) 363.13 (4+).

^{1}H and ^{13}C NMR showed multitude of peaks preventing assignment as discussed in "Chapter 2: Cage Assembly: Direct Assembly using Alternative Pd Salts"

C17·4OTf



L17 (23 mg, 0.075 mmol) and Pd[MeCN]$_2$(OTf)$_2$ (10 mg, 0.018 mmol, 133 mg/mL, 75 µL) into were sonicated together in d_3-MeCN (0.6 mL) for 2 h to give **C17** in quantitative yields (0.0375 mmol).

^{1}H DOSY NMR (600 MHz, d_6-DMSO): D 22 = 1.01 x 10^{-10} m^2 s^{-1}; calculated hydrodynamic radius = 10.5 Å. ESI-MS (m/z): 1052.1 (2+) 651.7 (3+) 451.6 (4+)

^{1}H and ^{13}C NMR showed multitude of peaks preventing assignment as discussed in "Chapter 2: Cage Assembly: Direct Assembly using Alternative Pd Salts"

4.4.3 Cage Stability Tests

Time Dependent Stability

To observe the stability of the cage in solution, 6 ^{1}H NMR spectra were recorded over a 48 hour time period of **C13** (0.08 mM, 0.5 mL d_6-DMSO, 1mM internal standard). The ^{1}H NMR spectra were then stacked and normalised against the internal standard, and the integrals compared over the time points to ensure the cage concentration remained constant throughout.

Salt Stability Tests

General Procedure for Salt Stability Studies

Pd_2L_4 cage (1 mM, 0.5 mL) was dissolved in d_6-DMSO (with internal standard) and the ^1H NMR spectrum recorded. Saline solution (NaCl 1% in D_2O, 50 µL) was then added to the complex solution and the mixture sonicated for 30 seconds before recording ^1H NMR spectra at interval time points. ^1H NMR spectra stacked and normalised against the internal standard. Compared integrals over the time points to work out cage consumed.

Glutathione Stability Tests

Procedure for Glutathione Stability Studies

Pd_2L_4 cage (1 mM, 0.5 mL) was dissolved in d_6-DMSO (with internal standard) and the ^1H NMR spectrum recorded. Reduced glutathione (3 mg, 10 µmol) was then added to the complex solution and the mixture was sonicated for 30 seconds before recording a ^1H NMR spectra at interval time points over a 9 h time period. ^1H NMR spectra stacked and normalised against the internal standard. Compared integrals over the time points to work out cage consumed.

4.4.4 Host-Guest Chemistry

Anionic Guests

In Water Based Solvents

General Procedure for Host-Guest Chemistry

To a solution of Pd_2L_4 cage (1.2 mg, 0.75 µmol) in D_2O:CD_3CN (0.376 mL: 0.125 mL with internal standard) was added salt guest (0 Eq, 0.5 Eq, and 1 Eq) and the ^1H NMR spectra recorded. Pseudo titration ^1H NMR spectra was stacked and normalised according to the internal standard.

In Organic Solvents

General Procedure for Host-Guest Chemistry

To a solution of Pd_2L_4 cage (1 mg, 0.5 µmol) in MeOD (0.5 mL with internal standard) was added salt guest (0 Eq and 1 Eq) and the ^1H NMR spectra recorded. Pseudo titration ^1H NMR spectra stacked and normalised against the internal standard.

Neutral Guests

General Procedure for Host-Guest Chemistry

To a solution of Pd_2L_4 cage (1 mg, 5×10^{-4} mmol) in deuterated solvent (0.5 mL with internal standard) was added fluorinil (0 Eq. and 1 Eq.) and the ^1H NMR spectra recorded. Pseudo titration ^1H NMR spectra stacked and normalised against the internal standard. Divide the integrals over the time points to work out cage consumed.

4.5 References

1 J. K. Clegg, J. R. Nitschke, V. Gouverneur, J. Brown, J. Cremers, A. J. Hogben, B. Breiner, M. M. J. Smulders and J. D. Thoburn, *Chem. Sci*, 2013, **4**, 89.

2 T. Kusukawa and M. Fujita, *J. Am. Chem. Soc.*, 2002, **124**, 13576–13582.

3 A. Kalra, E. Yetiskul, C. J. Wehrle and F. Tuma, *Physiology*, StatPearls Publishing, 2022.

4 D. A. McMorran and P. J. Steel, *Angew Chem Int Ed Engl*, 1998, **37**, 3295–3297.

5 J. E. M. Lewis, E. L. Gavey, S. A. Cameron and J. D. Crowley, *Chem. Sci.*, 2012, **3**, 778–784.

6 J. Han, A. Schmidt, T. Zhang, H. Permentier, G. M. M. Groothuis, R. Bischoff, F. E. Kühn, P. Horvatovich and A. Casini, *Chem. Commun.*, 2017, **53**, 1405–1408.

7 R. Cosialls, C. Simo, S. Borros, V. Gçomez-Vallejo, C. Schmidt, J. Llop, A. Belen Cuenca and A. Casini, *Chem. – A Eur. J.*, 2022, **29**, 1–7.

8 B. Woods, R. D. M. Silva, C. Schmidt, D. Wragg, M. Cavaco, V. Neves, V. F. C. Ferreira, L. Gano, T. S. Morais, F. Mendes, J. D. G. Correia and A. Casini, *Bioconjug. Chem.*, 2021, **32**, 1399–1408.

9 X. Schaapkens, J. N. H. Reek and T. J. Mooibroek, *Inorg. Chem. Commun.*, 2022, **139**, 109284.

10 S. M. McNeill, D. Preston, J. E. M. Lewis, A. Robert, K. Knerr-Rupp, D. O. Graham, J. R. Wright, G. I. Giles and J. D. Crowley, *Dalt. Trans.*, 2015, **44**, 11129–11136.

11 D. Preston, A. Fox-Charles, W. K. C. Lo and J. D. Crowley, *Chem. Commun.*, 2015, **51**, 9042–9045.

12 J. Wang, T. A. Young, F. Duarte and P. J. Lusby, *J. Am. Chem. Soc*, 2020, **142**, 17743–17750.

13 S. Sudan, D. W. Chen, C. Berton, F. Fadaei-Tirani and K. Severin, *Angew. Chem. Int. Ed.*, 2023, **62**, 1–5.

14 V. Martí-Centelles, A. L. Lawrence and P. J. Lusby, *J. Am. Chem. Soc.*, 2018, **140**, 2862–2868.

15 R. L. Spicer, A. D. Stergiou, T. A. Young, F. Duarte, M. D. Symes and P. J. Lusby, *J. Am. Chem. Soc.*, 2020, **142**, 2134–2139.

16 K. J. Kilpin, M. L. Gower, S. G. Telfer, G. B. Jameson and J. D. Crowley, *Inorg. Chem*, 2011, **50**, 1123.

17 A. M. Johnson, O. Moshe, A. S. Gamboa, B. W. Langloss, J. F. K. Limtiaco, C. K. Larive and R. J. Hooley, *Inorg. Chem.*, 2011, **50**, 9430–9442.

18 D. Preston, S. M. Mcneill, J. E. M. Lewis, G. I. Giles and J. D. Crowley, *Dalt. Trans.*, 2016, **45**, 8050.

19 B. P. Burke, W. Grantham, M. J. Burke, G. S. Nichol, D. Roberts, I. Renard, R. Hargreaves, C. Cawthorne, S. J. Archibald and P. J. Lusby, *J. Am. Chem. Soc.*, 2018, **140**, 16877–16881.

20 M. El Bakkari and J. M. Vincent, *Org. Lett.*, 2004, **6**, 2765–2767.

21 A.-C. C. Carlsson, K. Mehmeti, M. Uhrbom, A. Karim, M. Bedin, R. Puttreddy, R. Kleinmaier, A. A. Neverov, B. Nekoueishahraki, J. Gräfenstein, K. Rissanen and M. Erdélyi, *J. Am. Chem. Soc.*, 2016, **138**, 9853–9863.

22 P. Liao, B. W. Langloss, A. M. Johnson, E. R. Knudsen, F. S. Tham, R. R. Julian and R. J. Hooley, *Chem. Commun.*, 2010, **46**, 4932–4934.

23 E. G. Percástegui, T. K. Ronson and J. R. Nitschke, *Chem. Rev.*, 2020, **120**, 13480–13544.

24 D. P. August, G. S. Nichol and P. J. Lusby, *Angew. Chemie Int. Ed.*, 2016, **55**, 15022–15026.

25 T. Tateishi, Y. Yasutake, T. Kojima, S. Takahashi and S. Hiraoka, *Commun. Chem. 2019 21*, 2019, **2**, 1–9.

26 S. Freye, J. Hey, A. Torras-Galán, D. Stalke, R. Herbst-Irmer, M. John and G. H. Clever, *Angew. Chemie Int. Ed.*, 2012, **51**, 2191–2194.

27 F. Ibukuro, T. Kusukawa and M. Fujita, *J. Am. Chem. Soc.*, 1998, **120**, 8561–8562.

28 H. Ehrsson, I. Wallin and J. Yachnin, *Med. Oncol.*, 2002, **19**, 261–265.

29 N. Nezafati, F. Moztarzadeh and S. Hesaraki, *Biotechnol. Bioprocess Eng.*, 2012, **17**, 746–754.

30 H. Jay Forman, H. Zhang and A. Rinna, *Mol Asp. Med.*, 2009, **1**, 1–12.

31 K. Shahzad, A. S. A. Majid, M. Khan, M. A. Iqbal and A. Ali, *Rev. Inorg. Chem.*, 2021, **41**, 151–198.

32 W. Grantham and P. J. Lusby, University of Edinburgh, 2019.

33 V. Martí-Centelles, F. Duarte and P. J. Lusby, *Isr. J. Chem.*, 2019, **59**, 257–266.

34 R. L. Spicer and P. J. Lusby, University of Edinburgh, 2020.

35 M. El Bakkari and J. M. Vincent, *Org. Lett.*, 2004, **6**, 2765–2767.

Chapter 5 : Conclusion and Future Outlook

Chapter 5 : Conclusion and Future Outlook

This work explored the biomedical applications of coordination cages, with particular focus on their potential as delivery vessels. This included investigating the interactions of complexes with bioprevalent species, notably proteins, alongside synthesising novel cages to alter biodistribution and host-guest chemistry.

The previously synthesised **C1** tetrahedra was found to strongly interact with Human Serum Albumin (HSA) via a specific 1:1 binding model. This was evidenced using equilibrium dialysis, isothermal titration calorimetry and computational modelling, amongst other techniques. Whilst the other tetrahedral **C2** also showed significant interactions with the protein, the two smaller complexes **C3** and **C4** indicated weaker associations. This validates the theory that the binding between **C1** and HSA is a result of multiple electrostatic interactions between the cationic cage vertices and complimentary amino acid residues.

Ongoing research within the group and with our collaborators at the University of Hull, focusses on repeating the previous SPECT imaging studies following the biodistribution of **C1**·[^{99m}Tc]TcO_4^- but including HSA in the formulation stage. The strong interactions between the cage and the protein are hypothesised to further alter the biodistribution of the encapsulated radioisotope. Future investigations continuing on from this study should probe the specificity of the **C1**:HSA interactions and whether these can be emulated with different proteins, maximising the potential for protein mediated cage delivery.

In an attempt to actively gain control of the *in vivo* biodistribution of **C1**, different methods of bioconjugation were carried out including a comparison between pre- and post-assembly modifications. Given the complexity of robust cage assembly, including the reliance on transligation as opposed to spontaneous self-assembly, bioconjugation was favoured on the fully formed cage. Three novel Co^{III} tetrahedra; **C5, C6** and **C7**, were synthesis and exhibited different functionality of the cage exterior. **C5**, which exhibits *para*-methoxyl groups in comparison to the *para*-amines in **C1**, was found to be insufficiently stable in the presence of biologically abundant species and disassembled in mouse serum. Comparatively **C6,** which features *para*-aminoethanol substitutions, exhibited good stability under biological conditions and retained the encapsulated [^{99m}Tc]TcO_4^- when transferred into mouse serum.

The terminal alcohol present on **C6** serves as a scaffold for bioconjugation with targeting peptides via simple a simple esterification. Ongoing radiochemical research by our collaborators in the Archibald group, focuses on the radiosynthesis of [^{18}F]PF_6^- and [^{124}I]IO_4^- whereby preliminary NMR investigation with the corresponding cold analogues highlighted the radioisotopic anions as good guests for PET

imaging. Further investigations surrounding this research should focus on tethering targeting peptides to the terminal alcohol of **C6** and completing both SPECT and PET imaging experiments to follow the biodistribution of the novel cage system.

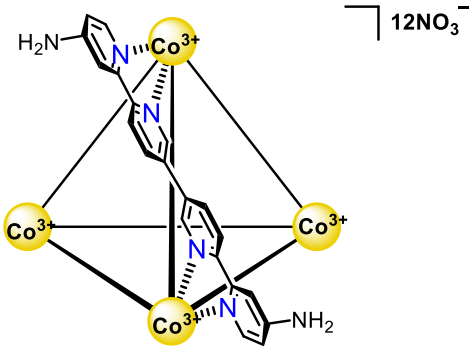
The final chapter of this project focussed on probing whether Pd_2L_4 systems could exhibit sufficient stability and solubility to have potential as biomedical delivery vessels. This was achieved through the generation of a library of novel cages where the stability and host-guest chemistry was investigated through a range of NMR studies. It was evidenced that significant cage modifications were required to provide water solubility which was not reliant on hydrating counter anions, such as NO_3^-. Moreover, the cage cavities were not found to provide strong encapsulation to the anions used to determine radiolabelling potential. Future research should focus on derivatisation of the cages with both *para* and *ortho* amine substitutions to improve stability and solubility, alongside the identification of more appropriate guests.

In summary, this project provides a strong insight into the behaviours of coordination cages under biological conditions. A comprehensive method for determining protein:complex interactions is defined, enabling future research to probe whether protein binding could alter cage biodistribution. A novel robust Co^{III} tetrahedral is synthesised with exterior functionality capable of simple bioconjugation.

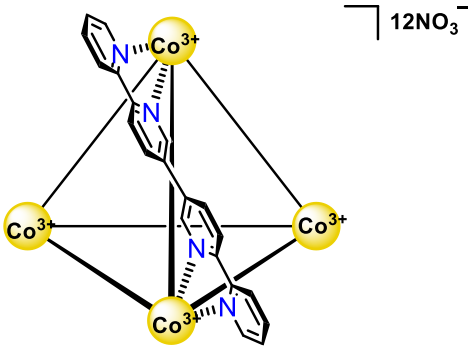
Appendix

Compound Numbering

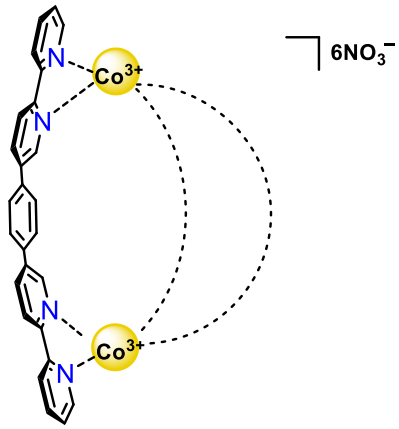
Chapter 2



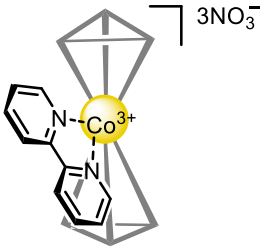
C1



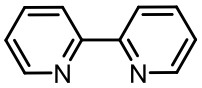
C2



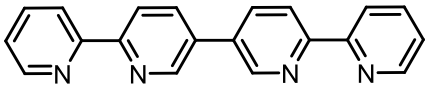
C3



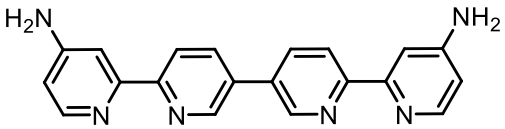
C4



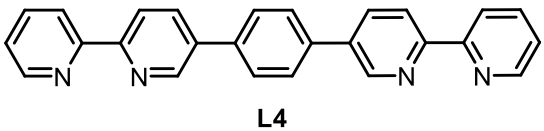
L1



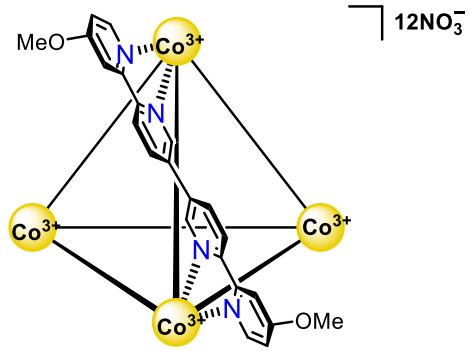
L2



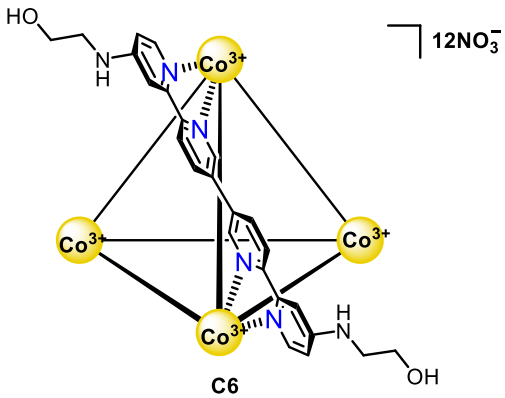
L3



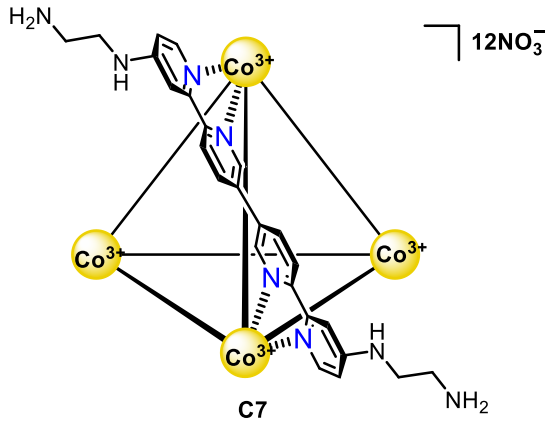
L4



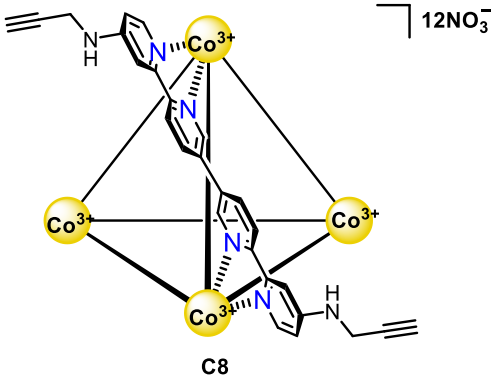
C5



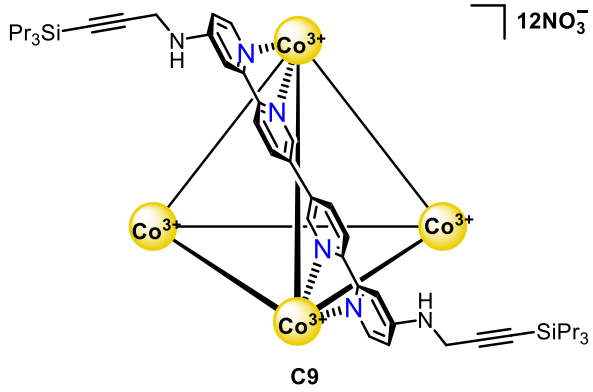
C6



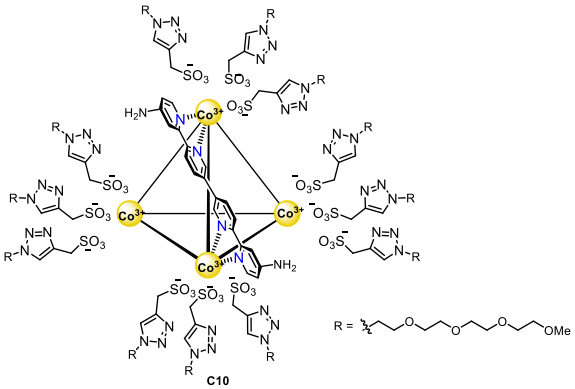
C7



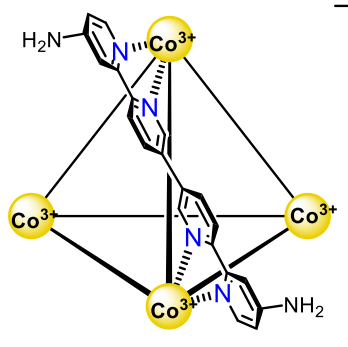
C8



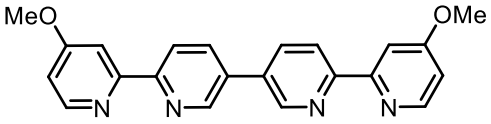
C9



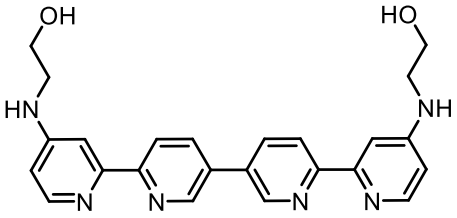
C10



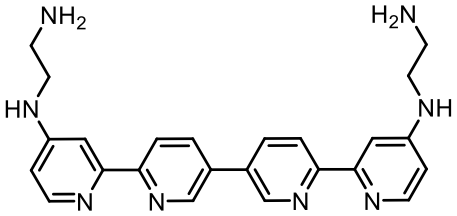
C11



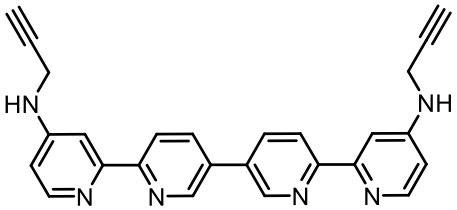
L5



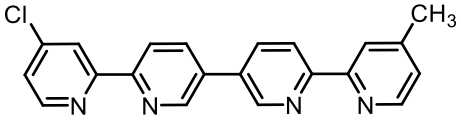
L6



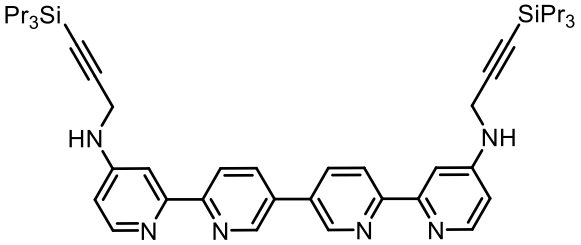
L7



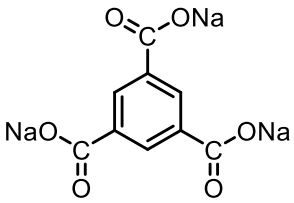
L8



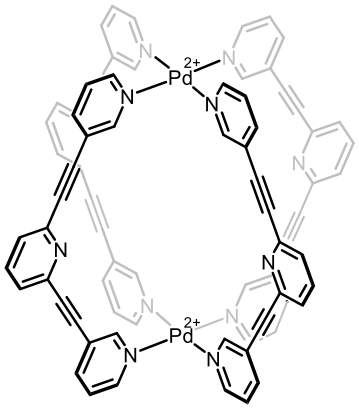
L9



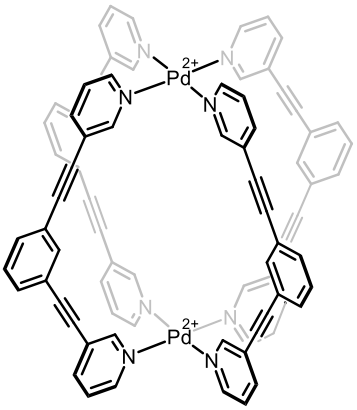
L10



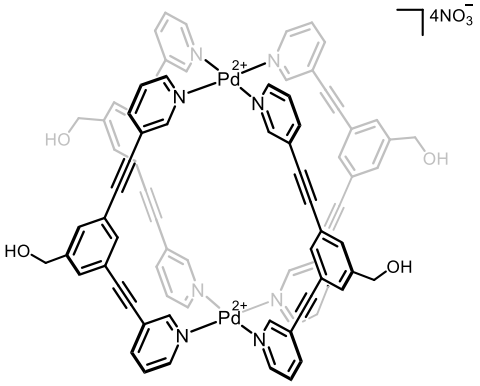
L11



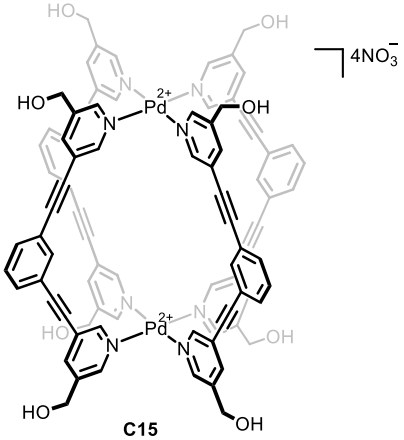
C12



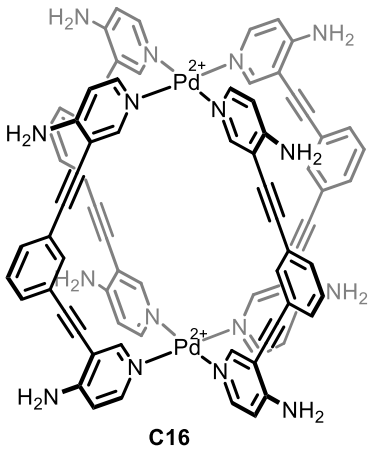
C13



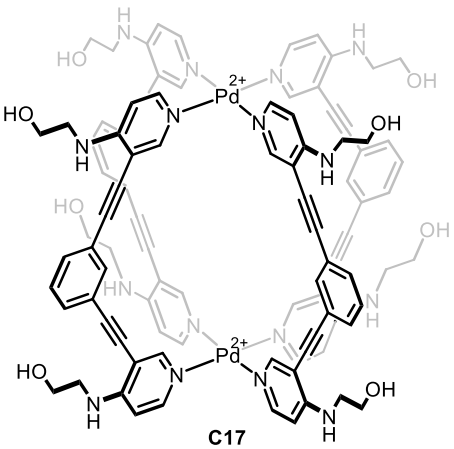
C14



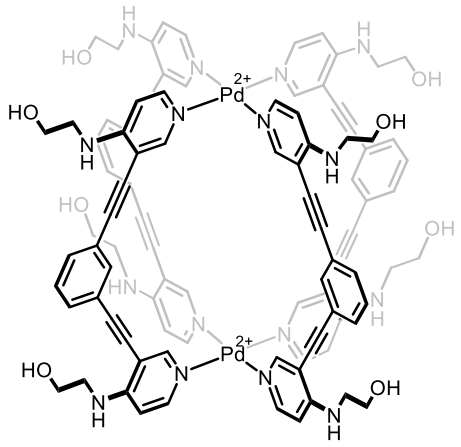
C15



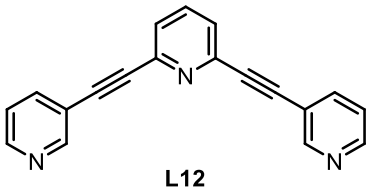
C16



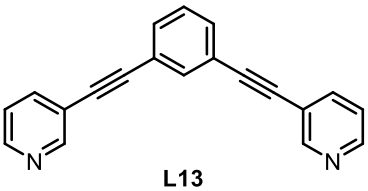
C17



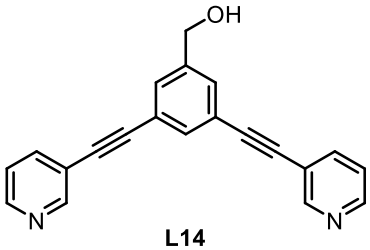
C18



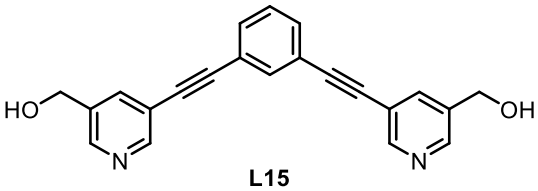
L12



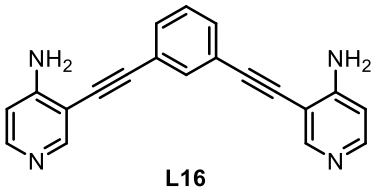
L13



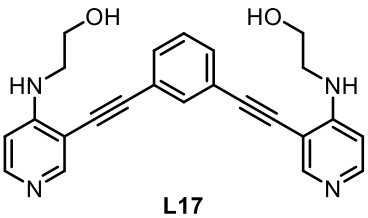
L14



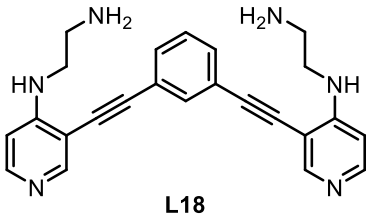
L15



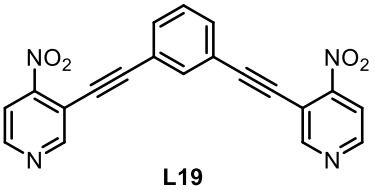
L16



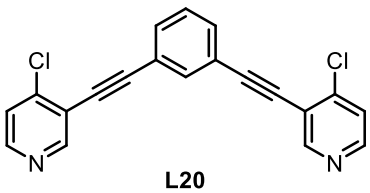
L17



L18



L19



L20

Supplementary Information

Chapter 2

Characterisation Data - C4

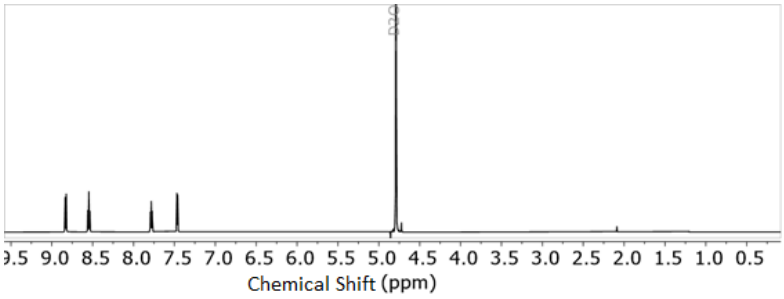


Figure SI 1. ^{1}H NMR (500 MHz, D_2O) of **C4**

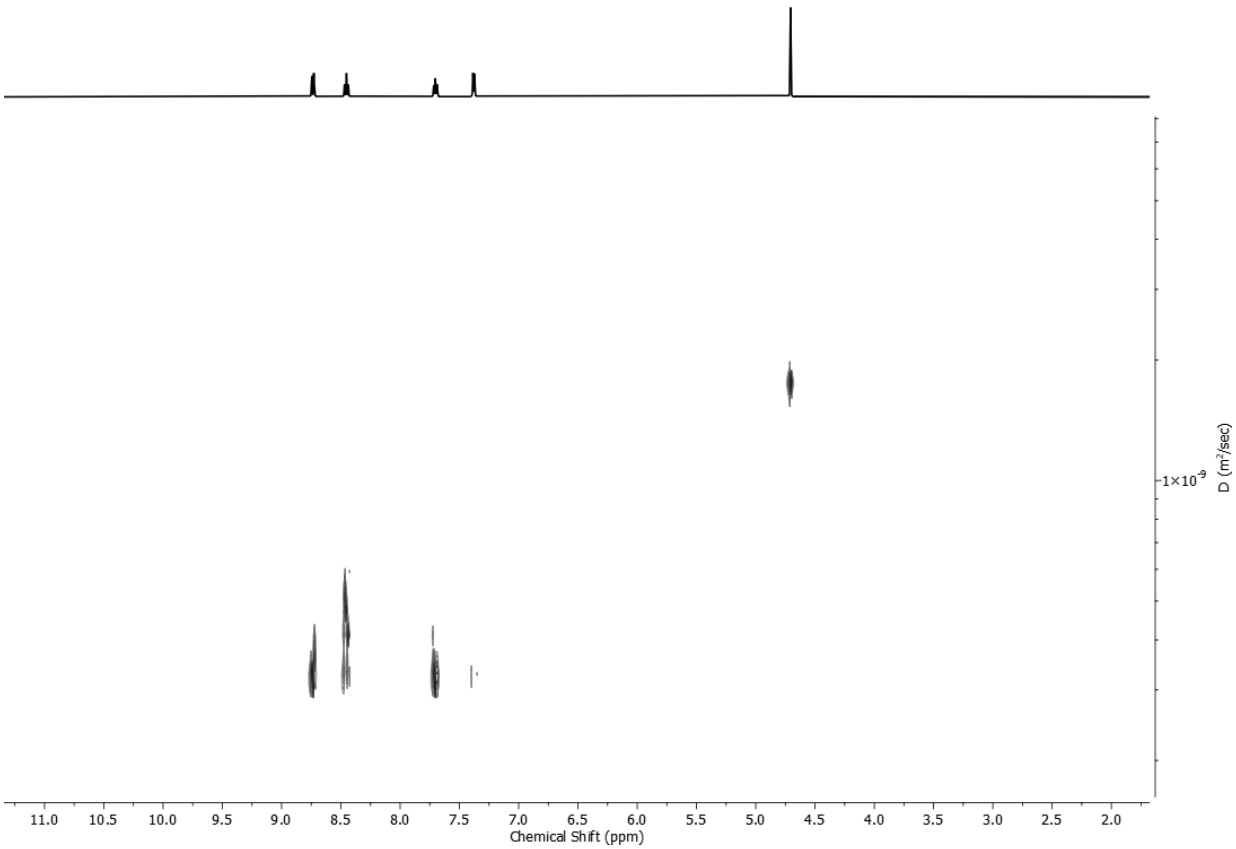


Figure SI 2. ^{1}H DOSY (500 MHz, D_2O) of **C4**

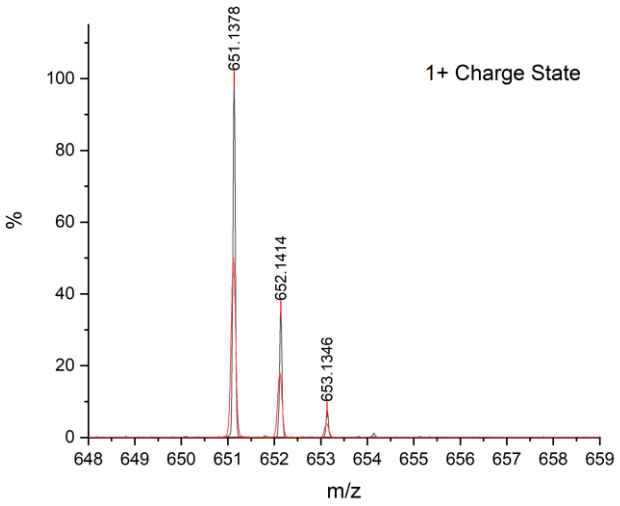


Figure SI 3. ESI Mass Spectrum of **C4** 1+ Charge State

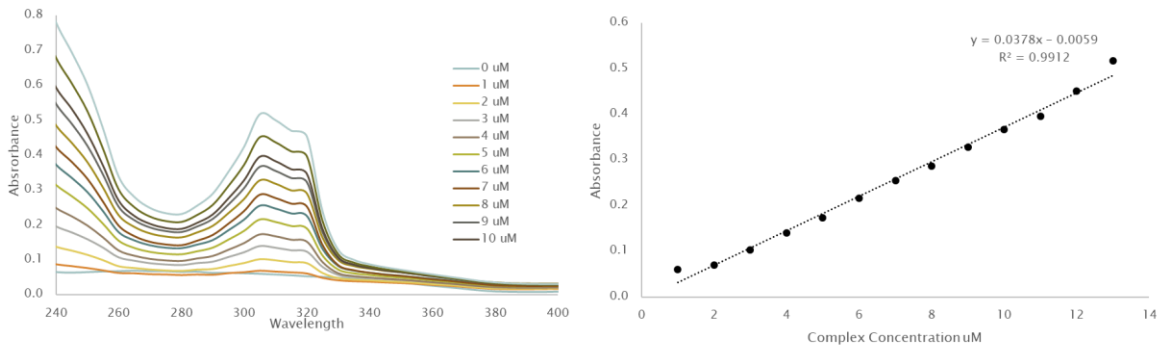


Figure SI 4. UV-Vis Spectroscopy for **C4** extinction coefficient determination

Characterisation of C3

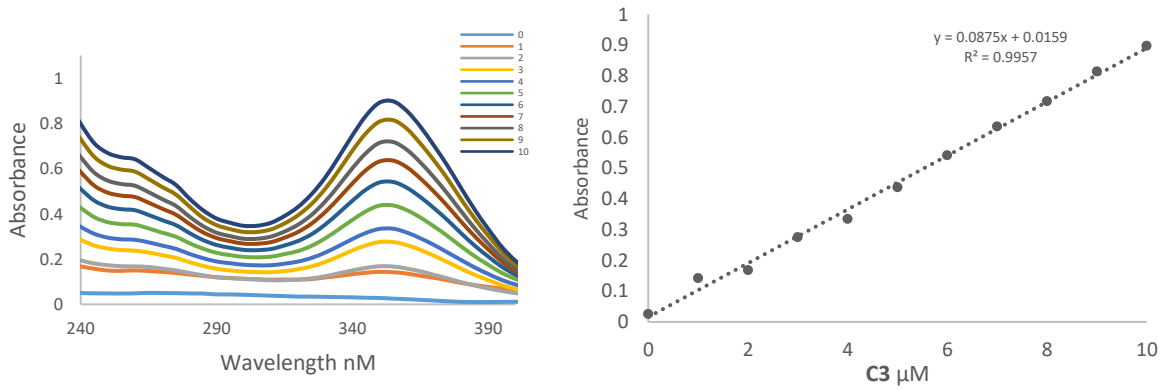


Figure SI 5. UV-Vis Spectroscopy for **C3** extinction coefficient determination

Characterisation of C2

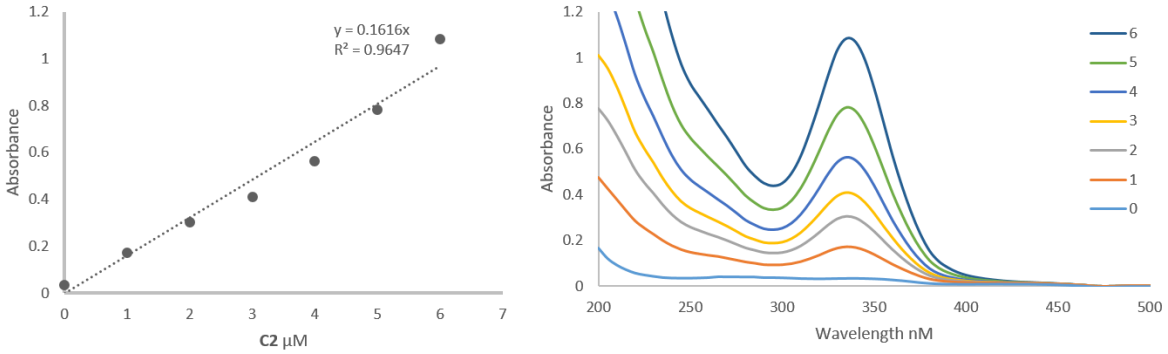


Figure SI 6. UV-Vis Spectroscopy for **C2** extinction coefficient determination

Characterisation of C1

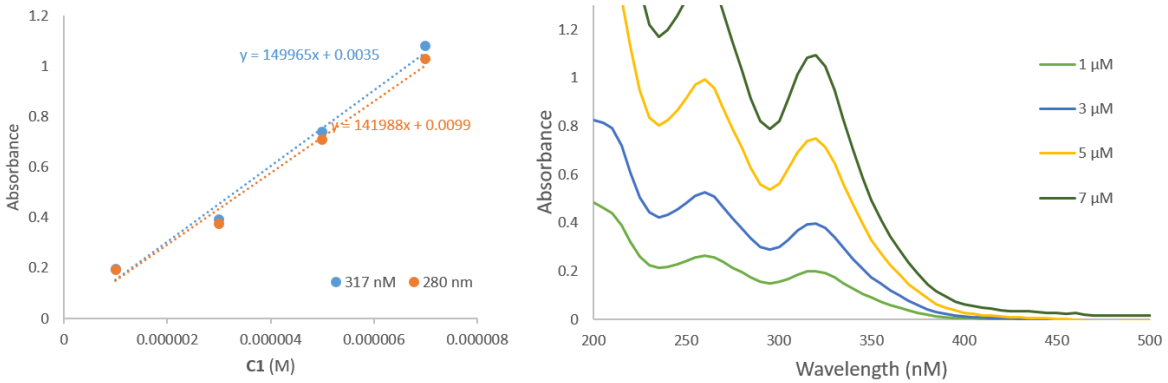


Figure SI 7. UV-Vis Spectroscopy for **C2** extinction coefficient determination

Chapter 3

Characterisation of 2

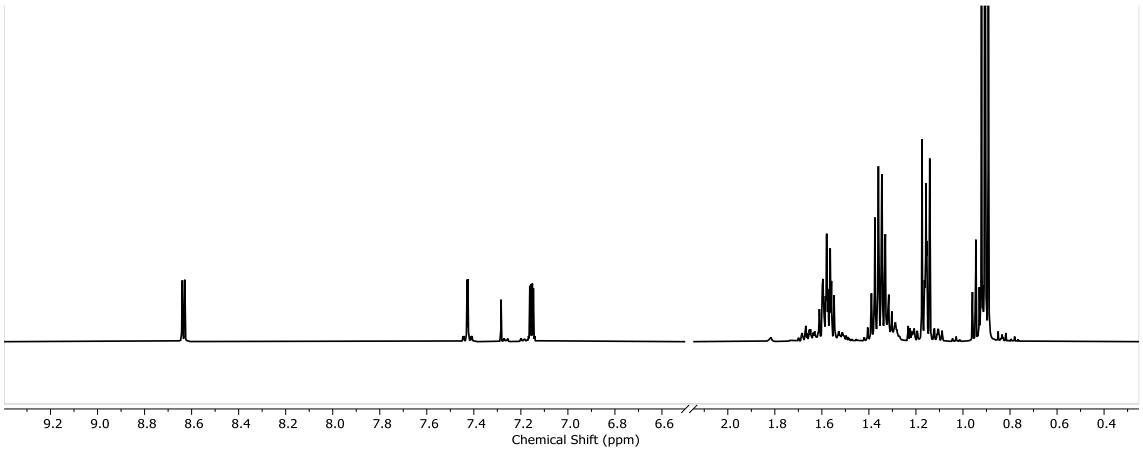


Figure SI 8. ^{1}H NMR (500 MHz, $CDCl_3$) of **2**

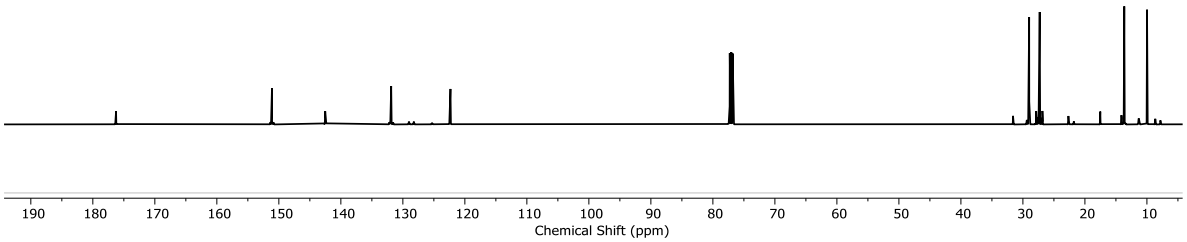


Figure SI 9. ^{13}C NMR (126 MHz, $CDCl_3$) of **2**

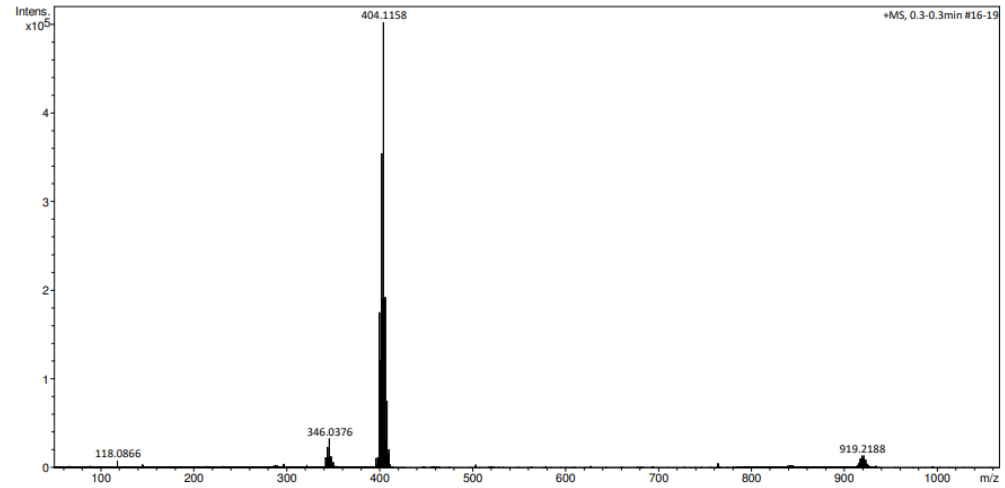


Figure SI 10. ESI Mass Spectrum of **2**

Characterisation of 4

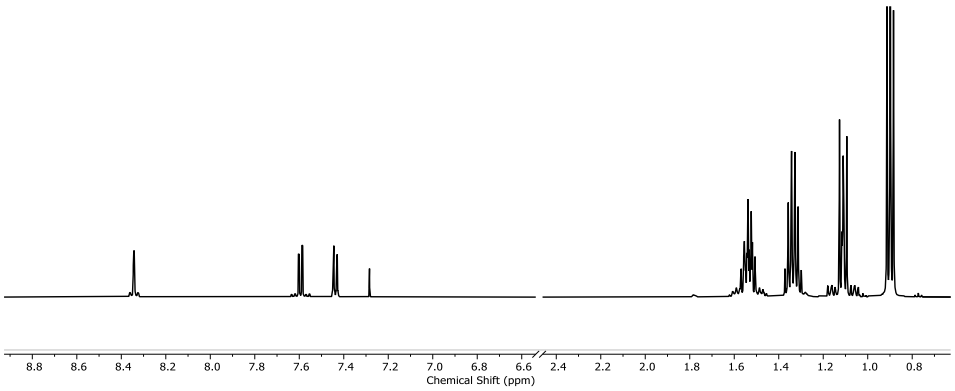


Figure SI 11. ^{1}H NMR (500 MHz, $CDCl_3$) of **4**

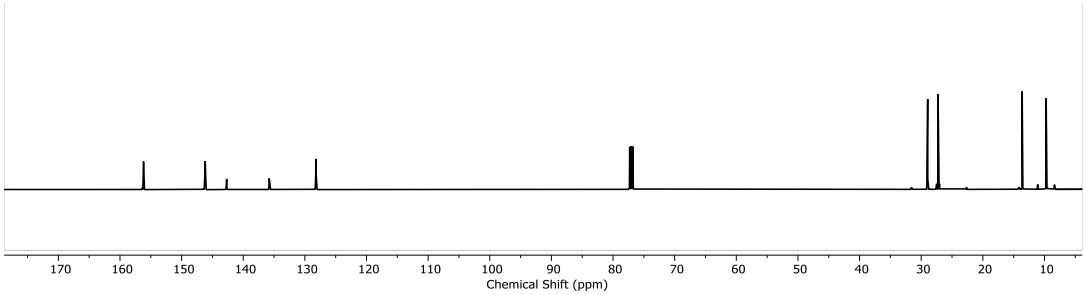


Figure SI 12. ^{13}C NMR (126 MHz, $CDCl_3$) of **4**

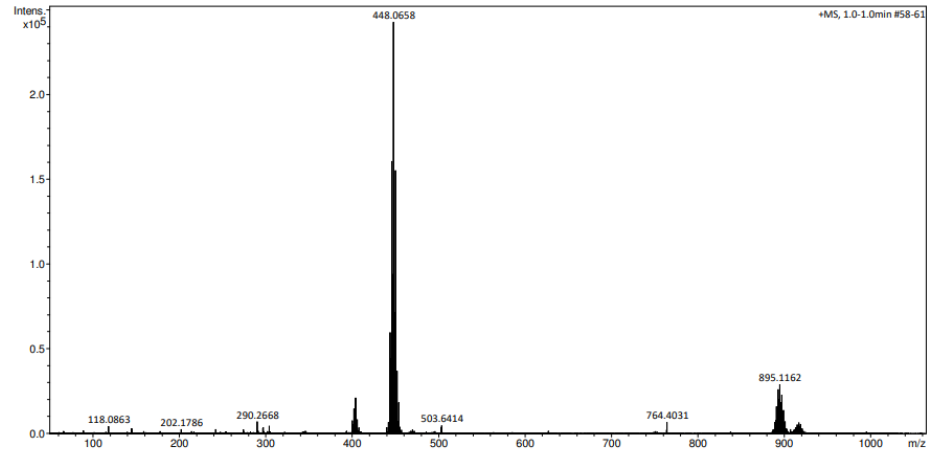


Figure SI 13. ESI Mass Spectrum of **4**

Characterisation of 3

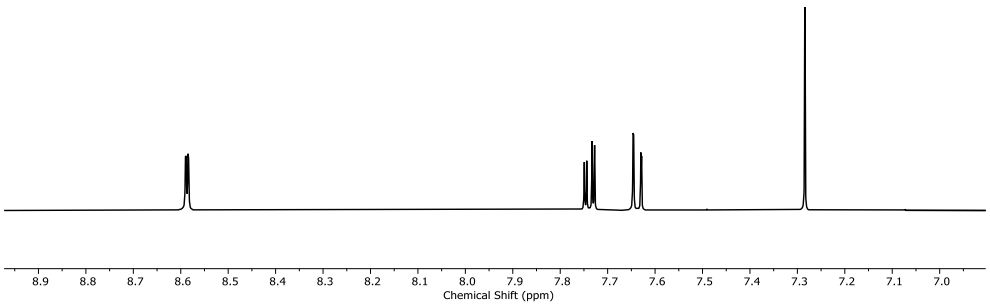


Figure SI 14. ^{1}H NMR (500 MHz, $CDCl_3$) of **3**

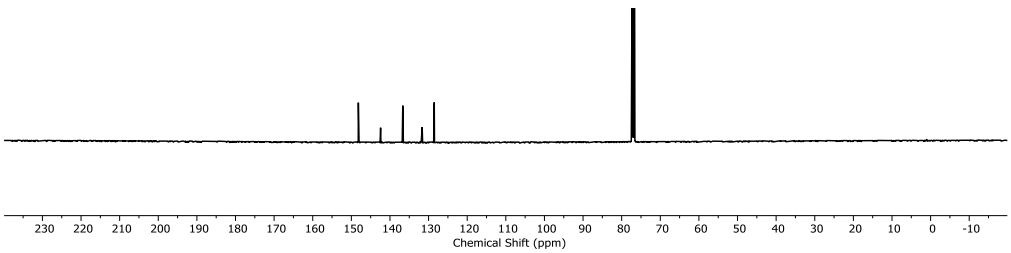


Figure SI 15. ^{13}C NMR (126 MHz, $CDCl_3$) of **3**

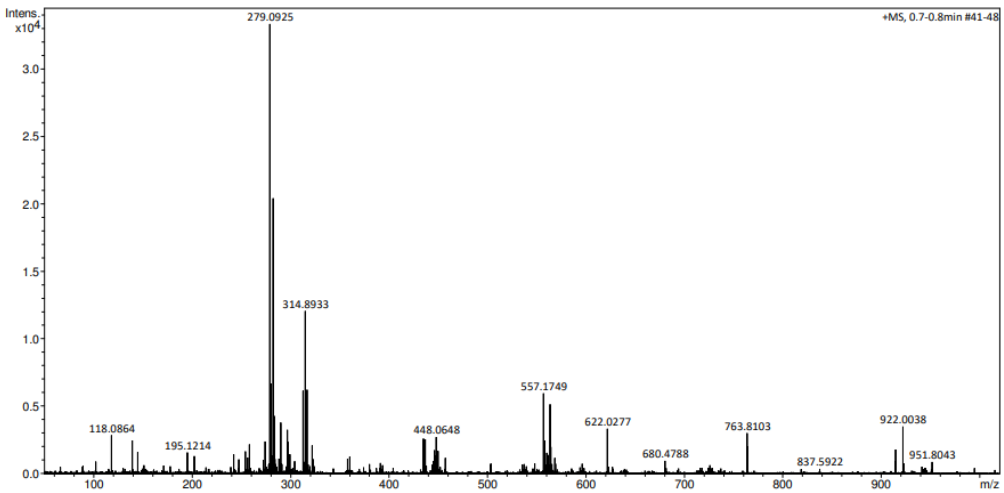


Figure SI 16. ESI Mass Spectrum of **4**

Characterisation of L9

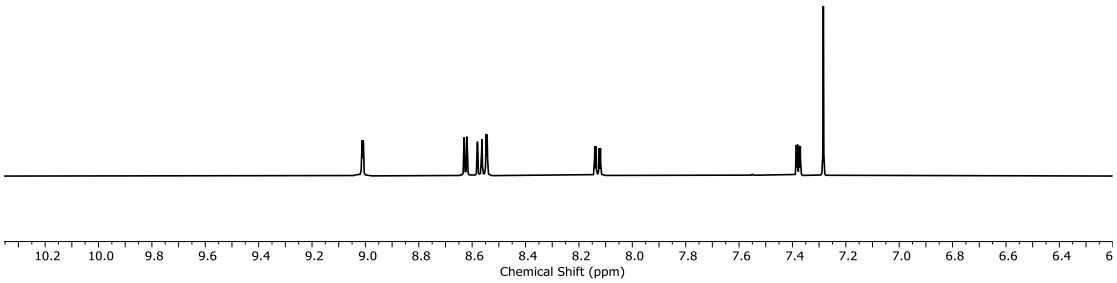


Figure SI 17. ^{1}H NMR (500 MHz, $CDCl_3$) of **L9**

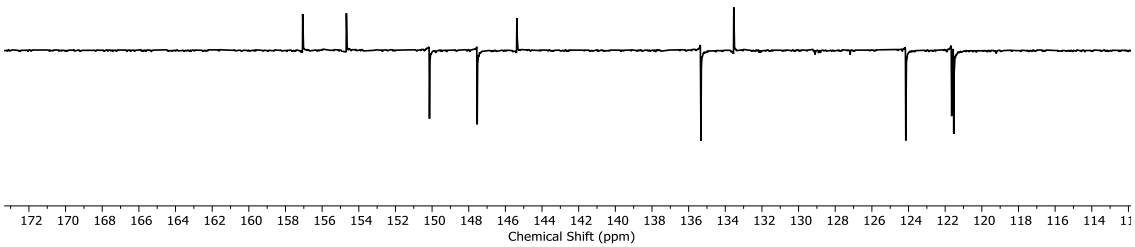


Figure SI 18. ^{13}C DEPT NMR (126 MHz, $CDCl_3$) of **L9**

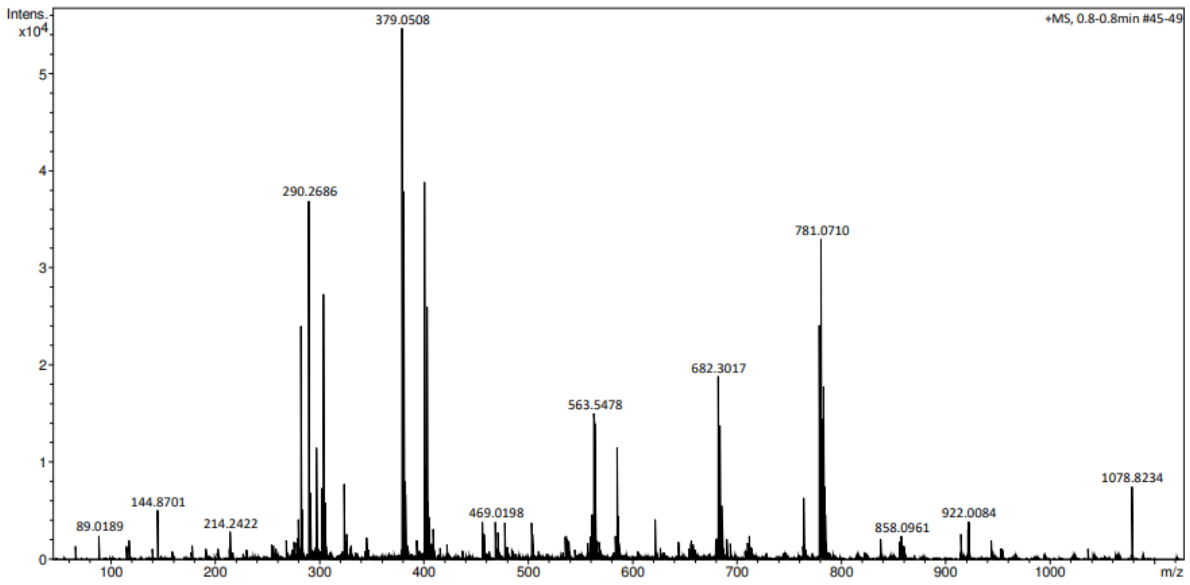


Figure SI 19. ESI Mass Spectrum of **L9**

Characterisation of L5

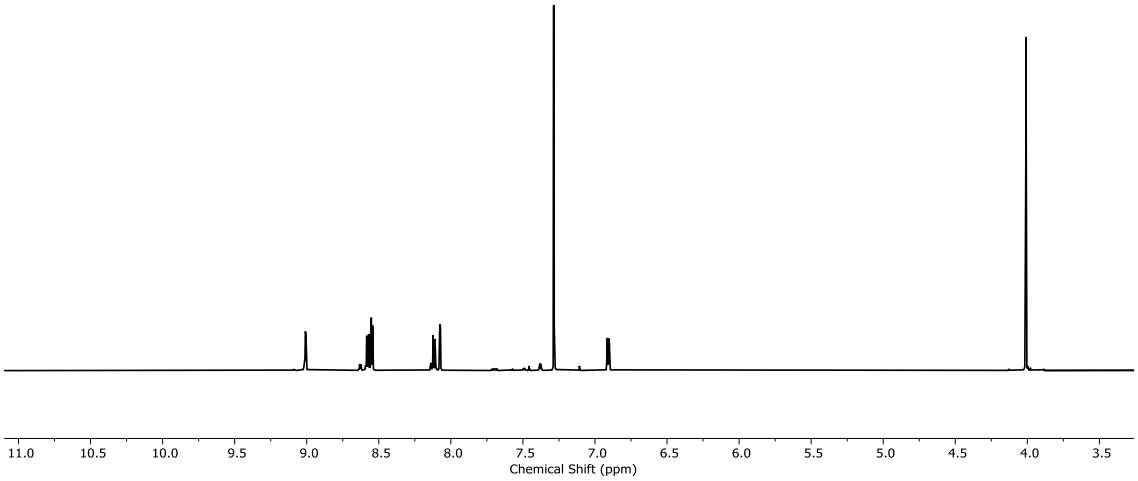


Figure SI 20. ^{1}H NMR (500 MHz, $CDCl_3$) of **L5**

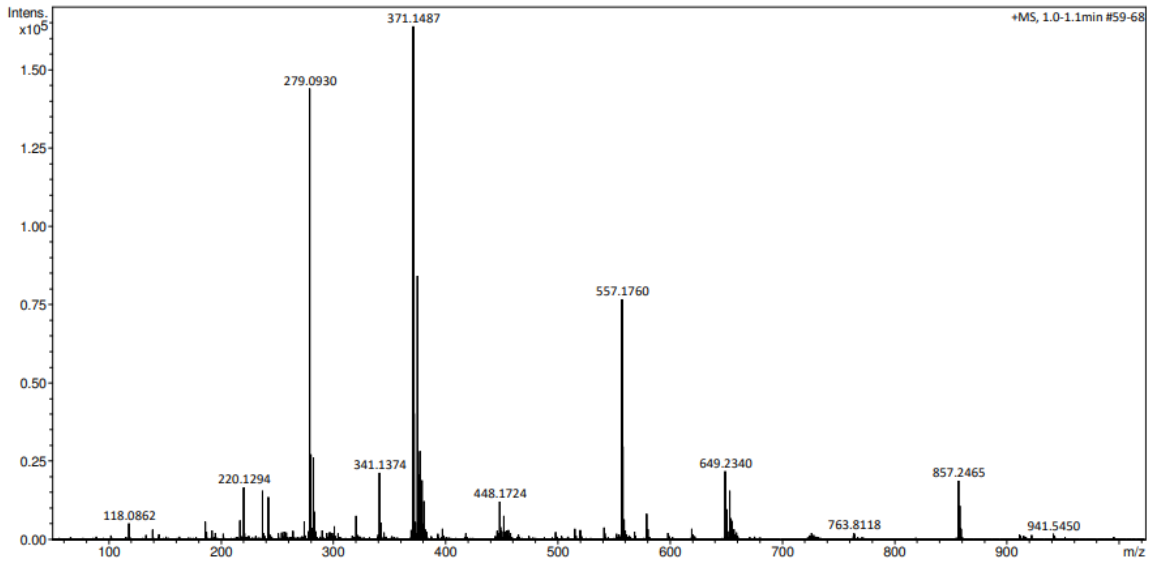


Figure SI 21. ESI Mass Spectrum of **L5**

Characterisation of L6

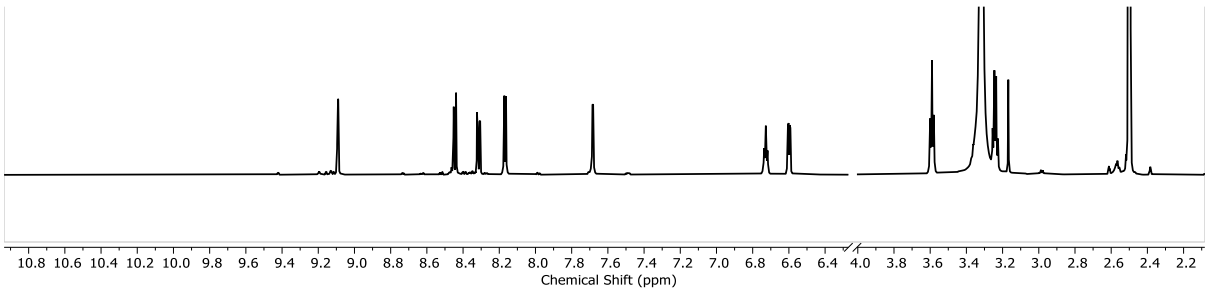


Figure SI 22. ^{1}H NMR (500 MHz, d_6-DMSO) of **L6**

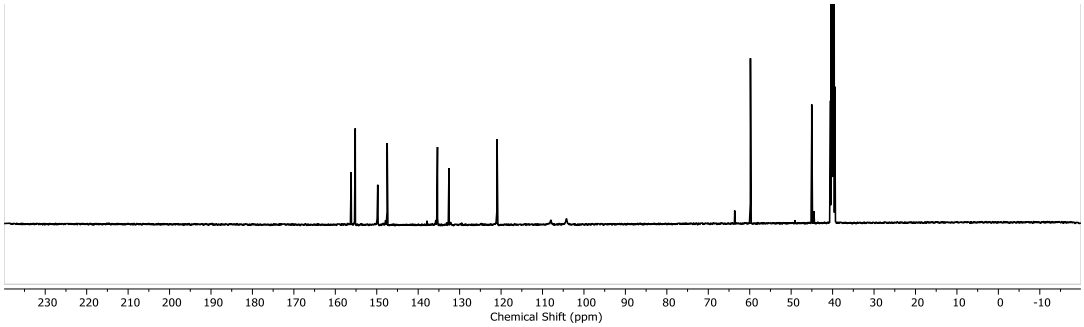


Figure SI 23. ^{13}C NMR (500 MHz, d_6-DMSO) of **L6**

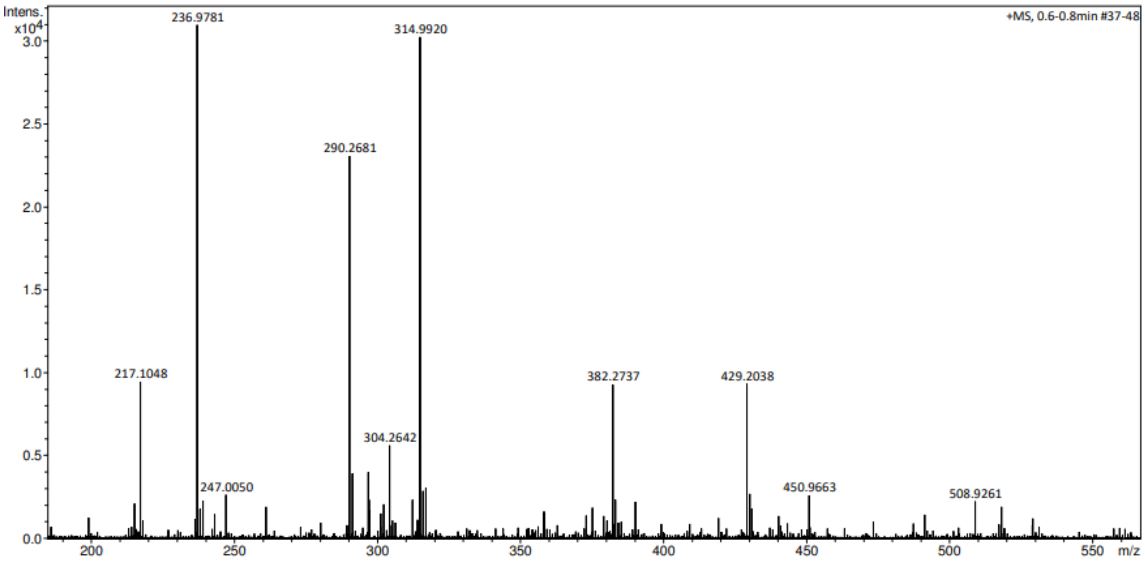


Figure SI 24. ESI Mass Spectrum of **L6**

Characterisation of L7

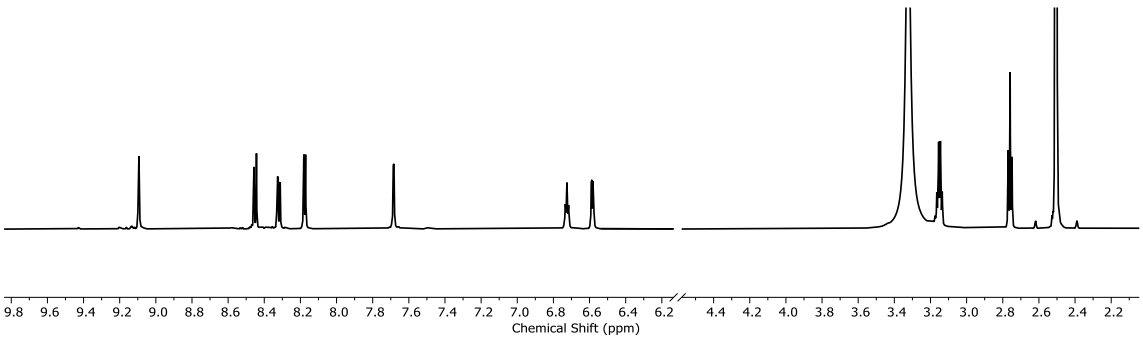


Figure SI 25. ^{1}H NMR (500 MHz, d_6-DMSO) of **L7**

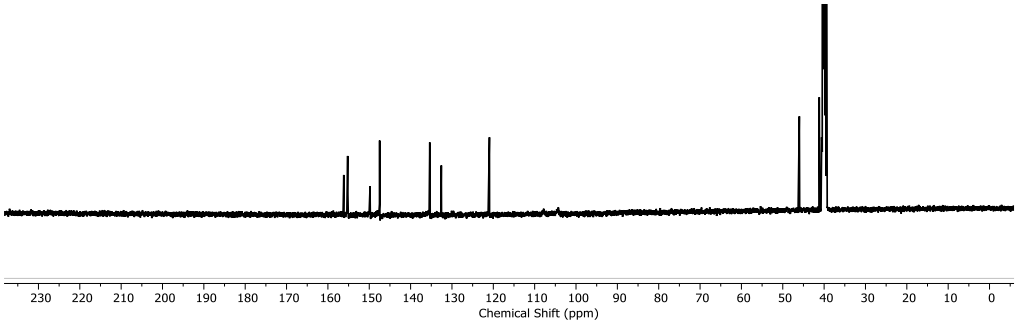


Figure SI 26. ^{13}C NMR (500 MHz, d_6-DMSO) of **L7**

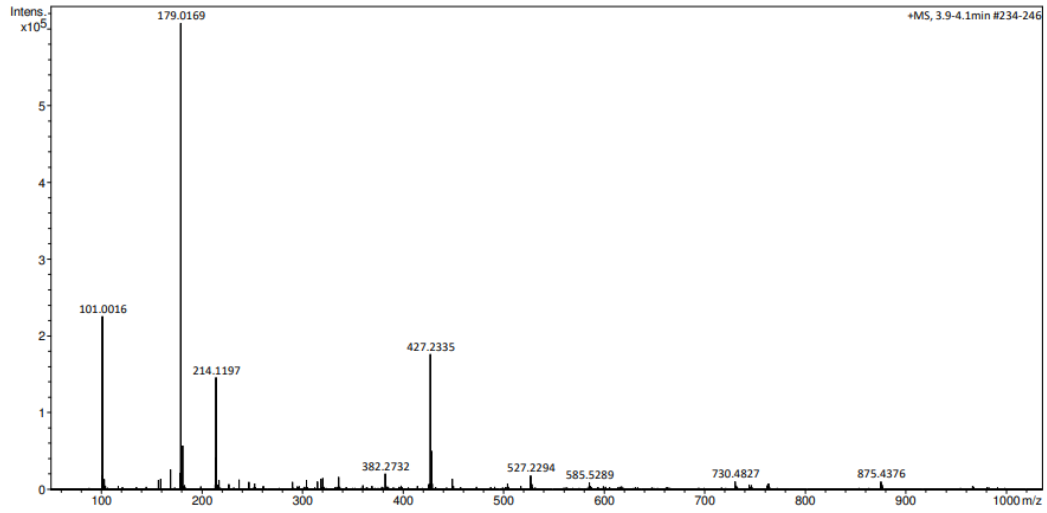


Figure SI 27. ESI Mass Spectrum of **L7**

Characterisation Data of L10

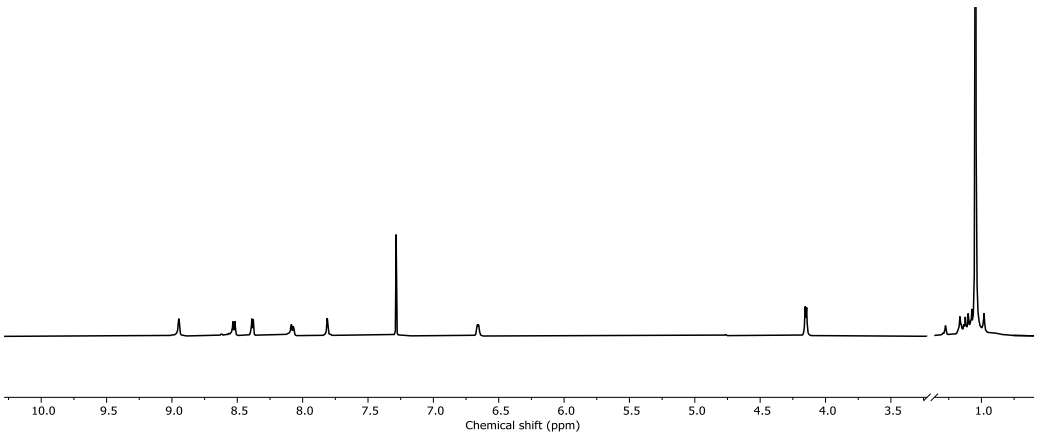


Figure SI 28. ^{1}H NMR (500 MHz, $CDCl_3$) of **L10**

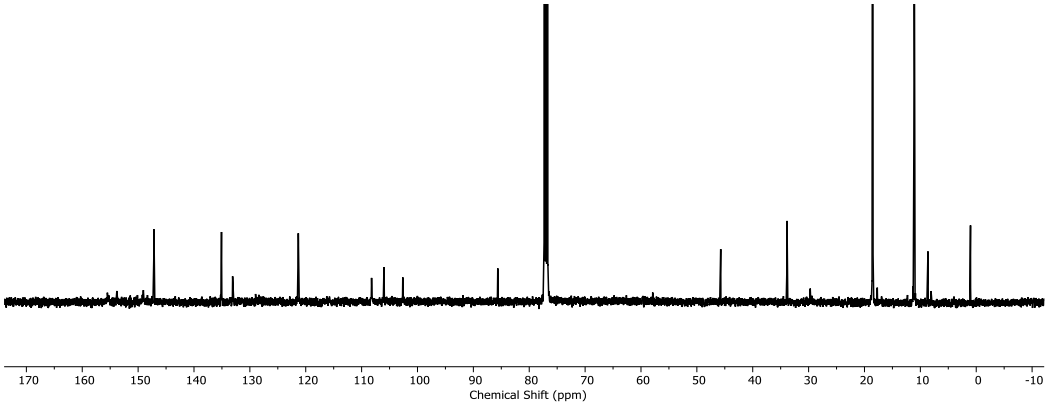


Figure SI 29. ^{13}C NMR (126 MHz, $CDCl_3$) of **L10**

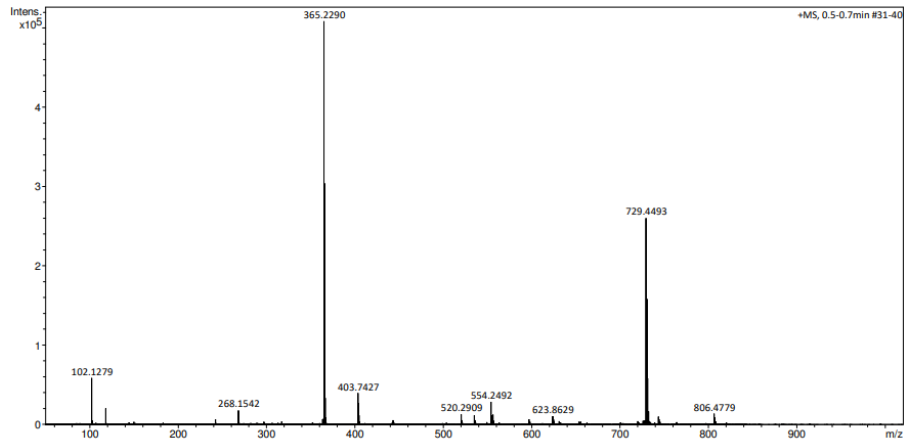


Figure SI 30. ESI Mass Spectrum of **L10**

Characterisation Data – L8

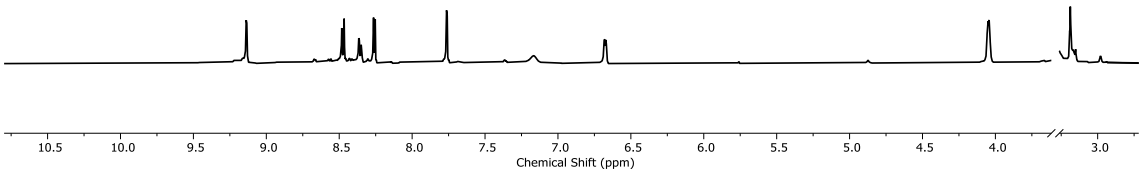


Figure SI 31. ^{1}H NMR (500 MHz, d_6-DMSO) of **L9**

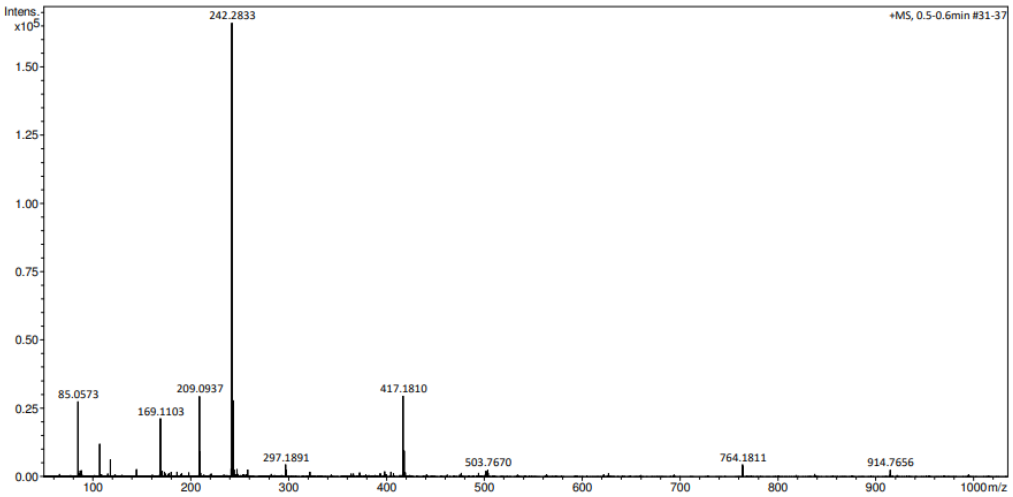


Figure SI 32. ESI Mass Spectrum of **L8**

Characterisation Data – C5

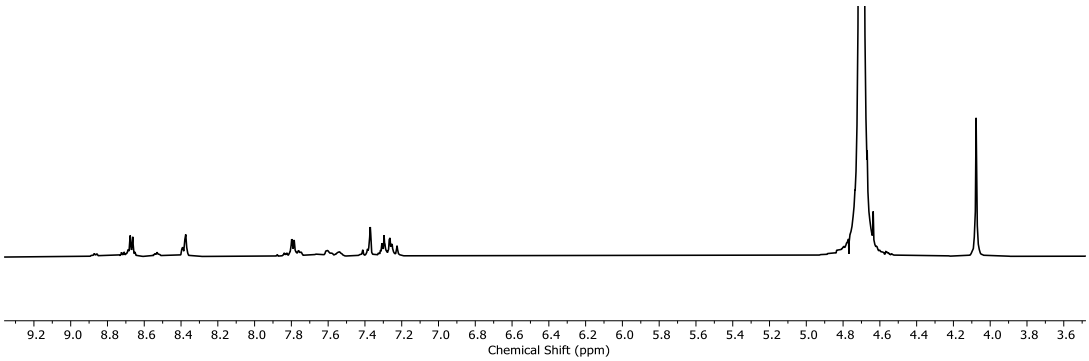


Figure SI 33. ^{1}H NMR (500 MHz, D_2O) of **C5**

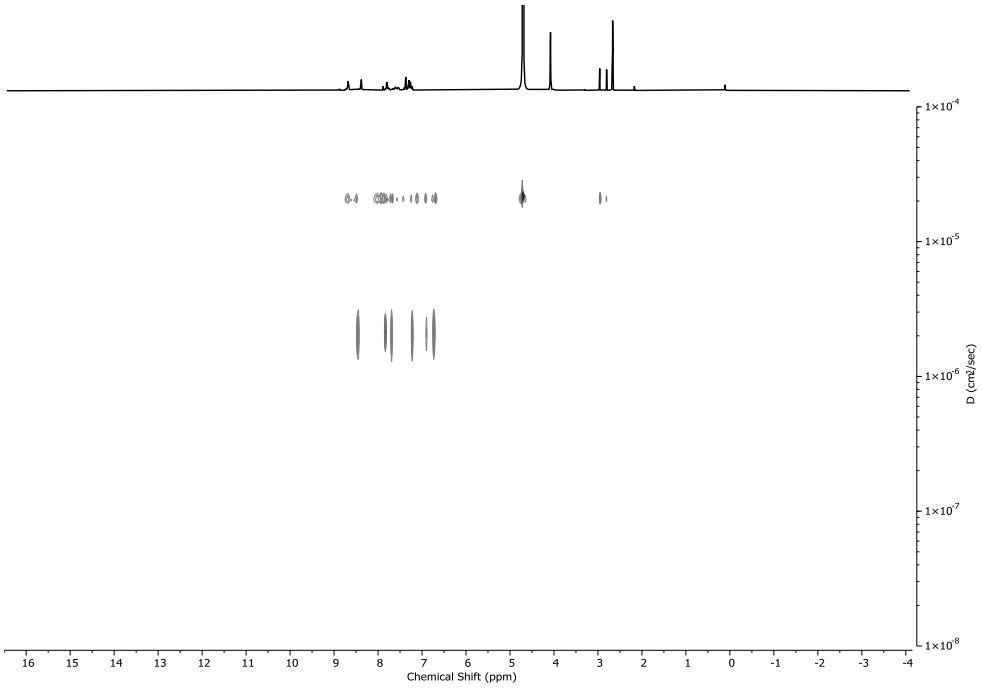


Figure SI 34. ^{1}H DOSY (500 MHz, D_2O) of **C5**

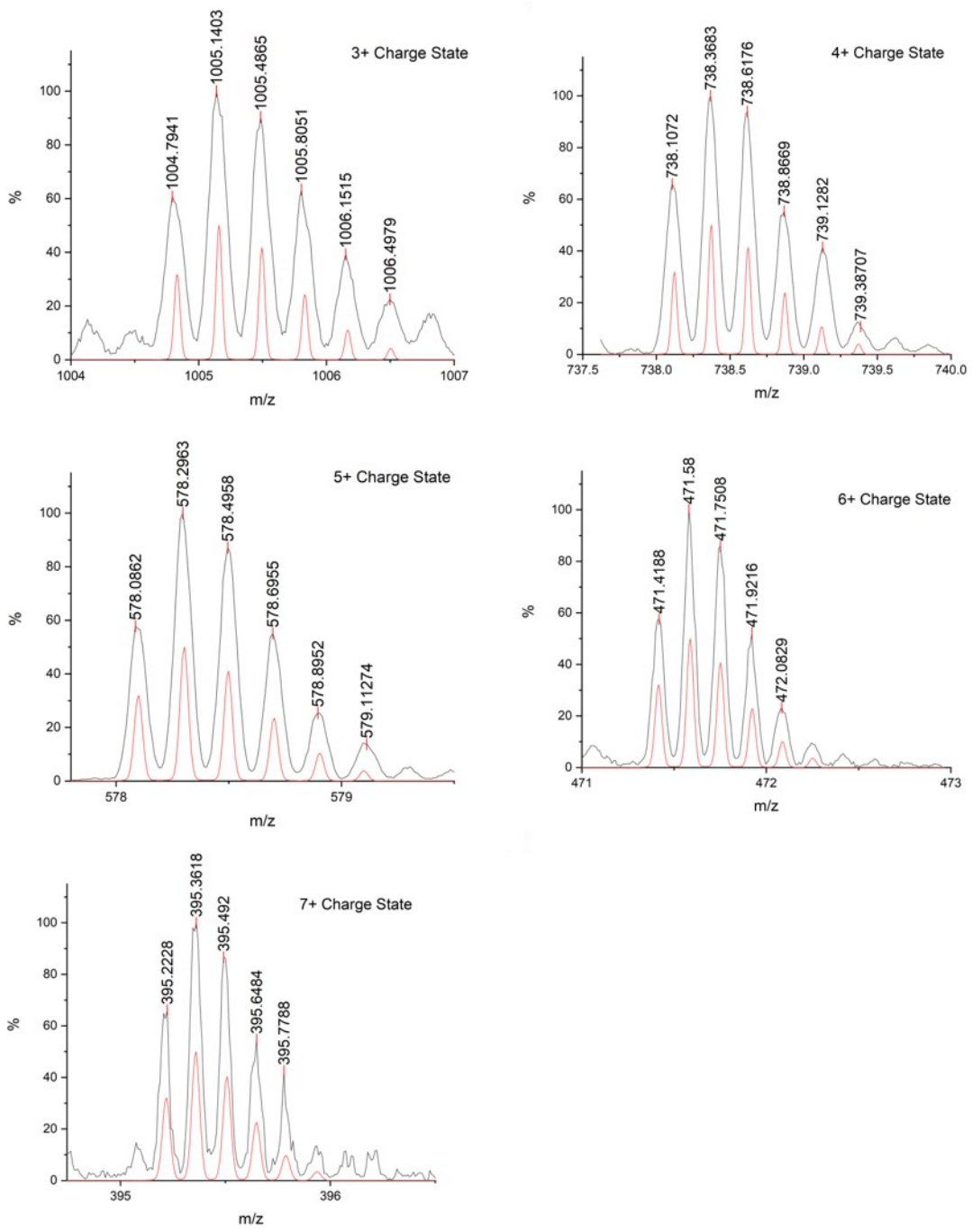


Figure SI 35. ESI Mass Spectrum Charge States for **C5.** Red shows predicted and black shows experimental data.

Characterisation Data – C7

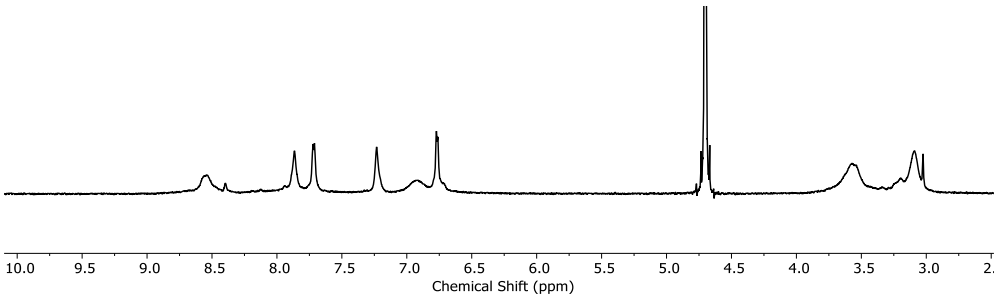


Figure SI 36. ^{1}H NMR (500 MHz, D_2O) of **C7**

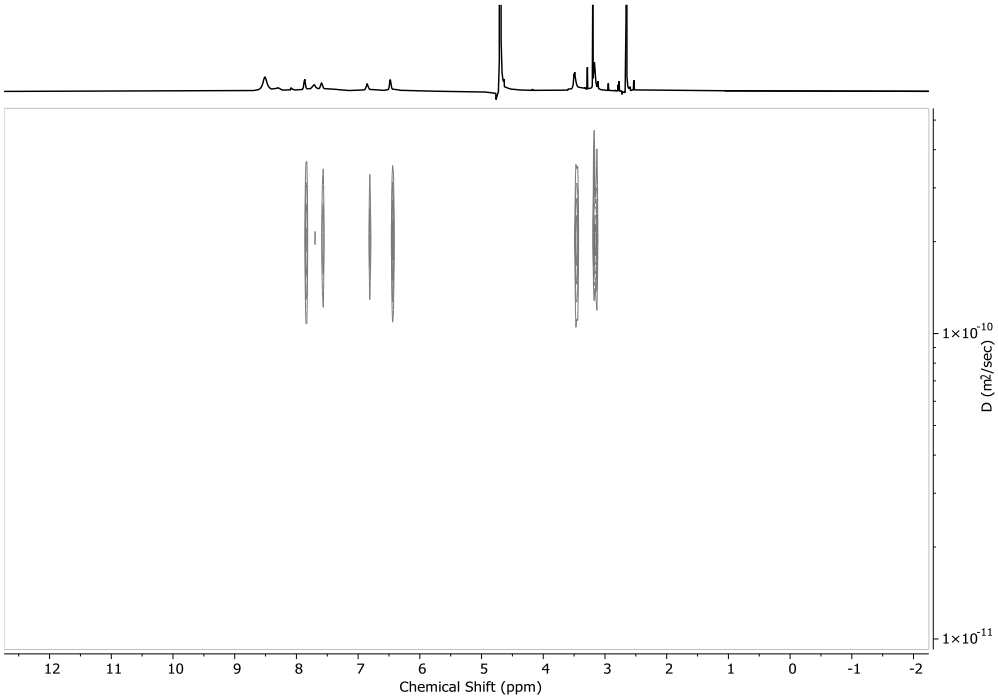


Figure SI 37. ^{1}H DOSY (500 MHz, D_2O) of **C7**

Characterisation Data – C6

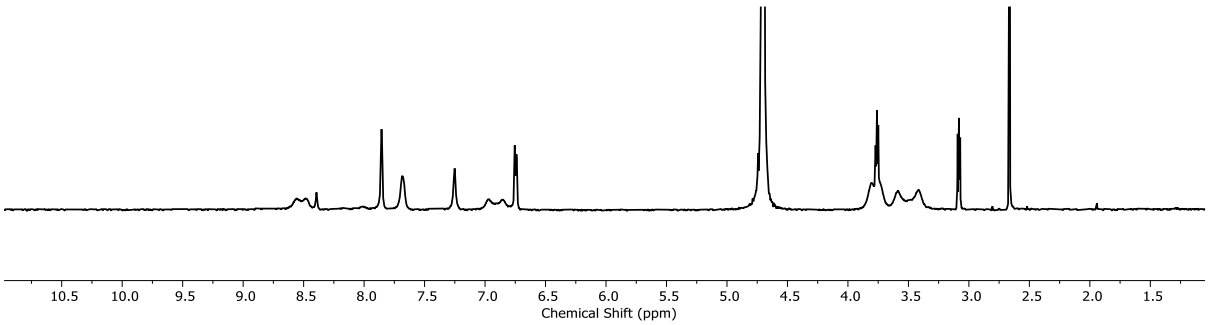


Figure SI 38. ^{1}H NMR (500 MHz, D_2O) of **C6**

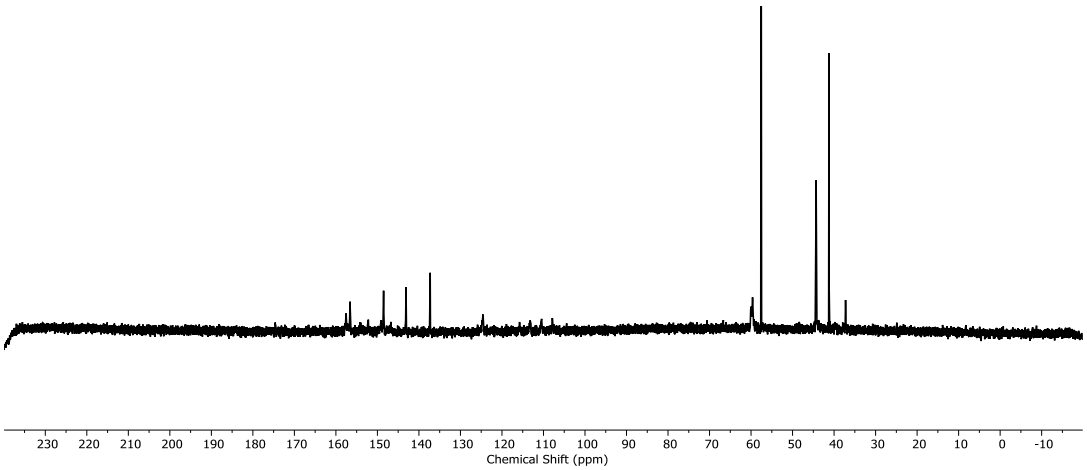


Figure SI 39. ^{13}C NMR (126 MHz, D_2O) of **C6**

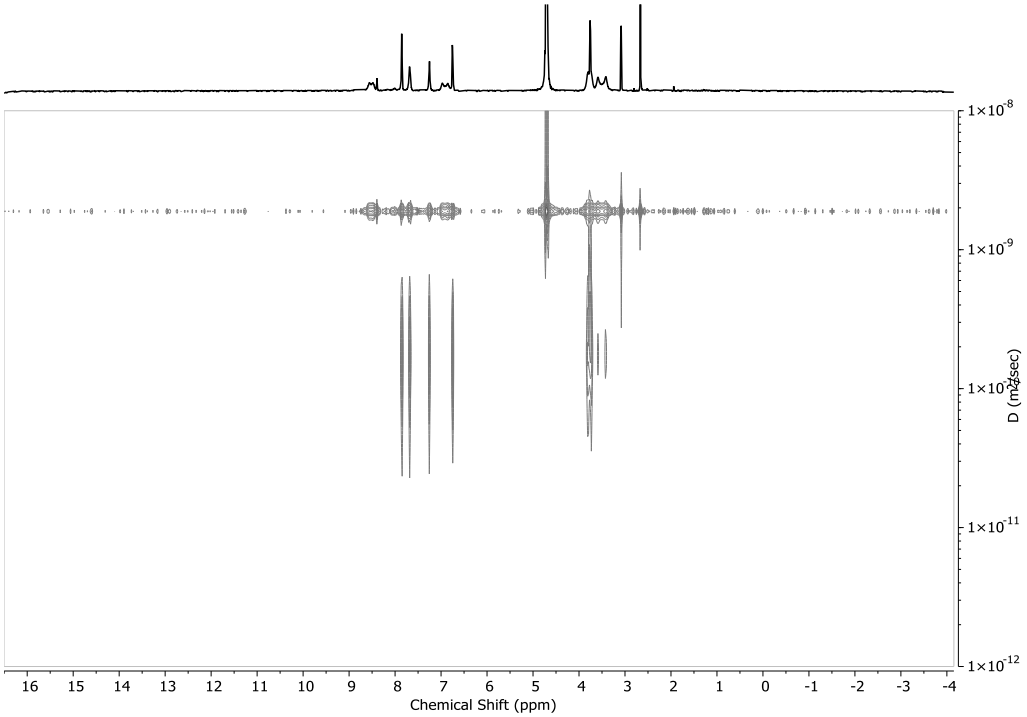


Figure SI 40. ^{1}H DOSY (500 MHz, D_2O) of **C6**

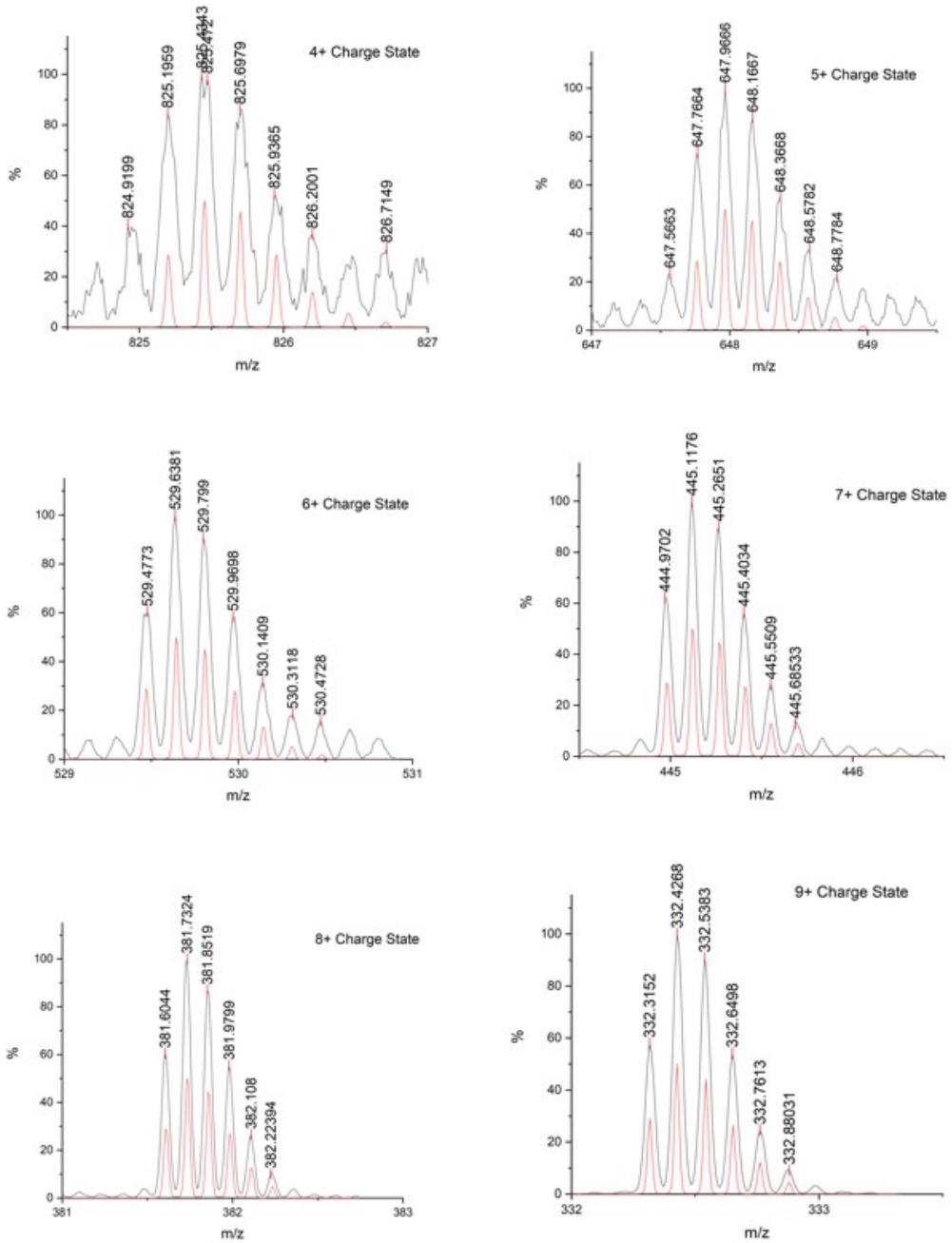


Figure SI 41. ESI Mass Spectrum Charge States for **C6.** Red shows predicted and black shows experimental data.

Characterisation Data – C9

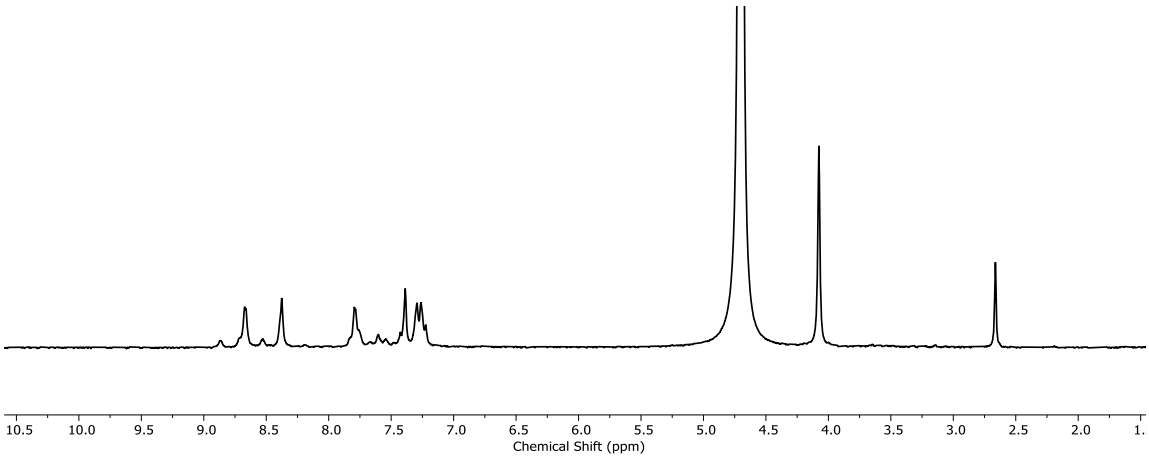


Figure SI 42. ^{1}H NMR (500 MHz, D_2O) of **C9**

Characterisation of 8

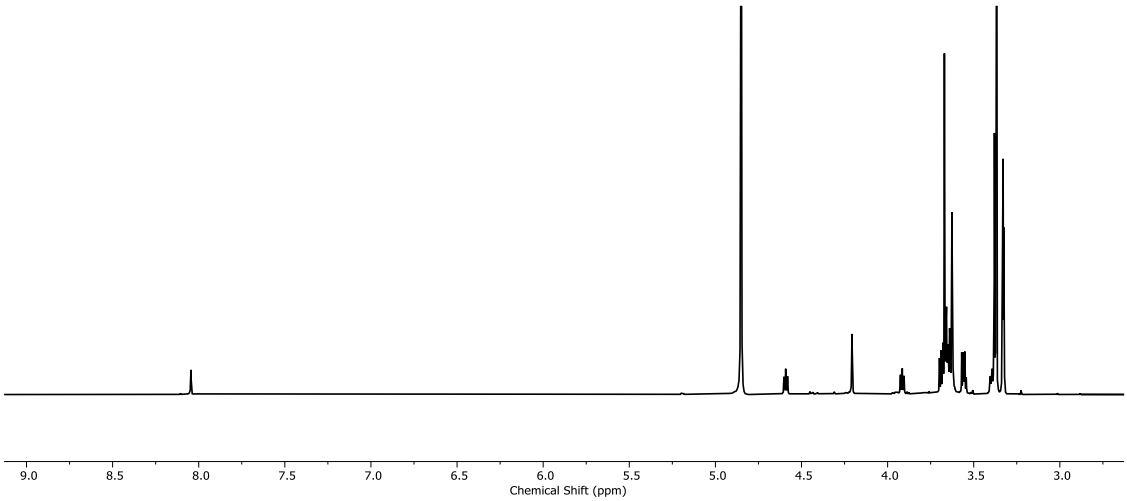


Figure SI 43.^{1}H NMR (500 MHz, MeOD) of **8**

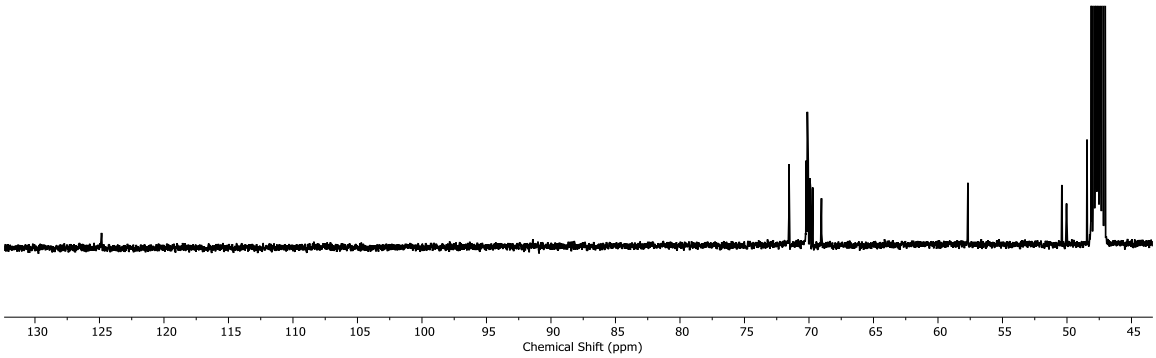


Figure SI 44. ^{13}C NMR (126 MHz, MeOD) of **8**

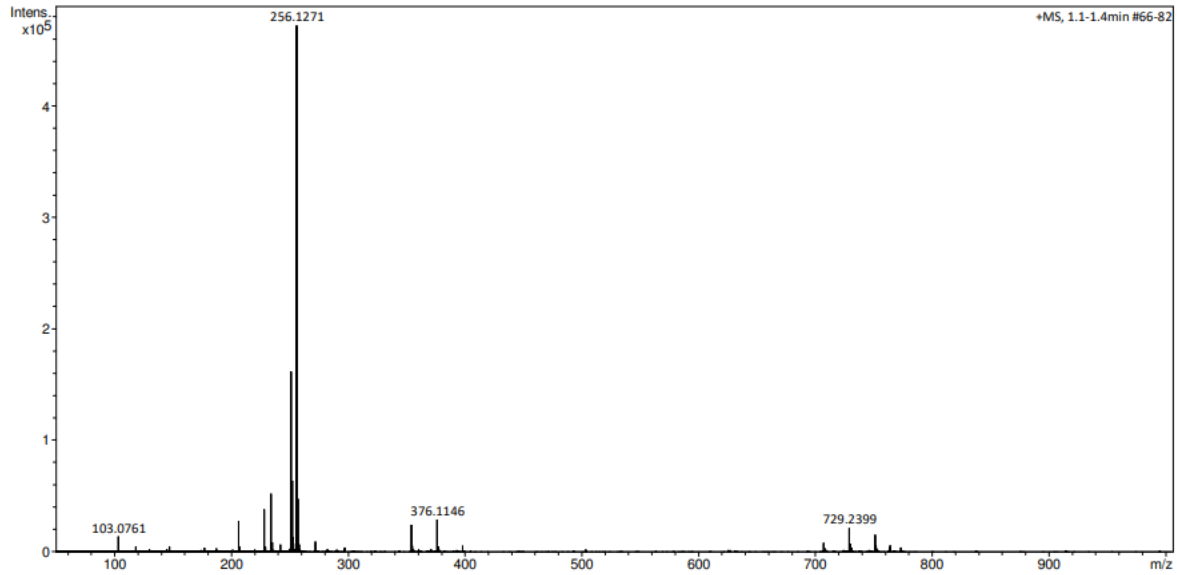


Figure SI 45. ESI Mass Spectrum of **8**

Characterisation of C10

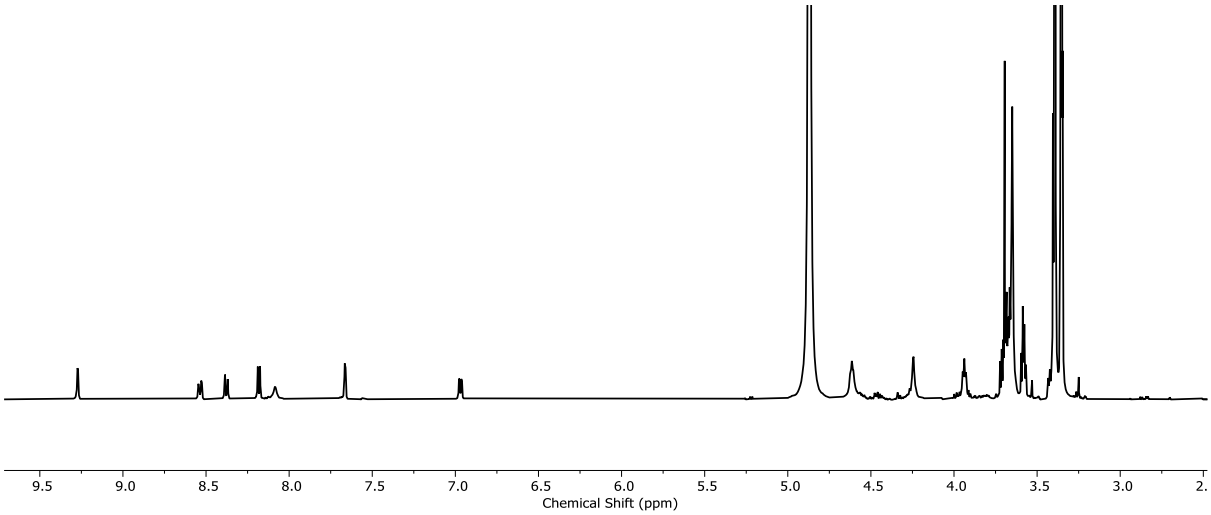


*Figure SI 46. ^{1}H NMR (500 MHz, MeOD) of **C10***

Binding Data for C1

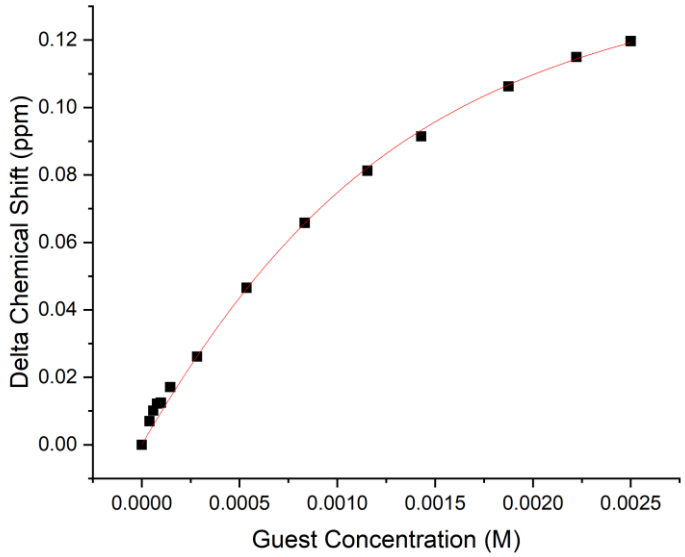


Figure SI 47. Fitted data for change in chemical shift of internal **C1** proton with increasing BF_4^- concentration.

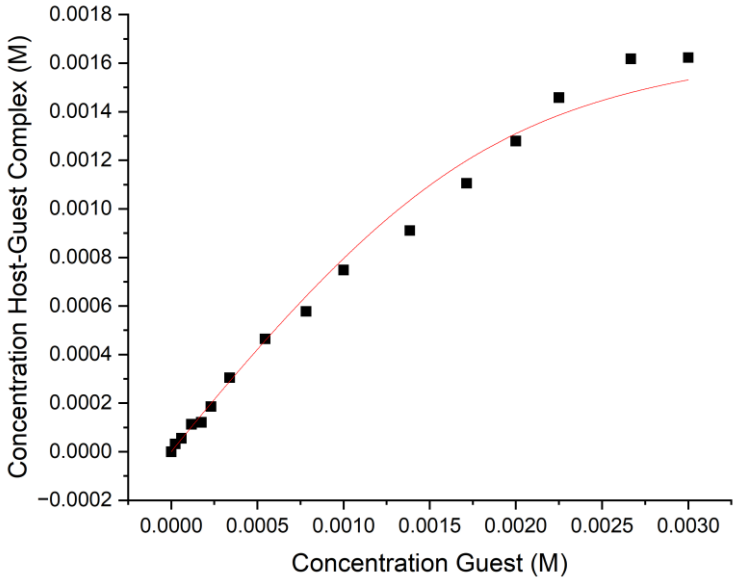


Figure SI 48. Fitted data for change in concentration of host-guest comples **C1**·IO_4^- with increasing IO_4^- concentration.

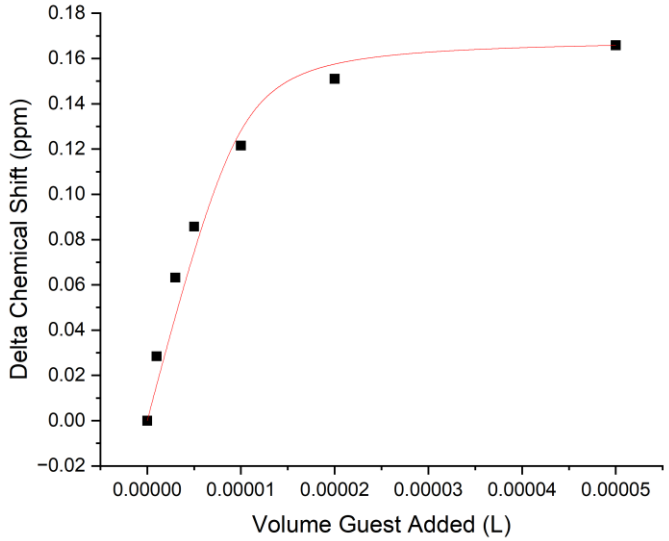


Figure SI 49. Fitted data for change in chemical shift of internal **C1** proton with increasing additions of SO_3F^-.

Binding Data for C6 and C7

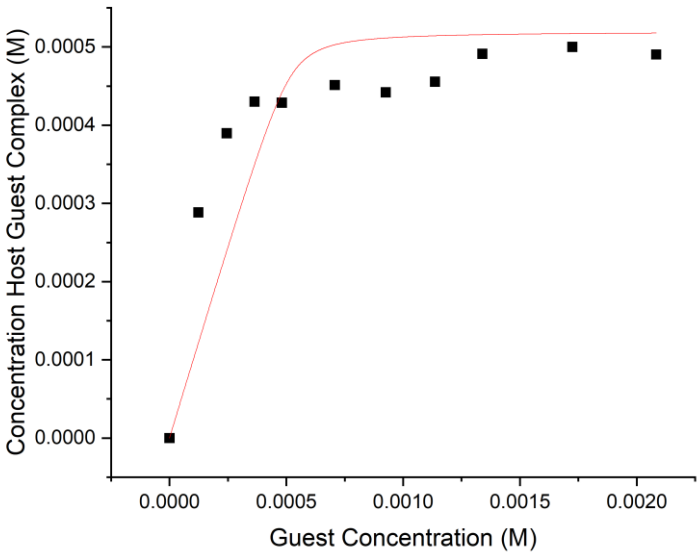


Figure SI 50. Fitted data for change in concentration of host-guest comples **C6**·IO_4^- with increasing IO_4^- concentration

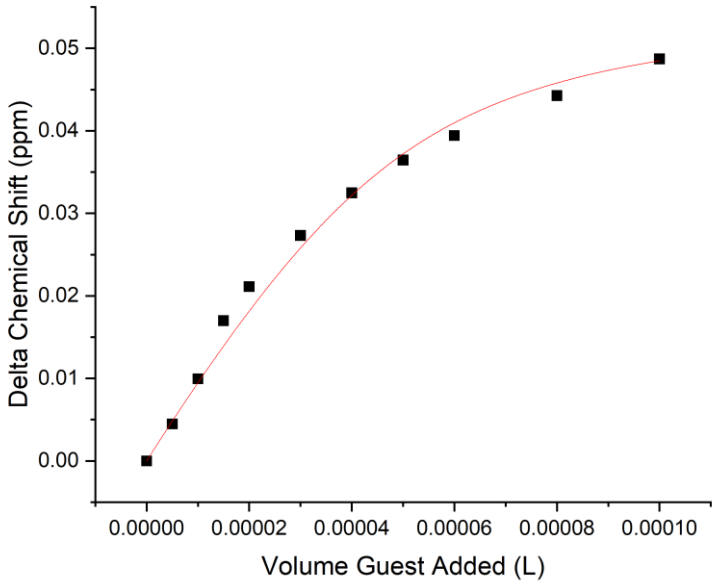


Figure SI 51. Fitted data for change in chemical shift of internal **C6** proton with increasing additions of PF_6^-.

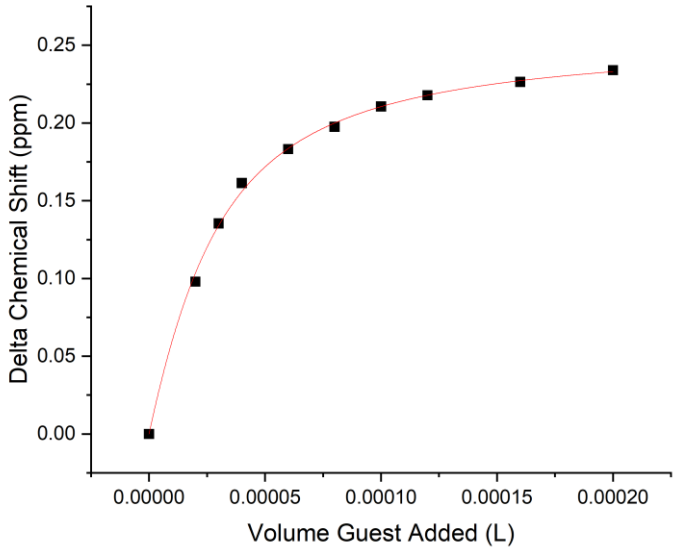


Figure SI 52. Fitted data for change in chemical shift of internal **C6** proton with increasing additions of ReO_4^-.

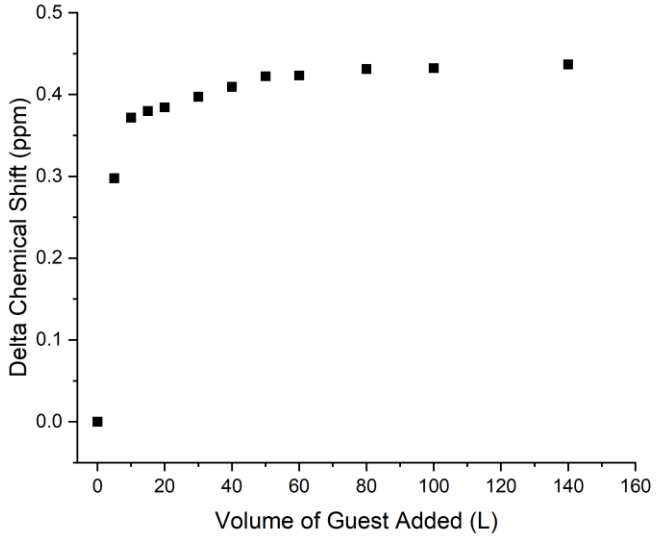


Figure SI 53. Non-fitted data for change in chemical shift of internal **C7** proton with increasing additions of PF_6^-.

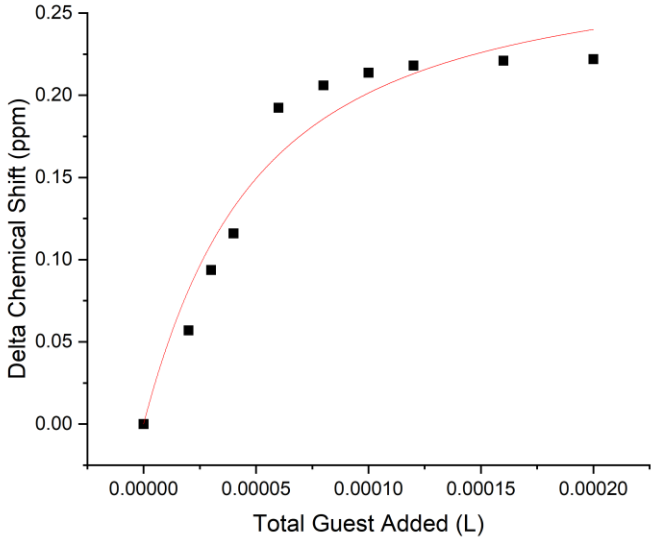


Figure SI 54. Fitted data for change in chemical shift of internal **C7** proton with increasing additions of ReO_4^-.

Characterisation of L14

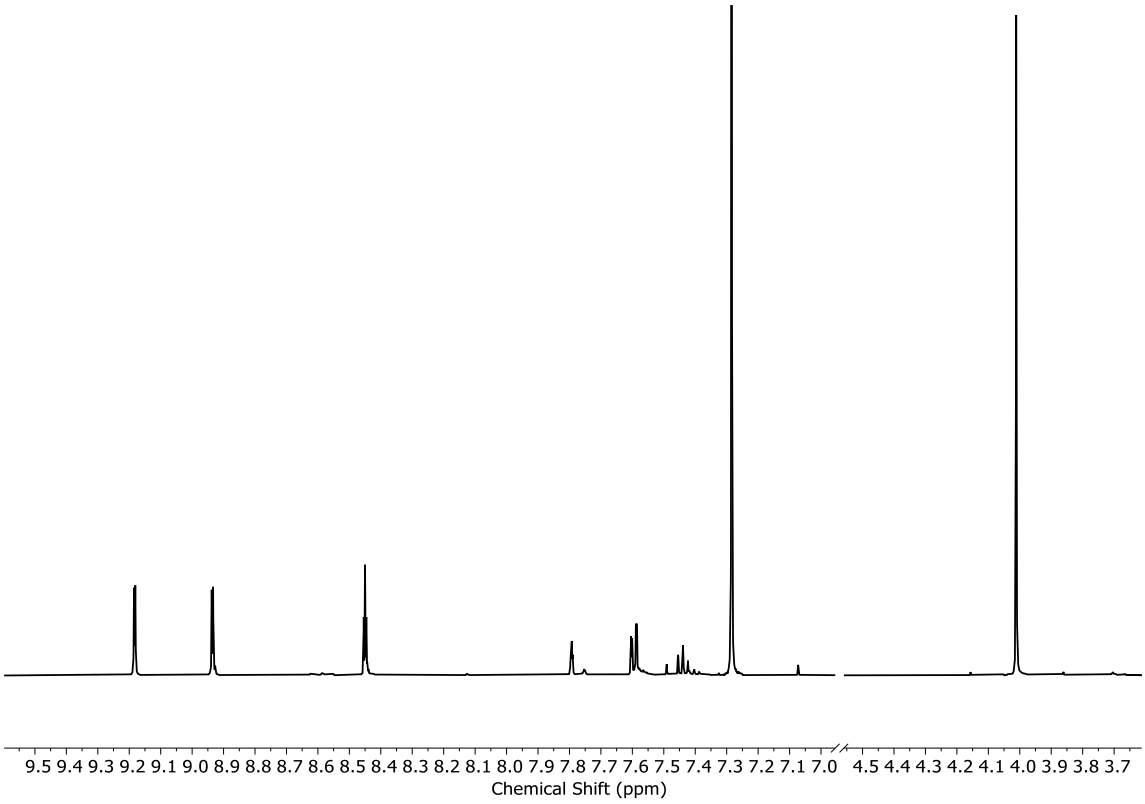


Figure SI 55. ^{1}H NMR (601 MHz, $CDCl_3$) of **L14**

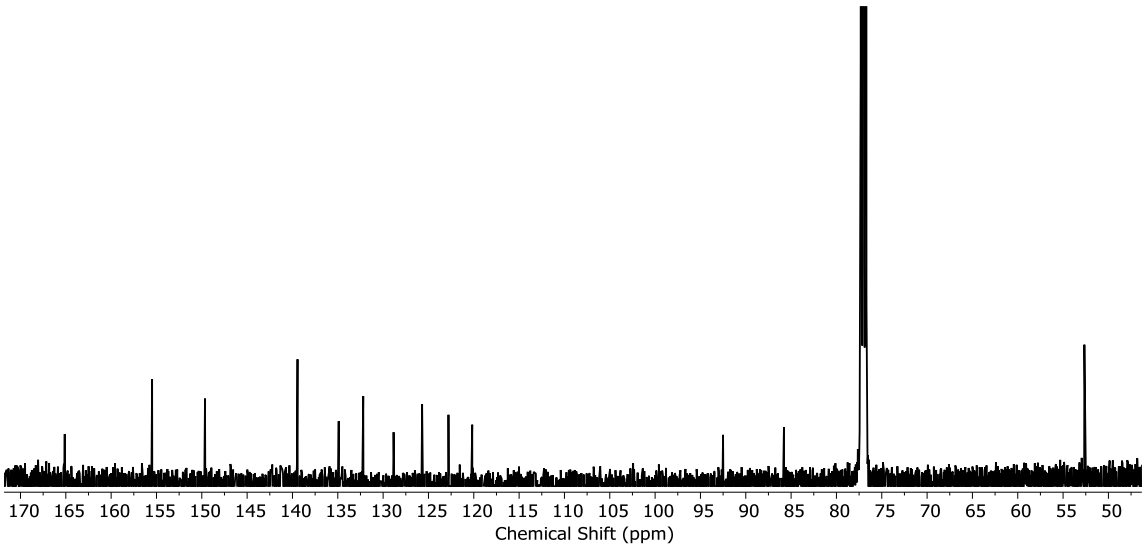


Figure SI 56. ^{13}C NMR (126 MHz, $CDCl_3$) of **L14**

Characterisation of L20

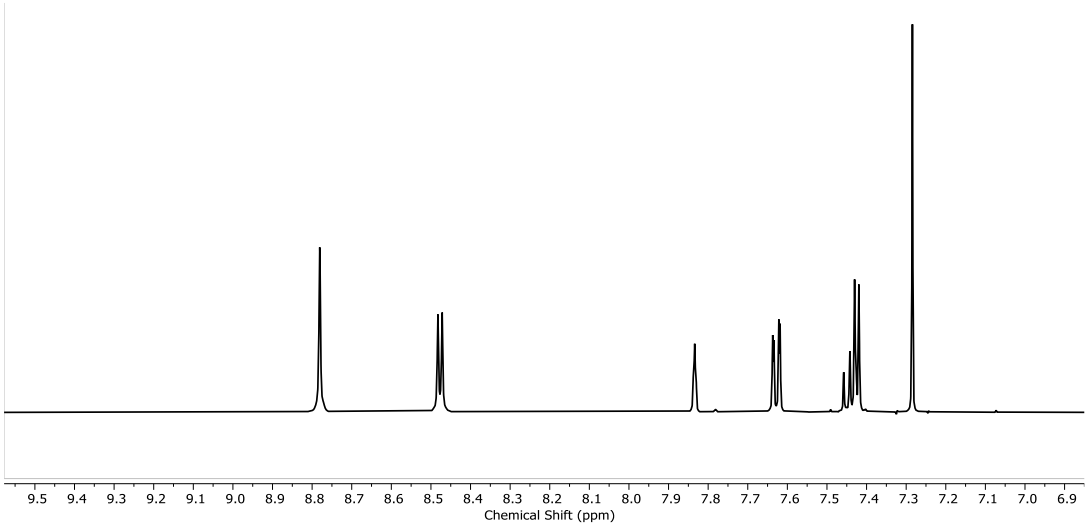


Figure SI 57. ^{1}H NMR (500 MHz, $CDCl_3$) of **L20**

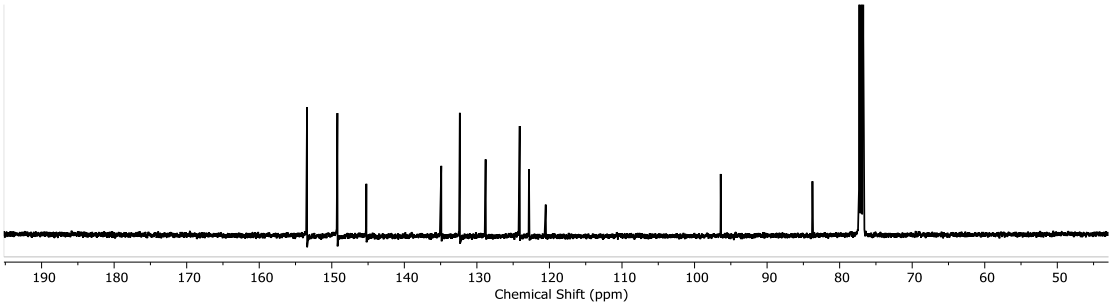


Figure SI 58. ^{13}C NMR (126 MHz, $CDCl_3$) of **L20**

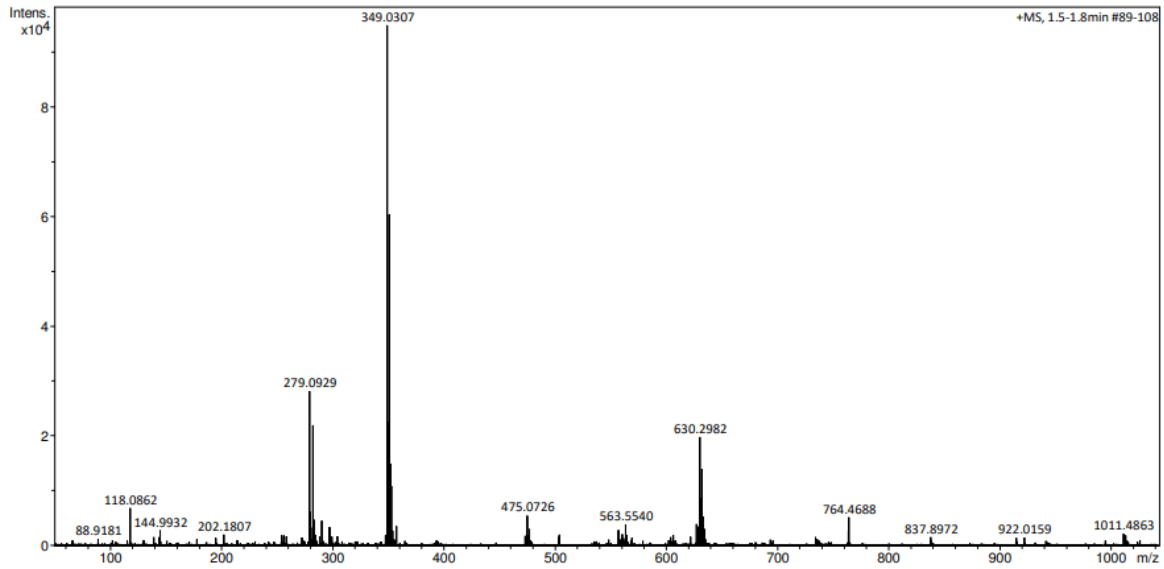


Figure SI 59. ESI Mass Spectrum of **L20**

Characterisation of L17

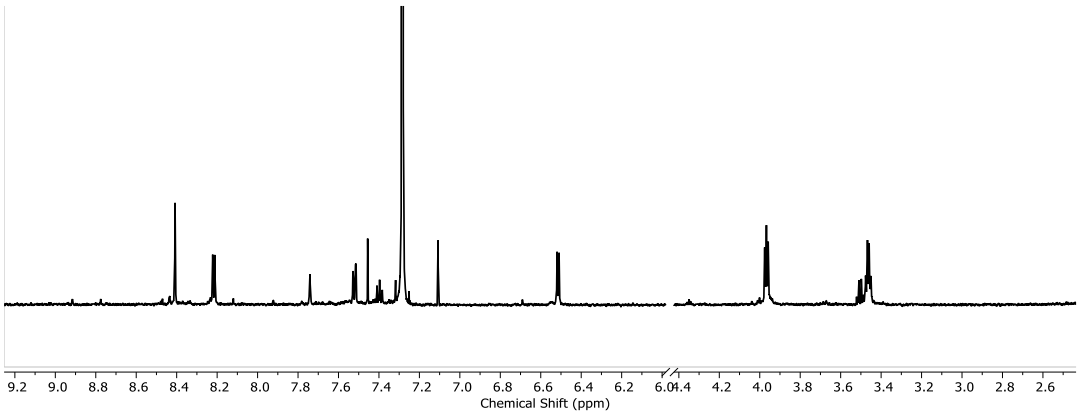


Figure SI 60. ^{1}H NMR (601 MHz, $CDCl_3$) of **L17**

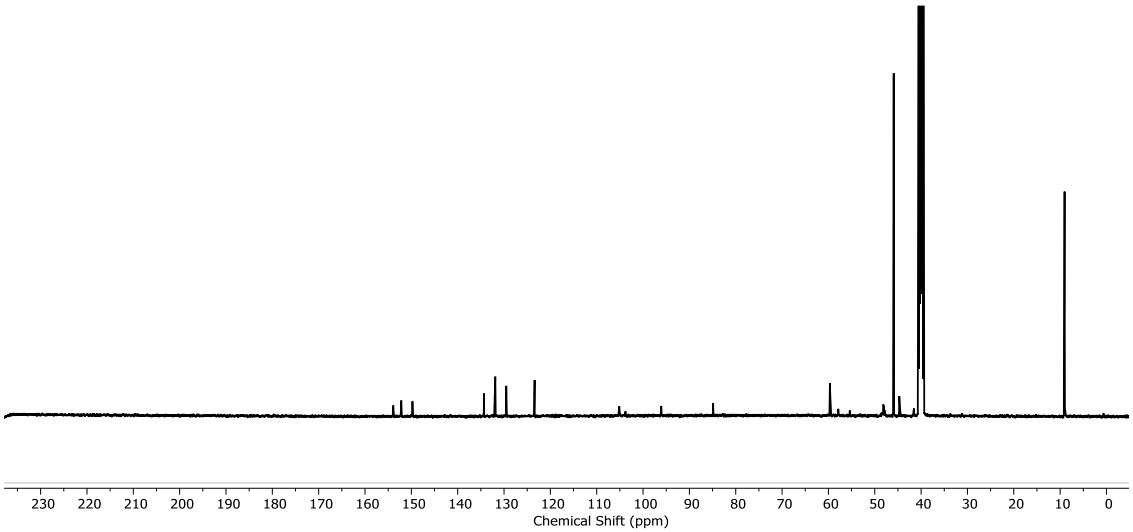


Figure SI 61. ^{13}C NMR (126 MHz, d_6-DMSO) of **L17**

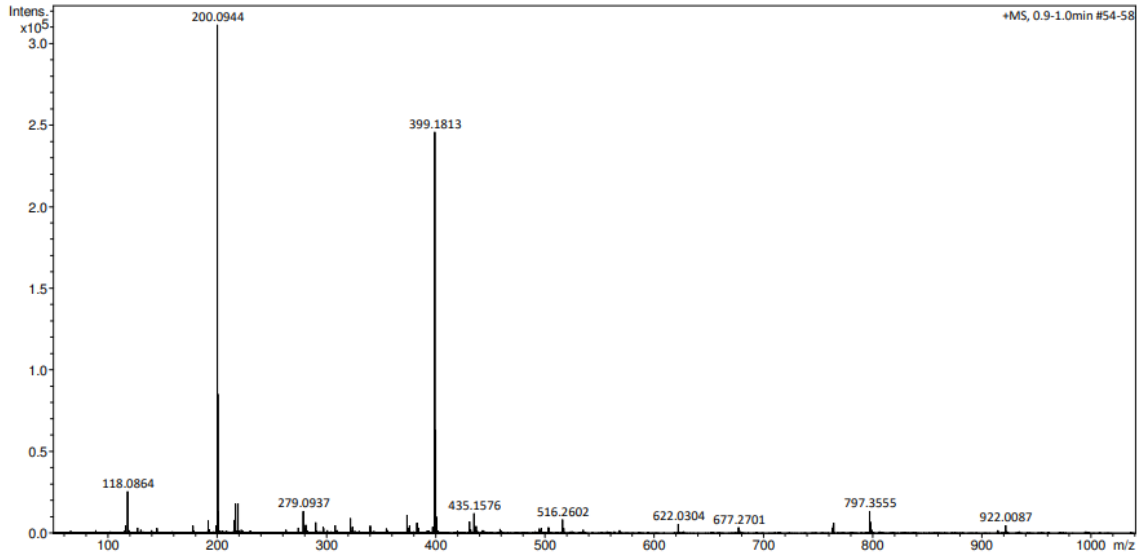


Figure SI 62. ESI Mass Spectrum of **L17**

Characterisation of L18

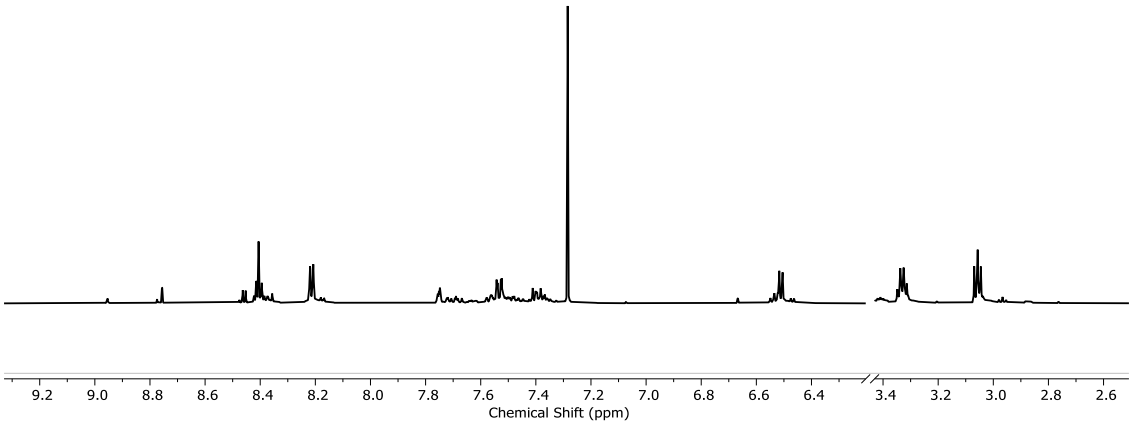


Figure SI 63. ^{1}H NMR (601 MHz, $CDCl_3$) of **L18**

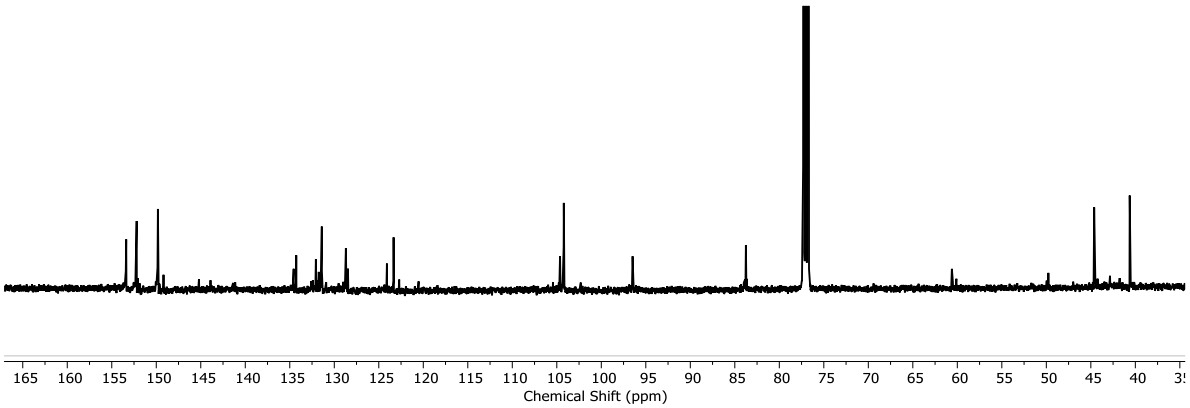


Figure SI 64. ^{13}C NMR (126 MHz, $CDCl_3$) of **L18**

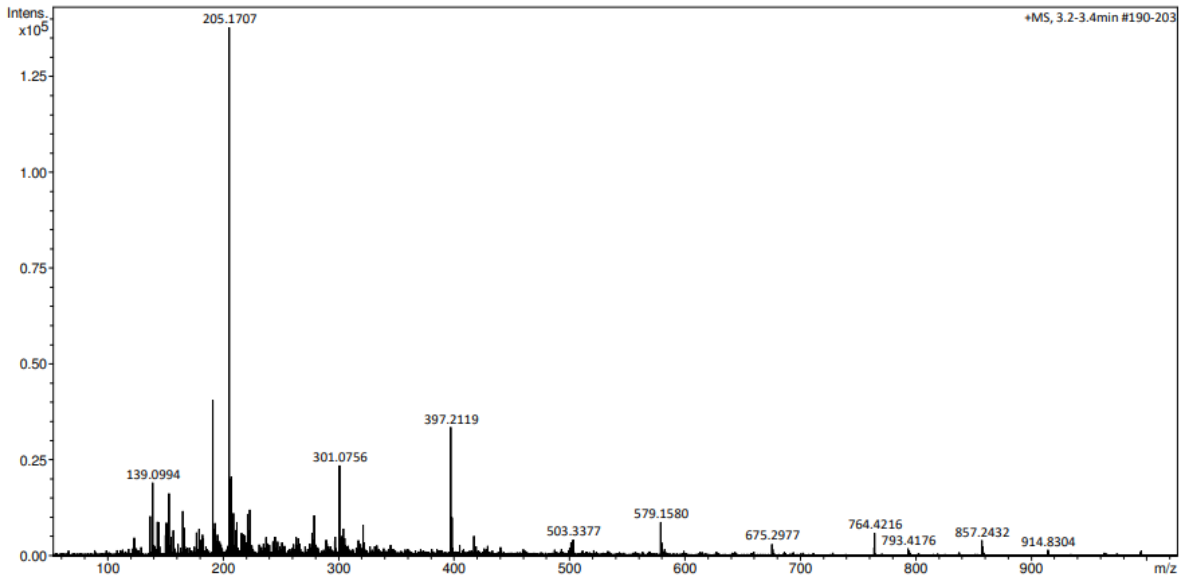


Figure SI 65. ESI Mass Spectrum of **L18**

Characterisation of L19

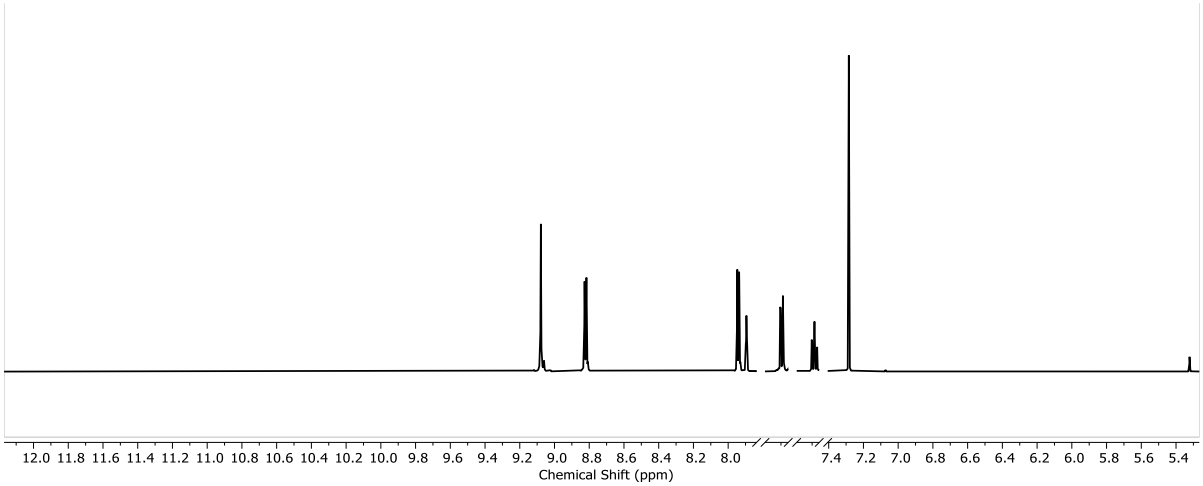


Figure SI 66. ^{1}H NMR (500 MHz, $CDCl_3$) of **L19**

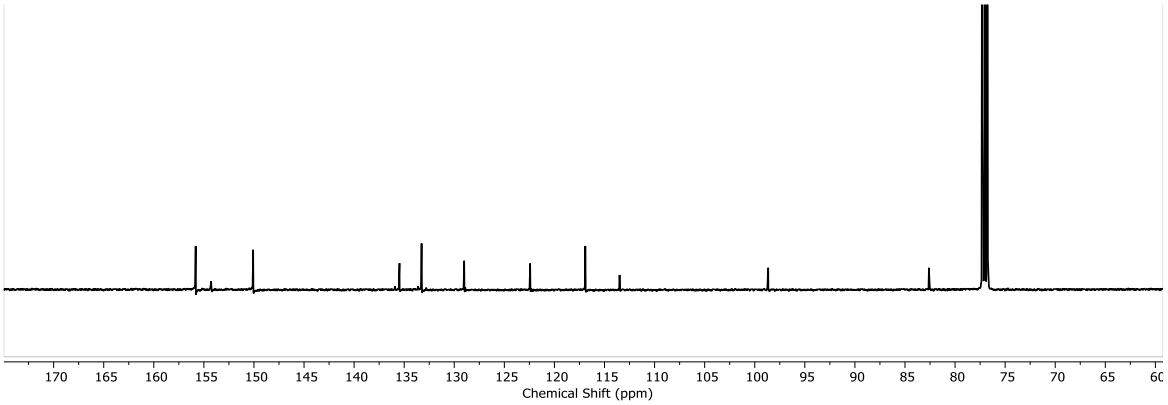


Figure SI 67. ^{13}C NMR (126 MHz, $CDCl_3$) of **L19**

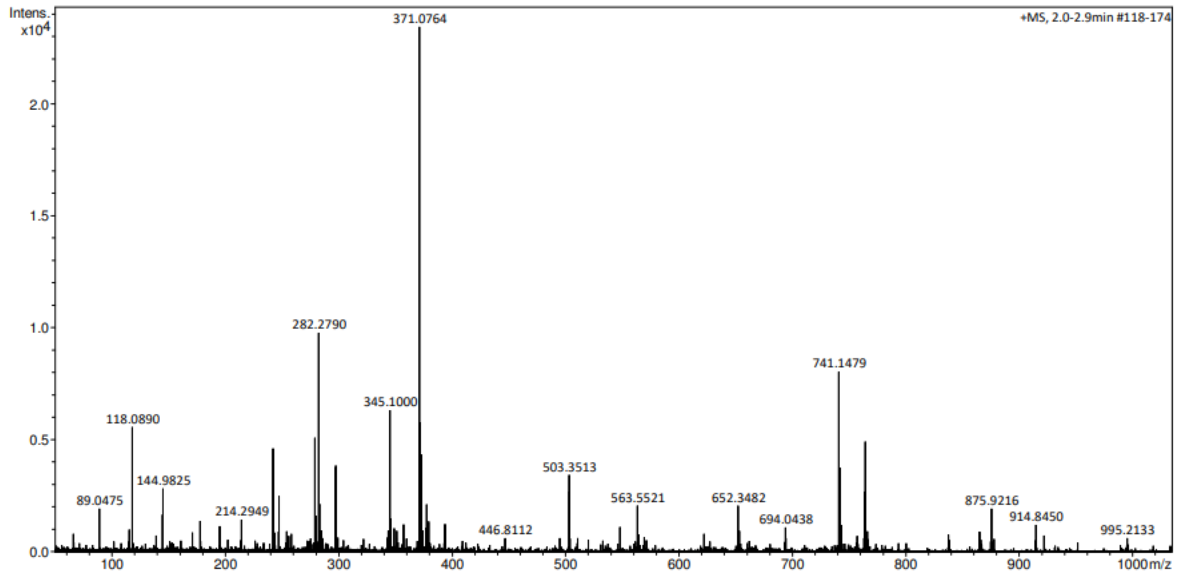


Figure SI 68. ESI Mass Spectrum of **L19**

Characterisation of L16

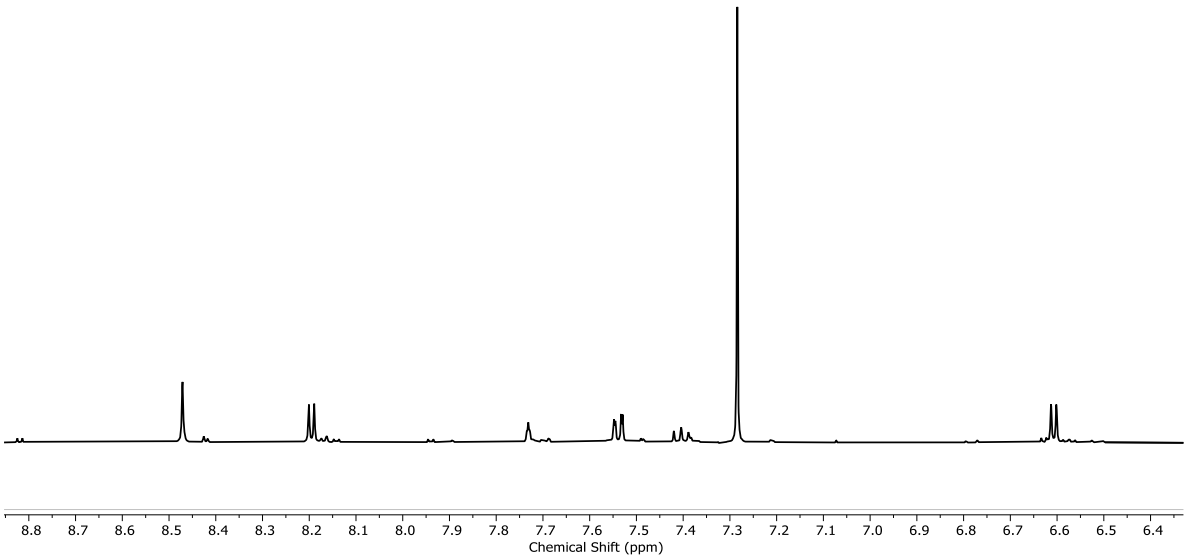


Figure SI 69. ^{1}H NMR (500 MHz, $CDCl_3$) of **L16**

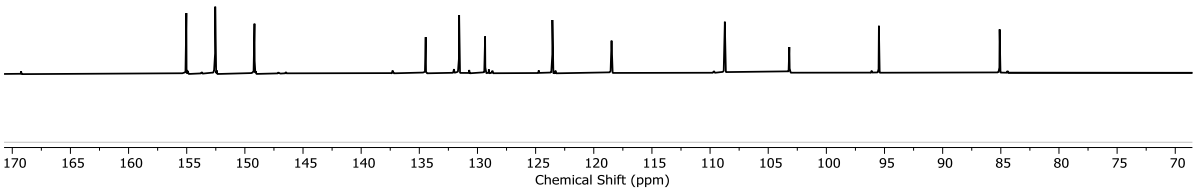


Figure SI 70. ^{13}C NMR (126 MHz, $CDCl_3$) of **L16**

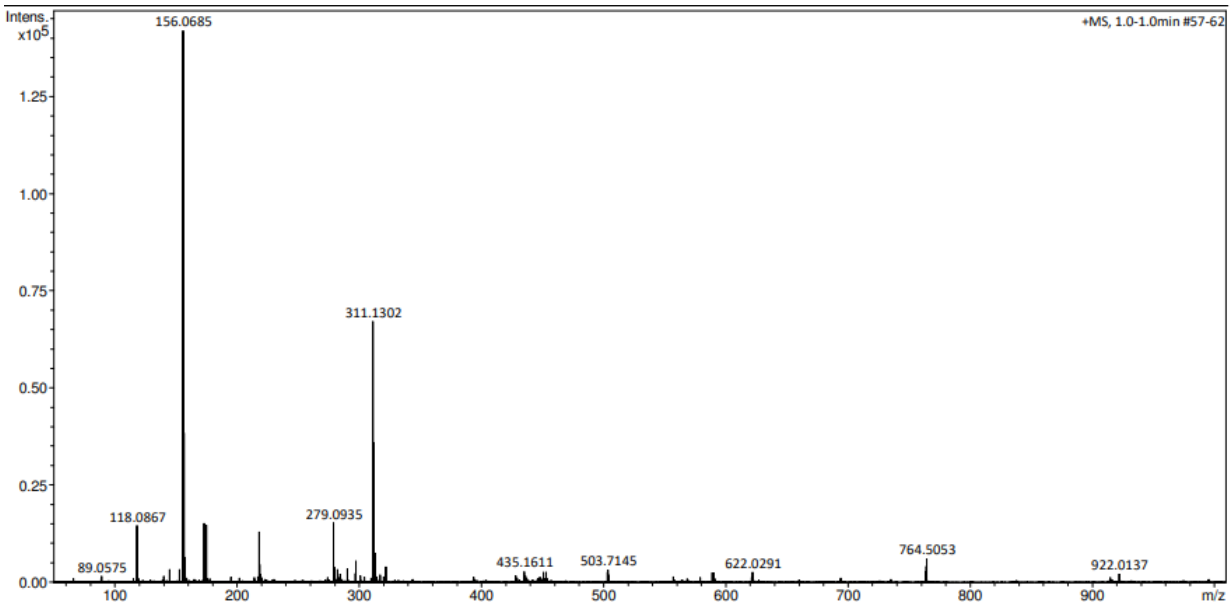


Figure SI 71. ESI Mass Spectrum of **L16**

Characterisation of C13

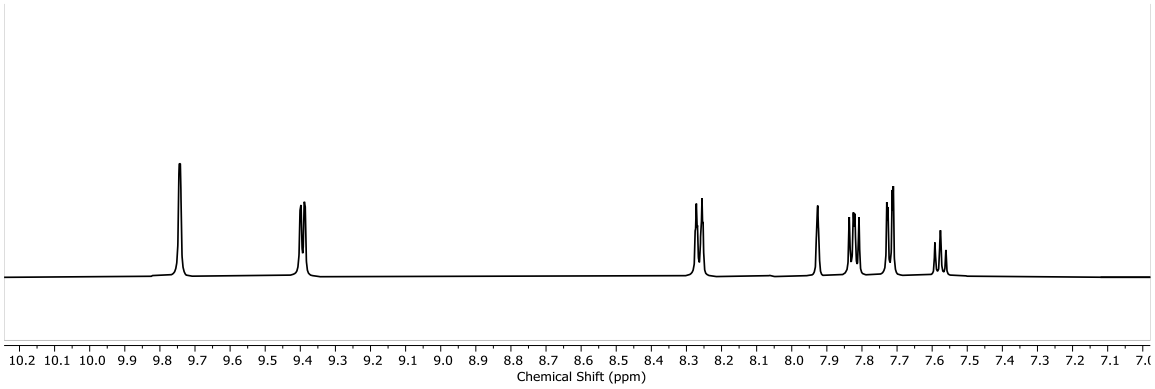


Figure SI 72. ^{1}H NMR (500 MHz, d_6-DMSO) of **C13**

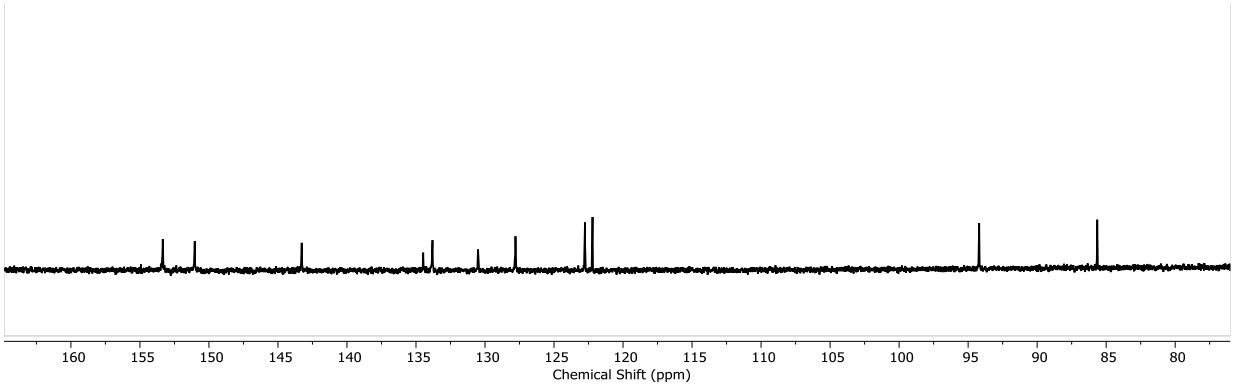


Figure SI 73. ^{13}C NMR (126 MHz, d_6-DMSO) of **C13**

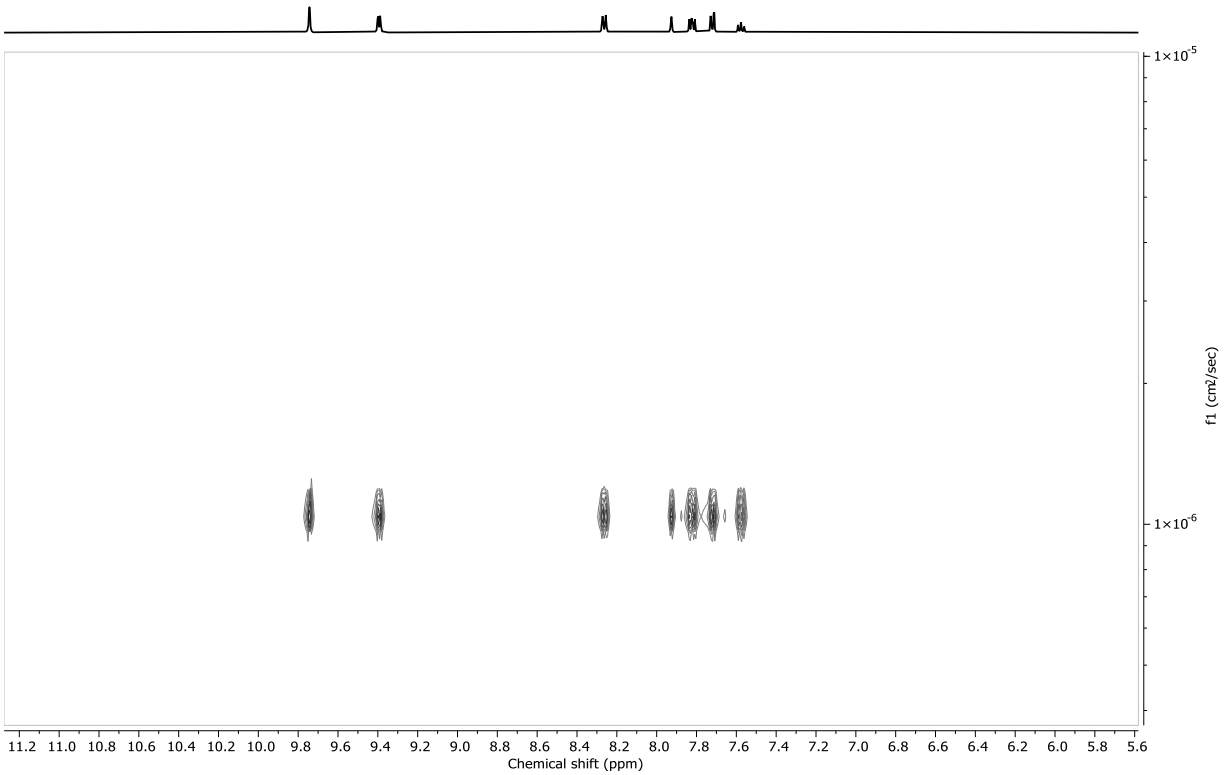


Figure SI 74. ^{1}H DOSY (500 MHz, d_6-DMSO) of **C13**

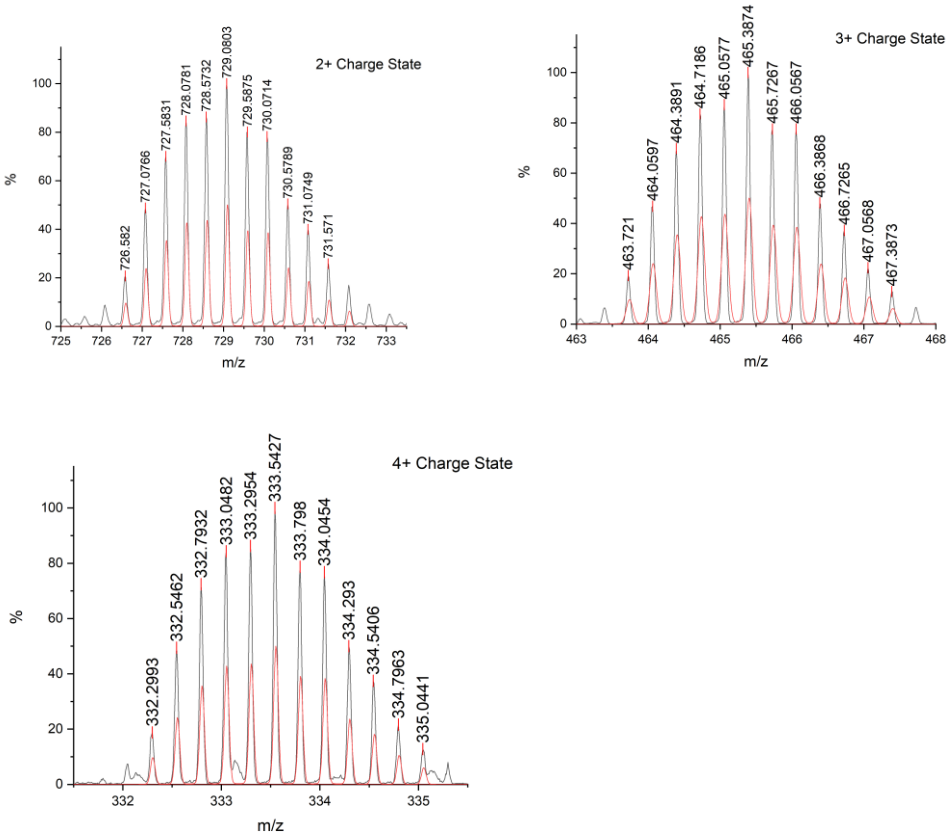


Figure SI 75. ESI Mass Spectrum Charge States for **C13.** Red shows predicted and black shows experimental data.

Characterisation of C12

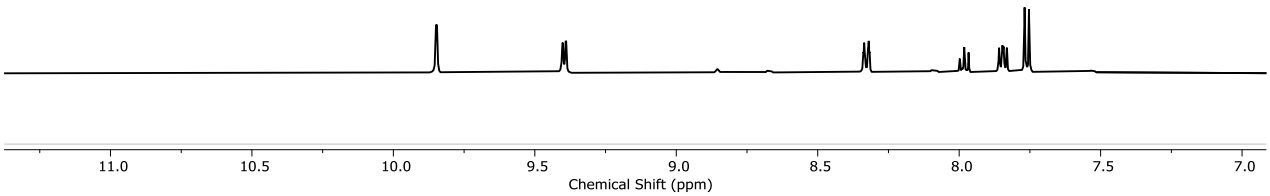


Figure SI 76. ^{1}H NMR (500 MHz, d_6-DMSO) of **C12**

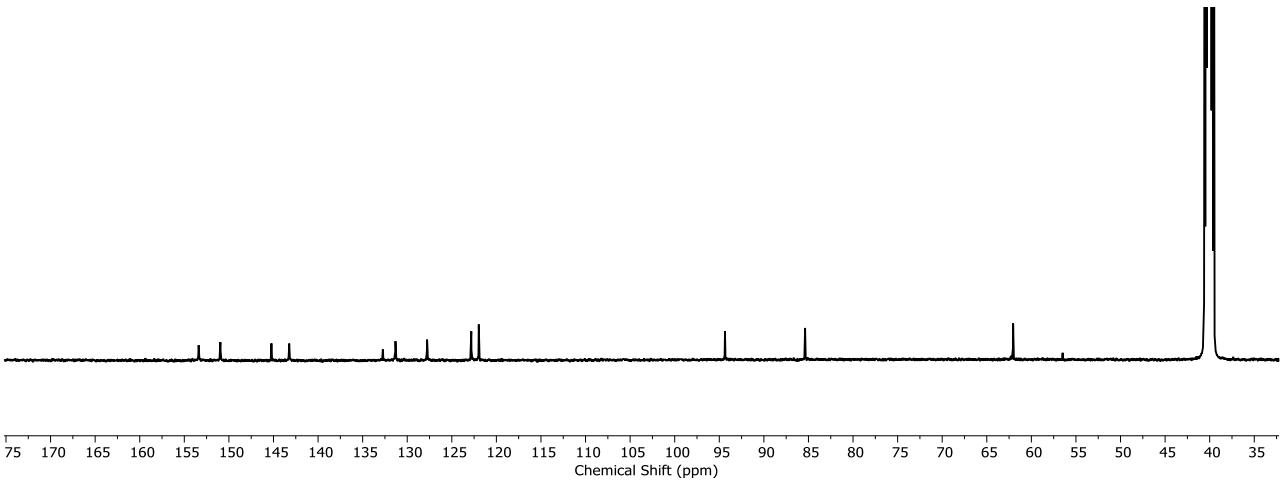


Figure SI 77. ^{13}C NMR (126 MHz, d_6-DMSO) of **C12**

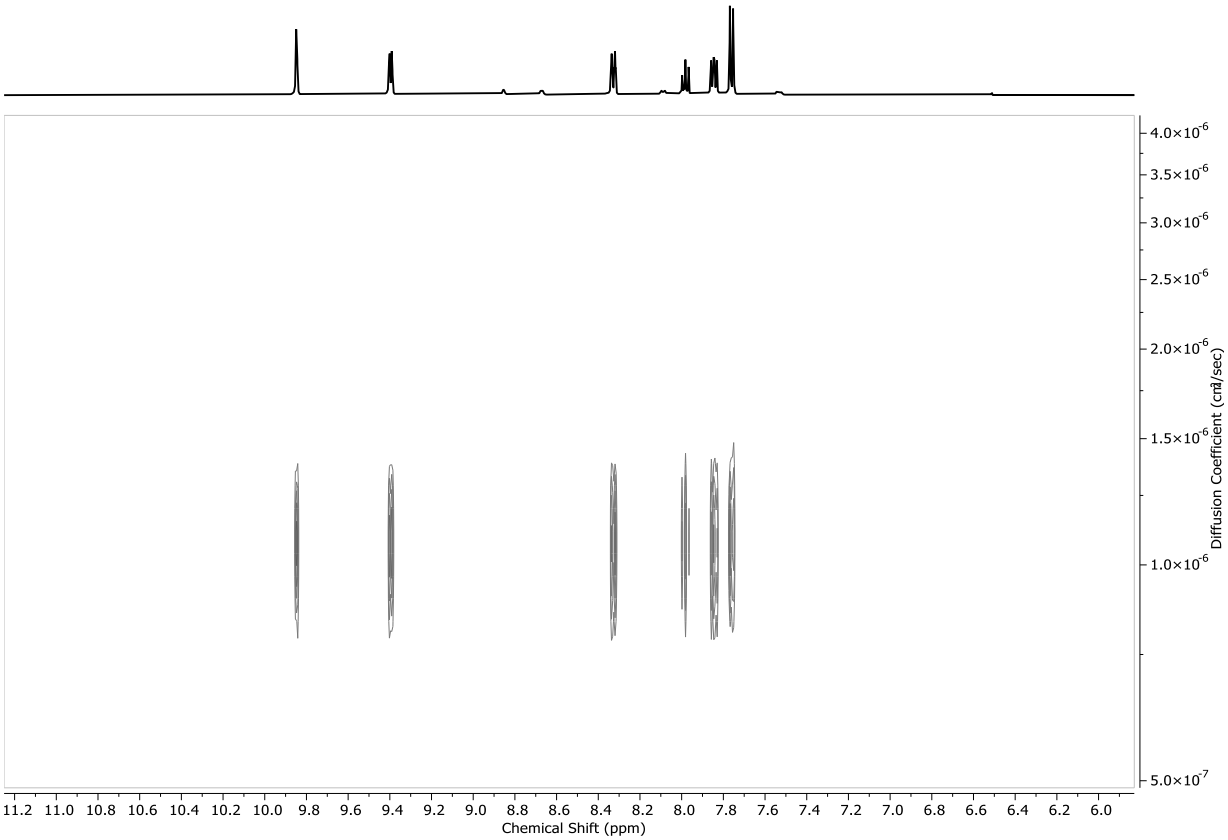


Figure SI 78. ^{1}H DOSY (500 MHz, d_6-DMSO) of **C12**

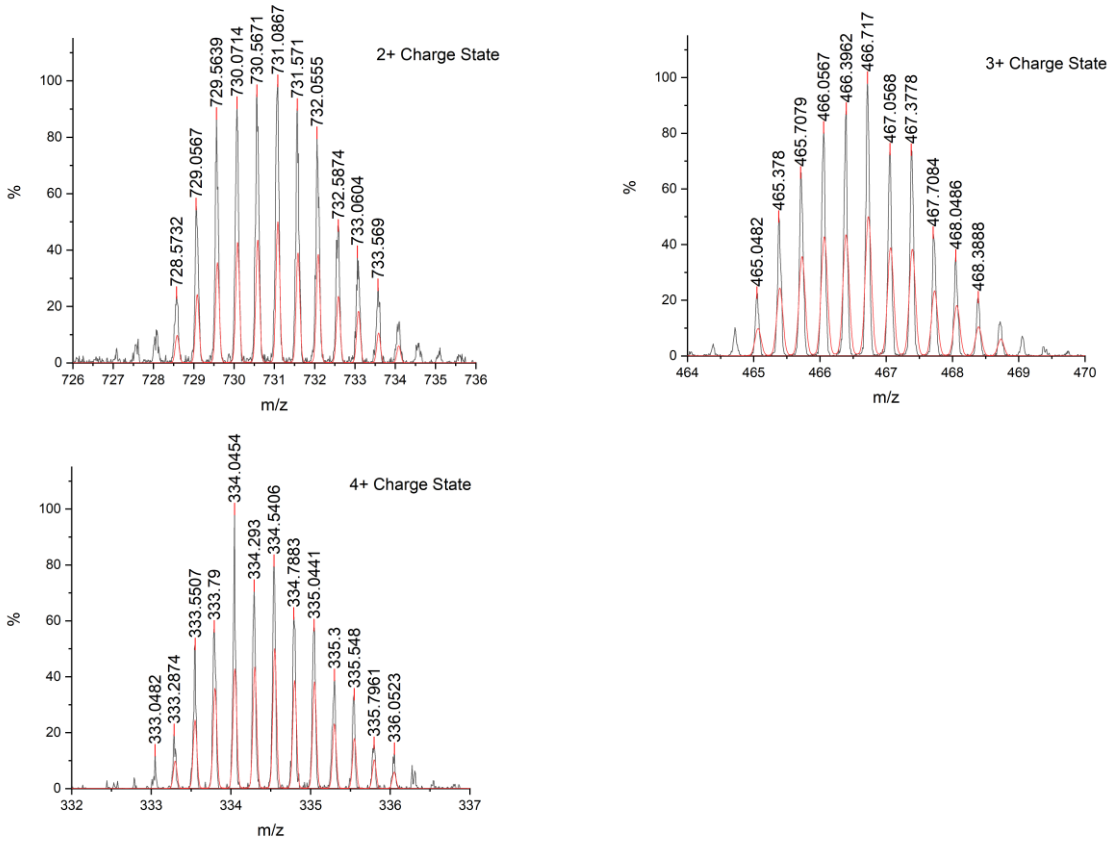


Figure SI 79. ESI Mass Spectrum Charge States for **C12.** Red shows predicted and black shows experimental data.

Characterisation of C14

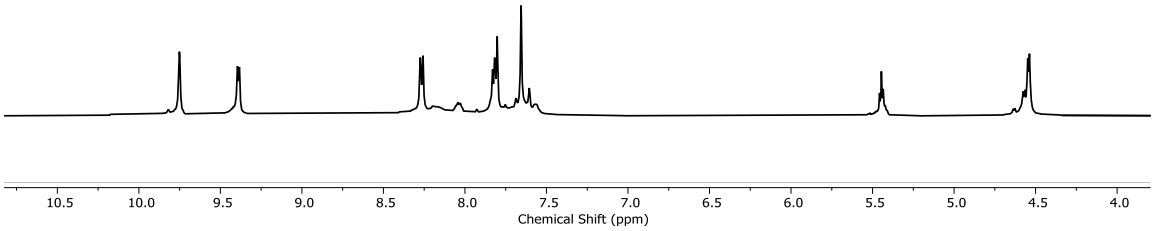


Figure SI 80. ^{1}H NMR (601 MHz, DMSO) of **C14.**

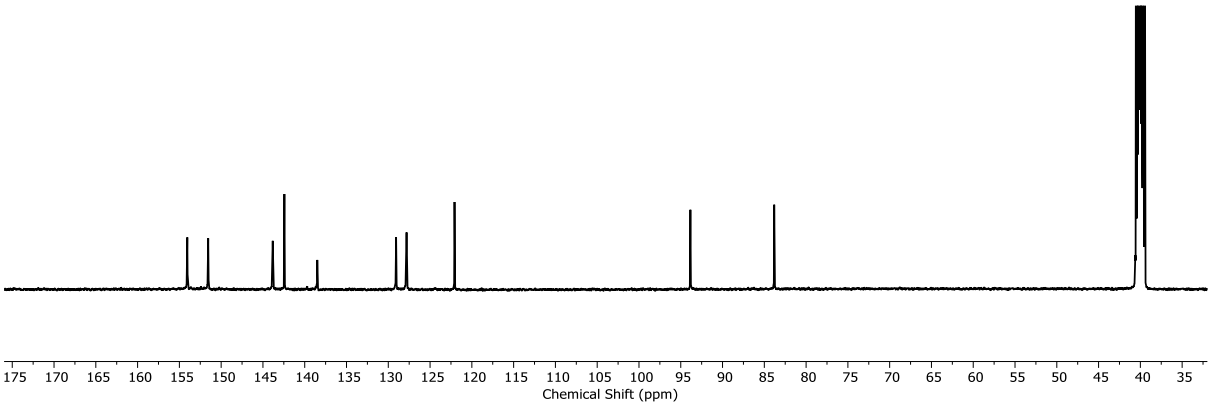


Figure SI 81. ^{13}C NMR (126 MHz, d_6-DMSO) of **C14.**

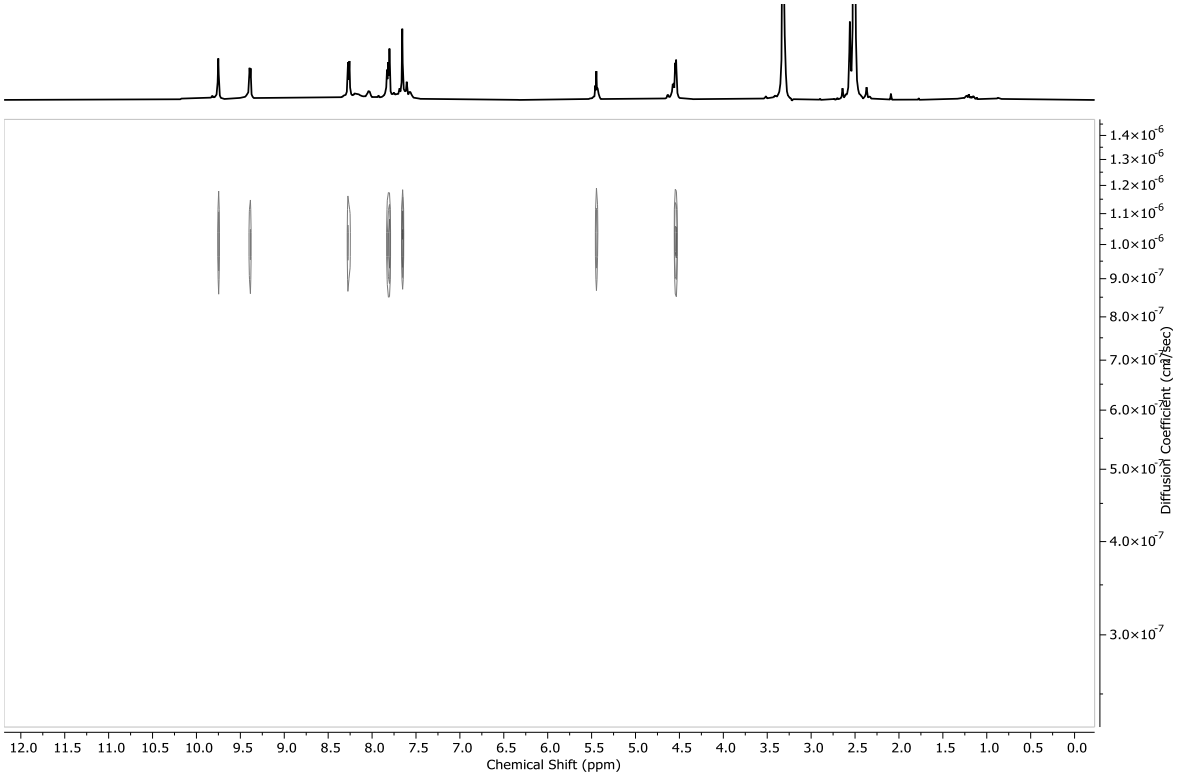


Figure SI 82. ^{1}H DOSY (601 MHz, d_6-DMSO) of **C14**

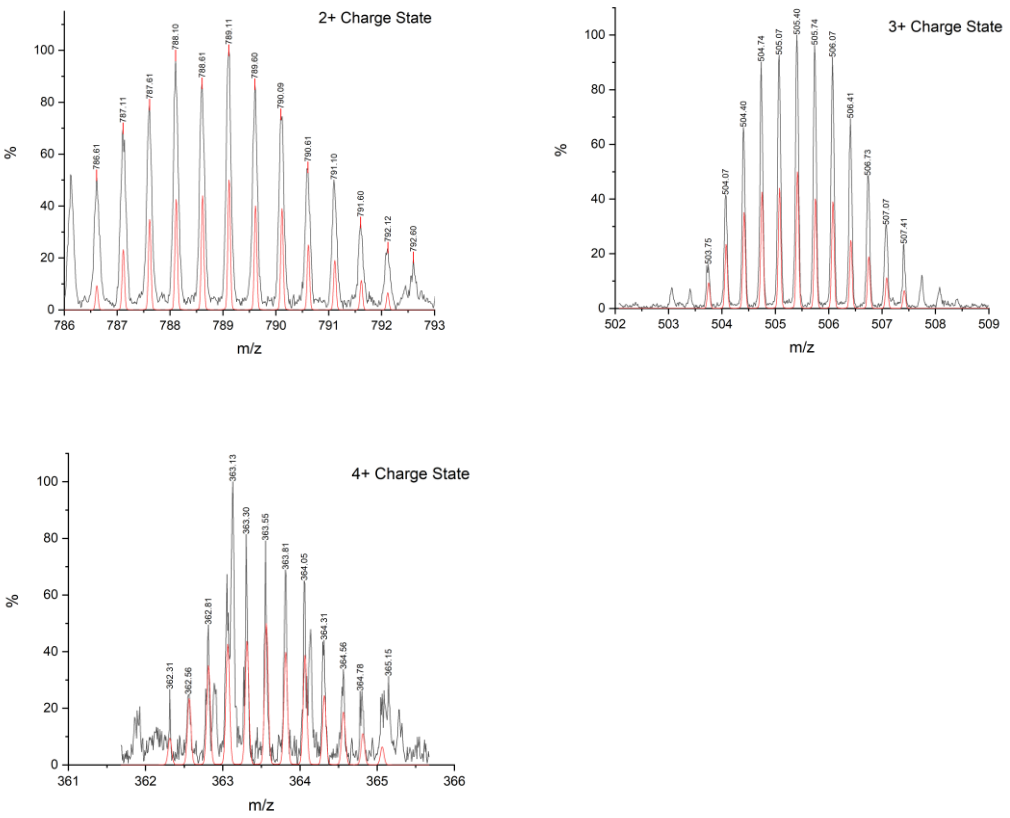


Figure SI 83. ESI Mass Spectrum Charge States for **C14.** Red shows predicted and black shows experimental data.