THE SYNTHESIS AND CHARACTERISATION OF NOVEL SUPRAMOLECULAR STRUCTURES AND THEIR BINDING TO DNA

by

CATHERINE AMELIA JOSEPHINE HOOPER

A thesis submitted to the University of Birmingham for the degree of MASTER OF RESEARCH

Department of Chemistry
College of Engineering and Physical Sciences
University of Birmingham
September 2019

Abstract

The aim of this thesis was to synthesise and characterise the DNA binding of novel structures based on a supramolecular cylinder. These structures are designed to retain the three way junction binding properties of the parent cylinder, while making modifications which have the potential to change DNA binding properties such as binding capacity, selectivity and control.

The first chapter provides an introduction to non-canonical DNA structures and the ligands which have been developed to bind them. In chapter 2, the synthesis of new supramolecular cylinders, a trinuclear helicate and an asymmetric helicate, is investigated. A method for post assembly modification of a cylinder analogue is developed in chapter 3, which is then used to synthesise the first rotaxane from a supramolecular helicate via a capping mechanism. These structures are analysed by mass spectrometry and NMR. Finally, the DNA binding of a cylinder analogue and the rotaxane is investigated using UV Visible spectroscopy, circular dichroism and linear dichroism in chapter 4.

Acknoledgements

Firstly I would like to thank Professor Hannon for giving me the opportunity to work as a member of such a fantastic research group. His advice and constant support has been thoroughly appreciated throughout the completion of my degree.

The Hannon Group have taught me everything I know about working in a research environment. It has been a pleasure to work with a group of people so ready to help and support each other. A huge thank you to them all for being such brilliant role models and having such a positive impact on my experience. Also to Dayra, who has always been so willing to help in any way she can, even from a different continent. I count myself as truly lucky.

Thank you to Zoe Pikramenou and her research group for their helpful feedback and suggestions, especially during group meetings.

Finally, thank you to my family and friends, who have picked me up and kept me going on so many occasions. A special thanks to the Cheaney family for donating their office while I wrote this thesis, and my parents and sister, who have endured every high and low of this experience with me. This thesis is dedicated to you all.

List of Abbreviations

(prep/RP)HPLC	(Preparatory/Reverse Phase) High Pressure Liquid Chromatography
3H(Sn)	Three Way Junction (with n unpaired bases at the junction)
3WJ	Three Way Junction
4WJ	Four Way Junction
5'-UTR	Untranslated Region, at the 5' end of a DNA strand.
A	Adenine, DNA base
AFM	Atomic Force Microscopy
C	Cytosine, DNA base
CBn	Cucubit[n]uril
CD	Circular Dichroism
COSY	Correlated Spectroscopy
ctDNA	Calf Thymus DNA
DCM	Dichloromethane
DIPEA	*N,N*-Diisopropylethylamine, Hunig's base
DMF	Dimethyl Formamide
DMS	Dimethyl Suphide

DMSO	Dimethyl Sulphoxide
DNA	Deoxyribonucleuic Acid, preceeded by alternative letters to indicate a specific structure (commonly A, B or Z). Can also be preceeded by ss (single stranded) or ds (double stranded).
DOSY	Diffusion Ordered Spectroscopy
ES	Electrospray Ionisation
G	Guanine, DNA base
GQ	G-Quadruplex
HMBC	Heteronuclear Multiple Bond Correlation
HSQC	Heteronuclear single quantum coherence
i-motif	Intercalated motif
ICD	Induced Circular Dichroism
ILD	Induced Linear Dichroism
LCMS	Liquid Chromotography Mass Spectrometry
LD^r	Reduced LD Signal
mCPBA	meta-Chloroperoxybenzoic Acid
MEC	Molar Extinction Coefficient
MLCT	Metal Ligand Charge Transfer

N	(Within a base sequence) refers to an unspecified base.
NMR	Nuclear Magnetic Resonance
NOESY	Nuclear Overhauser Effect Spectroscopy
RNA	Ribonucleic Acid, preceeded by a letter to indicate a specfic structure and function (commonly m, messenger, or t, transfer).
ROESY	Rotating Overhauser Effect Spectroscopy
SWNTs	Single Walled Carbon Nanotubes
T	Thymine, DNA base
$TMPyP_x$	Tetra-(N-methyl-x-pyridyl)porphine, where x=2 or 4
TOF MS EI	Time of Flight Mass Spectrometry Electron Ionisation
UV Vis	Ultra-Violet Visible
VdW	Van der Waals

Ligands

L	Parent Ligand (C25H20N4)
Lasym	Asymmetric Ligand (C22H118N6)
Lhalf	Half Ligand (C18H16N4)
Lim2	2-Imidazole Ligand (C21H18N6)
Lim4	4-Imidazole Ligand (C21H18N6)
Lmod	Modified (Alkylated) Ligand (C32H28N8)
Ltri	Trinuclear Ligand (C42H36N6)

Contents

Chapter 1:

BINDING NON-CANONICAL STRUCTURES OF DNA

1.1. Introduction

Double helical B-DNA is the prominent form of deoxyribonucleic acid (DNA) in our cells and so is by far the most thoroughly studied. It is important to remember that DNA is a dynamic molecule with a plethora of possible structures, some of which have been well characterised while others we are still debating the existence of in vivo. Sequences with the ability to form non-canonical structures of DNA have been found throughout the human genome[1,2] and are often found at specific functional sites such as telomeres[3,4] and gene promotor regions[5,6], suggesting a possible regulatory role in cells. By developing novel DNA-binding molecules that target specific structures of DNA, we may be able to manipulate the regulation and expression of DNA without having to enforce sequence specificity.

This review aims to survey examples of these non-canonical structures and some of the molecules that have been developed to bind to them.

1.2. Helical DNA

1.2.1. Structure of B-DNA

DNA is known for its double stranded, helical structure (see Figure 2b). The backbones are formed from deoxyribose sugars linked by phosphodiester groups and run antiparallel to generate the right handed helix[7,8]. Bound to the deoxyribose moieties are nitrogen bases which lie perpendicular to the helical axis and hydrogen bond to the bases of the adjacent backbone to hold the chains together. The four possible nitrogen bases (adenine, A, cytosine, C, guanine, G, and thymine, T) are split into two groups. Larger purine bases (A, G) have two aromatic rings while the smaller pyrimidine bases (C and

T) have only one aromatic ring. In their keto tautomeric form each base has a specific pattern of hydrogen acceptors and donors which is complimentary to one other base (A to T, C to G). Consequently, Watson-Crick base pairing always binds a purine to a pyrimidine base so the duplex is maintained at a relatively constant width, with minor changes dependent on the exact base sequence. The base pairs hold the backbones at an angle forming a wide major groove and a narrow minor groove[9].

Watson-Crick binding is not the only possible hydrogen bonding between bases, although it is the most common in DNA. Hoogsteen base pairing (see Figure 1) uses a different side of the bases to the Watson Crick edge, allowing bonds to be formed between two bases which are not necessarily purine and pyrimidine[10]. While this base pairing is not used in B-DNA, it is extremely relevant in non-canonical structures, as will be discussed later.

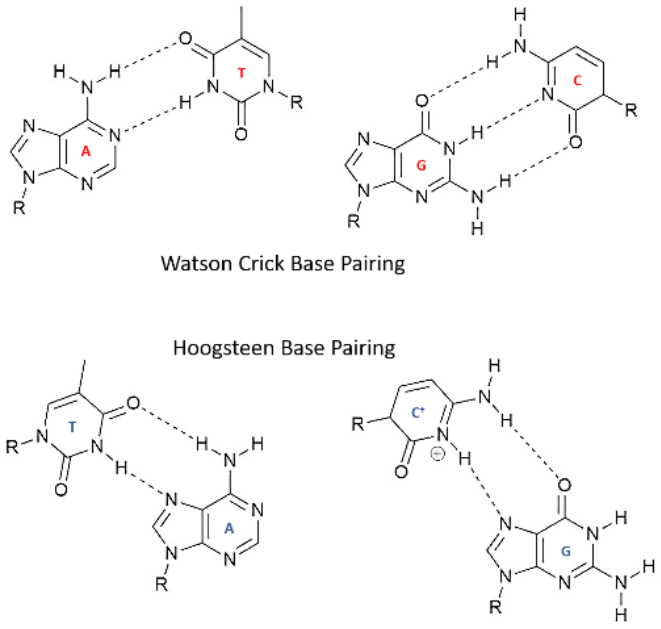


Figure 1: Watson-Crick base pairing (top) and Hoogsteen base pairing (bottom). GC^+ Hoogsteen base pairing requires a protonated cytosine. R is a deoxyribose sugar of the DNA backbone.

1.2.2. **Other Helical DNA Structures**

Alternative helical structures are less common in the nucleus. A-DNA is another right handed double helix with a smaller pitch than B-DNA and 11 base pairs per helical turn (compared to 10 base pairs in B-DNA). To preserve the Van der Waals separation in the more condensed A-DNA, the bases are tilted by 20^{o} [9]. Unlike B-DNA, where all the bases lie on the helical axis, the bases are displaced towards the minor groove which creates a 3 Å central cavity, decreases the depth of the minor groove, and increases the depth of the major groove (see Figure 2a). A-DNA is formed by ribonucleic acids (RNA) and by DNA at low humidity[11].

Z-DNA is a left-handed helix and so differs dramatically from A-DNA and B-DNA[9]. The backbone forms a characteristic zigzag as alternating sugar moieties have a C2' and C3' endo pucker while the glycosidic bonds take a syn or anti conformation. The minor groove is deep enough to contain the helical axis, forcing the major groove to take a convex shape which exposes the bases to the environment. Z-DNA has an extended helical pitch compared to B-DNA with 12 base pairs per helical turn (see Figure 2c) and is stabilised by high salt conditions.

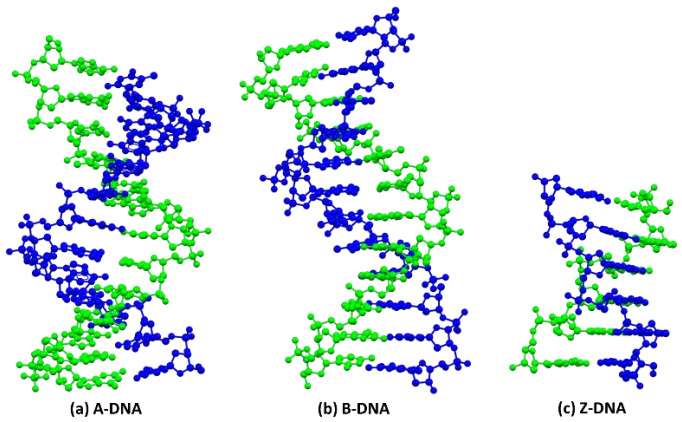


Figure 2: Different helical conformations of DNA; a) A-DNA with small helical rise and central hydrophobic cavity (PDB: 2dcg)[12], b) B-DNA with base pairs sitting on the helical axis (PDB: 1bna)[13], and c) Z-DNA showing the zigzag of the backbone and longer helical axis (PDB: 142d)[14]. Inspired by figure in Reference 15.

1.2.3. **Binding to DNA**

There are three main constituents of DNA which can be targeted by binding molecules: the major and minor grooves (groove binding), the backbone (backbone binding) and the bases (intercalation and covalent modification). Any combination of these modes of binding can be used by a single molecule[16,17].

The major groove shows significant structural variation. The width and charge distribution in the groove are dependent on the local base sequence which allows ligands capable of hydrogen bonding to distinguish between specific binding sites. Proteins may contain alpha-helical units that are too large to bind in the minor groove but fit snugly into the major groove[17]. For minor groove binding, proteins normally use a smaller beta-sheet structure which can fit into the deeper groove. Van der Waals (VdW) and electrostatic forces are utilised in binding due to the decreased variation in charge distribution, however sequence specific minor groove binders which use hydrogen bonding have been discovered[18].

Backbone binding is often used in conjunction with other binding mechanisms. Cationic residues cause electrostatic attraction and hydrogen bonds to form with the phosphodiester bond of the backbone. Backbone tracking, where a molecule relies primarily on backbone binding, has been observed in a synthetic inorganic complex[19,20].

Opening up the hydrophobic gap between two neighbouring base pairs allows intercalation of a planar molecule with an extended aromatic surface[17,21]. The aromatic rings of the base pairs can then π-stack onto the aromatic ligand much like they would onto an adjacent base pair. Distortion of the base pairs next to the intercalating molecule leads to slight unwinding of the double helix and is a possible cause of the neighbour

exclusion principle[9]. A similar type of binding occurs when a ligand with a wider aromatic surface inserts at a mismatch site. This prevents unwinding of the duplex but causes greater overall distortion as the two mismatched bases are ejected from the helix and the binding molecule takes their place[22]. Generally organic ligands do not show sequence specificity, however metal complexes allow co-ordination of further ligands capable of groove or backbone binding. Unsaturated metal complexes are also capable of directly co-ordinating to electron donor atoms in the bases, which is the binding mode of cis-platin[20,23].

1.3.G-Quadruplexes

1.3.1.Structure

G-quadruplexes (GQs) are stacked structures found in G-rich sections of DNA. Sets of four G bases held together by Hoogsteen bonds form tetrads in which the bases are related by a four-fold symmetry axis (see Figure 3a)[2,9]. These tetrads are π-stacked forming the GQ structure [4,9] which has a partially negatively charged channel running down the centre, lined by the O-6 of the G carbonyls.[9] The exact configuration of the GQ is dependent on the base sequence of the G motif, number of strands involved (mono-, di- or tetramolecular) and the orientation of the DNA backbone (parallel, antiparallel, hybrid)[2,4,24].

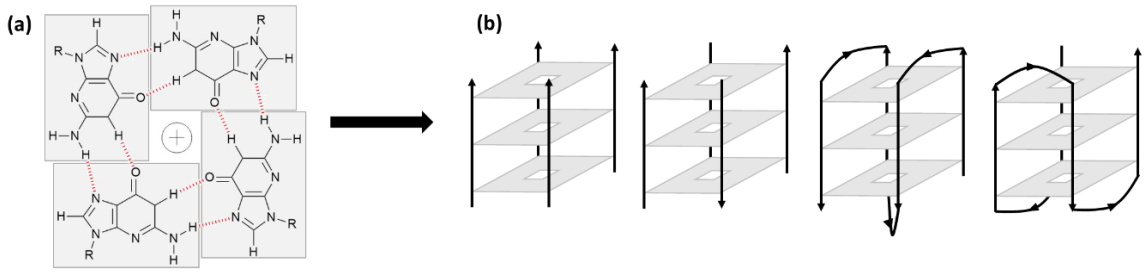


Figure 3: The structure of a G-quadruplex. a) the G bases form a tetrad via Hoogsteen base pairing which then stack to form the quadruplex, and b) some examples of the different possible topologies of a G-Quadruplex. Inspired by image from from Reference 2.

GQs can also be stabilised by a number of internal factors, such as loop length and number of G-tracts, or external factors, such as monovalent ions (Na^+, K^+), which electrostatically bind the negatively charged cavity[2,24,25], as well as GQ binders[4].

Sequences capable of forming GQs are summarised by the consensus sequence ($G_{3+}N_{1-7}G_{3+}N_{1-7}G_{3+}N_{1-7}G_{3+}$), which describes four G tracts each containing at least 3 guanine bases to form the tetrads. The sequence is used to screen for GQ forming sequences (or G motifs) in the human genome, and over 300,000 have been found.[2,4] These sequences tend to cluster in functional areas such as promotor regions[6], intron/exon boarders, at regions coding for 5'-UTRs, and most commonly at telomeres[2,4]. Considering these sequences are conserved in human populations[4], it is likely that they play a functional role in the human genome.

One of the best studied G motifs is the human telomere repeat sequence (TTAGGG). The telomere region is of interest as many tumours have shown increased telomerase activity which is a cause of cancer cell immortality. Telomerase activity can be inhibited by stabilising the GQ[2,26–29].

Another well studied G-motif is within the MYC promotor[2,6] which is often used as a model to investigate the role of GQs in promotor regions. In addition, the MYC

protein is a broad transcriptional activator upregulated in many cancer cell lines. Formation of GQs by negative supercoiling in the promotor region of the MYC gene can reduce proliferation and MYC expression[6].

1.3.2. **G-Quadruplex Binding Ligands**

There are three main modes of binding to G-quadruplexes: intercalation, where planar aromatic molecules stack between the tetrads of the G-quadruplex; outer binding, where the ligand is able to bind to the grooves or loops of the G-quadruplex; and end-stacking, in which a planar aromatic molecule π-stacks onto a terminal tetrad of the G-quadruplex. The latter is the only mode that cannot be used as a method of binding to duplex DNA, and therefore provides selectivity in binding[29].

1.3.2.1. Porphyrins

Porphyrins are macrocyclic ligands which rely heavily on π-stacking interactions to bind to GQs, as demonstrated by $TMPyP_4$ (tetra-(N-methyl-4-pyridyl)porphine)[28]. The extended planar aromatic face of the $TMPyP_4$ molecule (see Figure 4a) is able to π-stack onto the terminal tetrad of the GQ where the aromatic faces of the G bases are exposed. Additional hydrophobic interactions between the aromatic surfaces and electrostatic interactions between the methylpyridinium of $TMPyP_4$ and backbones of the DNA help stabilise the G-quadruplex.

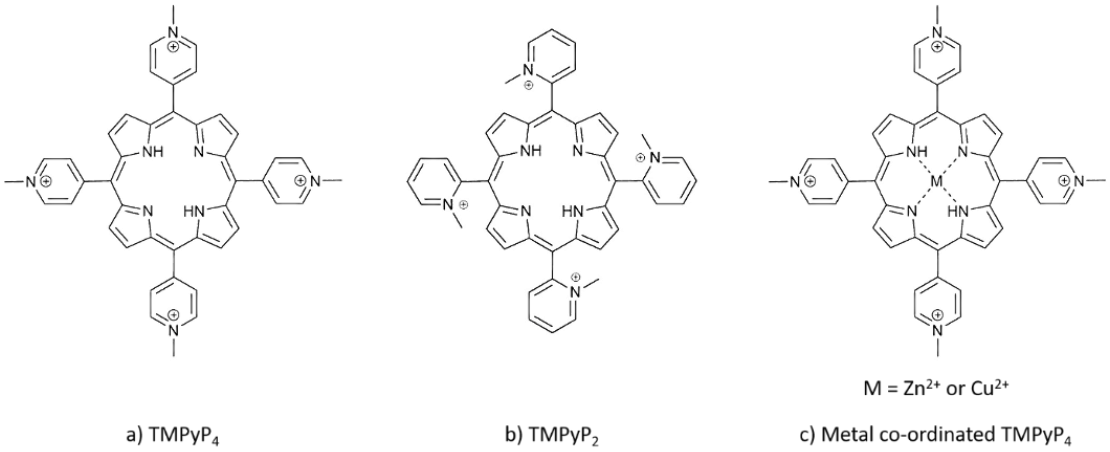


Figure 4: The structure of TMPyP4 (a), TMPyP2 (b) and the metal co-ordinated TMPyP4 complexes (c).

By investigating analogues of TMPyP4, including $TMPyP_2$, it has been shown that stacking interactions are vital for successful GQ binding, and that up to two of the cationic methylpyridinium units can be substituted by pyridine rings containing hydrogen bonding groups, such as alcohol or nitro groups, for increased binding ability[27]. The importance of π-stacking is demonstrated by $TMPyP_2$, a $TMPyP_4$ isomer, in which the N-methyl groups are in the ortho rather than the para position of the meso pyridine ring (see Figure 4b). Changing the position of the methyl groups to ortho causes the pyridine rings to twist out of the planar arrangement which sterically blocks the aromatic surface of the porphyrin, preventing π-stacking[24,27]. These methylpyridinium moieties must be in the meso position to lie in the GQ groove while the cationic groups are aligned with the DNA backbone. Moving these groups causes steric clash of the ligand with the backbone so hinders binding[27]. Unfortunately $TMPyP_4$ shows little selectivity for GQs and can bind to multiple secondary structures of DNA[30].

Metalloporphyrins bind to G-quadruplexes provided that the geometry of the metal ion does not sterically block π-stacking of the porphyrin onto the terminal tetrad of the G-quadruplex structure. Square pyramidal $ZnTMPyP_4$ and square planar $CuTMPyP_2$

are both capable of binding to GQs with $ZnTMPyP_4$ showing better stabilisation than $CuTMPyP_4$[24,27]. While $CuTMPyP_4$ can intercalate, potentially destabilising the GQ, $ZnTMPyP_4$ has an axial ligand which prevents intercalation and could explain its greater stabilising ability[24,27,29].

In cells, $TMPyP_4$ and its complexes with zinc and copper have been used to stabilise GQ structures in the telomeres causing a reduction of telomerase activity. This has worked most successfully for $ZnTMPyP_4$[26–28].

1.3.2.2.Pyridostatin

Pyridostatin is a well-established GQ binder with good selectivity for GQ structures over duplex DNA[31]. The large electron dense aromatic surface π-stacks onto a terminal tetrad, while hydrogen bonding groups able to bind to the DNA backbone or bases in GQ loops are connected by flexible linkers. The aromatic surface is held in a planar conformation by hydrogen bonding to a water molecule (see Figure 5). Treating cells with pyridostatin causes an increase in the population of cells in the G2-phase due to growth arrest[32]. At low concentration pyridostatin is able to bind to GQ forming sites throughout the genome such as the SRC proto-oncogene promotor where it represses transcription. At higher concentrations, pyridostatin also binds to telomeres and destabilises the shelterin complex which protects the telomeres from digestion[31].

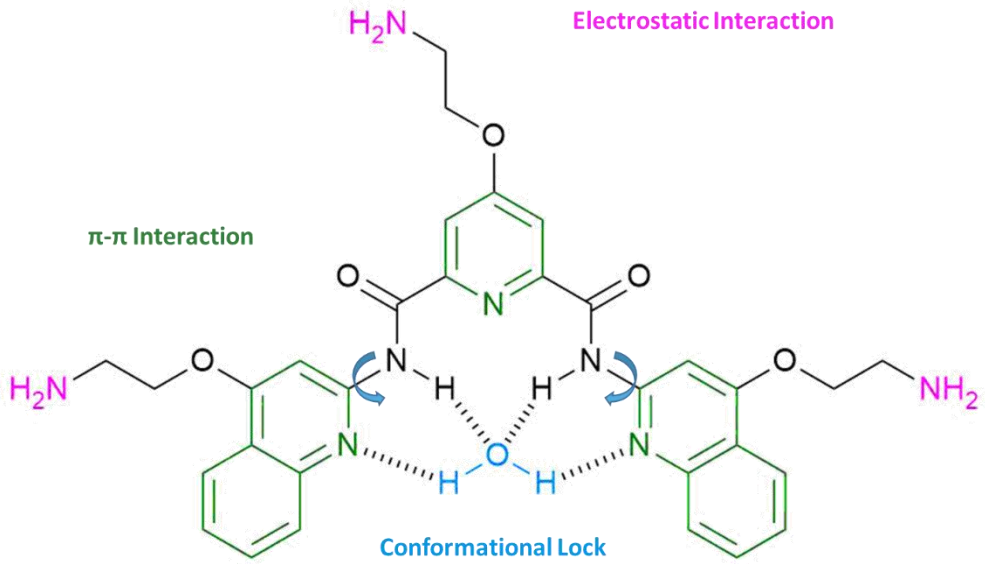


Figure 5: Structure of pyridostatin. The key binding features of the molecule have been highlighted. A water molecule is co-ordinated showing the locked planar structure. Inspired by figure in Reference 33.

Pyridostatin analogues with varying side chain lengths and functionalities have been tested for GQ binding ability[33]. None of these analogues bind to duplex DNA, demonstrating the incredible selectivity of the bis(quino-linyl)pyridine-2,6-dicarboxamide template, however their binding to GQs was varied. Amine, halogen and sugar side chains were able to stabilise the G-quadruplex structure through electrostatic interactions and potential hydrogen bonding. Substitutions such as fluorobenzyloxyl and triazole groups without flexible linking chains were less able to stabilise the GQ due to steric clash (see Figure 6). The side arm chains therefore significantly contribute to the binding of pyridostatin and demonstrate that π-stacking alone does not stabilise the GQ structure.

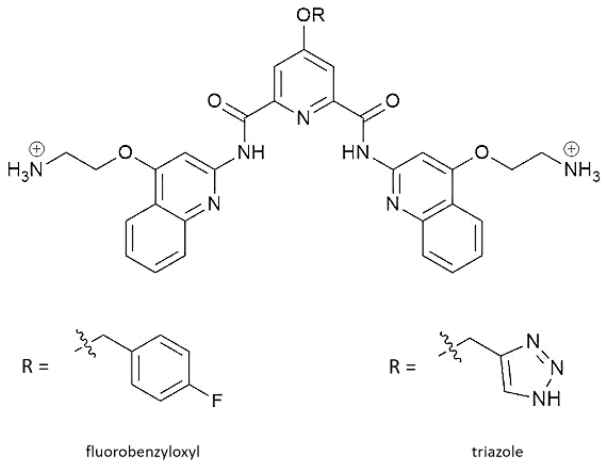


Figure 6: Pyridostatin analogues which show reduced binding to GQs despite the inclusion of electron withdrawing heteroatoms. This could be caused by steric clash with the G-quadruplex structure preventing tight binding.

The selective binding of pyridostatin to G-quadruplexes over duplex DNA has been exploited by the Balasubramanian group by the attachment of different functional groups such as fluorescent labels[34] and biotin[32] (see Figure 7). These have been used to demonstrate the binding of pyridostatin to GQ forming sequences in cells.

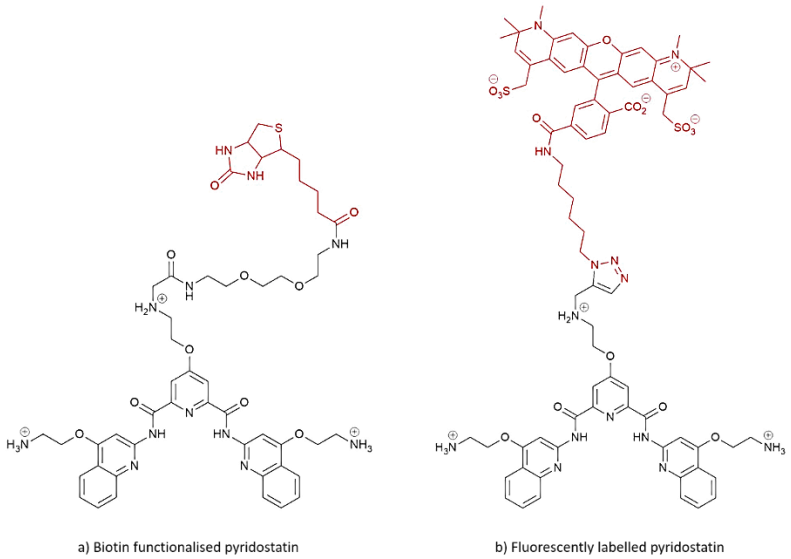


Figure 7: Examples of functionalising the pyridostatin scaffold: a) Biotin used to pull G-motifs from genomic DNA, and b) fluorescently labelling so pyridostatin can be tracked in cells.

Pyridostatin derivatives selectively bind GQ structures, but do not distinguish between different GQs. Specific binding to telomeric GQs has been achieved by covalently binding pyridostatin to pyrrole-imidazole polyamide chains, which are known for their sequence selective minor groove binding to dsDNA (see Figure 8)[18,35–37]. The polyamide chains have been designed to target the TTAGGG repeat sequences flanking the telomeric GQs, causing the pyridostatin derivative to favour telomeric GQ binding over other GQ structures flanked by alternative sequences elsewhere in the genome. The binding strength was also found to be increased when compared to the individual DNA ligands, pyridostatin and the polyamide chain[37].

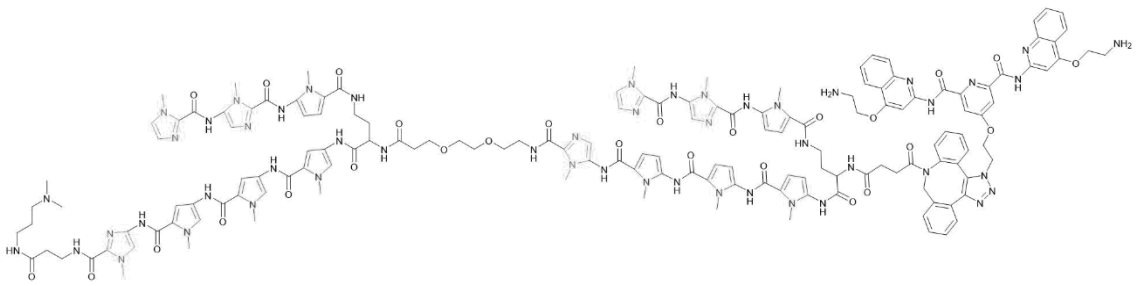


Figure 8: The polyamide-pyridostatin structure, and it's binding to telomeric DNA[37].

Recently, the selective binding and stabilising of GQs by pyridostatin has been utilised to map the GQ structures in the genomes of 12 different species, ranging from bacteria to humans[38].

1.3.2.3.Metal-Salphen Complexes

A variety of metal-salphen complexes have been shown to stabilise GQ structures[39]. The salphen ligand has three aromatic rings which are spaced to allow π-stacking on to three guanine bases of a G tetrad (see Figure 9b). Piperidine functionalised side chains are protonated at physiological pH so increase both the water solubility and

the binding affinity due to electrostatic attraction between the positively charged groups and the loops, grooves and backbone of the GQ DNA[40]. Co-ordination to a square planar metal ion such as Ni(II)[39–41] or Cu(II)[39,41] forces the salphen ligand to take the required planar structure while the complex cation is held over the channel where the partially negatively charged carbonyls meet. The favourable electrostatic interactions increases the stabilising effect of the ligand binding. These complexes show excellent binding and selectivity for GQ DNA, with a preference for parallel structures over antiparallel.

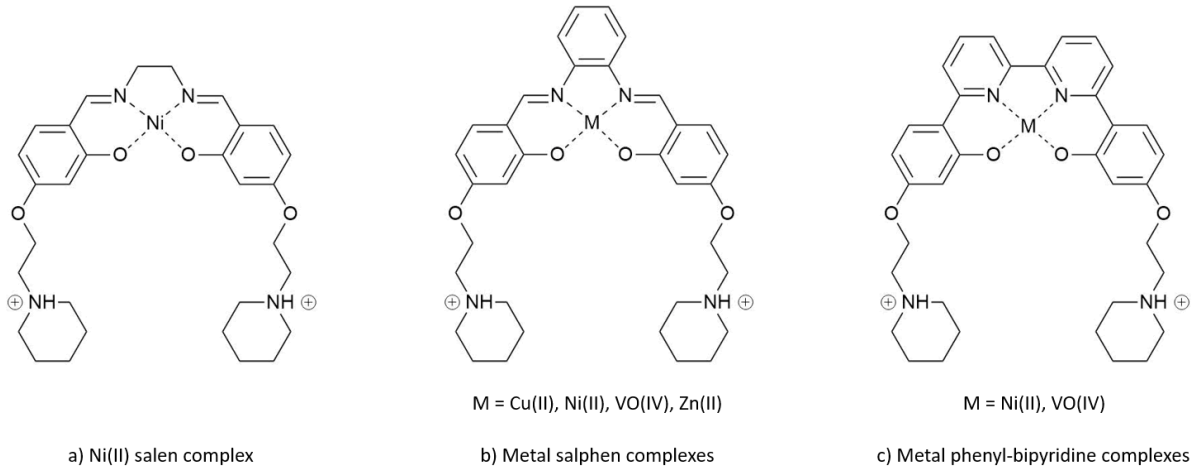


Figure 9: The structures of the investigated salphen complexes: a) the Ni(II) salen complex, b) the salphen ligand which can be co-ordinated to Cu(II), Ni(II), VO(IV) or Zn(II), and c) the phenyl-bipyridine ligand complexes with Ni(II) or VO(IV). Note the increasing number of aromatic rings in each ligand.

Similar complexes have been tested using metal centres with square pyramidal geometries, such as Zn (II) or VO (IV). These complexes show reduced affinity for duplex DNA when compared to the square planar complexes as they were unable to intercalate, but also showed a reduced affinity for the GQ suggesting that π-stacking requires both faces of the complex to be accessible. This may explain the weaker binding of Cu(II) compared to Ni(II) complexes, despite aromatic surface of the Ni(II) complex not being completely planar, as Cu(II) is able to co-ordinate a water molecule in the axial position which obstructs one face of the complex[41].

Extending or reducing the aromatic surface of the complexes has been investigated using salen[41] or bipyridine[39] ligands respectively. The salen ligand, with its reduced number of aromatic rings, shows weaker binding to the GQs, whereas the bipyridine ligands show stronger binding due to the extra aromatic ring available for π-stacking.

1.4.I-Motifs

1.4.1.Structure

Intercalating motifs, or i-motifs, were first discovered by Gehring, Leroy and Gueron in 1993[3]. They require C-rich sequences where a tract of protonated cytosines form hydrogen bonds to a second tract of unprotonated cytosines (see Figure 10) giving a parallel duplex structure. Two such duplexes are held antiparallel to one another by their intercalated base pairs (see Figure 11), the exact conformation of which can be classed as R, S or T[1]. Similar to GQs, stability of the i-motif is dependent on the number of base pairs as well as the length and sequence of the single stranded loops found at either end.[42] These loops allow i-motifs to be roughly grouped into two classes; class I has shorter loops while class II has longer loops[1].

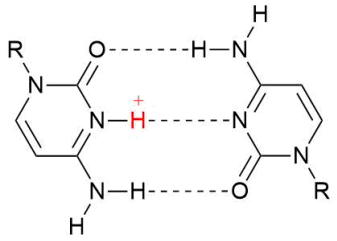


Figure 10: A CC^+ base pair. The extra proton is highlighted in red. R represents a sugar moiety of the DNA backbone.

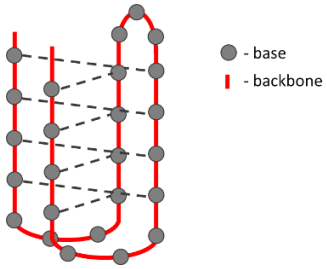


Figure 11: A possible i-motif conformation indicating the intercalating base pairs. Inspired by figure in Reference 1.

The four grooves of an i-motif show a much larger difference in width compared to the two grooves of B-DNA, with two wide major grooves and two very narrow minor grooves. The crystal structure of a simple d(C_4) i-motif (see Figure 12) illustrates the flat, ribbon like structure with two sets of two backbones in close enough proximity to be stabilised by Van der Waals contacts. Unlike other secondary structures of DNA, π-stacking interactions do not play a large role in stabilising the i-motif as the rings of the bases do not significantly overlap. Only the N4 and O2 atoms of the bases overlap and are held in close proximity which may allow the electron clouds to interact and compensate for the loss of π-stacking[43].

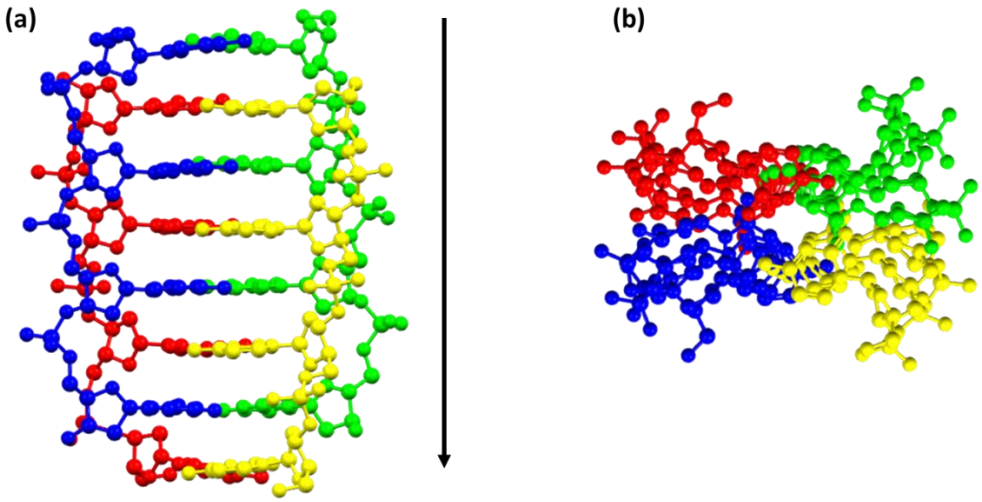


Figure 12: d(C_4) i-motif viewed from the side (a) and down the central axis (b), made of two parallel duplexes (red and yellow, green and blue) in an antiparallel configuration. The narrow grooves and close proximity of the backbones of the two duplexes can be seen (PDB: 190D, as in Reference 43).

Other factors, such as epigenetic modifications, molecular crowding, cationic conditions and negative supercoiling, have also been shown to stabilise i-motifs[1,42].

Sequences capable of forming i-motifs have been found throughout the genome with particularly high abundance in functional regions such as the telomeres, centromere and oncogene promotors[1], including the c-MYC promotor and the BCL-2 promotor[42]. The requirement for a protonated cytosine to form the CC^+ base pairs has lead to speculation over whether the i-motif is able to form at physiological pH. Recently, Zeraati *et. al.* have presented evidence of i-motif formation in vivo using an i-motif specific antibody[42]. They conclude the formation of i-motifs are cell cycle dependent, with i-motif structures being most predominant in the G1 phase when significant transcription and growth occurs.

1.4.2. **I-Motif Binding Ligands**

1.4.2.1. Cyclic Tetraoxazole Compounds

Masoud *et. al.*[44] have identified large cyclic hexa and hepta oxazoles as G-terminal quartet binders, however the smaller tetraoxazole is a similar size to a CC^+ base pair which lead to the investigation of L2H2-4OTD (see Figure 13a) and its i-motif binding potential[44]. The cyclic structure and primary amines of the side chains of L2H2-4OTD were proven critical for binding as both analogues lin-L2H2-4OTD and L2A2-4OTD were unable to bind to the i-motif (see Figure 13b-c). NMR and CD studies using increasing sequence to ligand ratio suggest two binding sites for the ligand, both of which cause distortions in neighbouring base pairs[44].

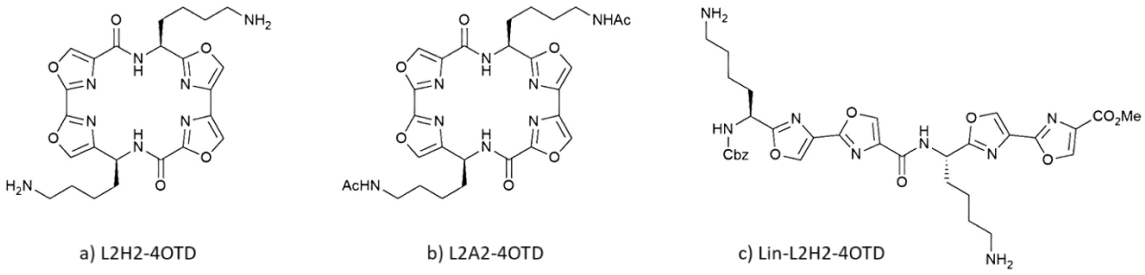


Figure 13: Structure of L2H2-4OTD (a), the acetylated side chain analogue L2A2-4OTD (b), and the linear analogue lin-L2H2-4OTD(c).

1.4.2.2. Acridone Derivatives

Derivatives of 2-methylacridin-9(10H)-one have been synthesised and tested for i-motif binding and stabilising ability[5]. Product B19 (see Figure 14) shows high selectivity for i-motifs without binding to GQs or duplex DNA. When tested on the c-myc promotor, known to contain an i-motif forming sequence[1,5], increasing the dosage of the ligand caused a reduction in both the abundance of c-MYC mRNA and the expression of c-MYC protein. This suggests that stabilising the i-motif in the c-myc promotor region could inhibit transcription and expression of the the c-myc gene.

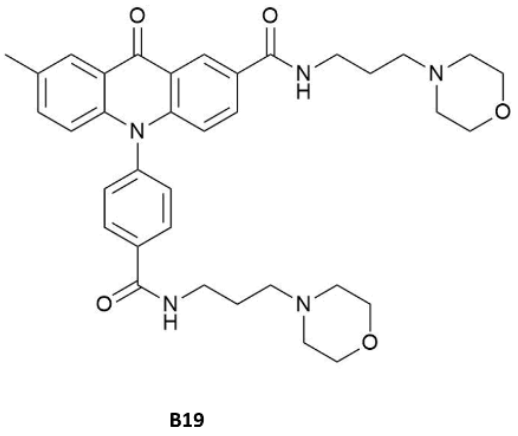


Figure 14: Structure of B19.

1.4.2.3.PBP1

The para (PBP1) and meta (PBP2) isomers of a peptidomimetic ligand are able to selectively bind to i-motifs and GQs respectively (see Figure 15)[45].

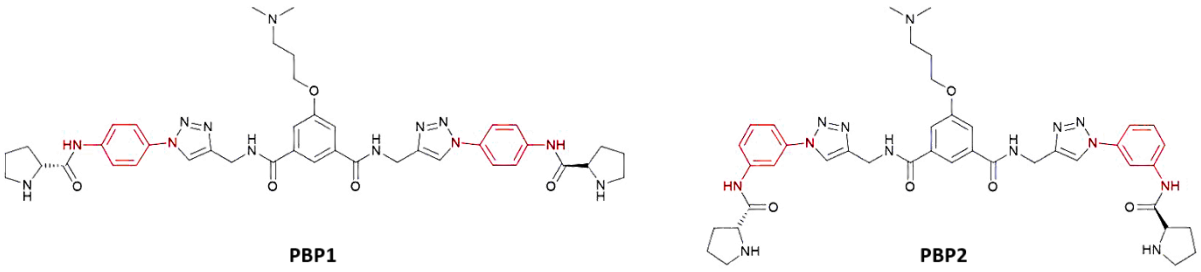


Figure 15: The structural isomers PBP1 and PBP2. The coloured rings show the meta and para conformations.

The C- and G-rich sequences BCL-2 and c-MYC (both oncogene promotor regions) and h-TELO (from the telomeric region) were chosen to test the ability of PBP1 and PBP2 to stabilise two different secondary DNA structures, i-motif and GQ. PBP1 showed greater stabilisation of all of these i-motif forming sequences compared to their complimentary G-motifs, with a particularly high ability to stabilise the BCL-2-C sequence (ΔT_m = 29°C). PBP1 showed greater affinity and binding strength for the i-motif forming sequences than the GQ compliments which was demonstrated by a thiazole orange displacement assay and fluorescence binding titrations. Again, PBP1 showed particularly high affinity for BCL-2 (DC_{50} = 0.9 µM, K_d = 0.3 µM). The meta isomer, PBP2, showed similar results for stabilising GQ sequences. The BCL-2 expression in cancerous cells treated with PBP1 lead to the upregulation of BCL-2 expression, while the same cell line treated with PBP2 showed the downregulation of BCL-2 expression. This could be due to the stabilisation of the i-motif or GQ in the promotor region of BCL-2 upon addition of PBP1 and PBP2 respectively, implying a regulatory role for the non-canonical secondary structures[45].

1.4.2.4.SWNTs

Nanoscience has also been employed to find possible ligands for the i-motif structure. There is evidence to suggest that carboxyl-modified single-walled carbon nanotubes (SWNTs) are able to bind and stabilise i-motif structures[46]. CD studies show no significant disruption to the DNA structure upon SWNT binding, ruling out minor groove binding due to the large size of the SWNTs. Thus it is hypothesised that the SWNTs bind the 5' end of the major groove of the i-motif in agreement with the 2-Ap fluorescence quenching assay.

Stabilisation by the nanotubes could be due to a number of factors. The carboxyl (COO^-) modifications allow electrostatic interactions between the SWNTs and the CC^+ base pairs, and replacement of the carboxyl group by a cationic amino group causes destabilisation of the i-motif. Increased stacking of the TTA loop bases upon SWNT binding could also increase the stability of the structure, as could VdW interactions with the DNA backbone[46].

1.5.Junctions

1.5.1.Structure of three way and four way junctions

DNA can form junctions from which multiple arms of duplex or single stranded DNA can branch off. The Holliday junction, or four way junction (4WJ), is formed when two double helices are held together by exchanging DNA strands. 4WJs are intermediates during key cellular processes, such as DNA repair and recombination, and take one of two forms. The double helical arms can be extended in a perpendicular arrangement with all the bases at the junction facing in one direction. The sides are assigned the major and

minor groove, and give a pyramidal structure. The second form of a 4WJ is the X-structure, where the arms are helix-helix stacked. The angle between the stacked branches is roughly 60^{o} but can vary. The structure adopted by each 4WJ is dependent on the base sequence but also the electrostatic repulsion of the backbones (favouring the open form) and the stabilising base pair stacking interactions (favouring the X form). Base pair stacking alone does not compensate for the destabilising backbone repulsion so the open form dominates unless divalent metal cations are present which alleviate the electrostatic repulsion and allow the X-structure to form[47,48].

A three way junction (3WJ) has three arms extending from a central branch point. The exact conformation of the junction is dependent on multiple factors including the local sequence of the DNA and the ionic environment[47,49]. The main energetic contributions to the structure are steric contributions and electrostatic interactions[49,50].

A perfect 3WJ, or 3H structure[51], occurs when all the bases at the junction point are Watson-Crick base paired. This is referred to as an open or Y shaped structure, as all the arms are approximately equally spread (see Figure 16)[50,52]. The restricted conformation creates a hydrophobic cavity in the centre of the junction leaving one side of the bases exposed to the environment[52]. The junction is not planar as each helical arm possesses a major and minor groove which combine to form one concave and one convex major groove[50,53].

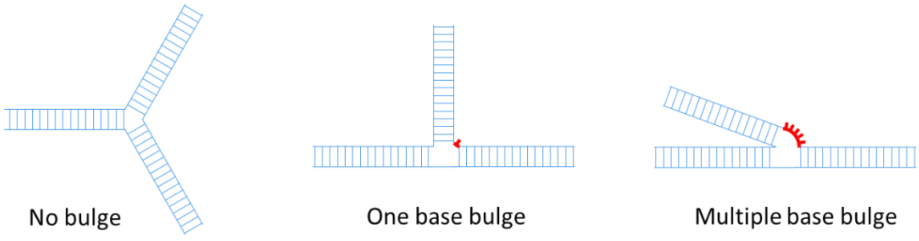


Figure 16: Structure of three way junctions with and without bulges. Inspired by figure from Reference 49.

3WJs can incorporate between two and five unpaired bases which are capable of stabilising the junction[54]. These are classified as $3HS_n$ where n is the number of unpaired bases in the bulged structure[51]. The steric constraint that prevents folding of 3H junctions is lifted by increasing the number of unpaired bases at the bulge[50]. In the presence of metal cations, a folded structure can form where two arms which do not contain bulges align so they are parallel to one another and show helix-helix stacking[49]. The third arm is suspected to be bent away from the unpaired bases at a reduced angle[50]. Stable junctions form when there are 2 or 3 unpaired bases on one,[49,54] while structures with only one unpaired base are less stable (see Figure 16)[55].

1.5.2. **Three Way Junction Binding Ligands**

1.5.2.1. Cylinders

The Hannon group have been investigating the properties of cylinders, a type of supramolecular helicate so called for their cylindrical shape. When first designed, the parent cylinder, a dinuclear, three ligand helicate, was unlike any other supramolecular helicates as it is relatively simple to synthesise and uses commercially available starting materials (an aldehyde and an amine) to produce the imine metal co-ordinating groups[56]. Whereas other helicates had flexible linking regions, the diphenyl spacers in the parent cylinder face-edge π-stack in two sets of three aromatic rings, holding the ligand in a rigid conformation with restricted flexibility from the central methyl group (see Figure 17)[56–58]. The rigid design of the ligand inhibits the co-ordination of the two imine groups to the same metal centre, aiding the synthesis of the helicate, while conjugation of the imine bonds with the terminal pyridine rings helps to prevent hydrolysis of the bonds[56]. The

chiral structure of the helicate gives two possible enantiomers of the cylinder, the left handed "M" enantiomer, or the right handed "P" enantiomer[59].

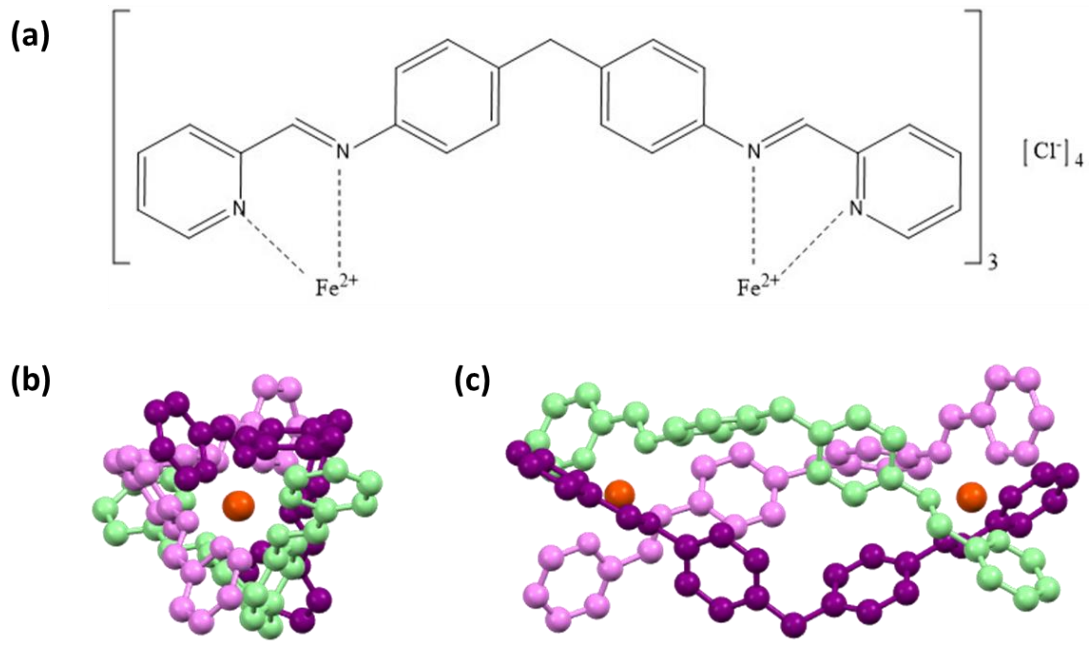


Figure 17: Molecular structure (a) and ball and stick model of the parent cylinder down the molecular azis (b) and parallel to the axis (c). (PDB: 3i1d, as seen in Reference 60).

Complexing of the ligand has been tested with many transition metals cations including Fe^{2+} [20,56,59,61–63], Ru^{2+} [20,58,64], Ni^{2+} [56], Cu^{+} [56] and Ag^{+} [56]. These metals have varied characteristics suitable for different purposes[58]. The well-studied iron cylinder has a deep purple colour corresponding to a metal ligand charge transfer (MLTC) band lying within the visible region allowing its binding to be monitored spectroscopically[53,62]. Meanwhile the ruthenium cylinder can cleave DNA by photoactivation[58].

The parent cylinder measures 1 nm in diameter by 2 nm in length[20] so is too large for the minor groove of B-DNA (5.7 Å wide, 7.5 Å deep)[9] , however the size is comparable to that of an alpha helix unit (1-2 nm in length)[65] used to bind to the major groove of DNA[58]. The electrostatic interactions between cationic cylinder and negatively charged backbone, as well as the size of the cylinder allows binding to the major groove

of DNA, causing intrastrand coiling[59,61]. Coiling appears to be co-operative, with atomic force microscopy (AFM) imaging showing some bent DNA strands while others are completely coiled.[61] This has been compared to the coiling of DNA found in the nucleus of cells, mediated by histones[58,59,66]. The M enantiomer causes a greater degree of coiling than the P enantiomer which could be due to the helical turn of the cylinder and the DNA being in phase[59].

Interest in the supramolecular parent cylinder however is not in its binding to duplex DNA, but its ability to bind non-canonical structures of DNA and particularly 3WJs. The size and shape of the cylinder allows it to sit in the cavity of a Y-shaped 3WJ without distortions to either the cylinder or the DNA (see Figure 18)[53,58]. Aside from the electrostatic interactions between the cylinder and DNA, as seen in the major groove binding, the central cavity of the 3WJ creates a hydrophobic environment that can bind the aromatic core of the cylinder. The two central phenyl rings of each helicate ligand are able to face-face π-stack on to the aromatic rings of the exposed bases on one strand of the 3WJ. The pyridine rings on the end of the cylinder stick out into the convex minor groove of the junction orientated so that they fit between the backbones of the minor groove of each double helical arm. This is not possible on the concave major groove side of the 3WJ as the pyridine rings are facing away from the groove. The pyridine ring binding is similar to intercalation, however cylinders do not cause the same disruption to the base pairs[58]. The M enantiomer of the cylinder a more stabilising effect on the 3WJ, again because of the reverse helical twist in the P enantiomer[63].

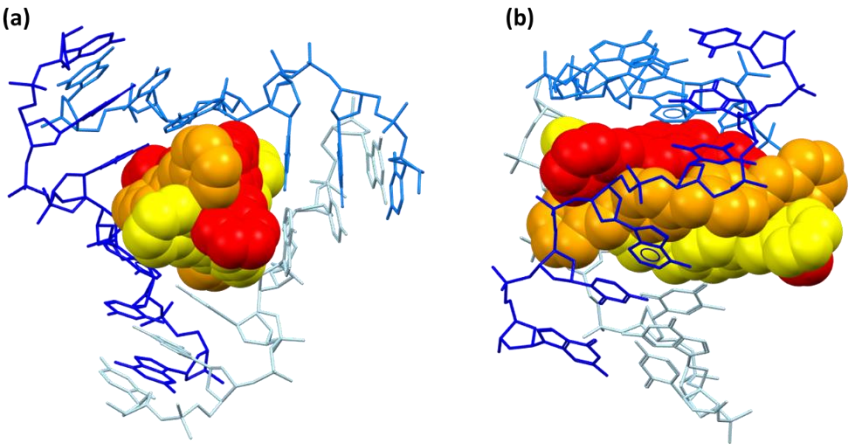


Figure 18: The parent cylinder sits in the three way junction viewed down (a) and parallel (b) to the molecular axis (PDB: 2ET0, as seen in References 53 and 58).

The exact size and shape of the cylinder has proven critical to its DNA binding. Alterations of the parent cylinder structure, such as the addition of two CF_3 groups to the central methyl group or phenyl groups to the imine carbon, have shown a decreased stabilising ability for a 3WJ[63]. Ketimine groups between the pyridine and spacer unit to elongate the cylinder cause a decreased coiling effect on the cylinder[66].

Studies into the cytotoxicity, genotoxicity and mutagenicity of the iron cylinder suggest that the helicate has a cytotoxicity similar to carboplatin but shows no evidence of direct genotoxicity or mutagenicity[62]. The cytotoxicity of a double stranded form of the ruthenium cylinder has also been tested[64]. The unsaturated cylinder was designed to allow coordination to DNA in a similar way to cisplatin, and so has two chloride ion leaving groups which can be arranged cis/cis, cis/trans or most commonly, trans/trans. The trans/trans isomer that has the highest cytotoxicity against HBL100 and T47D cell lines, greater than the iron cylinder and cisplatin.

1.5.2.2.Triazacyclononane Ligands

An important feature of the supramolecular cylinder is its three fold symmetry, which allows the π-stacking onto the bases of the 3WJ. This principle was used by Vuong

et. al. to produce ligands based on a 1,4,7-triazacyclononane (TACN) framework; TACN-quinoline (TACN-Q), TACN-acridine (TACN-A), and TACN-naphthalene (TACN-N) (see Figure 19). TACN-Q shows the greatest similarity to the parent cylinder as, unlike TACN-A and TACN-N, it is able to bind a cation (Li^+) using the lone pairs on the nitrogen atoms of the side arms. The ligand is able to take a helical conformation with face-edge π-stacking between the aromatic rings of the quinoline groups and an overall diameter similar to that of the cavity in a 3WJ (10 Å). This allows TACN-Q to bind to 3WJs with reasonable affinity and very high selectivity[30].

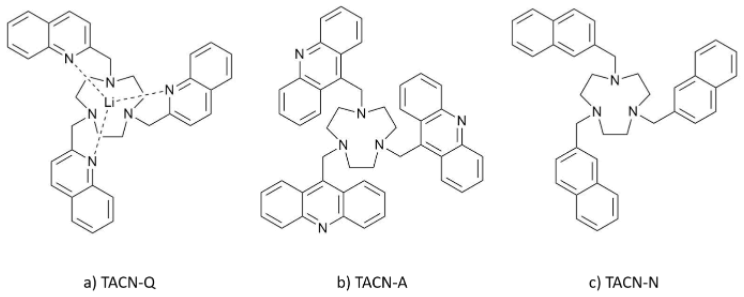


Figure 19: The structures of: a) TACN-Q co-ordinated to a Li^+ cation, b) TACN-A which cannot co-ordinate to an ion due to the position of the nitrogen atoms, and c) TACN-N which has no nitrogen atoms for co-ordination.

Cationic azacryptands (see Figure 20) have also been shown to bind to the central cavity of the 3WJ. The aromatic arms are not restricted by metal complexing so are able to position themselves for optimal π-stacking onto the aromatic rings of the bases. Of the investigated ligands, 2,7-TrisNP and 3, 3'-TrisBP showed the highest affinity for 3WJs and 3, 3'-TrisBP gave the greatest selectivity[67].

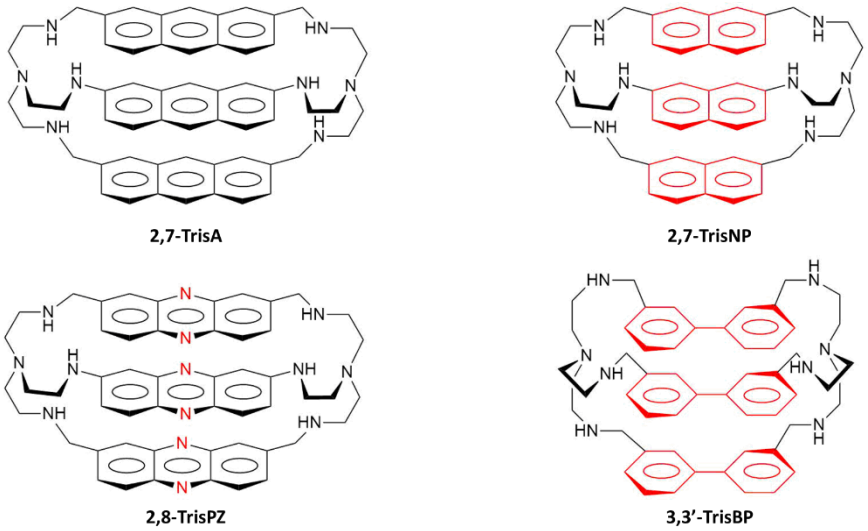


Figure 20: Monchaud's azacryptands, where modifications to 2,7-TrisA have been highlighted (red). Their size is similar to the parent cylinder allowing them to sit in the three way junction. Inspired by Reference 67.

1.5.2.3.Calix[3]carbazole Ligands

Another ligand which utilises the three fold symmetry of the 3WJ is the calix[3]cabazole ligand (see Figure 21)[68]. The calix[3]carbazole was shown to stabilise the 3WJ by denaturing and polyacrylamide gel electrophoresis experiments, and a 1:1 binding ratio was suggested by titrations monitored using ultra-violet visible spectroscopy (UV Vis). Interaction between the aromatic rings of the calix[3]carbazole and the bases is evidenced by the appearance of an induced circular dichroism (ICD) band in the circular dichroism (CD) spectrum, indicating the preferential absorbance of a specific direction of circularly polarised light (left or right) due to the binding of an achiral ligand within a chiral environment. Meanwhile, the peak at 280 nm is largely unaffected by the increasing DNA to ligand ratio, suggesting that the helical structure is unaffected by the binding mode. 3WJ binding is therefore likely for such a large DNA ligand, as it allows access to the bases in the junction without disturbing the helical structure of the DNA.

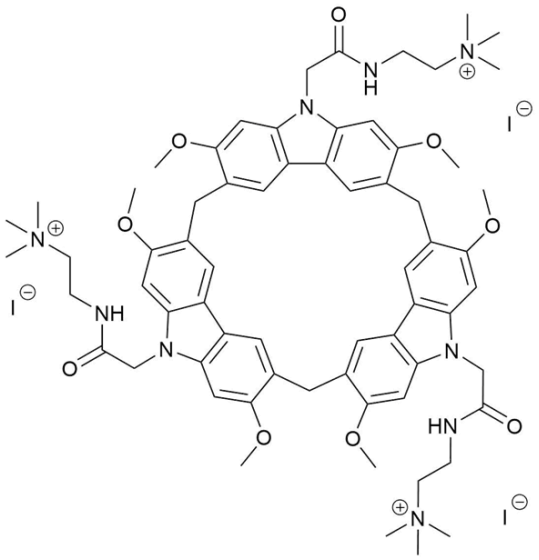


Figure 21: The structure of the calix[3]carbazole ligand[68].

Calix[3]carbazoles are the first example of a fluorescent ligand selective for 3WJ binding so has potential for application as a 3WJ probe. Electrostatic interactions between the negatively charged backbone of DNA and the cationic ammonium moieties of the ligand quench the ligand fluorescence. This means that the ligand is only able to fluoresce in the presence of DNA when bound in the hydrophobic environment of the 3WJ[68].

1.6.Conclusion

Our knowledge of non-canonical structures of DNA and their binding ligands is extremely varied depending on the structure we consider. G-quadruplexes have been well characterised and their existence and function in vivo widely accepted. Multiple ligands with clear binding modes are available, and research into ligands which bind only certain conformations of G-quadruplexes has now begun. Conversely, while the different conformations of i-motifs have been characterised, evidence of their existence and

possible functions in vivo is still under debate. Highly selective ligands able to stabilise the intercalated structure are rare and the binding modes are unclear.

The biological relevance of the Y-shaped junctions during transcription and replication makes them very interesting non-canonical DNA structure compared to i-motifs and G-quadruplexes. Though there is a lack of molecules which specifically bind to DNA, their binding modes have been well defined, making the design of new binding ligands a realistic goal.

Chapter 2:

SYNTHESIS OF NOVEL CYLINDERS

2.1.Introduction

As discussed in Chapter 1, supramolecular helicates consist of one or more ligand molecules wrapped around a helical axis which is defined by two or more metal cations[69]. These complexes can be catagorised based on the number of coordinated ligands (i.e. single-stranded, double-stranded, triple-stranded), and are often further subdivided dependent on the number of cation centres (dinuclear, trinuclear)[70].

To synthesise any supramolecular structure, the components must be carefully designed to encode the intrinsic information required for each building block to be recognised and self assembled[71]. The formation of bonds between the component parts must also be reversible as the different molecules must organise themselves to find the thermodynamic minimum. Metal complexes are the ideal candidate for the study of supramolecular systems as, unlike covalent bonding, coordinate bond formation is reversible and still strong enough to form stable structures. In addition, the encoding of molecular information into the ligand and the choice of metal centre allow supramolecular structures to be designed with relative ease.

Transition metals (particularly the first row transition metals, titanium to copper) have many interesting physical and chemical properties which can be exploited by chemists, such as their formation of highly coloured complexes and multiple available oxidation states[72]. These properties arise from their partially filled d-orbital electron configurations. Upon ligand binding, the d-orbitals of the metal centre lose their degeneracy and the new splitting is dependent on the geometry of the ligand binding. Metal ions favour certain binding geometries, as their electrons can be arranged to give the overall lowest energy. This is the crystal field theory.

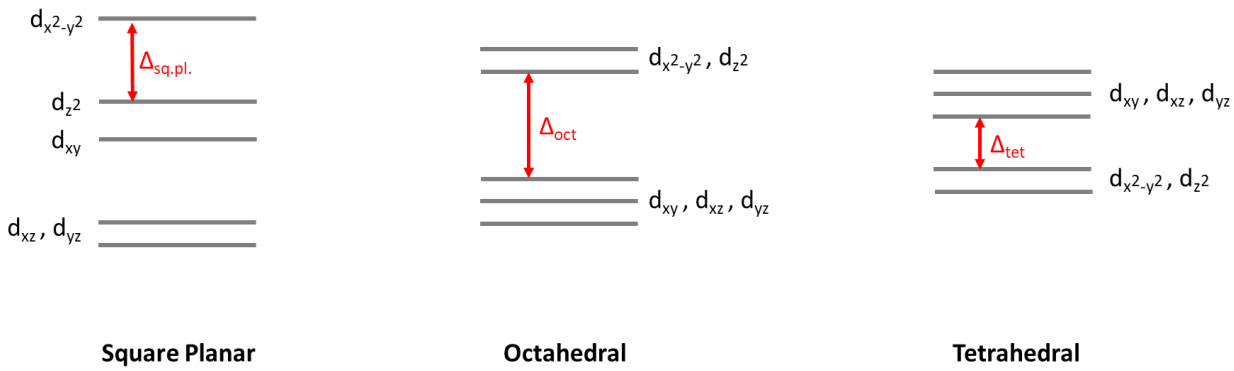


Figure 22: A diagram to show the crystal field splitting of the valence orbitals in different geometries.

The splitting between the non-degenerate d-orbitals is referred to as the crystal field splitting energy, Δ (see Figure 22). For most transition metal ions, the octahedral (6-coordinate) or the tetrahedral (4-coordinate) geometry is favoured with the exception of some d^8 metals, which prefer the square planar geometry. The preferred geometry of a metal ion is a clear example of the intrinsic information encoded by the metal ion, for example iron(II) complexes tend to favour the octahedral conformation.

A further consideration is the choice of ligand. Generally, the design of a ligand can be considered in two parts: the selection of the coordination sites and the choice of spacer unit.

As discussed earlier,the binding groups must be chosen so that they will form strong but reversible bonds. Typically, a bidentate or tri-dentate group is chosen as their binding is strengthened by the chelate effect. The number of binding groups per ligand must also be considered as this can alter the nuclearity of the helical product. Finally, the binding groups do not have to be identical, which can lead to the production of asymmetric ligands, with the possibility of heterometallic helicate formation (see Figure 23) or strand orientation control (either head-head-tail, HHT, or head-head-head, HHH)[73].

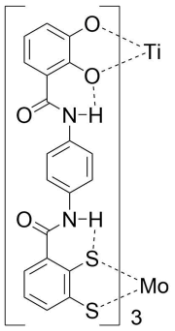


Figure 23: An example of a heterobimetallic helicate[73].

The role of the spacer may be less obvious, however it is no less important. The spacer, which refers to the section of the ligand molecule between the binding groups, must be flexible enough that the ligand can twist to form the double helix[69,74]. At the same time, the spacer region must be rigid enough to prevent the ligand from folding and binding to the same metal centre with multiple binding groups.

Depending on the way the ligands wrap round the helical axis (or rather the stereochemistry at each metal centre) a helicate can be defined as either an M enantiomer, with left handed twist, or a P enantiomer, with right handed twist. Here it should be mentioned that there are also a type of structure called a mesocate, where the stereochemistry at each metal centre is opposite (for a dinuclear system) meaning that the ligands do not technically wrap around the helical axis[70]. This will not be discussed further here.

By careful choice of metal centre and ligand it is possible to predict the form of the supramolecular structure produced. Selecting an octahedral metal centre and a ligand with two bidentate binding groups and an appropriate spacer, would be expected to generate a triple stranded helicate, whereas by selecting a tetrahedral metal centre or switching from bidentate to tridentate binding sites, you would expect a double stranded helicate.

It should be noted that the choice of counterion can also be important for the formation of a supramolecular architecture. Hasenknopf *et al*. were the first to report a circular helicate, in which each octahedral iron(II) centre was co-ordinated to a bipyridine binding group of three separate ligands. By using different counterions (chloride of sulphate), the number of metal centres involved in the circular helicate could be controlled (either 5 or 6). The circular helicates were able to form around the anion in solution, meaning that the cavity of the helicate had to be widened for the sulphate anion due to its larger size, and resulted in the hexanuclear circular helicate instead of the pentanuclear helicate. In addition, alterations to the ligand spacer group found that by increasing the length and flexibility of the spacer group, a tetranuclear circular helicate could also be formed[75].

2.2.Aims

Much of the work in the Hannon group has been focussed on the design and synthesis of novel supramolecular helicates which are often analogues of the parent cylinder (see Figure 24)[76–79]. This chapter will focus on the synthesis of new supramolecular helicates that have been designed to include certain features of interest.

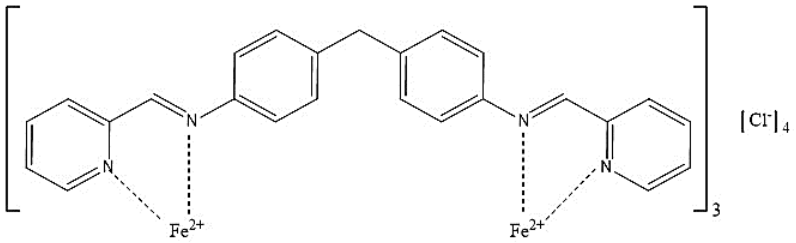


Figure 24: Structure of the parent iron cylinder.

So far, many of the cylinders investigated by the Hannon group have been dinuclear helicates with one central, aromatic methylenedianiline group which facilitates recognition of three way junctions. The synthesis of a trinuclear helicate based on the

parent cylinder is investigated in this chapter, with two spacer units which have the potential to bind two distinct three way junctions to one cylinder.

There are also few examples of cylinder analogues that are asymmetrical[79] which is due to the symmetry of the 4,4'-methylenedianiline precursor which is difficult to react only on one side. The synthesis of asymmetric helicates consisting of two different metal binding sites is attempted and reported.

To achieve either of these aims, the half ligand was synthesised following the work of previous Hannon Group members Laura Childs[80], Jennifer White[81] and Susana Vitorino[82]. The purification of the half ligand is also considered in some detail.

2.3. The Half Ligand (L_{half})

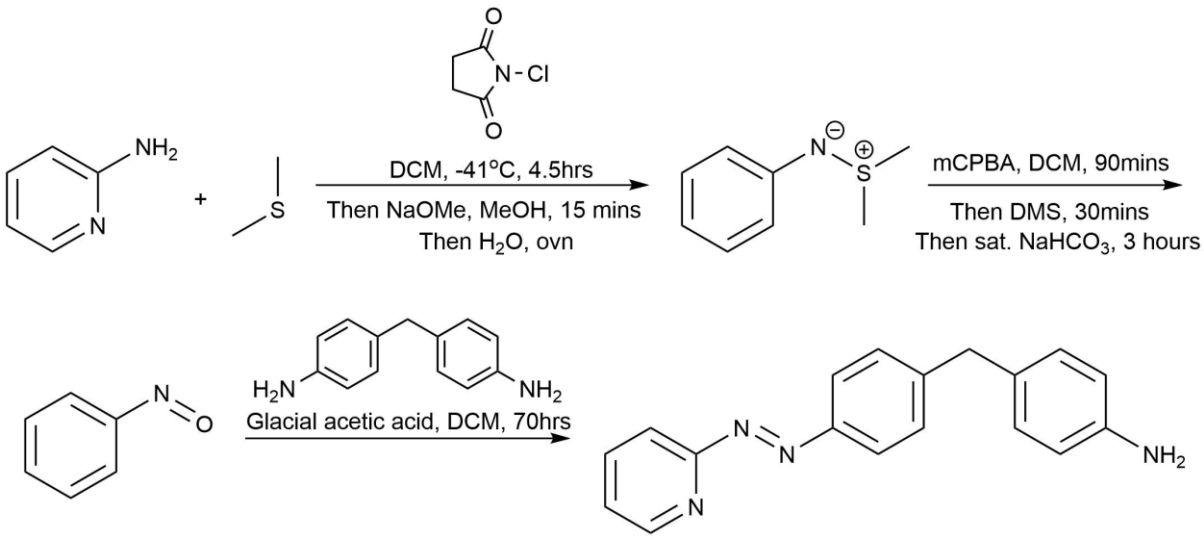
2.3.1. Introduction

Although very similar to the parent ligand, the half ligand has an azo bond instead of an imine bond between the pyridine and central phenyl groups. The azo bond is non-polar so it is less likely to be degraded by hydrolysis which is an issue with imine based ligands. The imine nitrogen of the parent ligand is involved in coordination to the metal ions, so changing the imine bond to an azo bond also alters the metal binding site of the half ligand.

Most importantly the half ligand still has one amine group available for imine bond formation. This combined with the increased bond strength, resilience again hydrolysis and slight change in binding site make the half ligand a useful precursor for the production of novel ligands.

2.3.2.**Synthesis**

The half ligand has been synthesised before by previous group members with varying degrees of success[80–82].



Scheme 1: Synthesis of the half ligand.

In order to synthesise the half ligand, 2-aminopyridine is reacted with dimethyl sulphide (DMS) in dichloromethane (DCM) to produce S,S-dimethyl-N-(2-pyridyl)sulphilimine. The sulphilimine is then oxidised to 2-nitrosopyridine by meta-chloroperoxybenzoic acid (mCPBA). Finally, the 4,4'-methylenedianiline spacer is added to the 2-nitrosopyridine with a small amount of glacial acetic acid to produce the half ligand (see Scheme 1).

The S,S-dimethyl-(N-pyridyl)sulphilimine is produced in relatively high yield (77%), however some of the starting material (2-aminopyridine) peaks are clearly visible in the NMR spectra with an integration for each proton roughly a quarter of the integration of the product integration.

The synthesis of the 2-nitropyridine is a much lower yielding reaction (34%), especially after recrystallization from hot ethanol (19%), which is an important

purification step to prevent any unwanted impurities being carried into the final stage of the synthesis, such as 2-aminopyridine which will readily react with the methylenedianiline to produce the unwanted parent ligand. The reduction of impurities is clearly visible when comparing between the baseline of the crude and recrystallized product NMR (see Figure 25).

The ^{1}H NMR also reveals that two species with very similar coupling are present. The peaks have been assigned to each species dependent on their integrations. Due to the similar coupling patterns it has been suggested that the two sets of peaks are caused by two tautomeric isomers of 2-nitrosopyridine[80].

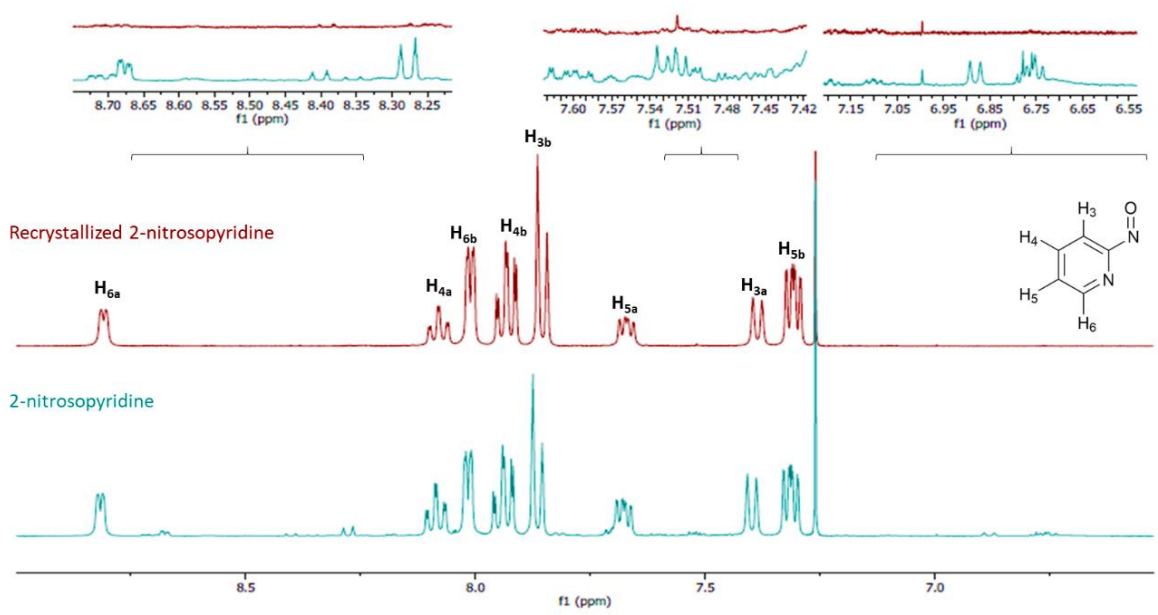


Figure 25: NMR (400 MHz, chloroform-d) of 2-nitrosopyridine product without recrystallising (blue spectra) and after recrystallising (red spectra). Insets (at the top of top of the spectra) are to demonstrate how impurities in the baseline are significantly reduced. The two tautomeric species are assigned (a) and (b) and are distinguished by their integrations.

The final stage of the synthesis requires a 1:4 ratio of 2-nitrosopyridine to 4,4'-methylenedianiline which is double the stoichiometric ratio of 1:2 nitrosopyridine to methylenedianiline. Excess 4,4'-methylenedianiline is necessary to minimise the amount of full ligand produced by the reaction but also means that there is a large amount of the

starting material in the product. Both methylenedianiline and the half ligand are primary amines which, despite their difference in size, makes them very difficult to separate. The amine groups would hydrogen bond to the alcohol groups in the stationary phase of a silica column, preventing the products from passing through the column.

A separation was therefore attempted using a reverse phase column. In the C18 silica used for a reverse phase column, the alcohol groups are replaced with long alkyl chains, 18 carbons in length. This lowers the affinity for the primary amines and allowing them to be eluted with a 3:1 water to acetonitrile solution. Unfortunately, this method gave varying degrees of success and never managed to completely remove the 4,4'-methylenedianiline due to a considerable amount of streaking causing the spacer and half ligand bands to overlap in the column.

Alternative solvent systems such as dichloromethane and 3:1 water to methanol were also attempted to see if the separation could be improved however it seemed that acetonitrile still gave the best separation. Decreasing the amount of sample being separated increased the purity of the final product, but the mass of product yielded is often very low. For this reason, the half ligand could be isolated by preparative reverse phase high pressure liquid chromatography (prep-RP HPLC). The UV visible spectrum at 265 nm wavelength was used to identify the products as they were eluted from the column (see Figure 26). The second, third and fourth peaks were analysed by mass spectrometry, allowing each to be identified as 4,4'methylenedianiline, the half ligand and the full azo ligand respectively (see Figure 27).

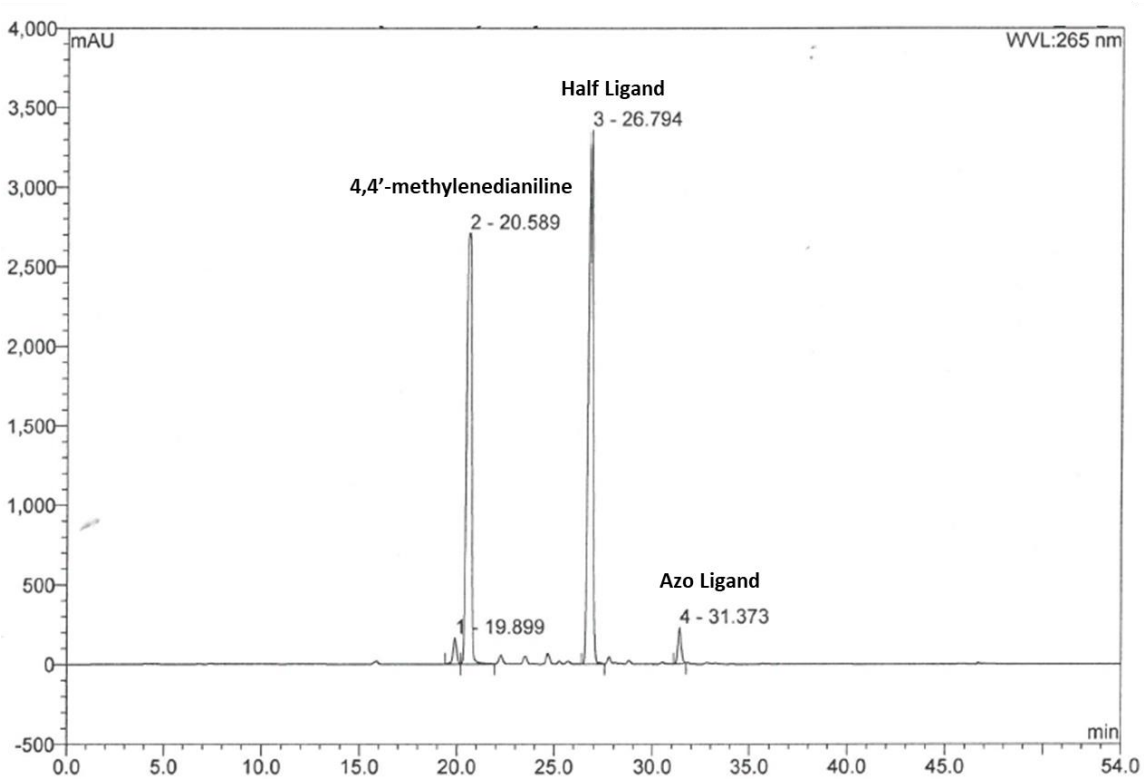


Figure 26: Absorbance of light at 265 nm wavelength as products are eluted from reverse phase column. RP HPLC was completed by Callum Campbell.

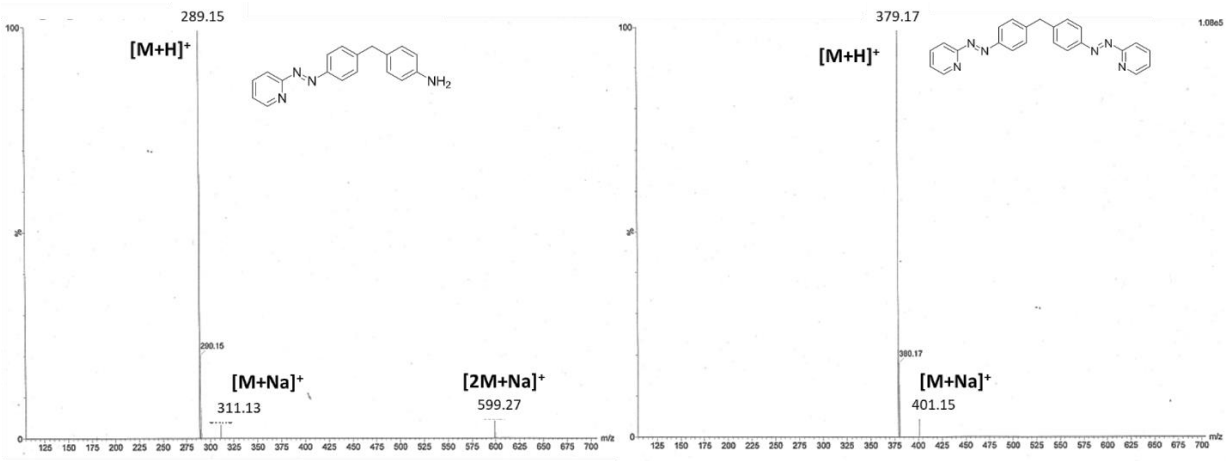


Figure 27: Mass spectra (both TOF MS ES+) of the half ligand (a) and the azo ligand (b) separated by prep-RP HPLC.

As a manual reverse phase column is a very slow method of purification with a low product yield per column, a different method was attempted. The crude product was boiled in water while acetonitrile was added until the half ligand had dissolved and then

the solution was chilled to precipitate. This gave a cleaner ^{1}H NMR spectrum without methylenedianiline present, however there still appeared to be a second species present as small peaks with regular integration values were observed in the baseline (see Figure 28 and Figure 29).

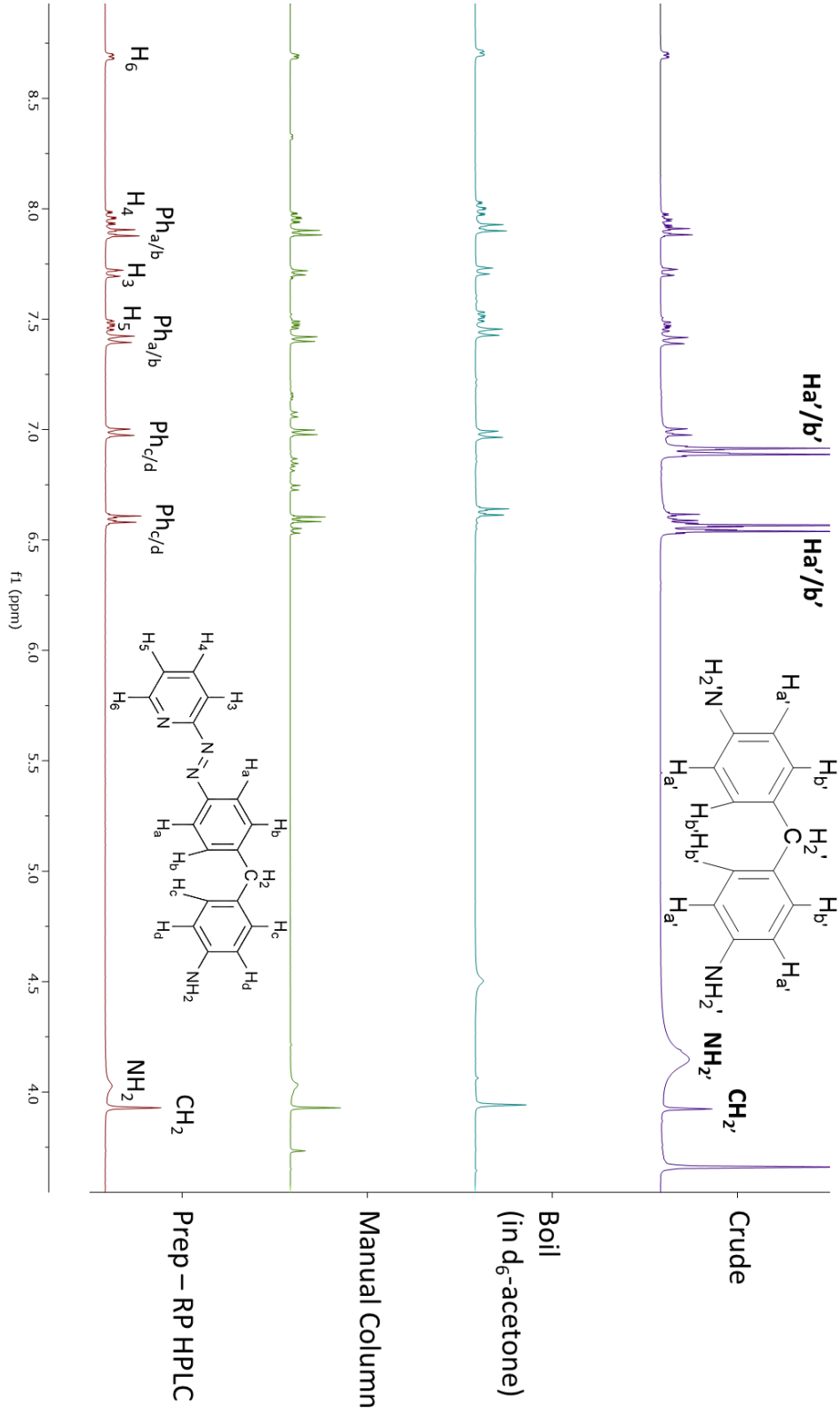


Figure 28: [1]H NMR of the half ligand as a crude product (purple – 300 MHz), after boiling in acetonitrile (blue – 300 MHz), after a manual reverse phase column (green – 400 MHz) and when isolated by prep-HPLC (maroon – 300 MHz). Samples were dissolved in d_3-acetonitrile unless stated otherwise.

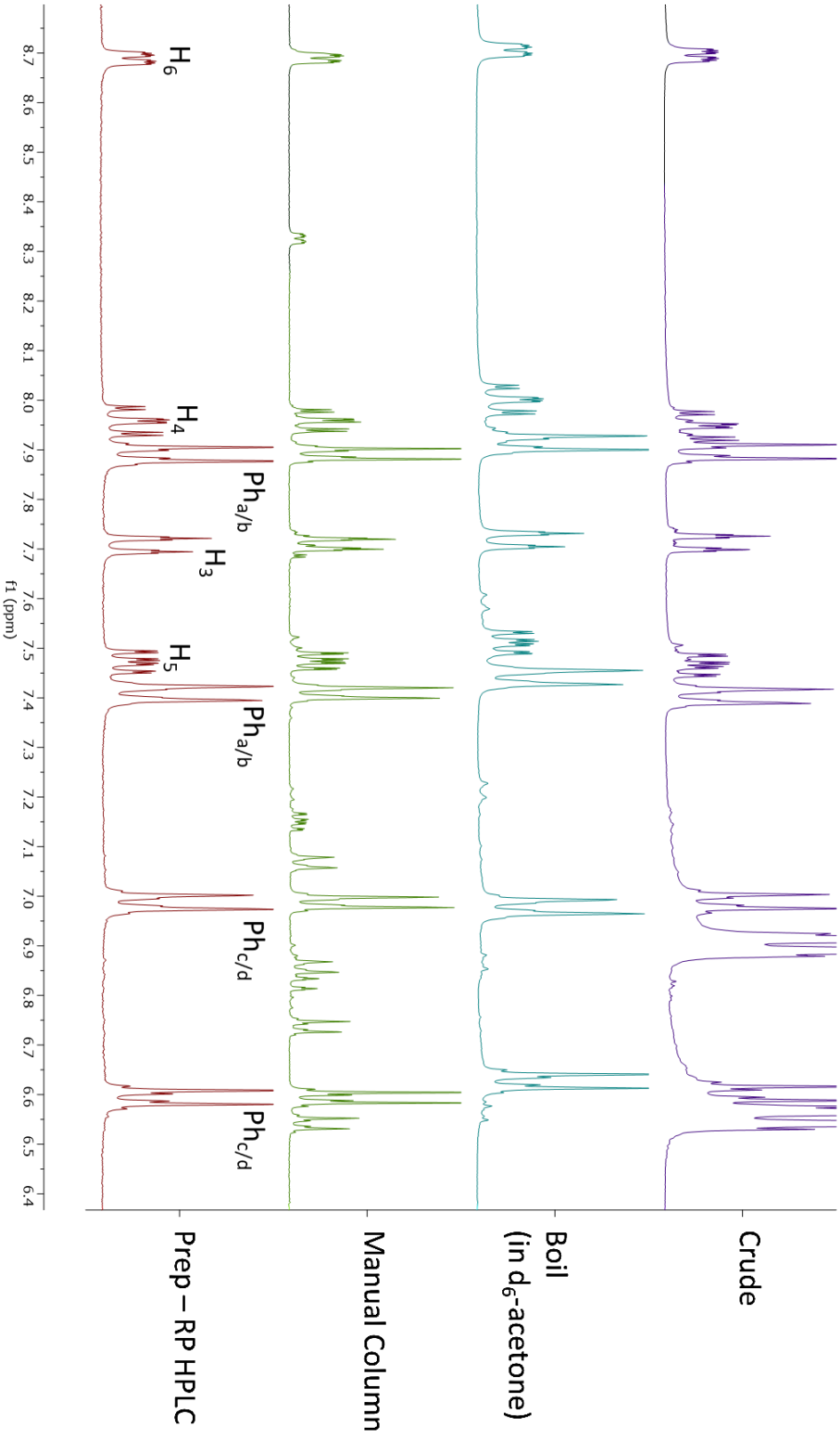


Figure 29: [1]H NMR between 6.3 and 8.8 ppm of the half ligand as a crude product (purple – 300 MHz), after boiling in acetonitrile (blue – 300 MHz), after a manual reverse phase column (green – 400 MHz) and when isolated by prep-HPLC (maroon – 300 MHz). Samples were dissolved in d_3-acetonitrile unless stated otherwise.

From this data it is clear that, despite the low yield of product, HPLC is the best method for the isolation of the half ligand. The boiling method may also be useful to get remove a majority of the half ligand before using a column to increase the yield.

2.4.Trinuclear Complex Synthesis

To synthesise a ligand with two three way junction binding sites, a suitable linker had to be chosen. 2,3-Butanedione was selected as it is a small, rigid molecule with two aldehyde sites available for imine formation, separated by a single C-C bond. This separation is significant as the imine nitrogens in the linker would be separated by three conjugated bonds which is very similar to the imine-pyridine binding sites at the end of the ligand (Figure 30). This synthesis has been attempted by Laura Childs[80] without success.

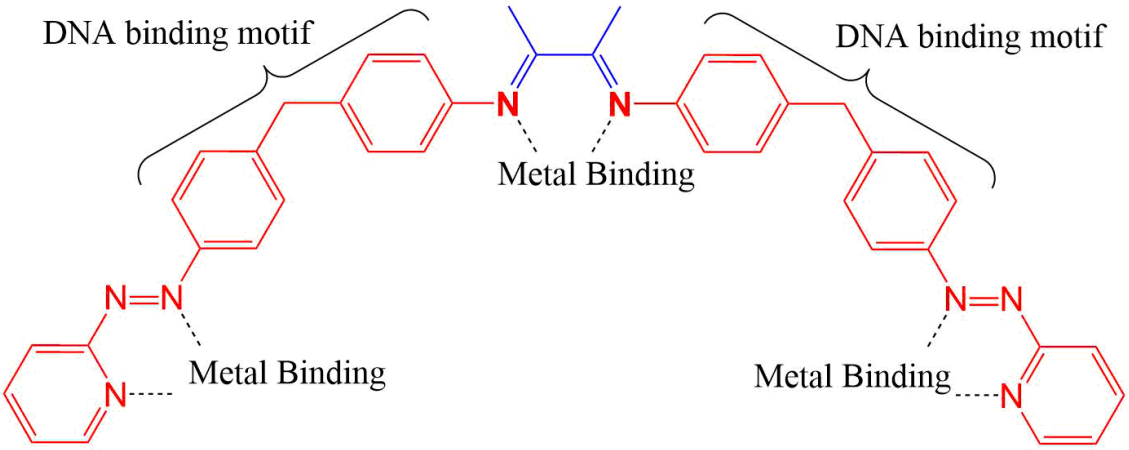
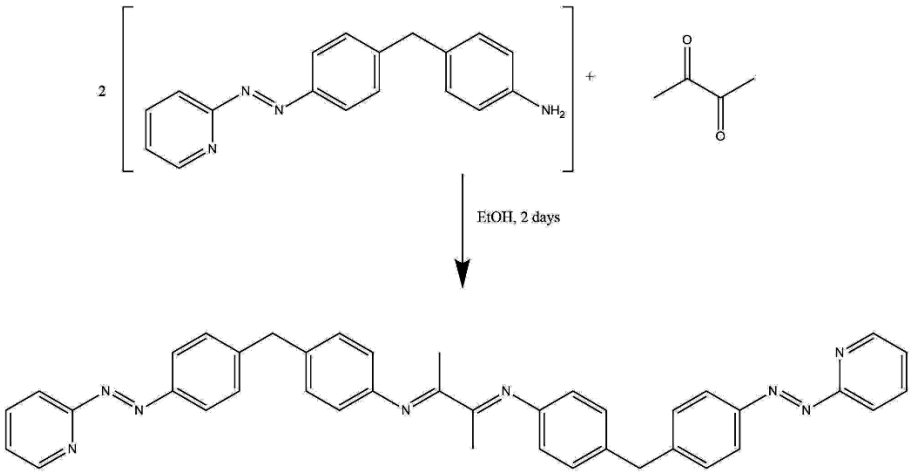


Figure 30: The key motifs in the design of the trinuclear ligand. The half ligand is highlighted in red whilst the 2,3-butanedione linker is highlighted in blue. Note that the metal binding nitrogens are all seperated by a similar distance.

The synthesis of the trinuclear ligand was attempted via two routes: a two-step synthesis in which the ligand was produced first and then complexed to the metal, and a

one-pot synthesis, in which the metal is used to aid the synthesis of the ligand by templating.



Scheme 2: Synthesis of the trinuclear ligand.

In an initial attempt to synthesise the ligand, the half ligand and 2,3-butanedione were dissolved in ethanol and stirred for two days (see Scheme 2). An NMR of the brown oil product suggested that the half ligand was left mostly unreacted and many peaks appeared between 2.2-2.4 ppm and 1.2-1.8 ppm.

An excess of 2,3-butanedione was added, a longer time was given for the reaction to occur, and later the solvent was switched to dry dichloromethane with the use of molecular sieves in order to push the reaction to completion. Under these conditions, the product NMR (see Figure 31) appeared to show the appearance of new phenyl peaks with a similar integration to the unreacted half ligand peaks. The pyridine peaks have double the expected integration value, making the integrations very similar to the phenyl peaks, and the coupling patterns are much less well defined with extra peaks appearing. This could indicate some of the trinuclear ligand is formed as the pyridine peaks would not be expected to shift compared to the half ligand, while the phenyl peaks would be the most

affected. The multiple peaks below 4 ppm prevent the methyl peaks being identified, which would give strong evidence of trinuclear ligand formation.

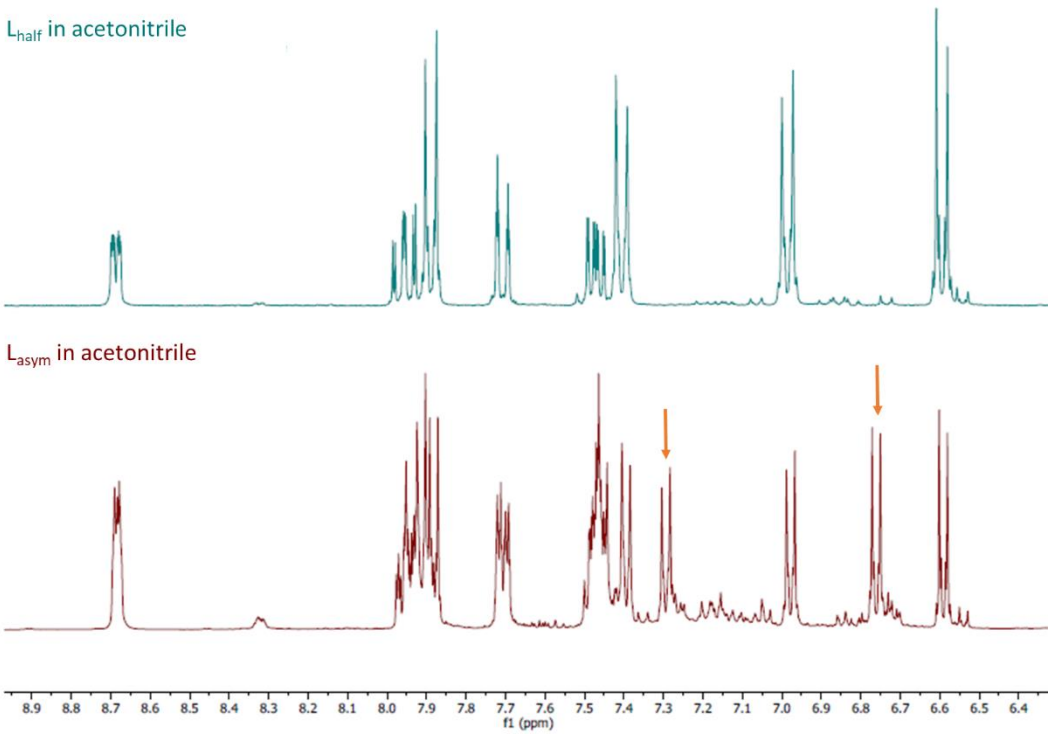


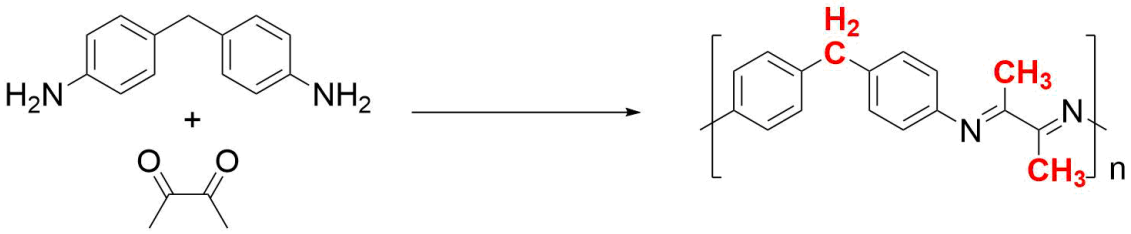
Figure 31: ^{1}H NMR (300 MHz, d_3-acetonitrile) of the half ligand (L_{half}) and the asymmetric ligand product when reacted in dry solvent using molecular sieves (L_{asym}).

As the mass spectra also showed that the half ligand was present, it was concluded that a one-pot templating method should be attempted. Templating is the use of a molecule, ion or counterion to direct the synthesis of a certain product. This method is often used in supramolecular chemistry to aid the synthesis of complex and knotted molecules which would not otherwise organise themselves in the desired orientation or placement[75,83]. Iron (II) chloride tetrahydrate was added to supply a metal ion template for the synthesis of the full trinuclear ligand.

Although there was clearly still half ligand present, there was a clear peak in the mass spectrum at *m/z* = 627 suggests the presence of the full ligand. None of the ligand was found complexed to the metal ion either in a fully or partially formed helicate.

The same reaction was attempted with tetrakis(acetonitrile) copper(I) tetrafluoroborate as copper(I) prefers tetrahedral geometry and the formation of a two-stranded helicate may be simpler than the formation of the triple stranded equivalent, however no further evidence was gathered for the formation of the ligand or the helicate as the low yield prevented a thorough analysis of the product.

Interestingly, the templating method did not show the multiple downfield NMR peaks seen in the ligand synthesis reaction. This could be due to the formation of polymeric species of different lengths caused by the reaction between excess 4,4'-methylenedianiline still present in the half ligand sample and 2,3-butanedione. Both the spacer unit and the butanedione have alkyl groups which could correspond to the signals observed at the low shifts (see Scheme 3).



Scheme 3: The reaction between 4,4'-methylenedianiline and 2,3-butanedione to form polymers of various lengths which could cause multiple peaks in the downfield region of the NMR spectrum.

In conclusion, the synthesis of the trinuclear ligand was achieved by a templating method using iron(II) chloride tetrahydrate, however the isolation of the ligand and the conditions for complexing require optimisation.

2.5.Asymmetric Ligand Synthesis

Another approach to producing novel ligands is to synthesise an asymmetric ligand from the half ligand and a suitable aldehyde. Imidazolecarboxaldehyde was chosen

as imidazoles had been worked with previously within the Hannon group and have been shown to support cylinder formation[84,85]. Imidazoles are heteroatomic, aromatic rings with a nitrogen atom which is able to act as an electron donor and therefore coordinate to a transition metal centre. Unlike pyridine rings, imidazoles also possess a second nitrogen atom which is available for hydrogen bonding and can react as a Lewis base. This will be discussed further in the next chapter.

An initial reaction was attempted by dissolving the half ligand and 2-imidazole caboxaldehyde in methanol in a 1:1 ratio. Acetic acid was added and the solution stirred at 70ºC before a dark brown solid was isolated.

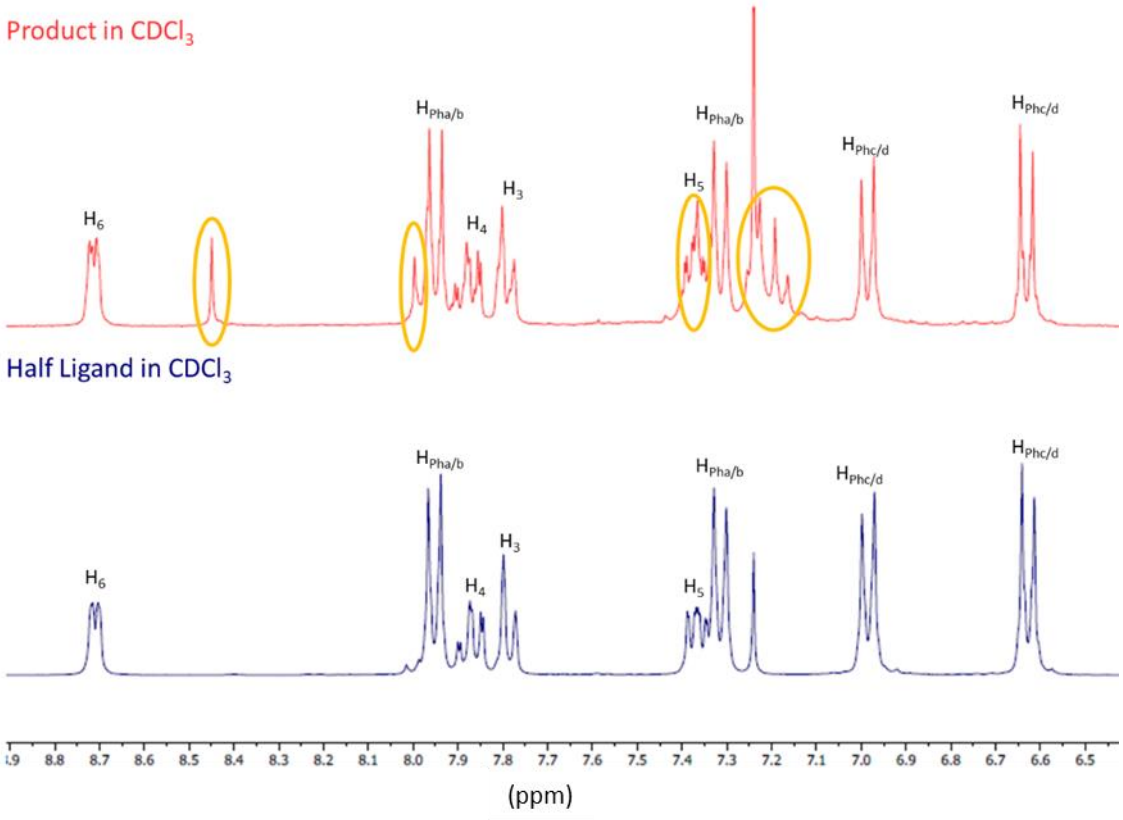


Figure 32: The ^{1}H NMR (300HMz) spectra of the initial product (top – pink) and the half ligand starting material (bottom – blue). Both samples were dissolved in d-chloroform. Peaks appear in the product spectra only are circled.

An initial comparison between the NMR of the dark brown product and the half ligand spectrum (see Figure 32) suggested that the reaction had not occurred as most of the peaks in the product were an exact match to the peaks in the half ligand spectrum. After closer inspection, a few novel peaks could be seen in the product spectrum, suggesting a small amount of product may be present but is being largely hidden by the excess of half ligand present in the sample.

As a consequence of this, the reaction was repeated using dry methanol and heated overnight under reflux in an attempt to push the reaction to completion. Filtering the solution and removing the solvent gave a green powder and an orange powder, which was suspected to be the half ligand. The powders were heated overnight under reflux again in dry dichloromethane to make sure all of the half ligand had reacted. When the solvent was removed, only a green powder remained giving a 47.6% yield.

The mass spectrum of the green powder product revealed a small peak at *m/z* = 289, suggesting the presence of the half ligand, however the remaining larger peaks (*m/z* = 367, 389, 733, 755, 1010) all correspond to fully reacted ligand species.

The ^{1}H NMR and ^{13}C NMR were then collected in chloroform.

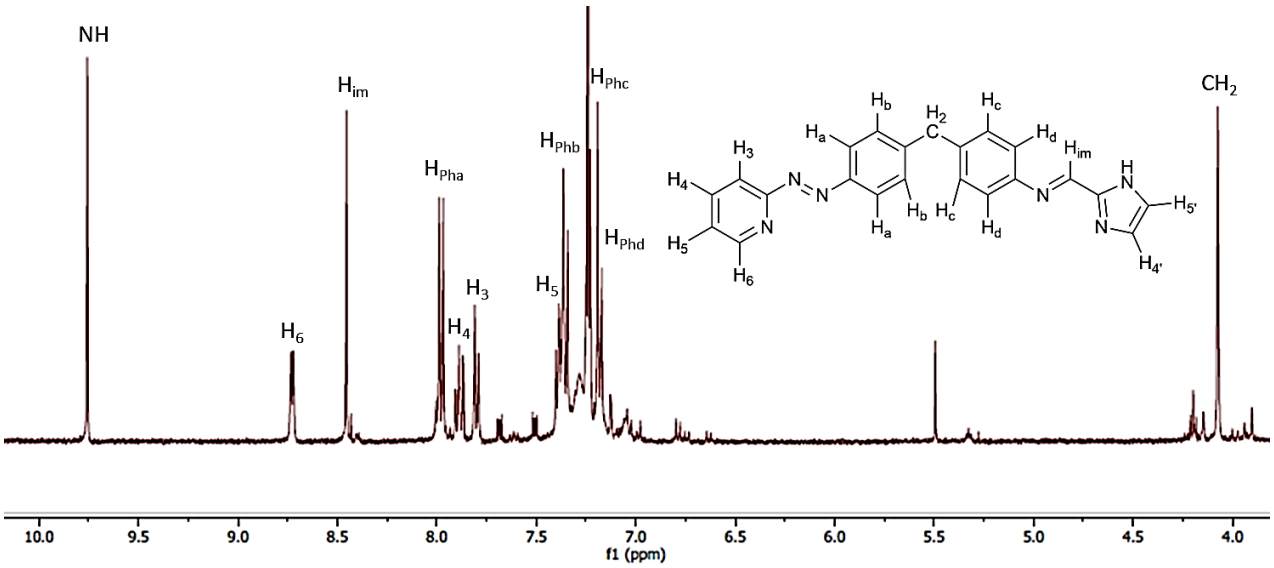


Figure 33: ^{1}H NMR (400 MHz) of the asymmetric ligand in d_3-chloroform.

The ^{1}H NMR suggests that the asymmetric ligand is the main product of the reaction (see Figure 33). There are clearly some impurities, such as excess half ligand as observed in the mass spectrum or methylenedianiline which was probably introduced with the half ligand. The integrations of these impurity peaks is much lower (less than 20%) than the product species. The imidazole protons were not visible in the spectrum, however the integration of H_5 (3H) suggests that two more protons are hidden beneath the H_5 and HPh_b signals. The imidazole protons ($H_{4'}$ and $H_{5'}$) have shifts between 7.15 and 7.35 ppm in the 2-imidazole ligand (L_{im2}, see Figure 35) so it is very likely that the imidazole protons are underneath other signals within a similar region.

A correlated spectroscopy (COSY) NMR was also attempted to see if any further signals could be observed to indicate the presence of the imidazole protons. Unfortunately, the strong coupling between the phenyl and pyridine protons in this region and the presence of strongly coupled impurities meant that no further peaks could be assigned. The COSY spectrum did agree with the assignments of the proton NMR, and

the intensity of the coupling of CH_2 to the neighbouring phenyl protons helped differentiate the phenyl peaks.

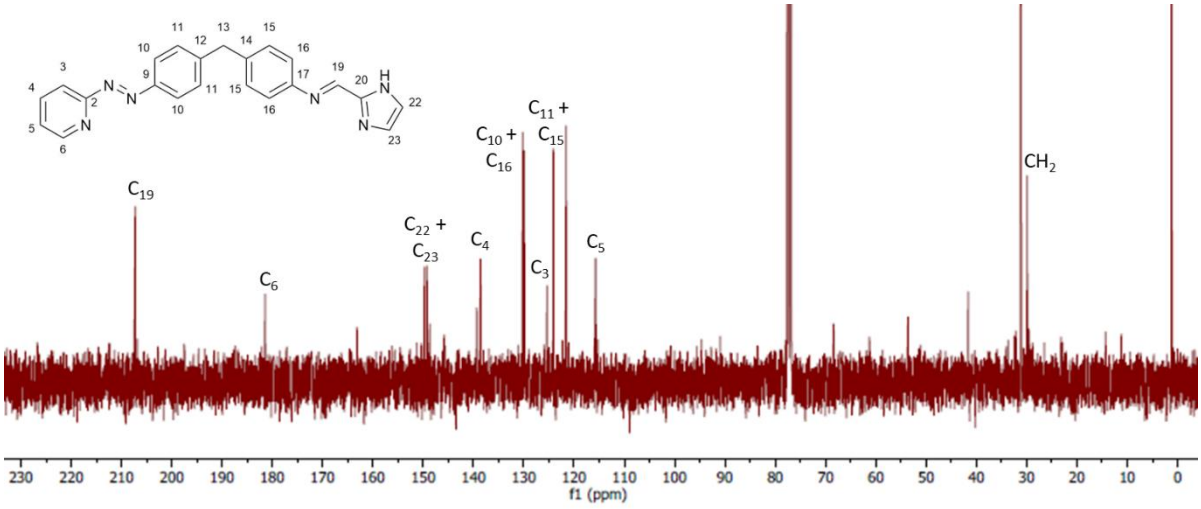


Figure 34: ^{13}C NMR (400 MHz) of the product in d-chloroform.

The ^{13}C NMR was also recorded to see if the imidazole carbons could be identified. All the peaks which have a large enough integration have been assigned (see Figure 34). Excluding the methylene carbon, all the peaks appear within the 110-210 ppm range as is expected for sp^2 hybridised carbons. There are three sets of two peaks which have been assigned to protons that are expected to be in a similar environment (C_{22} and C_{23}, C_{10} and C_{16}, and C_{11} and C_{15}). The imidazole carbons were identified as one of these sets as they have half the integration value of the other peaks. A few very small peaks can be seen in the spectrum and are likely to correspond to the quaternary carbons (H_2, H_9, H_{12}, H_{14}, H_{17} and H_{20}) but can not be assigned as the signal is too weak. Although a heteronuclear single quantum coherence (HSQC) and a heteronuclear multiple bond correlation (HMBC) were taken, the signal is too weak in both to provide further evidence for these assignments.

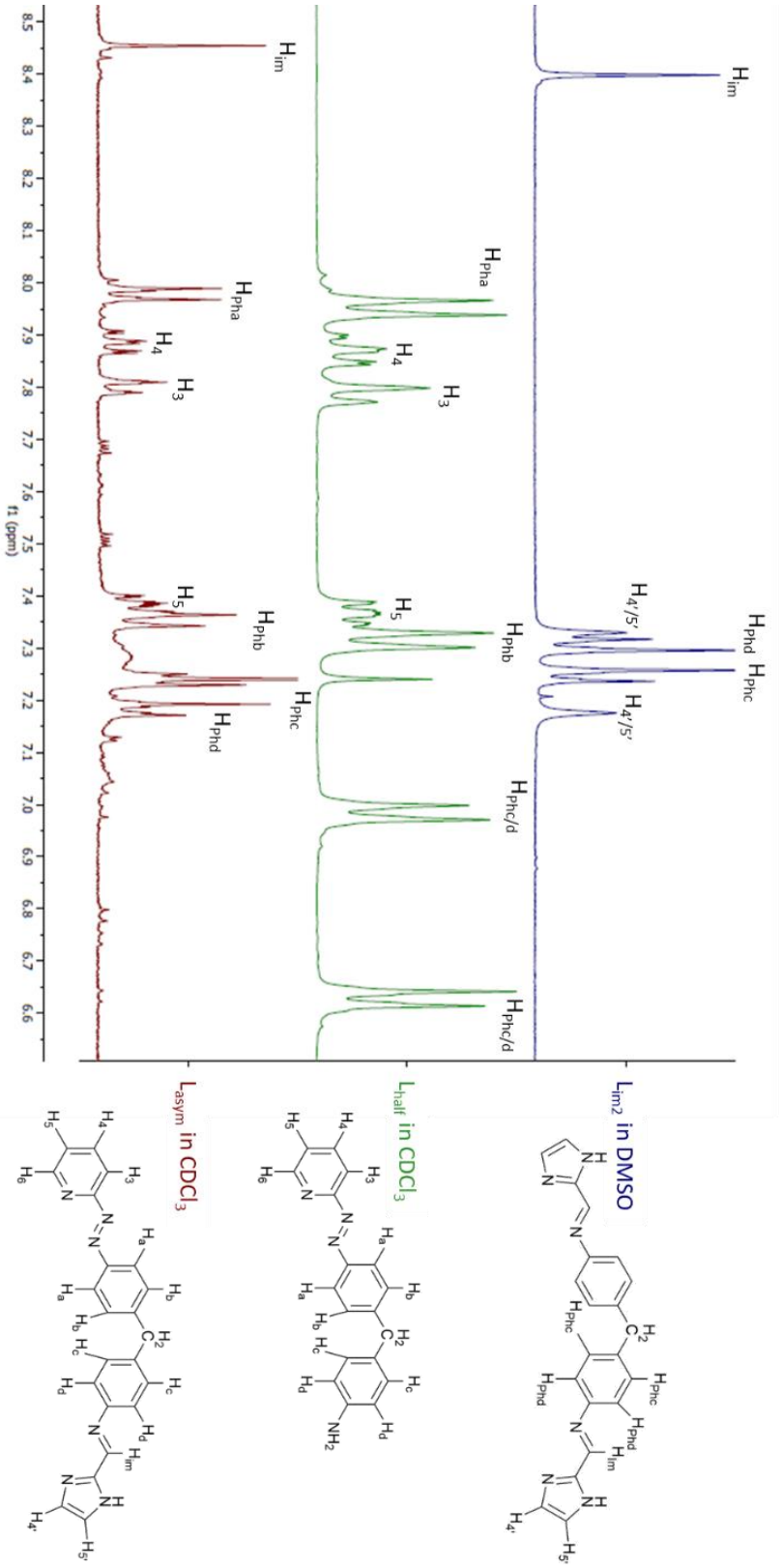


Figure 35: [1]H NMR (400 MHz) of the full 2-imidazole ligand (L_{im2}) in d_6-dimethyl sulphoxide (top, blue), the half ligand (L_{half}) in d-chloroform (middle, green) and the asymmetric ligand (L_{aysm}) in d-chloroform (bottom, red). Note that the assignment of L_{im2} is different here to later chapters in order to simplify the comparison.

From the ^{1}H NMR, it is clear to see that the starting materials are no longer the predominant species in solution (see Figure 35). While the pyridine and phenyl peaks have a similar splitting pattern to the peaks of the half ligand, each peak is slightly deshielded in the asymmetric ligand compared to the half ligand. The phenyl protons closest to the imidazole group in the asymmetric ligand are particularly shifted upfield and are much closer to where the phenyl peaks are expected in the imidazole ligand.

2.5.1. **Attempted Asymmetric Copper Complex Synthesis**

Once the asymmetric ligand had been synthesised, an attempt to complex the ligand was made. Copper(I) was chosen for this as it is expected to form double stranded di-nuclear helicates, rather than the triple stranded helicates formed previously.

In order to achieve this, the asymmetric ligand and tetrakis(acetonitrile) copper(I) tetrafluoroborate were dissolved in dry methanol and heated under reflux for 2 hours 30 minutes. The solution was cooled at room temperature before being chilled for an hour at 4^{o}C. A dark, black powder was collected by vacuum filtration.

The product was insoluble in most common solvents including methanol, acetonitrile and chloroform. When dissolved in DMSO a slight green colour was observed, however DMSO is known to complex metals and it is likely that the green colour observed was from the ligand once released from the copper ion. When an NMR was attempted, no signal was observed.

The mass spectrum for this compound shows one predominant peak at *m/z* = 890 with a 2+ charge (see Figure 36). The mass and isotope profile for this peak appears to match a complex with chemical formula $Cu_5C_{88}H_{68}N_{24}^{2+}$ (see Figure 37).

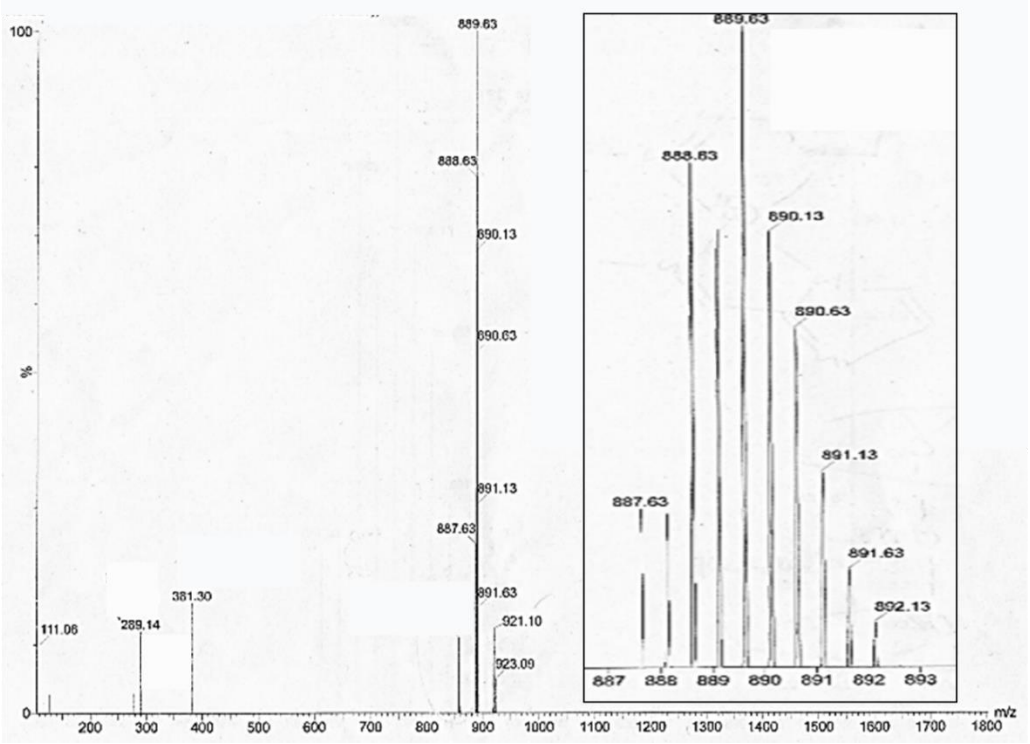


Figure 36: Mass spectrum (TOF MS ES+) of copper asymmetric copper complex synthesis.

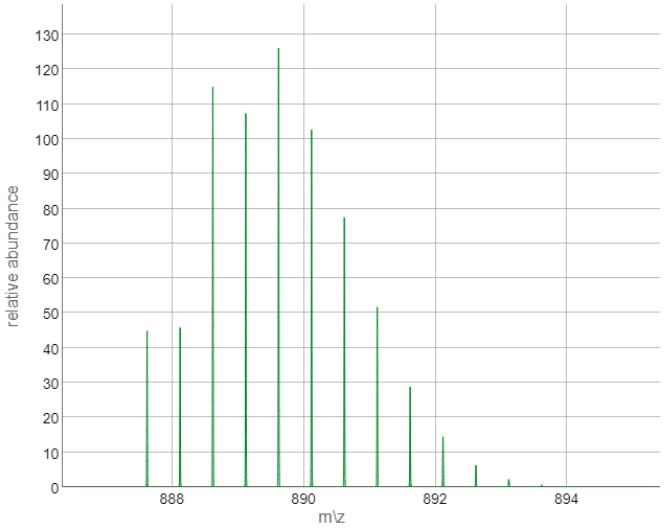


Figure 37: Calculated mass spectrum profile of $Cu_5C_{88}H_{67}N_{24}^{2+}$. Created usimg enviPat Web 2.4 isotope pattern calculator, Reference 86.

Instead of forming a double stranded helicate, the formula suggests that four asymmetric ligands have arranged themselves the head of one ligand and the tail of another binding one of four Cu(I) centres, using the nitrogen binding domains which

were expected to form the helicate. The remaining nitrogen moieties of the imidazole rings are deprotonated and bind the final metal ion, Cu(II) (see Figure 38).

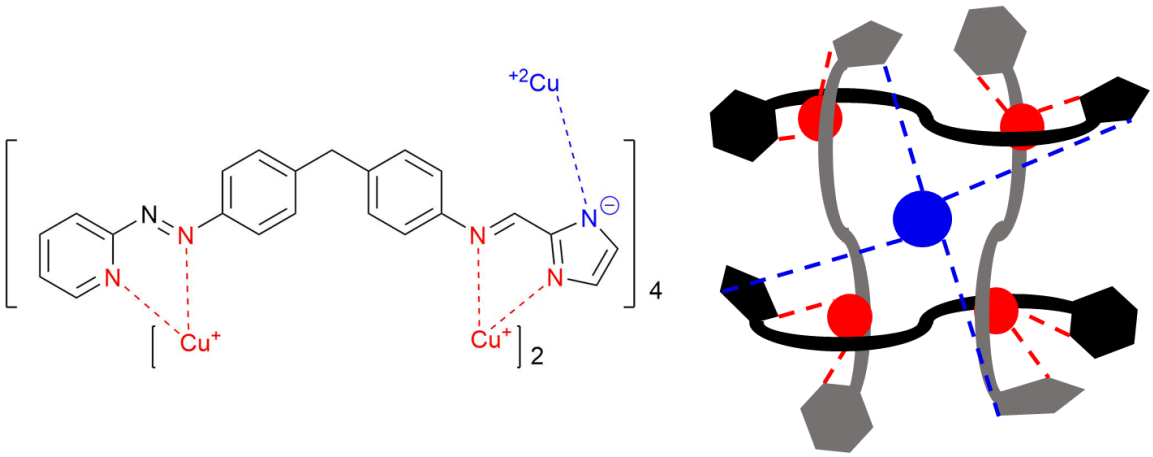


Figure 38: Showing the possible structure of the suspected Cu_5L_4 complex.

2.6.Conclusions and Further Work

The half ligand was synthesised and isolated with a high purity, however isolation by a column is time consuming and limits the amount of crude product which can be processed at once. Future work should be directed at improving the purification process, potentially using a faster purification method to remove a significant amount of the half ligand before isolation by column, as this could increase the yield from the column. It is important that the half ligand is pure before use as the methylenedianiline impurity has the same reactivity as the half ligand and is therefore able to interrupt any attempted reactions.

The trinuclear ligand showed signs of formation although this reaction could not be pushed to completion meaning a lot of the half ligand was still visible in the product. A potential cause of this is the presence of methylenedianiline in the half ligand sample which formed polymeric species with the 2,3-butanedione. If the purity of the half ligand

could be increased further, synthesis of the trinuclear ligand should be attempted again in dry solvent, with molecular sieves, excess butanedione and under an inert atmosphere. This may be sufficient to push the reaction to completion.

After increasing the reaction time and using dichloromethane as the solvent, the asymmetric ligand was synthesised. In future, this ligand should be purified before attempting to complex to a metal, and complexing to different metals should be investigated. If possible, a crystal structure of the copper complex should be obtained in order to indicate the structure of the expected five copper ion complex observed by mass spectrometry.

Chapter 3:

POST ASSEMBLY MODIFICATION AND ROTAXANATION OF A SUPRAMOLECULAR HELICATE

3.1. Introduction

Within the Hannon Group a range of analogues of the parent supramolecular cylinder have been synthesised and characterised, each with their own unique properties and purposes. The DNA binding of these cylinders has been investigated and used in order to bind more complex nucleic acid architectures, such as the ruthenium parent cylinder binding a DNA tetrahedron[87]. In this chapter, the synthesis of a higher order supramolecular structure, a rotaxane, is achieved using an analogue of the parent cylinder and a cucurbit[10]uril.

This introduction includes a brief summary of the cucurbituril family with a focus on the less well studied cucurbit[10]uril. Rotaxane synthesis using cucurbiturils is also discussed in more depth to add context.

3.1.1. Cucurbiturils

Cucurbiturils are a family of macrocycles with the general formula $[C_6H_6O_2N_4]_n$ (n = 5, 6, 7, 8, 10, 13, 14, or 15) (see Figure 39)[88–93]. Cucurbit[6]uril (CB6) was the first of this family to be identified as the product of the reaction between glycoluril and formaldehyde, and the host-guest properties of these macrocyclic structures became their main point of interest[90].

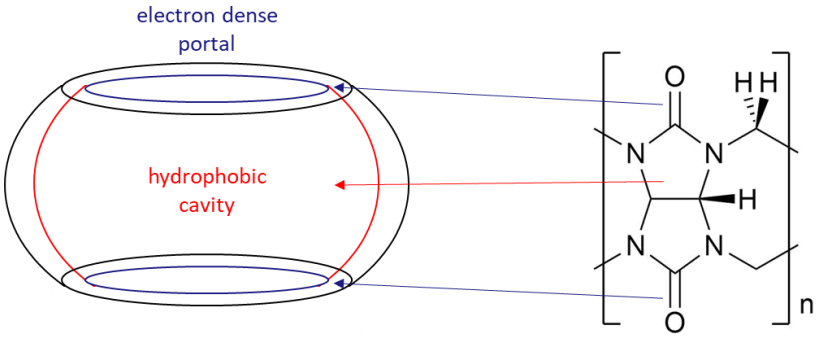


Figure 39: General structure of a cucurbituril macrocycle unit (right) and the 3D structure of the cucurbituril family (left).

A cucurbituril cavity contains no functional groups and has a very low polarizability creating an ideal environment for non-covalent binding of hydrophobic molecules. Insertion of guest ligands also releases water molecules trapped inside the cavity which provides both an enthalpic and entropic driving force for the hydrophobic binding[94,95].

Two portals positioned either side of the central cavity provide an entry and exit point for guest species. The portals are lined by electron dense ketone groups that provide an additional opportunity for both hydrogen bonding and ion-dipole interactions with cationic species[94,95].

Because of their unique structure, cucurbiturils are able to bind selectively depending on the size and shape of a potential guest molecule. A guest molecule may be too large to enter the cavity of the rigid cucubituril, or may be too small for favourable binding interactions. Studies by Mock and Shih also demonstrated that the distance between two ammonium ions linked by a hydrophobic alkyl chain can also affect whether binding occurs, as the cations must be held at a distance where binding to both portals is possible[96,97].

3.1.2.**CB10**

Cucurbit[10]uril (CB10) was not discovered until 2002 by Day *et. al.* who characterised CB10 as part of a new supramolecular structure termed a gyroscane[98]. This structure consisted of a cucurbit[5]uril (CB5) ring in the cavity of a CB10 ring with an angle of 64° between the axes of the two molecules. In 2005, Liu *et. al.* were able to isolate CB10 by exchanging the CB5 guest molecule for an alternative organic molecule (see Figure 40(**1**)) which could be removed from the CB10 cavity by heating in acetic anhydride and washing[99]. A number of potential guest molecules were later screened to increase the efficiency of CB10 isolation. This was achieved by using a guest molecule (Figure 40(**2**)) which bound CB10 and formed a water insoluble precipitate that was then collected and washed to remove the guest molecule[100].

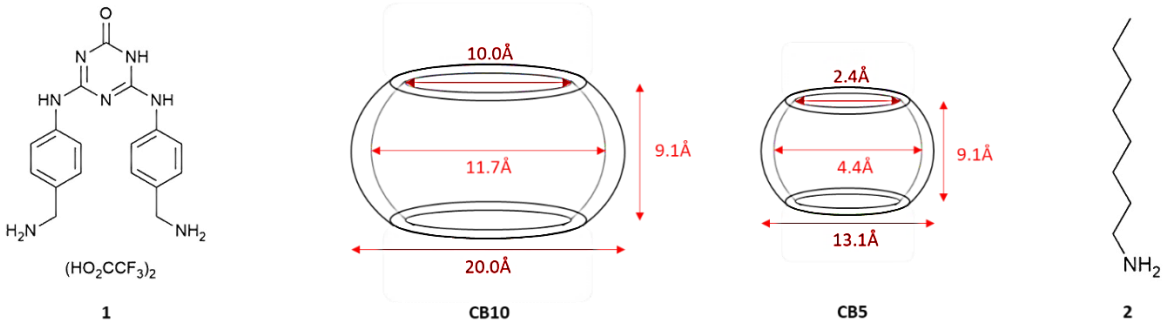


Figure 40: The dimensions of CB10 and CB5 (values taken from references[93,94]) and two guest molecules able to displace CB5 from CB10[99,100].

CB10 has the largest cavity of the cucurbituril family, approximately 870 Å^3. This increases the flexibility of the macrocycle when compared to other cucurbiturils enabling CB10 to encapsulate much larger guest molecules. As well as CB5, other macrocycles and other structures have been shown to fit inside the cavity of CB10[101].

3.1.3. **Cucurbituril based rotaxanes**

The binding and applications of CB6 and CB7 in particular have been extensively investigated. The cucurbiturils have been used to form a variety of supramolecular architectures, including molecular switches, grids and rotaxanes, and even nanostructures[102]. The internal cavity of the cucurbiturils have also been used in reaction catalysis[91,92,103].

The structure of interest for this study is the rotaxane, a supramolecular structure of at least two components that are physically but not chemically bound. The first of these components is an axle, which is a long, thin molecule with two bulky moieties at either end to give an overall dumbbell-like appearance. The second component, a macrocycle, surrounds the mid-section of the axis but can not fit over the bulky sections at either end so the macrocycle is mechanically trapped in place. These two types of component are often found in a 1:1 ratio (one ring around one axle) however there is no requirement for this, and examples can be found of rotaxanes with multiple rings[104].

In many cases the rotaxanes produced are specially designed in order to act as a molecular switch, a structure which can convert between two different configurations in response to a change in the environment. The two modes of non-covalent binding offered by cucurbiturils, hydrophobic and ion-dipole, are desirable for these systems as different binding modes can be activated or more favourable in certain conditions.

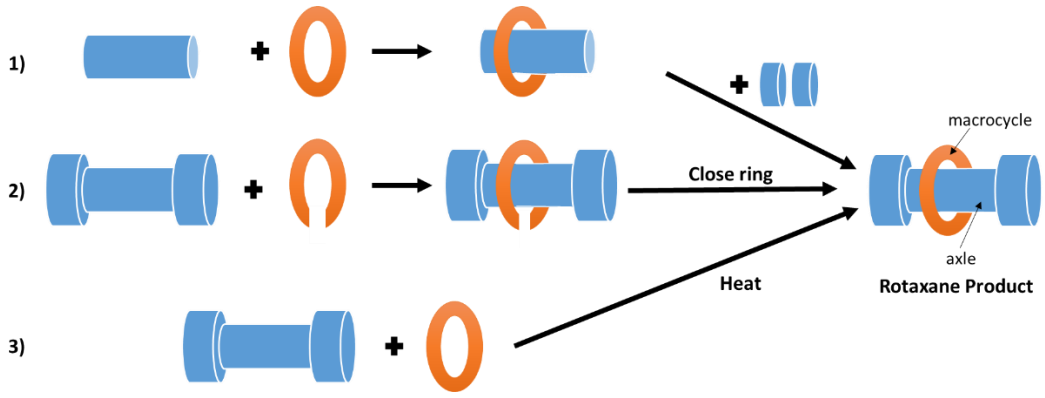


Figure 41: The main synthetic routes of rotaxane formation: 1) capping, 2) clipping, and 3) slippage. The general structure of a rotaxane is shown as the product.

There are three main synthetic routes for rotaxane formation (see Figure 41 and Figure 42)[105].

In the first approach, the long slim section of the axle can be passed through the centre of the macrocycle and the bulky groups are then attached to prevent the macrocycle falling off. This is referred to as the capping mechanism.

An example of this uses alkylammonium chains that are known for binding in cucurbiturils as the alkyl chain is able to bind the hydrophobic cavity while the ammonium cations can interact with the polar portal regions. By attaching such a chain to pyridine rings, pseudo-rotaxanes able to bind to ruthenium complexes have been achieved. As the metal complex is too large and highly charged to pass through the central hydrophobic cavity of the cucurbituril, the complexes act as stoppers and if they are attached to each end of the alkylammonium chain it can form a rotaxane. Given that some ruthenium complexes have been investigated as alternatives to platinum drugs, and the increased water solubility of cucurbiturils along with the protection they provide from the cellular environment could mean that such rotaxanes might find application in the pharmaceutical field[106].

In the second synthetic approach, if the axle molecule has already been synthesised with the bulky end groups, a rotaxane can be made by forming the ring structure around the centre of the axle. Cucurbiturils are usually isolated as a pre-formed macrocycle and then used so clipping is not often used. Wittenberg *et. al.* were the first to form a clipped rotaxane using a CB10 analogue (bis-nor-seco-cucurbituril) which is similar to CB10 but without two bridging methylene groups. This increases the flexibility of the macrocycle allowing a slight inward fold at the two missing bridge points. Using two alkylammonium chains with terminal t-butyl groups, a pseudorotaxane was formed with one guest molecule in each half of the macrocycle. The two halves of the macrocycle could then be closed around the guests molecules using 2-imidazolidone and formaldehyde. This forms two smaller linked cavities, each binding a guest molecule, which is too small for the t-butyl groups to slip through[107].

The final synthetic pathway, slipping, is used when both the macrocycle and the axle structures are already formed. By heating the two components together, the activation energy required to slip the macrocycle over the bulky group may be overcome, however when cooled to room temperature, the macrocycle no longer has the energy to slip off the axle, trapping it in the rotaxane formation.

Rotaxane synthesis by slippage was achieved by Yu *et. al.* using CB10. A central 4,4'-bipyridine unit was used to bind the hydrophobic cavity while the pyridinium cations were positioned at a distance which allowed both to engage in ion-dipole interactions with the portals. Either side of the central unit are two tetraphenylethylene stoppers which were identified as just small enough to allow the slippage mechanism to occur with only the largest cucurbituril, CB10. A methylated analogue was also synthesised but did not form the rotaxane species even after heating. Despite the guest molecule not showing any prior

fluorescent properties in DMSO, once encapsulated as the rotaxane a clear emission peak was observed at 558 nm[108].

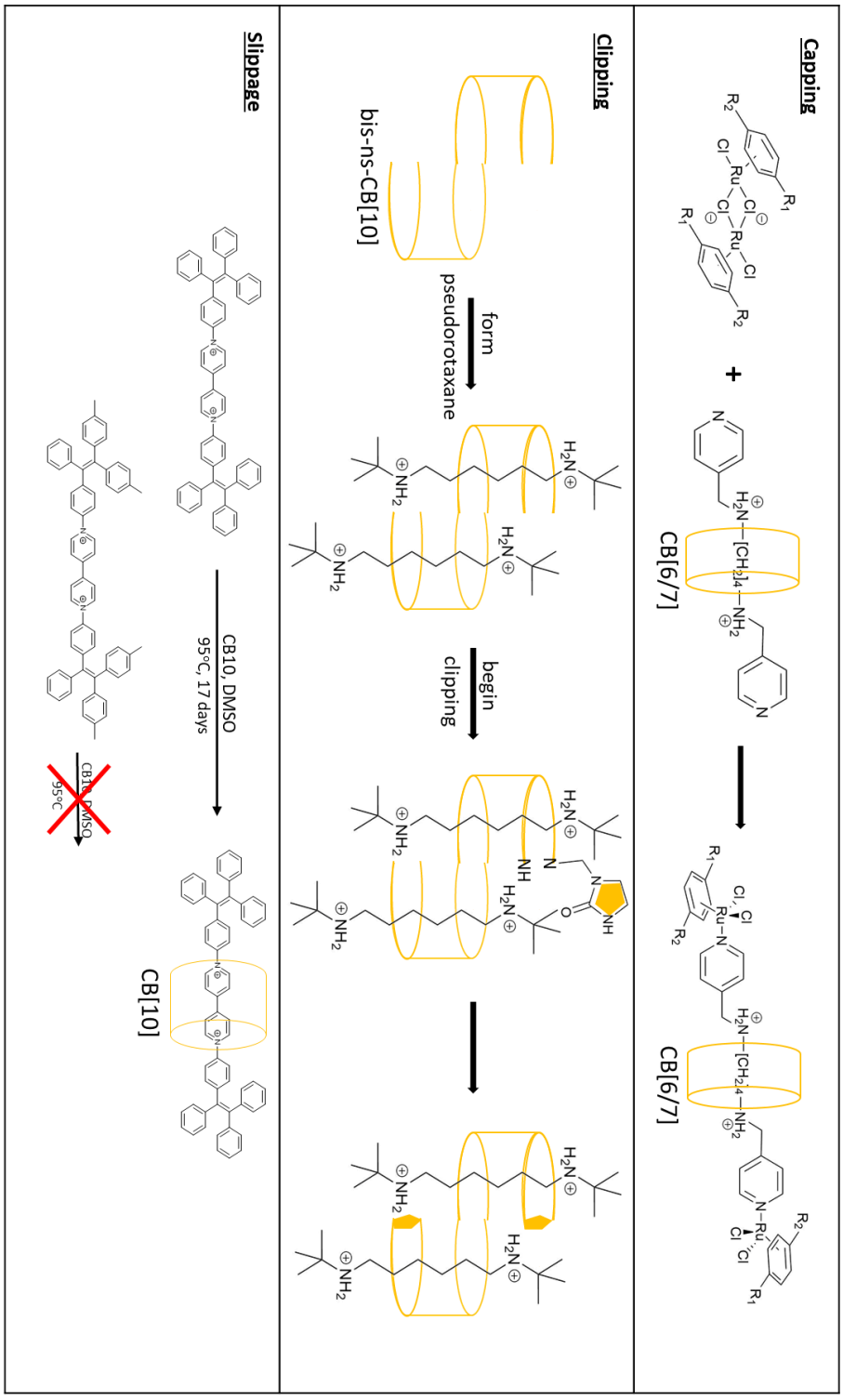


Figure 42: An example of each method of rotaxane synthesis using cucurbiturils. Images are inspired by figures from references 106, 107 and 108.

Finally, it is important to note that the ability of cucurbiturils to act as an alkyne-azide click chemistry catalyst has been utilised to form rotaxanes by an additional synthetic route. In some cases, the click reaction is used to add additional bulky groups to the end of the axle in a very similar manner to the capping mechanism seen previously[104], however in some cases the triazole ring is formed in the centre of the new axle molecule and the cucurbituril molecule traps itself between two bulky groups (see Figure 43)[109]. The formation of the axle from two "half-dumbbell" components seems significantly different for click reactions to be considered an alternative fourth route of rotaxane formation. (See reference[110] for rotaxane formation by copper catalysed click reaction).

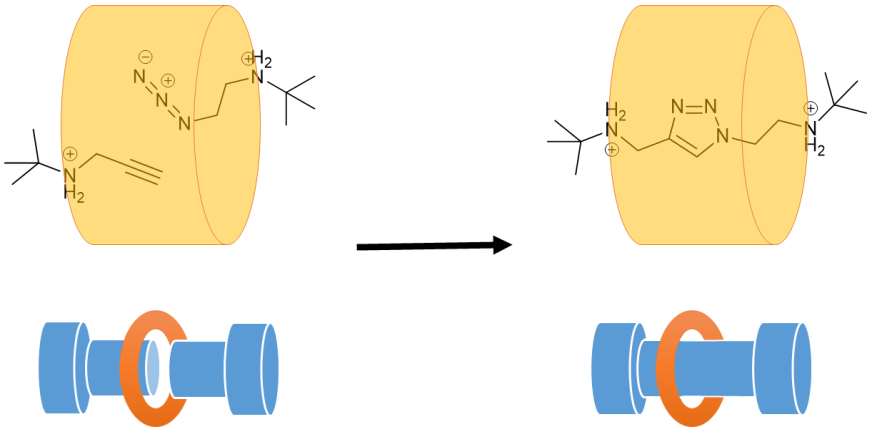


Figure 43: Using a CB6 catalysed click reaction to form a rotaxane. The t-butyl groups are bulky and lock the macrocycle into position[109].

There are a wide variety of applications for cucurbituril rotaxanes and so despite CB10 based structures being investigated significantly less than other cucurbituril constructions, current publications range from using CB10 as a drug delivery system to the potential of CB10 for molecular machinery[111].

3.1.4. **Previous Work**

The imidazole cylinder is an analogue of the parent cylinder where the terminal pyridine rings are replaced with a 2- or 4-imidazole rings (see Figure 44). This simple modification introduces an extra nitrogen atom into the terminal aromatic rings which is not involved in metal binding and is therefore free to bind or react. Tuna *et. al.* began investigating the spin crossover potential of the 4-imidazole iron, nickel, cobalt and manganese cylinders as the ability to hydrogen bond via the new nitrogen atom of the imidazole allowed the triple helicates to form sheet-like arrays. Dependent on the anion used, cooperativity between the cylinders in these arrays allowed a fast transition between the spin states[85].

Later the 2-imidazole iron cylinders were also studied, again with a focus on their spin-crossover behaviours. Using the 2-imidazole cylinder allowed the effect of the NH position on packing, and therefore helicate cooperativity, to be investigated. The DNA binding of the 4-imidazole cylinder was also investigated[84,112].

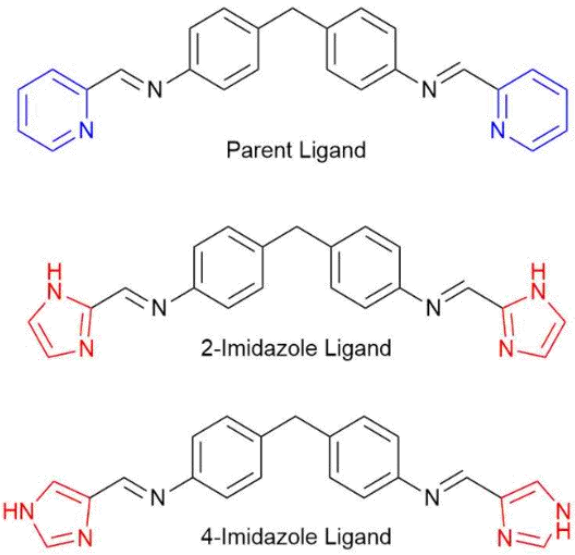


Figure 44: The structure of the parent and imidazole ligands.

3.2.Aims

The aim of this chapter is to synthesise a supramolecular rotaxane from the macrocyclic CB10 and an imidazole cylinder.

Following the work of Mirela Pascu[84], as well as Tuna *et al.*[85], 2- and 4-imidazole ligands are synthesised and complexed to iron(II) and nickel(II), creating a series of di-nuclear triple helicates. These helicates are then tested and selected for two key properties: the ability to thread through the CB10 ring, and whether the cylinder can be modified after assembly. A cylinder is then selected in order to attempt the formation of a rotaxane by a capping mechanism.

The imidazole cylinders retain the important structural factors of the parent cylinder, the binding of two metal ion centres and the core motif for DNA recognition, while introducing another nitrogen atom into the terminal aromatic rings which is available for reaction. This nitrogen has been used to attach three pyridine rings on to each end of the cylinder which are able to act as stoppers and prevent macrocyclic CB10 from dissociating from the cylinder.

3.3.Imidazole Cylinders

Both the 2- and 4-imidazole ligands were synthesised following the procedure previously described by F. Tuna *et al.*[85] and M. Pascu[84]. The synthesis is simple and uses commercially available starting materials. The spacer (4,4'-methylenedianiline) and the appropriate imidazolecarboxaldehyde are heated under reflux together in methanol with a small amount of glacial acetic acid. After two hours, an off white precipitate is collected in high yield (87%).

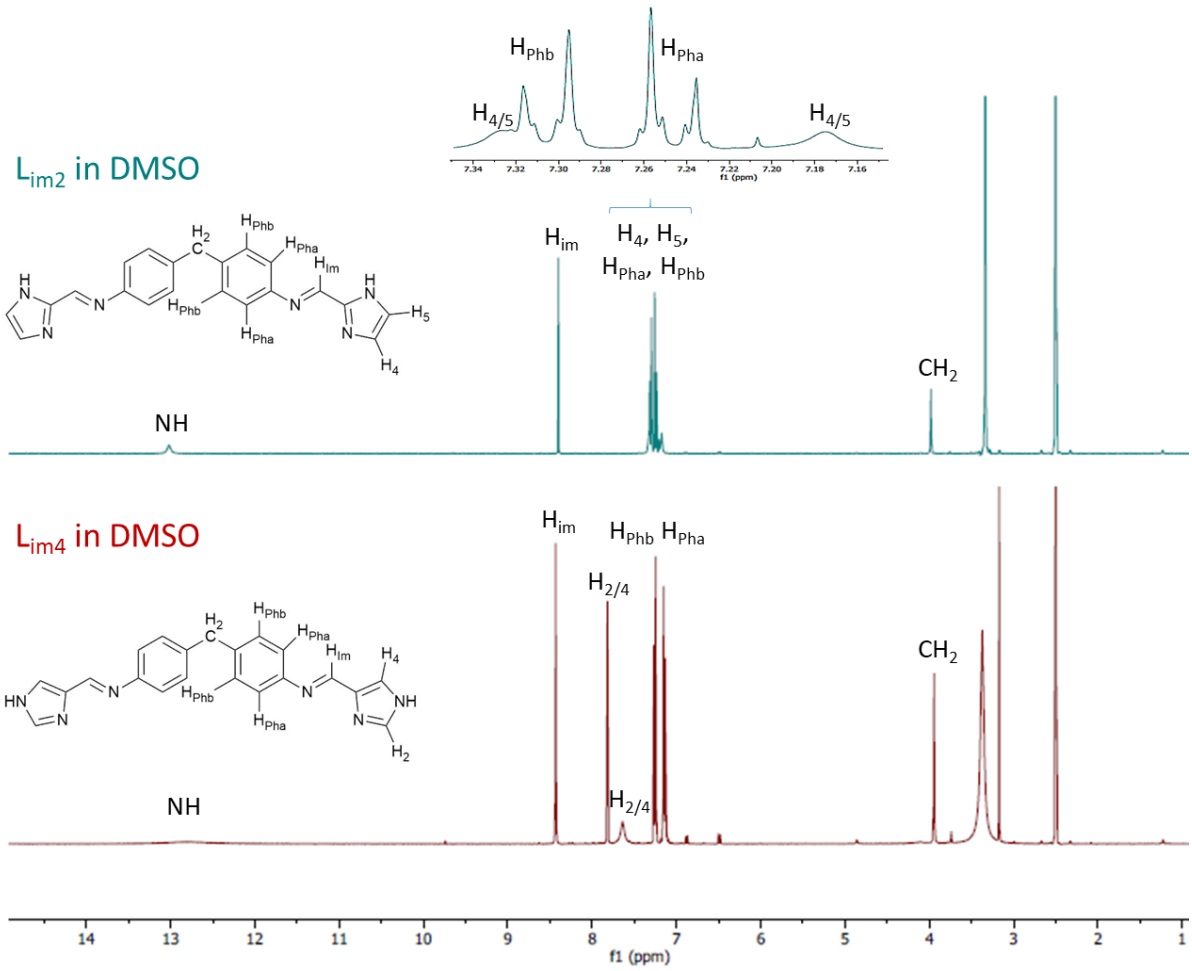


Figure 45: [1]H NMR (300 MHz) of the imidazole ligands in d_6-DMSO. Inset: The phenyl and imidazole peaks of L_{im2} between 7.16 and 7.34 ppm.

The chemical shifts of the ligand protons in DMSO closely match the expected shifts for ^{1}H NMR (see Figure 45)[84]. A COSY spectrum was also taken for both ligands which identified the phenyl protons based on their interaction with the CH_2 protons. As COSY measures through-bond coupling intensity, H_{Phb} showed a much stronger signal than H_{Pha} which are an extra bond (5 compared to 4) away from the CH2 protons (see Figure 46).

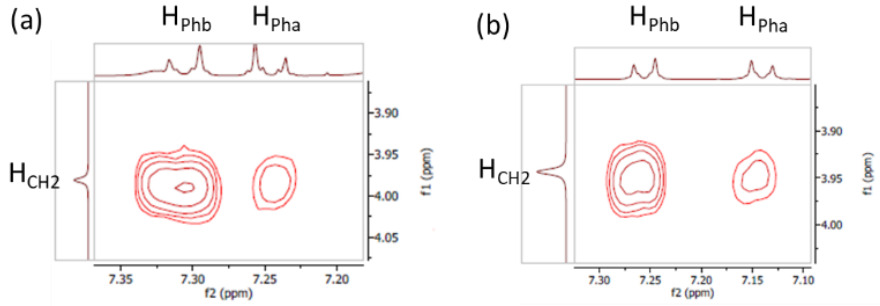


Figure 46: ^{1}H COSY NMR (400 MHz, DMSO) of (a) L_{im2} and (b) L_{im4} showing the stronger coupling of H_{CH2} to H_{Phb} than H_{Pha}.

The iron(II) cylinder was synthesised as both a tetrafluoroborate and a chloride salt for each ligand, producing 4 iron cylinders in total. To produce the tetrafluoroborate salt, each ligand was stirred in acetonitrile with iron(II) tetrafluoroborate hexahydrate. The chloride salt was synthesised by the same method using iron(II) chloride hexahydrate in methanol. The solvents were chosen to increase the solubility of the respective iron salts. Unlike the published procedure[85], the cylinder had to be precipitated in diethyl ether in both cases as no crystals appeared to form.

The same process was repeated with nickel(II) chloride hexahydrate in methanol to yield the nickel cylinder. The cylinder was precipitated by addition of methanoic ammonium hexafluorophosphate and was converted back to a chloride by stirring with Dowex (chloride form) beads in methanol until the cylinder was fully dissolved.

3.3.1. **Paramagnetic NMR**

For octahedral complexes, iron(II), in its high spin state, and nickel(II) both possess unpaired electrons (Figure 47). Iron(II) is in a high spin state when bound to both L_{im2} and L_{im4} meaning all the imidazole cylinders studied are paramagnetic, which significantly affects the ^{1}H NMR of these compounds.

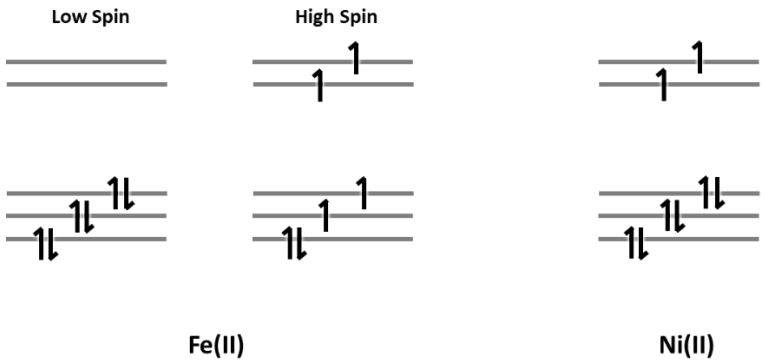


Figure 47: The electronic configurations of Fe(II) and Ni(II) in their octahedral coordination.

Electrons have an associated magnetic field hundreds of times stronger than the magnetic field of a nucleus[113]. While the variety of proton shifts in a molecule is usually due to shielding by local electron density, proximity to the unpaired electron has a much greater effect. The peaks in this case are now spread between -6.9 and 229.0 ppm, compared to 3.9 and 13.0 ppm observed for the ligand. This is called a paramagnetically shifted NMR spectrum. The spectra have been broadened as the presence of the unpaired electron decreases the relaxation time, which in turn prevents any coupling to be observed in the spectra. Assignment of the spectra is therefore very difficult, but an attempt has been made by comparison to the assignments by Floriana Tuna *et al.*[85] (see Table 1).

In some cases where a high shift was expected, particularly H_2 and NH protons for L_{im4}, the peaks were not observed. This can be rationalised by the close proximity to the unpaired electron, which not only causes the increased shift but is likely to have a stronger broadening effect.

	$[Fe_2(L_{im4})_3]$			**$[Ni_2(L_{im2})_3]$**		
(ppm)	Literature - $(BF_4)_4$	$(BF_4)_4$	$(Cl)_4$	Literature – $(PF_6)_4$	$(PF_6)_4$	$(Cl)_4$
$H_{Pha/b}$	-5.6	-5.7	-5.7	-6.7	-6.7	-6.9
$H_{Pha/b}$	14.7	14.7	14.7	15.4	15.5	15.6
CH_2	24.8	24.9	25.4	27.5	27.7	28.4
H_4	38.1	38.2	37.2	40.9	41.3	41.1
H_{im}	42.9	43.0	41.9	58.7	59.7	60.0
NH	92.7	92.9	*	93.1	94.1	*
H_2	159.2	159.5	157.5	229.0	*	*

*Table 1: A table to compare the ^1H NMR (300 MHz) values from the paramagnetic spectra of the L_{im4} cylinders. *Values were too weak and broad to be seen in the spectrum. Tetrafluoroborate and hexafluorophosphate salts were dissolved in d_3-actotnitrile while chloride salts were dissolved in d_4-methanol. All values in red were taken from the literature for comparison, see Reference 85.*

There are no previously published NMR spectra for the 2-imidazole cylinders so assignments were made by comparison to the respective 4-imidazole equivalents and proximity to the metal ion centre (the unpaired electron). The integration values were particularly useful when identifying the phenyl peaks which had integrations which were double the value of the other proton peaks.

	$[Fe_2(L_{im2})_3]$		**$[Ni_2(L_{im2})_3]$**	
(ppm)	$(BF_4)_4$	$(Cl)_4$	$(PF_6)_4$	$(Cl)_4$
$H_{Pha/b}$	0.6	0.6	-5.4	-5.3
$H_{Pha/b}$	17.8	17.4	15.2	15.6
CH_2	22.1	22.0	25.2	28.4
H_{im}	56.4	61.6	66.2	41.1
H_5	65.3	67.8	74.8	60.0
NH	71.2	*	*	*
H_4	124.2	118.7	*	*

*Table 2: A table to compare the ^1H NMR (300 MHz) values from the paramagnetic spectra of the L_{im2} cylinders. *Values were too weak and broad to be seen in the spectrum. Tetrafluoroborate and hexafluorophosphate salts were dissolved in d_3-actotnitrile while chloride salts were dissolved in d_4-methanol.*

3.3.2. **Interactions with CB10**

Previously L. Cardo[114] was able to show that the parent iron and ruthenium cylinders were able to thread through and bind the macrocyclic CB10. NMR and mass

spectroscopy were used to investigate whether the same threading interaction would occur using the imidazole cylinders.

$[Fe_2(L_{im4})_3]Cl_4$ is the most water soluble out of the imidazole cylinders so was used for the initial threading tests. The cylinder was dissolved in D_2O (1 mM) and 0.9 equivalents of CB10 were added so all the CB10 could be non-covalently bound. The solution was stirred until the white CB10 powder was completely dissolved.

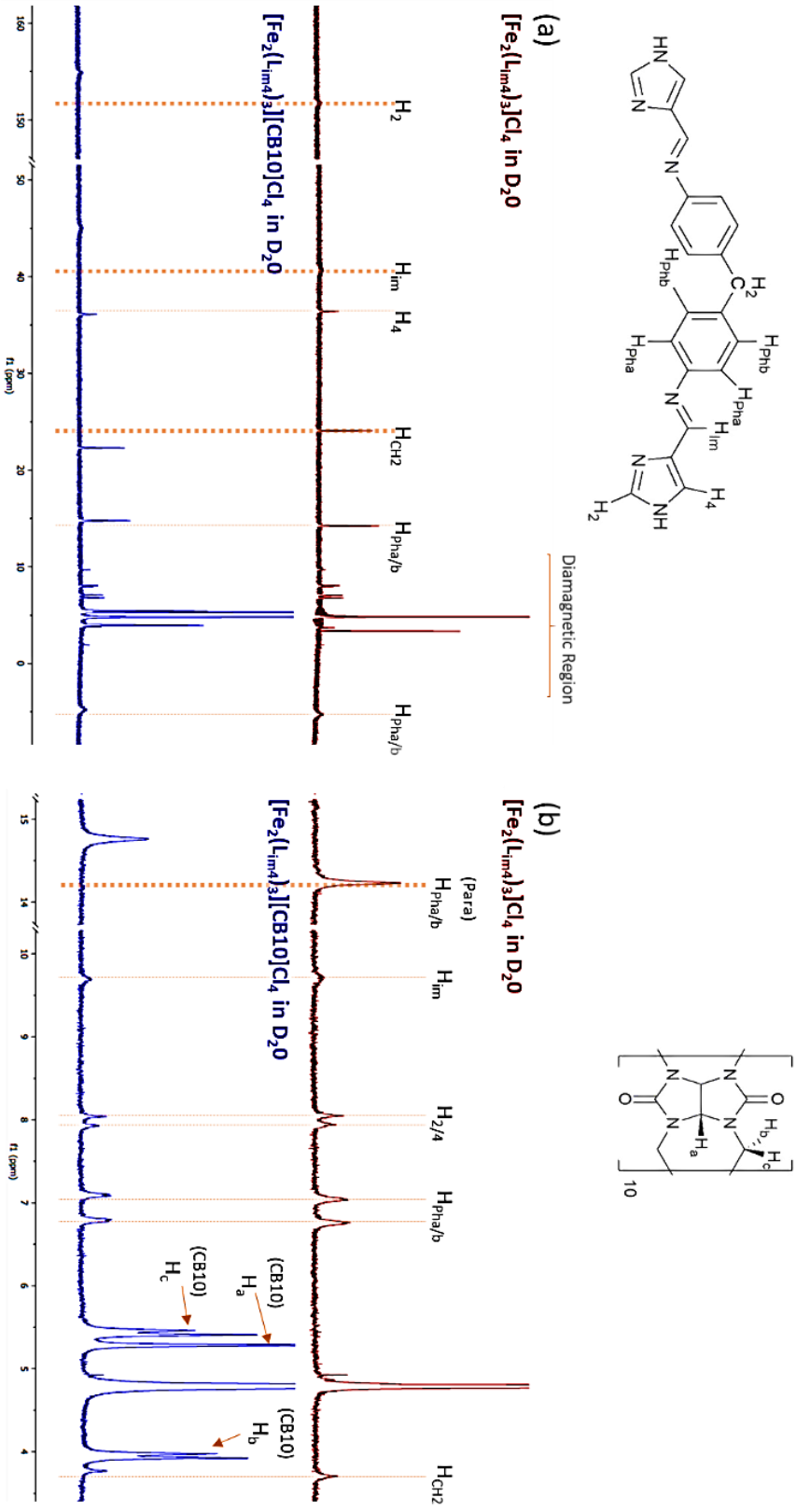


Figure 48: Comparing the chemical shifts of $[Fe_2(Lim_4)_3]Cl_4$ in D_2O both with and without CB10. The stacked spectra show (a) the paramagnetic shifts and (b) the diamagnetic shifts with the CB10 protons.The structures and labelling of the ligand (top left) and CB10 (top right) are included for clarity. (^1H NMR, 300 MHz)

In the ^{1}H NMR there are three clear, broadened peaks to indicate the presence of CB10 near the paramagnetic centres. The largest changes in the chemical shift of the cylinder protons upon addition of CB10 appear in the paramagnetic region of the ^{1}H NMR spectra (see Figure 48(a)). Protons H_2, H_{im} and H_{CH2} all show a large shift (>2 ppm). Both H_2 and H_{im} are shifted downfield so are deshielded. This could show that the protons are within the portal of CB10 which is surrounded by polar ketone groups and able to withdraw electron density from the already deshielded protons. The H_{CH2} protons are shielded by CB10 binding, shown by the downfield shift, as the hydrophobic aromatic core of the cylinder is now shielded by the CB10 macrocycle.

Despite being in the centre of the macrocyle cavity, the phenyl protons ($H_{Pha/b}$) do not show the same shielding as the H_{CH2} protons and appear to be slightly deshielded. As it is at the end of the cylinder and has very little contact with the CB10 ring, H_4 shows the smallest change in shift (0.31 ppm).

When the parent cylinder is thread through CB10, similar protons are shifted[114]. The CH_2 protons are also shielded, causing an upfield shift, and H_{im} is shifted downfield. For the parent cylinder, H_3 is also shifted downfield as it is pointing back into the macrocycle portal. This is not observed for H_4, possibly because the imidazole ring alters the angle at which H_4 is pointed directed. The protons of both the parent and imidazole cylinders are not split into different environments, as peak splitting is not observed in the NMR spectra, which can only be fulfilled if the CB10 is bound by threading.

The iron cylinder can also exist in a low spin diamagnetic state which is in equilibrium with the high spin paramagnetic state. Conversion between these two states is slow, especially when compared to an NMR timescale, so the diamagnetic species gives

a separate set of lower intensity peaks between 3.5 and 10 ppm in deuterated water (see Figure 48(b)). These peaks do not show the same change in chemical shift which could imply that the macrocycle binds the paramagnetic species preferentially. To study this further, a titration could be completed with increasing amounts of CB10 and the intensities of the paramagnetic peaks should be compared to the diamagnetic peaks. An increase in the ratio of the intensity of bound paramagnetic species to diamagnetic peaks would suggest preferential binding.

Threading of $[Ni_2(L_{im2})_3]Cl_4$ through CB10 was also tested. The nickel cylinder is not soluble at high enough concentrations in water alone for an NMR to be taken, so 20% deuterated methanol solution was used. The cylinder was dissolved in solution and 0.9 equivalents of CB10 were added. Unlike the iron cylinder, not all of the CB10 dissolved after stirring so any excess was removed by filtration.

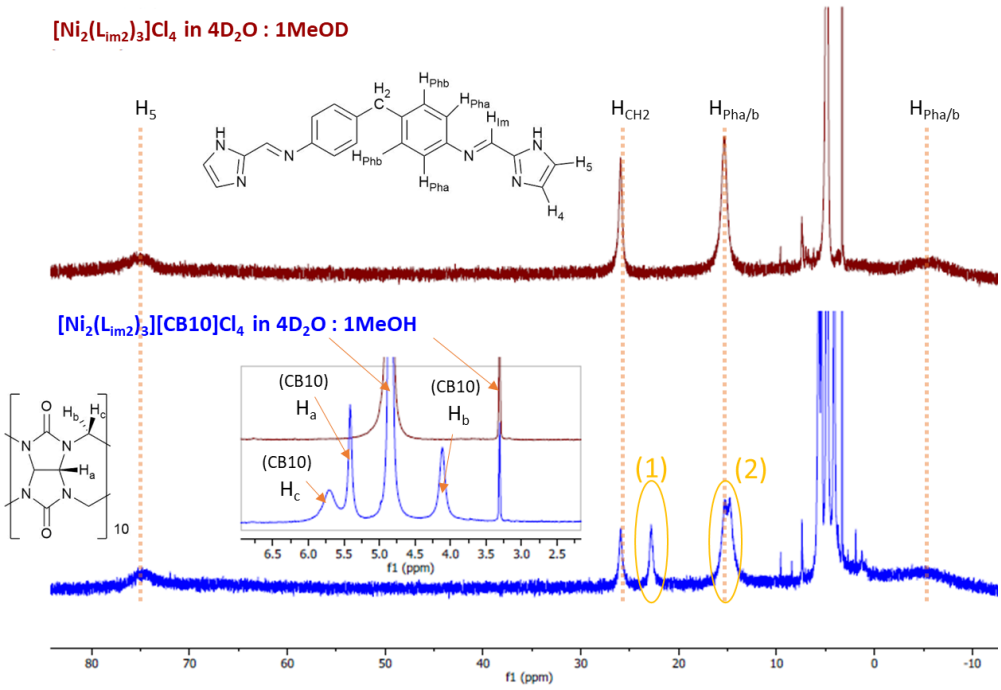


Figure 49: Comparison of the ^1H NMR paramagnetic spectra of $[Ni_2(L_{im2})_3]Cl_4$ with (bottom) and without (top) CB10. The insert gives a more detailed view of the spectra between 2.5-6.5 ppm to show the CB10 proton peaks. The structures of L_{im2} (top) and CB10 (left) are shown with labelled protons. (^1H NMR, 300 MHz, Gaussian function (1GB) applied to both spectra to reduce background noise).

Unlike the iron cylinder the nickel cylinder is a d^8 transition metal and does not have a diamagnetic state so only the paramagnetic region is shown (see Figure 49). The broader signals (H_5, $H_{Pha/b}$) appear unchanged by the addition of CB10 however the sharper signals (circled in Figure 49) show that the CH_2 and phenyl proton shifts have been split into two peaks. Comparison to the spectra before addition of CB10 suggests that this is because only half of the nickel cylinder has been bound by the CB10. This is confirmed by the equal integrations of the two CH_2 peaks (0.46 for both) and the CB10 proton integrations (3.32 and 1.58) being very close to the expected value for half of the cylinder bound (3.06 and 1.53 respectively). It is likely that the other peaks do not show this splitting as they are too broad. Again, the CH_2 protons seem to be most affected by the non-covalent macrocycle binding.

The mass spectrum (see Figure 50) also shows the presence of the unbound cylinder (*m/z* = 410 $[Ni_2(L_{im2})_3Cl^{3+} + H_2O]$) and the CB10 threaded cylinder (*m/z* = 710 $[Ni_2(L_{im2})_3Cl^{4+} + CB10]$).

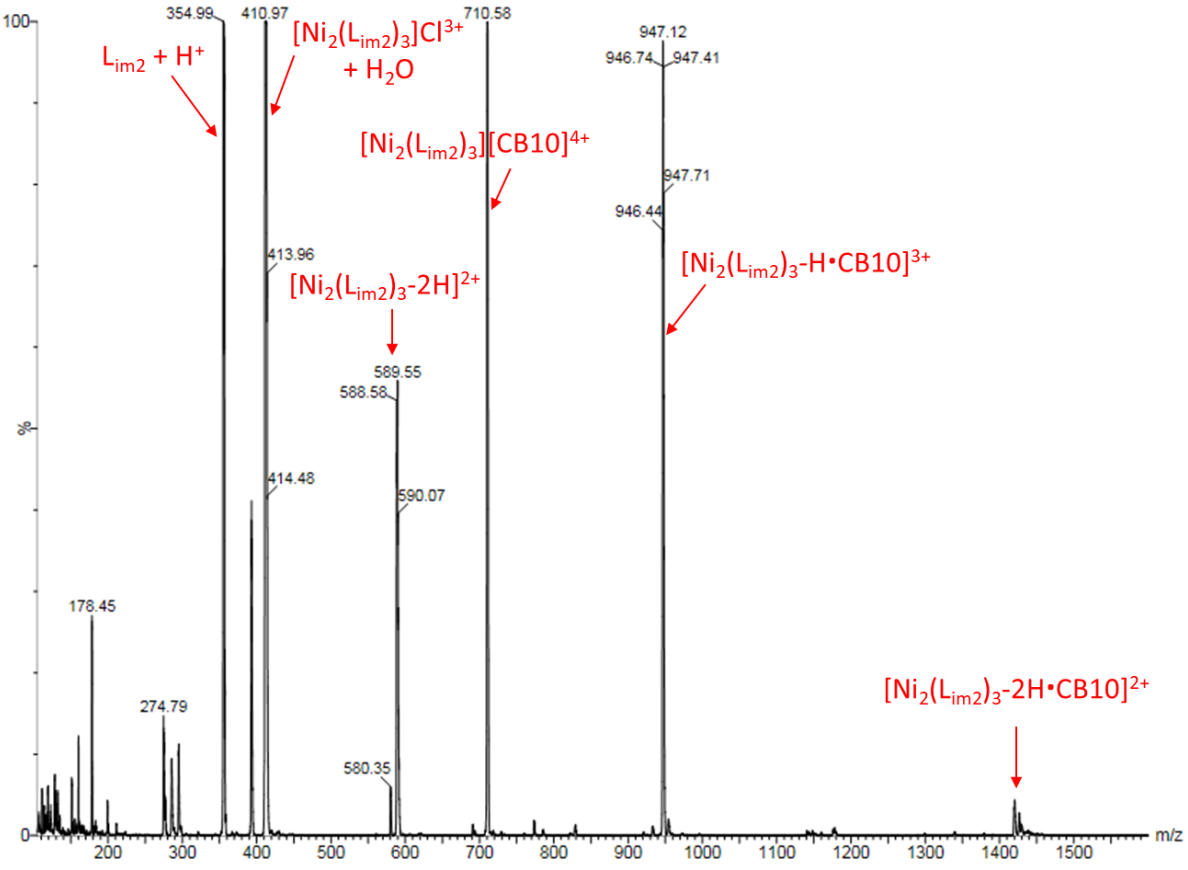


Figure 50: Mass spectrum (SCAN ES+) of the 2-imidazole cylinder and CB10 in solution.

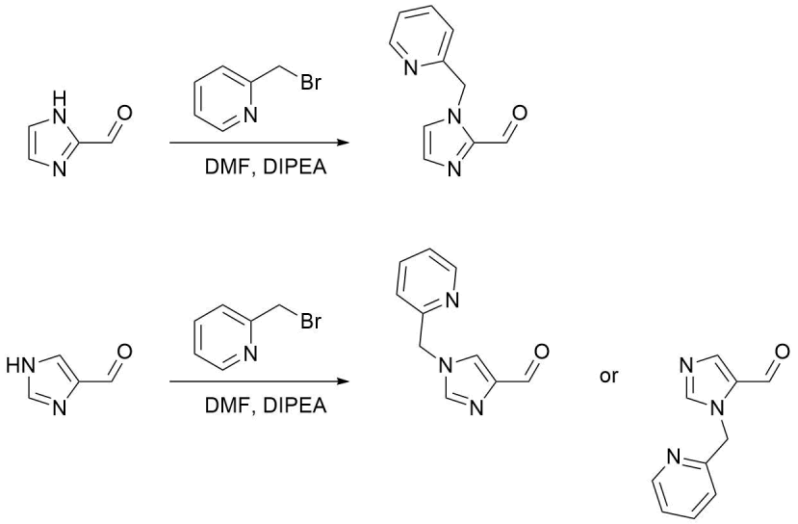
The data suggests that the imidazole cylinders are able to bind to CB10. It is likely that these cylinders thread through the CB10 much like the parent cylinder, however the data presented here is not enough to show that the cylinder is within the central cavity of the CB10 ring.

3.4.Post Assembly Modification

3.4.1.Alkylation of an imidazole carboxaldehyde

Viktoriia Sadovnikova followed a published method[115] to alkylate an imidazolecarboxaldehyde which could then be used to synthesise a modified imidazole ligand via imine formation. To alkylate the imidazolecarboxaldehyde, 2-

(bromomethyl)pyridine hydrobromide is dissolved in DMF and added to the aldehyde before the addition of Hunig's base (*N*,*N*-diisopropylethylamine). The reaction is quenched with aqueous sodium hydrogen carbonate before being extracted with dichloromethane[116].



Scheme 4: Alkylation of an imidazole carboxaldehyde, originally using the 2-imidazole carboxaldehyde[116] *(top) and repeated with the 4-imidazole carboxaldehyde (bottom).*

Originally the 2-imidazolecarboxaldehyde was used[116], so the reaction was repeated with the 4-imidazolecarboxaldehyde to see if the alkylation would occur preferentially at one nitrogen site over the other (see Scheme 4). The product was analysed by liquid chromatography – mass spectrometry (LCMS), where the sample is passed through a reverse phase C18 column to separate the products which are then immediately analysed by mass spectrometry. This process is monitored by UV-Vis spectroscopy which is used to detect products as they are eluted by measuring the light intensity at a wavelength that the product is expected to absorb. Pyridine has a strong absorbance at 254 nm[117] so this wavelength was selected to monitor the separation. The mass spectra corresponding to two large absorption peaks both showed evidence of the protonated product at m/z = 188. The mass spectrum corresponding to the smaller absorption peak did not have a peak at

m/z = 188 suggesting that the imidazolecarboxaldehyde is not present and explaining the weaker absorbance.

A separation was then attempted by preparatory high pressure liquid chromatography (prep-HPLC). Of the four peaks that were eluted, peaks three and four were shown to contain the product peaks by mass spectrometry. The two products were thought to be the two isomers produced by alkylation of the different nitrogen atoms and NMR spectra of the two separated products were taken to elucidate which product was which isomer.

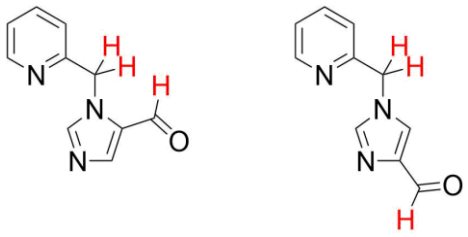


Figure 51: Highlighting the key difference between the two possible alkylation products. The proximity of the aldehyde proton to the carbons allows the isomers to be distinguished by COSY-NMR and NOESY-NMR.

The key difference between the two isomers is the proximity of the aldehyde proton versus the imidazole proton to the methylene linker between the imidazole and pyridine rings (see Figure 51). An attempt to distinguish the two isomers was done by COSY-NMR, which investigated the through bond coupling of protons. In both cases, the two imidazole proton peaks (H^9 and H^{12}) overlap, meaning that coupling to the CH_2 protons is expected regardless of the position of the second imidazole proton. Fortunately, the aldehyde proton is easily distinguished with the highest chemical shift. The coupling to the CH_2 proton in the spectra for the product of peak 3 is much more intense than the small signal observed in the COSY spectra for the product of peak 4, which could indicate that the product of peak 3 is isomer A while the product of peak 4 is isomer B (see Figure 52).

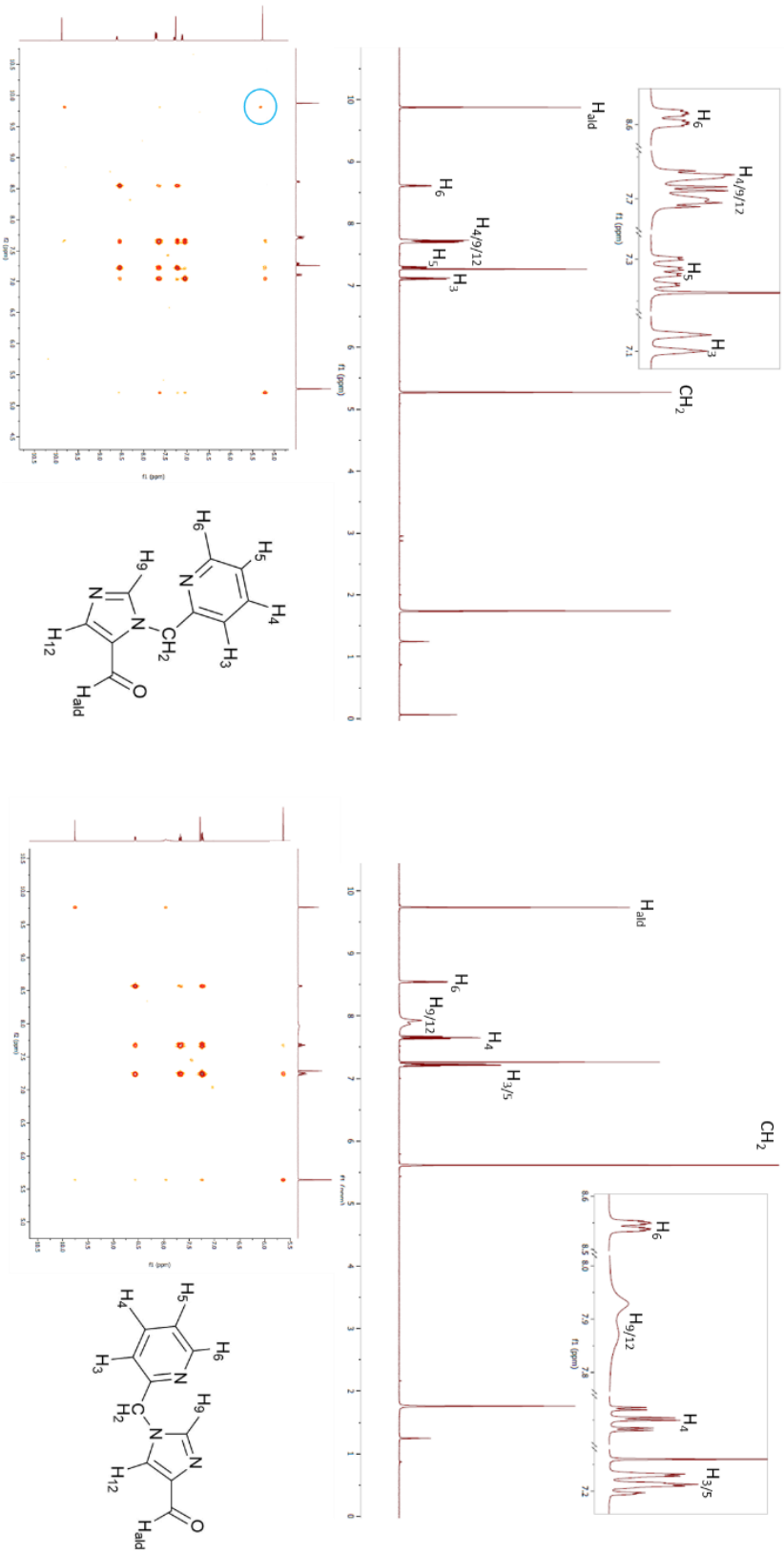


Figure 52: The [1]H NMR and COSY NMR (400 MHz, chloroform-d) of fraction 3 (left) and fraction 4 (right) isolated by prep-HPLC with their suggested asssignments. Circled in blue is the COSY signal between CH_2 and H_{ald} which is not observed for the second isomer.

HSQC was also attempted, however the signal from the carbon NMR was too weak to show significant coupling between the carbon atoms and their protons.

To confirm this assignment, further studies such as nuclear overhauser effect spectroscopy (NOESY) and HMBC should be done. HMBC may be able to show coupling of the imidazole carbons, which should not overlap, and the methylene protons. A more significant result may come from NOESY as it shows through space rather than through bond coupling. This means that the imidazole or the aldehyde proton would have a stronger correlation to the methylene protons which could quite easily distinguish the two isomers. The aim of this experiment was both to show that the procedure worked and that it could be repeated with the 4-imidazole as well as the 2-imidazole, so no further studies were completed at this point.

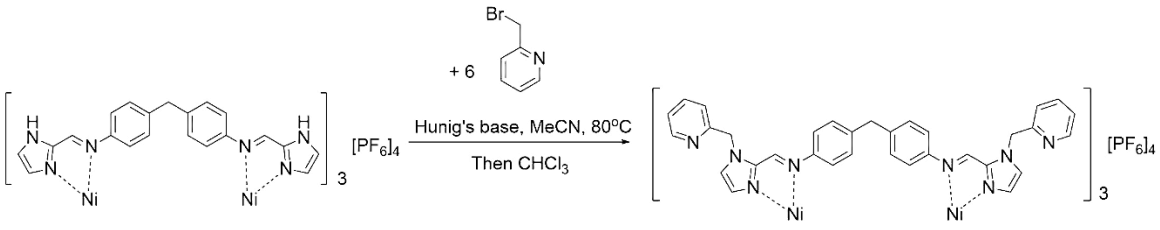
3.4.2. **Post assembly modification of an imidazole cylinder**

After successfully modifying the 4-imidazole carboxaldehyde, post assembly modification of the 2-imidazole nickel cylinder, $[Ni_2(L_{im2})_3]PF_6$, was attempted. A 2-imidazole cylinder was chosen as the 4-imidazolecarboaxaldehyde reaction demonstrated that alkylation can occur at either nitrogen position on the imidazole ring. When coordinated to a metal centre in a supramolecular cylinder, one of the nitrogen atoms of the ligand imidazole ring should be unavailable for alkylation, however it is also possible that the ligand temporarily dissociates from the metal ion, allowing either nitrogen to react. If the nitrogen atom in the 1 position of the 4-imidazole were to react, the imidazole would no longer be able to coordinate to the metal ion causing the breakdown of the cylinder. For this reason, the 2-imidazole ligand was chosen to attempt post assembly

modification as there is only one possible alkylation product, and the ligand will be able to bind the metal centre after the reaction has occured.

The nickel cylinder was chosen as generally nickel analogues are more stable than iron cylinders. UV studies of $[Fe_2(L_{im2})_3]Cl_4$ and $[Ni_2(L_{im2})_3]Cl_4$ in methanol (see sections 4.3.3 and 4.3.4) support this as the signal for $[Fe_2(L_{im2})_3]Cl_4$ decreases at a faster rate than the signal $[Ni_2(L_{im2})_3]Cl_4$ at their respective λ_{max} (0.16% per minute at 328 nm and 0.03% per minute at 319 nm).

Dimethylformamide is able to co-ordinate to metal centres so the solvent was changed to acetonitrile for the cylinder alkylation to allow the cylinder to be isolated after the reaction had occurred.



Scheme 5: Post assembly modification of the nickel 2-imidazole cylinder in acetonitrile.

The nickel cylinder and 2-(bromomethyl)pyridine hydrobromide were dissolved in acetonitrile before Hunig's base was added and the solution was stirred at 80°C overnight. The cloudy yellow solution initially became pink on addition of the base, which then turned to a cloudy orange after stirring. A small sample of this solution was allowed to settle, showing that the cloudy orange colour came from a yellow precipitate in a pink solution. The orange solution was filtered to remove a pale brown powder before the solvent was removed from the filtrate. The resulting sticky orange solid was washed with chloroform to remove any excess organic reagent or by-product which yielded a pale orange powder product (see Scheme 5).

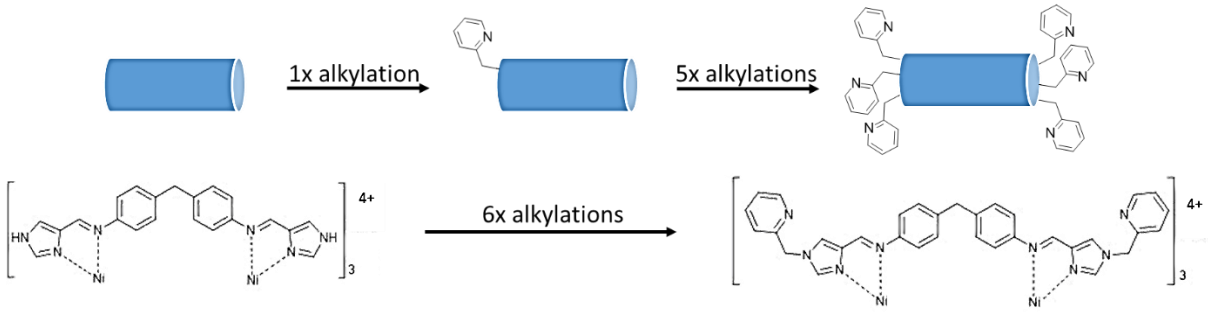


Figure 53: Illustrating that the fully alkylated product must go through six separate alkylation reactions.

The cylinder must undergo six alkylation reactions in order to form the fully alkylated product (see Figure 53). This means that the side products of this reaction include any partially reacted cylinder species which are referred to by their number of alkylations or modifications. By pushing this reaction to completion both the yield and purity of the product benefit, as it would be difficult to separate the fully reacted product from a partially reacted side product.

Mass spectrometry was primarily used to identify the product as a clear peaks at *m/z* = 432 ($[Ni_2(L_{im2})_3]^{4+}$), 624 ($[Ni_2(L_{im2})_3](PF_6)^{3+}$), and 1009 ($[Ni_2(L_{im2})_3](PF_6)^{2+}$) were easily assigned and all of the identifiable peaks corresponded to either completely modified cylinders or cylinders with 5 alkylations (see Figure 54).

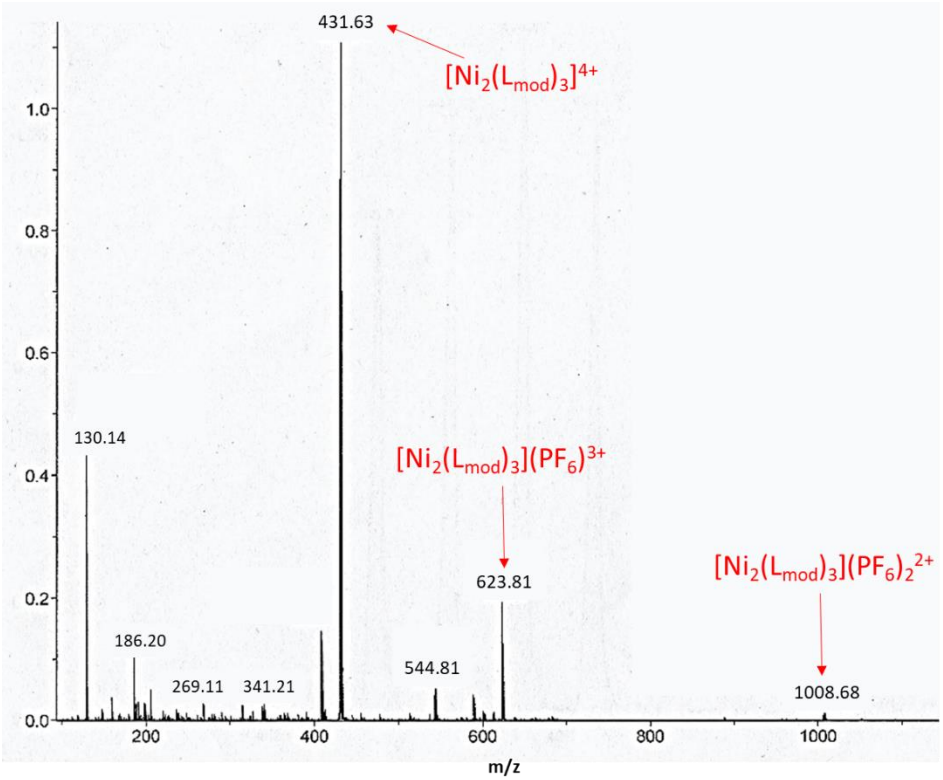


Figure 54:Mass spectrum (TOF MS ES+) of orange powder product containing the fully alkylated cylinder.

The pale brown precipitate appeared to contain many partly modified (and some fully modified) cylinder species but was not soluble in acetonitrile (see Figure 55).

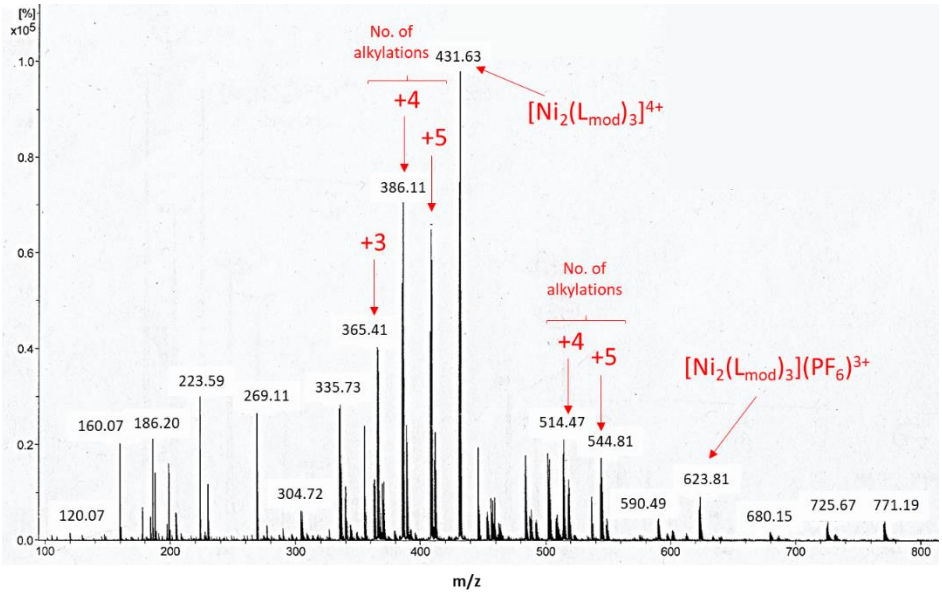


Figure 55: Mass spectrum (TOF MS ES+) of the brown precipitate containing many partially reacted species. Some of the larger peaks have been assigned by the number of alkylations before the fully alkylated species.

The NMR of the product (see Figure 56) showed similar peaks in the paramagnetic region to the nickel 2-imidazole cylinder. Between 0-10 ppm new peaks appear which are not observed in the spectrum prior to alkylation (see Figure 56 inset). Four peaks between 7-9 ppm each with an integration value of 1 can be attributed to the pyridine protons on the alkylated cylinder (H_3-H_6). Two further peaks can be found at 1.64 ppm and 4.20 ppm. Both peaks are broad and have an integration value 1. These peaks could be due to the CH_2 groups between the imidazole and pyridine rings although it is unclear why these protons do not have the same shift.

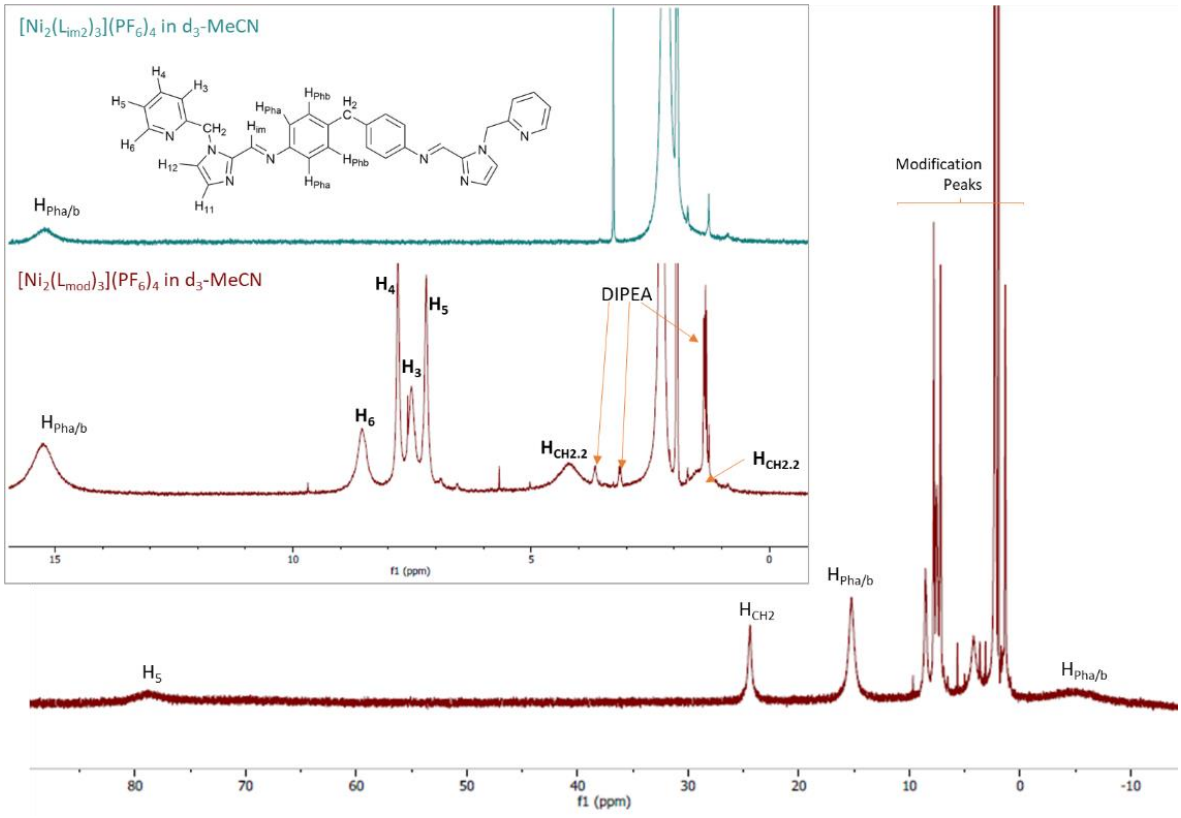


Figure 56: ^{1}H NMR (300 MHz) of $[Ni_2(L_{mod})_3](PF_6)_4$ in deuterated acetonitrile. Insert shows magnified view of spectrum between 0-15 ppm (bottom) with the spectrum for the cylinder prior to alkylation, $[Ni_2(L_{im2})_3](PF_6)_4$, stacked (top) for comparison.

As the alkylation had been successful in acetonitrile, the reaction was repeated in the presence of CB10 in an attempt to form the rotaxane. An orange, flakey precipitate (28.2 mg, 25% yield) was isolated however it was clear from the ^{1}H NMR and mass spectrum that there was very little if any cylinder threaded through the macrocycle (see Figure 57 and Figure 58). The polarity of acetonitrile may not be high enough to drive the cylinder through the CB10 ring resulting in favourable hydrophobic interactions.

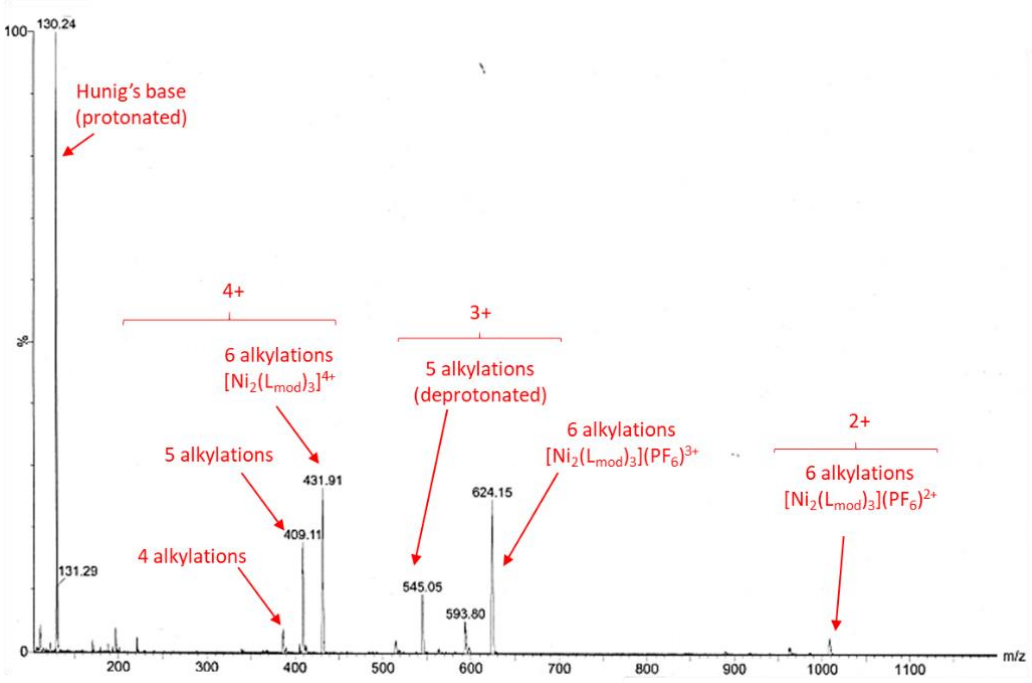


Figure 57: Mass spectrum (SCAN ES+) of the orange powder product from the alkylation reaction in the presence of CB10.

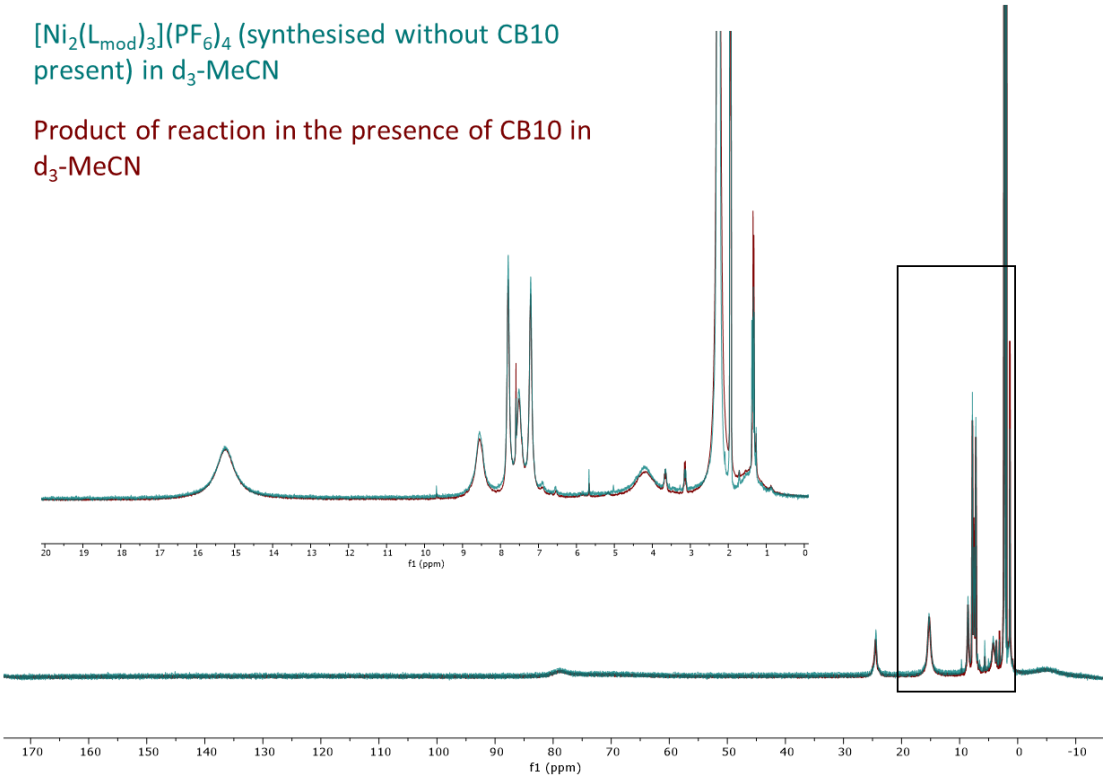


Figure 58: ^{1}H NMR (300 MHz) of the orange powder precipitate product of the alkylation in the presence of CB10 (maroon) overlaid with the alkylation product of the post assembly modification. The inset image is a more detailed view of the spectrum between 0-20 ppm as outlined in the spectrum.

The nickel 2-imidazole cylinder has been shown to non-covalently bind CB10 in a 20% methanol solution by ^{1}H NMR and mass spectrometry (see Section 3.3.2). Two higher polarity solvents, ethanol and water, were tested to see if the alkylation reaction could still occur. In both cases, $[Ni_2(L_{im2})_3]Cl_4$ and 2-(bromomethyl)pyridine hydrobromide were dissolved in the solvent and Hunig's base was added.

After cooling to room temperature, the reaction in ethanol was filtered and the ethanol removed from the filtrate under vacuo. The solid product was then suspended in chloroform to remove any organic products and the resulting precipitate was isolated by filtration. The mass spectrum of the precipitate showed that a mixture of modified cylinder species were present, with the 4 times alkylated cylinder giving the greatest peak intensity (see Figure 59(a)).

The alkylation reaction in water was also filtered after cooling to remove a dark brown precipitate. As chloroform and water are immiscible, filtrate could then be washed with chloroform without having to remove the solvent and the aqueous phase was collected. The cylinder was precipitated from the aqueous solution as a hexafluorophosphate salt. The mass spectrum showed that the 6 times (fully) alkylated cylinder species had the greatest intensity, demonstrating that the cylinder could be completely alkylated in water (see Figure 59(b)).

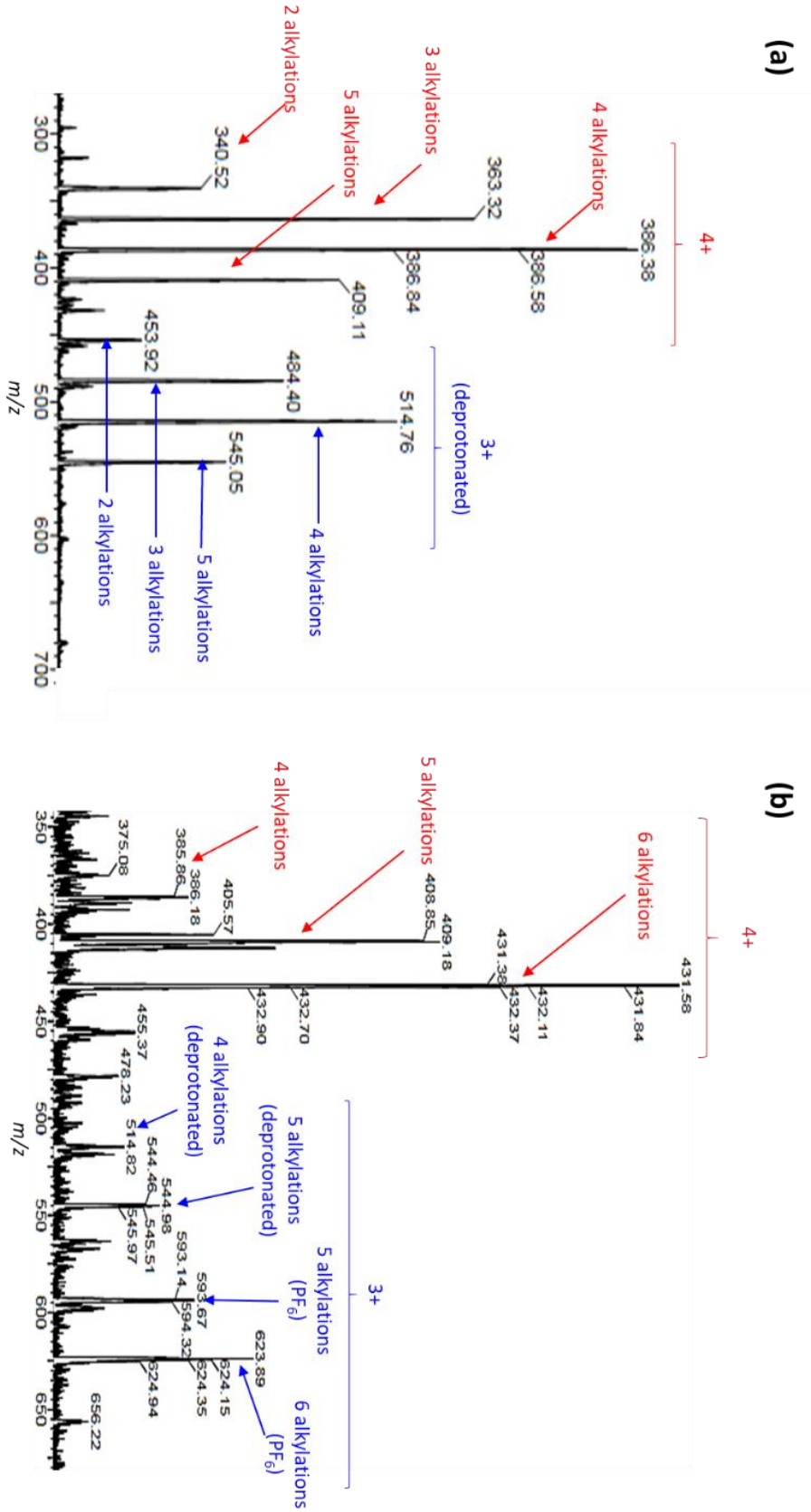


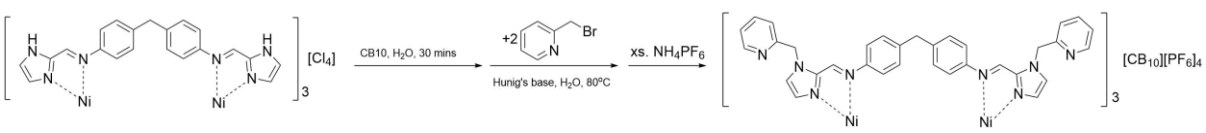
Figure 59: Mass spectrum (Scan ES+) of the product of the alkylation reaction in (a) ethanol and (b) water.

3.5.Rotaxanes

3.5.1.Synthesis of a rotaxane

The alkylation of the nickel 2-imidazole cylinder had already been attempted in the presence of the CB10 but failed to produce the desired rotaxanated species. For this reason, the alkylation of the cylinder in alternative, higher polarity solvents was tested. The reaction in water showed positive results with the fully modified cylinder being the main species detected by mass spectrometry. For this reason, water was chosen as the solvent for further attempts at rotaxane formation (see Scheme 6).

Initially the nickel cylinder, $[Ni_2(L_{im2})_3]Cl_4$, and macrocycle, CB10, were suspended in deionised water in stoichiometric amounts and heated slowly to 40°C while stirring to help the threading process. Once the solution had become clear, the alkylation reagents, 2-(bromomethyl)pyridine hydrobromide and Hunig's base, were added and the solution stirred overnight at 80°C. Like the alkylation without CB10, the reaction was cooled to room temperature, filtered to remove a cream precipitate, and the filtrate washed with chloroform. The product was precipitated from the aqueous phase as a hexafluorophosphate salt overnight at 4°C before the fine powder was isolated by filtration using a 0.45 μm nylon membrane. While the mass spectrum suggested a small amount of the rotaxanated species was present, the number of modifications was varied (see Figure 60a).



Scheme 6: Formation of a pseudorotaxane followed by alkylation of the cylinder in water.

To investigate whether the reaction required a greater amount of time to fully alkylate, the fine cream powder product was redissolved in water and the alkylation reagents were added again. The solution was stirred at 80°C overnight before a small sample was taken for analysis.

The new mass spectrum (see Figure 60b) showed that only the species with four or more modifications were present, with some peaks (*m/z* = 805 and 828) suggesting that a more modified cylinder had fragmented at the methylene group, which most likely occured while recording the mass spectrum. Unfortunately, the low mass of the pyridine ring fragment was not recorded as this was outside of the spectrum limits (*m/z* = 100-1200).

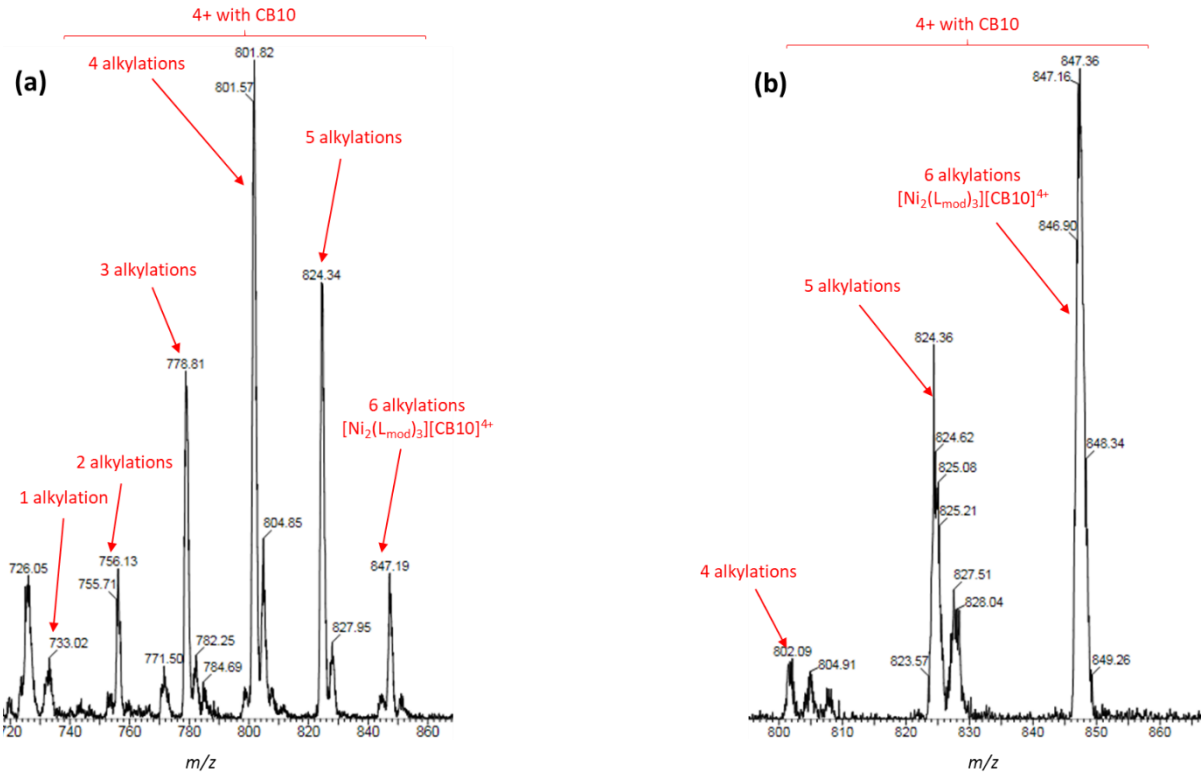


Figure 60: The mass spectrum (SCAN ES+) of the product of (a) the first overnight stir in water and (b) the second overnight stir in water.

The ^{1}H NMR spectrum of both products was also recorded (see Figure 61). A different solvent was used to take each spectrum as the first product was precipitated as a hexafluorophosphate salt and the second product was isolated as a chloride. Acetonitrile and deuterium oxide were used respectively to dissolve the products for NMR. Although this prevents comparison of the exact shifts of the NMR peaks, the CB10 proton peaks can be identified in each spectrum. The definition of the pyridine peaks also seems to improve in the spectrum of the second product, which agrees with the conclusion that after the second stir, more of the fully alkylated product and fewer partially alkylated side products are present.

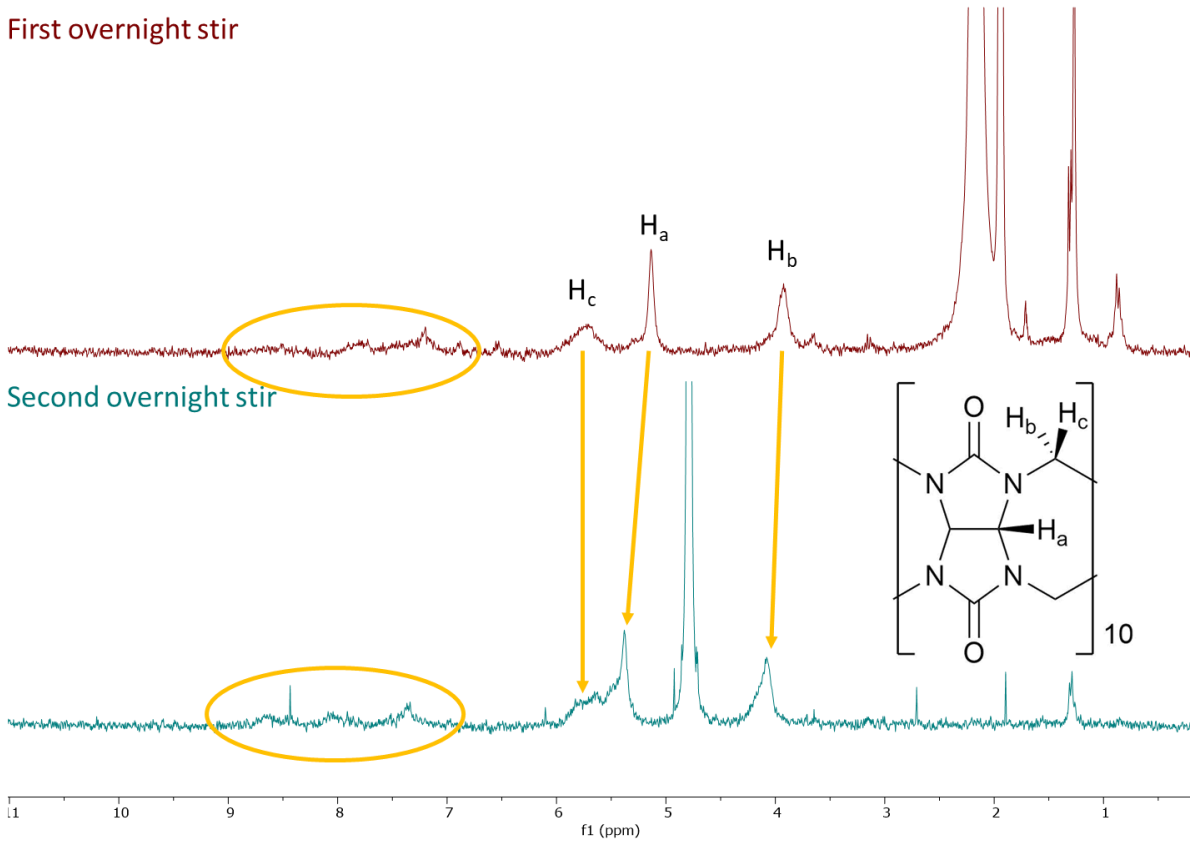


Figure 61: ^{1}H NMR spectra of the product of the first overnight stir dissolved in d_3-aceonitrile (maroon)) and (b) the second overnight stir dissolved in deuterium oxide (green). The CB10 peaks are labelled and the pyridine peaks of the alkylated cylinder are circled in each spectrum. A Gaussian (1GB) and exponential (0.3Hz) function has been applied to both spectra.

The rest of the reaction mixture was allowed to stir at 80^{o}C for a further two days to see if all the cylinder would be completely alkylated with only the 6 modified species left. Although the mass spectrum does appear to only show a peak for the 6 modified species, the peak is small and much larger peaks are observed for the unmodified ligand bound to different ions. Stirring for long periods at high temperature is therefore likely to cause the cylinder to breakdown.

In every product presented so far, peaks in the NMR (1.35 ppm) and mass spectra (*m/z* = 130) indicate that the base is still present. The original reaction suggested by Viktoriia Sadovnikova included a quenching step with sodium hydrogen carbonate, which converted any unreacted base into a salt. To isolate the cylinder without excess base being present, an attempt to quench the reaction with sodium hydrogen carbonate was made.

The alkylation reaction in water was repeated and after a 24 hour stir, a saturated solution of sodium hydrogen carbonate was added to quench the reaction. The solution allowed to settle before being filtered under vacuum to remove a brown precipitate and a bright yellow filtrate was collected. The solvent was removed under vacuo and the resulting salts suspended in excess methanol as the rotaxane should be weakly soluble in methanol but the salts should not be. The salts were removed by filtration and the solvent was removed from the filtrate to yield a pale beige powder solid.

The product was shown to be soluble in water, but not at the concentration required for a clear ^{1}H NMR, especially as the unpaired nickel electrons cause severe broadening of the paramagnetic peaks. The NMR shown are for demonstration only.

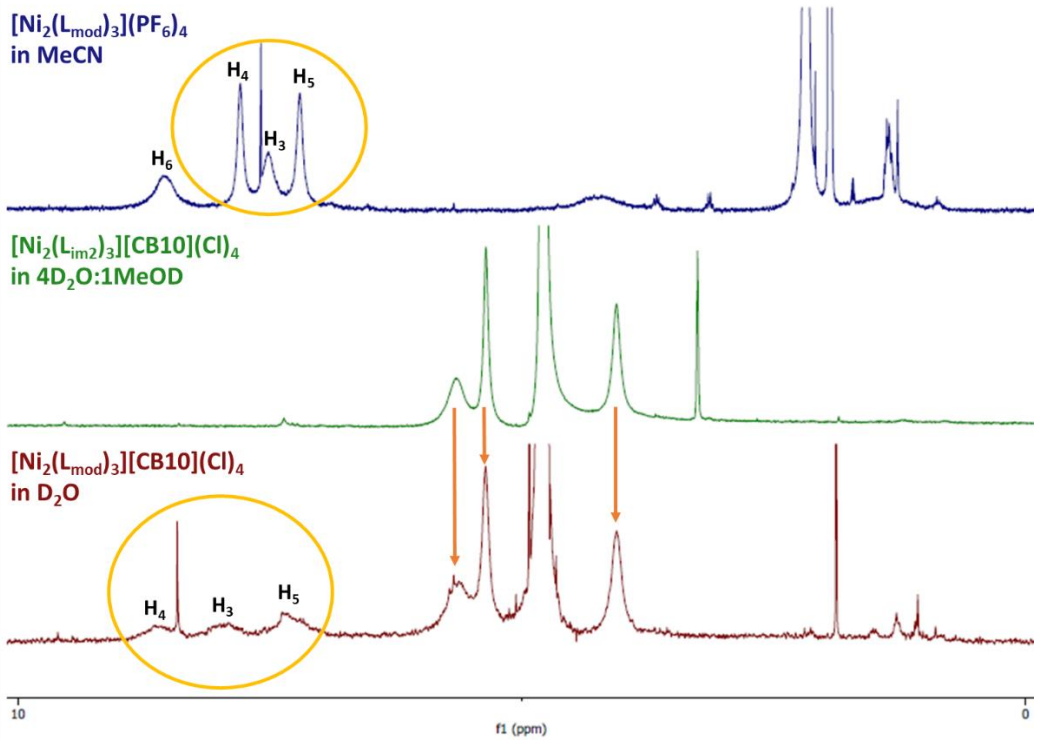


Figure 62: ^{1}H NMR between 0-10 ppm of (top-blue) the nickel cylinder after alkylation, (middle-green) the nickel cylinder thread through CB10 and (bottom-red) the rotaxane. A Gaussian function (1GB) and exponential (0.3 Hz) has been applied to the final spectrum ($[Ni_2(L_{mod})_3][CB10](Cl)_4$ in D_2O) so the peaks can be seen clearly.

The NMR of the rotaxane (bottom spectrum, Figure 62) has two main features which have been seen before in previous spectra.

The spectrum of the modified cylinder (top spectrum, Figure 62, also seen in the inset image, Figure 56), $[Ni_2(L_{mod})_3](PF_6)_4$, contains four peaks for each of the pyridine protons attached during the alkylation reaction. Three of these peaks are circled in yellow (H_3, H_4 and H_5), and the fourth peak (H_6) is seen next to them at a slightly higher shift. The same pattern of circled peaks is repeated in the rotaxane spectrum although they are broadened and shifted downfield. This deshielding is due to the change in solvent, from acetonitrile to water, which allows the pyridine nitrogen to hydrogen bond to the solvent. The fourth peak is the proton of the carbon neighbouring the pyridine nitrogen (H_6) which would be most affected by this hydrogen bonding. This peak was already much broader

than the other pyridine protons (Figure 62, blue - top) so is probably too broad to be observed in the rotaxane spectrum.

The spectrum of the 2-imidazole cylinder (middle spectrum, Figure 62, also seen in the inset image, Figure 49) clearly shows three peaks surrounding the water solvent peak which are assigned to the protons of CB10. These peaks are also observed in almost exactly the same position in the rotaxane spectra (shown by the orange arrows). These peaks have very similar shifts in both spectra, suggesting that the macrocyclic proton environments are not affected by the alkylation of the cylinder.

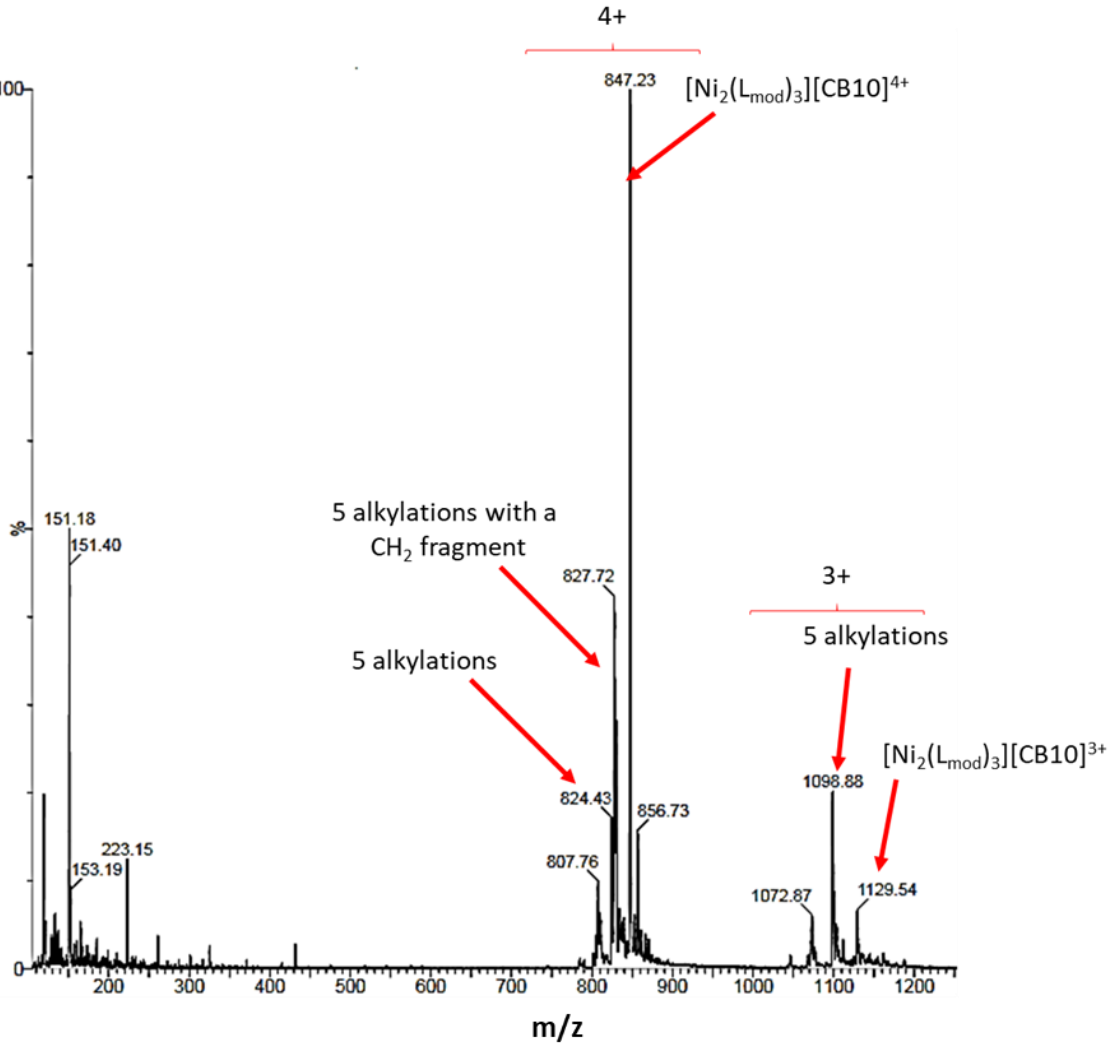


Figure 63: Mass spectrum (SCAN ES+) of the rotaxane product after quenching with $NaHCO_3$

The mass spectrum suggests that mostly the six times alkylated cylinder bound to CB10 is found in the product (m/z = 847.23). Again, some of the product appears to have been fragmented whilst recording the mass spectrum (see Figure 63).

3.6. Conclusions and Further Work

This chapter describes the successful synthesis of a rotaxane from a cylinder analogue and cucurbit[10]uril. This is the first time a supramolecular helicate has been rotaxanated.

Firstly, a post assembly modification strategy has been developed, building on previous work with the imidazole cylinders[84,85] and a parent cylinder analogue[116]. Here it has been demonstrated that while the imidazole ligand is coordinated in a helicate, the extra nitrogen of the imidazole ring is able to act as a Lewis base in an alkylation reaction with 2-(bromomethyl)pyridine hydrobromide. The alkylated product has been characterised by mass spectrometry and paramagnetically shifted ^{1}H NMR spectroscopy. This is the first example of post assembly modification using this cylinder system and, once optimised, could offer a quick synthetic pathway to novel cylinders as an alternative to long organic ligand synthesis.

A previous binding study found that the parent cylinder was a good shape and size to thread through the cavity of a CB10 macrocycle. As imidazole cylinders have similar dimensions, their binding to CB10 was also investigated by ^{1}H NMR and mass spectrometry. On addition of CB10 to and NMR sample, the obvious shifting of certain peaks in the NMR spectra indicated a change in the cylinder environment caused by binding. The lack of peak splitting suggested that every proton in each individual

environment was shifted by the same amount, retaining their equal chemical shift. The lack of peak splitting therefore suggests that the cylinder is in a symmetrical environment, and therefore must be threaded through the centre of the macrocyclic ring.

Cylinder threading followed by post assembly modification was utilised to form the first rotaxanated supramolecular helicate. After some optimisation, the mass spectrum indicated the minimum number of pyridine moieties covalently attached by alkylation onto each cylinder was four, which should be enough to prevent the CB10 ring dissociating. The low water solubility of the supramolecular structure and the broadening caused by the paramagnetic metal centres meant that a clear paramagnetically shifted ^{1}H NMR could not be recorded, however attempts at lower concentration lead to the identification of the CB10 and pyridine protons together, which gives further evidence for the formation of the rotaxane.

In future, the post assembly modification reaction should be attempted with the nickel 4-imidazole cylinder. As shown by the alkylation of 4-imdiazole carboxaldehyde, the 4-imidazoles form two different products depending on which nitrogen site is alkylated. When bound to a metal ion, one of these sites (N_1, nearest the imine bond) should be unavailable for reaction. Investigating which sites are alkylated during post assembly modification of the 4-imidazole cylinder would indicate whether the ligand remains bound to the cylinder, in which case only one alkylation site should be found, or whether the ligand is alkylated in an unbound state, allowing both nitrogen atoms of the imidazole to be alkylated. A simple way to check for multiple products would be to dissolve the modified cylinder in DMSO which is a coordinating solvent and should be able to remove the metal ions from the cylinder whilst also dissolving the ligand. The

solution could then be separated by HPLC, similar to the isolation of the modified carboxaldehyde, and the reaction products investigated by NMR.

Post assembly modification using a variety of different alkylating agents should also be attempted to see if certain moieties or characteristics are favoured. It may also be possible to form a wide range of cylinder analogues this way.

As with the parent cylinder[114], further NMR studies may be able to provide further evidence that the imidazole cylinder is thread through the CB10 cavity. In particular, through space NMR techniques such as NOESY and rotating overhauser effect spectroscopy (ROESY)[118] could be used to see which cylinder protons show strong interactions with CB10. Alternatively, diffusion ordered spectroscopy (DOSY)[119] may be used.

Finally, future studies should test whether the CB10 ring is able to dissociate from the cylinder after alkylation. Higher temperature or competitive binders, such as CB5[111], should be applied to the rotaxanated species in solution to see if the macrocycle can be released. It would also be interesting to investigate whether the CB10 ring is able to stabilise the cylinder against degradation over long periods of time or under high temperature conditions where it is likely that the imine bond will be hydrolysed.

Chapter 4:

DNA BINDING STUDIES

4.1. Previous Work and Current Aims

Although the interest in the 4-imidazole cylinders was originally their spin cross-over behaviours, their ability to bind to DNA was still tested[84,85,112]. $[Fe_2(L_{im4})_3]Cl_4$ was shown to bind to DNA at a more perpendicular angle, causing stiffening.

Similarly, the DNA binding of $[Fe_2(L_{mod})_3]Cl_4$ and $[Ni_2(L_{mod})_3]Cl_4$, which have the same structure as the alkylated cylinder, to DNA has been investigated[116]. The iron cylinder appears to be too unstable in solution to bind the DNA. Only minor changes to the DNA absorbance spectra (UV-Vis, CD, LD) are seen and it is suggested that these are due to the leftover ligand interfering with the DNA after the cylinder degrades. The nickel cylinder acts very differently in solution and the binding to DNA is confirmed by the appearance of an ICD band and MLCT band in the CD spectrum. In addition, the LD spectrum shows that the absorbance of polarised light by DNA is reduced as the base pair to $[Ni_2(L_{mod})_3]Cl_4$ ratio increases. This suggests that the modified nickel cylinder is able to bind and stiffen the DNA similar to $[Fe_2(L_{im4})_3]Cl_4$.

Unlike the 4-imdiazole cylinders and the alkylated cylinder, the DNA binding ability of the 2-imidazole cylinders has not been investigated. The aims of this chapter are to show that the 2-imidazole cylinder is able to bind DNA, and that this binding can be prevented by the CB10 macrocycle of the rotaxane. DNA binding of the 2-imidazole cylinder is investigated followed by the rotaxane binding. By comparison of the 2-imidazole binding, the rotaxane binding and previous binding studies[84,112,116], it should be possible to show whether threading the cylinder through a CB10 ring effects the DNA binding mode or efficiency.

The binding of DNA is investigated by UV-Vis spectroscopy, CD and LD. A brief introduction is provided for each method to give context to the results and discussion.

4.2. DNA Binding Studies: Solvent Systems

In order to study DNA binding, a complex must be soluble in a stock solution which has a higher concentration, 0.5-1 mM, than in the sample being measured, 2-50 μM. This is so that only small amounts of the stock solution are required to raise the concentration of the measured sample to the desired level. In some cases the complex being investigated has a low water solubility, for example $[Ni_2(L_{im2})_3]Cl_4$, which requires an alternative solvent system to be used, here a 20% methanol solution.

Although the volume of additional solvent is kept to a minimum, it is necessary to make sure any alternative solvent system does not have an effect on the DNA when added to the measured sample. Each $[Ni_2(L_{im2})_3]Cl_4$ titration has therefore been repeated with a 20% methanol stock solution as a control titration to investigate any effect that the presence of methanol may have. The volume of 20% methanol solution used for each addition exactly matches the volume of cylinder stock solution that would have been used in the original titration. The concentration and percentage volume of methanol has been measured for each sample (see Table 3).

UV Vis and CD			
bp : cyl	20% MeOH / μL	MeOH conc. / mM	MeOH / % volume
60:1	2	8.2	<0.1
40:1	3	12.3	<0.1
20:1	6	24.5	0.1
10:1	12	48.5	0.2
8:1	15	60.3	0.2
6:1	20	79.7	0.3
4:1	30	117.7	0.4

LD			
bp : cyl	20% MeOH / μL	MeOH conc. / mM	MeOH / % volume
60:1	1.3	63	0.1
40:1	2	95	0.2
20:1	4.1	187	0.4
10:1	9	337	0.8
8:1	11	486	1.0
6:1	15	570	1.3
4:1	25	824	2.0

Table 3: The volume of 20% methanol solution added for each measurement of the control titration. The base pair to cylinder ratio each addition corresponds to is provided.

4.3.UV-Vis Spectroscopy:

4.3.1.Introduction to UV-Vis

UV-Vis spectroscopy measures the absorption of different wavelengths of light by a chromophore within the ultraviolet (UV, around 200-400 nm) and visible light regions (around 400-800 nm) of the electromagnetic spectrum[120]. All wavelengths of light, λ, have an associated energy, E, which can be calculated using Equation (1)[121]. When light is absorbed, the associated energy is able to promote an electron from a low lying, filled molecular orbital (the ground state) into a higher energy, empty molecular orbital (the excited state). As the energy difference between two molecular orbitals is discrete, only the specific wavelengths of light with these energies are absorbed.

$$(1) \qquad E = \frac{hc}{\lambda}$$

The absorption of each wavelength of light is measured by a spectrophotometer, in which a beam of each wavelength of light is passed through the sample in a cuvette, and any light which is not absorbed is measured by a detector. The absorption of light by a choromophore can be directly related to its concentration by the Beer-Lambert Law (2)[121]:

$$(2) \qquad A = log\frac{I_0}{I} = \varepsilon\, c\, l$$

The intensity of the incident light, I_0, is the intensity of the specified wavelength of light emitted by the lamp, which depends on the instrument used. To measure the absorbance, A, the incident light intensity is compared to the intensity of the transmitted light, I, which is the light that is not absorbed and therefore able to pass through the sample. The path length, l, refers to the thickness of the sample or the distance over which the light can be absorbed which is dependent on the cuvette used, usually 1 cm. The molar extinction coefficient (MEC), ε, is the inherent ability of a chromophore to absorb the specific wavelength of light under stated conditions. The absorbance is therefore dependent on the concentration of a sample, the wavelength of light being measured and the external conditions of the experiment.

As the absorbance is directly related to the concentration, c, the concentration of a sample can be calculated from the absorption spectrum if the molar extinction coefficient is known. This is particularly useful for measuring the concentration of a DNA sample as the molar extinction coefficient at 260 nm is known for ctDNA, ε_{260} = 6600 $M^{-1}cm^{-1}$ and so 13200 $M^{-1}cm^{-1}$ per base and base pair respectively[122,123].

4.3.2. **Absorbance of the imidazole helicates**

The absorbance of each of the imidazole complexes was measured at room temperature. The helicate solutions were measured immediately after the sample was dissolved in order to minimise any breakdown of the cylinder that could occur in solution.

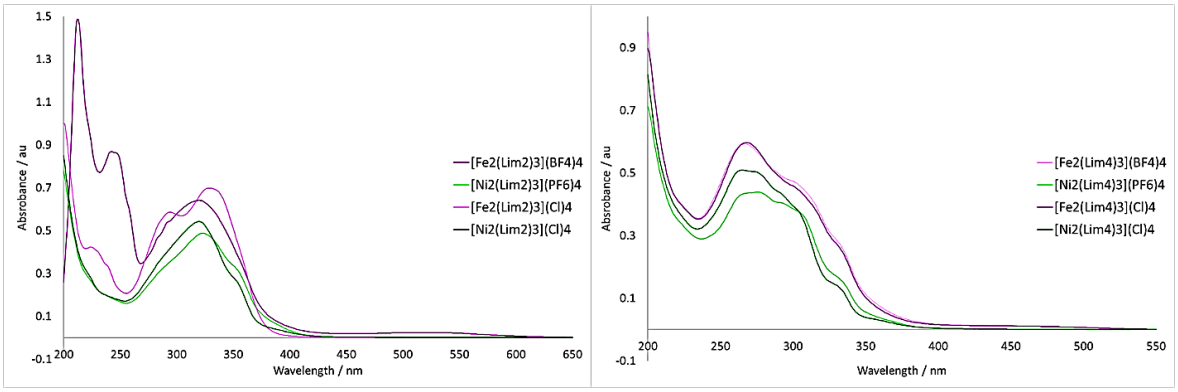


Figure 64: The absorbtion spectra of the 2-imidazole cylinders (left) and of the 4-imidazole cylinders (right). Iron cylinders are plotted in pink and nickel cylinders are plotted in green for comparison. Each inset is a more detailed view of the main absorbance peak for each spectrum. Chloride salts were dissolved in methanol while tetrafluoroborate and hexafluorophosphate salts were dissolved in acetonitrile.

The 2-imidazole ligand absorbance appears to vary dramatically with different solvents, anions and complexed metal ions (see Figure 64, left). The two nickel cylinders have similar absorption pattern with two main peaks around 320 nm and 360 nm. The absorbance of these two peaks is dependent on the solvent or anion present. The iron cylinders show two very different absorption spectra. The tetrafluoroborate salt (dark pink) has two large peaks around 210 and 245 nm and a smaller, broader peak at 320 nm. This is in contrast with the chloride salt (light pink), which has a smaller peak at 230 nm and two larger peaks at 290 nm and 330 nm. The peaks at lower wavelength (200-250 nm) and the peaks at higher wavelength (290-330 nm) could correspond to the two different spin states which are accessible by the iron helicate. When dissolved in acetonitrile, the iron cylinder forms a dark red solution, whereas when dissolved in methanol the iron cylinder begins to change colour to a pale yellow-orange colour. As this transition occurs very slowly, the spin transition may not have completely finished when the spectrum for the iron cylinder is recorded causing both species to be observed in the absorption spectrum.

When compared to the 2-imidazole cylinders, absorbance of the 4-imidazole ligand appears less dependent on the solvent, anion and complexed metal ion (see Figure 64,

right). The absorbance peak for all cylinders has a consistent shape with two main peaks around 305 nm and 330 nm. The iron cylinder peaks at 268 nm, while the nickel peak appears to split into a peak at 265 nm and 278 nm. The absorbance for the nickel cylinders is also slightly lower than for the iron cylinders.

4.3.3. **Stability and MEC of $[Fe_2(L_{im2})_3]Cl_4$**

Before completing DNA binding studies, the stability and MEC of the cylinders were investigated to find out how the cylinders behave in solution over time.

$[Fe_2(L_{im2})_3]Cl_4$ is insoluble in water and a 4:1 water to methanol solution so the stability was measured in methanol. The absorbance spectra (see Figure 65, left) shows that the signal from the cylinder appears to decrease over time. The percentage decrease in signal at 293 nm and 328 nm is 4% and 8% respectively and appears to be linear with increasing time (Figure 65, right). This could be caused by degradation of the cylinder when dissolved in methanol. It is also possible that the cylinder is slowly changing spin states, however this seems less likely, as another peak for the alternative spin state would be expected to increase if this were the case.

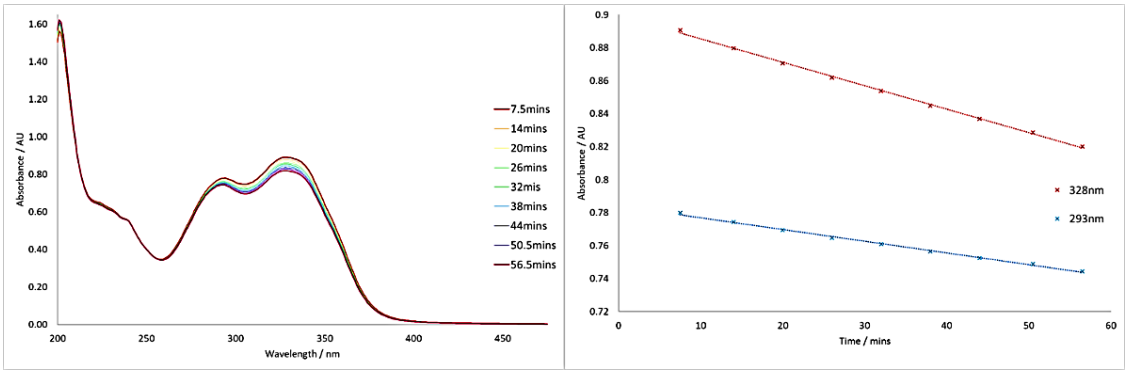


Figure 65:The absorption of a 0.99 mM sample of $[Fe_2(L_{im2})_3]Cl_4$ in methanol taken at various time points after stock cylinder solution was made (left) and the change in absorbance as a function of time at 293 and 328 nm.

The molar extinction coefficient was calculated at 293 nm and 328 nm, ε_{293} and ε_{328}. Small volumes of methanolic $[Fe_2(L_{im2})_3]Cl_4$ stock solution (1.04 mM) were diluted in methanol, the concentration calculated and the absorbance recorded. Plotting the absorbance against the concentrations gives the molar extinction coefficient as the gradient which, at 293 nm and 328 nm, is 48500 $M^{-1}cm^{-1}$ and 50700 $M^{-1}cm^{-1}$ respectively (Figure 66). This experiment did not take into account the possible breaking down of the cylinder over time. In a repeat experiment, the time between making the stock solution and recording the spectrum of each diluted sample should be recorded so that the error in each measurement could be predicted, based on the stability test of the cylinder in methanol (Figure 65). As all measurements were taken within an hour of making the stock solution in this reaction, the percentage breakdown after an hour (3.8% and 8.8% for 293 nm and 328 nm respectively) has been used to estimate the error at each point.

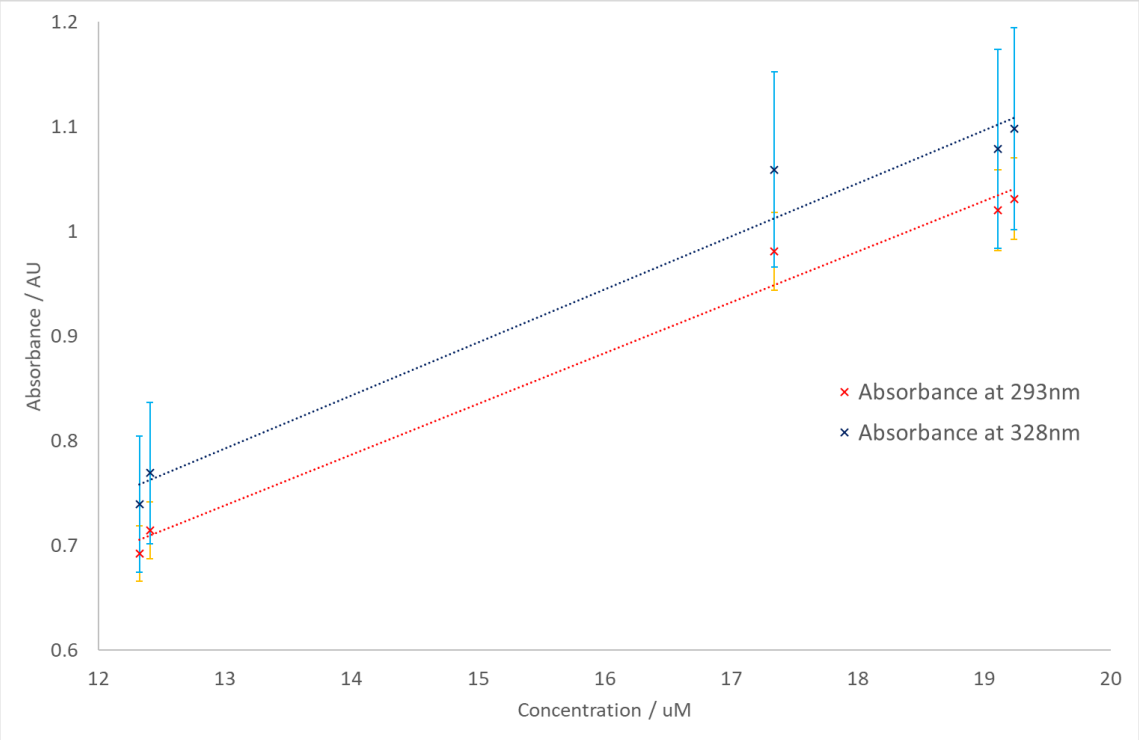


Figure 66: A plot of the absorbance of different concentrations of $[Fe_2(L_{im2})_3]Cl_4$ in methanol at 293 nm and 328 nm to calculate the molar extinction coefficient, which is equal to the gradient.

Owing to the poor solubility of this iron 2-imidazole cylinder in water, no further studies with DNA were completed.

4.3.4. **Stability and MEC of $[Ni_2(L_{im2})_3]Cl_4$**

Unlike $[Fe_2(L_{im2})_3]Cl_4$, the nickel analogue, $[Ni_2(L_{im2})_3]Cl_4$, is soluble in 20% methanol solution. The absorbance spectrum shows how the absorbance intensity of $[Ni_2(L_{im2})_3]Cl_4$ in solution decreases over a one hour period (see Figure 67). The percentage decrease in signal is only 2% at 319 nm and <1% at 353 nm which suggests that the nickel cylinder is more stable in 20% methanol solution than the iron cylinder in methanol. The slight decrease in signal appears to be regular (see Figure 68).

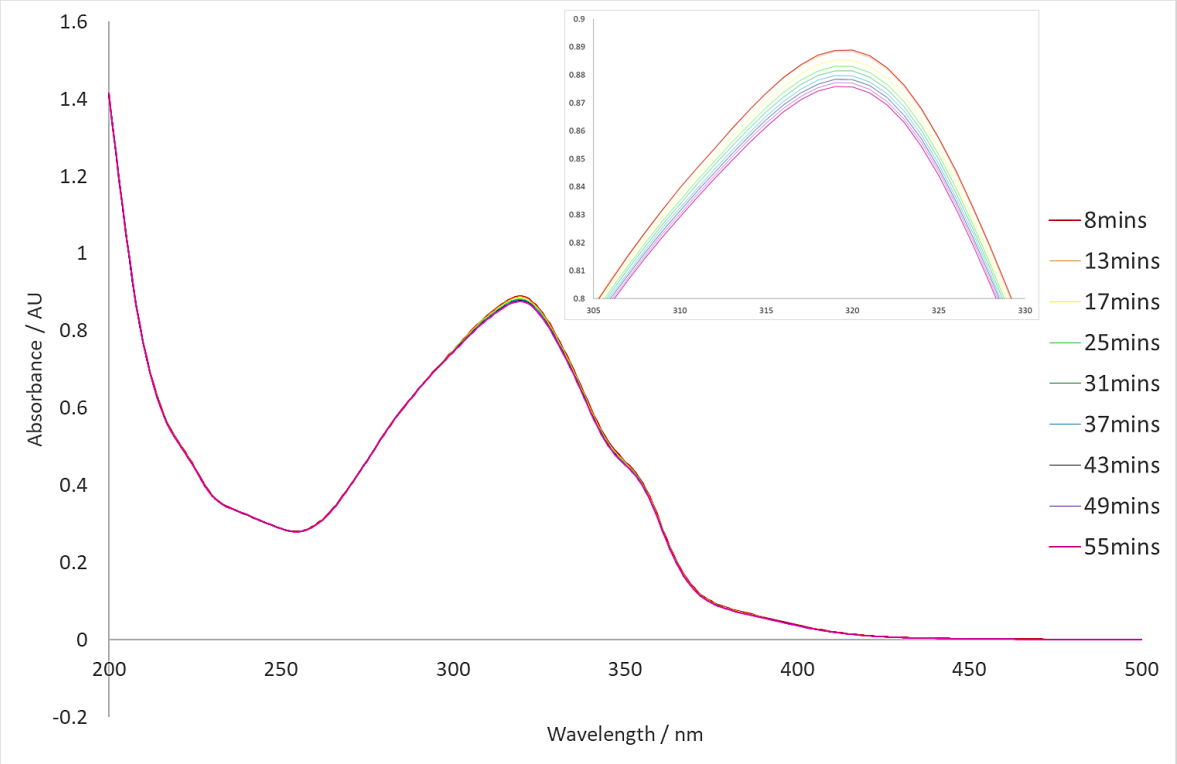


Figure 67: Absorbance of the nickel imidazole cylinder diluted from 20% methanol solution over time and, inset, the decreasing absorbance at 319 nm.

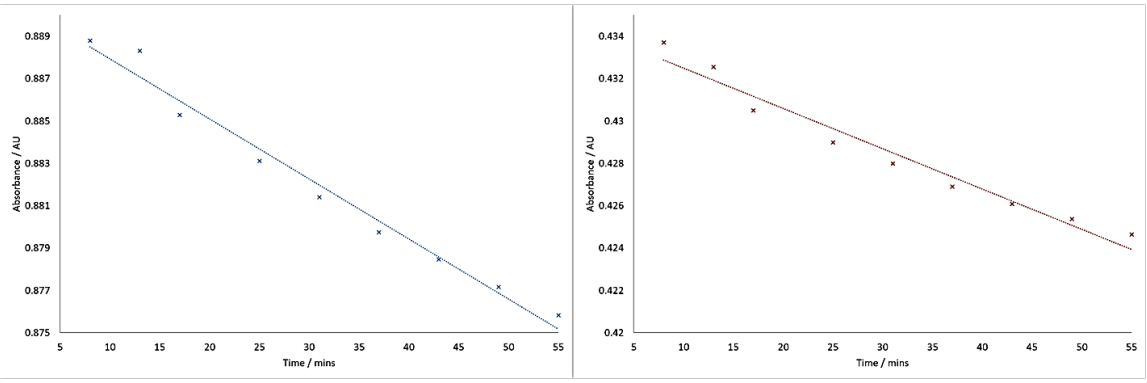


Figure 68: The absorbance of the nickel 2-imidazle cylinder decreases as a function of time at 319 nm (blue, left) and 353 nm (red, right).

The molecular extinction coefficient was also estimated for the nickel cylinder at 319 nm (767 $M^{-1}cm^{-1}$) and 353 nm (377 $M^{-1}cm^{-1}$). In an attempt to reduce the effect of possible cylinder degradation, each nickel sample was made fresh and was used immediately. This method appears to introduce a greater amount of error into the results as the data does not appear to match the line of best fit as well as for the iron cylinder (see Figure 69) which was completed by a series of dilutions. This error could originate from the weighing of small samples or not allowing the cylinder enough time to fully dissolve in solution.

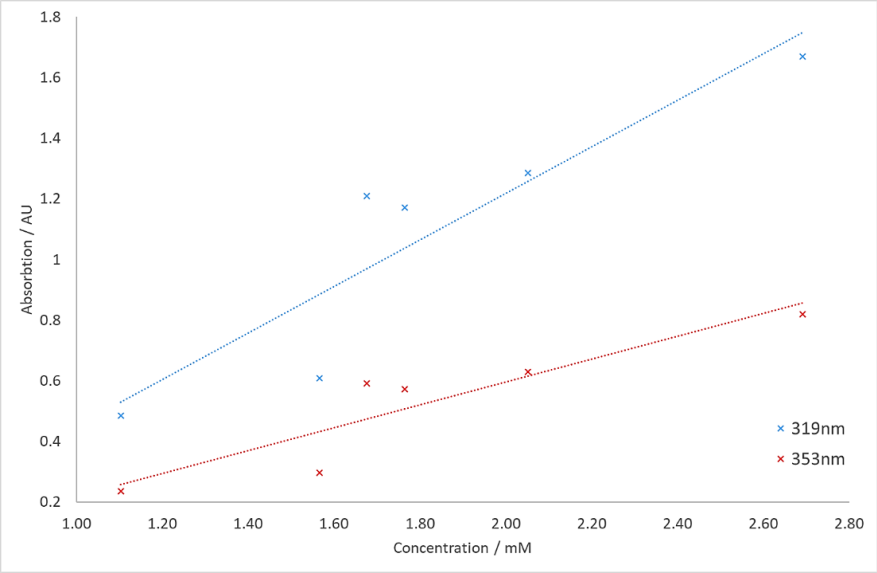


Figure 69: The absorption of $[Ni_2(L_{im2})_3]Cl_4$ as a function of cylinder concentration in 20% methanol solution.

4.3.5. **UV Vis with DNA**

As the nickel cylinder was found to be soluble in water and methanol, absorbance of DNA in the presence of the cylinder was then tested. This was done by slowly increasing the concentration of the cylinder in solution while the concentration of the DNA remained constant (100 μM). The same experiment was also repeated with increasing methanol concentration see if the small amount of methanol used to dissolve the cylinder had any effect on the DNA (Figure 70).

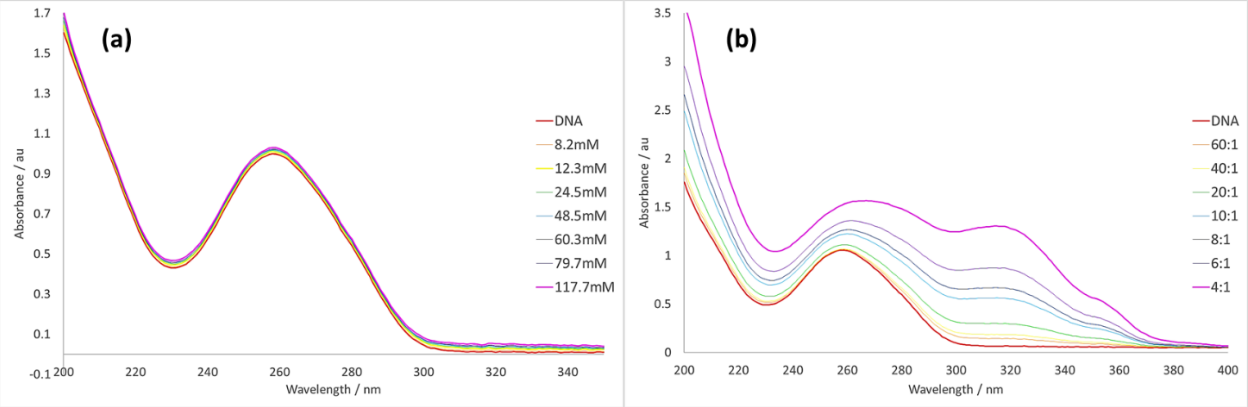


Figure 70: Absorbance spectrum of ctDNA (100 μM) in increasing methanol concentration (a) and increasing base pair to cylinder ratio (b).

Increasing the concentration of methanol in the sample solution appears to cause a small increase in the ctDNA absorbance, 258 nm, however, there are no significant changes in the spectrum (see Figure 70(b)). Increasing the base pair ratio to cylinder ratio causes an increase in both the peak at 258 nm, where ctDNA absorbs, and a peak at 319 nm, where the nickel 2-imidazole cylinder absorbs (see Figure 70(b)). The ctDNA peak also appears to shift to higher wavelengths (from 258 nm to 267 nm) as the ratio is increased from 60:1 to 4:1.

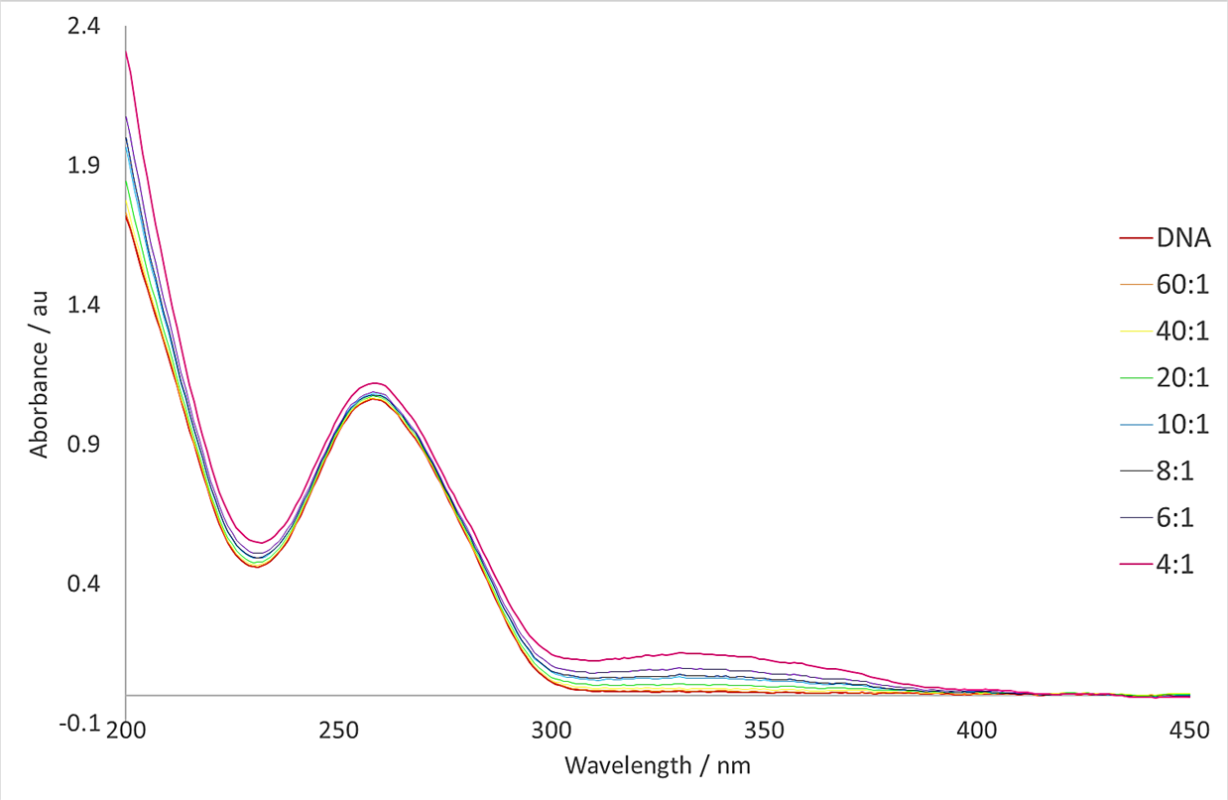


Figure 71: The UV Vis spectrum of ctDNA (100 μM) with increasing base pair to rotaxane ratio in water.

The UV Vis spectrum was also recorded for the rotaxane, $[Ni_2(L_{mod})_3][CB10]Cl_4$, in water (see Figure 71). Increasing the base pair to rotaxane ratio shows a clear increase in the intensity of a peak at 330 nm, which is likely to be where the rotaxane absorbs. The ctDNA peak intensity at 258 nm increases by 12% upon increasing cylinder to base pair ratio to 4:1, but when compared with the 2-imidazole nickel cylinder, $[Ni_2(L_{im2})_3]Cl_4$, where the absorbance peak shifts by 10 nm and increases by 48% at 258 nm, the change appears much less significant.

4.4.Circular Dichroism (CD)

4.4.1.Introduction to CD

For chiral molecules, the absorption of circularly polarised light is not only dependent on the wavelength and external conditions, but also the direction of rotation. Chiral molecules have a different molar extinction coefficient, ε, for left- and right-handed circularly polarised light causing a difference in absorbance which can be measured by circular dichroism, and this information can reveal certain structural changes of chiral molecules[121].

Helical molecules, including both DNA and the cylinders, can exist as one of two enantiomers: left handed or right handed, known as the M (Λ Λ) or P (Δ Δ) enantiomers for helicates[58,69,124]. While the cylinders synthesised by the Hannon group have to be separated into enantiomers[59], DNA primarily exists in its right handed, B-DNA form in cells. The secondary structure of B-DNA has been well studied by CD and gives a characteristic signal which is only slightly varied between different sequences of DNA. Typically a negative peak around 240 nm and a positive peak around 280 nm can be observed for randomly sequenced DNA[125]. Although there is a small dependence on sequence, such as the decrease in the peak at 280 nm with increasing A and T percentage[126], much larger changes in CD signal are caused by a change in the structure of the DNA. B-form, A-form, and Z-form DNA can be distinguished by their CD spectra due to their different structures. As mentioned in chapter 1, B-DNA and Z-DNA have opposite helicity which is shown by their CD spectra as the two signals are almost the mirror image of one another (Figure 72)[126,127].

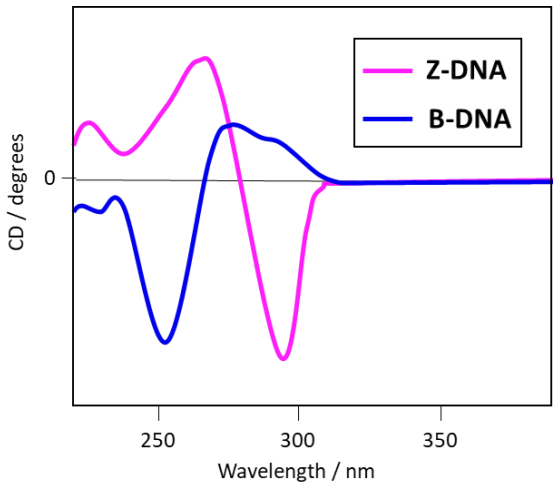


Figure 72: An example to show how the structure of DNA (here B-form and Z-form) can alter the CD spectrum to give a characteristic signal. Sketch of figure from Reference 127.

The binding of ligands to DNA can cause structural changes, and these changes have been monitored extensively by circular (and linear) dichroism. In particular, the effect of classical DNA binding modes such as intercalation and groove binding has been studied. It is sometimes possible to see large alterations to the B-DNA signal upon addition of a DNA binding agent which suggests large changes to the secondary structure of the DNA such as kinking or twisting of bases due to intercalation[128].

Another key indicator that a DNA binding event has occurred is the appearance of an induced CD band (ICD). The chiral surroundings of the bound molecule alters the ability of the molecule to absorb right and left circularly polarised light. ICD bands have a much smaller ellipticity than the DNA signal (the peaks are not as large), but allow the detection of binding modes which do not significantly perturb the secondary structure of DNA, such as groove binding[128].

Groove binders such as DAPI[129] and berenil[130] are expected to show a strong, positive ICD band as these molecules bind in a specific orientation in the groove of the DNA. Major groove binders may also show a smaller magnitude ICD signal due to the greater width of the major groove allowing greater variation in the orientation of the ligand[128]. In contrast, intercalators such as ethidium bromide[130] are expected to show a negative ICD band of much smaller magnitude[128].

The use of CD to probe DNA binding has been particularly useful in the detection of a change in binding mode due to both environmental conditions and base sequence. The CD spectra of thionine (see Figure 73), a very close relative of the DNA staining dye methylene blue, show that when bound to $[poly(dG\text{-}dC)]_2$ sequences a weak bisignate ICD band can be observed (magnitude between -1.0 to 1.5 millidegrees) indicating intercalation as a possible binding mode. In contrast, the CD spectra for binding to poly(dA).poly(dT) shows a large positive ICD band which appears to increase with increasing ionic strength (comparing "low" 5 mM to "high" 200 mM phosphate buffer), suggesting groove binding is the mode of binding for this sequence. These two different binding modes could explain the ICD band observed for the $[poly(dA\text{-}dT)]_2$-thionine complex, which has a weak bisignate shape at low ionic strength suggesting some intercalation is still occurring. As ionic strength is increased, the magnitude of the negative band decreases while the positive band becomes slightly broader suggesting a change in binding mode to groove binding[131].

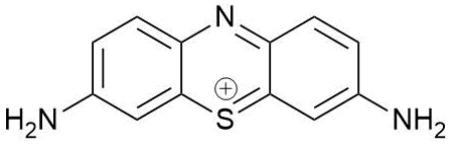


Figure 73: The structure of thionine[131].

Phenanthridine biguanides also appear to have a sequence dependent bind mode as shown by Stojković et. al.[132]. When bound to poly(dA).poly(dT), both DNA binders give a strong, positive ICD signal indicating groove binding is the preferred binding mode, whereas binding to poly(dG).poly(dC) gives a weak, negative ICD band which suggests intercalation is preferred for this sequence. The binding to ctDNA is also reported to give a strong, positive ICD with the di-substituted derivate (see Figure 74(b)) suggesting it acts as a groove binder, whereas no ICD band can be seen for the singly substituted species (see Figure 74(a)). The suggested explanation for this is that both binding modes are present for the singly substituted species, and their signals effectively cancel one another[132].

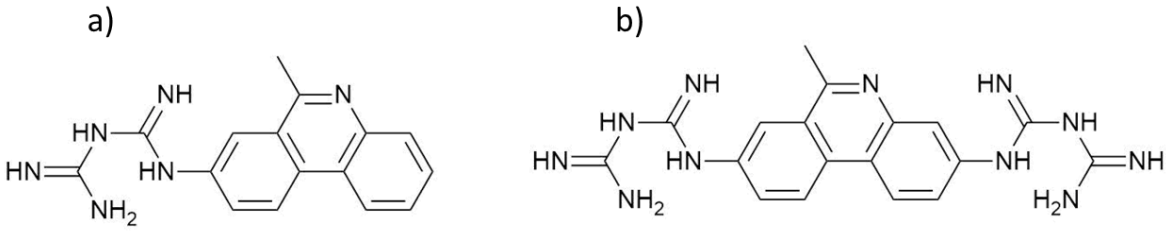


Figure 74:The structure of 1-(6-methylphenanthridin-8yl)biguanide (a) and 1-(6-methylphananthridine-3,8-diyl)biguanide (b).[132]

Modes of binding have characteristic ICD signals at wavelengths above the absorption of DNA, as well as showing changes to the expected DNA signal which can be used to help interpret CD data.

4.4.2. CD of $[Ni_2(L_{im2})_3]Cl_4$

The nickel imidazole cylinder was dissolved in a 20% methanol solution, as was used for the UV Vis studies. The cylinder stock was added to a solution of ctDNA, which was kept at constant concentration, to vary the ratio of base pairs to cylinder.

As mentioned previously, the cylinder can exist in two forms (M and P) and each of these enantiomers is expected to absorb circularly polarised light differently. This was shown in a study by Cardo *et. al.* in which CD was used to demonstrate that the attachment of L- or D- arginine enantiomers to the cylinder ligand allowed enantiospecific synthesis of the corresponding iron cylinder (the P or M enantiomer respectively)[76]. Similar experiments have also been completed for the parent cylinder (see Figure 75)[61]. The CD signals from each enantiomer are equal and opposite to one another meaning that for a racemic mixture, as used in the following experiments, the signals from the two enantiomers cancel and there is no net CD spectra. Any signal from the addition of cylinder must therefore be due to the binding of the cylinder to the DNA.

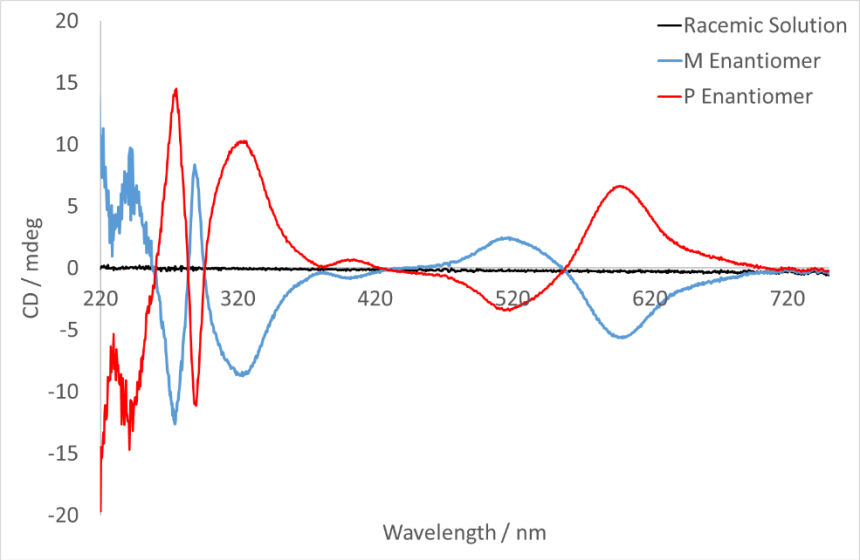


Figure 75: The difference in molar extinction coefficient of the M (solid line) and P (dashed line) enantiomers of the parent iron cylinder. Data collected by James Craig.

Between 200 nm and 300 nm the CD signal indicates that a B-DNA structure of the ctDNA is retained throughout the titration (see Figure 76). Unlike the UV Vis spectra, the peaks at around 245 nm and 275 nm decrease in magnitude with increasing cylinder ratio. A similar decrease has been seen with the parent cylinder[59] and its analogues[133].

An induced CD signal (ICD) can also be observed between 300 and 400 nm. A positive peak around 320 nm and the negative peak at 360 nm both increase in magnitude with increasing cylinder concentration.

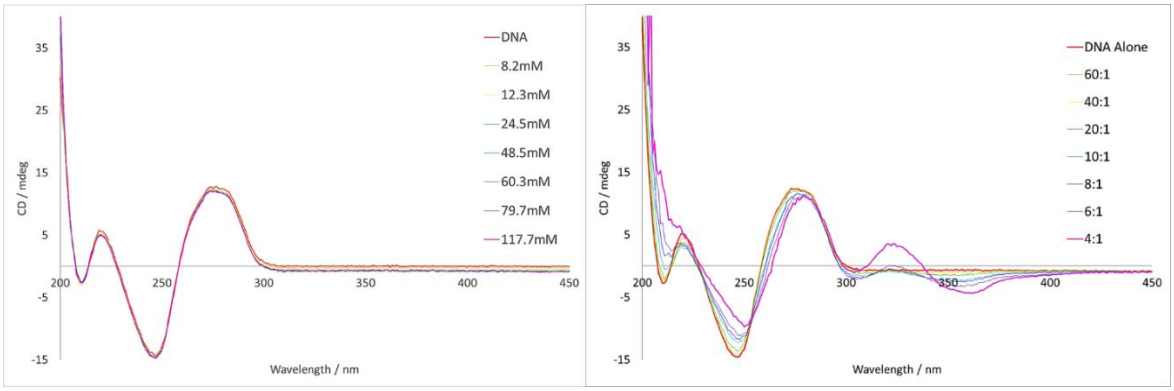


Figure 76: CD spectra of ctDNA (100 μM) with increasing methanol concentration (left) and with increasing ratio of base pairs to cylinder (right).

The presence of an ICD and the alteration in CD signal below 300 nm indicates that binding to DNA has occurred, although it is not possible to identify an exact binding mode from this data alone. The data suggests that binding by intercalation is unlikely as the ICD band is weak and bisignate. This could be due to interactions bewtween cylinders according to exciton coupling[128] which is not usually observed for intercalators due to the neighbor exclusion principle[125]. In addition, the cylinder is not expected to intercalate as it has a very different structure to the extended planar aromatic structure of most classical intercalators.

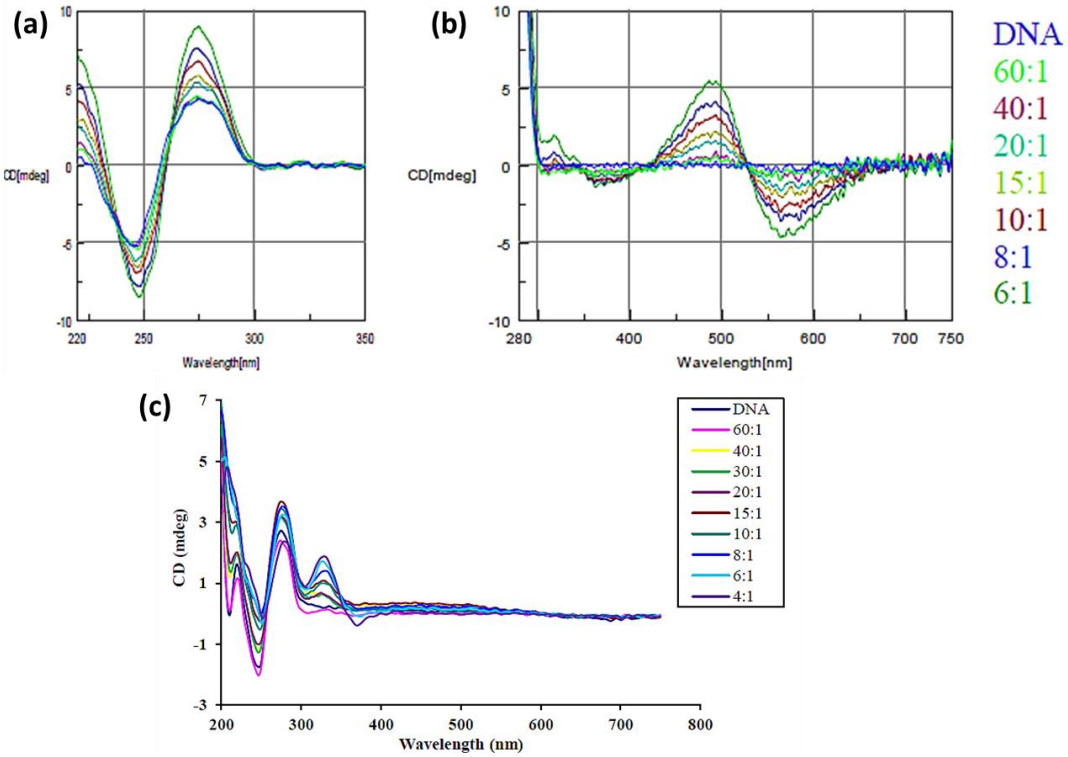


Figure 77: (a) and (b) show the binding of $[Fe_2(L_{im4})_3]Cl_4$ to ctDNA. Taken with permission from Reference 112. (c) shows the binding of $[Ni_2(L_{mod})_3]Cl_4$ to ctDNA. Taken with permission from Reference 116.

The CD spectrum for the 2-imidazole nickel cylinder very closely resembles the spectra taken for the 4-imidazole iron cylinder[84,112] (see Figure 77(a) and (b)), $[Fe_2(L_{im4})_3]Cl_4$, as well as the nickel alkylated cylinder[116] (see Figure 77(c)), $[Ni_2(L_{mod})_3]Cl_4$. The spectrum for the alkylated cylinder in particular has an ICD band similar to the 2-nickel imidazole cylinder. This suggests that the imidazole and alkylated cylinders all have a similar binding mode, implying that the the overall size and shape of the helicates are more important for DNA binding than the specific structure.

4.4.3. **CD of $[Ni_2(L_{mod})_3][CB10]Cl_4$**

The CD spectrum of the rotaxane was recorded in water (see Figure 78). Like the previous spectra, the DNA concentration was kept constant (100 μM) as the base pair to rotaxane ratio was increased from 60:1 to 4:1.

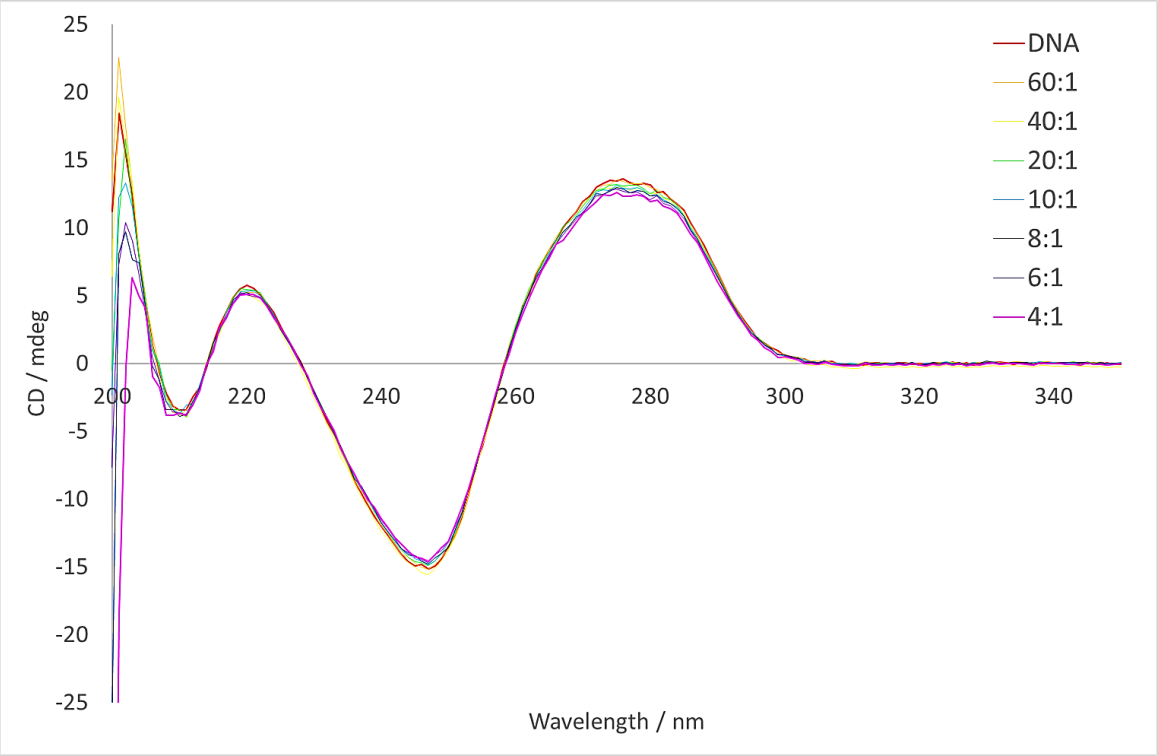


Figure 78: CD spectrum of ctDNA (100 μM) with increasing base pair to $[Ni_2(L_{mod})_3][CB10]Cl_4$ ratio in buffer solution.

There is a slight decrease in the magnitude of the ctDNA peaks upon addition of the rotaxane, with the largest percentage decrease (excluding values below 205 nm) at 276 nm (9%). This is similar to the percentage decrease in the ctDNA spectrum with methanol (5%), suggesting that the decrease is not significant. Considering there is no ICD band, there is very little evidence for a binding mode which is able to affect the helicity, or the chirality, of the DNA. This is not unexpected as CB10 is a large macrocycle which covers most of the helicate structure preventing the cylinder from

interacting with the DNA. In addition, although the parent helicate is thought to be a major groove binder[59,61], it is possible that the CB10 ring (20 Å wide by 9.1 Å high[94]) is too large to lie in the major groove of B-DNA (11.7 Å wide by 8.8 Å deep[9]).

4.5.Linear Dichroism

4.5.1.Introduction to LD

The absorption of plane polarised light is dependent on the orientation of the chromophore. Linear dichroism (LD) is the difference in absorption between plane polarised light which is parallel or perpendicular to the axis of the chromophore (3)[121]:

$$LD = A_{\parallel} - A_{\perp} \quad (3)$$

Molecules are randomly orientated in solution so the absorbance of light is an average of these orientations. It is therefore necessary to orientate a sample prior to measuring the absorbance in order to probe the effect of orientation[121]. For polymeric samples such as DNA, flow LD is used to orientate the chromophores. The sample is contained between a cylindrical rotating couette cell and a stationary internal rod which is parallel to the external cylinder. The motion of the outer cell causes the DNA within the couette to stretch so that the DNA axis aligns with the direction of flow (see Figure 79)[128].

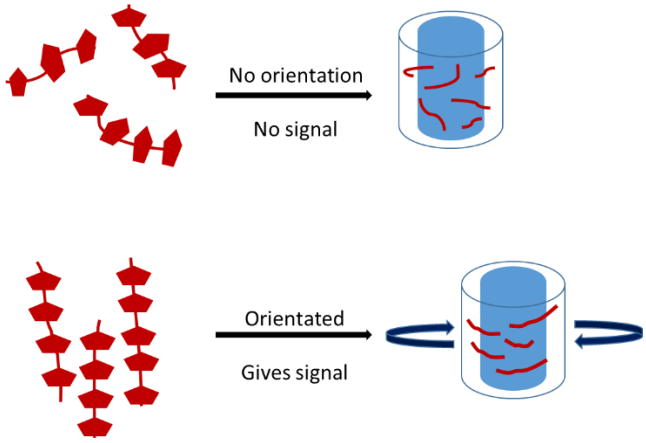


Figure 79: The effect of the rotating couette in flow LD. Inspired by figure from Reference 135.

While it is necessary to orientate a sample in order to observe an LD signal, the randomly oriented molecules in solution are still able to absorb the incident light. This means that the intensity of the incident light on the orientated sample is effectively reduced, leading to a reduced LD signal (LD^r). The LD^r can be calculated using the following equation (4), where S a measure of how aligned the sample is to the flow direction (0 = random, 1 = parallel) and α is the angle between the flow direction and the polarisation of the sample[134]:

$$(4) \qquad LD^r = \frac{A_{\parallel} - A_{\perp}}{A} = \frac{3}{2} S(3\cos^2\alpha - 1)$$

The absorption of light by the nucleobases of DNA induce a π-π electronic transition. As the bases are perpendicular to the helical axis of DNA, the absorbance of perpendicularly polarised light is greater than the absorbance of light parallel to the DNA axis. This causes the characteristic negative absorbance band seen in the LD spectrum of DNA[128,135].

Unlike CD, which is used to probe the secondary structure of DNA (for example B-DNA, A-DNA), LD gives more information about the overall structure of the DNA and its orientation. DNA binding agents and their effects on DNA structure can also be

monitored by LD. When a species binds DNA in a specific orientation it is no longer randomly oriented in solution and so an LD signal will be observed. The sign and intensity of an induced LD (ILD) peak above 300 nm wavelength provides information about the DNA binding of a sample. The extended aromatic surfaces of classical intercalators binds between the base pairs of DNA, orientating the intercalators perpendicular to the helical axis of DNA. Intercalation is therefore expected to give a negative ILD peak due to their greater absorption of perpendicularly polarised light. In comparison, rigid groove binders lie along the major and minor grooves of the DNA which are at a slight angle to the helical axis. This allows molecules bound in either groove to absorb the light polarised parallel to the helical axis, resulting in a positive ILD. The ILD bands of more flexible, non-planar binding agents is more difficult to characterise and should be considered on an individual basis[128,134].

Backbone binding and other binding modes which do not force the binding molecule into a specific orientation may not show an ILD, however binding may still be monitored by changes in the DNA LD spectrum below 300 nm wavelength. An increase in the intensity of the DNA signal implies an increase of DNA oriented parallel to the flow direction. This could show the lengthening of the DNA, possibly due to a stiffening effect caused by DNA binding. Alternatively, a reduction in the DNA intensity indicates that the amount of DNA in the pararllel orientation is decreasing and suggests that the DNA is bent, compacted or coiled by binding[128].

4.5.2. **Binding of $[Ni_2(L_{im2})_3]Cl_4$ to ctDNA**

As the nickel 2-imidazole showed poor solubility in water alone, the helicate was dissolved in a 4:1 water to methanol solution (by volume) to make a 500 μM stock solution. Before the cylinder binding to ctDNA could be tested, the effect of the methanol in solution was investigated. This was done by adding aliquots of 20% methanol solution (also containing 20% buffer) to a ctDNA solution at constant concentration (200 μM) in buffer and measuring the absorbance.

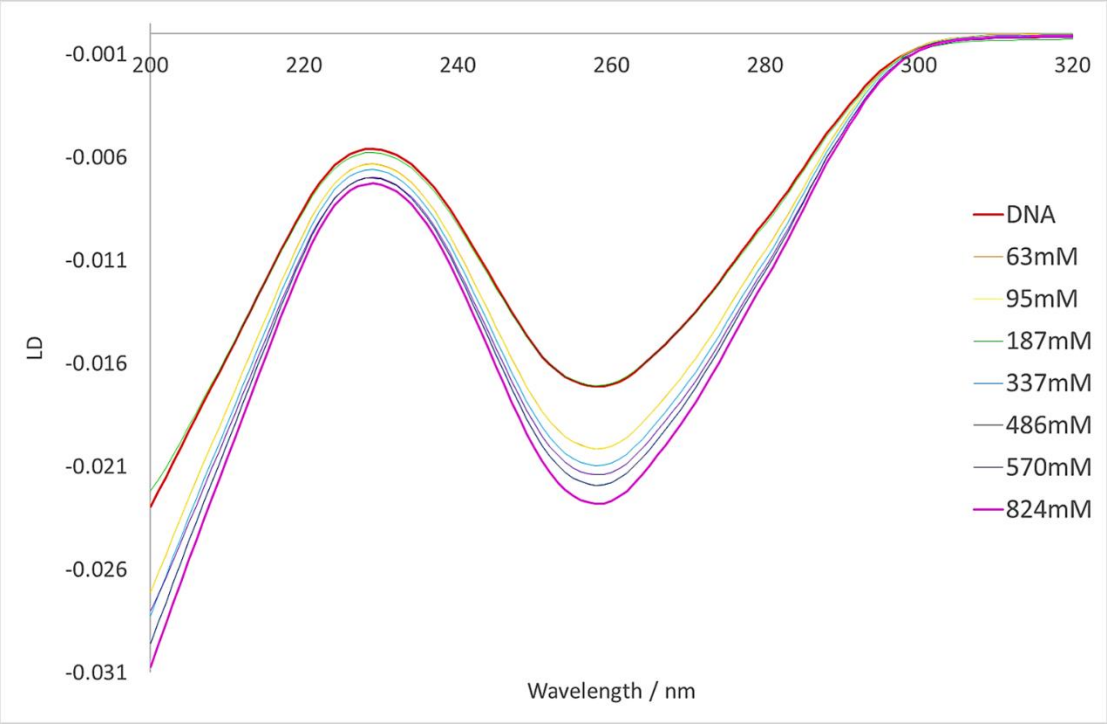


Figure 80: The LD of ctDNA (200 μM) with increasing methanol percentage (by volume).

The negative LD signal is typical of DNA however the intensity appears to be generally increasing with increasing methanol percentage (see Figure 80). This could suggest that the presence of methanol causes the ctDNA to lengthen or stiffen. An alternative explanation for this is that the amount of orientated DNA increases as the amount of time the DNA has been spun for increases. This could be tested by recording

the LD spectrum of the DNA at regular time points without any addition of ligand molecules or solvents, and should be done in future.

The intensity does not increase consistently with the percentage of methanol (or the cumulative time the DNA has been spun for). The data for 187 mM and 570 mM in particular appear to show less of an effect than the previous run (95 mM and 486 mM respectively). As expected, no absorbance was recorded above 300 nm.

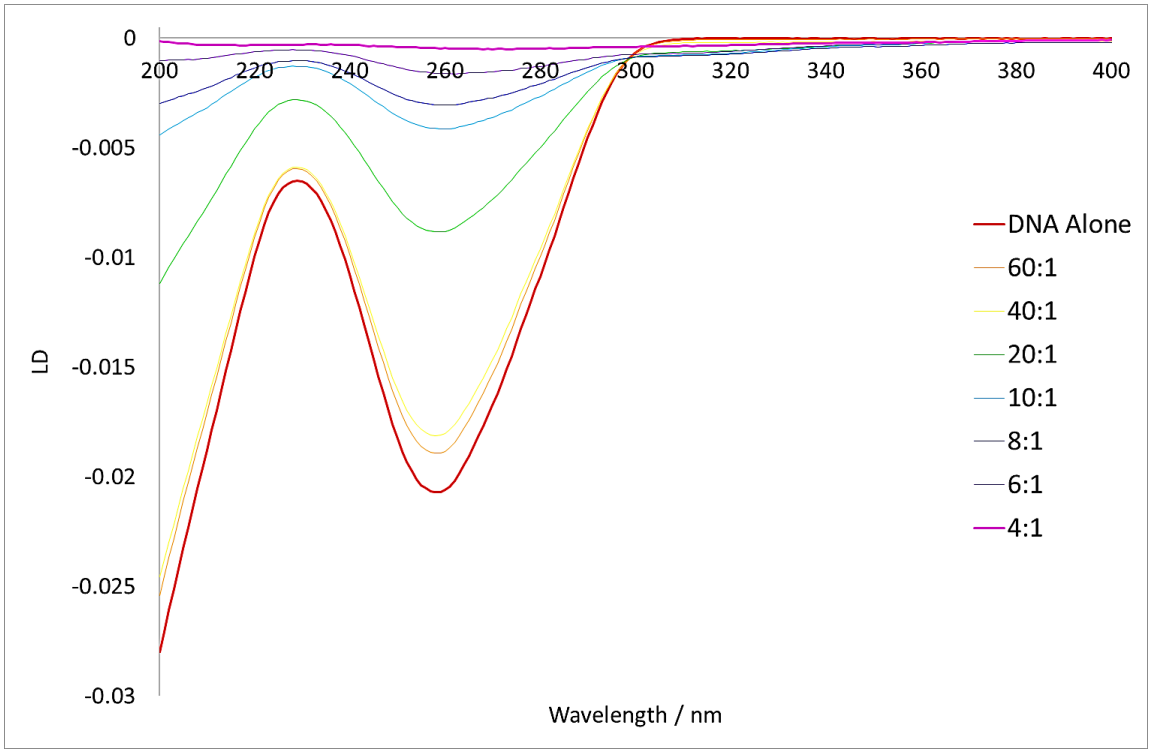


Figure 81: LD of ctDNA (200 μM) with increasing nickel 2-imidazole cylinder to base pair ratio from 60:1 to 4:1.

Unlike the addition of methanol, the magnitude of the negative DNA peak is significantly decreased with increasing cylinder to base pair ratio (see Figure 81). This suggests that the cylinder is bending or kinking the DNA, decreasing the amount of DNA aligned with the direction of flow and the absorbance of perpendicularly polarised light. Given the dramatic affect the cylinder appears to have on the ctDNA, it is possible that the helicate is coiling the DNA similar to the parent analogue[61].

A negative band also appears at 313 nm. As DNA does not absorb at this wavelength, this must be an ILD peak due to the arrangement of nickel cylinder in the sample. The intensity is very weak which could suggest very poor arrangement of the cylinder in solution or that the cylinder is orientated at an angle very close to 54.7°, at which point LD = 0 (see (4))[134].

While the methanol data shows a 33% increase (5.68×10^{-3} increase in absorbance of perpendicularly plane polarised light) in the DNA signal at 258 nm, the titration with the cylinder shows a 98% decrease (20.27×10^{-3} decrease in absorbance of perpendicularly plane polarised light) from DNA to the final measurement (95 mM methanol or 4:1 base pair to cylinder ratio). This shows that the change in absorbance is due to the addition of the cylinder and not due to the addition of methanol to the DNA.

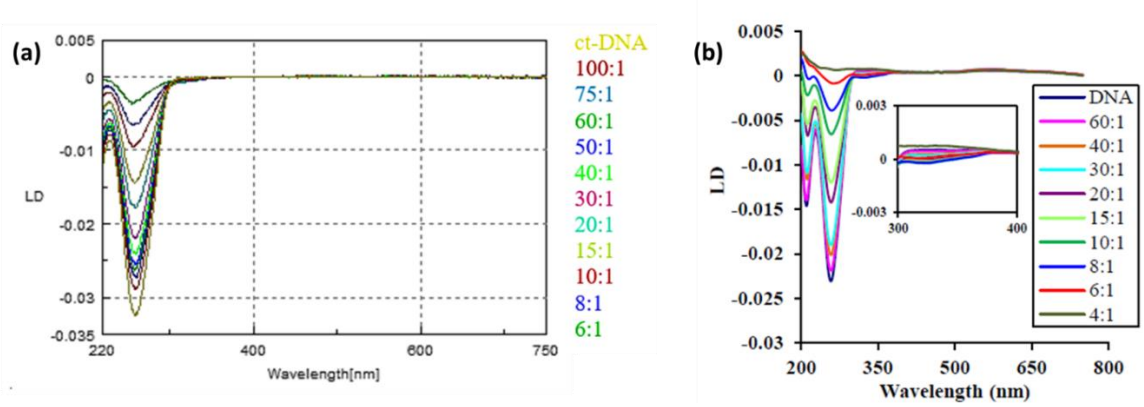


Figure 82: LD spectra of (a) $[Fe_2(L_{im4})_3]Cl_4$ and (b) $[Ni_2(L_{mod})_3]Cl_4$. Figures taken with permission from References 112, (a), and 116, (b).

These results are consistant with previous DNA binding studies using $[Fe_2(L_{im4})_3]Cl_4$ which show that the iron 4-imidazole cylinder also binds to the DNA causing a strong bending effect (see Figure 82(a))[112]. Again, the results show an even greater similarity to the previous DNA binding studies of $[Ni_2(L_{mod})_3]Cl_4$ in which the alkylated nickel cylinder causes a significant reduction in the ctDNA peak at 258 nm as

well as an induction band in a similar region to its unmodified counterpart (see Figure 82(b))[116].

4.5.3. **Binding of $[Ni_2(L_{mod})_3][CB10]Cl_4$ to ctDNA**

The binding of the rotaxanated cylinder to ctDNA was also investigated by LD. The rotaxane is soluble so no background run was required for this sample. No changes were made to the ctDNA and buffer concentrations compared with the previous experiments. The amount of sample added for each run and the concentration of the sample stock solution was also kept the same (500 μM).

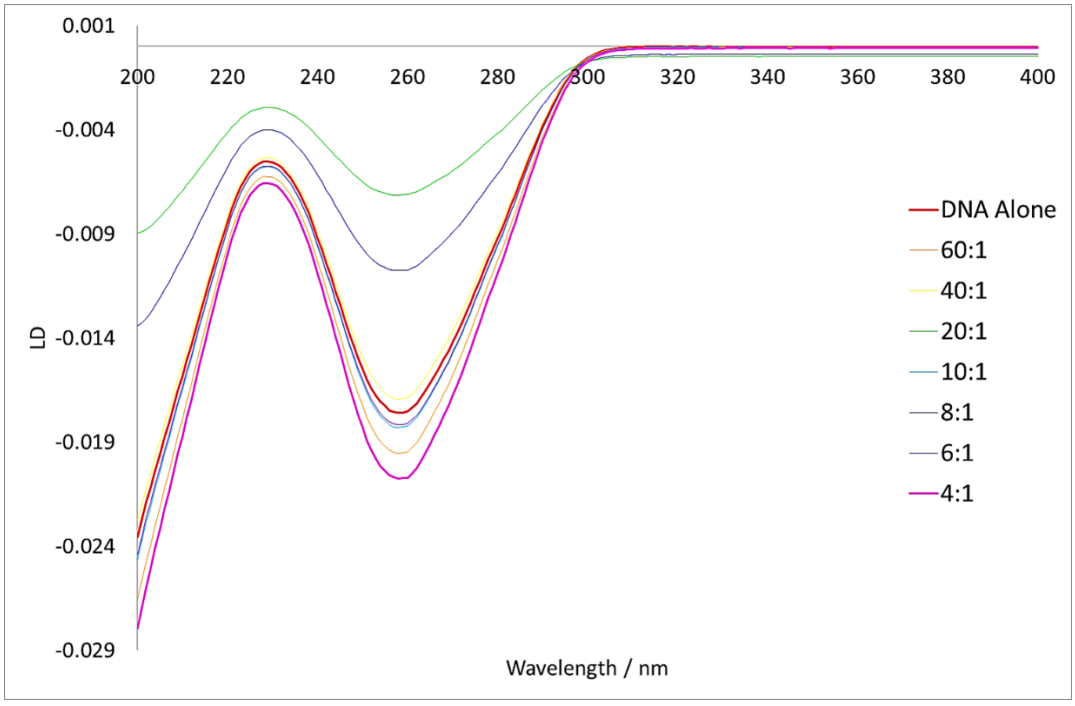


Figure 83: LD of ctDNA (200 μM) with increasing ratio of base pairs to $[Ni_2(L_{mod})_3][CB10]Cl_4$ from 60:1 to 4:1.

The LD of the ctDNA (see Figure 83) with increasing amounts of $[Ni_2(L_{mod})_3][CB10]Cl_4$ is inconclusive. The ctDNA peak intensity does not appear to

follow a pattern with increasing rotaxane to base pair ratio. Instead the signal fluctuates around the DNA absorbance of -0.018 between -0.007 (20:1) and -0.021 (4:1) at 258 nm.

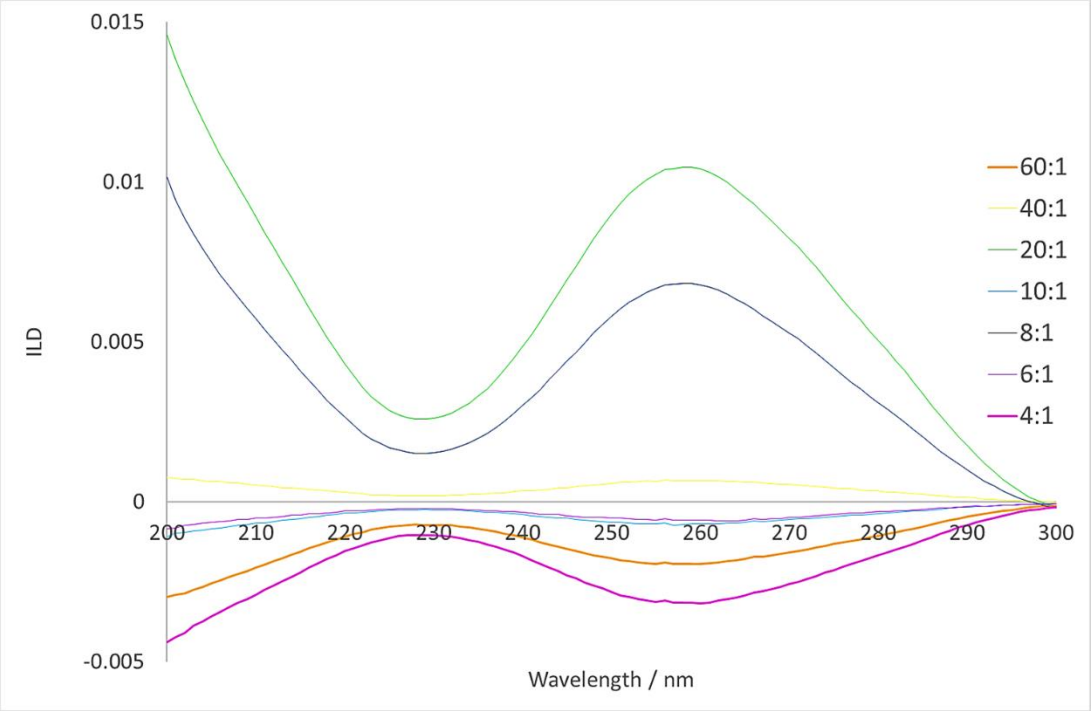


Figure 84: The change in LD, or the ILD, of the ctDNA at 258 nm as the base pair to $[Ni_2(L_{mod})_3][CB10]Cl_4$ ratio is increased from 60:1 to 4:1.

The change in LD demonstrates the lack of pattern between the data points (see Figure 84). The difference in LD at 258 nm ranges from an 18% increase to a 59% decrease in the intensity of the ctDNA signal.

4.6. Conclusions and Future Work

UV Visible spectroscopy was used to measure the absorbance of the imidazole cylinders in methanol and acetonitrile. The 2-imidazole cylinders, $[Fe_2(L_{im2})_3]Cl_4$ and $[Ni_2(L_{im2})_3]Cl_4$, were then studied in more detail to investigate their DNA binding

potential. The stability and molecular extinction coefficient of both cylinders was measured, with the results suggesting that the nickel cylinder is more stable in solution than the iron cylinder. As the iron cylinder was not soluble in aqueous solution, the DNA binding of the helicate was not investigated. While the nickel cylinder was not soluble in water at the required 1 mM concentration, the complex could be dissolved in a 20% methanol solution which was then used for DNA binding studies by UV Vis, CD and LD.

$[Ni_2(L_{im2})_3]Cl_4$ has a large effect on the DNA absorbance spectrum when increasing the base pair to cylinder ratio, increasing the absorbance by 48% at 258 nm and shifting the maximum intensity by 10 nm. CD shows that the B-DNA structure of the ctDNA is retained throughout the titration with the nickel cylinder, however the peaks shift to higher wavelengths and the magnitude of the peaks is reduced with increasing cylinder concentration. A bisignate ICD band appears with a positive peak at 360 nm and a negative peak at 320 nm which suggests that a DNA binding event is occurring. In addition, the LD studies show that the imidazole cylinder binds the DNA causing the magnitude of the LD signal of ctDNA to decrease. This suggests a binding mode which causes shortening, kinking or potentially coiling of the DNA. DNA binding of the nickel 2-imidazole cylinder has been compared to binding by the iron 4-imidazole cylinder and the alkylated nickel cylinder, both of which have been studied[84,112,116]. The spectra show that the binding of all three cylinders to ctDNA is very similar.

The DNA binding of the nickel rotaxane, $[Ni_2(L_{mod})_3][CB10]Cl_4$, was investigated to see whether the CB10 ring would prevent the same binding efficiency as the nickel 2-imidazole cylinder. A clear peak at 330 nm was seen in the absorbance spectrum which increased in intensity as the concentration of rotaxane was increased. This peak is likely due to absorbance by the rotaxane. The absorbance of ctDNA at 258 nm showed very

little increase on increasing base pair to rotaxane ratio. The CD spectrum of the ctDNA also remained largely unchanged by the addition of rotaxane and an ICD band was not recorded as the rotaxane concentration was increased. There is no evidence to suggest that the rotaxane is binding to the DNA, and the data confirms that the rotaxane does not share a binding mode with the nickel 2-imidazole cylinder. The LD spectrum fluctuates dramatically with increasing base pair to rotaxane ratio. It is possible that the rotaxane has multiple binding modes which cause the rapid changes in the LD intensity, however there is no ILD band to suggest that binding is occurring.

The data strongly suggests that the rotaxane is not binding to the ctDNA in a specific mode, and is unlikely to be binding at all. While increasing the concentration of the 2-imidazole cylinder showed consistant trend during the titration experiments, the rotaxane appears to have no effect when studied by UV Vis or CD, and an inconsistent and varied effect during the LD titration. This was expected, as the addition of such a large macrocyclic component not only changes the size and shape of the cylinder, but also prevents the cylinder from interacting with the DNA in close proximity.

The rotaxane demonstrates a new method of DNA binding control. Trapping a CB10 macrocycle around the 2-imidazole cylinder in the form of a rotaxane prevents the DNA binding interaction observed for the cylinder alone. This property could be utilised by investigating different systems which enable the CB10 ring to be released from the cylinder in response to an external stimulus, allowing DNA binding ability of the cylinder to be switched on.

The next stage of experiments should aim to gather more information about the interaction, or lack of interaction, between the rotaxane and DNA. The UV Vis spectrum

should be recorded and compared to the CD and LD to confirm if any form of binding is taking place. The LD titration should also be repeated to see if the fluctuating binding pattern is repeated.

To aid the analysis of the UV Vis, CD and LD spectra, the nickel 2-imidazole cylinder and rotaxane should be studied by agarose gel electrophoresis with ctDNA. The results should give a clear indication of which complexes are able to bind to the DNA.

As the parent cylinder has been shown to bind to DNA three way junctions[53,63], the binding of the 2-imidazole nickel cylinder and the rotaxane to three way junctions should also be investigated. This study would be particularly relevant for the rotaxane, as the CB10 ring directly covers the aromatic core of the cylinder that usually facilitates π-stacking onto the free bases of the three way junction. Without this recognition site, the rotaxane should not be able to bind the three way junction.

Chapter 5:

CONCLUSIONS AND FUTURE APPLICATIONS

Post assembly modification of a close parent cylinder analogue, the imidazole cylinder, has been achieved by alkylation. This is the first successful modification of a pre-formed cylinder and offers an alternative synthetic pathway for the production of novel cylinders.

The development of post assembly modification has granted access to new supramolecular architectures, including rotaxanes, and the first rotaxane from a supramolecular helicate has been synthesised. Initial DNA binding studies suggest that the presence of a trapped CB10 ring prevents the rotaxane from binding to ctDNA. It is likely that the macrocycle will also impede the cylinder binding to three way junctions, as the central recognition site cannot interact with the DNA.

An investigation into the synthesis of novel cylinders lead to the production of an asymmetric pyridine-imidazole ligand that offers two different binding sites for metal coordination. Complexes formed from this ligand require further investigation due to their poor solubility. The trinuclear ligand appeared to form in the presence of iron(II) chloride, but no evidence of the complex could be found. The half ligand has great potential for the synthesis of new ligands, but the purification process and yield must be improved before further work continues. This could be done by trying alternative purification processes, or the design of a new synthetic method to yield the half ligand.

The applications of this work are varied. The alkylation reaction of the post assembly modification should not require the 2-(bromomethyl)pyridine hydrobromide. The pyridine ring could be replaced by substituents that increase the ability of the cylinder to cross the cell membrane or allow specific targeting of the cylinder in cells.

Alternatively, suitable fluorescent or radioactive probes could be used to image the cylinder.

To produce a rotaxane, the only requirement of the added moiety is that it is large enough to prevent the ring of choice from being removed. Alternative bulky alkylating agents should be investigated for their ability to achieve post assembly modification. Particular attention should be paid to those that possess a functional group or bond which can be degraded in response to the local environment or external stimuli. These functional groups would include an enzyme cleavable bond, such as an amide, or a photosensitive moiety which would facilitate the controlled removal of the bulky stopper groups of the rotaxane. As the data collected suggests that the CB10 ring prevents DNA binding by the cylinder, the controlled release of the cylinder from the centre of the macrocycle would enable DNA binding only in the correct environment or when triggered by a stimulus. In future, this method could have potential for cell and tissue targeting.

Chapter 6: EXPERIMENTAL

6.1.Standard Procedures

All reagents and solvents were purchased from Sigma Aldrich, Fischer Scientific and VWR. No further purification was required. All water was deionised for use in synthesis (as a solvent or when using an aqueous solution) and collected from an Egla Option 3 purifier.

Deuterated solvents for NMR were purchased from Sigma Aldrich and Goss Scientific. Bruker AVIII300 (300 MHz), Bruker AVIII400 (400 MHz) and Brüker NEO 400 (400 MHz) NMR spectrometers were used to record all NMR spectra. Any functions applied to the spectra (Gaussian/exponential) have been stated, chemical shifts are measured in parts per million (ppm) and coupling constants are measured in Hertz (Hz). ^{1}H COSY, ^{1}H-^{13}C HSQC and ^{1}H-^{13}C HMBC have been used to aid assignments. Multiplets have been stated with each assignment: singlet (s), doublet (d), triplet (t), quartet (q), doublet of doublets (dd), doublet of triplets (dt), triplet of doublets (td) and doublet of doublet of doublets (ddd).

Mass spectra have been recorded using a Waters GCT TOF (TOF MS EI+) spectrometer by staff at the University of Birmingham or a Waters SQD (SCAN ES+) spectrometer.

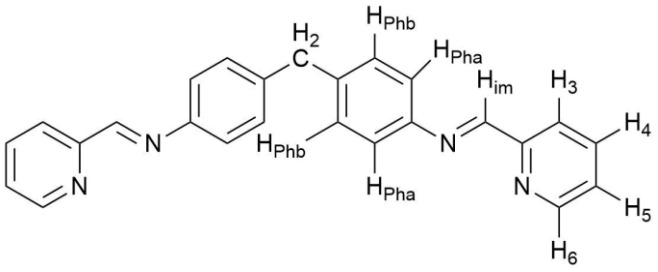
A Shimadzu LC20 Prominence system and reverse phase column (250 nm x 4.6 nm) was used for reverse phase high pressure liquid chromatography (RP HPLC). Dionex UltiMate 3000 pump and a reverse phase column (250 nm x 21.2 nm) was used for preparative RP HPLC (prep-RP HPLC). Both columns were purchased from Phenomenex, and solvents were purchased from Fischer Scientific. Solvents were degassed with helium by the staff at the University of Birmingham. RP HPLC of the half

ligand (see section 6.2.5) was performed by Callum Campbell, and data collection for the CD spectra of the parent cylinder enantiomers (see section 4.4.2) was performed by James Craig.

6.2.Novel Cylinders

6.2.1.Synthesis of the Parent Ligand, L

Based on synthesis by Hannon *et al.* [56].



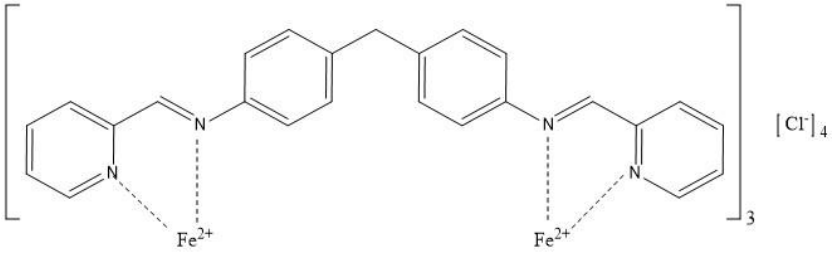
4,4'-Methylenedianiline (1.05 g, 5.31 mmol) was dissolved in ethanol (50 mL). Pyridine-2-carboxaldehyde (1.01 mL, 10.62 mmol) was added dropwise and was stirred for 6 hours. A white precipitate was collected by vacuum filtration and washed with ethanol (2 x 10 mL) before being dried *in vacuo* (1.85 g, 4.91 mmol, 93%).

^{1}H NMR: (300 MHz, Chloroform-*d*) δ 8.71 (H_6, ddd, *J* = 4.9, 1.8, 1.1 Hz, 1H), 8.62 (H_{im}, s, 1H), 8.20 (H_3, dt, *J* = 7.9, 1.1 Hz, 1H), 7.81 (H_4, ddd, *J* = 7.9, 7.5, 1.8 Hz, 1H), 7.36 (H_5, ddd, *J* = 7.5, 4.9, 1.1 Hz, 1H), 7.26 ($HPh_{a/b}$, s, 4H), 4.04 (CH_2, s, 1H).

Mass spectrum (TOF MS EI+): *m/z* = 337 $[C_{25}H_{20}N_4 + H]^+$

6.2.2. **Synthesis of the Parent Cylinder, $[Fe_2(L)_3]Cl_4$**

Based on synthesis by Hannon *et al.*[56].



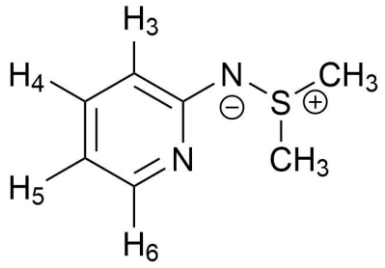
The parent ligand, L (0.40 g, 1.06 mmol) was dissolved in methanol (20 mL) and iron(II) chloride tetrahydrate (0.14 g, 0.71 mmol) in methanol (20 mL) was added dropwise. The solution was heated under reflux for 2 hours before the solvent was removed *in vacuo*. The resultant purple powder product was redissolved in minimal methanol, saturated methanolic ammonium hexafluorophosphate (4.62 g, 28.32 mmol) was added dropwise and the solution was stirred for 30 minutes. The purple precipitate was collected by filtration and washed with water (2 x 10 mL) then diethyl ether (4 x 25 mL). The precipitate was resuspended in methanol and stirred for 30 minutes with Dowex (1x8 chloride form, 200-400 mesh). The Dowex beads were then removed by vauum filtration and washed with methanol (2 x 10 mL) before the solvent was removed from the filtrate *in vacuo*. The filtrate was dissolved again in minimal methanol and added dropwise to diethyl ether (500 mL) whilst stirring. The precipitate was isolated by vacuum filtration and washed with ether (2 x 10 mL) before being dried *in vacuo* to leave a purple crystalline product (0.33 g, 0.24 mmol, 67% yield).

^{1}H NMR: (300 MHz, Methanol-d_4) δ 9.19 (H_{im}, s, 1H), 8.71 (H_3, d, J = 6.7 Hz, 1H), 8.51 (H_4, ddd, J = 7.6, 6.7, 1.2 Hz, 1H), 7.88 (H_5, ddd, J = 7.6, 5.6, 1.4 Hz, 1H), 7.48 (H_6, d, J = 5.6 Hz, 1H), 7.06 ($HPh_{a/b}$, s, 2H), 5.63 ($HPh_{a/b}$, s, 2H), 4.09 (CH_2, s, 1H).

Mass spectrum (TOF MS EI+): *m/z* = 310 $[Fe_2(C_{25}H_{20}N_4)_3]^{4+}$, 425 $[Fe_2(C_{25}H_{20}N_4)_3Cl]^{3+}$, 467 $[Fe_2(C_{25}H_{20}N_4)_3Cl_2]^{2+}$

6.2.3. Synthesis of S,S-dimethyl-N-(2-pyridyl)sulphilimine

Based on synthesis methods by L. J. Childs[80], J. C. White[81] and S. Vitorino[82].



2-Aminopyridine (5.65g, 0.06mol) and dimethyl sulphide (4.84 mL, 0.07mol) were dissolved in dichloromethane (40 mL) under an argon atmosphere and cooled to -41°C. N-chlorosuccinimide (8.01g, 0.06mol) in dichloromethane (40 mL) was added dropwise over 2 hours and stirred for a further 2 hours 30 minutes at reduced temperature. Sodium methoxide (3.68g, 0.07mol) in methanol (18 mL) was added and stirred for 15 minute at -41°C. Water (25 mL) was added and the solution stirred overnight. The organic phase was collected and the aqueous phase extracted with dichloromethane (2x10 mL). The organic phases were combined, washed with water (15 mL) and dried over magnesium sulphate. The solvent was removed *in vacuo* to yield a yellow oil (7.14g, 0.05mol, 77%).

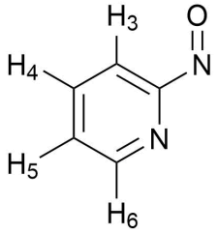
A trace amount of starting material was seen in the ^{1}H NMR. These peaks have been assigned "SM-".

^{1}H NMR: (300 MHz, Chloroform-*d*) δ 7.98 (SM-H_6, ddd, *J* = 5.2, 1.9, 0.9 Hz, trace), 7.92 (H_6, ddd, *J* = 5.3, 1.9, 1.0 Hz, 1H), 7.35 (SM-H_4, ddd, *J* = 8.3, 7.2, 1.9 Hz, trace), 7.25 (H_4, ddd, *J* = 8.4, 7.0, 1.9 Hz, 1H), 6.64 (H_3, dt, *J* = 8.4, 1.0 Hz, 1H), 6.56 (SM-H_5, ddd, *J* = 7.2, 5.2, 1.0 Hz, trace), 6.44 – 6.38 (SM-H_3 and H_5, m, 1H), 4.60 (SM-NH2, d, *J* = 23.1 Hz, trace), 2.66 (CH3, s, 6H).

Mass spectrum (TOF MS EI+): *m/z* = 80 $[C_5H_5N + H]^+$, 139 $[C_6H_7N_2S]$, 154 $[C_7H_{10}N_2S]$

6.2.4.<u>Synthesis of 2-nitrosopyridine</u>

Based on synthesis methods by L. J. Childs[80], J. C. White[81] and S. Vitorino[82].



S,S-dimethyl-N-(2-pyridyl)sulphilimine (7.00g, 0.05mol) in dichloromethane (25 mL) was added to 3-chloroperoxybenzoic acid (mCPBA, 13.33g, 0.08mol) in dichloromethane (25 mL) under an argon atmosphere. The solution was stirred for 90 minutes before dimethyl sulphide (1.69 mL) was added and stirred for a further 30 minutes. Saturated sodium carbonate solution (25 mL) was added and stirred for 3 hours before water (100 mL) was added. The green organic layer was collected, washed with water (2 x 30 mL) and dried over magnesium sulphate. The solvent was removed *in vacuo* to yield an orange solid product (1.66g, 0.02mol, 34%).

The product was recrystallized from hot ethanol to yield off-white, needle-like crystals (0.90g, 0.01mol, 19%).

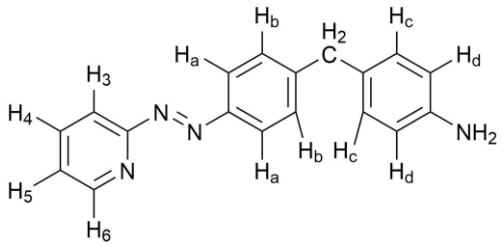
Tautomers have been assigned as a and b.

^{1}H NMR: (400 MHz, Chloroform-*d*) δ 8.76 (H6a, d, *J* = 4.6 Hz, 1H), 8.03 (H4a, td, *J* = 7.7, 1.8 Hz, 1H), 7.96 (H6b, ddd, *J* = 4.9, 1.8, 1.0 Hz, 2H), 7.88 (H4b, ddd, *J* = 8.1, 7.4, 1.8 Hz, 1H), 7.81 (H3b, dt, *J* = 8.1, 1.0 Hz, 2H), 7.61 (H5a, ddd, *J* = 7.7, 4.6, 1.0 Hz, 1H), 7.34 (H3a, dt, *J* = 7.7, 1.0 Hz, 1H), 7.25 (H5b, ddd, *J* = 7.4, 4.9, 1.0 Hz, 2H).

Mass spectrum (TOF MS EI+): *m/z* = 94 $[C_5H_6N_2]^+$, 108 $[C_5H_4N_2O]^+$

6.2.5. **Synthesis of half ligand, L_{half}**

Based on synthesis methods by L. J. Childs[80], J. C. White[81] and S. Vitorino[82].



2-Nitrosopyridine (0.87 g, 8.08 mmol) in dichloromethane (10 mL) was added dropwise to 4,4'-methylenedianiline (6.41 g, 32.31 mmol) dissolved in dichloromethane (20 mL) under an argon atmosphere. Glacial acetic acid (1 mL) was added and the solution was stirred for 70 hours. The solvent was removed *in vacuo* to yield a dark red oil which was dried overnight *in vacuo* (7.43 g crude product).

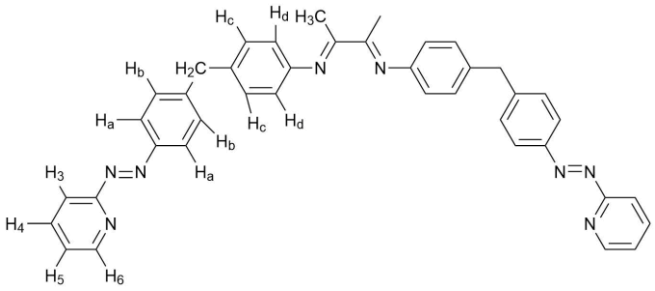
Prep-RP HPLC was used to purify the product. A water to acetonitrile gradient was applied over an hour. This was performed by Callum Campbell.

^{1}H NMR: (300 MHz, Acetonitrile-d_3) δ 8.69 (H_6, ddd, J = 4.8, 1.9, 0.9 Hz, 1H), 7.96 (H_4, ddd, J = 8.1, 7.4, 1.9 Hz, 1H), 7.89 ($HPh_{a/b}$, dt, J = 8.6, 2.0 Hz, 2H), 7.71 (H_3, dt, J = 8.1, 0.8 Hz, 1H), 7.47 (H_5, ddd, J = 7.4, 4.8, 0.9 Hz, 1H), 7.41 ($HPh_{a/b}$, dt, J = 8.6, 2.0 Hz, 2H), 6.99 ($HPh_{c/d}$, dt, J = 8.6, 1.9 Hz, 2H), 6.59 ($HPh_{c/d}$, dt, J = 8.6, 1.9 Hz, 2H), 4.03 (NH_2 s, 2H), 3.93 (CH_2, s, 2H).

Mass spectrum (TOF MS EI+): m/z = 289 $[C_{18}H_{16}N_4 + H]^+$, 311 $[C_{18}H_{16}N_4 + Na]^+$, 599 $[2(C_{18}H_{16}N_4) + H]^+$

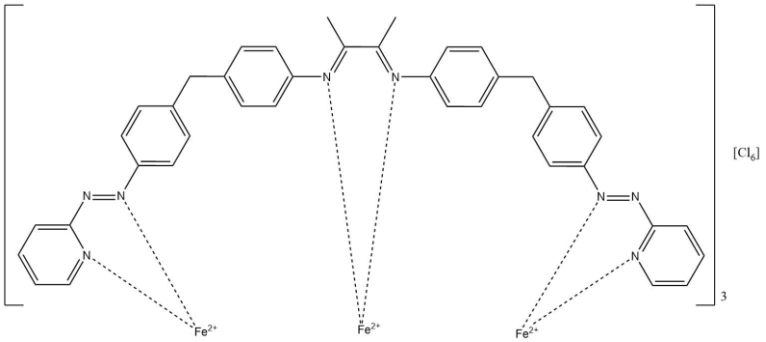
6.2.6. **Attempted synthesis of a trinuclear ligand, L_{tri}**

Based on synthesis methods by L. J. Childs[80].



Molecular sieved (2 A, powder) were activated by heating at 200°C for 2 days. The half ligand (0.02 g, 0.08 mmol) was dissolved in dry dichloromethane (5 mL) containing the molecular sieves under an argon atmosphere. A solution of 2,3-butanedione (0.1 mL, 1.14 mmol) in dichloromethane (9.9 mL) was made (0.11 M). The 2,3-butanedione solution (0.4 mL) was added to the half ligand solution and stirred for 2 days. The molecular sieves were removed by vacuum filtration and washed with dichloromethane (10 mL) to yield a clear orange solution. The solvent was removed *in vacuo* and the orange-brown oil product was dried *in vacuo* overnight (0.03 g, 0.06 mmol, 131%).

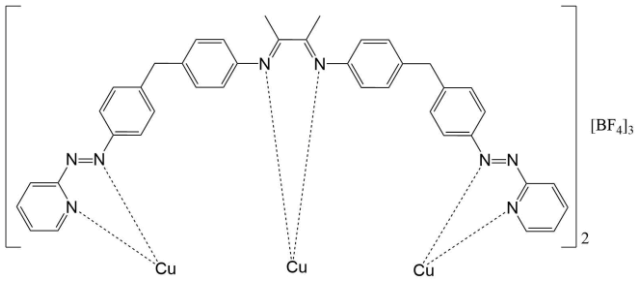
6.2.7.**Attempted synthesis of a trinuclear complex, $[Fe_3(L_{tri})_3]Cl_6$**



The half ligand (0.03 g, 0.12 mmol) was dissolved in methanol (5 mL) under an argon atmosphere. 2,3-Butanedione (0.01 mL, 0.09 mmol) was added dropwise to the solution whilst stirring. Iron(II) chloride tetrahydrate (0.02 g, 0.09 mmol) was dissolved in methanol (5 mL) and added dropwise to the ligand solution which was then heated under reflux for 24 hours. The solution was cooled to room temperature before being precipitated in excess diethyl ether and cooled overnight. A dark green precipitate was collected by vacuum filtration and washed with diethyl ether (10 mL). The precipitate was dried under *vacuo* (0.03 g, 0.01 mmol, 16%).

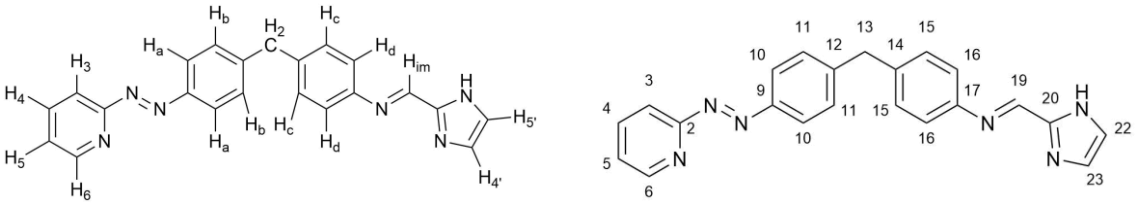
Mass spectrum (TOF MS EI+): *m/z* = 289 $[C_{18}H_{16}N_4 + H]^+$, 397 $[C_{22}H_{20}N_4O + Na]^+$, 413 $[C_{22}H_{20}N_4O + Na + MeOH]^+$, 627 $[C_{40}H_{34}N_8 + H]^+$

6.2.8.**Attempted synthesis of a copper(I) trinuclear helicate, $[Cu_3(L_{tri})_2](BF_4)_3$**



The half ligand (0.01g, 0.04 mmol) was dissolved in methanol (2 mL) under an argon atmosphere. A solution of 2,3-butanedione (0.1 mL, 1.14 mmol) in dichloromethane (9.9 mL) was made (0.11 M). The 2,3-butanedione solution (0.2 mL) was added to the half ligand solution. Tetrakisacetonitrile copper(I) tetrafluoroborate (0.01 g, 0.03 mmol) was dissolved in methanol (2 mL) and added dropwise to the ligand solution which was then heated under reflux for 3 days. A precipitate was removed by filtration and washed with methanol (2 mL). The precipitate was then dissolved in acetonitrile to remove from the filter. The solvent was removed *in vacuo* and the purple precipitate was dried overnight *in vacuo* (0.02 g, 0.01 mmol, 84%).

6.2.9. **Synthesis of an asymmetric ligand, L_{asym}**



2-Imdiazolecarboxaldehyde (0.02 g, 0.17 mmol) was suspended in dry methanol (6 mL) and added to a solution of half ligand (0.05 g, 0.17 mmol) in dry methanol (2 mL). Acetic acid (2 drops) was added after 10 minutes and the solution was heated overnight under reflux. The solution was filtered to remove any precipitate and the solvent was removed to give a green powder product (0.05g, 0.13 mmol, 80%).

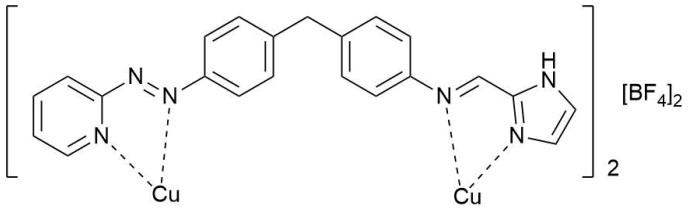
The green powder product was then dissolved in dichloromethane (10 mL). 2-Imdiazolecarboxaldehyde (0.02 g, 0.17 mmol) suspended in dry methanol (6 mL) with acetic acid (2 drops) and heated overnight under reflux. The solvent was removed under *vacuo* to yield a green powder solid (0.03 g, 0.08 mmol, 48%).

^{1}H NMR: (400 MHz, Chloroform-*d*) δ 9.78 (NH, s, 1H), 8.75 (H_6, ddd, *J* = 4.7, 1.8, 0.9 Hz, 1H), 8.47 (H_{im}, s, 1H), 8.00 (HPh_a, d, *J* = 8.5 Hz, 2H), 7.91 (H_4, ddd, *J* = 8.0, 7.3, 1.8 Hz, 1H), 7.82 (H_3, dt, *J* = 8.0, 1.0 Hz, 1H), 7.43 – 7.39 (H_5 + $H_{4'}$ + $H_{5'}$, m, 3H), 7.37 (HPh_b, d, *J* = 8.5 Hz, 2H), 7.26 (HPh_c, d, *J* = 8.5 Hz, 1H), 7.20 (HPh_d, d, *J* = 8.5 Hz, 2H), 4.09 (CH_2, s, 2H).

^{13}C NMR: (101 MHz, Chloroform-*d*) δ 207.24 (C_{19}), 181.44 (C_6), 149.74 ($C_{22/23}$), 149.25($C_{22/23}$), 138.58 (C_4), 130.19 ($C_{10/16}$), 129.92 ($C_{10/16}$), 125.36 (C_2), 124.12 ($C_{11/15}$), 121.56 ($C_{11/15}$), 115.72 (C_5), 29.90 (CH_2).

Mass spectrum (TOF MS EI+): *m/z* = 367 $[C_{22}H_{18}N_6 + H]^+$, 389 $[C_{22}H_{18}N_6 + Na]^+$, 733 $[C_{44}H_{36}N_{12} + H]^+$, 755 $[C_{44}H_{36}N_{12} + Na]^+$, 1100 $[C_{66}H_{54}N_{18} + H]^+$

6.2.10. **Attempted Synthesis of an asymmetric copper helicate, $[Cu_2(L_{asym})_2](BF_4)_2$**



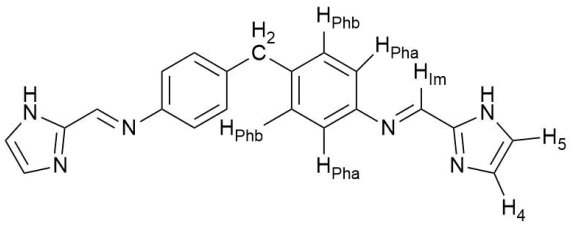
The asymmetric ligand (0.03 g, 0.08 mmol) and tetrakis(acetonitrile) copper(I) tetrafluoroborate (0.03 g, 0.08 mmol) was dissolved in dry methanol (10 mL) under an argon atmosphere and heated under reflux for 2 hours and 30 minutes. The solution was allowed to cool to room temperature and then chilled at 4°C for an hour. A black crystalline solid was collected by vacuum filtration (0.02 g, 0.02 mmol, 23.5%).

Mass spectrum (TOF MS EI+): *m/z* = 890 $[Cu_5C_{88}H_{67}N_{24}]^{2+}$

6.3.Imidazole Cylinders:

6.3.1.2-Imidazole Ligand (L_{im2}):

Based on synthesis by M. Pascu[84].



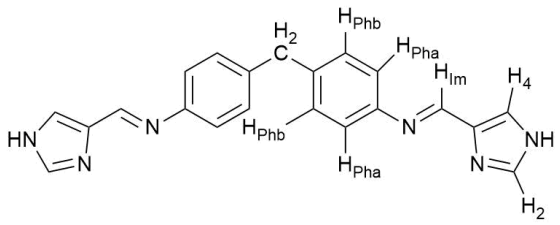
4,4'-Methylenedianiline (0.79 g, 8.00 mmol) dissolved in methanol (15 mL) was added to 2-imidazolecarboxaldehdye (0.77 g, 4.00 mmol) in methanol (15 mL) giving a cloudy white solution which was stirred for 10 minutes under a nitrogen atmosphere. Glacial acetic acid (3 drops) was added and the solution heated under reflux for 2 hours. The solution was cooled to room temperature before being vacuum filtered to collect the white precipitate which was washed with methanol (2 x 10 mL). The product was dried *in vacuo* (1.25 g, 3.53 mmol, 88%).

^{1}H NMR: (400 MHz, DMSO-d_6) δ 13.02 (NH, s, 1H), 8.40 (H_{im}, s, 1H), 7.32 ($H_{4/5}$, s, 1H), 7.30 ($HPh_{a/b}$, dt, J = 8.6, 2.1 Hz, 2H), 7.25 ($HPh_{a/b}$, dt, J = 8.6, 2.1 Hz, 2H), 7.18 ($H_{4/5}$, s, 1H), 3.98 (CH_2, s, 1H).

Mass spectrum (TOF MS EI+): m/z = 355 $[C_{21}H_{18}N_6 + H]^+$

6.3.2.**4-Imidazole Ligand (L_{im4})**

Based on synthesis methods by F. Tuna[85] and M. Pascu[84].



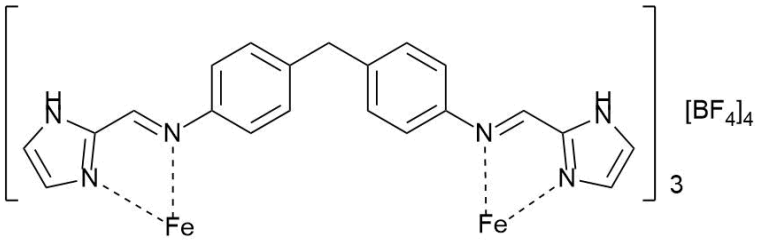
4,4'-Methylenedianiline (0.80 g, 8.33 mmol) in methanol (15 mL) was added to 4-imidazolecarboxaldehyde (0.77 g, 4.00 mmol) in methanol (15 mL) under a nitrogen atmosphere and stirred for 10 minutes. Acetic acid (2 drops) was added and the solution heated under reflux for 2 hours. The solution was cooled to room temperature before being vacuum filtered to collect the white precipitate. The white powder product was wahed with methanol (2 x 10 mL) and dried *in vacuo* (1.18 g, 3.33 mmol, 83%).

^{1}H NMR: (400 MHz, DMSO-d_6) δ 12.80 (NH, s, 1H), 8.43 (H_2, s, 1H), 7.82 (H_4, s, 1H), 7.64 (H_{im}, s, 1H), 7.26 (HPh_a, dd, J = 8.4, 2.4 Hz, 2H), 7.14 (HPh_b, dd, J = 8.4, 2.4 Hz, 2H), 3.94 (CH_2, s, 1H).

Mass spectrum (TOF MS EI+): *m/z* = 178 $[C_{21}H_{18}N_6 + 2H]^{2+}$, 277 $[C_{17}H_{16}N_4 + H]^+$, 299 $[C_{17}H_{16}N_4 + Na]^+$, 355 $[C_{21}H_{18}N_6 + H]^+$, 377 $[C_{21}H_{18}N_6 + Na]^+$, 575 $[2C_{17}H_{16}N_4 + Na]^+$, 653 $[C_{21}H_{18}N_6 + C_{17}H_{16}N_4 + Na]^+$.

6.3.3. **2-Imidazole iron complex ($[Fe_2(L_{im2})_3](BF_4)_4$)**

Based on synthesis method by M. Pascu[84].



The 2-imidazole ligand (L_{im2}, 0.11 g, 0.30 mmol) was suspended in acetonitrile (10 mL) under a nitrogen atmosphere. Iron(II) tetrafluoroborate hexahydrate (0.07 g, 0.20 mmol) was dissolved in acetonitrile (10 mL) and added to the ligand solution giving an instant colour change to light pink which gradually became dark purple. The solution was stirred for 24 hours before the solvent was removed *in vacuo*. The purple powder was redissolved in minimal acetonitrile (3 mL) and the solution was added dropwise to excess diethyl ether whilst stirring. The purple precipitate was isolated by vacuum filtration and was dried *in vacuo* (0.13 g, 0.10 mmol, 99%).

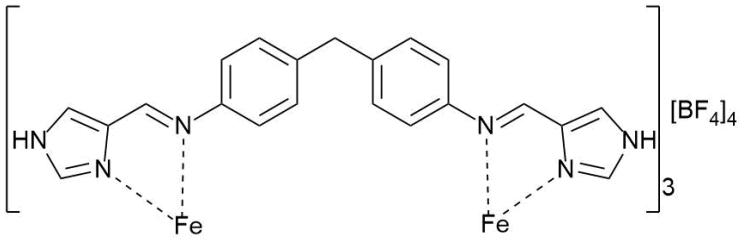
^{1}H NMR: (300 MHz, Acetonitrile-d_3) δ 124.15 (H_4, s, 1H), 71.16 (NH, s, 1H), 65.25 (H_5, s, 1H), 56.38 (H_{im}, s, 1H), 22.10 (CH_2, s, 1H), 17.79 ($HPh_{a/b}$, s, 2H), 0.63 ($HPh_{a/b}$, s, 2H).

Mass spectrum (SCAN ES+): *m/z* = 178 $[C_{21}H_{18}N_6 + H]^{2+}$, 277 $[C_{17}H_{16}N_4 + H]^+$, 294 $[Fe_2(C_{21}H_{18}N_6)_3]^{4+}$, 355 $[C_{21}H_{18}N_6 + H]^+$, 391 $[Fe_2(C_{21}H_{18}N_6)_3 - H]^{3+}$, 587 $[Fe_2(C_{21}H_{18}N_6)_3 - 2H]^{2+}$

UV/VIS (acetonitrile): 212, 242, 320, 447, 513 nm.

6.3.4. **4-Imidazole iron complex ($[Fe_2(L_{im4})_3](BF_4)_4$)**

Based on synthesis methods by F. Tuna[85] and M. Pascu[84].



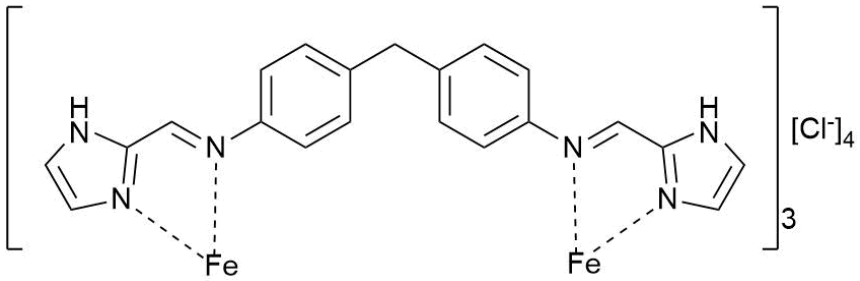
The 4-imidazole ligand (L_{im4}, 0.11 g, 0.30 mmol) was suspended in acetonitrile (10 mL) under a nitrogen atmosphere. Iron(II) tetrafluoroborate hexahydrate (0.07 g, 0.20 mmol) was dissolved in acetonitrile (10 mL) and added to the ligand solution giving a gradual colour change to an orange solution. The solution was stirred for 24 hours before the solvent was removed *in vacuo*. The powder was redissolved in minimal acetonitrile (3 mL) and the solution was added dropwise to excess diethyl ether whilst stirring. The orange precipitate was isolated by vacuum filtration, washed with diethyl ether (2 x 10 mL) and dried *in vacuo* (0.13 g, 0.10 mmol, 98.6%).

^{1}H NMR: (300 MHz, Acetonitrile-d_3) δ 159.53 (H_2, s, 1H), 92.86 (NH, s, 1H), 43.00 (H_{Im}, s, 1H), 38.15 (H_4, s, 1H), 24.86 (CH, s, 1H), 14.70 (HPh/b, s, 2H), -5.73 ($HPh_{a/b}$, s, 2H).

Mass spectrum (SCAN ES+): *m/z* = 178 $[C_{21}H_{18}N_6 + 2H]^{2+}$, 277 $[C_{17}H_{16}N_4 + H]^+$, 293 $[Fe_2(C_{21}H_{18}N_6)_3]^{4+}$, 355 $[C_{21}H_{18}N_6 + H]^+$, 391 $[Fe_2(C_{21}H_{17}N_6)_3 - H]^{3+}$, 420 $[Fe_2(C_{21}H_{18}N_6)_3(BF_4)]^{3+}$, 587 $[Fe_2(C_{21}H_{18}N_6)_3 – 2H]^{2+}$, 631 $[Fe_2(C_{21}H_{17}N_6)_3(BF_4) - H]^{2+}$, 675 $[Fe_2(C_{21}H_{18}N_6)_3(BF_4)_2]^{2+}$

UV/VIS (acetonitrile): 267, 445 nm.

6.3.5.**2-Imidazole iron chloride complex ($[Fe_2(L_{im2})_3](Cl)_4$)**



The 2-imidazole ligand (L_{im2}, 0.11 g, 0.3 mmol) was suspended in methanol (10 mL) under an inert atmosphere. Iron(II) chloride tetrahydrate (0.04 g, 0.2 mmol) was dissolved in methanol (10 mL) , added to the ligand solution and stirred for 20 hours causing a gradual colour change to dark purple. The solution was filtered through celite and washed with methanol (3 x 10 mL). The solvent was removed *in vacuo* to give a dark purple solid which was redissolved in minimal methanol and precipitated in excess diethyl ether. The purple precipitate was collected by vacuum filtration, washed with diethyl ether (10 mL) and dried *in vacuo* (0.08 g, 0.06 mmol, 60.6%).

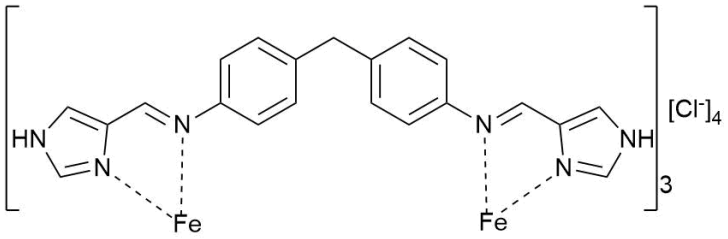
^{1}H NMR: (300 MHz, Methanol-d_4) δ 118.69 (H_4, s, 1H), 67.82 (H_{im}, s, 1H), 61.63 (H_5, s, 1H), 21.97 (CH_2, s, 1H), 17.43 ($HPh_{a/b}$, s, 2H), 0.79 ($HPh_{a/b}$, s, 2H).

Mass spectrum (SCAN ES+): *m/z* = 178 $[C_{21}H_{18}N_6 + 2H]^{2+}$, 277 $[C_{17}H_{16}N_4 + H]^+$, 309 $[C_{17}H_{16}N_4 + H + MeOH]^+$, 355 $[C_{21}H_{18}N_6 + H]^+$, 387 $[C_{21}H_{18}N_6 + H + MeOH]^+$, 631 $[(C_{21}H_{18}N_6) + (C_{17}H_{16}N_4) + H]^+$, 709 $[2(C_{21}H_{18}N_6) + H]^+$

UV/VIS(methanol): 224, 294, 329, 423 nm.

6.3.6.**4-Imidazole iron chloride complex ($[Fe_2(L_{im4})_3](Cl)_4$)**

Based on synthesis method by M. Pascu[84].



The 4-imidazole ligand (L_{im4}, 0.11 g, 0.30 mmol) was suspended in methanol (10 mL) under a nitrogen atmosphere. Iron(II) chloride tetrahydrate (0.04 g, 0.20 mmol) was dissolved in methanol (10 mL) , added to the ligand solution and stirred for 21 hours causing a gradual colour change to orange. The solution was filtered through celite and washed with methanol (3 x 10 mL). The solvent was removed *in vacuo* to give an orange solid which was redissolved in minimal methanol and precipitated in excess diethyl ether. The orange precipitate was collected by vacuum filtration, washed with diethyl ether (10 mL) and dried *in vacuo* (0.10 g, 0.08 mmol, 76%).

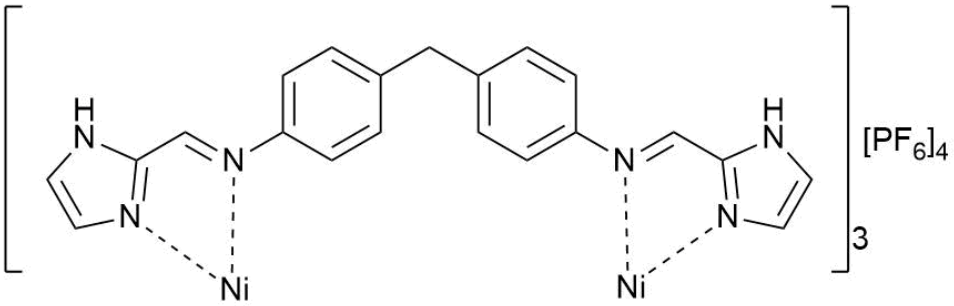
Peaks assigned "Dia-" are the diamagnetic cylinder.

^{1}H NMR: (300 MHz, Methanol-d_4) δ 157.50 (H_2, s, 1H), 41.92 (H_{im}, s, 1H), 37.20 (H_4, s, 1H), 25.38 (CH_2, s, 1H), 14.73 ($HPh_{a/b}$, s, 2H), -5.66 ($HPh_{a/b}$, s, 2H). (300 MHz, Deuterium Oxide) δ 151.79 (H_2, s, 2H), 36.41 (H_4, s, 2H), 24.07 (CH_2, s, 2H), 14.23 ($HPh_{a/b}$, s, 4H), 9.70 (Dia-H_{im}, s, 1H), 8.00 (Dia-H_2 + Dia-H_5, m, 2H), 7.04 (Dia-$HPh_{a/b}$, s, 2H), 6.76 (Dia-$HPh_{a/b}$, s, 2H), 3.71 (Dia-CH_2, s, 1H), -5.26 ($HPh_{a/b}$, s, 4H).

Mass spectrum (TOF MS EI+): *m/z* = 178 $[C_{21}H_{18}N_6 + 2H]^{2+}$, 277 $[C_{17}H_{16}N_4 + H]^+$, 293 $[Fe_2(C_{21}H_{18}N_6)_3]^{4+}$, 355 $[C_{21}H_{18}N_6 + H]^+$, 391 $[Fe_2(C_{21}H_{18}N_6)_3 - H]^{3+}$, 586 $[Fe_2(C_{21}H_{18}N_6)_3 - 2H]^{2+}$, 623 $[Fe_2(C_{21}H_{18}N_6)_3(Cl_2)]^{2+}$.

UV/VIS(methanol): 268 nm, 434 nm.

6.3.7. **2-Imidazole nickel complex ($[Ni_2(L_{im2})_3](PF_6)_4$)**



The 2-imidazole ligand (L_{im2}, 0.11 g, 0.30 mmol) was suspended in methanol (10 mL) under a nitrogen atmosphere. Nickel(II) chloride hexahydrate (0.05 g, 0.20 mmol) in methanol (10 mL) was added dropwise to the 2-imidazole ligand solution whilst stirring. After 45 minutes, excess ammonium hexafluorophosphate was added and the solution was immediately filtered through celite. The filtrate was cooled to room temperature and then chilled at 4°C for two days. An orange precipitate was collected by vacuum filtration and washed with cold methanol (10 mL) before being dried *in vacuo* (0.05 g, 0.03 mmol, 28%).

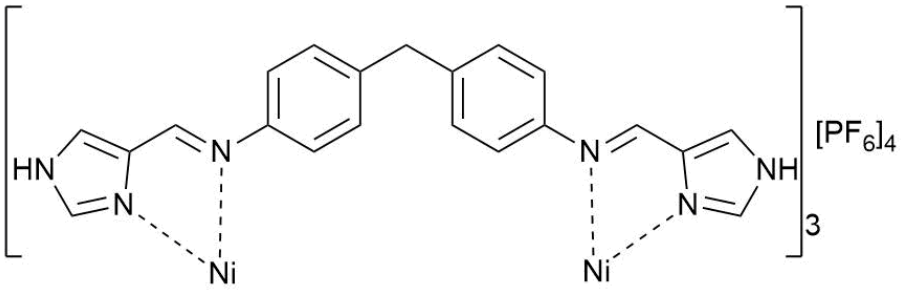
^{1}H NMR: (300 MHz, Acetonitrile-d_3) δ 74.77 (H_5, s, 1H), 66.15 (H_{im}, s, 1H), 25.16 (CH_2, s, 1H), 15.23 ($HPh_{a/b}$, s, 2H), -5.43 ($HPh_{a/b}$, s, 2H).

Mass spectrum (SCAN ES+): *m/z* = 295 $[Ni_2(C_{21}H_{18}N_6)_3]^{4+}$, 305 $[Ni_2(C_{21}H_{18}N_6)_3 + CH_3CN]^{4+}$, 316 $[Ni_2(C_{21}H_{18}N_6)_3 + 2CH_3CN]^{4+}$, 326 $[Ni_2(C_{21}H_{18}N_6)_3 + 3CH_3CN]^{4+}$, 392 $[Ni_2(C_{21}H_{17}N_6)_3 - H]^{3+}$, 442 $[Ni_2(C_{21}H_{18}N_6)_3(PF_6)]^{3+}$, 588 $[Ni_2(C_{21}H_{16}N_6)_3 - 2H]^{2+}$, 735 $[Ni_2(C_{21}H_{18}N_6)_3(PF_6)_2]^{2+}$

UV/VIS(acetonitrile): 323 nm.

6.3.8.**4-Imidazole nickel complex ($[Ni_2(L_{im4})_3](PF_6)_4$)**

Based on synthesis method by F. Tuna[85].



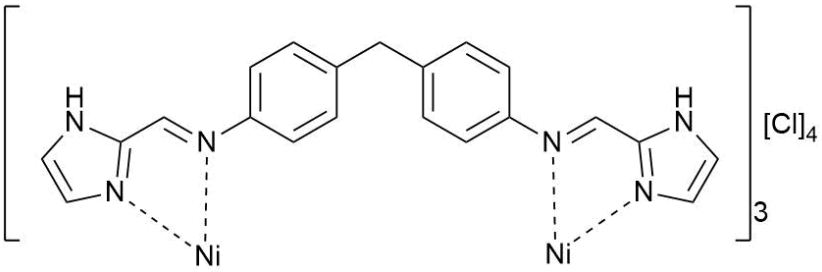
The 4-imidazole ligand (L_{im4},0.11 g, 0.30 mmol) was suspended in methanol (10 mL) under a nitrogen atmosphere. Nickel(II) chloride hexahydrate (0.05 g, 0.20 mmol) in methanol (10 mL) was added dropwise to the 2-imidazole ligand solution whilst stirring. After 45 minutes, excess ammonium hexafluorophosphate was added to the solution and the solution was immediately filtered through celite. The filtrate was allowed to cool and left to chill at 4^{o}C for two days. A pale green precipitate was collected by vacuum filtration and washed with cold methanol (10 mL) before being dried *in vacuo* (0.05 g, 0.03 mmol, 28%).

^{1}H NMR: (300 MHz, Acetonitrile-d_3) δ 94.10 (NH, s, 1H), 59.73 (H_{im}, s, 1H), 41.29 (H_4, s, 1H), 27.71 (CH_2, s, 1H), 15.46 ($HPh_{a/b}$, s, 2H), -6.71 ($HPh_{a/b}$, s, 1H).

Mass spectrum (SCAN ES+): *m/z* = 295 $[Ni_2(C_{21}H_{18}N_6)_3]^{4+}$, 305 $[Ni_2(C_{21}H_{18}N_6)_3 + CH_3CN]^{4+}$, 316 $[Ni_2(C_{21}H_{18}N_6)_3 + 2CH_3CN]^{4+}$, 393 $[Ni_2(C_{21}H_{17}N_6)_3 - H]^{3+}$, 442 $[Ni_2(C_{21}H_{18}N_6)_3(PF_6)]^{3+}$, 735 $[Ni_2(C_{21}H_{18}N_6)_3(PF_6)_2]^{2+}$.

UV/VIS(acetonitrile): 276, 409 nm.

6.3.9.**2-Imidazole nickel complex ($[Ni_2(L_{im2})_3](Cl)_4$)**

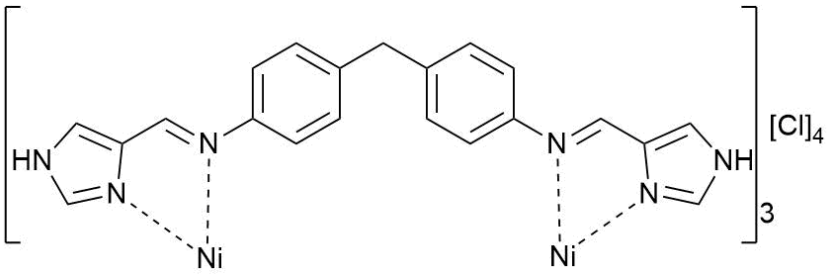


2-Imidazole ligand (L_{im2}, 0.11 g, 0.28 mmol) was suspended in methanol (10 mL) under a nitrogen atmosphere. Nickel(II) chloride hexahydrate (0.05 g, 0.20 mmol) in methanol (10 mL) was added dropwise to the 2-imidazole ligand solution whilst stirring. The orange solution was stirred for 35 minutes before being filtered through celite and washed with methanol (5 mL). Excess methanolic ammonium hexafluorophosphate was added and stirred for 30 minutes. The solution was then chilled at 4°C for two days to give an orange precipitate which was collected by vacuum filtration using a fine nylon membrane filter (0.45 μm, 25mm diameter) and washed with diethyl ether (10 mL). The precipitate suspended in methanol (20 mL) and stirred for 30 minutes with Dowex (1x8 chloride form, 200-400 mesh).The Dowex beads were removed by vacuum filtration and washed with methanol (10 mL). The solvent was removed from the filtrate under *vacuo* to yield an orange powder (0.10 g, 0.07 mmol, 74%).

^{1}H NMR: (300 MHz, Methanol-d_4) δ 74.16 (H_5, s, 1H), 65.71 (H_{im}, s, 1H), 26.11 (CH_2, s, 1H), 15.34 ($HPh_{a/b}$, s, 2H), -5.34 ($HPh_{a/b}$, s, 2H). (300 MHz, Methanol-d_4) δ 75.13 (H_5, s, 1H), 25.95 (CH_2, s, 1H), 15.35 ($HPh_{a/b}$, s, 2H), -5.37 ($HPh_{a/b}$, s, 2H). (Gaussian function: 1GB applied)

UV/VIS(methanol): 321 nm.

6.3.10.**4-Imidazole nickel complex ($[Ni_2(L_{im4})_3](Cl)_4$)**



4-Imidazole ligand (L_{im4}, 0.11 g, 0.28 mmol) was suspended in methanol (10 mL) under a nitrogen atmosphere. Nickel(II) chloride hexahydrate (0.05 g, 0.20 mmol) in methanol (10 mL) was added dropwise to the 2-imidazole ligand solution whilst stirring. The green solution was stirred for 45 minutes then gravity filtered. Excess methanolic ammonium hexafluorophosphate was added and stirred for 30minutes. The solution was chilled at 4^{o}C for two days to give a pale green precipitate which was collected by vacuum filtration using a fine nylon membrane filter (0.45 μm, 25 mm diameter) and washed with diethyl ether (10 mL). The precipitate was then suspended in methanol (20 mL) and stirred for 30 minutes with Dowex (1x8 chloride form, 200-400 mesh).The Dowex beads were removed by vacuum filtration and washed with methanol (10 mL). The solvent was removed from the filtrate under *vacuo* to give an green powder (0.03 g, 0.02 mmol, 20%).

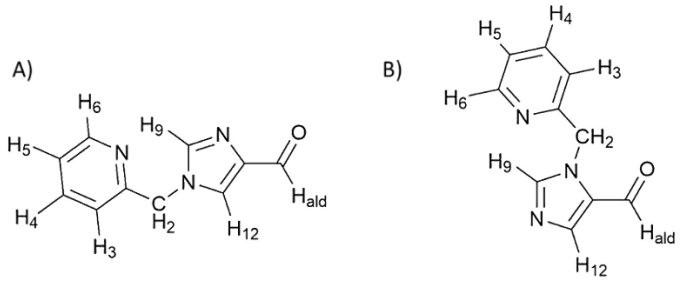
^{1}H NMR: (300 MHz, Methanol-d_4) δ 60.00 (Him, s, 1H), 41.11 (H4, s, 1H), 28.38 (CH2, s, 1H), 15.62 (HPha/b, s, 2H), -6.85 (HPha/b, s, 2H).

UV/VIS(methanol): 265 nm.

6.4.Rotaxanated Cylinders

6.4.1.Alkylating 4-imidazole carboxaldehyde

Based on method developed by V. Sadovnikova[116].



2-(Bromomethyl)pyridine hydrobromide (0.79 g, 3.12 mmol) and 4-imidazolecarboxaldehyde (0.27 g, 2.86 mmol) were dissolved in dimethylformamide (10 mL) under an argon atmosphere. Hunig's base (1.5 mL, 8.70 mmol) was added slowly and the reaction solution was heated to 80^{o}C overnight. The solution was cooled to room temperature before a saturated solution of sodium hydrogen carbonate (10 mL) was added. The solution was extracted with dichloromethane (4 x 10 mL) and the combined organic layers were washed with brine (20 mL). The organic layer was then dried over magnesium sulphate, filtered and the solvent removed from the filtrate under *vacuo* (0.20g, crude). Prep-RP HPLC was used to purify.

Peak 3:

^{1}H NMR: (400 MHz, Chloroform-*d*) δ 9.88 (H_{ald}, s, 1H), 8.61 (H_6, ddd, *J* = 4.9, 1.8, 1.0 Hz, 1H), 7.78 – 7.65 (H_4 + H_9 + H_{12}, m, 3H), 7.29 (H_5, ddd, *J* = 8.1, 4.9, 1.0 Hz, 1H), 7.11 (H_3, dt, *J* = 7.8, 1.0 Hz, 1H), 5.27 (CH_2, s, 2H).

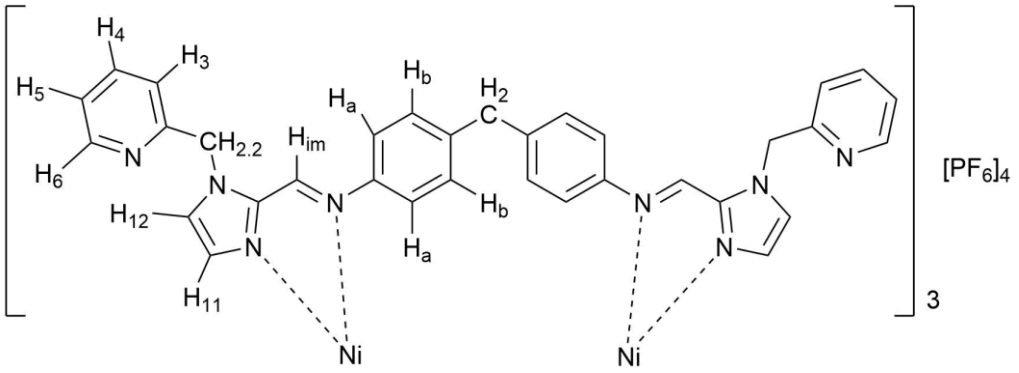
Mass spectrum (ES+): *m/z* = 130 $[C_8H_{10}N]^+$, 188 $[C_{10}H_9N_3O + H]^+$, 210 $[C_{10}H_9N_3O + Na]^+$, 397 $[2C_{10}H_9N_3O + Na]^+$

Peak 4:

^{1}H NMR: ^{1}H NMR (400 MHz, Chloroform-*d*) δ 9.74 (H_{ald}, s, 1H), 8.54 (H_6, ddd, *J* = 4.8, 1.8, 1.0 Hz, 1H), 7.93 ($H_{9/12}$, s, 1H), 7.87 ($H_{9/12}$, s, 1H), 7.65 (H_4, td, *J* = 7.7, 1.8 Hz, 1H), 7.25 – 7.18 ($H_{3/5}$, m, 2H), 5.61 (CH_2, s, 2H).

Mass spectrum (ES+): *m/z* = 130 $[C_8H_{10}N]^+$, 188 $[C_{10}H_9N_3O + H]^+$, 210 $[C_{10}H_9N_3O + Na]^+$, 220 $[C_{10}H_9N_3O + MeOH + H]^+$, 358 $[C_{10}H_9N_3O + MeCN + H]^+$

6.4.2. **Post assembly modification of the nickel 2-imidazole cylinder, $[Ni_2(L_{mod})_3](PF_6)_4$**



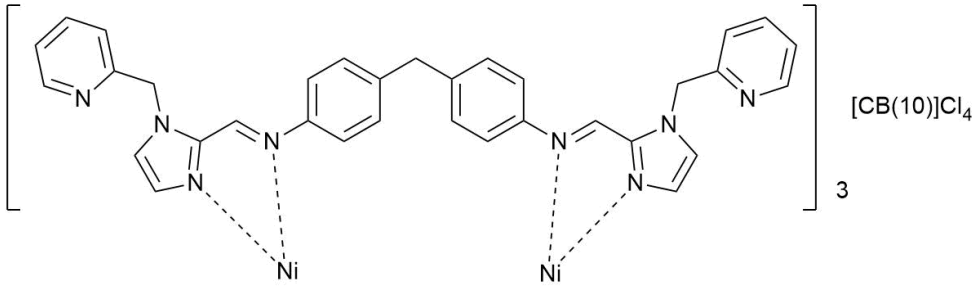
$[Ni_2(L_{im2})_3](PF_6)_4$ (0.10 g, 0.06 mmol) and 2-(bromomethyl)pyridine hydrobromide (0.09 g, 0.34 mmol) were dissolved in acetonitrile (20 mL) under an argon atmosphere. After 10 minutes, Hunig's base (0.18 mL, 1.02 mmol) was added and the solution was heated overnight under reflux. The solution was cooled to room temperature before being vacuum filtered and washed with acetonitrile (2 mL). The solvent was removed from the filtrate to give an orange, waxy crude product. The crude product was

washed with chloroform (10 mL), the precipitate collected by vacuum filtration and washed with chloroform (2 x 5 mL) to yield an orange powder (0.05 g, 0.02 mmol, 35%).

^{1}H NMR: 300 MHz, Acetonitrile-d_3) δ 78.77 (H_{12}, s, 1H), 71.92 (H_{im}, s, 1H), 24.42 (CH_2, s, 1H), 15.23 ($H_{Pha/b}$, s, 2H), 8.54 (H_6, s, 1H), 7.79 ($H_{3/4/5}$, s, 1H), 7.51 ($H_{3/4/5}$, s, 1H), 7.20 ($H_{3/4/5}$, s, 1H), 4.20 ($CH_{2.2}$, s, 1H), 1.43 ($CH_{2.2}$, s, 1H), -4.58 ($H_{Pha/b}$, s, 2H).

Mass spectrum (TOF MS EI+): *m/z* = 130 $[C_8H_{20}N]^+$, 409 $[Ni_2(C_{33}H_{28}N_8)_2(C_{27}H_{23}N_7)]^{4+}$, 432 $[Ni_2(C_{33}H_{28}N_8)_3]^{4+}$, 545 $[Ni_2(C_{33}H_{28}N_8)_2(C_{27}H_{23}N_7) - H]^{3+}$, 623 $[Ni_2(C_{33}H_{28}N_8)_3(PF_6)_4]^{3+}$

6.4.3.**Synthesis of a 2-imidazole nickel rotaxane by alkylation**



$[Ni_2(L_{im2})_3](Cl_4)$ (0.04 g, 0.03 mmol) was dissolved in deionised water (10 mL) and methanol (0.3 mL) and added to CB10 (0.05 g, 0.03 mmol) under an argon atmosphere. The solution was stirred for 20 minutes before 2-(bromomethyl)pyridine hydrobromide (0.05 g, 0.17 mmol) in deionised water (5 mL) was added. Hunig's base (0.09 mL, 0.51 mmol) was added to the solution after 10 minutes which was then heated to 80^{o}C and stirred overnight. The reaction was cooled to room temperature and quenched with saturated aqueous sodium hydrogen carbonate solution (15 mL). The solution was allowed to settle overnight before being filtered to remove a brown precipitate. The

solvent was removed *in vacuo* and the resultant solid was suspended in methanol (25 mL). The precipitated salt was removed by vacuum filtration and the filtrate collected. The solvent was removed *in vacuo* and to yield a fine, cream precipitate (0.04 g, 0.11 mmol, 40%).

Mass spectrum (TOF MS ES+): *m/z* = 801 $[Ni_2(C_{33}H_{28}N_8)(C_{27}H_{23}N_7)_2(CB10)]^{4+}$, 824 $[Ni_2(C_{33}H_{28}N_8)_2(C_{27}H_{23}N_7)(CB10)]^{4+}$, 828 $[Ni_2(C_{33}H_{28}N_8)_2(C_{27}H_{23}N_7)(CB10) + H_2O]^{4+}$, 847 $[Ni_2(C_{33}H_{28}N_8)_3(CB10)]^{4+}$, 851 $[Ni_2(C_{33}H_{28}N_8)_3(CB10) + H_2O]^{4+}$, 855 $[Ni_2(C_{33}H_{28}N_8)_3(CB10) + 2H_2O]^{4+}$, 1072 $[Ni_2(C_{33}H_{28}N_8)(C_{27}H_{23}N_7)_2(CB10) + OH]^{3+}$, 1098 $[Ni_2(C_{33}H_{28}N_8)_2(C_{27}H_{23}N_7)(CB10) - H]^{3+}$, 1104 $[Ni_2(C_{33}H_{28}N_8)_2(C_{27}H_{23}N_7)(CB10) + OH]^{3+}$, 1134 $[Ni_2(C_{33}H_{28}N_8)_3(CB10) + OH]^{3+}$

6.5. Threading Studies

Iron 4-imidazole cylinder, $[Fe_2(L_{im4})_3]Cl_4$, with CB10

$[Fe_2(L_{im4})_3]Cl_4$ (0.99 mg, 0.75 μmol) was dissolved in deuterated water (0.75 mL). 0.9 equivalents of CB10 (1.12 mg, 0.68 μmol) were added and the solution stirred until the CB10 had dissolved.

^{1}H NMR: (300 MHz, Deuterium Oxide) δ 154.92 (H2, s, 2H), 44.97 (Him, s, 2H), 36.10 (H4, s, 2H), 22.29 (Ch2, s, 2H), 14.76 (Pha/b, s, 4H), 9.69 (Dia-Him, s, 1H), 7.98 (Dia-H4 and H5, d, *J* = 33.6 Hz, 2H), 7.09 (Dia-Pha/b, d, *J* = 7.3 Hz, 2H), 6.79 (Dia-Pha/b, d, *J* = 7.3 Hz, 2H), 5.43 (CB10-Hc, d, *J* = 15.6 Hz, 6H), 5.28 (CB10-Ha, s, 6H), 3.95 (CB10-Hb, d, *J* = 15.6 Hz, 6H), 3.77 (Dia-CH2, s, 1H), -4.77 (Pha/b, s, 4H).

Nickel 2-imidazole cylinder, $[Ni_2(L_{im2})_3]Cl_4$, with CB10

$[Ni_2(L_{im2})_3]Cl_4$ (1.24 mg, 0.94 μmol) was dissolved in deuterated water (0.75 mL) and deuterated methanol (0.19 mL). 0.9 equivalents of CB10 (1.40 mg, 0.84 μmol) were added and the solution stirred until no more CB10 would dissolve. Excess CB10 was removed by filtration.

Both the CB10 bound cylinder and the free cylinder are observed in the NMR. Protons have been assigned CB- and cyl- for each species respectively.

^{1}H NMR: (300 MHz, Deuterium Oxide) δ 74.81 (CB-H_4 and cyl-H_4, s, 12H), 25.93 (cyl-CH_2, s, 6H), 22.81 (CB-CH_2, s, 6H), 15.33 (cyl-$HPh_{a/b}$, s, 12H), 14.75 (CB-$HPh_{a/b}$ s, 12H), 5.71 ($H_{CB10\text{-}c}$, s, 20H), 5.41 ($H_{CB10\text{-}a}$, s, 20H), 4.12 ($H_{CB10\text{-}b}$, s, 20H), -4.88 (CB-$HPh_{a/b}$ and cyl-$HPh_{a/b}$, s, 24H).

Mass spectrum (SCAN ES+): *m/z* = 355 $[C_{21}H_{18}N_6 + H]^+$, 410 $[Ni_2(C_{21}H_{18}N_6)Cl + H_2O]^{3+}$, 590 $[Ni_2(C_{21}H_{18}N_6)_3 - 2H]^{2+}$, 711 $[Ni_2(C_{21}H_{18}N_6)_3 + CB10]^{4+}$, 947 $[Ni_2(C_{21}H_{18}N_6)_3 + CB10 - H]^{3+}$, 1420 $[Ni_2(C_{21}H_{18}N_6)_3 + CB10 - 2H]^{2+}$, 1426 $[Ni_2(C_{21}H_{18}N_6)_3 + CB10 + H_2O - 2H]^{2+}$

6.6. Spectrometry

Nuclease free water purchased from Omega Bio-tek was used for all DNA binding experiments and was stored at 4°C.

A 5 times concentrated buffer stock solution was made from sodium cacodylate trihydrate (5 mM, Fluka Biochemika, >98.0%) and sodium chloride (100 mM, Fischer) in water (pH 7.44), and was stored at 4°C.

Calf thymus DNA sodium salt (ctDNA, 27.70 mgs, purchased from Sigma Aldrich) was dissolved in water (20 mL). All DNA concentrations are quoted in terms of base pairs.

All cylinder and rotaxane stock solutions were made fresh and used immediately.

UV Visible Spectrometry and Circular Dichroism

A Varian Carey 5000 UV Vis and NIR spectrometer was used for all UV Vis measurements. The absorbance of each sample was measured between 200-800 nm (average time: 0.5s, data interval: 1.0 nm) in a 3 mL cuvette with 1 cm path length.

6.6.1.1. Cylinder absorbance, stability and molar extinction coefficients

For absorbance measurements the imidazole cylinders were dissolved in acetonitrile or methanol depending on which solvent gave the best solubility. The absorbance of $[Ni_2(L_{im2})_3]Cl_4$ was also measured in a 4:1 water to methanol solution. The concentration of each sample measured is given in Table 4.

Cylinder	Solvent	Sample Concentration / uM
$[Fe_2(L_{im2})_3](BF_4)_4$	Acetonitrile	8.69
$[Fe_2(L_{im4})_3](BF_4)_4$	Acetonitrile	8.55
$[Ni_2(L_{im2})_3](BF_4)_4$	Acetonitrile	8.55
$[Ni_2(L_{im4})_3](BF_4)_4$	Acetonitrile	8.59
$[Fe_2(L_{im2})_3](Cl)_4$	Methanol	8.61
$[Fe_2(L_{im4})_3](Cl)_4$	Methanol	8.60
$[Ni_2(L_{im2})_3](Cl)_4$	Methanol	8.32
$[Ni_2(L_{im4})_3](Cl)_4$	Methanol	8.32

Table 4: Shows the solvent and concentration each measured imidazole cylinder sample.

The stability of the iron 2-imidazole cylinder, $[Fe_2(L_{im2})_3]Cl_4$, was measured in methanol (13.16 μM) by recording the absorbance spectrum at regular intervals (6

minutes). The same procedure was used to measure the stability of the nickel 2-imidazole cylinder, $[Ni_2(L_{im2})_3]Cl_4$, which was dissolved in a 4:1 water to methanol stock solution and then diluted for testing (13.30 μM). The molar extinction coefficient of these cylinders was estimated in the same solvent system by measuring the absorbance of cylinder solutions with different known concentrations.

6.6.1.2.DNA Stock Concentration

The DNA stock solution was used to make a series of diluted solutions, each with a known amount of the stock solution. The absorbance of each solution was calculated and concentration of the DNA stock solution (2.85 mM) was calculated using the Beer Lambert Law and the molar extinction coefficient (ε_{260} = 13200 $M^{-1}cm^{-1}$) per base pair of ctDNA.

6.6.1.3.DNA Binding Studies (UV Vis and CD)

The DNA binding ability of both the nickel 2-imidazole cylinder, $[Ni_2(L_{im2})_3]Cl_4$, and for the rotaxane, $[Ni_2(L_{mod})_3][CB10]Cl_4$ were investigated

All UV Vis and CD spectra were collected by an Applied Photophysics Chirascan Plus between 200-800 nm (bandwidth: 1 nm, step size: 1 nm, time per point: 0.5s) and three repeats were taken for each scan. A 3 mL cuvette with a 1 cm pathlength was used.

A stock solution of each complex was made (1 mM). The nickel 2-imdiazole cylinder was dissolved in a 20% (by volume) methanol solution, while the rotaxane was dissolved in water. The titration was repeated with a 20% methanol solution as a control experiment.

Each UV Vis and CD titration required two DNA solutions. A two times DNA solution (2x DNA, 200 μM ctDNA, 2 mM sodium cacodylate trihydrate, sodium chloride 20 mM) was made. 1 mL of the 2x DNA stock was then diluted to half concentration to make the starting DNA solution (1x DNA, 100 μM ctDNA, 1 mM sodium cacodylate trihydrate, sodium chloride 10 mM). The 1x DNA solution was used to measure the initial DNA spectrum and each addition of cylinder or rotaxane stock solution was matched in volume by an addition of the 2x DNA stock solution to maintain the DNA and buffer concentration (1x DNA, 100 μM ctDNA, 1 mM sodium cacodylate trihydrate, sodium chloride 10 mM). A background of buffer solution (1 mM sodium cacodylate trihydrate, sodium chloride 10 mM) was taken before each titration.

The concentration of cylinder or rotaxane was increased to a 4:1 base pair to complex ratio for each titration and the absorption and CD spectra was recorded after each addition. An equal volume of the 20% methanol solution was added in a separate control titration. The addition volumes and concentrations are recorded in Table 5 and Table 6.

BP:Cyl	1xDNA Volume / uL	Total volume added / uL	Methanol / %	Total Volume / uL
-	1200	0	-	1200
60.2	1200	2	0.03	1204
40.2	1200	3	0.05	1210
20.2	1200	6	0.10	1222
10.2	1200	12	0.19	1246
8.2	1200	15	0.24	1276
6.2	1200	20	0.30	1316
4.2	1200	30	0.44	1376

Table 5: The volume of methanol added and the percentage by volume of methanol in each sample.

BP:Cyl	1xDNA Volume / uL	Total volume added / uL	Complex conc. / uM	Total Volume / uL
-	1200	0	-	1200
60.2	1200	2	1.66	1204
40.2	1200	3	2.48	1210
20.2	1200	6	4.91	1222
10.2	1200	12	9.63	1246
8.2	1200	15	11.76	1276
6.2	1200	20	15.20	1316
4.2	1200	30	21.80	1376

Table 6: The volume of complex added and the increase in complex concentration in each sample.

6.6.2. **Linear Dichroism**

All LD spectra were collected by an Applied Photophysics Chirascan Plus with LD accessory between 200-800 nm (bandwidth: 1 nm, step size: 1 nm, time per point: 0.5s) and three repeats were taken for each scan. A 250 μM cuette was used (speed: 40rpm, acceleration: 2).

Each LD titration required a DNA stock solution (400 μM ctDNA, 1 mM sodium cacodylate trihydrate, sodium chloride 10 mM) and a complex solution (500 μM cylinder or rotaxane, 1 mM sodium cacodylate trihydrate, sodium chloride 10 mM). After a background absorption spectrum was recorded using buffer solution (1 mM sodium cacodylate trihydrate, sodium chloride 10 mM), half of the solution (100 μL) was removed and the DNA stock solution added (100 μL) to record the initial DNA spectrum. Each addition of complex stock, was matched in volume by the addition of the DNA stock solution to maintain the DNA and buffer concentrations(200 μM ctDNA, 1 mM sodium cacodylate trihydrate, sodium chloride 10 mM). The complex concentration was increased to a 4:1 base pair to complex ratio (Table 8).

A control titration was completed with a 20% methanol stock solution (Table 7).

BP:Cyl	Total volume added / uL	Methanol percentage / %	Total Volume / uL
-	0.0	-	200.0
62.34	1.3	0.13	202.6
40.80	2.0	0.30	204.0
20.31	4.1	0.39	208.2
9.69	9.0	0.83	218.0
8.07	11.0	0.99	222.0
6.13	15.0	1.30	230.0
4.00	25.0	2.00	250.0

Table 7: Control titration showing the volume of 20% methanol stock solution added and the percentage of methanol by volume in each sample.

BP:Cyl	Total volume added / uL	Complex conc. / uM	Total Volume / uL
-	0	-	200.0
62.34	1.3	3.21	202.6
40.80	2.0	4.90	204.0
20.31	4.1	9.85	208.2
9.69	9.0	20.64	218.0
8.07	11.0	24.77	222.0
6.13	15.0	32.61	230.0
4.00	25.0	50.00	250.0

Table 8: Volume of complex stock solution added to each sample in order to reach the desired base pair to cylinder ratio, and the concentration of complex in each sample.

Chapter 7: REFERENCES

1 H. A. Day, P. Pavlou and Z. A. E. Waller, *Bioorg. Med. Chem.*, 2014, **22**, 4407–4418.

2 D. Rhodes and H. J. Lipps, *Nucleic Acids Res.*, 2015, **43**, 8627–8637.

3 K. Gehring, J.-L. Leroy and M. Guéron, *Nature*, 1993, **363**, 561–565.

4 M. L. Bochman, K. Paeschke and V. A. Zakian, *Nat. Rev. Genet.*, 2012, **13**, 770–780.

5 B. Shu, J. Cao, G. Kuang, J. Qiu, M. Zhang, Y. Zhang, M. Wang, X. Li, S. Kang, T.-M. Ou, J.-H. Tan, Z.-S. Huang and D. Li, *Chem. Commun.*, 2018, **54**, 2036–2039.

6 O. Zaytseva and L. M. Quinn, *BioEssays*, 2018, **40**, 1700235.

7 J. D. Watson and F. H. C. Crick, *Nature*, 1953, **171**, 964–967.

8 J. D. Watson and F. H. C. Crick, *Nature*, 1953, **171**, 737–738.

9 G. M. Blackburn, *Nucleic acids in chemistry and biology*, RSC Pub, 2006.

10 K. Hoogsteen, *Acta Crystallogr.*, 1959, **12**, 822–823.

11 A. Travers and G. Muskhelishvili, *FEBS J.*, 2015, **282**, 2279–2295.

12 A. Mujeeb, S. M. Kerwin, G. L. Kenyon and T. L. James, *Biochemistry*, 1993, **32**, 13419–13431.

13 H. R. Drew, R. M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura and R. E. Dickerson, *Proc. Natl. Acad. Sci. U. S. A.*, 1981, **78**, 2179–2183.

14 A. H. Wang, G. J. Quigley, F. J. Kolpak, J. L. Crawford, J. H. van Boom, G. van der Marel and A. Rich, *Nature*, 1979, **282**, 680–6.

15 J. Carlos, R. Galindo-Murillo, F. Cortés-Guzmán and L. Ruiz-Azuara, *J. Mex. Chem. Soc.*, 2013, **57**, 245–259.

16 A. C. Komor and J. K. Barton, *Chem. Commun.*, 2013, **49**, 3617.

17 M. J. Hannon, *Chem. Soc. Rev.*, 2007, **36**, 280–295.

18 R. Baliga and D. M. Crothers, *Proc. Natl. Acad. Sci.*, 2000, **97**, 7814–7818.

19 S. Komeda, T. Moulaei, K. K. Woods, M. Chikuma, N. P. Farrell and L. D. Williams, *J. Am. Chem. Soc.*, 2006, **128**, 16092–16103.

20 M. J. Hannon, *Pure Appl. Chem.*, 2007, **79**, 2243–2261.

21 L. S. Lerman, *J. Mol. Biol.*, 1961, **3**, 18–30.

22 K. M. Boyle and J. K. Barton, *Inorganica Chim. Acta*, 2016, 452, 3–11.

23 A. M. J. Fichtinger-Schepman, J. L. Van der Veer, J. H. J. Den Hartog, P. H. M.

Lohman and J. Reedijk, *Biochemistry*, 1985, **24**, 707–713.

24 E. Boschi, S. Davis, S. Taylor, A. Butterworth, L. A. Chirayath, V. Purohit, L. K. Siegel, J. Buenaventura, A. H. Sheriff, R. Jin, R. Sheardy, L. A. Yatsunyk and M. Azam, *J. Phys. Chem. B*, 2016, **120**, 12807–12819.

25 S. N. Georgiades, N. H. Abd Karim, K. Suntharalingam and R. Vilar, *Angew. Chemie Int. Ed.*, 2010, **49**, 4020–4034.

26 E. Izbicka, R. T. Wheelhouse, E. Raymond, K. K. Davidson, R. A. Lawrence, D. Sun, B. E. Windle, L. H. Hurley and D. D. Von Hoff, *Cancer Res.*, 1999, **59**, 639–644.

27 D.-F. Shi, R. T. Wheelhouse, D. Sun and L. H. Hurley, *J. Med. Chem.*, 2001, **44**, 4509–4523.

28 R. T. Wheelhouse, D. Sun, H. Han, F. X. Han and L. H. Hurley, *J. Am. Chem. Soc.*, 1998, **120**, 3261–3262.

29 J. Pan and S. Zhang, *J. Biol. Inorg. Chem.*, 2009, **14**, 401–407.

30 S. Vuong, L. Stefan, P. Lejault, Y. Rousselin, F. Denat and D. Monchaud, *Biochimie*, 2012, **94**, 442–450.

31 R. Rodriguez, S. Müller, J. A. Yeoman, C. Trentesaux, J.-F. Riou and S. Balasubramanian, *J. Am. Chem. Soc.*, 2008, **130**, 15758–15759.

32 S. Müller, S. Kumari, R. Rodriguez and S. Balasubramanian, *Nat. Chem.*, 2010, **2**, 1095–1098.

33 S. Müller, D. A. Sanders, M. Di Antonio, S. Matsis, J.-F. Riou, R. Rodriguez and S. Balasubramanian, *Org. Biomol. Chem.*, 2012, **10**, 6537.

34 R. Rodriguez, K. M. Miller, J. V Forment, C. R. Bradshaw, M. Nikan, S. Britton, T. Oelschlaegel, B. Xhemalce, S. Balasubramanian and S. P. Jackson, *Nat. Chem. Biol.*, 2012, **8**, 301–310.

35 S. White, J. W. Szewczyk, J. M. Turner, E. E. Baird and P. B. Dervan, *Nature*, 1998, **391**, 468–71.

36 W. S. Wade, M. Mrksich and P. B. Dervan, *J. Am. Chem. Soc.*, 1992, **114**, 8783–8794.

37 S. Mandal, Y. Kawamoto, Z. Yue, K. Hashiya, Y. Cui, T. Bando, S. Pandey, M. E. Hoque, M. A. Hossain, H. Sugiyama and H. Mao, *Nucleic Acids Res.*, 2019, **47**, 3295–3305.

38 G. Marsico, V. S. Chambers, A. B. Sahakyan, P. McCauley, J. M. Boutell, M. Di Antonio and S. Balasubramanian, *Nucleic Acids Res.*, 2019, **47**, 3862–3874.

39 A. Łęczkowska, J. Gonzalez-Garcia, C. Perez-Arnaiz, B. Garcia, A. J. P. White and R. Vilar, *Chem. - A Eur. J.*, 2018, **24**, 11785–11794.

40 J. E. Reed, A. A. Arnal, S. Neidle and R. Vilar, *J. Am. Chem. Soc.*, 2006, **128**, 5992–5993.

41 A. Arola-Arnal, J. Benet-Buchholz, S. Neidle and R. Vilar, *Inorg. Chem.*, 2008, **47**, 11910–11919.

42 M. Zeraati, D. B. Langley, P. Schofield, A. L. Moye, R. Rouet, W. E. Hughes, T. M. Bryan, M. E. Dinger and D. Christ, *Nat. Chem.*, 2018, **10**, 631–637.

43 L. Chen, L. Cai, X. Zhang and A. Rich, *Biochemistry*, 1994, **33**, 13540–13546.

44 S. Sedghi Masoud, Y. Yamaoki, Y. Ma, A. Marchand, F. R. Winnerdy, V. Gabelica, A. T. Phan, M. Katahira and K. Nagasawa, *ChemBioChem*, 2018, **19**, 2268–2272.

45 M. Debnath, S. Ghosh, A. Chauhan, R. Paul, K. Bhattacharyya and J. Dash, *Chem. Sci.*, 2017, **8**, 7448–7456.

46 X. Li, Y. Peng, J. Ren and X. Qu, *Proc. Natl. Acad. Sci.*, 2006, **103**, 19658–19663.

47 D. M. Lilley, *Q. Rev. Biophys.*, 2000, **33**, 109–59.

48 Y. Liu and S. C. West, *Nat. Rev. Mol. Cell Biol.*, 2004, **5**, 937–944.

49 J. B. Welch, D. R. Duckett and D. M. J. Lilley, *Nucleic Acids Res.*, 1993, **21**, 4548–4555.

50 F. Stühmeier, J. B. Welch, A. I. H. Murchie, D. M. J. Lilley and R. M. Clegg, *Biochemistry*, 1997, **36**, 13530–13538.

51 B. Wu, *Nucleic Acids Res.*, 2004, **32**, 3228–3239.

52 D. R. Duckett and D. M. Lilley, *EMBO J.*, 1990, **9**, 1659–1664.

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

54 N. B. Leontis, W. Kwok and J. S. Newman, *Nucleic Acids Res.*, 1991, **19**, 759–766.

55 J. L. Kadrmas, A. J. Ravin and N. B. Leontis, *Nucleic Acids Res.*, 1995, **23**, 2212–2222.

56 M. J. Hannon, C. L. Painting, A. Jackson, J. Hamblin and W. Errington, *Chem. Commun.*, 1997, 1807–1808.

57 M. J. Hannon, *Chem. Soc. Rev.*, 2007, **36**, 280–295.

58 L. Cardo and M. J. Hannon, in *Metallo-Drugs: Development and Action of Anticancer Agents*, eds. A. Sigel, H. Sigel, E. Freisinger and R. K. O. Sigel, De Gruyter, Berlin, Boston, 2018, vol. 18, pp. 303–324.

59 I. Meistermann, V. Moreno, M. J. Prieto, E. Moldrheim, E. Sletten, S. Khalid, P. M. Rodger, J. C. Peberdy, C. J. Isaac, A. Rodger and M. J. Hannon, *Proc. Natl. Acad. Sci.*, 2002, **99**, 5069–5074.

60 D. R. Boer, J. M. C. A. Kerckhoffs, Y. Parajo, M. Pascu, I. Usón, P. Lincoln, M. J. Hannon and M. Coll, *Angew. Chem. Int. Ed. Engl.*, 2010, **49**, 2336–9.

61 M. J. Hannon, V. Moreno, M. J. Prieto, E. Moldrheim, E. Sletten, I. Meistermann, C. J. Isaac, K. J. Sanders and A. Rodger, *Angew. Chemie Int. Ed.*, 2001, **40**, 879–884.

62 A. C. G. Hotze, N. J. Hodges, R. E. Hayden, C. Sanchez-Cano, C. Paines, N. Male, M.-K. Tse, C. M. Bunce, J. K. Chipman and M. J. Hannon, *Chem. Biol.*, 2008, **15**, 1258–1267.

63 J. Malina, M. J. Hannon and V. Brabec, *Chem. - A Eur. J.*, 2007, **13**, 3871–3877.

64 A. C. G. Hotze, B. M. Kariuki and M. J. Hannon, *Angew. Chemie Int. Ed.*, 2006, **45**, 4839–4842.

65 S. Perticaroli, J. D. Nickels, G. Ehlers, H. O'Neill, Q. Zhang and A. P. Sokolov, *Soft Matter*, 2013, **9**, 9548–9556.

66 C. Uerpmann, J. Malina, M. Pascu, G. J. Clarkson, V. Moreno, A. Rodger, A. Grandas and M. J. Hannon, *Chem. - A Eur. J.*, 2005, **11**, 1750–1756.

67 J. Novotna, A. Laguerre, A. Granzhan, M. Pirrotta, M.-P. Teulade-Fichou and D. Monchaud, *Org. Biomol. Chem.*, 2015, **13**, 215–222.

68 Z. Yang, Y. Chen, G. Li, Z. Tian, L. Zhao, X. Wu, Q. Ma, M. Liu and P. Yang, *Chem. - A Eur. J.*, 2018, **24**, 6087–6093.

69 C. Piguet, G. Bernardinelli and G. Hopfgartner, *Chem. Rev.*, 1997, **97**, 2005–2062.

70 A. P. Paneerselvam, S. S. Mishra and D. K. Chand, *J. Chem. Sci.*, 2018, **130**, 96.

71 J.-M. Lehn, Perspectives in Chemistry: From Supramolecular Chemistry towards Adaptive Chemistry, https://www.youtube.com/watch?v=08RBLIQ8VPE, (accessed 9 September 2019).

72 J. A. McCleverty, *Chemistry of the first-row transition metals*, Oxford University Press, 1999.

73 F. E. Hahn, M. Offermann, C. Schulze Isfort, T. Pape and R. Fröhlich, *Angew. Chemie Int. Ed.*, 2008, **47**, 6794–6797.

74 F. Habib, J. Long, P.-H. Lin, I. Korobkov, L. Ungur, W. Wernsdorfer, L. F. Chibotaru and M. Murugesu, *Chem. Sci.*, 2012, **3**, 2158.

75 B. Hasenknopf, J.-M. Lehn, N. Boumediene, A. Dupont-Gervais, A. Van Dorsselaer, B. Kneisel and D. Fenske, *J. Am. Chem. Soc.*, 1997, **119**, 10956–10962.

76 L. Cardo, V. Sadovnikova, S. Phongtongpasuk, N. J. Hodges and M. J. Hannon, *Chem. Commun.*, 2011, **47**, 6575.

77 Y. Parajó, J. Malina, I. Meistermann, G. J. Clarkson, M. Pascu, A. Rodger, M. J. Hannon and P. Lincoln, *Dalt. Trans.*, 2009, 4868.

78 J. C. Peberdy, J. Malina, S. Khalid, M. J. Hannon and A. Rodger, *J. Inorg. Biochem.*, 2007, **101**, 1937–1945.

79 M. J. Hannon, S. Bunce, A. J. Clarke and N. W. Alcock, *Angew. Chemie Int. Ed.*, 1999, **38**, 1277–1278.

80 L. J. Childs, PhD Thesis, University of Warwick, 2002.

81 J. C. White, PhD Thesis, University of Birmingham, 2015.

82 S. Vitorino, PhD Thesis, University of Birmingham, 2008.

83 Y. Liu and O. M. Yaghi, *Bull. Japan Soc. Coord. Chem.*, 2018, **71**, 12–17.

84 M. Pascu, PhD Thesis, University of Birmingham, 2007.

85 F. Tuna, M. R. Lees, G. J. Clarkson and M. J. Hannon, *Chem. - A Eur. J.*, 2004, **10**, 5737–5750.

86 M. Loos, C. Gerber, F. Corona, J. Hollender and H. Singer, *Anal. Chem.*, 2015, **87**, 5738–5744.

87 P. J. Cail, PhD Thesis, University of Birmingham, 2017.

88 Q. Li, S.-C. Qiu, J. Zhang, K. Chen, Y. Huang, X. Xiao, Y. Zhang, F. Li, Y.-Q. Zhang, S.-F. Xue, Q.-J. Zhu, Z. Tao, L. F. Lindoy and G. Wei, *Org. Lett.*, 2016, **18**, 4020–4023.

89 X.-J. Cheng, L.-L. Liang, K. Chen, N.-N. Ji, X. Xiao, J.-X. Zhang, Y.-Q. Zhang, S.-F. Xue, Q.-J. Zhu, X.-L. Ni and Z. Tao, *Angew. Chemie Int. Ed.*, 2013, **52**, 7252–7255.

90 W. A. Freeman, W. L. Mock and N. Y. Shih, *J. Am. Chem. Soc.*, 1981, **103**, 7367–7368.

91 E. Masson, X. Ling, R. Joseph, L. Kyeremeh-Mensah and X. Lu, *RSC Adv.*, 2012, **2**, 1213–1247.

92 S. J. Barrow, S. Kasera, M. J. Rowland, J. del Barrio and O. A. Scherman, *Chem. Rev.*, 2015, **115**, 12320–12406.

93 K. Kim, J. Murray, N. Selvapalam, Y. H. Ko and I. Hwang, in *Cucurbiturils*, World Scientific (Europe), 2018, pp. 9–29.

94 K. I. Assaf and W. M. Nau, *Chem. Soc. Rev.*, 2015, **44**, 394–418.

95 C. Márquez, R. R. Hudgins and W. M. Nau, *J. Am. Chem. Soc.*, 2004, **126**, 5806–5816.

96 W. L. Mock and N. Y. Shih, *J. Org. Chem.*, 1986, **51**, 4440–4446.

97 W. L. Mock and N. Y. Shih, *J. Am. Chem. Soc.*, 1988, **110**, 4706–4710.

98 A. I. Day, R. J. Blanch, A. P. Arnold, S. Lorenzo, G. R. Lewis and I. Dance, *Angew. Chemie Int. Ed.*, 2002, **41**, 275.

99 S. Liu, P. Y. Zavalij and L. Isaacs, *J. Am. Chem. Soc.*, 2005, **127**, 16798–16799.

100 X. Yang, Z. Zhao, X. Zhang and S. Liu, *Sci. China Chem.*, 2018, **61**, 787–791.

101 K. Kim, J. Murray, N. Selvapalam, Y. H. Ko and I. Hwang, in *Cucurbiturils*, World Scientific (Europe), 2018, pp. 31–55.

102 K. I. Assaf, M. A. Alnajjar and W. M. Nau, *Chem. Commun*, 2018, **54**, 1734.

103 W. L. Mock and J. Pierpont, *J. Chem. Soc. Chem. Commun.*, 1990, 1509.

104 D. Tuncel, Ö. Özsar, H. B. Tiftik and B. Salih, *Chem. Commun.*, 2007, 1369–1371.

105 P. D. Beer, P. A. Gale and D. K. Smith, *Supramolecular Chemistry*, Oxford University Press, Oxford, 1999.

106 M. Sojka, M. Fojtu, J. Fialova, M. Masarik, M. Necas and R. Marek, *Inorg. Chem.*, 2019, **58**, 10861–10870.

107 J. B. Wittenberg, M. G. Costales, P. Y. Zavalij and L. Isaacs, *Chem. Commun.*, 2011, **47**, 9420.

108 Y. Yu, Y. Li, X. Wang, H. Nian, L. Wang, J. Li, Y. Zhao, X. Yang, S. Liu and L. Cao, *J. Org. Chem.*, 2017, **82**, 5590–5596.

109 W. L. Mock, T. A. Irra, J. P. Wepsiec and M. Adhya, *J. Org. Chem.*, 1989, **54**, 5302–5308.

110 K. D. Hänni and D. A. Leigh, *Chem. Soc. Rev.*, 2010, **39**, 1240–1251.

111 X. Yang, F. Liu, Z. Zhao, F. Liang, H. Zhang and S. Liu, *Chinese Chem. Lett.*, 2018, **29**, 1560–1566.

112 J. Malina and M. J. Hannon, *Unpublished Results*, .

113 I. Bertini, C. (Claudio) Luchinat, G. Parigi and E. Ravera, *Solution NMR of Paramagnetic Molecules*, Elsevier, 2017.

114 L. Cardo and M. J. Hannon, Unpublished Work, 2018.

115 J. L. Moore, S. M. Taylor and V. A. Soloshonok, *Arkivoc*, 2005, **2005**, 287.

116 V. Sadovnikova, PhD Thesis, University of Birmingham, 2011.

117 E. B. Hughes, H. H. G. Jellinek and B. A. Ambrose, *J. Phys. Colloid Chem.*, 1949, **53**, 410–414.

118 A. Sedaghat Doost, M. Akbari, C. V. Stevens, A. D. Setiowati and P. Van der Meeren, *Trends Food Sci. Technol.*, 2019, **86**, 16–24.

119 L. Avram and Y. Cohen, *Chem. Soc. Rev.*, 2015, **44**, 586–602.

120 Introduction to Ultraviolet-Visible Spectroscopy (UV), http://www.chemguide.co.uk/analysis/uvvisiblemenu.html#top, (accessed 13 August 2019).

121 D. Sheehan, *Physical Biochemistry: Principles and Applications*, Wiley-Blackwell, 2nd ed., 2009.

122 B. Nordén and F. Tjerneld, *Biopolymers*, 1982, **21**, 1713–1734.

123 S. Yousuf and I. V. M. V. Enoch, *AAPS PharmSciTech*, 2013, **14**, 770–781.

124 A. P. Paneerselvam, S. S. Mishra and D. K. Chand, *J. Chem. Sci.*, 2018, **130**, 96.

125 D. M. Gray, R. L. Ratliff and M. R. Vaughan, *Methods Enzymol.*, 1992, **211**, 389–406.

126 J. Kypr, I. Kejnovská, D. Renčiuk and M. Vorlíčková, *Nucleic Acids Res.*, 2009, 37, 1713–1725.

127 T. Qin, K. Liu, D. Song, C. Yang, H. Zhao and H. Su, *Int. J. Mol. Sci.*, 2018, **19**, 1071.

128 T. Šmidlehner, I. Piantanida and G. Pescitelli, *Beilstein J. Org. Chem.*, 2018, **14**, 84–105.

129 N. Holmgaard List, J. Knoops, J. Rubio-Magnieto, J. Idé, D. Beljonne, P. Norman, M. Surin and M. Linares, *J. Am. Chem. Soc.*, 2017, **139**, 14947–14953.

130 N. C. Garbett, P. A. Ragazzon and J. B. Chaires, *Nat. Protoc.*, 2007, **2**, 3166–3172.

131 E. M. Tuite and B. Nordén, *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.*, 2018, **189**, 86–92.

132 M. Radić Stojković, S. Miljanić, K. Mišković, L. Glavaš-Obrovac and I. Piantanida, *Mol. Biosyst.*, 2011, **7**, 1753.

133 C. Uerpmann, J. Malina, M. Pascu, G. J. Clarkson, V. Moreno, A. Rodger, A. Grandas and M. J. Hannon, *Chem. - A Eur. J.*, 2005, **11**, 1750–1756.

134 A. Rodger, *Sci. Prog.*, 2008, **91**, 377–396.

135 A. Rodger, G. Dorrington and D. L. Ang, *Analyst*, 2016, **141**, 6490–6498.