METALLO-SUPRAMOLECULAR HELICATES AND THEIR ABILITY TO BIND TO HIGHER-ORDER DNA STRUCTURES

By

HANNAH MARY PETERS STOCK

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

August 2024

Abstract

Targeting DNA in order to develop therapeutic agents which treat disease is commonplace as many disorders arise from the expression of DNA, where the amplification, and expression as RNA or proteins, of the DNA code has led to the disease manifestation. However, many DNA-binding agents are non-specific and bind non-selectively to DNA, affecting numerous genes and consequently, many cellular processes.

Despite this, metallo-supramolecular compounds have shown the ability to overcome the selectivity issues of other DNA-binding agents, through structure specific targeting of DNA. A metallo-supramolecular helicate, previously developed by the Hannon group, demonstrated the ability to selectively target higher-order DNA structures, such as DNA three-way junctions (3WJs), four-way junctions (4WJs) through structure specific binding mechanisms. Here a significantly longer metallo-supramolecular helicate was synthesised in order to investigate how length impacted the ability of a helicate to bind to higher-order DNA structures. A number of *in vitro* and *in silica* experiments were conducted, which demonstrated that the novel helicate bound to the central cavity of both 3WJs and 4WJs in a similar way to previously developed helicates, and furthermore, that the different enantiomers of this novel helicate showed differing binding activity.

Overall, the research presented in this thesis adds a new triple-stranded helicate to the current, known portfolio and also presents a metallo-supramolecular helicate that has demonstrated a strong ability to bind to higher-order DNA, suggesting its potential application to therapeutics.

Acknowledgements

Firstly, I would like to extend my thanks to Professor Hannon for not only giving me the opportunity to work in his group but also for his unwavering support throughout the year of research and during the write-up. His encouragement and advice have been instrumental in the completion of my degree, and I am truly grateful for the time and help he provided.

Secondly, I am grateful to all the members of the Hannon group for guiding me through this process and teaching me everything I know about research. Working with you all has been thoroughly enjoyable, and I wish the team the best in their future endeavours. I would especially like to acknowledge two members who had a significant impact on my time within the team – Cat Hooper and Sam Dettmer. Cat, thank you for being such a wonderful role model in the lab. Your enthusiasm and kindness helped me both professionally and personally, and I feel extremely fortunate to have had your guidance, especially during the synthesis phase. Sam, I am hugely appreciative of the time you took to teach me and your patience with my extensive questioning. Your support throughout this whole process has been invaluable.

Finally, I would like to give a huge thank you to my friends and family, without whom I would have struggled to push through to the end. A special thank you to my brothers, Tom and Dan, who both provided me with great practical advice, to my parents, who were there to pick me up in the lows of this experience and to my partner, Liam, who never let me give up – I dedicate this thesis to you all.

List of Abbreviations

3WJ	Three Way Junction
4WJ	Four Way Junction
A	Adenine (a DNA base)
ACN	Acetonitrile
CD	Circular Dichroism
COSY	Correlated Spectroscopy
C	Cytosine (a DNA base)
ctDNA	Calf Thymus DNA
DCM	Dichloromethane
DMS	Dimethyl Sulphide
DMSO	Dimethyl Sulphoxide
DNA	Deoxyribonucleuic Acid, where letters before the word indicate the specific structure of the DNA (e.g A, B, Z or ss for single-stranded and ds for double-stranded)
ESI	Electrospray Ionisation
G4	G-Quadruplex
G	Guanine (a DNA base)
LD	Linear Dichroism
MeOH	Methanol
NMR	Nuclear Magnetic Resonance
NOESY	Nuclear Overhauser Effect Spectroscopy
T	Thymine (a DNA base)
UV-Vis	Ultra-Violet Visible

Contents

Chapter 1: METALLO-SUPRAMOLECULAR COMPOUNDS AND THEIR ABILITY TO BIND HIGHER-ORDER DNA STURCTURES

1.1) Introduction

DNA has long been an appealing target for researchers attempting to find therapeutic and biomedical solutions to the most untreatable diseases.[1] One focus of DNA research is founded in the basis that many disorders and cancers arise through expression of the DNA code, where the amplification, and expression as RNA or proteins, of the code has led to the disease manifestation.[1] Therefore, drugs that selectively target DNA, preventing gene expression, can be used to treat and supress these diseases.[1] Cancer is one such disease that can be generated by erroneous gene expression. Mistranslation of DNA (or translation of damaged DNA) resulting in incorrect protein formation, under-expression of suppressor proteins or over- expression of growth factors can all lead to oncogenesis.[2,3] All three, in turn, can be succeeded by rapid uncontrolled cell proliferation which causes the formation of tumours.[3,4] Furthermore, these tumours are perpetual as cancer cells utilise telomerase to add telomeric segments to the ends of DNA enabling the cells to be immortalised. [3,4] Therefore, for any synthetic compound to be an effective anti-cancer drug through DNA recognition, it must be able to prevent both uncontrolled cell replication and immortalisation.

However, while carcinogenesis is often the consequence of erroneous gene expression it is important to clarify that other factors such as DNA mutation, chromosomal aberrations and epigenetic alterations can also lead to this inaccuracy in gene expression.[2] One such example is single nucleotide polymorphism (SNP). SNP is where a single nucleotide base in a genetic sequence has been changed and this alteration can give rise to faulty gene

expression.[5] The location of the SNP within the gene will impact its effect on gene expression, for instance, when located in the exonal region the SNP may supress transcription and translation of the gene, while if found in the intron, region splice variants of transcripts may be produced.[5] Nevertheless, the result is often consistent – gene expression is negatively impacted and consequentially causes tumorigenesis.

DNA binders traditionally focus on DNA recognition through five individual approaches:[6–9] intercalation between DNA bases, covalent coordination to DNA bases, hydrogen bond formation to the major groove, bond formation to the minor groove and bond formation to the sugar-phosphate backbone.[6] However, clinical anti-cancer drugs currently only act by targeting the minor groove, by base coordination/alkylation or by intercalation. Cisplatin is one such drug that works by coordinating to purine bases, precluding DNA replication, cell division and eventually causing cell death.[10] However, most DNA-binding drugs, cisplatin among them, are non-specific which causes two key issues. Firstly, it means these drugs do not discriminate between healthy and cancerous cells, and therefore they have large cytotoxic effects on all cells. Secondly, these drugs bind non-selectively to DNA, affecting many genes and thus many cellular processes. The consequences of this that the patient receiving treatment is likely to experience significant discomfort.[6] Attempts to structurally modify these drugs to enable them to bind selectively to specific DNA sequences have so far been unsuccessful.[6] Despite this, recent developments within the metallo-supramolecular field have opened the question of whether structure specific targeting of DNA by binders could be a more effective route to attaining specificity of anti-cancer drugs.[6] One of the most notable examples of this concept is a metallo-supramolecular cylinder developed by the Hannon group that bonds

preferentially to three-way-junctions (3WJs) and four-way-junctions (4WJs) over B-DNA structures. [6,11]

The focus of this thesis is to investigate the binding capabilities of a novel helicate to higher-order DNA structures, such as DNA 3WJs and4WJs as well as the more complexed folded DNA structure, the G-quadruplex. The latter DNA structure was chosen as an additional target in this thesis as it is known to a prevalent DNA structure in cancerous cells.The novel helicate has similar functional groups to the parent cylinder, however, is much longer, so a direct comparison between abilities of the two cylinders can be made in order to ascertain the full potential of novel cylinder. Firstly, this thesis will contextualise its findings through a review of the background literature that underpins the project.

1.2) DNA Structures and their Binders

1.2.1) DNA Three-Way-Junctions

In general, a double stranded helix (B-DNA) with Watson-Crick base pairing between the two polynucleotide strands is regarded as the standard structure of DNA.[12] However, structures other than the typical B-DNA duplex can be transiently and permanently formed during the initial replication and transcription phases within cells.[12] One such transient structure that forms during DNA replication is the 3WJ-like structure, termed a Y- fork.[6] The structure is generated when DNA helicases bind to and 'unzip' the double stranded DNA (dsDNA) strands leading to a system that has two single stranded DNA (ssDNA) branched arms and one dsDNA arm.[13] 3WJs are higher order DNA assemblies that are very similar to Y-forks and are constructed from three intersecting dsDNA strands as shown in Figure 1.2.2.1.[14]

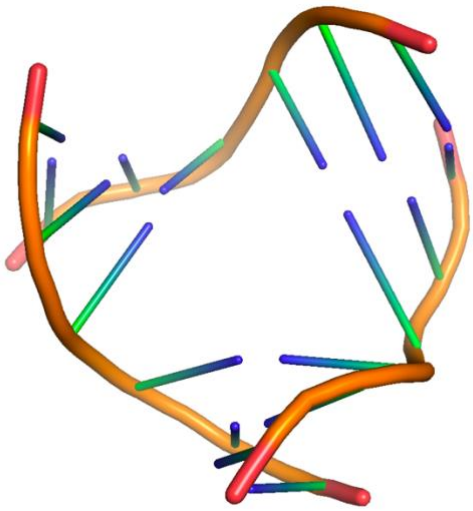


Figure 1.2.2.1 Structure of DNA 3WJ (from PDB ET20[6])

Torsional stress can cause 3WJs to form ahead of the Y-forks.[15] Their formation often presents an obstruction to the DNA replication process and consequently, it is unsurprising that they have been linked to numerous degenerative diseases, such as Huntington's disease.[15] The centre of the 3WJ assembly is a hydrophobic cavity approximately 12 Å in diameter and is known as the branching point.[14] Binding of certain molecules to the branching point stabilises the structures and this property is exploited by molecules with anti-cancer activity.

1.2.2) Three-way-junction (3WJ) Binders

3WJ binders are any molecular structures that are able to bind to the cavity within DNA 3WJs, which in turn has a stabilising effect on the junction. The next section will discuss the different types of 3WJ binders present within the literature.

1.2.2.1) Helical Binders to Three-Way-Junctions

Many 3WJ-binders have been developed in recent years with the purpose of inducing cell death in cancerous cells. By binding to 3WJs selectively, it theorised that these binders can interfere with the replication fork, eventually resulting in cell death of the cancerous cell.[16] One such 3WJ-binder with potential anti-cancer activity is the triple-stranded helicate developed by the Hannon group in 1997.[17] This supramolecular architecture is formed from ligands with pyridylimine co-ordination sites and octahedral metal ions in a 3:2 ratio.[17] The geometry of the compound displays axial chirality due to the helical twist induced by the ligands. The electrostatic attraction between the anionic 3WJ and the positive charge of the metal ions within the helicate drives the interaction between the

DNA and helicate.[6] However, the pi-stacking interactions between the phenyl groups with the ligands of the helicate and the DNA bases at the junction's centre are responsible for the helicate being bound within the junction.[6] Another helical structure that is able to bind to 3WJs is the polyazamacrocycle developed by the Monchaud group, shown in Figure 1.2.2.1.[16] The small molecule is comprised of 1,4,7-triazacyclononane with side-arms constructed from quinoline units and has the ability to bind mono- and bi- cationic metals to its centre. The innate ability to bind metal ions is what enables the small molecule to interact with 3WJs.[16] The presence of excess Li^+ ions in particular induces the macrocycle to undergo helicoidal folding in the cavity of the 3WJ. As previously discussed, the cavity of the 3WJ has a diameter of 12 Å and this means the induced helicate, which had a diameter of 10 Å, fits perfectly within the junction. Moreover, the quinoline side-arms form pi-stacking interactions with the DNA bases at the centre of the branched point, stabilising both structures and thus enabling the binding event to take place.[16]The work produced by Monchaud's group showed clearly that this polymacrocycle had real potential to cause an effective and efficient anti-cancer response in cells and further highlighted that polymacrocycles in general have an under-researched capability to act as effectual therapeutic agents.

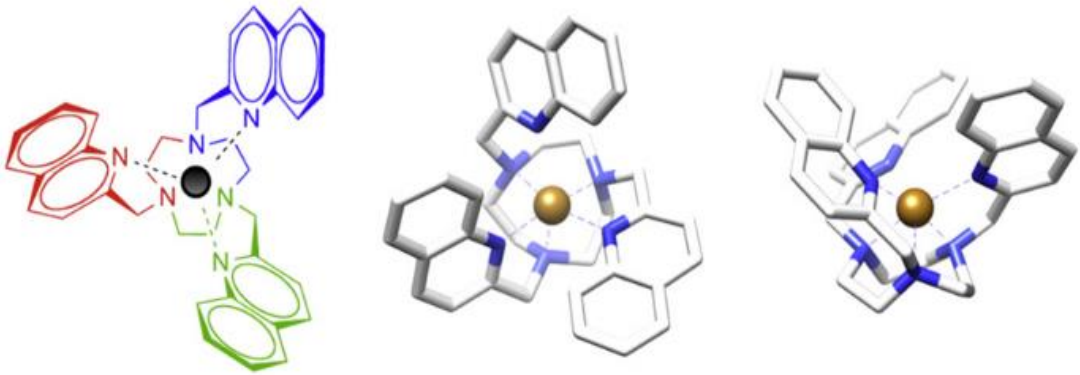


Figure 1.2.2.1 Structure of Monchaud's polymacrocycle including a 3D side view and top view[16]

However, Monchaud's group is not the only research group that recognized how effective helical structures are at interacting with 3WJs. The Vasquez group, based in Spain, has also been able to bind helical structures within the hydrophobic cavity of 3WJs.[18,19] Similarly to Monchaud's research, the organic structures developed by Vasquez fold into helical forms in the presence of metal ions, however, unlike Monchaud's polymacrocyclic molecules, Vasquez's structures are peptide based.[18,19] The principle behind Vasquez's use of peptide structures is that by carefully selecting the residues present within the peptide sequence, the group could selectively code for stereo-specific folding of peptide. Moreover, the deliberately constructed sequence within the ligand would enable the optimisation of the chemical and physical properties of metallo-peptides developed.[18] In 2021, the group reported the successful development of such oligocationic peptides.

These peptides were able to coordinate two Fe(II) ions through bipyridine units within the ligand and presence of the metal ions consequently resulted in the folding of the single-stranded peptide to produce a triple-stranded helicate. [18] The group further reported

that these helicates were able to bind DNA 3WJs with high affinity and specificity. From this, Vasquez theorised that the same technique of encoding chiral and structural information into a peptide through careful selection of the residues in the sequence) could be applied to helicates formed from three shorter peptide strands to force the formation of a specific helical enantiomer over the other.[19] Within the same year, Vasquez' group achieved this through the modification of a short β-annulus motif to contain bipyridine units.[19] Three of this shorter peptide sequence can come together in the presence of Fe(II) ions to form a triple-stranded helicate that has solely P chirality as shown by Figure 1.2.2.2.[19] Furthermore, the helicate displayed a high affinity and selectivity to DNA 3WJs.[19] As a result of these papers, the Vaquez group have confirmed beyond all doubt that combing peptide engineering with coordination chemistry creates a pragmatic approach for developing metallopeptides that have palpable anti-cancer activity.

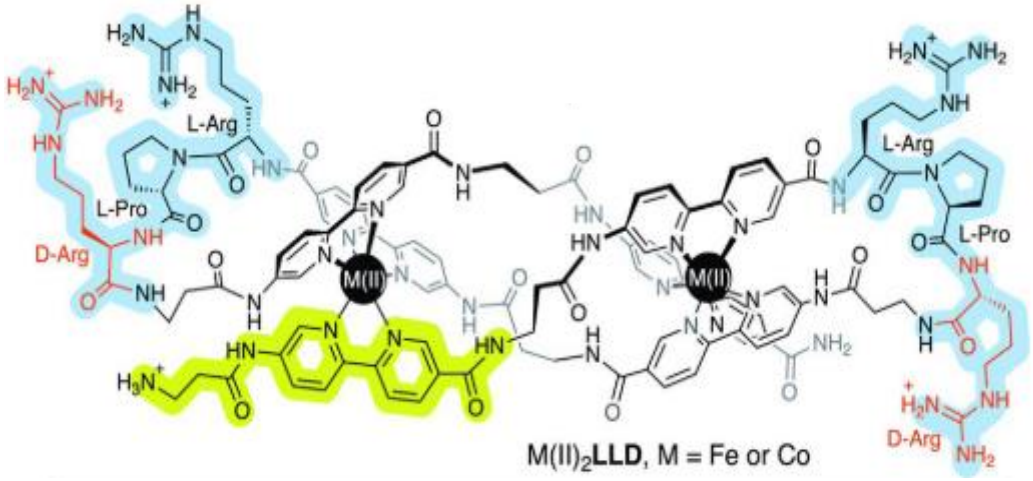


Figure 1.2.2.2 Structure of Vasuqez's helicate where the ligand is generated from peptide synthesis. b-turn sequences (RPr) are highlighted in blue and onbe bAlaBpy building block is highlighted in green.[19]

1.2.2.2) Non-helical Binders to Three-Way-Junctions

It is evident that metallo-supramolecular molecules with a helical structure have a strong affinity for 3WJs, as shown by the research of Hannon, Monchaud and Vasquez. However, it is not the helical nature of these structures that enables them to bind so effectively to 3WJs, but rather the three-fold symmetry that these helical structures possess. In 2019, the Nitschke group identified this property and utilised it to synthesise a non-helical tetrahedron that was able to bind to 3WJs, shown in Figure 1.2.2.3.[20] The tetrahedral cage was formed from four Fe(II) ions coordinated to modified tri-aniline molecule and had a trigonal, three-dimensional shape which gave it the desired three-fold symmetry. Binding studies completed by the group showed the cage had a clear preference for cavity binding of the 3WJ.[20] The group theorised this was due to the nature of the central cavity, namely that the DNA bases located nearest to the cavity were much less stable and therefore, had a higher probability of being unpaired.[20]

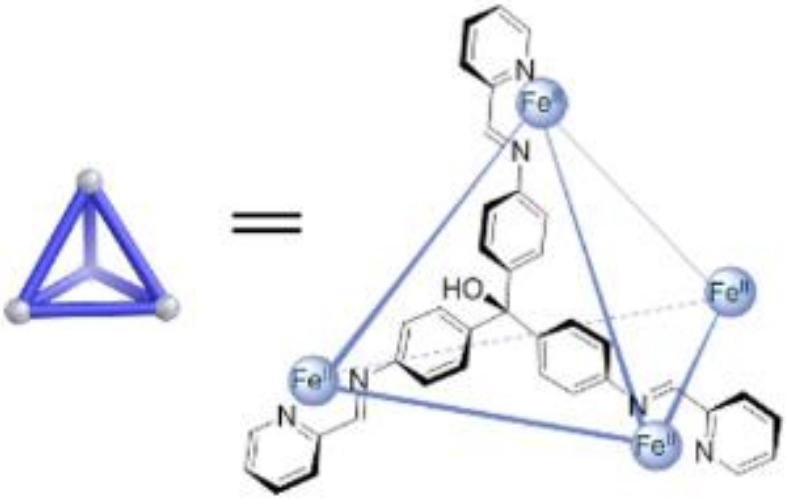


Figure 1.2.2.3 Structure of Nitschke's tetrahedron $[Fe_2L_4]^{4+}$ [20]

Consequently, the group suggested that the dynamic arrangement of these bases could induce the cavity to form a trigonal conformation that is much more suitable for binding the tetrahedral cage.[20] This theory is widely supported within the literature[6,21–24] and while group does report the tetrahedron has an affinity for unpaired bases it is likely that 3WJ cavity conformation has large impact on successful binding of the cage.

Another class of 3WJ binders that do not possess a helical structure are triptycene-based molecules developed by the Barros group, shown in Figure 1.2.2.4.[25] Like Nitschke, Barros identified that three-fold symmetry was a critical characteristic required by a molecule for it to successfully bind to a 3WJ.[25] Additionally, Barros recognized that triptycene had this three-fold symmetry as well as comparable dimensions to the 3WJ cavity.[25] On further analysis in the form of binding studies, Barros confirmed that modified triptycene molecules successfully recognized and bound to 3WJs and he further suggested that, like Nitschke's tetrahedron, the triptycene-based units would have a conformation-selective fit with the 3WJ.[25] This hypothesis was founded on the basis that aromatic rings of the triptycene molecules were not planar and therefore, could interact with the DNA bases closest to the cavity centre. These interactions would likely be in the form of pi-stacking or pi-buckled interactions and would drive a conformation-selective binding event between the triptycene molecule and the junction.[25] Similarly to Nitschke, the hypothesis is widely supported within the literature[6,21,26,27] and provides a new perspective on how 3WJs may bind different molecules through conformation-selective fit.

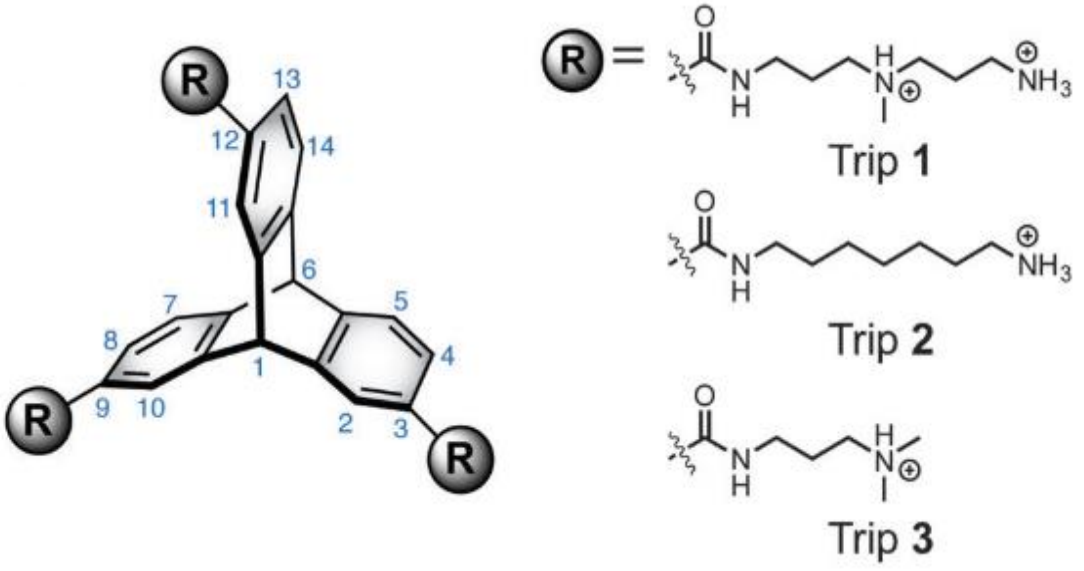


Figure 1.2.2.4 Structure of Barros's triptycene-based units[25]

Overall, while all these 3WJ binders are formed from different approaches and have very different base molecules, there are many parallels that can be drawn between all of them. Most importantly, they all have three-fold symmetry and all have conformational and dimensional compatibility with the cavity of the 3WJ. Moreover, they all display clear potential as future therapeutic agents and present significant progress in field of bioactive innovation.

1.2.3) DNA Four-Way-Junctions

Another higher order DNA arrangement is 4WJs or Holliday junctions, shown in Figure 1.2.3.1. This structure is formed as an intermediate during homologous recombination and its formation is vital for the cellular process to be successful.[6] 4WJs share many

structural similarities with 3WJs such as dsDNA segments converging to a form a system with a central cavity.[28]

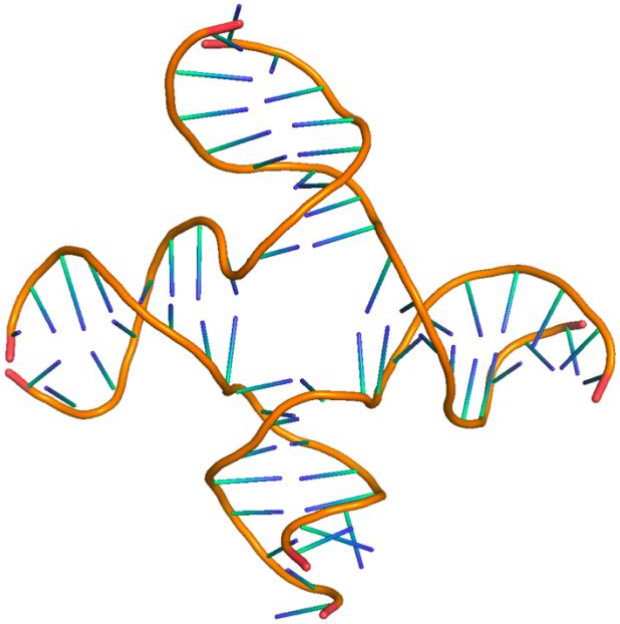


Figure 1.2.3.1 Structure of DNA 4WJ (adapted by collegue from PDB 1XNS as seen the Hannon group's pillarplex paper[11])

However, there are key compositional differences between 3WJs and 4WJs. Firstly, 4WJs are formed from 4 duplex DNA strands intersecting rather than three. Secondly, 4WJs can exist in two different conformations, an open square-planar shaped conformation and a more compact x-stacked conformation, however, it should be noted than when in its bound state 4WJs exist in the planar-cruciform conformation exclusively.[28] Furthermore, it is important to recognise that 4WJs are much more stable than 3WJs so while both can exist without a molecule bound in their central cavity, this is much more likely to be seen for the 4WJs (hence they exist as intermediates). Despite the differences in structure and stability, the cavity of 4WJs remains as a binding point for molecules and therefore, these DNA structures can also be exploited by anti-cancer drugs. Interestingly, although 4WJs have long represented a realistic target for anti-cancer drugs, only the Hannon group have

successfully confirmed binding a molecule within the 4WJ cavity. In a recently published paper, the group report binding an organometallic pillarplex to the centre of open cruciform conformation of 4WJs, in a mode of binding that had never before been replicated by other synthetic agents.[11] The results provide a new metallo-supramolecular method of targeting higher order DNA and RNA junctions. [11] These findings are particularly exciting as while targeting other non-canonical forms of DNA (like G-quadruplexes) has been thoroughly researched, the selective binding of therapeutic molecules to junction-like structures is still widely unexplored.

1.2.4) Four-way-junction (4WJ) Binders

As discussed, the organo-metallic pillarplex, shown in Figure 1.2.4.1, developed by the Hannon group is the only molecule confirmed to bind to the cavity of a 4WJ in its open cruciform form.[11] The Hannon group carefully designed the molecule to have properties that would promote binding of the pillarplex with planar 4WJs. For example, the pillarplex contains 12 aryl rings (4 sets of 3) within its structure which enables the potential for stabilising pi-stacking interactions to form between the molecule and the

DNA bases closest to the cavity, thus favouring binding between the pillarplex and the junction.[11]

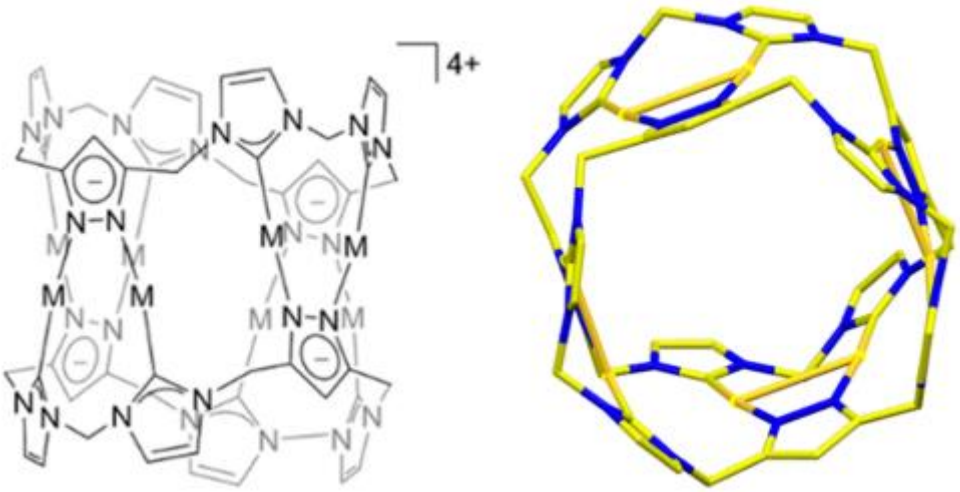


Figure 1.2.4.1 Structure of Hannon's pillarplex, including 3D view along cylindrical axis[11]

As well as this, the pillarplex has a tetracationic charge due to the 8 gold centers coordinated within the structure. The high positive charge enables stabilising electrostatic interactions between the pillarplex and the anionic DNA, again favouring a binding event.[11] Binding studies completed by the group confirmed that the pillarplex successfully bound to the cavity with the 4WJs, stabilising the open form. Moreover, molecular dynamic simulations produced by the group provided evidence for the binding mechanism. The simulations showed that due to the girth of the pillarplex, two base-pairs closest to the junction transiently break to enable the molecule to move into the cavity. Once the pillarplex moves into the cavity, these base-pairs reform and the open-cruciform structure of the 4WJ is stabilised.[11] Additionally, all the simulations show that the pillarplex stays within the cavity for the entirety of the simulation (2 μs) clearly indicating the favourability of the interaction. The success of this research provides a strong foundation for the development of more molecules that can bind to open-cruciform 4WJs

but more importantly, highlights the real potential of selective targeting of 4WJs to induce anti-cancer activity.

While the Hannon group is the only group to have confirmed binding a molecule to the cavity of an open-cruciform style 4WJ, two other groups have been able to design molecules that generate 4WJ-like structures from their respective interactions with DNA.[29–31] In 2019, Bonnet's group reported a platinum(II) complex, shown in Figure 1.2.4.2, that was able to force a DNA oligomer 5'-d(CGTACG)-3' to adopt a 4WJ-like structure.[29] The mononuclear complex interacts with four strands of oligomer through pi-stacking and hydrogen bonds to induce the 4WJ-like motif, however, remarkably the platinum centre itself does not form any coordination bonds with the duplex DNA.[29] Similarly to the Hannon group, Bonnet discusses the importance of careful complex design to ensure that the complex has molecular properties that that favour interaction with DNA.[29]

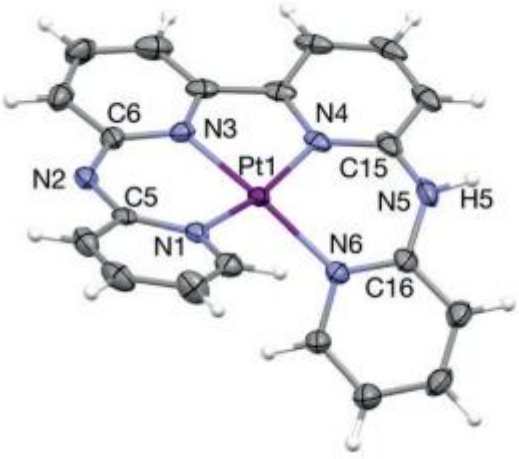


Figure 1.2.4.2 Displacement ellipsoid plot (50% probability level) of Bonnet's 4WJ binder [Pt(Hbapbpy)]$^{4+}$ excluding counterions[29]

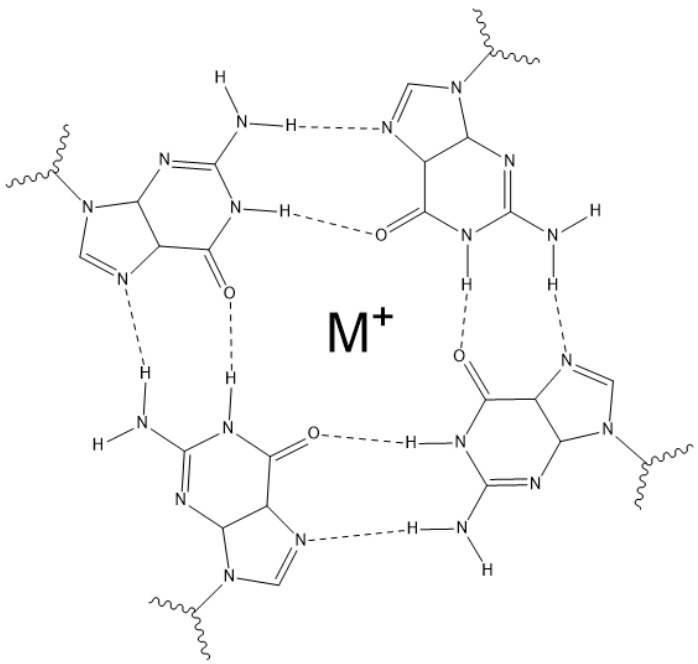
For example, the platinum complex has a cationic charge so there are strong electrostatic forces of attraction between the complex and the anionic DNA. Moreover, the aryl rings within the complex enable pi-stacking interactions to form between the complex and DNA. Both these factors help to drive the formation of the 4WJ-like motif from the oligomer strands and are a result of deliberate decision making in design of the complex.[29]

Like the Hannon group's pillarplex, Bonnet's 4WJ binder is a metallo-supramolecular molecule. However, the final 4WJ binder is a small organic molecule that binds to DNA by acting as a bis-intercalator.[30,31] In 2009, the Searcey group reported that they had synthesised an asymmetrical bis-acridine molecule and that the oligonucleotide d(TCGGTACCGA) had crystallized into a 4WJ in the presence of this small molecule.[31] This was surprising discovery as until that point the sequence had only crystalized to a 4WJ in the presence of metal ions.[31] On further analysis the group found that the small organic molecule bound across the junction in a non-covalent fashion and that molecule interacts with the DNA in a like manner to Bonnet's binder. Searcey's bis-acridine molecule is positively charged so, like Bonnet's complex, has the ability to neutralise the anionic charge of the DNA which is particularly high at the centre of the junction.[31] Moreover, the chromophores within the bis-acridine molecule are able to stabilise pi-stacking interactions with the DNA base pairs and like Bonnet's complex, this helps to drive formation of the 4WJ.[31] It is important to note Searcey's molecule does not bind to the cavity inside the 4WJ junction and therefore, binds to the X-stacked form of the 4WJ. This is alternate binding mechanism to the pillarplex which, as discussed, binds to the cavity of the open cruciform form of the 4WJ.

Currently, few 4WJ binders exist, however, in the last decade much progress has been made in developing such molecules. While the current array of 4WJ-binders interact with DNA by wholly dissimilar mechanisms, there are still strong parallels between all molecules, namely that all these compounds have been deliberately designed to have properties that favour DNA binding. Additionally, this has provided a clear framework for development of future 4WJ-binders.

1.2.5) G-quadruplexes

G-quadruplexes, shown in Figure 1.2.5.1, are an alternative higher order DNA structure that exist in two main forms, the parallel and the anti- parallel, although occasionally other hybrid forms are seen. These are complex folded DNA structures that are found at



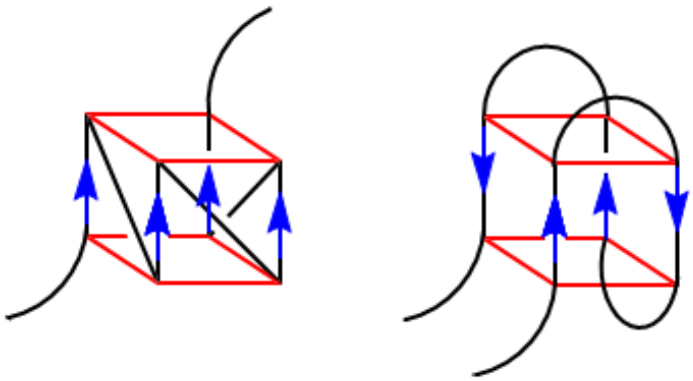


Figure 1.2.5.1 Top image shows a single G4 tetrad with hydrogen bond between the bases. The bottom image is a schematic depiction of G-quadruplex in both its parallel (left hand side) and anti-parallel (right hand side) forms. The blue arrows show the direction of the DNA strand, the red squares illustrate the guanine tetrameres and the black lines show the DNA backbone.

the ends of chromosomes in telomeric regions and in transcriptional regions of a few key oncogenes.[32] The G-quadruplexes are constructed of stacked tetrads stabilised by a Na^+ or K^+ located at the centre of the arrangement.[33] The tetrads themselves are formed from the planar interaction of four guanines, such that they are arranged cyclically with hydrogen bonds between them.[32] This is also known as a Hoogsteen arrangement. [32]Although G-quadruplexes do not have a hydrophobic cavity within which molecules can bind, their chirality means that chiral molecules can interact and stabilise the structure and this property is utilised by anti-cancer drugs to enable binding.[32,34]

In vivo, over 30 naturally occurring proteins have been found to bind to G4-structures, either promoting of disrupting G-quadruplex formation[35] and selective targeting of these G4 structures by synthetic molecules has been extensively researched.[11] However, designing therapeutic molecules to recognise these structures has proved challenging. Most recently, research has focused on developing metallated G4 binders due to the

advantages structural and electronic properties that metal complexes possess.[35] For example, metal complexes can change their coordination geometries based on the ligands bound to them. Exploitation of this characteristic can optimise G4-binding by facilitating the lowest energy arrangement of the metal complex with the G-quadruplex, ultimately ensuring successful interaction.[35] Moreover, metal centres are electron withdrawing, hence form complexes that tend towards being electron poor. These electron deficient systems are then able to bind much more strongly to the electron-rich G4 structures through π-interactions.[35] It must be noted that the majority of these metallated G-quadruplex binders interact with G4 quartets by π-stacking to the top of the quadruplex structure.[36] For example, metalloporphyrins, metallophthalocyanines and metallocorroles all favour this form of π-stacking.[36] Additionally, the electropositive nature of metal complexes means that other mechanisms of binding to G-quartets are accessible.[35] Namely, that the metal can replace the alkali metal at situated at the centre of the G4 structure, stabilising the G4 arrangement.[35] However, while this mechanism has been well theorised, it has yet to have been achieved experimentally with metal complexes.

1.2.6) G-quadruplex Binders

Although metal complexes as G4 binders have been the focus of significant research over the last few years, there has also been remarkable progress made in the development of other types of G4 binders. Most notably, Janez Plavec and his research group managed to intercalate a U-shaped ligand named Phen-DC$_3$, shown in Figure 1.2.6.1, into the centre of a G-quadruplex, thus substituting the alkali metal usually located there (as discussed in section 1.2.5 G-quadruplexes).[37] Once more, while this intercalation mechanism had

been theorised by molecular dynamic models it had not been achieved until now.[37] Interestingly, the intercalation of the Phen-DC_3 into the G4 quartet forces the adoption of the anti-parallel G4 structure, where the quinolinium rings within the ligand drive the process due to their favourable interactions within the quadruplex.[37] The two quinolinium rings' positioned within the G-quadruplex enable them to form favourable stacking interactions with the guanine rings while also optimising π-π overlap with the G-quartets, thus they have a stabilising effect on the overall arrangement.[37] The development of such a ligand that interacts with G-quadruplexes in this way is highly exciting as shows the real potential for therapeutic targeting this non-canonical DNA structure.

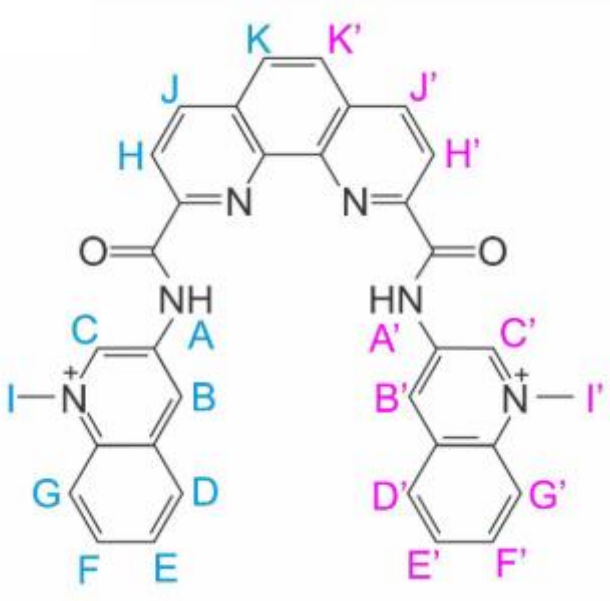


Figure 1.2.6.1 Chemical structure of Plavec's G4 binder (Phen-DC_3) where atomic numbering (and colouring) shows the asymmetry of the compound[37]

Another G4 binder that is of significant note is the metallo-supramolecular cylinder, termed the parent cylinder.[38] This cylinder was developed by the Hannon group in 1997,

however, it was the Qu group from China that managed to successfully bind the P enantiomer of the cylinder to a G-quadruplex in 2008.[38] The group reported that the cylinder stacked externally to the G-quadruplex, rather binding within the G4 structure like Plavec's ligand. The binding of the cylinder to the DNA was driven by the stabilising electrostatic interactions between the positively charged cylinder with the anionic DNA – something that has been seen consistently throughout the literature. Additionally, the group further reported that in the presence of sodium ions, binding of the P-enantiomer to a G-quartet results in conversion of the G4 from its anti-parallel form to a more unique hybrid structure.[38] The evidence presented by the Qu group indicates that the potential of DNA recognition by metallo-supramolecular cylinders is not limited to simply B-DNA and that therapeutic applications of these cylinders may extend much further than initially realised.

Overall, while many G4 binders exist, the two examples presented represent the most exciting developments in recent years. Plavec ligand's ability to coordinate to the centre of the G4 is a unique binding mechanism that had not been previously achieved, while until the Qu group's research, metallo-supramolecular cylinders had not been shown to interact with these higher-order DNA structures. The therapeutic possibilities of both these molecules represents an exciting prospect for future research.

1.2.7) How binding to higher-order DNA structures results in anti-cancer activity

As discussed, the DNA structures above do not represent the majority structure of DNA within cells but rather structures that are formed briefly during key DNA processes.[6,14]

The stabilisation of these transient structures can cause cytotoxic effects and can lead to cell death. Therefore, it is through the targeted binding and stabilisation of these noncanonical DNA structures that anti-cancer activity can be induced.[14]

Overall, the literature definitively concludes that stabilisation of noncanonical DNA structure leads to DNA breakage and cell apoptosis.[14,39] Nevertheless, the mechanism by which these stabilised DNA structures induce cell apoptosis remains uncertain. It is theorised in the literature, that the stabilisation of noncanonical DNA causes an obstruction to the translocation of the polymerase enzyme.[39,40] In turn, the obstruction of these key enzymes leads to important cellular processes (such as the replication and translation of the cellular DNA) being severely hindered, resulting in protein machinery stalling or collapse, DNA deterioration and eventually cell death. [39,40] This hypothesis is widely regarded as the most logical explanation for how anti-cancer activity is achieved through the stabilisation of 3WJs and 4WJs.

However, there is a debate within the literature as to whether this theory is applicable to G-quadruplexes. A paper written in 2019 by Tracy Bryan asserts that there is mounting evidence that G-quadruplexes can act as 'potent barriers to DNA replication machinery' [41] and many other papers conclude similarly.[39,42] A good example of this is the presence of a G4 rich regions within the promoter of the MYC oncogene.[43] These G4 structures are known as C-MYC quadruplexes. Over 70% of cancers cases in the US have identified the overexpression of MYC within the individuals and there are strong links between increased MYC levels and cancer metastasis. Therefore, it is logical to assume that targeting and stabilising these C-MYC quadruplexes located upstream of transcription

initiation site would induce anti-cancer activity through the down-regulation of the MYC oncogene.[43]

However, the alternative theory for how stabilised G-quadruplexes induce anti-cancer activity is also widely supported and rests on the fact that telomeric G-quadruplexes can inhibit the production of the enzyme telomerase.[28,44] These G4 quadruplexes are known as H-telo quadruplexes. Telomerase is an enzyme that is responsible for catalysing the elongation of the 3' end of telomeric DNA. However, it has been identified that the over-expression of this enzyme is strongly connected to oncogenesis.[28,45] Therefore, stabilising the formation of H-telo quadruplexes in telomeric structures could induce anti-cancer activity through the consequential inhibition of telomerase.[28,46,47]

Both hypotheses presented for how the stabilisation of G-quadruplexes results in anti-cancer action have evidential support from the literature, with a number of resources quoting both as potential mechanisms. There is insufficient evidence for either theory to promote it over the other, however, it could be inferred that both theories hold enough weight to suggest that the stabilising of G-quadruplexes can bring about anti-cancer activity in a number of different ways simultaneously. This would make the G-quadruplex a more appealing target for anti-cancer drugs (over other noncanonical DNA structures) as dually induced stabilisation could be deemed more cytotoxic to cancer cells.

1.3) Metallo-supramolecular structures as anti-cancer agents

1.3.1) History of Metallo-Supramolecular Helicates

In order to fully comprehend the structure and action of the current array of supramolecular helicates, one must return to origins of this type of chemistry. In 1987, Lehn first proposed the term 'helicate' to describe a helical double-stranded, poly-metallic complex which he had syntheised.[48] It is from this work that we define supramolecular helicates as complexes where a minimum of two metal ions define an axis from which one or more organic ligands (coordinated to the ions) wrap around.[48] However, while helicates were first defined in the late 1980s, their precursors were first synthesised much earlier on. In the 1960s, Busch published numerous papers documenting the synthesis of nickel and cobalt complexes with a variety of different organic ligands.[49,50] The organic ligands used by Busch were often formed from Schiff base amines and mercaptoamines due to their ability to easily coordinate to many different metal centres and these types of ligands are still routinely used to synthesise supramolecular structures because of this property. Although Busch did not manage to synthesise a helicate as traditional defined, he did report the synthesis of helix-like metal complexes and overall, his work laid the foundation for metallo-supramolecular chemistry.[51]

20 years later, the first helical structures were synthesised by Lehn. These helicates were all formed of two bipyridine based ligands with two to three metal ions contained within the structures.[52] Lehn identifies specifically using bipyridine units due to the favourable ability of this unit to interact with the majority of metal ions, therefore, it was likely analogues of this unit would also possess the same favourable properties.[52] Moreover, a

dimeric conformation was selectively targeted by Lehn through optimising the properties of both the ligand and the metal. Lehn references synthesising ligands with 'bridges' of a certain length between the coordination sites. He further stated that these 'bridges' were short enough to prevent a single Cu(I) metal centre from binding in a tetrahedral fashion to both coordination sites within a ligand but long enough to ensure dimeric coordination that was strain free.[52] This concept of a bridge between the coordination sites of a ligand is now more commonly known as a spacer. Lehn's idea of fine-tuning the length and shape of the spacer in order to optimise helicate structure is a platform for the variation of spacers in this thesis.

While Lehn's work has been foundational to the field of helicates, he initially remained focussed on the synthesis of dimeric helical structures.[53] It was not until a full year later, that the first triple-stranded helicates were synthesised by Elliott in 1988. These complexes were also synthesised using bipyridine based ligands, however, Fe(II) was used over Cu(I). Elliott realised the potential of using Fe(II) to force triple-stranded helicate formation due to the metal's d^6 electron configuration. Accordingly, the metal favours an octahedral geometry, the most energetically stable configuration for Fe, as all d orbitals are fully occupied. To achieve this d^{10} geometry, an Fe(II) atom must coordinate 6 electron donor ligands. Since bipyridine-based ligands are bidentate ligands, meaning they have two donor sites, when mixed in a 2:3 ratio of metal to ligand, the most electronically stable product is the triple-stranded helicate as both metal centres are electronically saturated by 1 donor site from each ligand. It should also be noted that in Elliott's synthesis, the bidentate ligands are displacing monodentate ligands, an entropically favourable process due to the chelate effect. Therefore, Elliott perceived that

the formation of these triple-stranded helicates had both electronic and entropic drivers and consequently, he leveraged these drivers to successfully synthesise the first triple-stranded helicates. Moreover, he reported that while his research focussed mainly on complexes of Fe(II), other metals, such as Co(II), Mn(II) and Ni(II), which also favour an octahedral geometry would also be capable of forming triple stranded helicates.

The successful synthesis of self-assembling triple-stranded helical structures created a new and exciting branch of supramolecular chemistry that was un-explored. However, little progress was made in the field over next decade, largely due time-consuming nature of the syntheses of heliactes and the expense of the starting materials.[17] The biggest contribution during this period was in 1993 when Lehn reported the successful synthesis of a trinuclear triple-stranded helix.[54] Like Elliott, Lehn recognised the importance of leveraging the natural electronic and entropic contributions of metal ions that favour octahedral geometry and bidentate ligands to drive his synthesis, although Lehn favoured using Ni(II) over Fe(II). A major turning point came in 1997, when the Hannon group then at the University of Warwick developed a simplistic and inexpensive method of synthesising helical supramolecular architectures.[17] The group designed a novel ligand with pyridylimine binding sites instead of traditionally used bis-pyridine based units.[17] The decision to use pyridylimine units was based on the facile nature of synthesising imine ligands from simple iterative reactions as well as the vast array of aldehydes and amines that are cheaply available from which imines can be synthesised.[17] The group went on to report the successful synthesis of a variety of triple helices from both Ni(II) and Fe(II) salts. These helicates had the general formula $[M_2L_3][X]_4$ where X tended be counterion from the metal salt and were later termed the ‘parent cylinders’.

The final historical discovery fundamental to the research is thesis was also reported by Hannon group. After successfully synthesising the parent cylinders discussed above, the group redirected its focus onto ascertaining the applications of these novel cylinders. In 2001, the group confirmed selectively binding their parent cylinder into the major groove in B-DNA resulting in bending and intramolecular coiling of the DNA.[55] The group used molecular modelling to predict that these cylinders were the too large to bind the minor groove in B-DNA but the right size to bind the major groove and then went on to successfully prove a binding event through DNA binding studies such as Linear-Flow Dichroism and Circular Dichroism.[55] The ability of these cylinders to recognise and bind to the major groove in B-DNA evidenced their potential to have far-reaching therapeutic applications and stimulated a renewed interest in the field. More recently, research has concentrated on the anti-cancer activity metallo-supramolecular helicates may possess and this is the fundamental aspect of the research in this thesis.

1.3.2) How metallo-supramolecular helicates bind to DNA

As discussed, metallo-supramolecular helicates, as shown in Figure 1.3.2.1, can exist in many different forms with a varying number of strands. The ligands present within these diverse architectures tend to be bis-pyridylimine or bis-pyridylazo based structures that when coupled with metal ions help to the give the cylinders their unique binding properties.[6,56] For example, the positively charged metal ions positioned at each end of the cylinder cause the cylinder itself to possess a high positive charge which induces an electrostatic driving force during the binding of the helicate to the DNA (due to the negatively charged phosphate present in the DNA backbone).[6] Moreover, the large

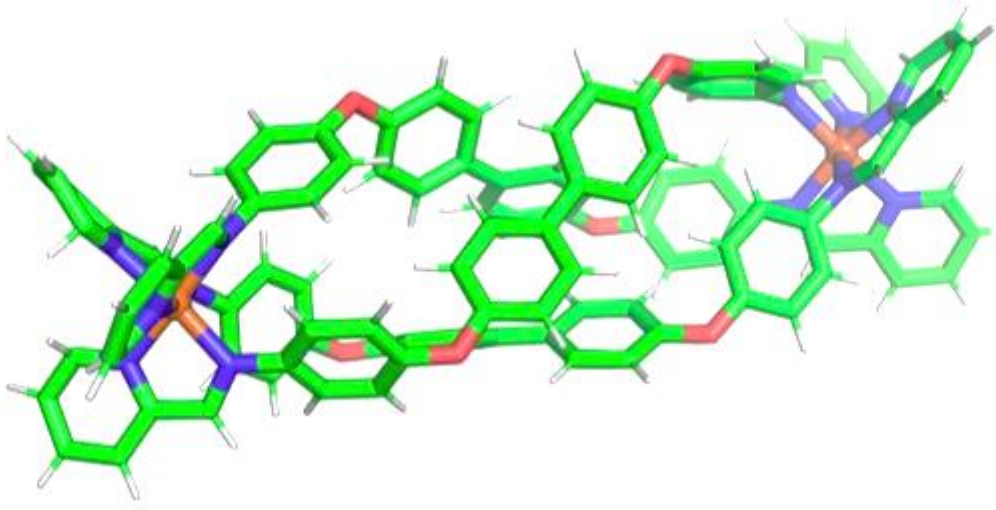


Figure 1.3.2.1 Structure of a metallo-supramolecular helicate known as parent cylinder $[Fe_2L_3]^{4+}$ (where L is $C_{25}H_{21}N_2$). Visualised by PyMol.

number of aromatic rings that tend to be present in the ligands causes the exterior of the cylinder to be hydrophobic. [6]The hydrophobicity of the helicate has an attractive interaction with the hydrophobic cavities within 3WJs and 4WJs further driving the cylinder to bind these noncanonical DNA structures. [6]

It is important to note that although G-quadruplexes do not have hydrophobic cavities to drive a binding event with a helicate, the helicates themselves are chiral molecules and can exist in either *M* or *P* enantiomers, where *M* represents the left-handed helix and *P* the right handed helix.[34,38,57] As previously discussed, chiral molecules are able to interact and stabilise G-quadruplexes, thus, there is an energetic driving force for a binding event between G-quadruplexes and these helicates.[34] Furthermore, the phenyl rings within the ligands of the cylinders can form pi-stacking interactions with the adenine and thymine bases within the DNA[6], providing another driving force for a binding event between these helicates and G-quadruplexes. It should be noted that this particular stabilising interaction is integral to the cylinder binding to 3WJs and 4WJs as well.

Considering that these types of helicates can form two separate enantiomers, it could be anticipated that the noncanonical DNA structures could show a binding preference to one enantiomer over the other.[57] In fact, numerous papers within the research report chiral discrimination by these DNA structures. Many research groups including the Hannon group have found that for both 3WJs, their respective triple-stranded helicate in the *M* chiral form have a greater stabilising effect on the DNA junctions than the corresponding *P* form, which means that the *M* enantiomer binds in preference to the *P* enantiomer and thus will cause greater anti-cancer effects within a cell.[57–59] Interestingly, the Scott group also found that for their triple-stranded flexicates the *M* enantiomer had a greater affinity for interacting with higher-order structures of DNA and therefore, also had greater anti-cancer activity than the corresponding *P* enantiomer.[15] However, when these metallo-supramolecular helices interact with G4s the results are more diverse. The Scott group has found that both enantiomers their flexicate are show strong affinity for G4 binding

and stabilisation.[60] This is the opposite for helicates, with two other papers that state for the parent cylinder, the *P* enantiomer has a greater stabilising effect than the *M* and moreover, *P* enantiomer selectively targets G-quadruplex DNA over duplex DNA.[34,38] The differences between these compounds evidences that binding preference differs from molecule to molecule and that there is no clear rule that can be applied to all supramolecular structures when binding to DNA. What is evident is the clear theme that within literature there are gaps in regard to G-quadruplexes and how they interact with metallo-supramolecular structures. Therefore, there is a distinct need for further research into this field to resolve these gaps.

1.3.3) Helicates and similar molecules that show Biological Activity

Since Lehn's work in 1987 first classifying the term 'helicate' significant progress has been made in the supramolecular field to develop a vast array of helicate molecules. More recently, research has focused on developing helicates that display biological activity often driven by the concept of using the helicates as therapeutic agents.

The Crowley is one such group that has been researching the biological activity of helicates. In 2015 the group reported the synthesis of two bioactive cylinders, one that displayed anti-microbial activity and the other anti-cancer activity, as shown in Figure 1.3.3.1.[61,62] The anti-microbial helicate is a ruthenium based, triple-stranded cylinder with the formula $[Ru_2L_3]^{4+}$.[61] The ligand used was a bis-bidentate ‘‘click’’ pyridyl-1,2,3-triazole and had similar dimensions to the pyridylimine based ligands in the iron parent cylinder, which suggests helicates formed from this ligand could have the ability to interact with higher-order DNA sturctures.[61] When tested in vitro for anti-microbial activity against both Gram positive (*S. aureus*) and Gram negative (*E.coli*) bacteria strains, the cylinder showed a small amount of activity against both. However, the minimum inhibitory concentrations (MIC) were not significant to suggest that this cylinder could be used effectively as an anti-bacterial agent.[61] The group have hypothesised that by making the cylinder more hydrophobic, and therefore more lipophilic, they may be able to improve the MIC of the cylinder and have begun research in this direction.[61]

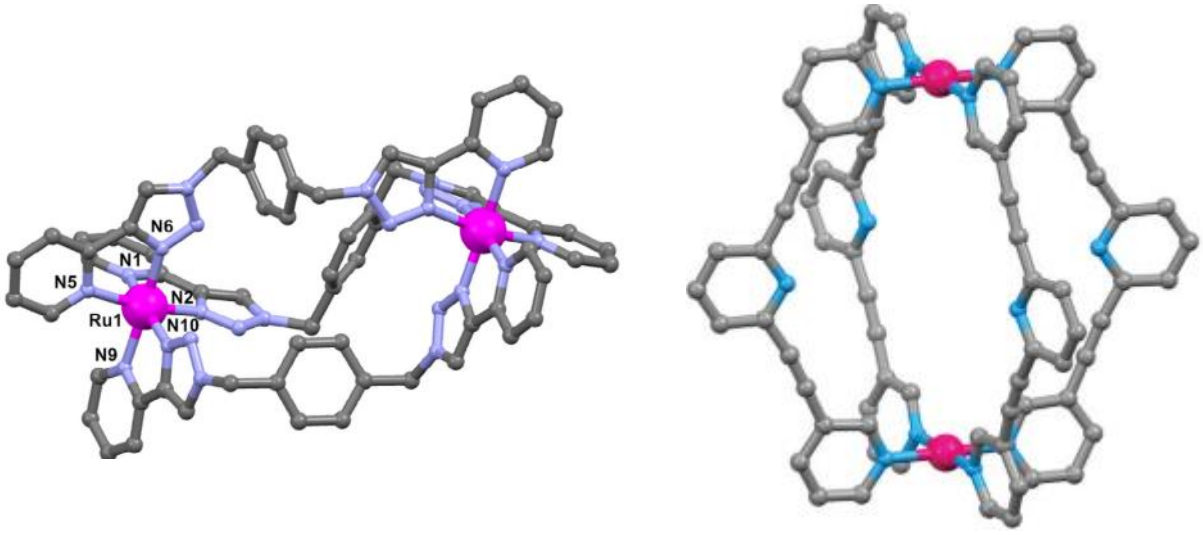


Figure 1.3.3.1 Left-hand image shows structure of Crowley's triple-stranded ruthenium-based helicate. Right-hand image shows Crowley's quadruply stranded palladium-based helicate[61,62]

The other helicate developed by the Crowley group was somewhat more successful in acting as a bioactive agent. Unlike the ruthenium cylinder, this helicate had 4 ligand strands coordinated between two palladium ions.[62] Excitingly, the helicate proved to be highly effective as an anticancer agent, displaying 7 times more activity than cis-platin in the breast-cancer cell line.[62] Furthermore, the cylinder displayed a very unique mechanism by which it induced anti-cancer activity. Rather than interacting with higher order DNA structures to induce cell death, the cylinder induced cell death through compromising cell membrane integrity (as evidenced by both release of intracellular lactate dehydrogenase (LDH) into the extracellular medium and uptake of cell impermeable propidium iodide), indicating that is this instance the helicate is acting as a metallo-detergent.[62] This mode of action has never before been seen for metallo-supramolecular helicates and while the exact mechanism through which membrane damage occurs is unclear, it is apparent that quadruply-stranded helicates present a completely new way forward for inducing anti-cancer activity. Unfortunately, while this palladium helicate clearly shows strong anti-cancer activity, the cylinder does not show any selectivity towards cancerous cells over healthy cells.[62] Therefore, its applications as an effective anti-cancer agent are currently limited due to the large cytotoxic effects that it would have on all cells.

Another group known for researching the biological activity of metal helices is the Scott group. In particular, the Scott group have focused on a slightly different class of helicates termed ‘flexicates’ due to their flexible nature that arises due to the use of non-rigid

ligands.[63] The group decided to concentrate on these flexicates as a result of their favourable chemical properties such as good solubility in water and stability in a wide range of solvents and biological media.[63,64] In 2019, the group reported the development of a new range of triplex metallo-helices, shown in Figure 1.3.3.1, that were able to undergo post-assembly click-modifications.[64]

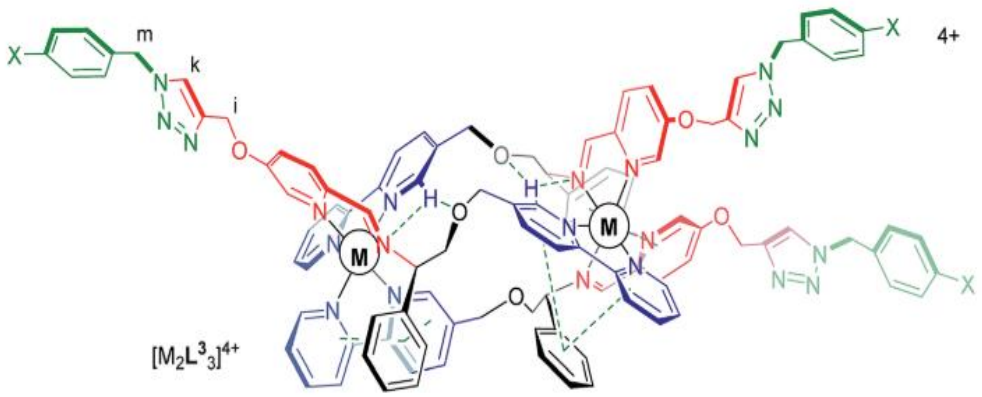


Figure 1.3.3.1 Structure of Scott's flexicates $[M_2L_3]^{4+}$ [63]

Excitingly, these flexicates showed two forms of biological activity, anti-cancer activity as well as anti-metastatic activity. The main flexicate discussed by Scott was an iron-based helix with the formula $[Fe_2L^{3a}{}_3]Cl_4$. Both enantiomers of this clicked helix showed potency against cancer cell lines, with the Δ enantiomer showing greater anti-cancer activity, however, it should be noted that this is the opposite case for the unclicked version.[64] Critically though, the helix Δ-$[Fe_2L^{3a}{}_3]Cl_4$ showed clear selectivity for cancerous cells over health cells.[64] This achievement cannot be understated as not only does it the potential of this helix to act as an effective anti-cancer agent it also goes beyond all current anti-cancer research as well as anti-cancer drugs by being able to target

cancer cells. The ability of this helix to selectively target cancer cells is likely due to electrostatics as the positively charged helix will be attracted to the anionic outer leaflet possessed by cancer cells[64]. However, the group reported that the Δ isomer shows better selectivity than the Λ isomer which suggests that the mechanism for selectivity is defined by more than electrostatics.[64] Another exciting property of Δ-$[Fe_2L^{3a}_3]Cl_4$ helix is that it shows activity against p53 mutated cell line. This cancer cell-line shows resistance to cis-platin, and this further points to the real potential of this helix to act as an anti-cancer agent.[64] Finally, the Scott group reported that Δ-$[Fe_2L^{3a}_3]Cl_4$ showed anti-metastatic action, through its capacity to limit the current functions performed by cancer cells.[64] For example, the presence of the helix diminished the cancer cells' ability to detach, migrate, relocate and reattach to other healthy cells, in turn, preventing metastasis.[64] Overall, the possibilities and applications of this flexicate developed by the Scott group are vast and highlights the potential of flexicates as biological agents.

Helicates represent one class of supramolecular structures that show biological activity, however, there are many classes of similarly formed compounds that also exhibit the same activity. One such example, is the platinum quadrangular boxes developed by the Terenzi group, shown in Figure 1.3.3.3.[65] The platinum boxes have the formula Pt_2L_2 where the ligands (L) are based on 4,4′-bipyridine units of increasing length. These boxes showed the capability to act as anti-cancer agents by interacting with H-telo G4s mainly through stacking at the top ends of the DNA.[65] Moveover, the group saw a general trend, where the larger the platinum box, the better the anti-cancer activity of the respective box.[65] This is because the larger boxes whave a larger affinity for G4 structures, so are more likely to interact with and bind to the G4 structures and therefore, have a better ability to down-

regulate cancer-inducing genes.[65] Interestingly, some of these platinum boxes also showed potency against the VM-1 cancer cell line that is resistant to cisplatin, thus evidencing the potential of these boxes to act as anti-cancer agents. Terenzi does not discuss the selectivity of these boxes for targeting cancerous cells over normal health cells except to say that the selectivity needs to be refined and that is the focus of future research. Nonetheless, these supramolecular structures show strong anti-cancer activity and have good potential as future anti-cancer agents.

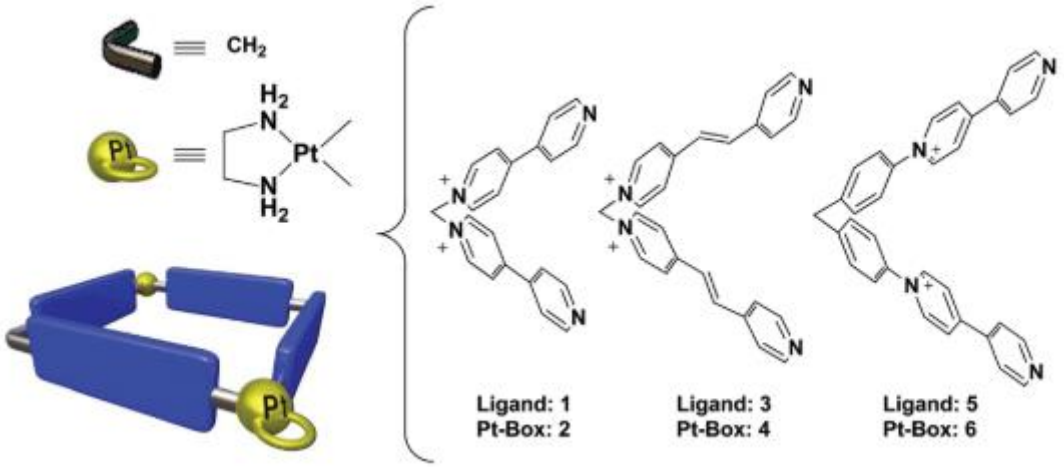


Figure 1.3.3.3 Structure of Terenzi's boxes including the ligands that can be used to form the boxes. The blue rectangles in the box on the left-handside of the image represent backbone of the different L-shaped 4,4'-bipyridine ligands [65].

Many research groups have focussed on developing metallo-supramolecular structures, helicates in particular, due to their capability to exhibit biological activity. The drive behind research is often to develop these structures into biological agents that can treat a range of diseases including cancer. The examples given above highlight the vast progression of the current research to achieving its goal but also the challenges facing it. The most exciting developments have come from the Scott group that rapidly seems to be overcoming the challenges of selective recognition of cancerous cells over healthy

cells, suggesting that helicates and flexiactes could become the next biological agents used to treat cancer.

1.4) The link between metallo-supramolecular helicates and zinc finger proteins

Although different ligands and metals ions can be used to form a variety of these triple-stranded supramolecular helicates, all of these structures tend to have similar dimensions to a specific class of proteins classified as zinc fingers.[38,45,66,67]Zinc fingers are proteins that coordinate at least one zinc ion. The zinc ions stabilise the structure enabling the protein to fulfil its role. These zinc finger protein structures can act as transcriptional factors and therefore, are very important in the regulation of gene expression.[68,69] Both the helicates and zinc fingers contain a cylindrical unit that shares structural similarities to an alpha helix and this means that both can bind to DNA in similar ways.[1,67] This is clearly evidenced by the fact both metallo-supramolecular helicates and zinc fingers can bind to the DNA major groove but are far too big for the minor groove.[1,56,67] However, there are several differences between them both. One key difference resides partly in how each structure interacts with the DNA major groove; helicates span a larger number of base pairs than zinc fingers do as they are slightly longer.[67] Moreover, it is also important to note that while supramolecular helicates bind to DNA through structure-specific recognition, zinc fingers bind through sequence-specific mechanisms.[68,70] Nonetheless, the similarities between zinc fingers and metallo-supramolecular helicates are more significant than the differences, yet within the literature, the similarities are mentioned briefly and are not deeply explored. While the link between them may seem arbitrary,

recent research into the effect of zinc fingers on oncogenesis progression suggests that a better understanding of how these naturally produced proteins interact with DNA in cancer cells could significantly impact the development and research into the anticancer activity produced by supramolecular helicates.

In terms of cancer progression, zinc fingers can have both a positive and negative effect. The over-expression and amplification of certain zinc finger proteins can promote cell migration and prevent apoptosis leading to the formation of tumours while other zinc finger proteins can act in the opposite way, thus having a suppressor effect.[71,72] Fundamentally, it is their control over the transcription of 'downstream genes' that causes these proteins to have either cancerous or anti-cancerous effects.[71] Once again, the relevance of this may seem somewhat arbitrary as zinc fingers bind to DNA through sequence-specific recognition while metallo-supramolecular helicates bind through structure-specific recognition mechanisms. However, it was reported by Boer et al. that protein recognition of DNA is controlled by enthalpic contributions and entropic contributions from both the protein and DNA and while this is mostly reliant on the DNA sequence, the spatial arrangement of the different functional groups can have a large effect in the binding energy and thus structure plays some role in binding.[12] The paper goes on to further conclude that the design of a molecule that can target structure and sequence may be possible and could represent another approach to the creation of DNA-binding drugs.[12] This raises questions as to whether research into the binding mechanism of zinc finger proteins that act as suppressors of cancer could provide vital information into how triple-stranded cylinders could be adapted to make them more selective at targeting cancerous cells. The correct functionalisation of these structures could adapt

them to recognise both structure and sequence and therefore could enable them to have much more effective anti-cancer activity.

1.5) Future outlook for metallo-supramolecular compounds

Over the last few years, the field of metallo-supramolecular chemistry has seen significant development and yet, there is still much further research needed before viable anti-cancer agents can be produced from these compounds. The first major issue that needs to be overcome is how the compounds can be delivered to the cancerous cells safely without having an effect on normal cells.[12] This concept is vital as it prevents the drug from causing widely toxic effects across a system. Another major issue arises in connection to the former; how can the drug now be activated once it is within the target cell to bring about the desired result – in this case, cancer cell apoptosis. Within the literature a few methods to combat these issues have been suggested, the most recent and interesting being rotaxantion of the metallo-supramolecular cylinders.[73]

1.5.1) Rotaxanation of metallo-supramolecular cylinders

The utilisation of host-guest chemistry for safe delivery of anti-cancer drugs to target cells has been a highly investigated concept since it was first suggested in 1987.[74,75] The theory suggests that hosts, often macrocyclic molecules, can bind anti-cancer drugs as guests into their cavities to create a host-guest assembly.[74] The interactions between the host and guest are non-covalent, namely, hydrophobic, electrostatic and hydrogen-bonding

interactions, which means that although the guest is firmly bonded within the macrocycle, only a small amount of energy is needed for the host-guest structure to disassemble.[74]

Furthermore, the host-guest system is soluble due to an external property of the macrocycles, which enables them to interact with solvent molecules and thus, lends them ability to travel through bioactive systems.[76] Therefore, if anti-cancer drugs could be bound within the hydrophobic cavity of a host, there is the potential that this arrangement could act as a safe delivery system.

Rotaxanes are mechanically interlocked molecules that take advantage of the host-guest chemistry. They are created by a linear guest being threaded through a macrocyclic ring, like a thread through a needle eye.[77] Stoppers exist on the ends of the linear molecule which are larger than the macrocyclic ring and these prevent de-threading, the process by which the ring comes off the linear molecule.[77] The literature suggests that by utilising the metallo-supramoleuclar cylinders as the linear molecule, a macrocyclic molecule can be threaded onto them to create this rotaxane. The macrocycle would need to have the appropriate proportions to enable the large cylinder structures to thread through and the literature suggests that cucurbit[10]uril, as shown in Figure 1.5.1.1, would have a cavity large enough to enable threading.[78,79]

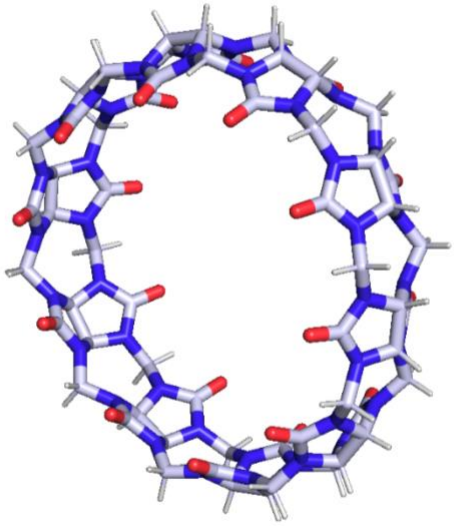


Figure 1.5.1.1 Structure of the macrocycle, cucurbit[10]uril. Visualised in PyMol.

A recent paper by C. Hooper et al. showed that they had successfully threaded supramolecular helicates through these cucurbit[10]uril macrocycle rings to create a rotaxane like structure.[73] To prevent the macrocycle from de-rotaxanating from the cylinder, the paper suggests that the ends of the cylinder can be functionalised to create branches, as shown in Figure 1.5.1.2.[73] The more branched the ends, the more difficult for the macrocycle to de-thread. Furthermore, the paper suggests that if the cylinder is functionalised with incomplete branched points, the macrocycle can de-thread in the presence of a competitor guest.[73] This concept has the potential to resolve both previously stated issues that affect anti-cancer drugs as with the cucurbit[10]uril ring the cylinders can be safely delivered to the target site and also in the correct conditions can be activated at these sites. While very promising, this mechanism of action still has issues it would need to overcome, namely ensuring that there is a suitable guest present within the target cells that can outcompete the cylinder for the cavity of the macrocycles and thus enable the anti-cancer drug to be activated.

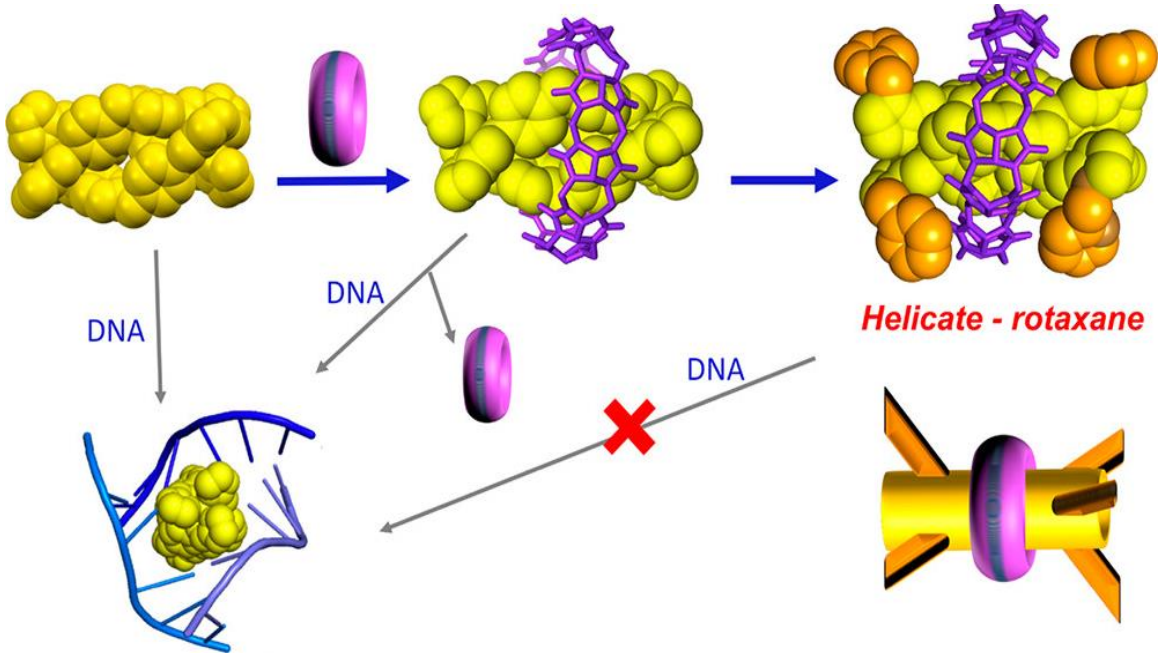


Figure 1.5.1.2. Infographic detailing how rotaxantion of a helicate (that is then modified to prevent the ring from de-threading) prevents it from intercalating between DNA bases.

A much less explored option that may solve this issue is the use of a self-opening macrocycle.[80] A β-galactosidase-responsive macrocycle with enzymatic (galactoside) trigger would enable the cylinder to be protected in the blood stream, however, once inside cells intracellular β-galactosidase would activate the release of the cylinder through triggering a series of reactions that cause the macrocycle to open.[80] This mechanism of action overcomes the issues presented by competition-based rotaxanation strategy as well as issues related to delivery and activation that affect most anti-cancer drugs. Moreover, since most cancer cells over-express the enzymes β-galactosidase, this method is also intrinsically selective to cancer cells.[80–83] The main issue with using this macrocycle is whether the helicate systems would fit within its cavity. The advantage of the curbit[10]uril macrocycle is that it has a large enough cavity in which a helicate system could reside, however, the β-galactosidase-responsive macrocycle suggested in R. Barat et al. is much smaller so it is unlikely it could be used as a host for supramolecular

structures. That being said, the self-opening macrocycle warrants further research as it has real potential to revolutionise delivery of anti-cancer drugs.

1.6) Research objectives

Overall, while the supramolecular field of chemistry has developed significantly within the last decade, there are still many areas of uncertainty within the field that need further in-depth exploration. A large variety of metallo-supramolecular compounds have been developed that successfully interact with higher-order DNA structures, however, these often suffer with either specificity or solubility issues, limiting their potential application as therapeutic agents. Moreover, few of these compounds display the ability to successfully bind to G4 quadruplexes, despite this DNA structure presenting a real target for potential anti-cancer agents.

This research will detail the development of a novel helicate, building on previous research of the Hannon group, that seeks to overcome some of the challenges faces the current portfolio of metallo-supramolecular compounds. This helicate will have increased length over the traditional parent cylinder developed by the Hannon group and the research will explore how the larger length of the helicate affects its ability to bind higher-order DNA structures.

Chapter 2: SYNTHESIS OF NOVEL CYLINDERS

2.1) Introduction to Designing Novel Cylinders

Chapter 1 provided an overview of the literature that is relevant to this project. Within the chapter a supramolecular helicate was defined as a complex where a minimum of two metal ions define an axis from which one or more organic ligands (coordinated to the ions) could wrap around.[48] Design, and consequently synthesis, of helicates and other supramolecular structures discussed within the literature showed that researchers were able to carefully encode specific physical and chemical properties into these structures through the deliberate selection of the individual components that make up these structures, such as metal atoms, binding units and spacers. Therefore, in order to contextualise fully the synthesis decisions made in this thesis, each component will be discussed.

The first and arguably most critical component of metallo-supramolecular helicates is the metal ion. Transition metals are often used synthesis of the majority of supramolecular helicate structures due to their practical physical and chemical properties.[84] For example, the geometries preferred by transition metals depend on the innate properties of the ion such as charge, size and electronic structure.[48,85] When ligands coordinate to a transition metal ion the d-orbitals of the metal split into non-degenerate orbitals according to the favoured binding geometry of the metal which is driven by the lowest energy arrangement of electrons. The difference in energy between the higher energy and lower energy orbitals is known as crystal field splitting energy. The majority of transition metals favour the splitting of their d-orbitals into an octahedral or tetrahedral arrangement (influencing their corresponding geometry) although some d^8 ions' orbitals will split to a square planar

configuration. This enables the structures of metallo-supramolecular compounds to be predicted.

Spacers represent another key structure within supramolecular helicates. The purpose of the spacer is to ensure that the two coordinating ends of a ligand do not coordinate to the same metal ion but rather to separate ions. Consequently, it is a requirement of spacers to have some rigidity to achieve this.[48,86] On the other hand, spacers must have enough flexibility within their structure to enable the characteristic twist of a helicate.[48,86] Therefore, suitable spacers for helicates need to possess a careful balance of both flexibility and rigidity, thus caution and consideration must be taken when selecting a spacer.

The final key unit within a metallo-supramolecular helicate is the binding unit. Deliberate selection of the binding unit is crucial to ensuring self-assembly of the helicate structure over other less thermodynamically stable structures. In order to guarantee formation of the helicate, the interaction between the metal ion and the coordination groups must be reversible, to allow the system to settle into an equilibrium with the most thermodynamically stable product.[84] This means that the binding units must be capable of forming bonds to from that are both labile and strong to facilitate both reversibility and product formation.[84] Commonly used binding units often involve Schiff bases, such as imines, as they are able to afford the dynamic nature of the coordination bonds.[35] Bi-pyridine and phenanthroline units are also widely used in metallo-supramolecular chemistry while catechol units have been explored by Markus Albrecht for the same reasons – that these structures are able to form coordination bonds. Moreover, binding

units that involve bi-dentate (and sometimes even tri-dentate) groups are popular in helicate synthesis as these groups enable stronger bond formation through the chelate effect.

While consideration of the key units that form a metallo-supramolecular helicate is critical to successful synthesis of such a structure, there are other factors that influence the overall structure of these cylinders. As discussed in chapter one, these helicates can be present in two different chiral forms, the *M* and the *P* enantiomers. In *M* enantiomer the cylinder has a left-handed twist while *P* enantiomer has a right handed twist.[34,38,57] However, it is also important to note that these metallo-structures can exist in a mesocate form instead of the helicate. The mesocate form is caused by an internal helix reversal, resulting in a structure where one where one coordination site shows *M* chirality and the other *P* chirality.[86] This means that the mesocate structure generated is a-chiral and therefore, exists only in one form.[86] Mesocates remain less well explored than helicates, however they are relevant to this thesis as there is evidence to suggest one of the cylinders developed in this project exists as both a helicate and mesocate.

Overall, through careful construction of the ligand and choice of metal ion it is possible to predict the structure of a supramolecular helicate and moreover, encode specific intrinsic chemical properties into the cylinder.

2.2) Aims

Since the synthesis of the parent cylinders in 1997, the Hannon group has focussed its research on developing further analogues of this cylinder as well as other novel supramolecular structures. Additionally, the group has investigated whether these supramolecular compounds bind to duplex DNA and higher order DNA structures with a view to present these structures as potential anti-cancer agents. This chapter will explore this further and seek new understanding by studying new cylinders that are similar to the parent cylinder but modify its structure.

In 2020, the Li group based at Western Sydney University reported the synthesis of a Fe(II) triple helicate that contained a 4,4′-(1,1′-biphenyl-4,4′-diyldioxy)dianiline spacer with thiazole-based binding units.[87] The resulting cylinder was significantly longer than the parent cylinders synthesised by the Hannon group and consequently will have a lower charge density, but also has larger aryl surfaces around its central core. Inspired by this structure, herein the synthesis of related larger cylinders based on pyridylimine and imidazolimine ligands is investigated in this chapter with a view to exploring their DNA-binding which is studied in Chapter 3. During the final stages of writing-up of this work, the Li group also reported synthesis of the pyridylimine cylinder.[88]

2.3) Ligand Synthesis

To synthesise a much larger ligand a suitable spacer needed to be chosen, as the larger sized ligand would cause the resulting helicate to experience a significant intramolecular

strain. Therefore, the ligand would need to have a substantial amount of flexibility in order to reduce the strain exerted on the cylinder and as discussed, flexibility comes from the spacer. Consequently, 4,4′-(1,1′-biphenyl-4,4′-diyldioxy)dianiline was chosen as a spacer as the oxygen linkers within the structure would enable greater conformational freedom in the backbone of the ligand and therefore, an increase in flexibility and a decrease in intramolecular strain.[87]

2.3.1) Pyridine-based ligand synthesis, L^1

The synthesis route of the larger ligand shown in Figure 2.3.1, involves a condensation reaction between 4,4′-(1,1′-biphenyl-4,4′-diyldioxy)dianiline and two pyridine-2-carboxyaldehyes as shown in Figure 2.3.1. This condensation reaction produces two pyridylimine components that will act as binding units to metal centres.

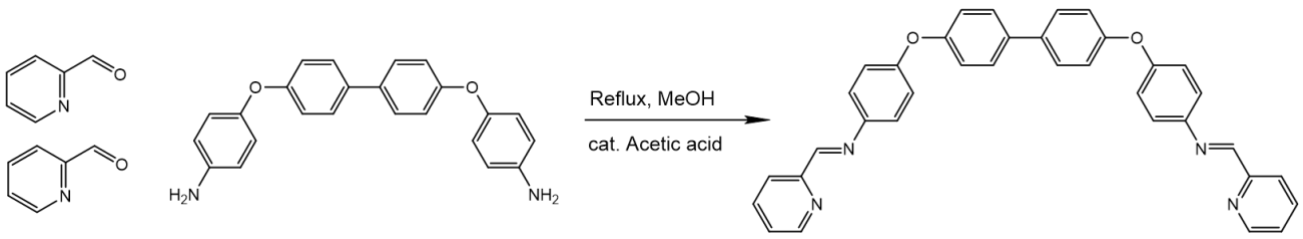


Figure 2.3.1 Synthesis of L^1

In the initial ligand synthesis attempt, similar conditions to the parent ligand synthesis were used. The spacer was dissolved in ethanol and pyridine-2-carboxyaldehyde was added dropwise. The solution was then stirred for two hours at room temperature. It was observed that the spacer showed insolubility in ethanol and while some of the desired ligand was formed from this reaction (NMR showed some potential product peaks), there were also strong peaks relating to the reactants. To overcome the solubility issue of 4,4′-(1,1′-biphenyl-4,4′-diyldioxy)dianiline the reaction was stired at room temperature for 24

hours using methanol as the solvent as within the literature Li reported that this enabled the spacer to successfully dissolve and react with a thiazolealdehyde.[87] Once again, the NMR only showed peaks consistent with the reactants, with the majority of the peaks occurring between 6.5-7.5 ppm. On reviewing the Li group's synthesis further, the reaction was repeated with reflux under inert conditions and a few drops of acetic acid were added to act as a catalyst. These conditions resulted in successful synthesis of the ligand L^1, shown in Figure 2.3.2, at 80% yield and this was confirmed by NMR and mass spectrometry.

The NMR spectrum, shown in Figure 2.3.3, shows the H_3 proton is shifted the furthest at 8.69 ppm and this is expected as this proton is experiencing electron de-shielding from both the pyridine ring and the imine group. The singlet peak at 8.63 ppm is consistent with the expected shift of an imine, therefore is the biggest indicator of successful product formation.

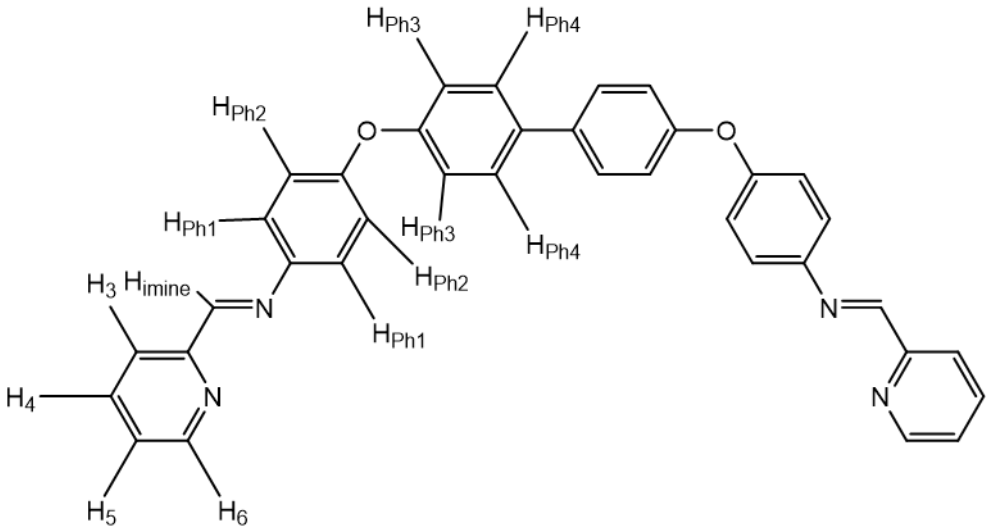


Figure 2.3.2 The structure of the ligand (L^1)

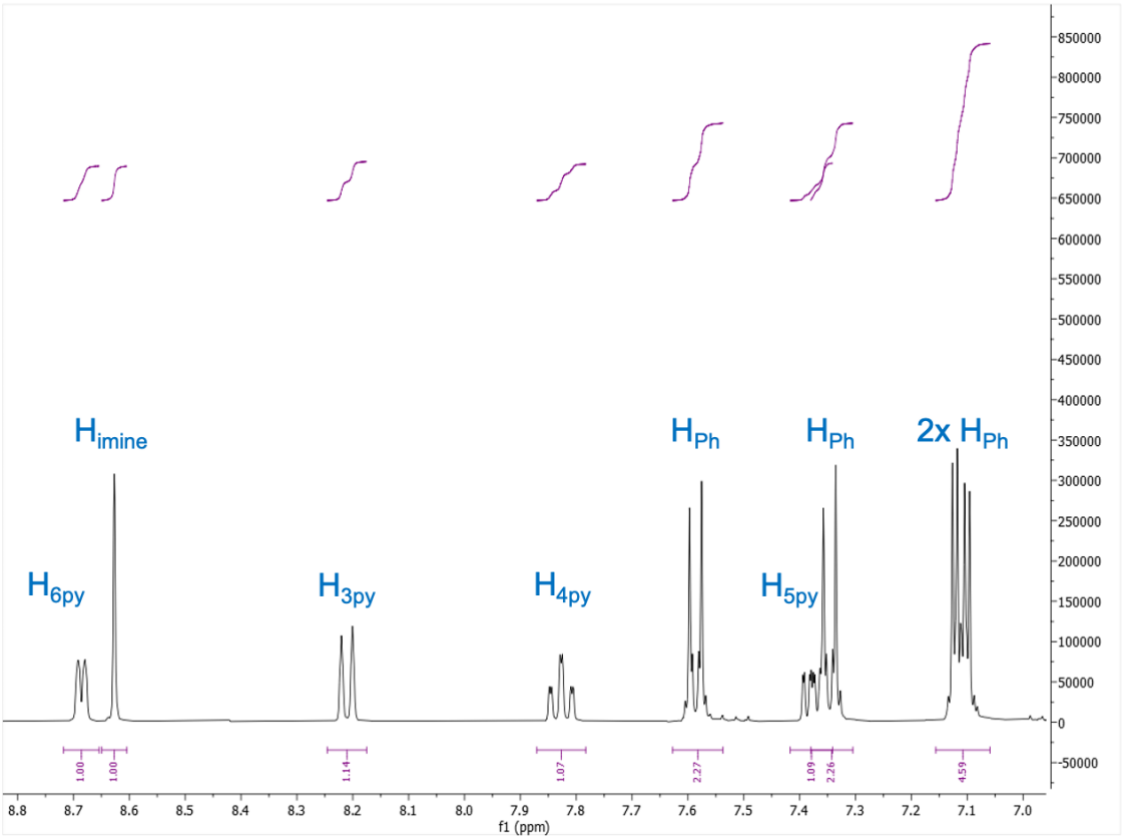


Figure 2.3.3 NMR of L^1 ^1H NMR (400 MHz, CD_2Cl_2) δ 8.69 (br d, J = 4.7 Hz, 1H, H_{6py}), 8.63 (s, 1H, H_{imine}), 8.21 (br d, J = 7.8 Hz, 1H, H_{3py}), 7.83 (td, J = 7.6, 1.7 Hz, 1H, H_{4py}), 7.59 (d, J = 8.6 Hz, 2H, H_{Ph}), 7.38 (ddd, J = 7.5, 4.7, 1.2 Hz, 1H, H_{5py}), 7.35 (d, J = 8.9 Hz, 2H, H_{Ph}), 7.12 (m comprising two overlapping d, J~ 8.8 Hz, 4H, 2x H_{Ph}).

Moreover, the pyridine and spacer protons are present in the correct ratio (with respect to each other). Further indications of successful product formation come from the splitting pattern of the protons on the pyridine ring as they are in agreement with the splitting of the parent cylinder ligand.

Additionally, the mass spectrum shown in Figure 2.3.4, shows a strong peak at *m/z* = 547 $[M+H]^+$ which is in accordance with the expected molecular weight of the ligand. There are no strong peaks at higher molecular weights shown in the spectrum and the peaks shown below *m/z* = 547 are in accordance with the fragmentation of the ligand within the mass spectrum.

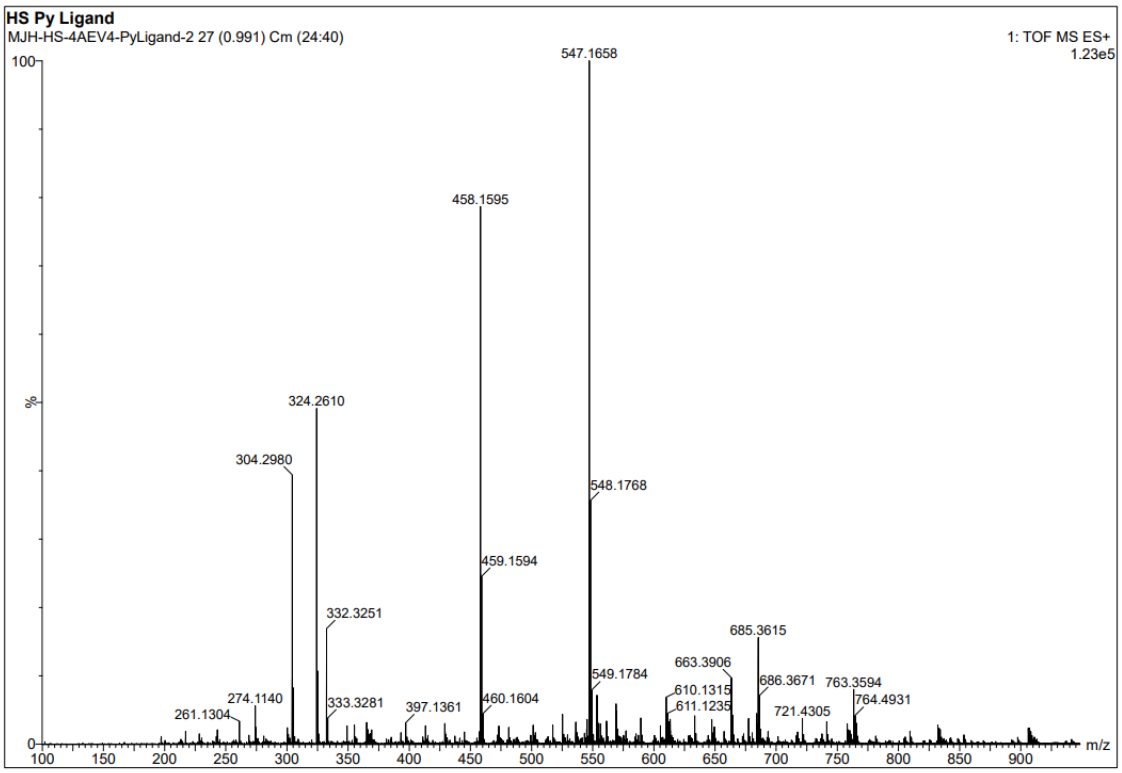


Figure 2.3.4 Mass Spectrum, of L^1 (ESI+): m/z 547 ($[M+H]^+$, 100), m/z 458 ($[C_{30}H_{23}N_3O_2]^+$, 40.

In conclusion, it is clear that synthesis of this longer ligand requires milder conditions and an acid catalyst to drive the reaction to completion. The data showed that these conditions enabled the desired ligand to be synthesised to a high yield and purity. The very recent report of Li uses a similar synthetic approach (but did not use acetic acid).[87]

2.3.2) Imidazole-based ligand, L^2

Similarly, to the first large ligand discussed, the synthesis of this imidazole-based ligand involves a condensation reaction, as shown in figure 2.3.5. The condensation reaction produces two imidazolimine components that are act as binding units to metal centres, and takes place between 4,4′-(1,1′-biphenyl-4,4′-diyldioxy)dianiline and two imidazole-2-carboxaldehydes.

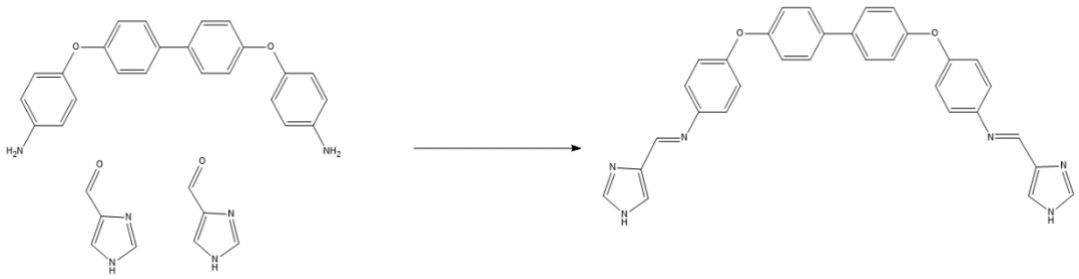


Figure 2.3.5 Synthesis of L^2

The initial ligand synthesis attempt followed the final conditions that resulted in successfully synthesis of the pyridine-based ligand, namely stirring at room temperature under inert conditions for 24 hours with the addition of a couple drops of acetic acid. This attempt was successful at producing the ligand to yield of 42 %, confirmed by the *m/z* = 525 in ESI mass spectrometry, as shown in Figure 2.3.6, which is in accordance with the ligand's molecular weight. While this is a reasonable yield, it does show that this synthesis is much less efficient than synthesis of pyridine-based ligand which formed to 80 % yield under the same conditions.

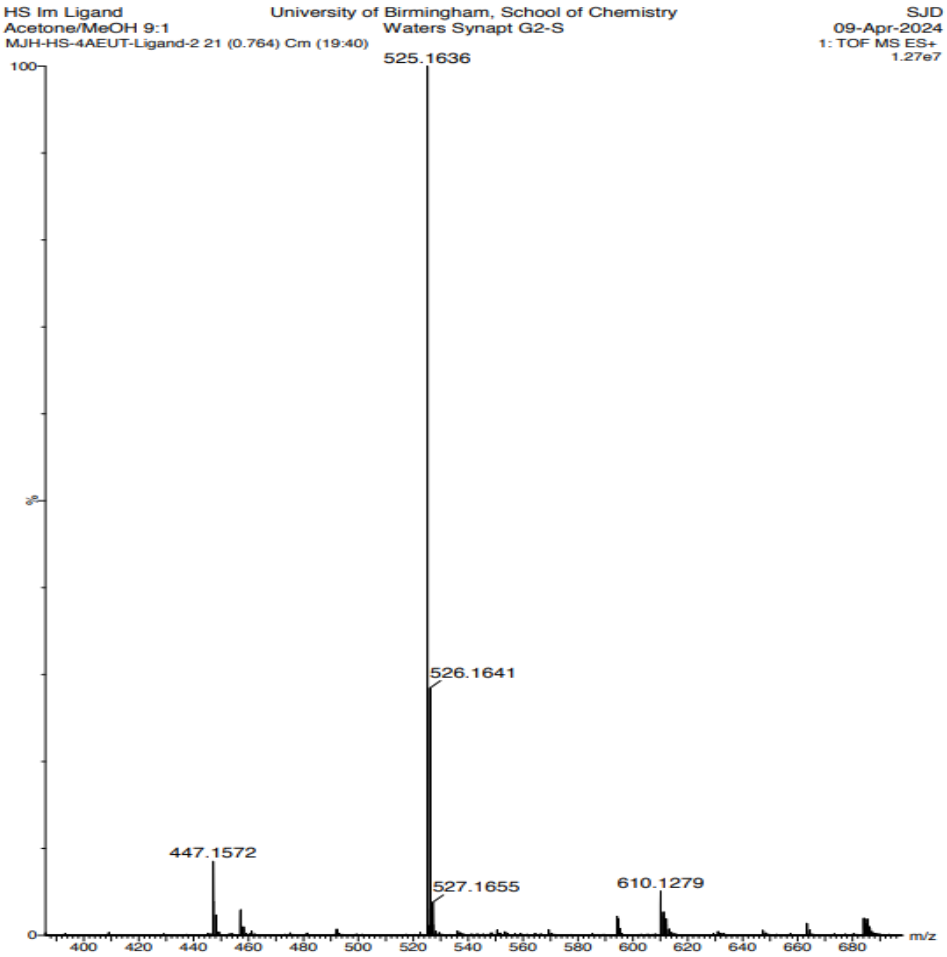


Figure 2.3.6 Mass spectrum of L^2 (ESI^+): m/z 525 ($[M+H]^+$, 100), m/z 447 ($[C_{28}H_{19}N_3O_2+H]^+$, 10).

NMR did confirm synthesis of the imidazole-based ligand L^2, shown in Figure2.3., however, the peaks were much further into the baseline of the spectrum. This was likely due to the insoluble nature of the ligand, which made it difficult to dissolve a concentrated sample. It is likely that the insolubility of the ligand comes from the ability of the imidazole unit to easily form hydrogen bonds in solid state and therefore, remain out of solution.

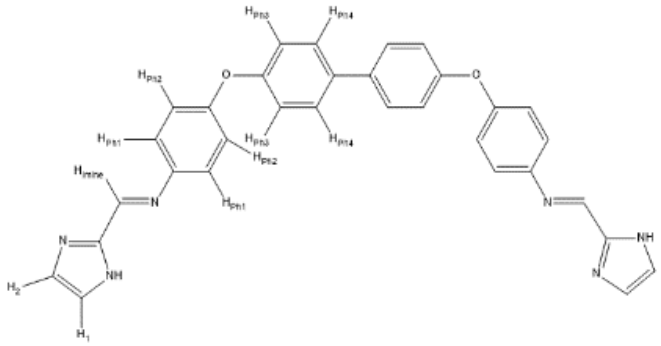


Figure 2.3.7 Structure of L^2

The spectrum generated by NMR, shown in Figure 2.3.8, verifies the desired product formation as both the expected splitting patterns and shifts are seen. For example, both the N-H imidazole proton and imine proton are both connected to strong electron withdrawing groups, so are particularly de-shielded. Therefore, it can be assumed that these protons experience the magnetic field most strongly so are expected to have the highest shifts in the spectrum and this is seen in their shifts of 8.55 ppm and 8.46 ppm respectively. Additionally, the protons on the imidazole rings exhibit the splitting pattern

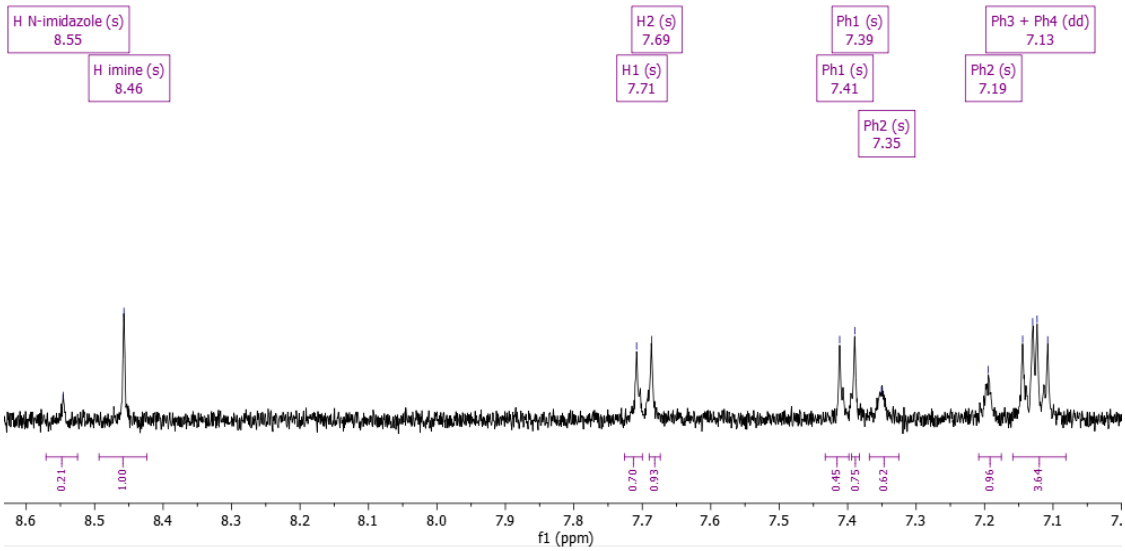


*Figure 2.3.8 NMR of L^2 (400 MHz, DMSO-d) δ 8.55 (s, 2H, NH), 8.46 (s, 2H, H **imine**), 7.71 (s, 2H, H_1), 7.69 (s, 2H, H_2), 7.41 (s, 2H, Ph_1), 7.39 (s, 2H, Ph_1), 7.35 (s, 2H, Ph_2), 7.19 (s, 2H, Ph_2), 7.13 (dd, 8H, Ph_3 + Ph_4).*

expected of an imidazole group as shown by other groups that have synthesised imidazole-based ligands.[89]

Overall, the data showed that mild conditions enabled the imidazole-ligand to be synthesised, however, the reaction was not as efficient as the pyridine-ligand synthesis even though the same conditions were used. NMR confirmed successful of the desired product, however, due to solubility issues related to the ligand, the peaks of the NMR were very weak as it was difficult to generate a concentrated sample.

2.4) Synthesis of triple-stranded helicates

2.4.1) Triple-helicate with pyridine-based ligands, L^1

After synthesis of the pyridine-based ligand to a high purity and yield, design of the helicate synthesis began. The initial decision was made to follow the reaction conditions similar to that of the parent cylinder, due to the similarity of the functional groups between the larger ligand synthesised in this paper and the ligand used in the parent cylinder. Moreover, Fe(II) was selected as for this synthesis due to its preference to form octahedral coordination and this assemble triple-stranded di-nuclear cylinders.

In the first synthesis attempt, the L^1 ligand was refluxed at 70^{o}C in a 3:2 ratio with iron(II) chloride tetrahydrate in methanol for 1 hours. The yellow-coloured mixture was filtered to remove the excess solid present in the solution and the solvent was removed from the product *in vacuo*. The resulting cream solid was obtained in very small amounts (>1%

mass recovery) and further analysis through ESI mass spectrometry showed no peak at *m/z* = 438 which would be expected for the $[Fe_2L^1_3]^{4+}$ ion if the product was successfully formed. Additionally, the mass spectrum showed a strong peak at *m/z* = 547 which corresponds to the mass of ligand, suggesting that no reaction took place. This is further supported by the fact that during the filtration step a large quantity of residue remained on the filter paper and when analysed by mass spectrometry this proved be unreacted ligand. During the synthesis it was observed the ligand had significant insolubility in methanol and that the higher reaction temperature did not improve this. Consequently, the decision was made to use a different solvent to methanol in the next synthesis attempts.

On analysis of the Li group's thiazoleimine helicate synthesis, it was found the Li's larger thiozole-based ligand showed solubility in acetonitrile.[87] Therefore, it was decided to use acetonitrile instead of methanol as a solvent for this synthesis. The added benefit of using acetonitrile in this case is that it is a polar, aprotic solvent meaning it is unable to form hydrogen bonds with nucleophiles yet is still able to dissolve polar organic molecules. Consequently, by using acetonitrile over methanol, the yield of the reaction should increase as the solvent will not compete with the metal centres for the nucleophilic binding sites of the ligand, and therefore, should raise the reactivity of the ligands with the metal ions. However, while iron(II) chloride tetrahydrate is known to be soluble in acetonitrile, in practice the solid struggled to dissolve well in the solvent and so iron(II) tetrafluoroborate hexahydrate was used as the starting material instead.

In this second synthesis attempt using acetonitrile with the same temperature conditions as the first, a dark-purple reaction mixture was produced, which was the first indication the reaction was a success as Fe complexes with pyridylimine usually produce yellow or red coloured compounds. The reaction mixture was filtered to remove any unreacted ligand and the solvent was removed *in vacuo* to give a dark-purple solid. The solid was confirmed by ESI and NMR as product and was yielded at 68 %. ESI mass spectrometry, shown in Figure 2.3.9, showed a strong peak at *m/z* = 438 consistent with the molecular weight and 4+ splitting pattern expected of the $[Fe_2L^1{}_3]^{4+}$ ion.

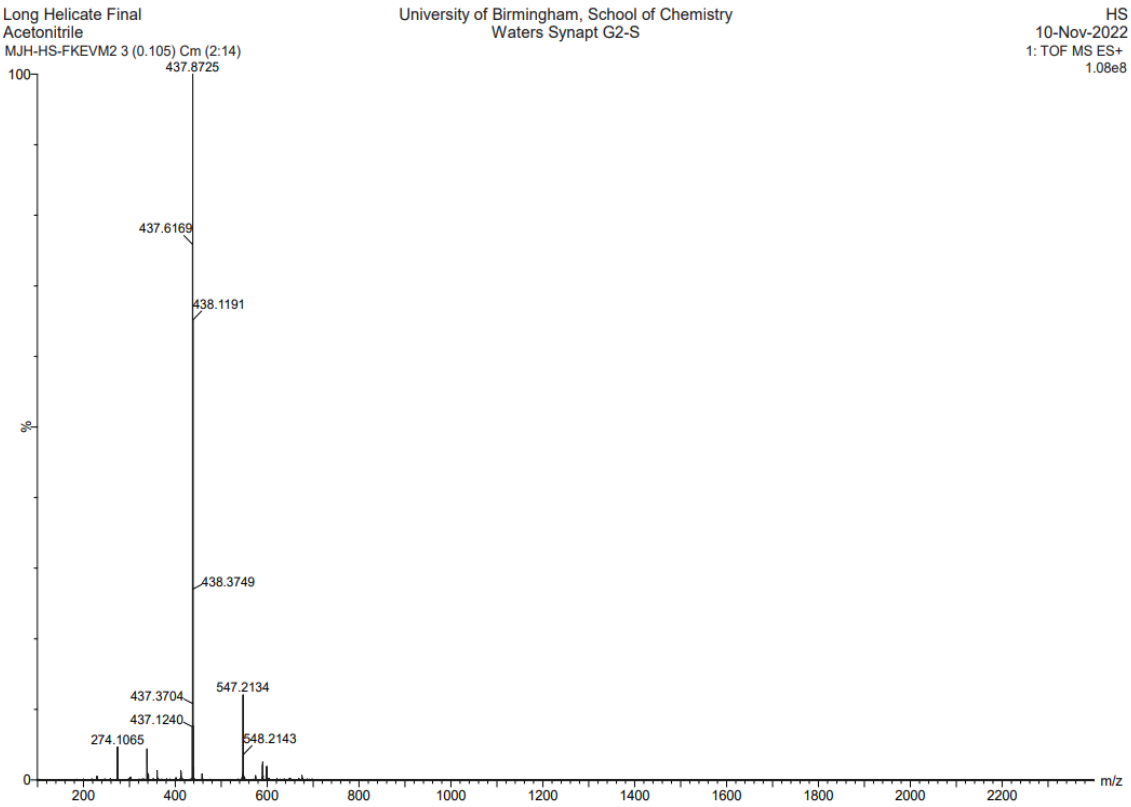


Figure 2.3.9 Mass spectrum $[Fe_2L^1{}_3][BF_4]_4$ (ESI^+): m/z 437 (M^{4+}, 100), m/z 547 ($C_{36}H_{27}O_2N_4^+$, 10), m/z 274 ($C_{18}H_{13}ON_2^+$, 5).

NMR analysis, shown in Figures 2.3.10, 2.3.11 and 2.3.12, was completed to confirm the successful synthesis of the desired helicate $[Fe_2L^1{}_3][BF_4]_4$. Unexpectedly, the spectrum (Figure 2.3.9) showed double the number of peaks that were predicted by the number of proton environments that existed with the desired helicate, with one set of peaks at lower integration than the other. The correct stoichiometry cylinder had been formed (as confirmed by MS), however, two forms were present, the helicate and mesocate, hence double the number of peaks were seen.

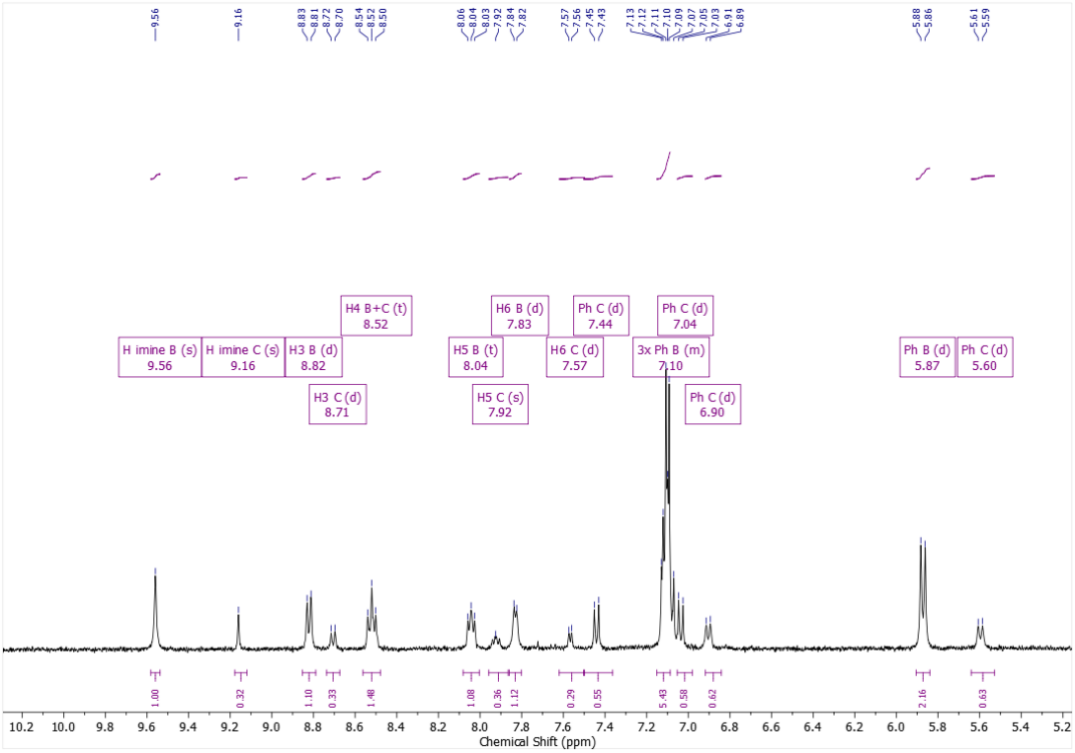


Figure 2.3.10 ^{1}H NMR of $[Fe_2L^1{}_3][BF_4]_4$ (500 MHz, Acetonitrile-d) δ 9.49 (s, 7H, $H_{imine}B$), 9.06 (s, 2H, $H_{imine}C$), 8.71 (d, J = 7.6 Hz, 7H, $H_{3py}B$), 8.58 (d, J = 7.6 Hz, 2H, $H_{3py}C$), 8.38 (m, comprising two overlapping t, J ~ 7.6Hz, 9H, $H_{4py}B$ + $H_{4py}C$), 7.93 (br t, J = 6.6 Hz, 7H, $H_{5py}B$), 7.80 (br t, J = 6.6 Hz, 2H, $H_{5py}C$), 7.74 (br d, J = 5.3 Hz, 7H, $H_{6py}B$), 7.45 (br d, J = 5.5 Hz, 2H, $H_{6py}C$), 7.31 (d, J = 8.6 Hz, 4H, $H_{Ph}C$), 6.98 (m, 42H, 3x $H_{Ph}B$), 6.90 (d, J = 7.6 Hz, 4H, $H_{Ph}C$), 6.77 (br d, J = 8.2 Hz, 4H, $H_{Ph}C$), 5.73 (d, J = 8.3 Hz, 14H, $H_{Ph}B$), 5.46 (br d, J = 8.3 Hz, 4H, $H_{Ph}C$).

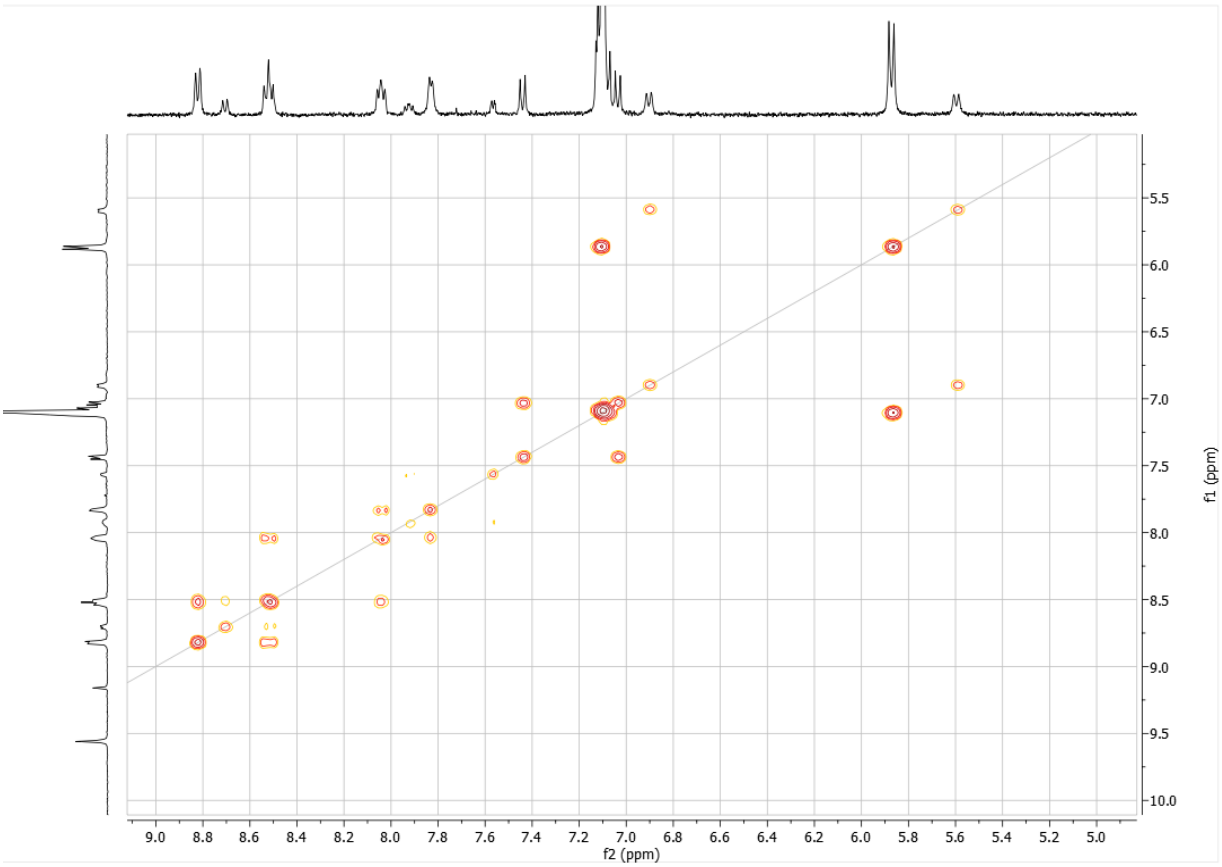


Figure 2.3.11 COSY NMR of $[Fe_2L^1{}_3][BF_4]_4$ (Acetonitrile-d, 500 MHz)

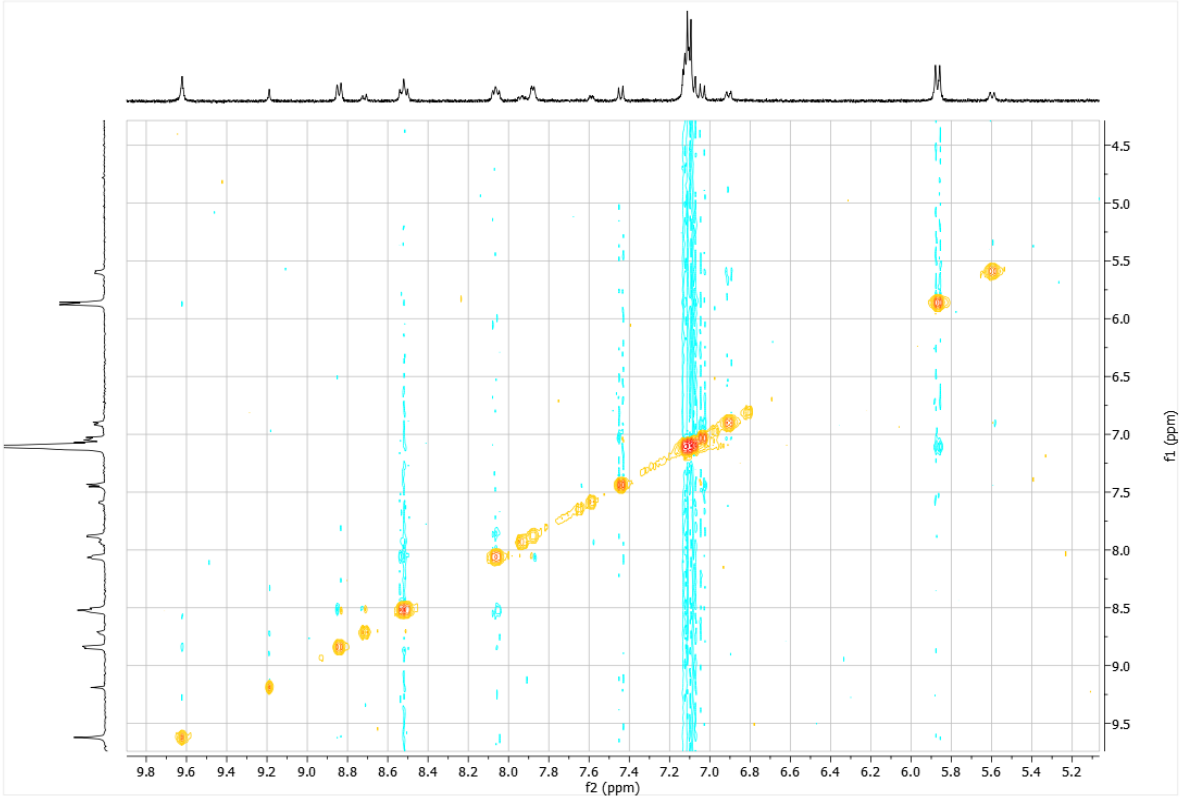


Figure 2.3.12 NOSEY NMR of $[Fe_2L^1{}_3][BF_4]_4$ (Acetonitrile-d, 500 MHz)

As discussed previously in section 2.1, a mesocate is caused by internal helix reversal resulting in an achiral structure where the two coordination sites show different chirality.[86] Within the literature, mesocates have been reported in triple-stranded catechol structures where there are an odd number of methylene groups within the spacer of ligand (though not with the parent ligand).[86,90] However, since the pyridine ligand used in the synthesis of $[Fe_2L^1_3][BF_4]_4$ does not contain methylene units within its spacers, therefore, this theory is not applicable in this case. Nonetheless, since the parent ligand with just one methylene (or oxy/thio) group between pyridiyliminephenyl groups is clearly conformationally more rigid than the more flexible bi-phenyldiol spacer, the integration of the longer spacer in the ligand affords the formation for both isomers.

It is evident from the normal ^{1}H NMR as well as the COSY NMR that two isomeric species, the mesocate and helicate, are forming. The clear 4:1 ratio between the two peaks that relate to the imine proton shifts suggests that the major species is formed 4 times more preferentially to the minor species, however, it is difficult to predict which isomer is the major species as the ligand does not contain a methylene group within its spacer. Despite this, previous studies involving similar structures have shown the mesocate to form in preference to the helicate and it could be speculated that this is also the case for $[Fe_2L^1_3][BF_4]_4$.[91] The majority of the ^{1}H NMR peaks were assigned the either the major (B) or minor (C) species using the cross peaks in the COSY spectrum, however, the full assignment of the major species has been provisionally assigned on the basis of the NOE peak at 5.88 ppm – 9.62 ppm, which shows the imine coupling the phenyl group 1 and the NOE peak at 6.90 – 7.00 ppm which shows coupling between phenyl group 2 and 3.

As a result of the successful synthesis of $[Fe_2L^1_3][BF_4]_4$, the synthesis was repeated using nickel(II) tetrafluoroborate hexahydrate. The reaction conditions were slightly changed; the mixture was refluxed at 55°C for 2 hours and this produced a higher yield of 94 %. Mass spectrometry was used to confirm formation of the desired product. The spectrum showed a strong peak at *m/z* = 439 consistent with the molecular weight and splitting pattern of expected $[Ni_2L^1_3]^{4+}$ ion. Moreover, within the spectrum no other strong peaks were seen suggesting that this helicate/mesocate was formed to a high purity. To confirm this, experimental elemental analysis was performed. The results showed the product contain 59.5 % carbon, 4.02 % hydrogen and 7.63 % nitrogen, which was similar but slightly deviated from the theoretical CHN values (61.64 % carbon, 3.74 % hydrogen and 7.99 % nitrogen). It is likely that any deviation from the theoretical analysis came from the presence of water in the sample as the presence of even a trace amounts of water can cause significant variation in values.[92] Further analysis using the Universities of Illinois's elemental composition calculator showed that $[Ni_2L_3][BF_4]_4.5H_2O$ had the composition of 59.11 % carbon, 4.04 % hydrogen and 7.66 % nitrogen, which is much more consistent with the experimental data, confirming that the intended product had been made, however the sample had not been completely dry when analysed.

To further analyse the helicate/mesocate, NMR spectroscopy was run on the sample. Since $[Ni_2L^1_3][BF_4]_4$ has the d^8 nickel ion in an octahedral geometry with two unpaired electrons, the cylinder is a paramagnetic material, therefore, only paramagnetic NMR can be used to analyse the helicate. Three paramagnetic NMR were run for the helicate, each involving different counterions ($[Ac]^-$, $[BF_4]^-$, $[PF_6]^-$). The spectrums produced, shown in Figure 2.3.13 and 2.3.14, were not extremely clear as the signal peaks generated were

relatively weak, however, this is to be expected for Ni complexes. Despite this, all three spectrum showed a high level of symmetry, with consistent peaks at 15 ppm and 50 ppm confirming the presence of helicate/mesocate.

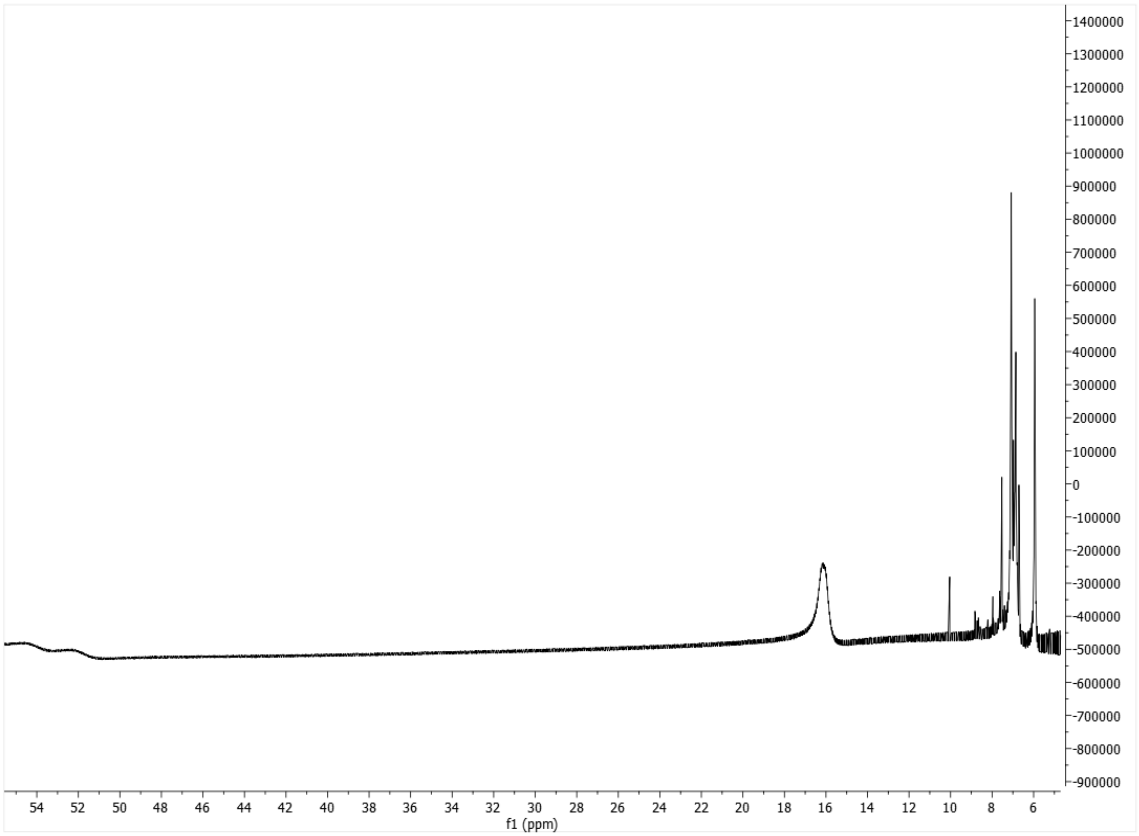


Figure 2.3.13 Paramagnetic NMR $[Ni_2L^1_3][PF_6]_4$ (500 MHz, Acetonitrile-d)

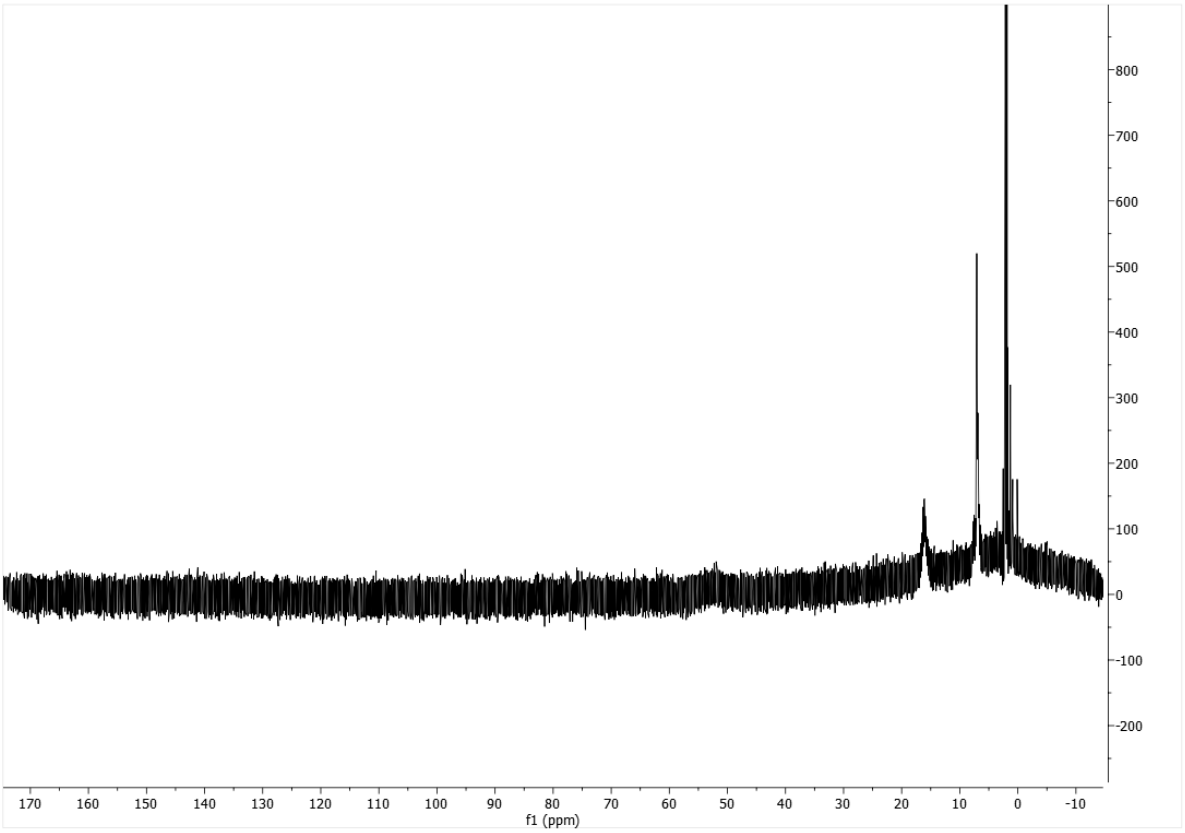


Figure 2.3.14 Paramagnetic NMR $[Ni_2L^1_3][BF_4]_4$ (500 MHz, Acetonitrile-d)

In conclusion, both $[Fe_2L^1_3][BF_4]_4$ and $[Ni_2L^1_3][BF_4]_4$ were successfully synthesised to a high yield. $[Fe_2L^1_3][BF_4]_4$ was confirmed by mass spectrometry and NMR, which further showed the cylinder existed as two species, the mesocate and helicate, in a 4:1 ratio between the major and minor species. Multiple attempts were made to separate the two species through column chromatography, using cellulose as the solid phase and either NaCl (20 mM), Na(Ac) (20 mM) or $NaNO_3$ (20 mM) as the solvent, however each of these attempts were unsuccessful. $[Ni_2L^1_3][BF_4]_4$ was confirmed mainly through mass spectroscopy with the paramagnetic NMR spectrum showing weak peaks and not very informative. The very recent Li paper contains crystal structures of these two cations (as BF4 salts) revealing that the helicate version crystallises, but it is notable that that study

does not study the solution NMR behaviour and does not identify that the second mesocate species can be present.

2.4.2) Triple-helicate with imidazole-based (L^2) ligands

Once the imidazole ligand had been successfully synthesised, the focus changed to synthesising a triple-stranded helicate using this imidazole ligand. The decision was made to follow similar reaction conditions to those with the PY ligand. Consequently, nickel(II) tetrafluoroborate hexahydrate dissolved in acetonitrile was heated at 55^{o}C for three hours with the imidazole based ligand in a 2:3 ratio. This reaction produced the desired product at a yield of 75 %.

Mass spectrometry confirmed formation of the $[Ni_2L^2_3][BF_4]_4$ with a strong peak at *m/z* = 422 which was consistent with both the splitting pattern and molecular weight of the expected $[Ni_2L^1_3]^{4+}$ ion. However, the mass spectrum generated showed other strong peaks that related to charged species. The majority of these peaks related to the desired-triple-stranded helicate, but where one two or all three of the ligands had been deprotonated thus reducing the overall charge of the complex. This is not an unexpected observation as the proton the N in imidazole unit shows acidic character and therefore, can be easily lost during mass spectrometry. However, there were other peaks that instead related to the formation of a double stranded helicate. This observation was unexpected as the ratio of ligand to metal was 3:2, which in theory should favour triple stranded helicate formation. Therefore, it could be assumed that these imidazole-based ligands also support a double-stranded dinuclear helicate formation. Further investigation into the

literature showed that in fact this uncertainty of architecture was a common occurrence with the synthesis of metallo-structures with imidazole-based ligands[93,94] due the ability of the imdazole unit to deprotonate. This deprotonation causes the ligand to become negatively charged, which, in turn, might help support the formation of the double-stranded helicate.

The yield of the desired helicate was moderately lower than the same synthesis using pyridine-based ligands. During filtration step, significant residue was removed from the filtrate and mass spectrometry showed that this solid was unreacted ligand. This could be a result of formation of the double stranded helicate, which in turn would cause a significant amount of ligand to remain unreacted or it could be due to the insoluble nature of imidazole-based ligands, which causes them to remain out of solution and therefore, not react. In order to analyse the helicate further, paramagnetic NMR was run on species. The spectrum (Figure 2.3.15) showed similar peaks to the analogous $[Ni_2L^1_3][BF_4]_4$ cylinder at 6 ppm and 16 ppm, which was consistent with expectations as both species contain the same spacer in the same conformation.

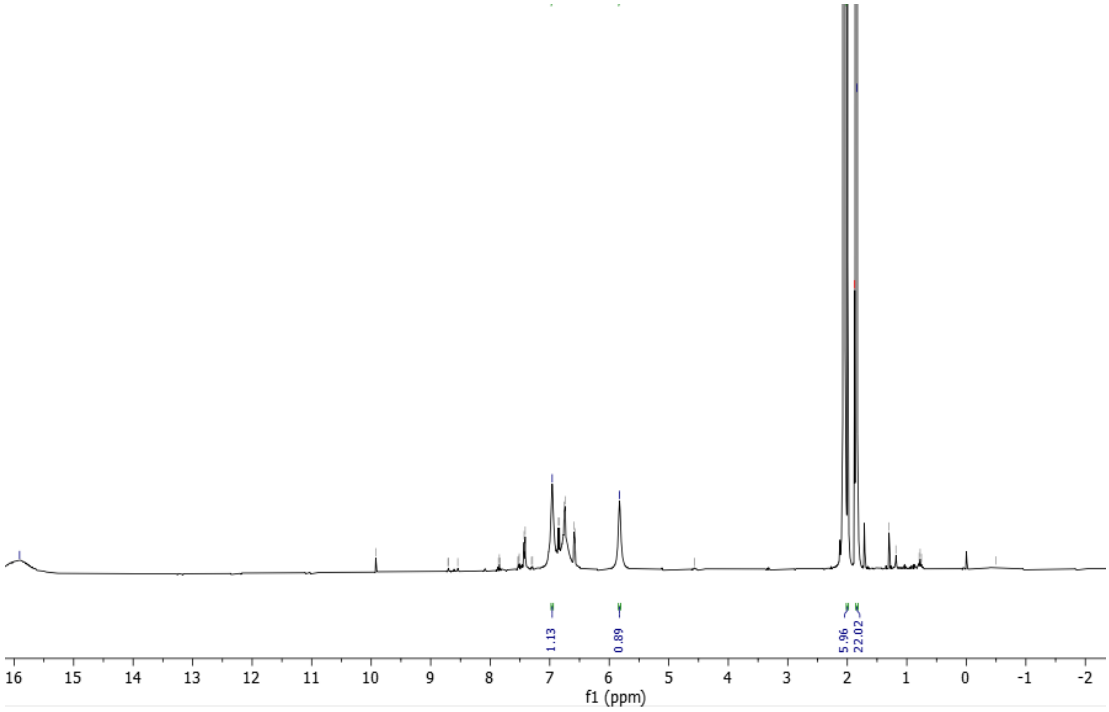


Figure 2.3.15 Paramagnetic NMR spectrum of $[Ni_2L^2_3][BF_4]_4$ (500 MHz, Acetonitrile-d)

In conclusion, $[Ni_2L^2_3][BF_4]_4$ was synthesised using a similar synthesis method to the corresponding L^1 helicate/mesocate and ESI mass spectrometry was used to confirm the synthesis of the desired product. Although the mass spectrum produced did show a peak for the 4+ product ion at *m/z* = 422 other strong peaks relating to charged species were also seen. These peaks related to both the desired triple-stranded helicate that had experienced different levels of deprotonation on the ligand but also the undesired double-stranded helicate species, which could be present due to dissociation of the major species in the mass spectrometer, however, the presence of the unreacted ligand could reflect the formation a real solution species. The variation in helicate assembly was found to be a common event in the literature and related to the ability of the imidazole unit to deprotonate.

2.5) Explorations of Spacer Modification

The helicates synthesised from both the pyridine-based ligands and imidazole-based ligands contain oxygen atoms as linkers between the aryl groups. As will be seen in chapter 3, these helicates displayed instability when in water but were stable when in acetonitrile. Moreover, the helicates dissolved easily in nitromethane but not in methanol and when in solution with methanol, quickly lost their colour. From this data it was deduced that these cylinder favour polar aprotic solvents over polar protic solvents. This is due to the fact that polar protic solvents are able to form hydrogen bonds so will attack the nucleophilic binding sites in the cylinder and compete with the metal centre for ligand binding. The use of oxygen as a linker within the spacer increases the instability of the spacer due to its electron-withdrawing nature.

Oxygen is a highly electro-negative atom due to the high-positive charge of the atom's nucleus and the small atomic radius. Therefore, oxygen has the ability to pull electron-density towards itself and in both the pyridine-based and imidazole-based ligands, the oxygen linkers pull this electron density away from the binding units. In turn, this reduces the electro-negativity of these binding units, making them less effective nucleophiles and consequently, the bonds formed between them and the metal centres are weaker. The Hannon group explored how changing the groups within the linker effected stability and found that oxygen < CH_2 < sulphur for thermal stability. This result is expected as although sulphur has a higher nuclear charge than oxygen, its atomic radius is much larger and therefore, it has a lower electronegativity.

In order to explore and improve the stability of the $[Ni_2L^1_3][BF_4]_4$ helicate previously synthesised, attempts were made to synthesise the spacers with sulphur in the linker as opposed to oxygen. The aim was to then react these spacers with pyridine-carboxaldehyde to create the same pyridine-based ligands with a modified spacer and then further react these ligands with metal ions to form metallo-supramolecular helicates. Ultimately this was unsuccessful but the approaches used are summarised below:

2.5.1) Attempts to synthesise a sulphur-linked ligand

The first step of synthesising the desired ligand with a sulphur linker instead of an oxygen linker required synthesis of a spacer with sulphur linkers as shown in Figure 2.3.16. The initial synthesis proposal was to react 4-fluoronitrobenzene in a 2:1 ratio with biphenyl-4,4'-dithiol to produce the dinitro compound shown below.

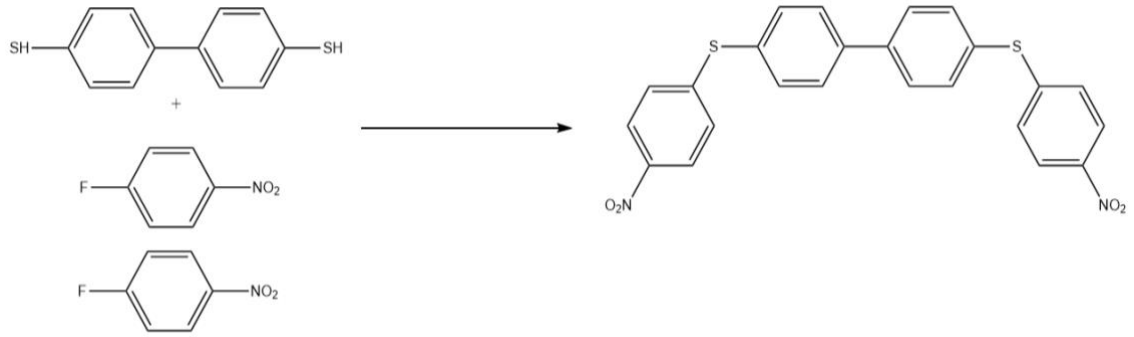


Figure 2.3.16 Synthesis of spacer for sulphur-based ligand

Investigation into the literature showed that few reactions of this type with thiols have been done, however, a thesis produced by Kirti Patel detailed a similar reaction taking place with a phenol.[95] In this reaction, which was completed successfully, Patel used DMSO as a solvent with NaOH as a base. Patel's method also involved mildly heating

the reaction to approximately 55ºC. The initial synthesis attempt involved reacting 4-fluoronitrobenzene in a 2:1 ratio with biphenyl-4,4'-dithiol in DMSO and NaOH at 55ºC for two hours. It was observed that on addition of 4-fluoronitrobenzene the yellow solution turned a dark orange/red suggesting a reaction was taking place, however, after two hours being heated the solution turned back to a yellow colour. This was the first indication that there were issues with this synthesis. The challenges associated with using DMSO as a solvent then became apparent in the fact that the product was soluble in this solvent. Since DMSO has a high boiling, the solvent could not be removed from the product *in vacuo*. Consequently, a different method had to be used. The solution to this problem was to add excess diethyl ether, centrifuge the mixture and then remove the supernatant. This process was be repeated until a precipitate was formed and isolated by vacuum filtration. Unfortunately, this is not a highly effective method of isolating the product as not only was the yield low but also the product was more of a paste than a powder suggesting it was still relatively wet.

Further analysis on the sulphur-based spacer, shown in Figure 2.3.16, was completed through mass spectrometry and NMR. The mass spectrum (Figure 2.3.17) produced was not clean and showed a large number of peaks that could not be assigned as they showed no relation to the desired product. On top of this the NMR (Figure 2.3.18) showed only solvent peaks and no peaks relating to any organic compound including the reactants. It was unclear from the data whether the reaction was unsuccessful due to the conditions,

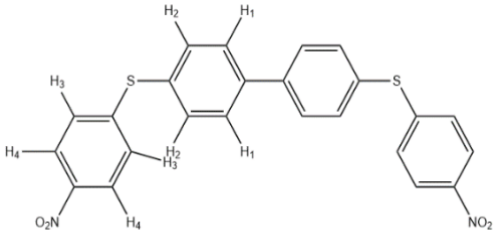


Figure 2.3.16 Structure of precursor for sulphur-based spacer

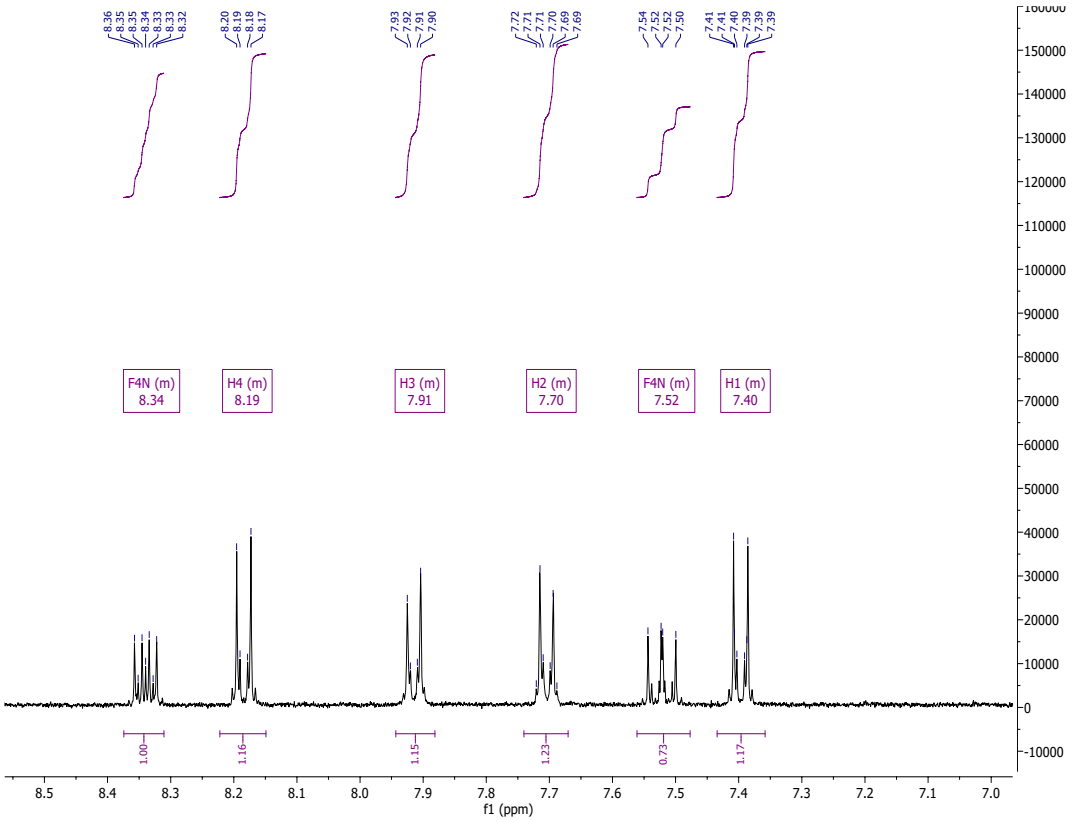


Figure 2.3.17 Mass spectrum of sulphur spacer

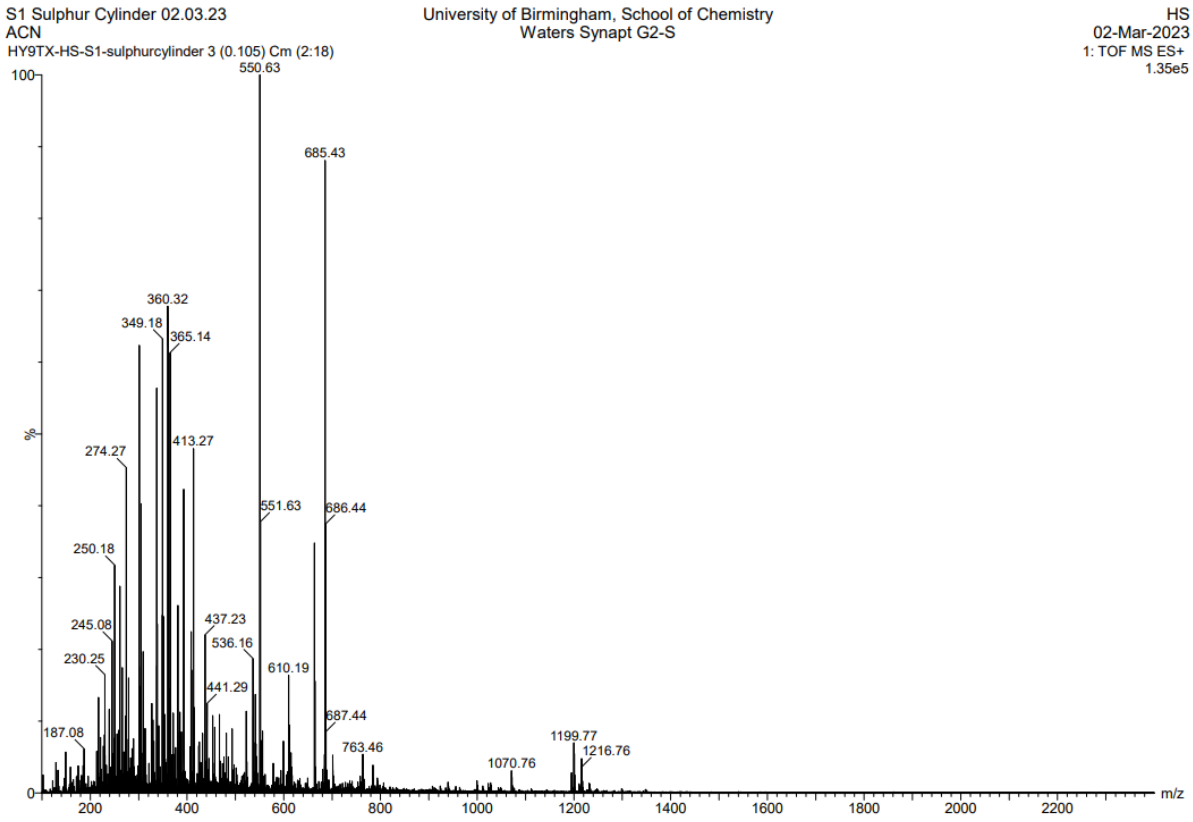


Figure 2.3.18 NMR of precursor for sulphur-based spacer (400 MHz, DMSO-d)

or whether the extraction method was unsuitable. This would be the case if the product had some solubility in diethyl ether, therefore, would be extracted along with the DMSO. Therefore, it was decided to repeat the experiment in deuterated DMSO which enabled a direct NMR to be taken of the reaction mixture without necessary extraction. Moreover, samples were taken immediately from the addition of the 4-fluoronitrobenzene and every 30 minutes thereafter for two hours. Interestingly, all the NMR spectra taken from the reaction showed clear formation of a new species consistent with the number and type of peaks that would be expected for the desired product, suggesting that the conditions were suitable, and instead, that the extraction method was the problem. However, factoring in the difficulties associated with using DMSO as a solvent it was decided that this synthesis method was unsuitable and that an alternative method should be used in further synthesis attempts.

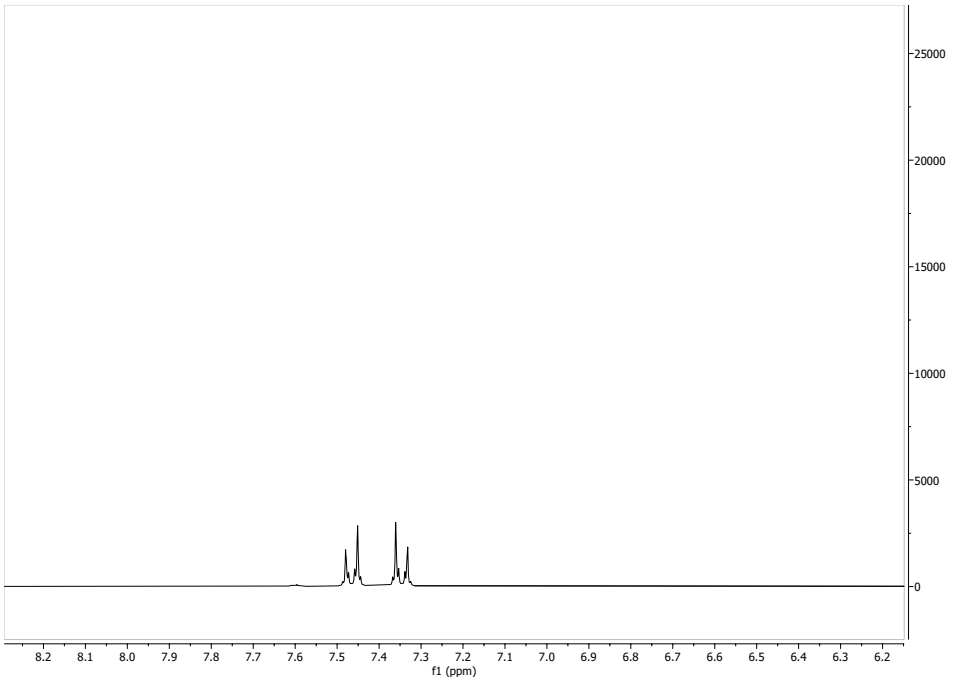


Figure 2.3.19 NMR of 4,4-bis-(4-nitro-phenylmercapto)-biphenyl (400 MHz, MeOH-d)

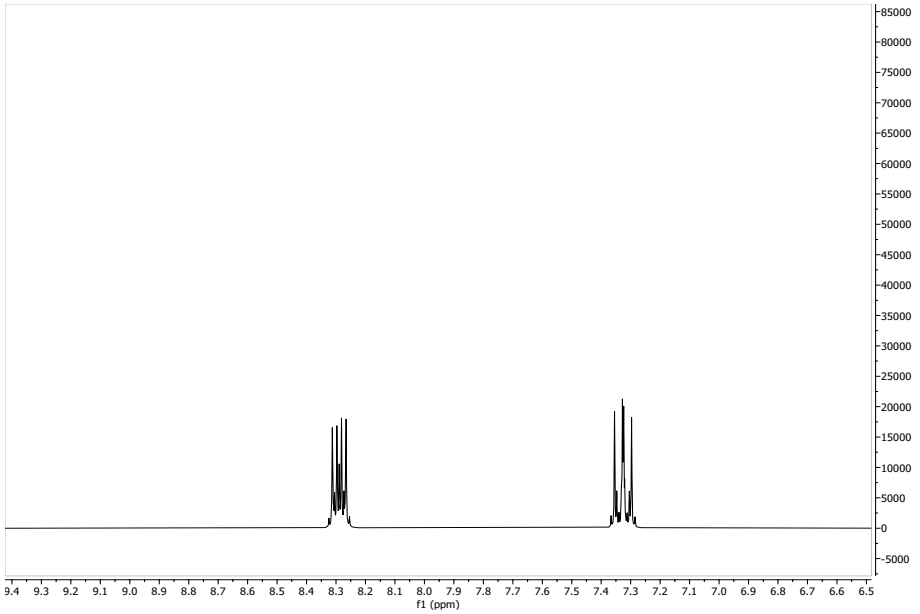


Figure 2.3.20 NMR of 4-fluoronitrobenzene (400 MHz, MeOH-d)

Further research led to the discovery of an alternative approach used by an Italian research group in 1967.[96] The group was able to prepare the desired 4,4-bis-(4-nitro-phenylmercapto)-biphenyl using ethanol as a solvent and potassium hydroxide as the

base. Due to the suitability of the solvent it was decided to use these conditions in the subsequent synthesis attempt. The reaction was heated to 55°C and on addiction of the 4-fluoronitrobenzene the reaction mixture turned orange. After two hours the reaction was filtered and the product isolated *in vacuo*. Both NMR and mass spectrometry were completed to analyse the product further.

The NMR (Figure 2.3.21) showed more proton environments than would be expected for the desired product. On further analysis a number of these peaks were assigned as peaks that related to the initial reactant biphenyl-4,4'-dithiol, however, no peaks relating to 4-fluoronitrobenzene were seen.

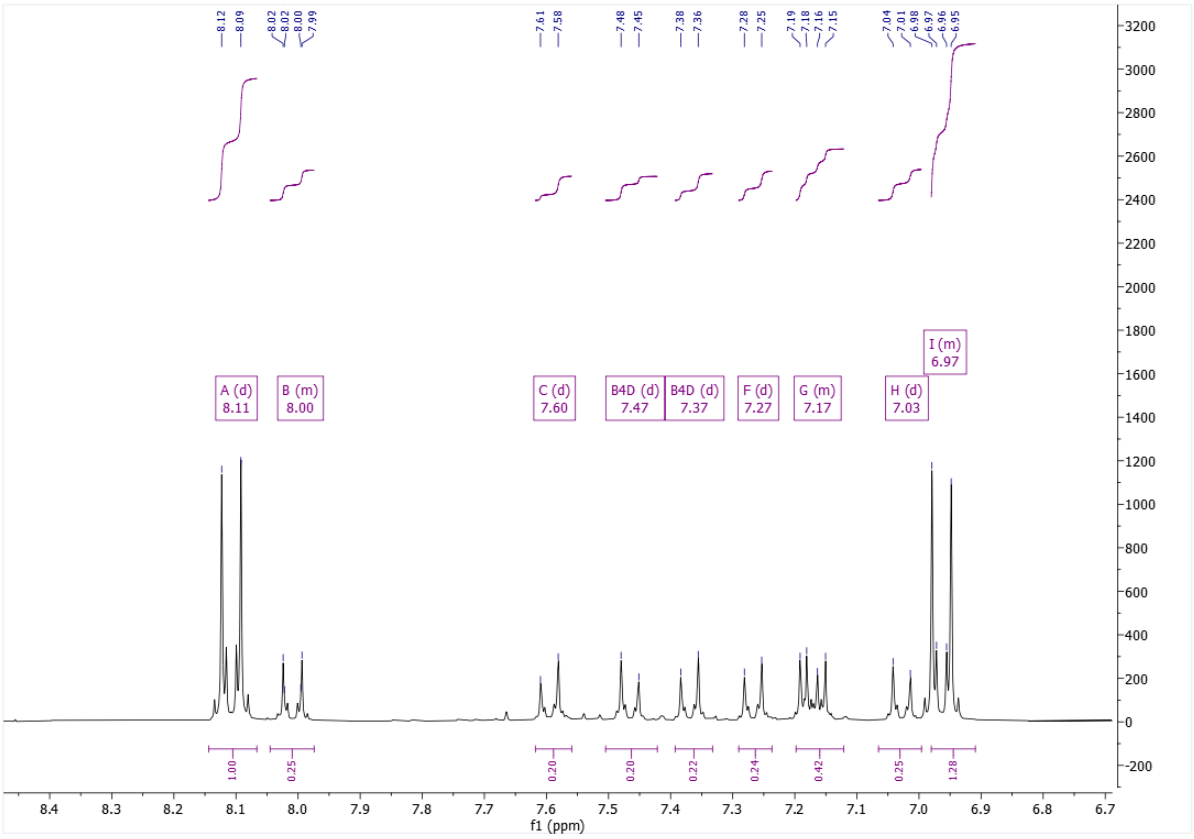


Figure 2.3.21 NMR of precursor of sulphur spacer (400 MHz, MeOH-d)

When comparing to the NMR spectrum generated by the first synthesis to the one below, it could be suggested that peaks labelled B, C, F and I are potentially product peaks as they have the same splitting patterns and similar shifts as the initially assigned product peaks. Unfortunately, this leaves two peaks, labelled A and H, that could not be assigned as product peaks. It is impossible that these peaks relate to the minor product, predicted to be potassium hydride, as that would not show up on an NMR spectrum. There is potential that peaks A and H could relate to 4-nitrophenol, which would be expected as a side a product for this reaction, however, there is no singular phenol peak in the NMR confirming 4-nitrophenol's presence. Although the mass spectrum was relatively clean it did not help clarify matters further. There was no peak at *m/*z = 453, which would relate to the desired product or any other peak relating to the product with a metal cation such as potassium of sodium. Moreover, the only peaks present could not be assigned to either starting material.

Further analysis was attempted using TLC of the reactants versus the reaction mixtures taken at each time interval. The TLC were run in 10:90 MeOH:DCM. Overall, although there was some suggestion of a new product being formed in the reaction mixture over time, the results were not hugely informative. Over time, it was observed that the samples collected grew darker in colour. NMR analysis (Figure 2.3.22) showed a change in peaks over this time period, where most recent spectrums were noisier and showed less defined peaks and splitting patterns than the initial spectrum. After considering all of these results it was decided to not pursue further the synthesis of this sulphur-based ligand due to the challenges presented in the initial synthesis steps and the time that might be required to overcome then.

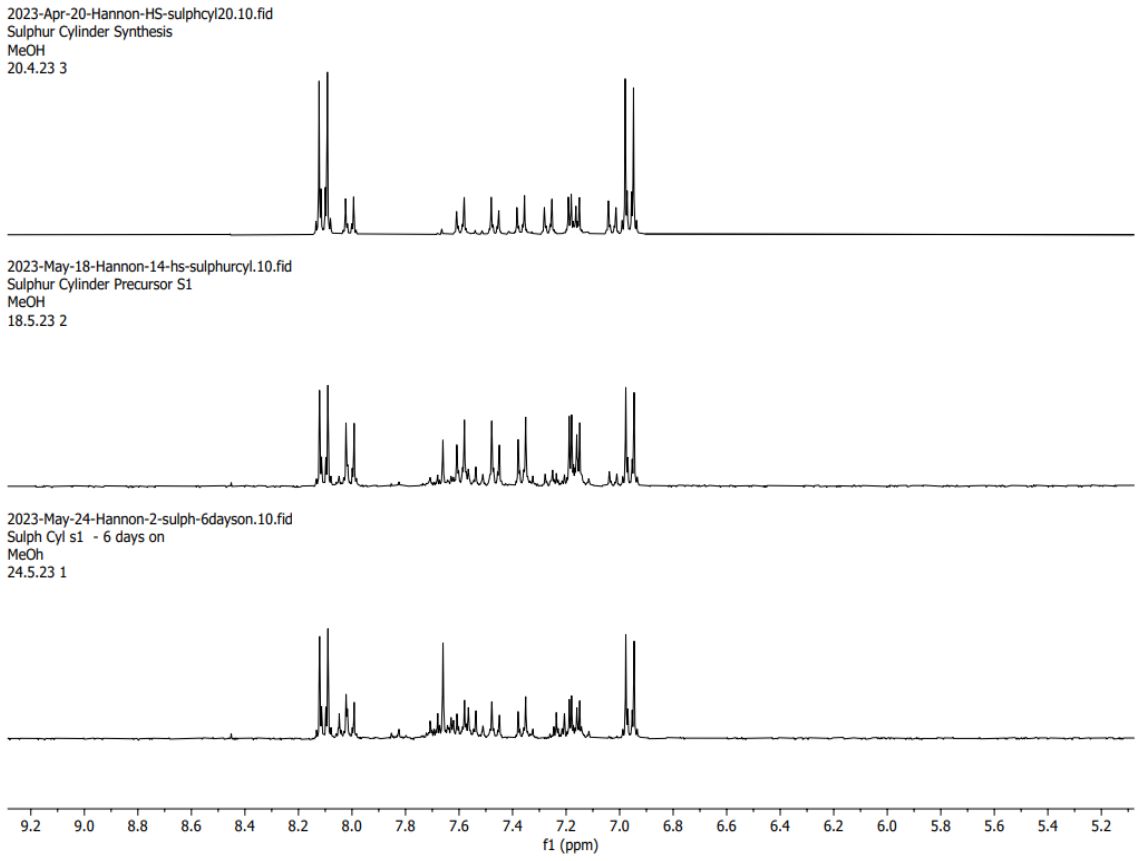


Figure 2.3.22 Comparative NMR showing degradation of precursor over time (400 MHz, MeOH-d)

2.5.2) Attempted Synthesis of an extended ligand with CH_2 group in the centre of the spacer

As discussed previously in this chapter, the initial cylinders synthesised from the longer pyridine-based ligand were shown to exist in two forms, the helicate and mesocate. While this does not necessarily present a major challenge in the therapeutic use of these metallo-supramolecular structures, it does mean there is a lack of clarity in terms of which structural form of the cylinder interacts best with higher-order DNA structures. In order to try and overcome this ambiguity in relation to structure an attempt was made to modify the initial pyridine-based ligand to force the resulting supramolecular structures to adopt the helicate structure. It was hypothesised that by adding a CH_2 between the two aryl groups in the spacer, a helical structure will be induced as seen with the Hannon group parent cylinder.

Figure 2.3.23 shows the first step in the synthesis of this CH_2 spacer. Further research into the literature showed only one other recorded attempt of this synthesis involving slightly different reactants.[97]

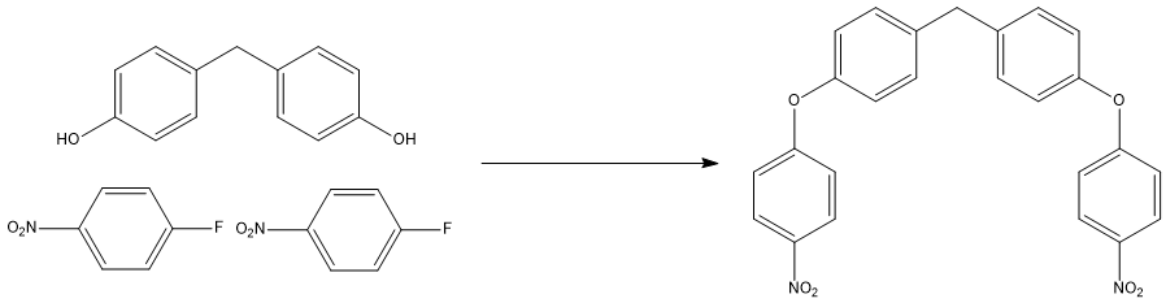


Figure 2.3.23 Synthesis of precursor to CH_2-based spacer

Due to the similarities between this initial synthesis step and the initial step in sulphur spacer synthesis it was decided to use similar synthesis conditions. Therefore, 4-

fluoronitrobenzene was heated under in a 1:2 ratio with bis(4-hydroxyphenyl)methane with ethanol as the solvent and potassium hydroxide as a base. MS showed a peak at *m/z* 483 which suggested the presence of the product with K^+ and ^1H NMR was able to confirm the product was successfully synthesised. The yield was 2.5 times the maximum theoretical yield which suggested that impurities and/or solvent were also present in product sample, however, the decision was made to not attempt to purify the sample further at this stage in case the desired product was degraded or lost.

The next step was to reduce the nitro groups on this half-ligand to amino groups, as these amino would eventually enable 2-pyridinecarboxaldehyde to react with the half-ligand to form the desired full-length ligand. Consequently, the ligand was suspended in dichloromethane and stirred with sodium borohydride for 24 hours, before the aqueous layer was extracted and dried. During the 24 hours the solution underwent two colour changes – initially from cream to black before going clear. The resulting solid that was extracted gave a yield of 115 %, once more suggesting the product contained impurities. Analysis by MS showed at peak at *m/z* 382 which resalted to the desired product, indicating the reduction had been a success., however, it also showed presence of the precursor, bis(4-(4-nitrophenoxy)phenyl)methane, suggesting that the reaction had not gone to completion. Nonetheless, as the reduction step had clearly taken place (at least partially), the synthesis was continued.

The final step of synthesising the full desired ligand involved a condensation reaction between pyridine-carbaldehyde and the reduced ligand formed in the previous step as shown in Figure 2.3.23.

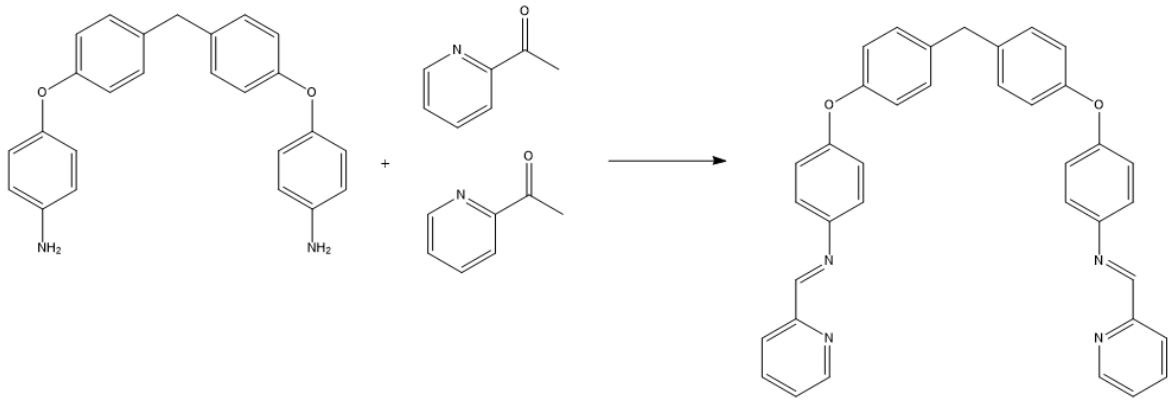


Figure 2.3.23 Synthesis of extended CH_2-based ligand

The reaction was carried out at 50°C in an inert atmosphere for 24 hours, resulting in a small amount of cream precipitate forming, which was isolated by vacuum filtration. Analysis of this precipitate by MS showed a strong peak at *m/z* 379 and a mass of small peaks at *m/z* 650 - 1000. None of these peaks could not be related to the desired product which had a known molecular weight of 560u, consequently, it was decided to remove the solvent from the remaining filtrate to assess whether the desired product had remained in solution. The resulting solid showed similar results to the initial precipitate in MS, suggesting the desired product had not been produced. After considering this result, it was decided to not pursue synthesis of this ligand further due to the challenges presented in the initial synthesis steps.

Chapter 3: DNA BINDING STUDIES

3.1) Group Research and Individual Aims

After successful synthesis of the parent cylinder in 1997, the Hannon group looked at the interaction of this cylinder with B-DNA as well as higher order DNA structures. The group found that the parent cylinder had the ability to bind to both the B-DNA in its major groove but also to the centre of 3WJs as well as other y-shaped structures.[55,57] As discussed this was a major development in field of supramolecular chemistry as it showed the potential of these structures to act as effective therapeutic agents.

Further research conducted by the group showed that the metal ion coordinated within the $[M_2L_3]^{3+}$ cylinder had an impact on the cylinder's properties. While the original $[Fe_2L^1_3]^{3+}$ cylinder showed the ability to cause dramatic intramolecular coiling as well as binding to higher-order DNA structures, the ruthenium(II) analogue showed additional abilities such as exhibiting fluorescence on DNA binding and activity in cancer cell lines.[56] Moreover, when attempting to observe the binding of the cylinder to higher-order DNA structures in polyacrylamide gel electrophoresis studies, the group found that $[Ni_2L^1_3]^{3+}$ was much more suitable than the iron(II) analogue as the Ni ion did not quench the fluorescent stain and thus enabled a binding event to be clearly seen. Additionally, the group showed that functionalisation and modification of the ligand within the parent cylinder impacted the cylinder's ability to bind to DNA. [56]

While the group has widely explored modifying and functionalising the original parent cylinder with the aim of seeing how these changes have impacted the cylinder's ability to

bind to DNA structures, there has not been investigation into how changing the length of these cylinders might impact DNA binding. Therefore, this chapter will aim to show that longer cylinders, that are analogues of the parent cylinder, have the ability to bind to higher order DNA structures, such as 3WJs, 4WJs and G-quadruplexes.

3.2) In vitro studies

3.2.1) UV-Visible Absorbance Spectroscopy

UV-Visible absorbance spectroscopy (UV-Vis) is a technique used to analyse a sample through the production an absorbance spectrum. Light is passed through a compound in solution and some energy is absorbed. This energy used to promote electrons from the compound's ground state to the first excited state and is represented in an absorbance spectrum by peaks that are characteristic to the compound in solution. The volume of light absorption (and thus the size of the peaks) is dependent on two concepts; the number of molecules in solution and the pathlength of the light through the solution. This relationship is clearly defined through the Beer-Lambert Law:

$$A = \varepsilon\, c\, l$$

Where A is absorbance, ε is the extinction coefficient, c is the concentration of the sample with units mol L^{-1} and l is pathlength of the cell (in cm).

Absorbance can also be determined through logarithmic ratio of the incident and emergent light intensity as shown:

$$A = \log\left(\frac{I_0}{I}\right)$$

Where I_0 is incident light intensity and I is emergent light intensity.

As discussed, absorbance peaks or bands are characteristic to the compound in the sample. For Fe complexes an absorbance band tends to occur around 580 nm while for DNA the characteristic band occurs with a minimum at 230 nm and a maximum at 260 nm.

3.2.2) Absorbance of pyridine-based helicates

During the synthesis of these pyridine-based supramolecular helicates, it became clear that successful formation of the desired helicate product was favoured by $[BF_4]^-$ as the counterion and acetonitrile as the solvent. Syntheses involving other combinations of counterions and solvents were attempted, however, none yielded the product to the same level of volume and purity. The cylinders also showed good stability in acetonitrile, hence, initial UV-Vis spectroscopic investigations were performed using this solvent.

The absorbance of all pyridine-based cylinders was recorded at room temperature. $[Fe_2L^1_3][BF_4]_4$ showed the characteristic band consistent with all Fe pyridylimine complexes at 580nm and shown in Figure 3.2.2.1, Graph 1. This band relates to the excitation of the electrons from the t_{2g} orbitals of the metal to the pi^* orbitals of the ligand and any variation seen in this band with time or solvent would be indication of the stability of the cylinder. In acetonitrile, the UV-vis spectrum (Figure 3.2.2.1 Graph 2) of the $[Fe_2L^1_3][BF_4]_4$ cylinder showed little change across a 72 hour period, however, this was

not the case for the cylinder in different solvents. Across the same time period, the expected peak relating to the cylinder disappeared, when the cylinder was dissolved in water, methanol and TrisHCl (buffer). Moreover, the $[Fe_2L^1_3][BF_4]_4$ cylinder showed rapid degradation when in buffer, with the characteristic peak not even being present at T = 0.

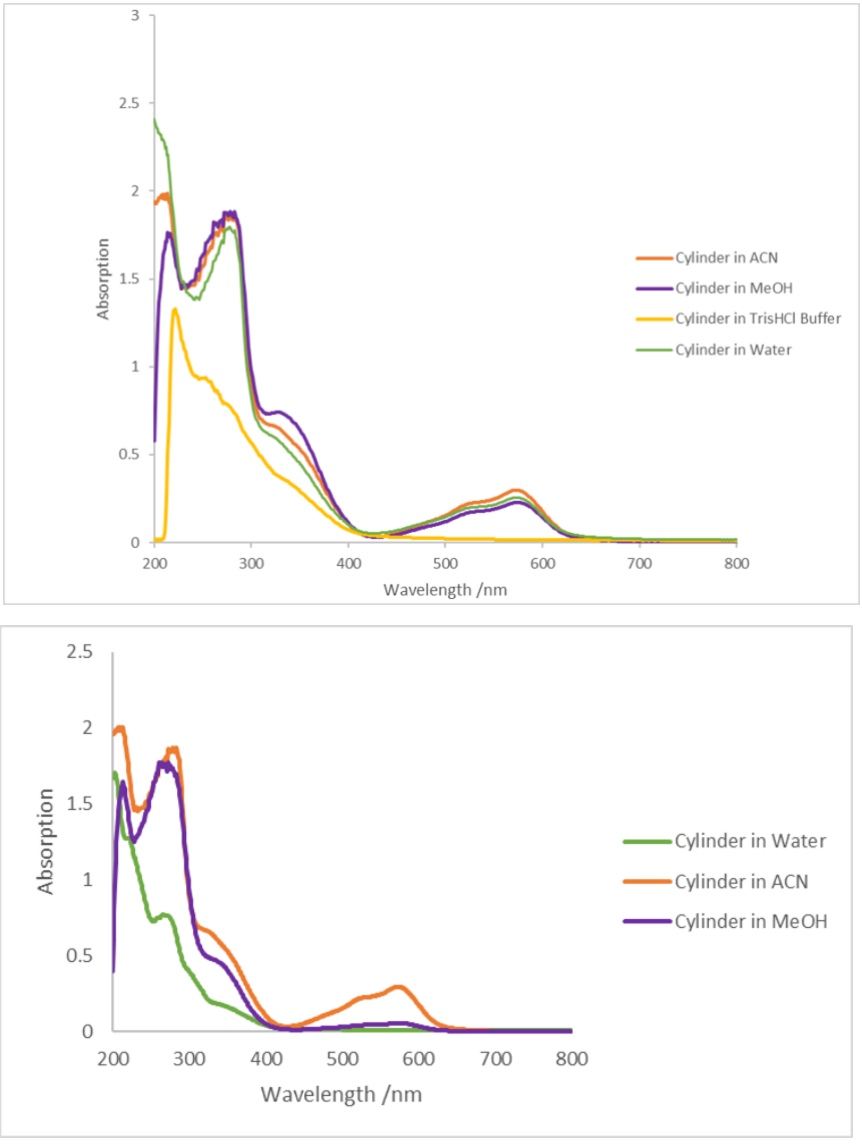


Figure 3.2.2.1 UV-VIS absorbance spectroscopy of 1 mM $[Fe_2L^1_3][BF_4]_4$ helicate in different solvents. Graph 1 (top) shows the spectrum at 0 hours and Graph 2 shows the spectrum after 72 hours

The loss of the characteristic peak was due to cylinder falling apart due to instability in the solvent. All 3 solvents have the ability to coordinate to metals so it is unlikely this is what results in the cylinder's instability in methanol and water. However, in water iron can be oxidised to form insoluble iron oxide (iron hydroxide) or rust, which is a thermodynamic sink. In turn, the formation of rust will cause the cylinder to degrade and since methanol has a high chance of containing water as an impurity, this is also the most likely reason for why the cylinder shows instability in water and methanol. On the other hand, ACN rarely contains high levels of water impurity, hence the cylinder is shown to be stable in this solvent. Unfortunately, acetonitrile is not a suitable solvent to use in biological systems as it exhibits toxic activity within cells. This is mainly due to the fact that the solvent is transformed into the toxic molecule cyanide by the cell's own metabolic machinery. It has also been shown that at high concentrations, acetonitrile negatively impacts cell membranes through its integration into the outer layer of the phospholipids. In turn, the incorporation of the solvent molecules into the cell membrane induces the lipid membrane to undergo a phase transition and hence the membrane will exhibit significant curvature or even burst. Consequently, using acetonitrile as the solvent when studying the biological activity of $[Fe_2L^1_3][BF_4]_4$ in cells would not be suitable.[98]

Since $[Fe_2L^1_3][BF_4]_4$ only showed good stability in acetonitrile this presented a real problem. To overcome this issue, it was necessary to develop the $[Fe_2L^1_3][BF_4]_4$ helicate further so that it was soluble and remained stable in a solvent that was biocompatibility with cells, for example methanol or water. The first modification made was to change the metal centres of the helicate from Fe(II) to Ni(II). The reason for this substitution is that Ni(II) based cylinders have previously demonstrated greater stability than the

corresponding Fe(II) cylinders.[73] Thus $[Ni_2L^1_3][BF_4]_4$ was synthesised from nickel (II) tetrafluoroborate using similar conditions to the corresponding iron cylinder. The $[Ni_2L^1_3][BF_4]_4$ helicate showed good stability in ACN as was expected, however, presented the same issue as $[Fe_2L^1_3][BF_4]_4$ as was insoluble in both methanol and water.

While overcoming the issue with stability was relatively straightforward, overcoming the solubility issues proved more challenging. One method of improving the helicates solubility in water/methanol was to change the counterion from $[BF_4]^-$, to a counterion that is more compatible with polar protic solvents such as chloride, hexafluorophosphate or acetate. As discussed in section 2.4.1), previous attempts to synthesise the chloride version of the helicate were unsuccessful as they were low yield and low purity. Therefore, it was necessary to develop an alternative way of generating the chloride version of this cylinder. The first approach explored was to use Dowex to attempt counterion conversion of the cylinder from the $[BF_4]^-$ to the $[Cl]^-$ in methanol solvent. The desired chloride product was then isolated by allowing the methanol to evaporate off. This method proved successful with MS confirming presence of the cylinder with a seemingly good purity. However, extracting the cylinder from the solution proved difficult and consequently, the resulting yield of the cylinder was extremely low, meaning further use of this cylinder was not viable.

Another counterion that presented a possible solution to the issue of solubility was $[PF_6]^-$. Generally, salts with this counterion show high solubility in polar organic solvents, traditionally acetone or acetonitrile, however some have shown solubility in methanol. Therefore, it is not implausible that the $[Ni_2L^1_3][PF_6]_4$ cylinder would have dissolved well

in methanol. Although the Dowex method explained above did produce the desired cylinder, the product was largely unusable due to the difficulties involved with extraction. Therefore, it was decided to not use this method but rather to attempt to synthesise the cylinder using similar conditions to the $[Ni_2L^1_3][BF_4]_4$ synthesis. Nickel (II) tetrafluoroborate was refluxed with the ligand and ammonia hexafluorophosphate. The desired cylinder was produced to a high yield and purity, as confirmed by mass spectroscopy, where the negative spectrum of the product showed only the $[PF_6]^-$ ion as present. However, similarly to the $[Ni_2L^1_3][BF_4]_4$ cylinder, showed little to no solubility to methanol.

The final counterion that was considered was the acetate counterion. Almost all salts containing this counterion are known to be soluble in water and consequently, it could be predicted that cylinders containing this counterion would also be soluble in water and/or methanol.[99] Synthesis of the cylinder was attempted using similar conditions to the synthesis of the $[Ni_2L^1_3][BF_4]$ cylinder, however, used nickel (II) acetate tetrahydrate as the starting material and methanol as the solvent instead of nickel (II) tetrafluoroborate and acetonitrile. It was expected that if the reaction was successful the yield would not be high as the ligand previously did not show good solubility in methanol. Fortunately, the synthesis was successful with the desired $[Ni_2L^1_3][Ac_4]$ cylinder being produced in a yield of 45 %. Mass spectrometry confirmed presence of the cylinder and showed a high level of purity. Moreover, the helicate showed good solubility in methanol and water, overcoming the problem of incompatibility with bioactive systems.

While the $[Ni_2L^1_3][Ac]_4$ cylinder was soluble in water, the cylinder showed instability in this solvent, which was not overcome by using the Ni over Fe as anticipated. The UV-Vis spectrum (see Figure 3.2.2.2 below) shows that within 6 hours the cylinder is significantly degraded. Both the peaks that relate to the Ni cylinder (seen at 254 nm and from 280 nm - 380nm for the 0 hour measurement) are almost completely lost by 6 hours, indicating that the cylinder is decaying rapidly to ligand. The instability of the cylinder presents a challenge for use of $[Ni_2L^1_3][Ac]_4$ as a cellular DNA binder as it may too quickly degrade in situ so will not be able to carry out its intended functions (although the DNA-binding may stabilise the cylinder towards degradation). It is important to note that the consistent peak seen at 270 nm in Figure 3.2.2.2 relates to ligand, however, it is shifted in comparison to the peak seen of the ligand on its own in Figure 3.2.2.3 due to solvent effects.

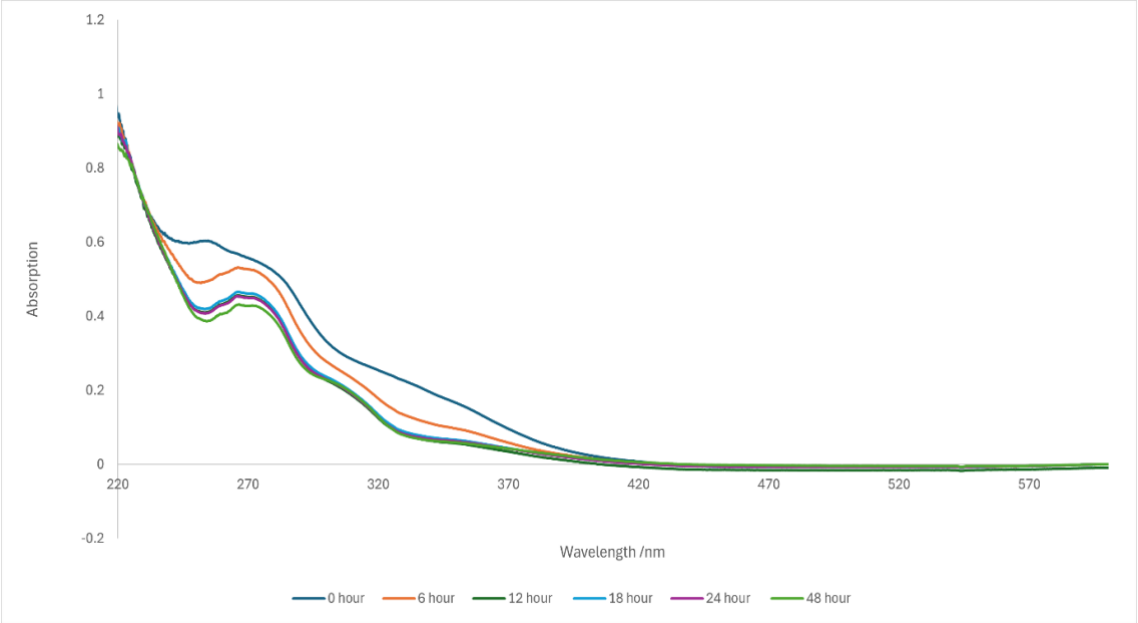


Figure 3.2.2.2 UV-VIS absorbance spectrum of $[Ni_2L^1_3][Ac]_4$ in water across 48 hours

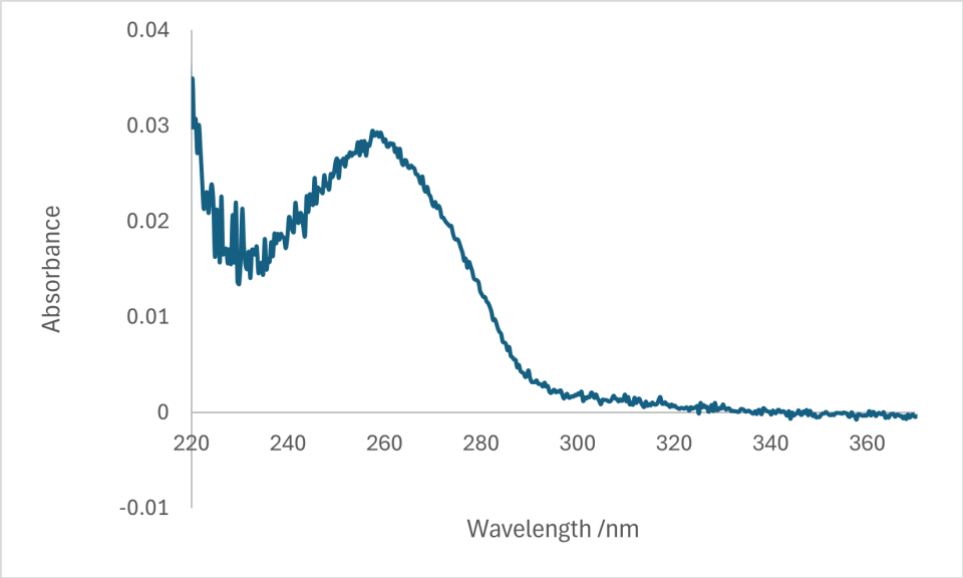


Figure 3.2.2.3 UV-VIS absorbance spectrum of Ligand L^1 in 10% ACN

In an attempt to overcome the stability issue, the same UV-Vis experiment was carried out using the $[Ni_2L^1_3][Ac]_4$ cylinder in the presence of 200 uM CT-DNA, where the cylinder was at a 1:50 ratio with the DNA. Interestingly, the presence of the CT-DNA stabilised the helicate for 24 hours shown by the peak between 280nm and 380 nm in Figure 3.2.2.4. It is likely the presence of the DNA slowed the rate of degradation so that

the complex was still stable at 24 hours, however after this point the cylinder degraded more rapidly to ligand. While the result does not completely solve the issue of stability for the helicate, it does suggest that any results from in vitro experiments involving the $[Ni_2L^1_3][Ac]_4$ in DNA can be trusted as it is likely the cylinder has been stabilised by the presence of DNA.

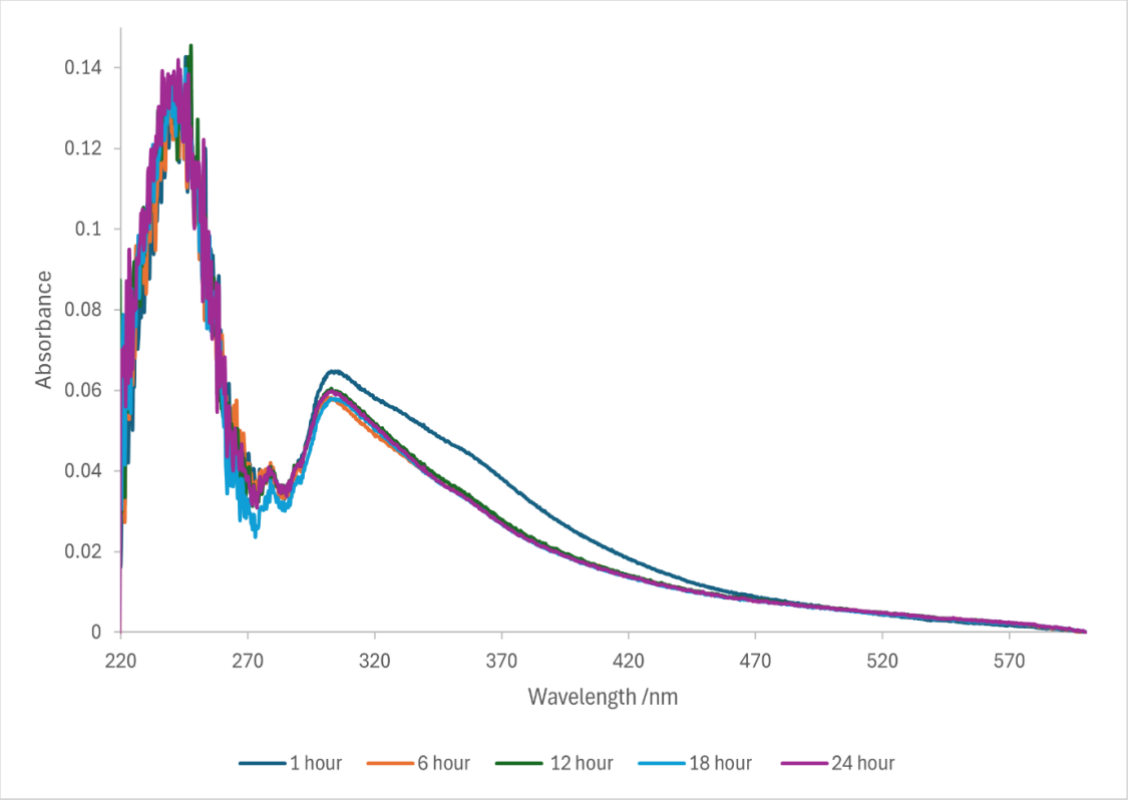


Figure 3.2.2.4 UV-Vis absorbance spectrum of $[Ni_2L^1_3][Ac]_4$ cylinder in the presence of 200 uM CT-DNA

Overall, the UV-Vis studies of the cylinders in this section presented the issues of instability and insolubility that impact the effectiveness of the cylinders to work at DNA binders. To overcome either one of these issues, the helicates were modified to contain different counterions. It was found that the $[Ni_2L^1_3][BF_4]_4$ showed strong stability in

acetonitrile, however, was insoluble in both methanol and water, meaning that its ability to act as a DNA binder is limited. However, the more biocompatible cylinder, $[Ni_2L^1_3][Ac]_4$, which showed good solubility in water and methanol, was relatively unstable and therefore, also showed a limited ability to act as a DNA binder. While both cylinders show potential, further modification and development of both is needed before they can be considered as real options for therapeutic action.

3.2.3) Circular Dichroism

Circular dichroism is an analytical technique that uses circularly polarised light to probe the structure of chiral compounds that are optically active. The spectroscopic technique works by quantifying the difference in absorption of right or left circularly polarised light by a compound. Optically active chiral compounds, like DNA, can selectively absorb either right (RCP) or left (LCP) circularly polarised light. The difference in absorption of RCP and LCP can be detected as it affects the amplitude of the incident light passing through the compound.

$$\Delta A = A_l - A_r$$

The Beer Lambert law stated above in equation 1 shows that absorption is equal to product of the molar extinction coefficient, pathlength and concentration. When we apply equation 3 to the Beer Lambert Law following equation is generated,

$$A = (\varepsilon_l - \varepsilon_r)cl$$

Or

$$A = (\Delta\varepsilon)cl$$

Where $\Delta\varepsilon$ is known as the molar absorptivity or molar circular dichroism.

This molar absorptivity of chiral compounds is different for RCP and LCP light. Consequently, these are absorbed by different amounts in a chiral compound, resulting in the superimposed light changing from being linearly polarised to elliptically polarised.

Therefore, circular dichroism measures this ellipticity of polarisation using the equation below,

$$tan\theta = \frac{E_l - E_r}{E_l + E_r}$$

Where E is the magnitude of electric field vector and θ is the degrees of ellipticity. Since DNA is made up of chiral units, the source being deoxyribose sugar, circular dichroism can be used to measure conformational changes of the DNA molecule due to changes in environment or due to binding of an external molecule. The ability of CD to detect structural changes in DNA due to binding is highly relevant and invaluable to DNA binder research as it enables researchers to detect whether their compound has successfully interacted with DNA.

3.2.4) Circular Dichroism (CD) of pyridine-based cylinders

CD analysis of all novel cylinders synthesised was conducting in order to ascertain whether these cylinders were able to bind to B-DNA and induce conformational changes. The first cylinders to be analysed by CD were the initial novel helicates synthesised, $[Fe_2L^1_3][BF_4]_4$ and $[Ni_2L^1_3][BF_4]_4$. 100 µM samples of these helicates were made up and

titrated against 100 µM ct-DNA in buffer. As these helicate samples were not soluble in water, the sample was initially dissolved in acetonitrile before water was added to make up the solution to 100 µM. While this was not the ideal, it enabled CD spectrums of the helicates to be taken and since the percentage of the acetonitrile in the titrated samples was less than 0.5%, the negative impact on the DNA was minimal.

The CD titration of $[Fe_2L^1_3][BF_4]_4$ with 100 µM ct-DNA is shown in Figure 3.2.4.1 below, where the ratio of DNA basepairs to complex is given. The graph clearly shows that as the ratio of basepairs to complex decreases, so as more complex is added to the ct-DNA, the characteristic peak of the ct-DNA decreases in intensity and shifted slightly right. The spectrum also becomes noisier at the higher concentrations of cylinder added. Initially, these observations were attributed to the ct-DNA coiling due to the presence of the cylinder, however, further analysis of the solution inside the cuvette showed that the solution had begun to turn cloudy. This, coupled with the fact that at the highest

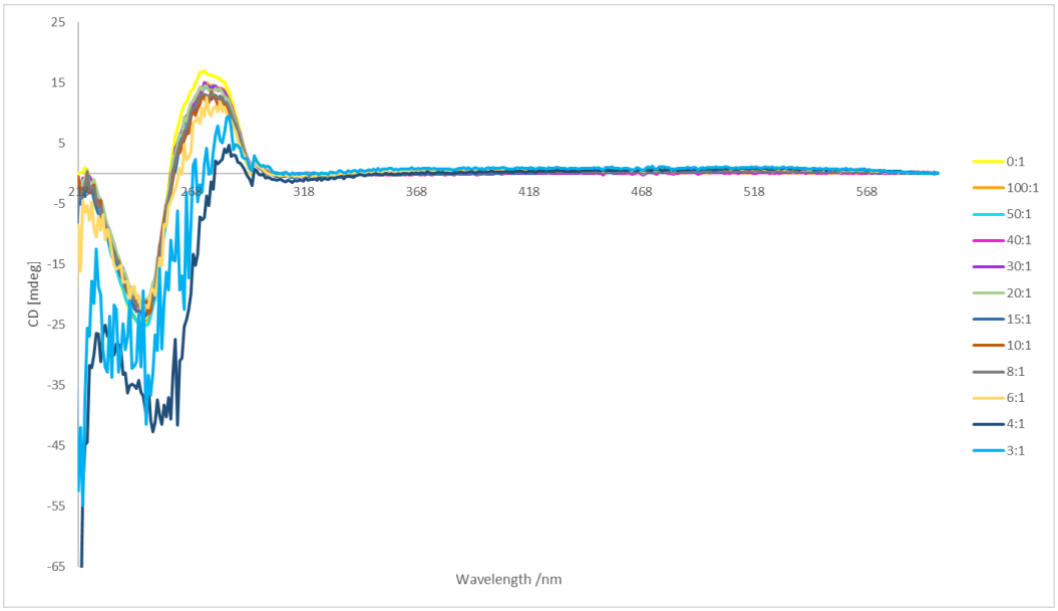


Figure 3.2.4.1 CD titration of $[Fe_2L^1_3][BF_4]_4$ against 100 µM ct-DNA in buffer (100 mM NaCl, 10 mM TrisHCl) where the ratio is DNA base pairs to complex.

concentrations of the cylinder the spectrum became excessively noisy at wavelengths below 220 nm, suggested that in-fact the presence of the cylinder was causing the DNA to precipitate out of solution rather than the desired solution binding and inducing of conformational changes.

Similar observations were observed for the titration of the $[Ni_2L^1_3][BF_4]_4$ with ct-DNA shown in Figure 3.2.4.2, however, the spectrum was significantly noisier at lower values of R, suggesting this cylinder caused higher degrees of DNA precipitation than its Fe(II) analogue.

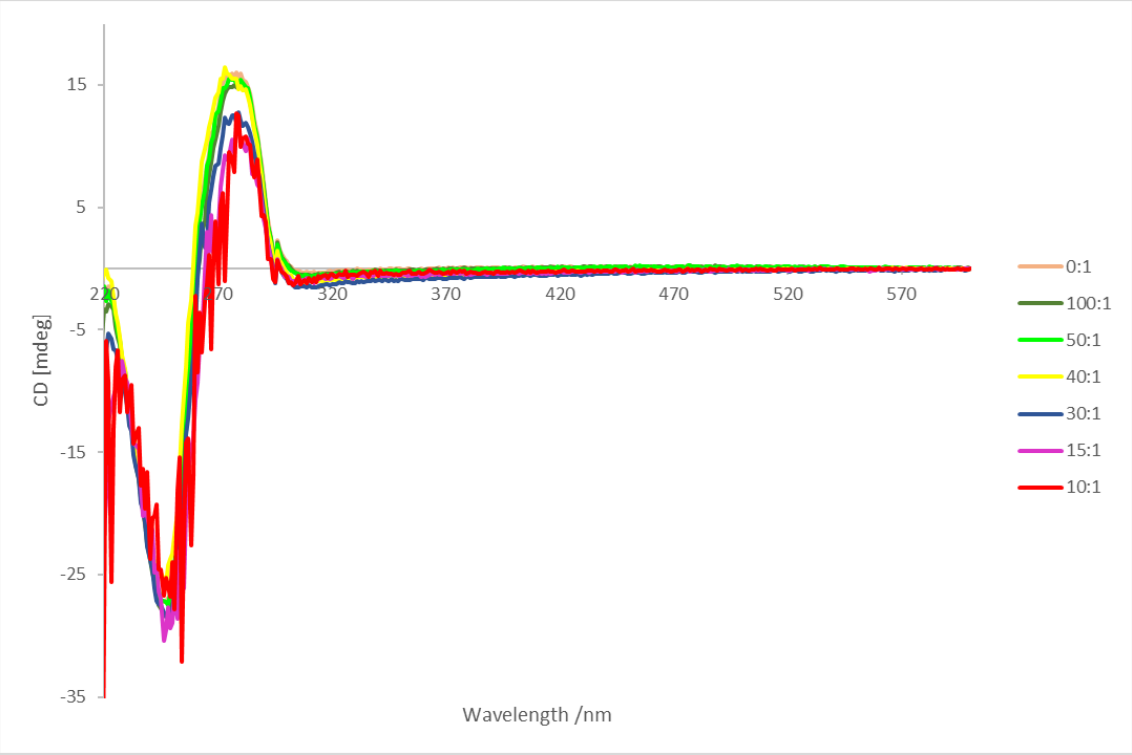


Figure 3.2.4.2 CD titration of $[Ni_2L^1_3][BF_4]_4$ against 100 µM ct-DNA in buffer (100 mM NaCl, 10 mM TrisHCl) where the ratio is DNA base pairs to complex.

It was unclear at this stage whether it was the increasing concentrations of cylinder or the solvent, acetonitrile, causing the ct-DNA to precipitate out of solution. In order to investigate this, the CD spectrum of the water soluble analogue $[Ni_2L^1_3][Ac]_4$ (Figure

3.2.4.3), binding ct-DNA was investigated. Similarly to the previous CD experiments, $[Ni_2L^1_3][Ac]_4$ was titrated against 100 µM ct-DNA in buffer, however, these titrations were performed more rapidly than those completed with the other cylinder analogues due to the known instability of this cylinder. Analysis of the CD spectrum produced showed the same spectral characteristics seen for both $[Ni_2L^1_3][BF_4]_4$ and $[Fe_2L^1_3][BF_4]_4$ confirming that the increasing concentrations of the cylinder caused the ct-DNA to precipitate out of solution.

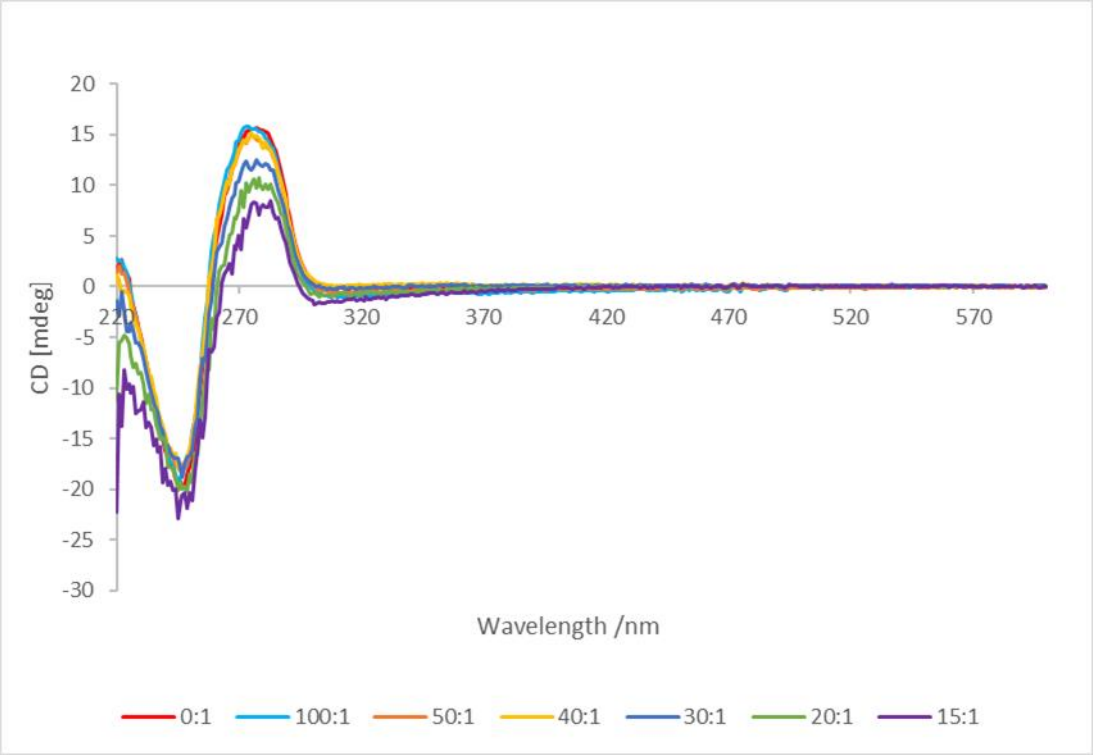


Figure 3.2.4.3: CD titration of $[Ni_2L^1_3][Ac]_4$ against 100 µM ct-DNA in buffer (100 mM NaCl, 10 mM TrisHCl) where the ratio is DNA base pairs to complex.

Overall, while it was clear from the various CD titrations completed that presence of the longer helicate (and its analogues) did affect ct-DNA, the spectra produced provided no direct evidence to indicate that the cylinders bound to the ct-DNA (though the

precipitation may indicate that) or that their presence induced conformational changes in the ct-DNA.

One further CD titration was completed using similar conditions to the previous CD experiments, involving C-MYC DNA quadruplexes instead of ct-DNA. As aforementioned, C-MYC quadruplex are G4 structures that exist within the promoter of the MYC oncogene. Stabilising these structures prevents the over -expression of the MYC gene and therefore, is likely to reduce cancer metastasis.[43] Consequently, if the longer helicate is able to bind to C-MYC DNA and induce conformational changes where the G4 structure is stable, it suggests the potential of the cylinder to act as an anti-cancer agent. CD titrations of the cylinder against C-MYC quadruplexes will provide evidence as to whether the cylinder is able to interact with this higher-order form of DNA and therefore, it is able to act as a therapeutic agent.

$[Ni_2L^1_3][Ac]_4$ was titrated against 2 uM C-YMC DNA in buffer (20 mM LiAc/KCL) to produce the CD spectrum shown in Figure 3.2.4.4. In this case R refers the to the ratio of C-YMC DNA to cylinder. Analysis of the spectrum shows that as the cylinder concentration increases, the concentration of C-MYC DNA decreases and shifts slightly to the left of the spectrum. Moreover, unlike previous CD spectrums there is not an increase in noise in the spectrum at higher concentrations of cylinder, suggesting that the DNA is not precipitating out of solution at these high concentrations as seen previously. On the whole, these spectral shifts suggest that the cylinder is interacting with the C-MYC DNA and potentially inducing small conformational changes, however, since there are no dramatic changes within the spectrum no definitive conclusions can be drawn.

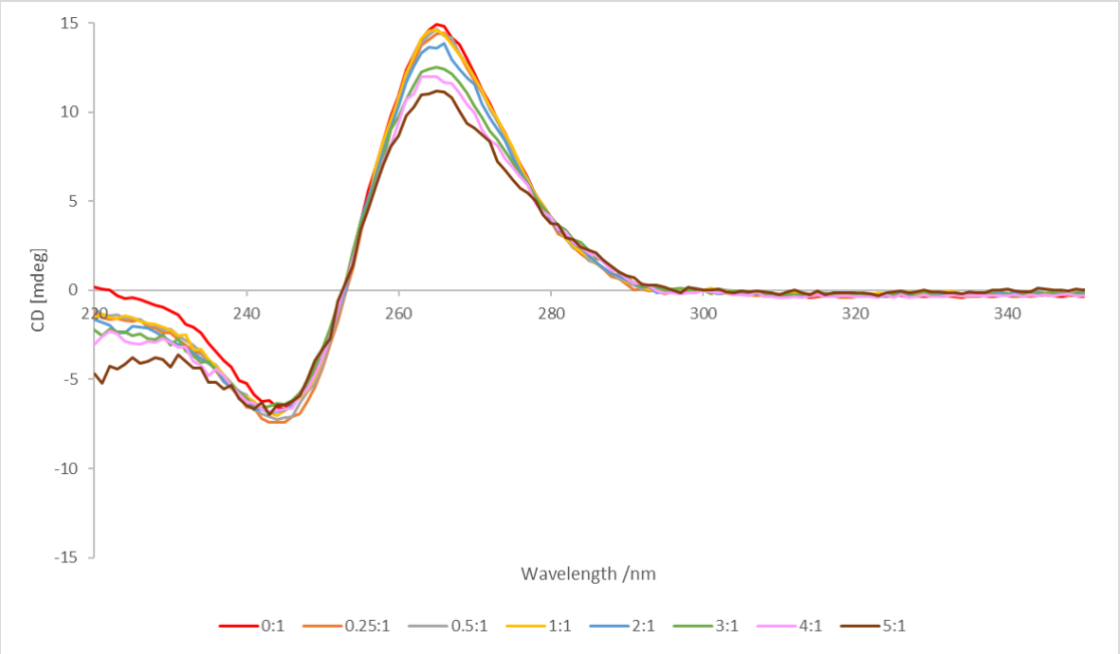


Figure 3.2.4.4 CD titration $[Ni_2L^1_3][Ac]_4$ with 2 uM C-MYC DNA in buffer (20 mM LiAc and KCl) where the ratio is DNA base pairs to complex

3.2.5) Linear Dichroism (LD)

Linear dichroism (LD) is defined as spectroscopy that measures the difference in absorption of light polarised parallel and light polarised perpendicular to an axis by a sample.[100]

The technique works by shinning plane polarised light through a sample, first parallel to the sample and then perpendicular. As the polarised light is shone through the sample, some of the light is absorbed by the sample material, promoting electrons within the material to an excited state. The movement of charge generated by this light absorption is known as an electronic transition. LD measures the direction in which the electronic transition moves, which is know as electronic transition polarisation. Since the orientation

of the molecule affects how much light is absorbed, where one dimensions requires more energy for photoexcitation than the other, and the difference gives a net polarisation in a specific direction. Therefore, LD is calculated using the following equation,

$$LD = A_{\parallel} - A_{\perp}$$

Where $A_{\parallel}$ represents absorbance of light parallel to the axis and $A_{\perp}$ represents the absorbance or light perpendicular to the axis.

The direction of polarisation (LD) can give an indication of the conformation of molecules when the orientation of flow of the molecules is known. A cuvette cell is used, in which the sample is placed in the gap between two concentric cylinders one of which rotates. Shear forces cause B-DNA to orient in the direction of the spin. Small molecules are not oriented (no LD signal) unless bound in specific orientation(s) to the DNA. B-DNA produces a characteristic LD spectrum, therefore, LD can be used to observe if B-DNA conformation alters due to the interaction of an external molecule. This is an extremely useful property as it enables researchers to identify whether potential DNA binders can successfully bind to B-DNA.

3.2.6) Linear Dichroism (LD) of pyridine-based cylinders

LD analysis of the novel longer helicate cylinder was completed in order to observe whether this cylinder was able to bind to B-DNA and consequently, induce conformational changes in the DNA strands. These further analytical tests were necessary

as analysis of the CD failed to provide any clear evidence that the cylinders were capable of interacting with ds-DNA.

$[Ni_2L^1_3][Ac]_4$ was titrated against 150 μL CT-DNA (100 μM) with the samples being analysed under rotation. ct-DNA alone is known to show a strong peak at 260 nm as the DNA bases align with the flow and this is seen on the graph below (Figure 3.2.6.1) for 0:1 indicating that the ratio of DNA base pairs to complex is zero, hence no complex has been added. Overall, on increasing additions of cylinder to the CT-DNA sample, there is a general increase in the magnitude in the peak at 260 nm, however this effect is inconsistent from each individual addition to the next. As seen in other LD studies completed by the Hannon group using an extremely similar method, if the complex was inducing conformational changes within the CT-DNA, there should be a discernible decrease in the magnitude of the peak at 260 nm with other peaks appearing within the spectrum.[11] However, these changes were not observed on additions of the $[Ni_2L^1_3][Ac]_4$ helicate to the DNA sample,. LD analysis of $[Ni_2L^1_3][BF_4]_4$ in 10% acetonitrile was also performed in order to ensure that the results observed were not due to instability of the cylinder in solution. The spectrum produced shared the same characteristic as the initial experiment; no significant changes in the characteristic ct-DNA peak were seen on addition of the helicate. The precipitation observed in the CD experiments, may partially explain the observations in LD.

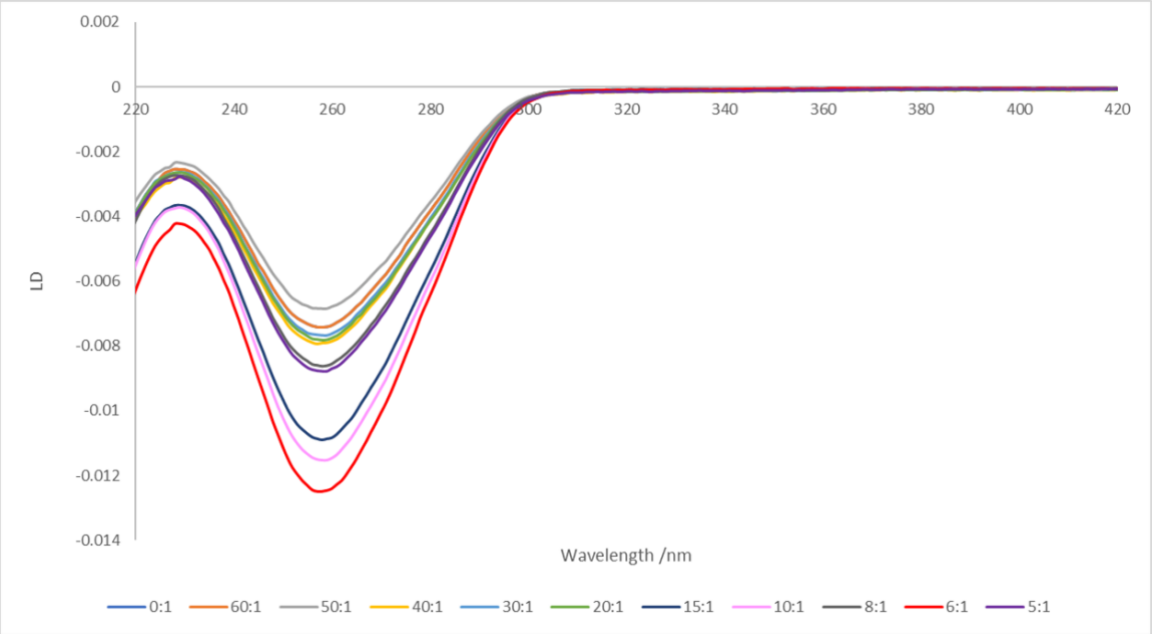


Figure 3.2.6.1 LD titration of $[Ni_2L^1{}_3][Ac]_4$ with 100 μM CT-DNA in water where the ratio is DNA base pairs to complex

Coupled with the CD data, the LD data does not provide strong evidence that these longer helicates are able to form a complex in solution with B-DNA, though they do seem to precipitate the DNA which does imply some binding or association. Initially, this conclusion was surprising as the parent cylinder developed by the Hannon group, from which these novel cylinders were developed, is able to bind to the major groove in B-DNA.[55] However, while the novel cylinders developed in this thesis shared extremely similar chemical properties to the Hannon parent cylinders, the novel cylinders are significantly longer (than the parent cylinder) due to ethoxy linker that they contain. Consequently, the increase in size, and thus the decrease in electrostatic attraction, makes it less favourable for the novel cylinder from binding to the major groove of B-DNA. The $[Ni_2L^1{}_3][Ac]_4$ is significantly larger than the parent cylinder but both complex have the same positive charge. Consequently, the $[Ni_2L^1{}_3][Ac]_4$ cylinder is less charge dense so

experiences a lower level of electrostatic attraction to the DNA, and therefore, is less likely to bind to the major groove.

3.2.7) Polyacrylamide Gel Electrophoresis (PAGE)

Polyacrylamide gel electrophoresis (PAGE) is an analytical technique that uses electrophoretic mobility, which is the capacity of an analyte to move in the direction in an oppositely charged electrode, to separate large molecules like proteins or DNA. The ability of macromolecules to travel through a polyacrylamide gel is affected by its charge, its molecular weight and its size, hence different molecules are differentiated by the distance they are able to advance down the gel. The macromolecules diffuse through the pores the gel, which is affected by the percentage of acrylamide used to make the gel. The larger the acrylamide percentage, the smaller the pore size, which in turn means the slower the molecules will move through the gel.

The gel itself is prepared by prepared by mixing 20 ml of 37.5:1 acrylamide/bis-acrylamide with 5 mL of 10x Tris-Boric acid buffer (890 mM each, pH 8.3) and 25 ml of deionised water. To this 400 μL of a 10% w/v ammonium persulfate/water solution and 40 μL of TEMED are added to initialise polymerisation. TEMED acts as a catalyst enabling the bis-acrylamide mix to polymerise in the presence of ammonium persulfate, creating a mesh-like structure. Wells within the gel, in which samples are eventually loaded, are created by a comb inserted into the gel as it sets. Once the gel is set, it is placed in an electrophoresis tank which contains a TB buffer which enables the gel to run under

constant pH conditions. As soon as the samples are loaded into the wells of the gel, an electric current is applied which induces the analyte to being to travel down the gel. When the current is removed, the analyte stops moving and the gel is then stained so that the results can be easily observed.

SYBR gold stain was used in the PAGE experiments within this research to visualise the gel. The stain enabled the DNA samples to be viewed as bands under UV light. If higher-order DNA structures were formed bands would appear higher up the gel as these structures would be larger than single strands of DNA and therefore, moved more slowly through the gel. In general, the larger the DNA structure, the higher its molecular weight and thus the higher up the gel a corresponding band would be observed.

3.2.8) Polyacrylamide Electrophoresis Gels of pyridine-based cylinders

The aim of the PAGE in this research was to observe whether the longer helicates developed were able to bind to and stabilise higher-order DNA structures, such as 3WJs, 4WJs and G-quadruplexes. Initial PAGE experiments were run using the first helicates synthesised, $[Fe_2L^1_3][BF_4]_4$ and $[Ni_2L^1_3][BF_4]_4$, however, since these were not soluble in water, requiring instead to first be dissolved in acetonitrile before being diluted, all further PAGE experiments were done with the water soluble $[Ni_2L^1_3][Ac]_4$. This was implemented in order to ensure that the results were not impacted by acetonitrile due to its aforementioned nature to disrupt biological systems. While the $[Ni_2L^1_3][Ac]_4$ cylinder was known to be less stable, it had shown reasonable stability in DNA solutions and therefore, any results could be assumed to be accurate.

The initial experiments that took place explored whether $[Ni_2L^1_3][Ac]_4$ was able to interact and bind to DNA 3WJs and these followed a very similar method to that used by the Hannon group for their PAGE experiments involving their gold pillarplex structure.[11] Three 14-mer DNA strands were exposed to increasing equivalents of the $[Ni_2L^1_3][Ac]_4$ helicate. At room temperature these strands can only be observed as single-stranded DNA, however, they are designed to form higher-order 3WJs when a compound binds to and stabilises them.[11] When this happens, a new band is observed in the PAGE gel along with the original band relating to the single stranded DNA. In these experiments, the 14-mer DNA strands were also exposed to the original Ni parent cylinder developed by the Hannon group as well as just the ligand that formed the novel helicate $[Ni_2L^1_3][Ac]_4$. The inclusion of the ligand was implemented as a control to see whether the ligand alone would be able to stabilise 3WJs, as if it was unable to do this then any new bands observed could be confirmed as a result of the presence of the helicate. The inclusion of the nickel parent was implemented as a positive control to confirm the experiment was running correctly. The nickel parent cylinder has been shown to successfully bind to 3WJs and induce an additional 3WJ band in the gel. If this does not happen, it indicates that the gel has failed for some reason.

Figure 3.2.8.1 shows the PAGE between the cylinder and DNA 3WJ. The positive control (lane 1) does indeed show 3WJ formation in the presence on nickel parent cylinder as anticipated. Lanes 3 – 6 show addition of 0.5, 1.0, 2.0 and 4.0 equivalents of $[Ni_2L^1_3][Ac]_4$ to the three strands of 14mer DNA.

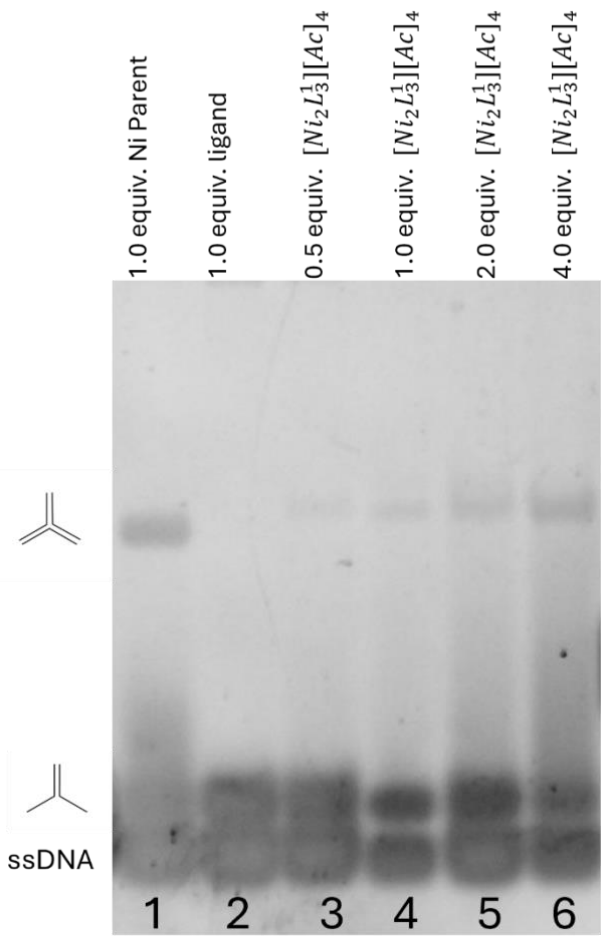


Figure 3.2.8.1 PAGE $[Ni_2L^1_3][Ac]_4$ with DNA 3WJ

It is clear from the higher up band seen on the gel in these lanes that 3WJs are also formed on addition of the $[Ni_2L^1_3][Ac]_4$ helicate, confirming that this cylinder is able to interact

with and stabilise this higher order structure. The band is retarded with respect to that of the Ni parent cylinder, which is consistent with the larger dimensions of the new complex which will protrude from the 3WJ increasing the hydrodynamic radius. The absence of any band relating to 3WJ formation in lane 2, demonstrates that the ligand on its own does not induce 3WJ formation and it is the cylinder complex that is responsible for the binding and stabilisation. This finding also removes concerns created by the known instability of the cylinder - it evidences that the cylinder is sufficiently stable under the conditions and time of the gel, to bind and stabilise the 3WJs. Interestingly, when comparing the results between the nickel parent and the novel cylinder, there is a clear difference in band intensity. 1 equivalent of nickel parent (as seen in lane 1) generates a corresponding 3WJ band with much stronger intensity than 1 equivalent of $[Ni_2L^1_3][Ac]_4$ (lane 4) and potentially shows even stronger intensity than 4 equivalents of $[Ni_2L^1_3][Ac]_4$ (lane 6). Since intensity approximately relates to concentration, it can be inferred that the nickel parent helicate is better able to stabilise 3WJs than $[Ni_2L^1_3][Ac]_4$ and this may be due to a couple of different reasons.

It is not unexpected that the nickel parent cylinder has a higher binding constant for 3WJs. Comparatively, the nickel parent cylinder is much smaller than $[Ni_2L^1_3][Ac]_4$ and, as discussed previously, has a diameter that fits perfectly into the cavity of 3WJs and the rings are able to stack perfectly with the bases in the 3WJ. It is probable that the larger $[Ni_2L^1_3][Ac]_4$ cylinder does not fit as perfectly into the cavity of 3WJs, and its charge density is lower with the metals further from the central binding site, so anion-cation electrostatic attraction may be reduced and the binding to 3WJs thus less strong. Alternatively (or in addition), the instability of $[Ni_2L^1_3][Ac]_4$ could cause some of the

cylinder to begin to fall apart during the PAGE process, thus reducing its overall concentration. As a result, less cylinder than expected could be available in each well to bind to the 14mer strands and thus, less 3WJs are generated. It is likely that both the reasons discussed are working in tandem to cause the reduced band intensity in comparison to the nickel parent. However, it is difficult to explore whether the cylinder's size is affecting its binding capabilities further through in vitro experiments. Computer generated simulations would enable this aspect to be investigated more deeply and the next chapter aims to do this.

After confirming the successful binding of $[Ni_2L^1_3][Ac]_4$ to 3WJ, potential binding of the cylinder to other higher-order DNA structures was investigated using PAGE. DNA 4WJs assemble from 4 complementary strands in a similar way to 3WJs. However, in contrast to the 3WJ experiments, the 4WJ is stable in absence of the cylinder, so the gel is not a simple on/off visualisation; instead gel shifts or changes in intensity are important. The complementary DNA strands that form 4WJs were exposed to $[Ni_2L^1_3][Ac]_4$ as well as

ligand and the nickel parent cylinder as control. The top band seen in Figure 3.2.8.2 is 4WJ.

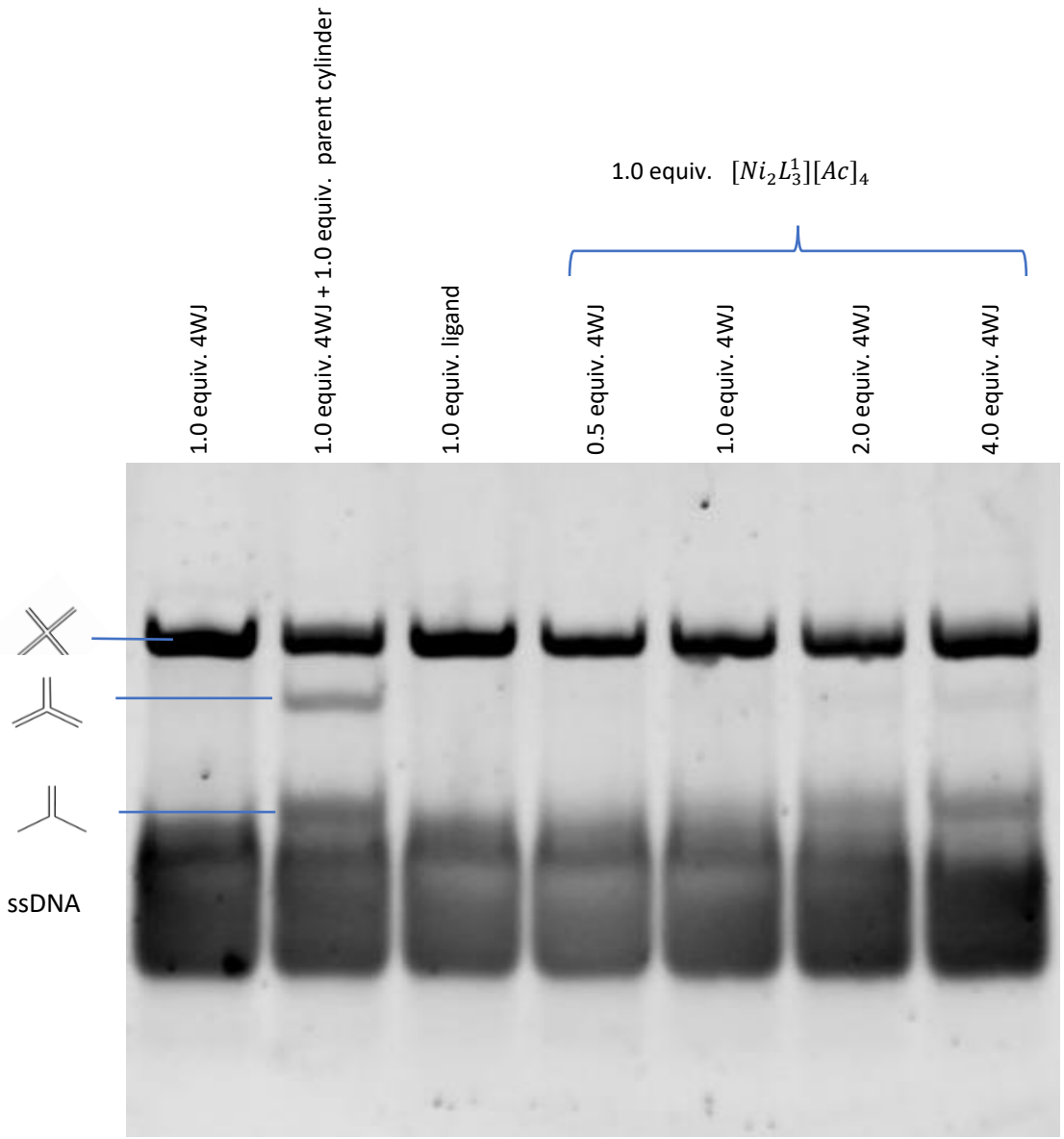


Figure 3.2.8.2 PAGE between DNA 4WJ and $[Ni_2L^1_3][Ac]_4$

The bound 4WJ has a slightly slower mobility than the unbound[11] so is retarded slightly higher up the gel as shown by the double banding in lane 2 by the nickel parent. This is interpreted as nickel cylinder binding the open cruciform form of the 4WJ, whereas in

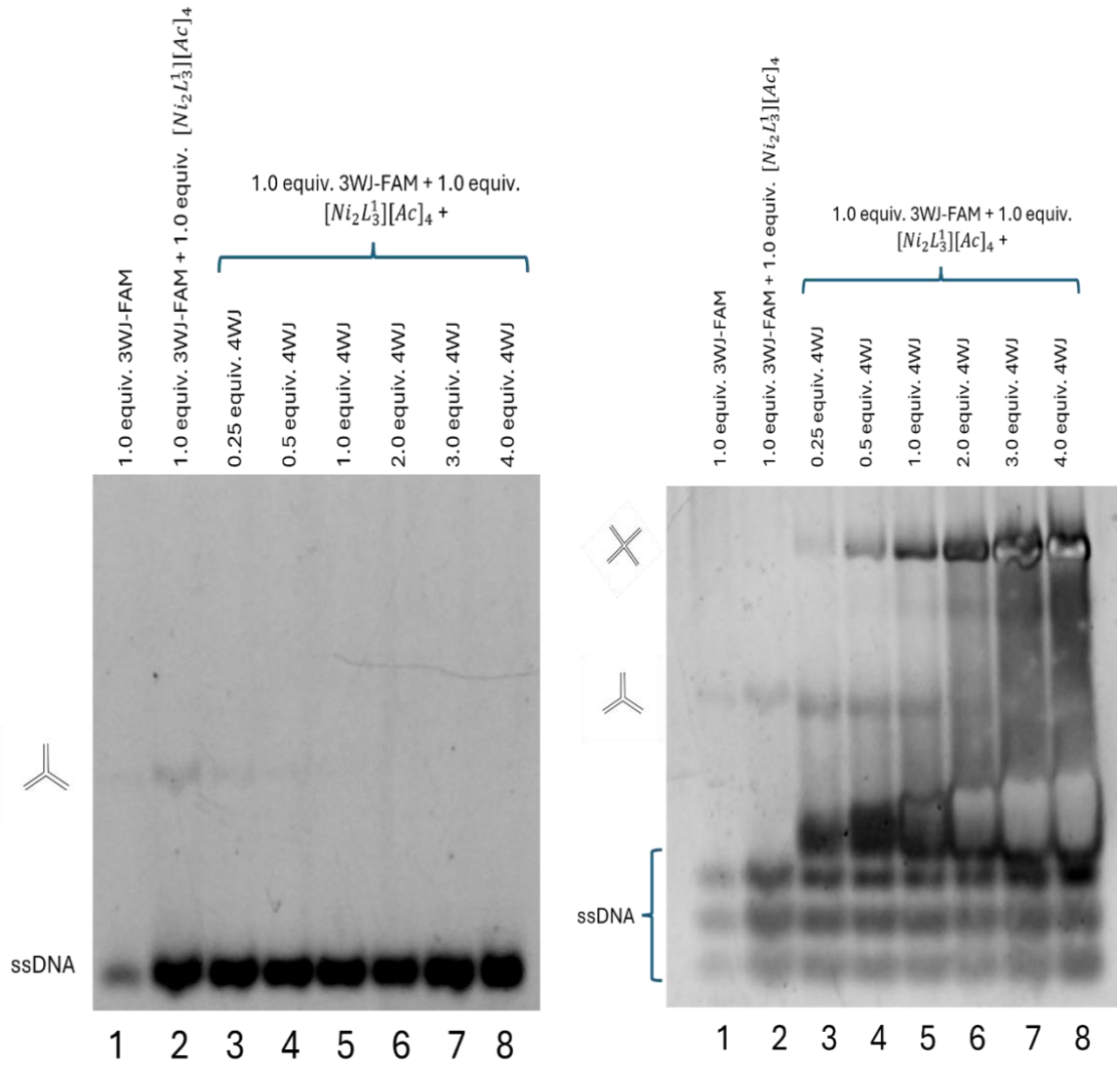
absence of cylinder the 4WJ will be in the closed arm-stacked X form. It is important to note that the strands also are potentially able to form pseudo-3WJs in the presence of a binder and, therefore, some 3WJ bands are also seen in lane 2 reflecting its strong preference for 3WJ. Finally, it can be seen that some Y-forks are also stabilised in the presence of parent Ni cylinder and these DNA structures have a higher mobility. There is some evidence from Figure 3.2.8.2 to suggest $[Ni_2L^1_3][Ac]_4$ might bind to 4WJs, as shown by the slight broadening of at 4WJ band at the top of the lanes 4 – 7 which contain equivalents of the novel cylinder, but there is no clear new band for a stabilised open-X 4WJ-conformation. At high loading, the new cylinder show subtle bands relating to 3WJs, providing further evidence that the novel cylinder is able to bind and stabilise these higher order 3WJ structures like the parent Ni cylinder.

Lane 3 shows no broadening of the 4WJ band suggesting that the ligand itself does not interact with these higher order structures and confirming that the bound 4WJs bands shown in lanes 4 – 7 are a result of the novel cylinder and not ligand. When comparing the nickel parent cylinder results in lane 2 with the novel cylinder, the results are highly similar. Both cylinders are able to bind to Y-forks and pseudo 3WJs, however, it is clear that the intensity of the 3WJ bands in the lanes 4-7 is lower, suggesting that the Ni parent cylinder has a greater ability to bind and stabilise pseudo 3WJs than the $[Ni_2L^1_3][Ac]_4$ does.

To probe whether the novel helicate had a binding preference for 4WJs over 3WJs, a competition gel was undertaken. A competition gel is where the DNA binder, in this case $[Ni_2L^1_3][Ac]_4$, is added to wells that contain one DNA structure at a constant

concentration (Constant) and another competitor DNA structure added in increasing equivalents, known as the Competitor. The competitor DNA structure competes with the DNA at constant concentration for binding of the cylinder. Since the Constant is fluorescently tagged, corresponding bands in the gel will show up without staining under UV light, therefore, by measuring intensity of the bands and plotting this against the concentration of the Competitor, the effect of the competitor on binding can be seen and binding preference of the complex can be established.

Using this concept, a competition gel between 3WJs and 4WJs using $[Ni_2L^1{}_3][Ac]_4$ was completed to investigate further whether the cylinder showed binding preference for 4WJs over 3WJs. In each well the concentration of 3WJ and $[Ni_2L^1{}_3][Ac]_4$ were kept constant while increasing equivalents of the competitor, 4WJ, were added. The left-hand image of Figure 3.7.8.3 shows formation of 3WJ in lanes 2-4, with the intensity of the 3WJ band decreasing across the lanes. Lanes 5 – 8, where the concentration of 4WJ was



higher, did not show any clear 3WJ formation. This showed that as concentration of 4WJ

Figure 3.2.8.3 Competition PAGE between 3WJ-FAM and 4WJ with $[Ni_2L_3^1][Ac]_4$. Gel on left is unstained while gel on right is stained

was increased, the formation of 3WJ decreased, confirming that the 4WJ competes with 3WJ for binding the helicate. Further evidence for this is provided by the right-hand image of Figure 3.2.8.4, which is the same gel stained with SYBR gold to reveal all the DNA species present. This shows increasing 4WJ formation and decreasing 3WJ formation across the lanes from left to right. However, these images alone cannot be used to conclude that $[Ni_2L^1{}_3][Ac]_4$ binds preferentially to 4WJ DNA structures over 3WJs as visually they are subjective. Therefore, further computational analysis of the images is needed in order to accurately confirm binding preference. This was done by plotting the percentage intensity of the 3WJ against the concentration of competitor in the well, the graph of which is shown below in Figure 3.2.8.4.

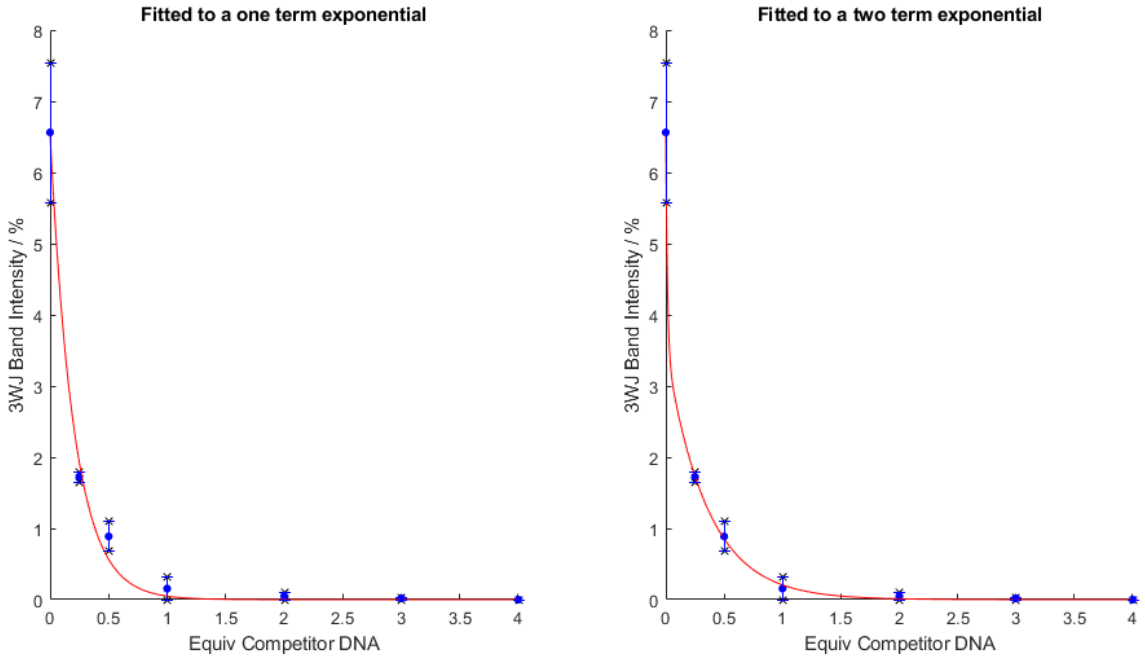


Figure 3.2.8.4 Intensity plot showing how 3WJ-FAM band intensity changes with addition of competing 4WJ. The error bars show the standard deviation across two repeats. The graph was fitted to both a one-term and two-term exponential

From the graph it can is clear that as 4WJ is added 3WJ formation is reduced. The effect of 4WJ can be evaluated by comparing the intensity percentage to the known concentration of 3WJ DNA. If the cylinder has no binding preference and in absence of competitor the 2WJ shows 5.5% intensity, at 1 equivalent of competitor there should be 2.75% intensity of 3WJ band (as percentage intensity is equivalent to concentration). However, from the graph, 1 equivalent of 4WJ corresponds to 0.31 % intensity of 3WJ band, suggesting that the complex binds preferentially to 4WJ.

Since the cavity of DNA 4WJs is slightly largely in size than 3WJs, it is likely that the cylinder preferentially binds to the 4WJs as it is able to accommodate the substantially sized cylinder better. However, since 4WJ is more anionic than 3WJ, electrostatics could be driving the preferential binding of the cylinder to 4WJ. Consequently, further computational analysis is required in order to accurately ascertain what may be causing the preference in binding.

A further competition gel between DNA 3WJ-FAM and duplex DNA with the helicate was conducted as shown in Figure 3.2.8.5. This was done to see whether the helicate was able to bind B-DNA, as the CD experiments had been inconclusive, and if so, did the helicate bind in preference to 3WJ. Both the unstained and stained gels show that with higher loading of duplex DNA there is no change in the band intensity of 3WJ suggesting that the presence of the duplex had no impact on the helicate binding to the 3WJ. Consequently, this suggested that the $[Ni_2L^1_3][Ac]_4$ binds more strongly to 3WJ than to duplex DNA.

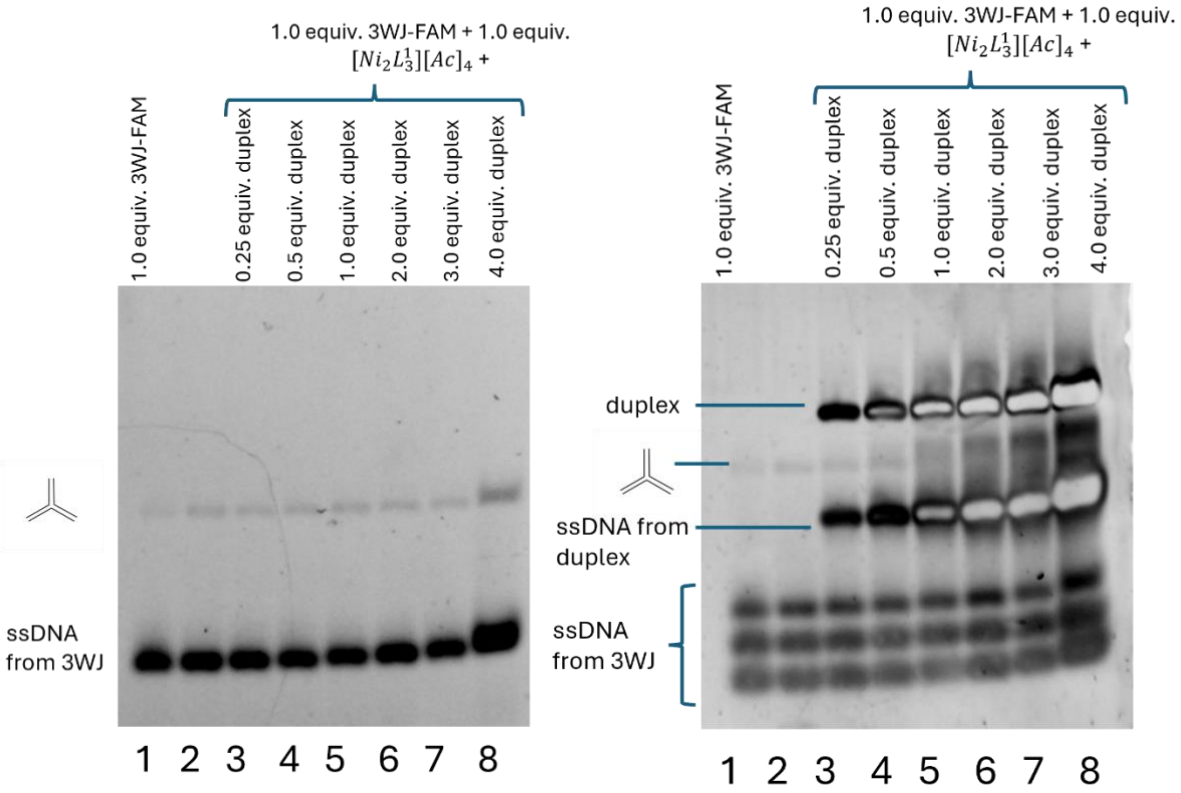


Figure 3.2.8.5 Competition PAGE between 3WJ-FAM and duplex with $[Ni_2L^1_3][Ac]_4$. Gel on left is unstained while gel on right is stained

The G4 quadruplex was the final higher order DNA structure investigated by PAGE as Hannon's parent cylinder previously showed the ability to bind to these structures.[38] Therefore, as the $[Ni_2L^1_3][Ac]_4$ helicate has close similarities to the parent cylinder it is not unlikely that it could display the same binding capabilities. Since the aims of this paper is to investigate the potential of $[Ni_2L^1_3][Ac]_4$ to bind to higher-order DNA structures in order to investigate whether the helicate has potential as an anti-cancer agent, the PAGE experiments were completed with H-telo and C-MYC G-quadruplexes as these are known to be associated with cancer. As with the previous PAGE experiments, H-telo and C-MYC were exposed to increasing equivalents of the cylinder. The gels were completed using Tris-borate (TB) (with 20 mM KCl) buffer and a water jacket, which enabled cold water to pass over the gel keeping it cool. These conditions were essential as K^+ ions are required in order to enable successful G4 formation, however, the movement of the K^+ ions to do this increases the current, which in turn increases the temperature of the gel. Therefore, by using the water jacket the temperature of the gel is kept constant throughout the experiment.

The gel between the G4 structure, H-telo, and the $[Ni_2L^1_3][Ac]_4$ cylinder showed one band only which related to single stranded DNA, suggesting that the cylinder was unable to interact and bind to the G4 structure and hence no H-telo structure was formed. However, the gel between the C-MYC G4 structure and the cylinder (Figure 3.2.8.6) was much more successful. Lane 1 shows that in the absence of complex, two bands are observed, which was interpreted as both the C-MYC and the C-MYC dimer being present (with the dimer being a minor fraction of the total). The dimer is not a stack of two monomers but an intertwined assembly of 2 DNA strands and it is favoured by high K+ or Na+. When

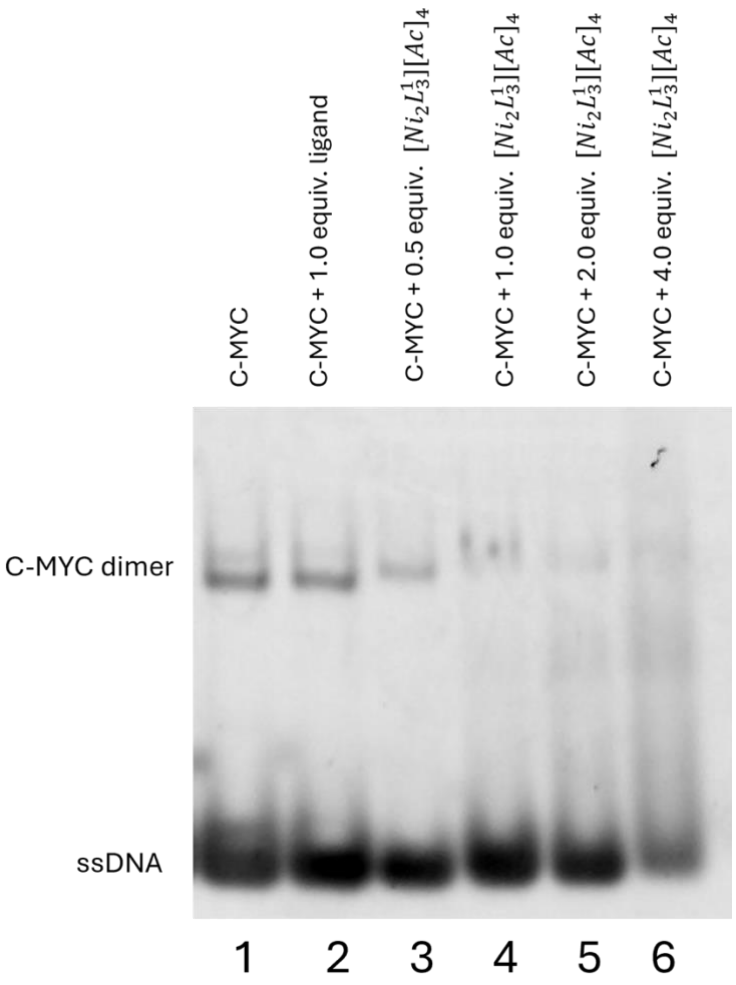


Figure 3.2.8.6 PAGE between C-MYC G4 structure and $[Ni_2L^1_3][Ac]_4$

complex is added there is a dramatic loss of intensity in the C-MYC dimer band, and a possible retardation of that band. It is difficult to assess whether there is retardation due to the loss of intensity, but if retardation is present it could indicate a binding of cylinder to the dimer. The loss of intensity could be a result of either the dimer being de-stablised by cylinder or because the cylinder binding stops the stain from binding or quenches the fluorescence of the stain. The former seems more likely and might imply that the cylinder binds and stabilises monomer preferentially over dimer, however, it is hard to exclude latter from this experiment.

When complex is added there is, at higher loading (lanes 5 and 6), both a general smearing and possible formation of a small amount of a band running between the monomer and dimer. It is unclear what this might be. At high loading (lane 6) there is a loss of staining in the C-MYC monomer band but also more smearing. This may suggest that at high loading the cylinder is binding the monomer - perhaps on both sides of the G4 stack - and preventing effective staining by the SYBR gold dye.

As the $[Ni_2L^1_3][Ac]_4$ showed the ability to interact with C-MYC G4 quadruplexes, a competition gel between C-MYC DNA and 3WJ-FAM with the helicate was completed in order to ascertain whether the helicate showed a preference to either DNA structure. The right-hand gel in Figure 3.2.8.7 shows the unstained gel. In lanes 3 – 6, as the loading of C-MYC DNA increases there is little change in band intensity relating to DNA 3WJ, while at high loadings of C-MYC in lanes 7 – 8, there is a very slight decrease in intensity of the 3WJ band and a slight retardation of the band (this is more pronounced in the

stained gel). This indicates that while $[Ni_2L^1_3][Ac]_4$ is able to interact with C-MYC DNA it will preferentially bind to 3WJs.

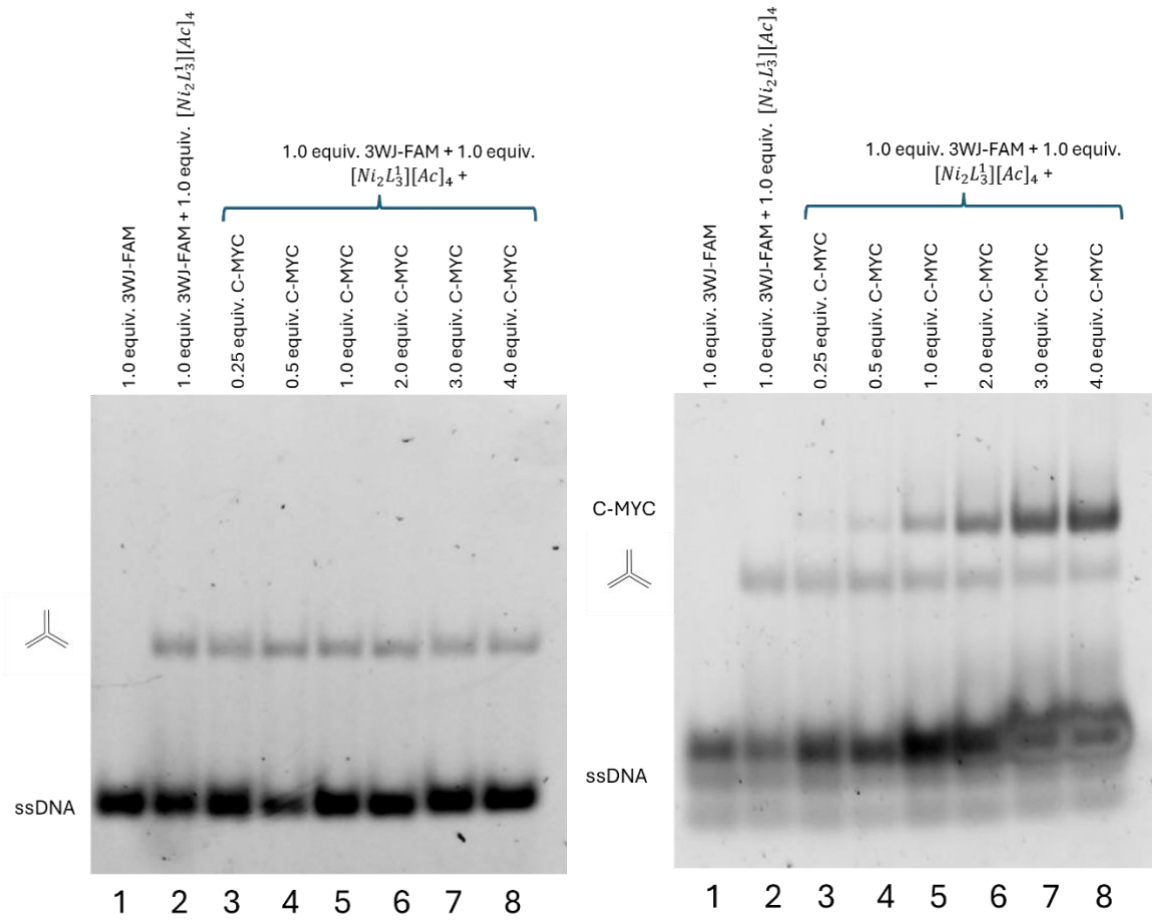


Figure 3.2.8.7 Competition PAGE between 3WJ-FAM and C-MYC G4 quadruplex with $[Ni_2L^1_3][Ac]_4$. Gel on left is unstained while gel on right is stained

Overall, it can be concluded from these PAGE experiments that cylinder is able to bind to 3WJs, 4WJs and potentially also (more weakly) to C-MYC G-quadruplexes. The potential preferential binding of the cylinder to 4WJs over 3WJs (observed in the competition gel) is likely due to the larger cavity size of the junction, making it

entropically preferentially for the cylinder to bind, however, this preference could also be driven by electrostatics.

3.3) In silico

The experiments carried out in vitro provided clear evidence that the novel cylinder, $[Ni_2L^1_3][Ac]_4$ in able to interact with the higher-order DNA structures, specifically DNA 3WJs, 4WJs and potentially even C-MYC G4 quadruplexes. However, these experiments are not capable of elucidating how the helicate might interact with these higher-order structures. Therefore, in silico analysis of the cylinder with these DNA structures is required in order to establish exactly how the helicate binds to the DNA. It should be noted that while the gels analysed the binding mechanism between DNA and $[Ni_2L^1_3]^{4+}$, the simulations will analyse the interaction of DNA with $[Fe_2L^1_3]^{4+}$. This is a result of the parametrisation method requiring simulation data from previous similar simulations, of which only the Fe version of the cylinder was used. Although this is not ideal, it is likely that this change in metal ion has little effect on the simulations themselves since we know that the Fe and Ni complexes of the original 'parent' ligand bind identically to DNA.

3.3.1) Parameterisation and Optimisation of Novel Helicate

The crystal structure of the novel iron (II) cylinder was generated using the crystal structure of a similar larger helicate that was previously synthesised by the Li group.[87] This involved converting the thiol groups in the ligand to a pyridine group using the software AVOGADRO. Initially, an attempt to parameterise the cylinder was made using

the MCPB.py pipeline as detailed by Hannon group.[101] The calculations failed to reach convergence meaning parameterisation could not be completed - the primary reasons this could happen is that a) there cylinder is large and therefore computationally expensive to parameterise, b) the initial starting geometry is too far away from the energy minimum geometry. Consequently, it was determined that the cylinder was too large for DFT parameterisation at the ωb97xd / DEF2SVP level of theory. However, this finding was not unexpected as although the Hannon group had previously had success parameterising the parent cylinder to this level of theory, the novel cylinder was significantly larger so has many more electrons for DFT to calculate. Moreover, this inability to parameterise the cylinder at this level of theory presented a significant problem for the eventual generation of the simulations as the cylinder would need to be parameterised at a comparable level to the forcefield used to generate the DNA crystal structures.

In order to try and overcome the inability of the MCPB.py workflow to parameterise the novel helicate, an attempt was made to optimise the cylinder using the software ORCA at B3LYP level of theory. The B3LYP functional was chosen due to its popularity and generate a zig-zag like structure for the cylinder. However, the software struggled to accurately calculate the energy frequencies of the cylinder due to its size. Still using ORCA, the following attempt at parameterising the cylinder involved first optimising at MO6 level, as it was known to be good for transition metals, then subsequently feeding this output into a second optimisation at the ωB97XD level.[102] This method was successful and parameterised the cylinder with no imaginary frequencies, indicating the structure was on an energy minimum. However, once the geometry was fed back into

MCPB, the software was unable to generate the topology file for the structure indicating that the attempt was unsuccessful.

The main issue that caused all these attempts to fail was the size of the cylinder. Due to its extensive size, two separate optimisation softwares, ORCA and MCPB, were unable to parameterise it properly. Therefore, to successfully optimise the cylinder, a simplification of the structure that reduced its overall size was needed. Discussions with the Hannon group simulation lead, Larry Melidis, lead to the hypothesis that the parameters for the parent cylinder could be used in the optimisation of the novel cylinder $[Fe_2Py_3][BF_4]_4$ as both helicates contained a pyridine to metal bonding system (Figure 3.3.1.1). The first step of this proposed method was to parameterise the parent cylinder modified to include an oxygen in the linker over CH_2 as this mimicked the oxygen

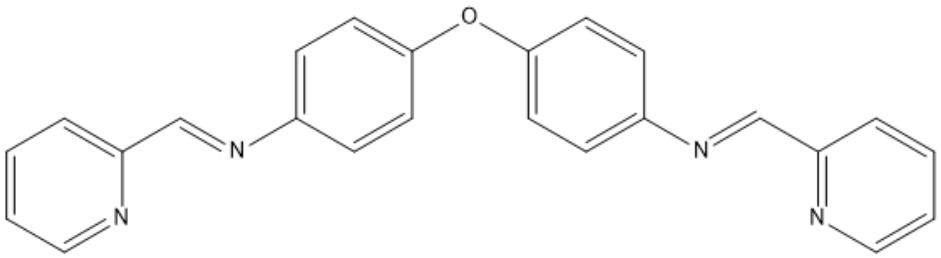


Figure 3.3.1.1 Structure of the Hannon Parent Cylinder Ligand

conjugated in the larger novel cylinder. This modified parent cylinder was successfully geometry optimised in ORCA at ωB97XD/DEF2SVP level of theory. The optimised structure was then fed into MCPB.py, which generated Forcemod files that enabled the bond lengths, bond angles and dihedral angles of the cylinder to be defined.

The next step of this method was to optimise the geometry of the larger novel cylinder in ORCA to calculate the distribution of the partial atomic charges (also known as the Mulliken charges) and to generate MOL2 files. For the $[Fe_2L^1_3]^{4+}$ cylinder, the expected overall charge is 4+, with each of Fe atoms having a charge of approximately 0.6 with the rest of the charge being distributed along the ligands. As the geometry optimisation of the $[Fe_2L^1_3]^{4+}$ cylinder calculated an overall charge of 4+, it suggested the optimisation in ORCA of the larger cylinder had worked correctly. The subsequent phase involved transferring the charges taken from the $[Fe_2L^1_3]^{4+}$ cylinder optimisation and the bond lengths, angles and dihedral angles taken from the modified parent cylinder parameterisation to the $[Fe_2L^1_3]^{4+}$ cylinder. First, the Muliken charges were copied from ORCA and pasted into the MOL2 files. Then, the MCPB generated Forecmod file from the modified parent cylinder was used to redefine the atoms in the MOL2 files, with all other atoms being defined by the general amber force field (GAFF). Amber T-leap was used to generate the Amber topology file from the adapted MOL2 files, which incorporated the Forecmod file from the modified parent cylinder. Finally, the Amber topology file was converted to the GROMACS topology file using the ParmEd function, resulting in the successful parameterisation of the $[Fe_2L^1_3]^{4+}$ cylinder. It should be noted that above methodology details how the *M* -enantiomer of the helicate was optimised. In order to generate the parameterised *P*-enantiomer, the chirality of parameterised *M*-enantiomer was inverted in AVOGADRO by collegue, Sam Dettmer.

3.3.2) Parameterisation of DNA

Several DNA structures have already been parameterised using the Hannon group's method including DNA 3WJ (one larger version and one smaller version), DNA 4WJ and duplex DNA and so could be used in this research.[11] The smaller 3WJ was prepared from PDB 2ET0, with the compounds removed.[6] The large 3WJ was adapted by colleague Sam Dettmer from PDB 1F44 as reported previously.[26] The extra solvent and protein were removed from the 3WJ crystal structure files before the backbone of the structures were joined up using Avogadro to create 3WJs formed of continuous DNA strands. Finally, these 3WJ structures underwent energy minimisation in GROMACS to produce the final structures used in the molecular dynamic (MD) simulations. The 4WJ was parameterised similarly to the 3WJs, where the crystal structure, taken from PDB 1XNS, had all water molecules and peptide strands removed before each arm strand was shortened to 9 base pairs as this reduced the modelling cost.[103] Again, this structure was energy minimised in GROMACS to produce the final structure used in the MD simulations. It is important to note that all these DNA structures were parameterised using parmbsc1, an AMBER forcefield. [104]

3.3.3) Molecular Dynamics (MD) method

The molecular dynamics simulations method used in this research closely follows that detailed by Larry Melidis in his 2021 paper.[105] In the initial simulations, the cylinder was placed in the centre of the cavity in either the 3WJ or 4WJ, while in the later simulations the cylinder was placed directly outside of the junctions. The system (the cylinder and DNA) was then placed in a dodecahedral box of 1.0 nm size and solvated using the TIP3P

water model. A 50 mmol concentration of NaCl ions was added to neutralise it. Subsequently the system was minimised such that the maximum force on any one atom was less than 500 kJ mol^{-1}, followed by NVT and NPT equilibration using V-rescale modified Berendsen thermostat at 310 K and for 1000 ps each, in which the cylinder was coupled with the DNA. The simulations were then run using a time-step of 2 fs. This procedure was also used to model the un-complexed (free) DNA by not including the parameters of the cylinder. Finally, the compressed trajectories were processed in GROMACS to remove rotational and translational motion, and to centre the DNA within the box for visualisation and analysis purposes, which were done in PYMOL.

3.3.4) Root Mean Square Deviation (RMSD) Calculations

RMSD calculations are used to show how much a system diverges from a known reference structure and is often used to investigate the stability of a binding interaction. Large conformational changes are expressed as spikes on the RMSD plot and, therefore, can indicate if an interaction is unstable. In this research, RMSD plots were used to explore how energetically stable the interactions between the novel cylinder and the higher-order DNA structures were and consequently, evaluate whether the novel cylinder exhibited a binding preference to a particular structure. The plots were calculated in GROMACS using a reference and trajectory file for each system.

3.3.5) Molecular Dynamics (MD) Results and Analysis

With the aim of gaining a comprehensive understanding of how the presence of the cylinder impacts the higher order DNA structures, simulations of unbound 3WJs, 4WJs and duplex DNA were generated. This enabled the performance of these free DNA structures investigated so that a direct comparison to the bound forms could be made. Following this, simulations where the novel cylinder was placed directly inside the junctions as well as directly in-front of them were produced. Further simulations where the cylinder was placed in close proximity to duplex DNA were also generated. By introducing the novel cylinder to the DNA structures in different positions, a comprehensive view of behaviour of the cylinder on simulation timescales was enabled.

3.3.5.1) 3WJ and $[Fe_2L^1_3][Ac]_4$

Unbound DNA 3WJ was modelled across 6 individual simulations, totalling 6μs of simulation. 3 of the simulations modelled the smaller DNA 3WJ which came from 2ET0 PDB file and 3 simulations modelled the larger DNA 3WJ that showed base pairing of the chains further away from the junction. The simulations generated by larger DNA 3WJ came from the PDB file 1F44.

The smaller 3WJs (Figure 3.3.5.1) are formed from three identical DNA chains made up of 6 bases with the sequence 5'CGTACG3', therefore, only show the base pairing at the centre of the 3WJ cavity and a couple bases along the chain. All three of the simulations were dynamic with the Watson-Crick base pairings at the centre of the junction fraying and reforming as the DNA chains wiggle. Two of these simulations showed one of the DNA chains flipping completely then reforming the junction, while the final simulation showed the end of two chains fray completely to leave structure with a half-formed cavity.

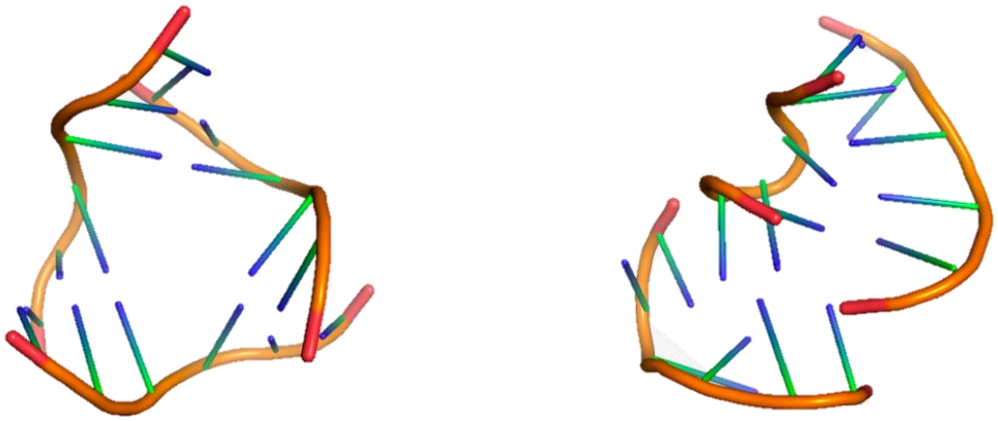


Figure 3.3.5.1 Molecular dynamic simulation of unbound small 3WJ. Image on right shows the 3WJ at the start of the simulation and the image on the left shows the 3WJ at the end of the simulation (1 μs)

The larger 3WJ (Figure 3.3.5.1) is formed from longer ssDNA (14 bases long), therefore, shows the DNA structure further along the chain from the cavity. The 3 simulations formed from this larger structure were also shown to be dynamic, the base pairing at the junction, $A_1$8 – $T_2$7, $A_2$8 – $T_3$7 and $A_3$8 – $T_1$7, continuously fray and reform throughout the simulation. However, in comparison to its smaller counterpart, the larger 3WJs were much less dynamic due to the extra length of the DNA chains, which allows for more Watson-Crick base pairing along from the junction, providing stability to the overall structure. It is important to note that for both free 3WJs the simulations show that the

cavity closed up. While this can indicate the unstable nature of these unbound structures, it is also partially an artefact of the molecular dynamic forcefield used in these simulations. As this forcefield favours pi-stacking, it causes the open conformation of the 3WJ to rapidly close into a stacked conformation – the re-opening of the 3WJ is not observed on the MD timescale. Therefore, while there is some evidence to indicate the unbound 3WJ is unstable in nature, this cannot be fully attributed to its conformation.

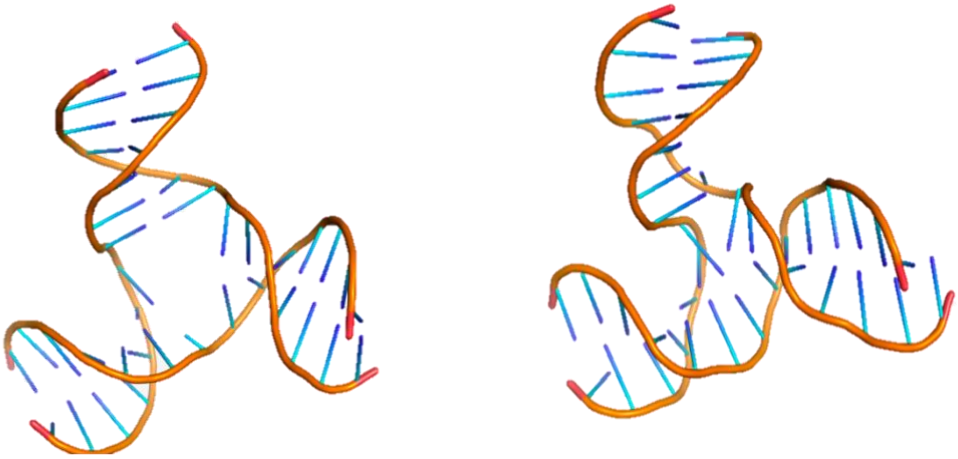


Figure 3.3.5.1 Molecular dynamic simulation of unbound large 3WJ. Image on right shows the 3WJ at the start of the simulation and the image on the left shows the 3WJ at the end of the simulation (1 μs)

In order to simulate the interaction between the novel cylinder and the small/large 3WJs, both the *M*- and *P*-enantiomers of cylinder were parameterised and introduced into the molecular dynamic models. Initially, 9 individual simulations totalling 9μs of simulation were generated, 6 from the larger 3WJ where both *M*- and *P*-enantiomers were placed at the centre of the junction (so that the biphenyl group within the cylinder was at the heart of the junction) and 3 from the smaller 3WJ, where only the *M*-enantiomer was used. A further 6 simulations were generated using the *M*- and *P*- enantiomers, where the cylinder was placed outside of the junction.

For both the large and small 3WJ, when the *M*-enantiomer of the cylinder was placed at the centre of the 3WJs, the dynamic movement of the chains seen in the free DNA structures was reduced significantly, suggesting that the cylinder had a stabilising effect on the 3WJ. While the cylinder remained within the cavity of the 3WJs for the entirety of the simulation time, it did move from being situated centrally within the junction, to having one end bound within the cavity. On further analysis of the larger 3WJ, it was clear that this was a result of how the cylinder interacted the DNA bases within the cavity in order to bind to the junction (Figure 3.3.5.2). The simulations showed that 3 permanent pi-stacking interactions form between the phenyl rings (bound to a nitrogen and oxygen) at one end of the ligands and the T7 bases on each DNA chain, each of which will have a stabilising effect on the junction. Moreover, 3 temporary pi-stacking interactions formed between linker phenyl rings of the ligand and the A8 bases on the junction enables one end of cylinder to provide stabilising interactions to the junction, by binding to the bases that were continuously forming temporary interactions in the free form of the DNA. Consequently, the presence of the cylinder in these simulations was shown to stabilise the open form of the 3WJ.

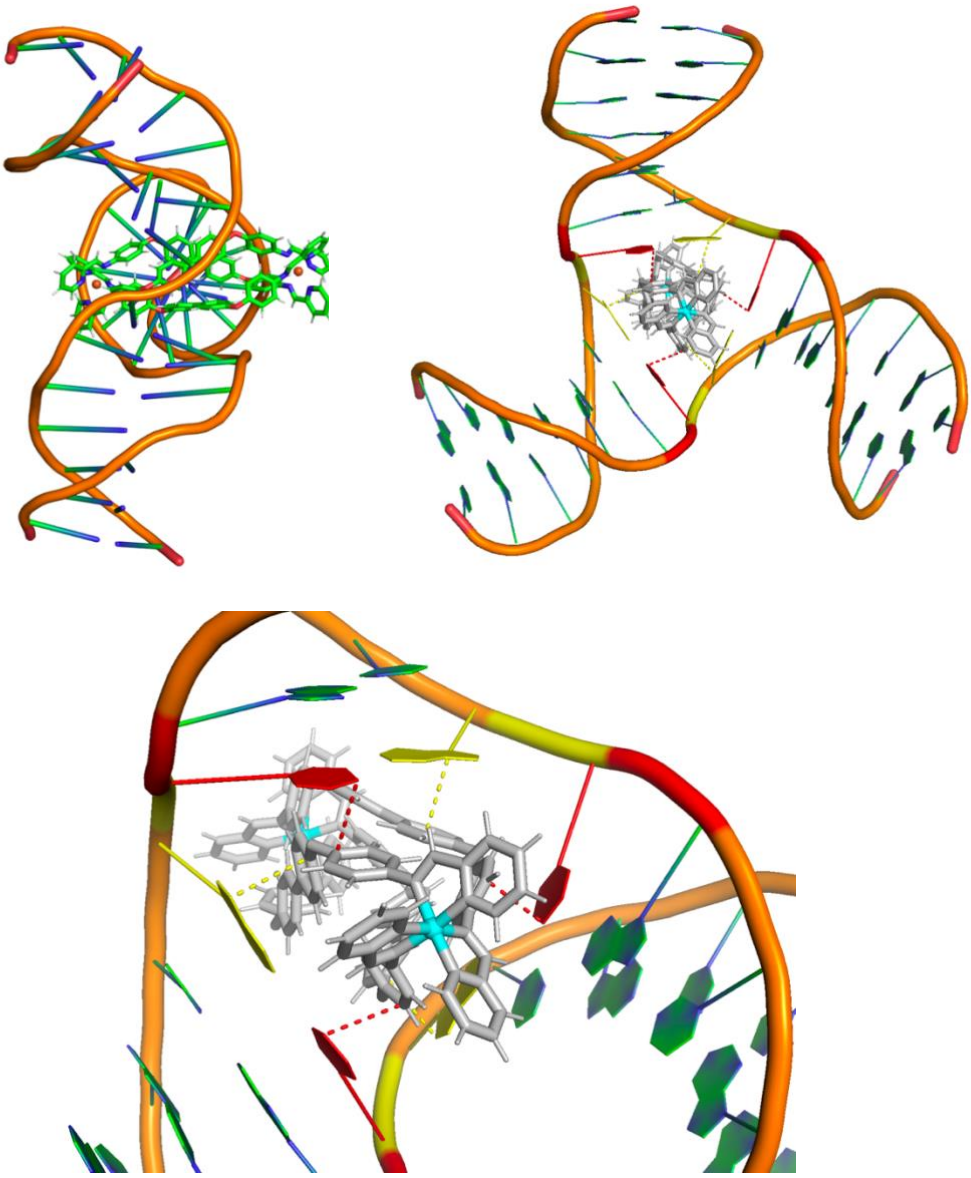


Figure 3.3.5.2 M-enantiomer bound within DNA 3WJ. Top right-hand image and bottom image show pi-stacking interactions between the bases of the DNA and the phenyl rings in the cylinder backbone. The dotted red lines show the permanent pi-stacking interactions and the dotted yellow line show the transient pi-stacking interactions

Similarly, to the *M*-enantiomer, the *P*-enantiomer of the cylinder was placed at the centre of both the large and small 3WJ as shown in Figure 3.3.5.3. Its presence at the junction, significantly reduced the movement of the 3WJ comparatively to the free structures, suggesting that the cylinder had a stabilising effect on the junction. As seen with the *M*-enantiomer, during each simulation the *P*-enantiomer moved from its initial central

position so that one end of the cylinder remained within the junction. This movement enabled the phenyl rings within the ligands of the cylinder to form stabilising pi-stacking interactions with the DNA bases at the centre of the 3WJ, which in turn stabilised the overall open form of the 3WJ (in an analogous process to that of the *M*-enantiomer). However, further analysis of the larger 3WJ simulation showed that the *P*-enantiomer formed a greater number of permanent pi-stacking interactions (6 vs 3) with the DNA than the *M*-enantiomer did. Both the A8 and T7 bases on each DNA chain were able to form permanent pi-stacking interactions with the phenyl groups from ligands, therefore providing six stabilising interactions at the cavity centre. Moreover, due to the position of the rings, the A8 bases were able to form temporary pi-stacking interactions with the linker phenyl groups of the ligand, which in turn provided more stability to the 3WJ structure. These findings indicate that for the $[Fe_2L^1_3]^{4+}$ helicate, the *P*-enantiomer has a better fit for the 3WJ cavity than the *M*-enantiomer, which directly contrasts the findings of previous studies that investigated the binding of smaller helicates to 3WJs.[57–59]

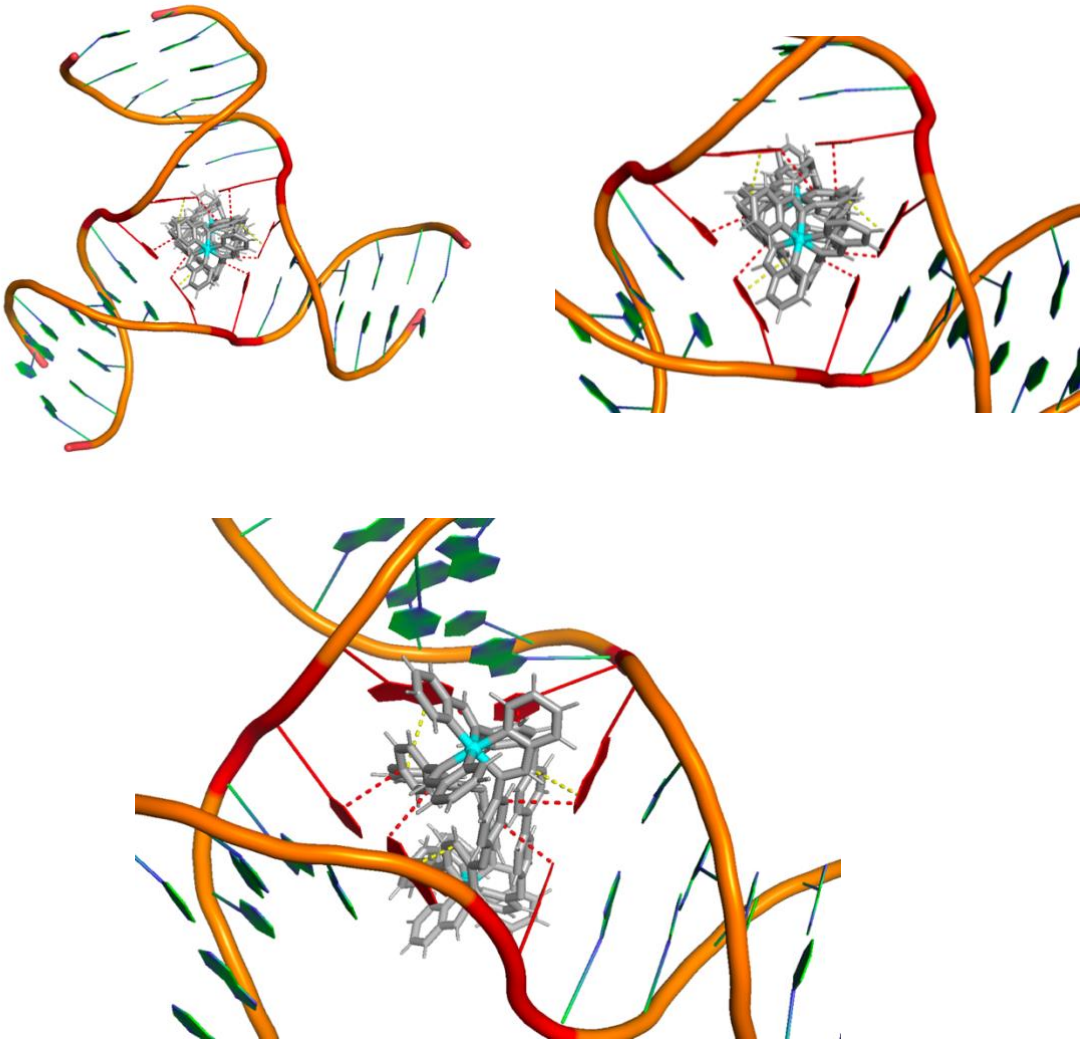


Figure 3.3.5.3 P-enantiomer bound within DNA 3WJ. Top and bottom images show pi-stacking interactions between the bases of the DNA and the phenyl rings in the cylinder backbone. The dotted red lines show the permanent pi-stacking interactions and the yellow dotted lines the transient pi-stacking interactions

It is likely that the larger size of the novel helicate is what causes the deviation of previous findings as only one end of this helicate binds to the 3WJ while with smaller helicates, the whole cylinder is situated within the junction cavity. Therefore, the difference in interaction could cause the difference in enantiomer affinity for 3WJs, however, this can only be assumed. To assess the difference in binding affinities of the two enantiomers further, in the following simulations the cylinder was placed outside of the junctions so that its interaction with the DNA could be modelled more realistically.

Three simulations were generated where the *M*-enantiomer was placed just outside the large 3WJ. In two of these simulations, the cylinder was shown to move away from the junction and interacted with the duplex DNA ends of junctions. The first showed the cylinder forcing itself between the DNA chains, breaking the base pairs, in order to bind to the DNA, whereas second showed the cylinder interacting with A1 – T14 base pair at the end of the junction. However, the remaining simulation showed movement of the cylinder into the junction (Figure 3.3.5.4), suggesting that the cylinder does have natural affinity to bind to the junction even when not placed directly inside. To analyse the movement of the cylinder further, the first 80ns of the simulation where the cylinder binds to the junction was investigated. This short timeframe showed that the cylinder approached the junction from an angle, initially binding parallel across the cavity, before threading through the cavity completely. From this central position within the cavity, the cylinder than moved back outwards so that only one half of it remained inside the junction and the system stayed in this conformation for the remaining duration of the simulation. It can be suggested from this observed simulated interaction of the *M*-enantiomer with the 3WJ, that while the enantiomer has some affinity for binding the 3WJ, it is conformationally challenging for the system to reach its most stable configuration hence the extensive movement of the cylinder within the cavity. Once more, this could be a consequence of the large size of the cylinder. The diameter of the novel helicate is approximately 10 Å while the diameter of 3WJs is estimated to be between 10 – 15 A, therefore, it is likely that cylinder may have to force the cavity wider in order to bind resulting considerable conformational change until the most stable conformation is reached.

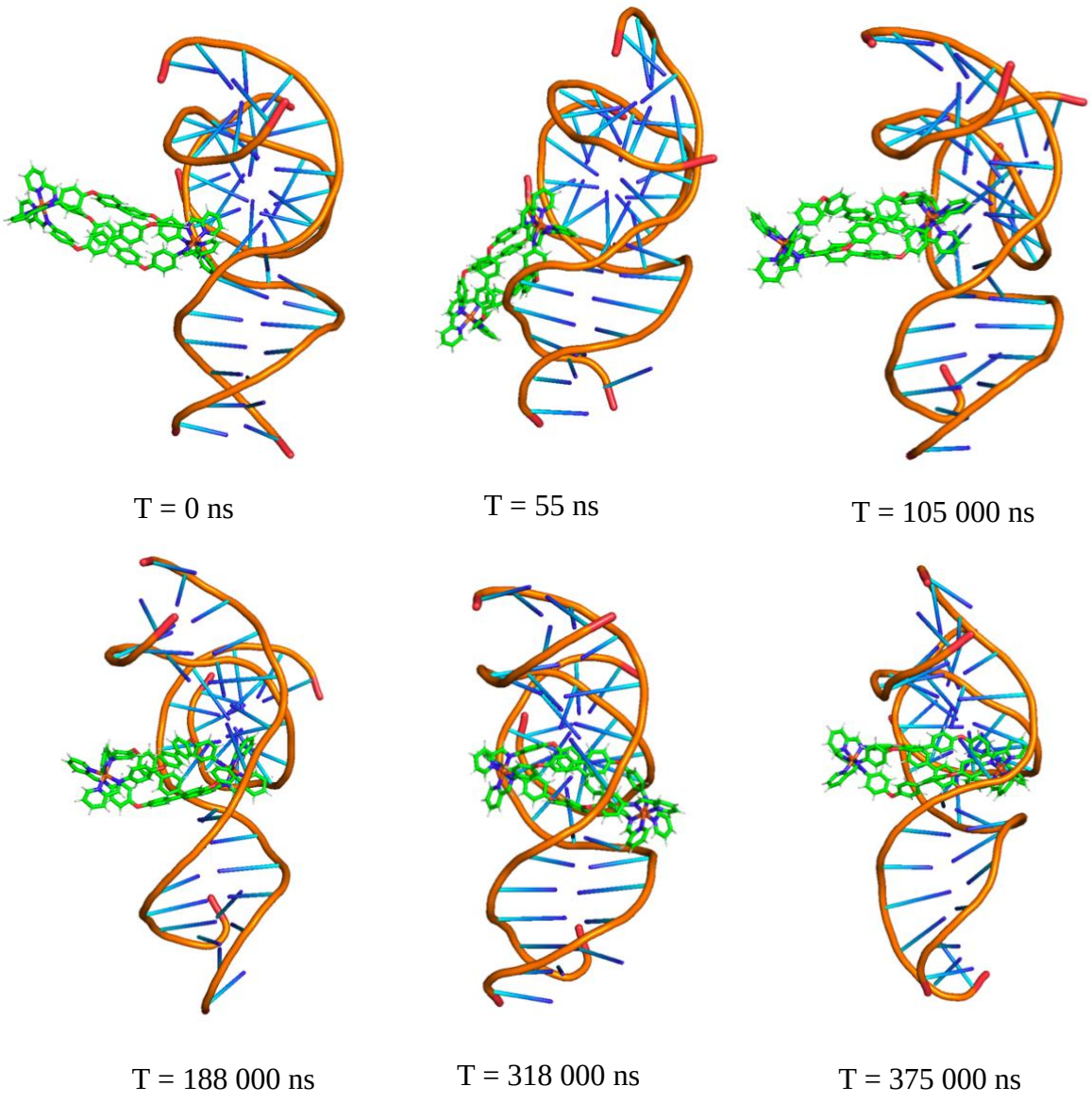


Figure 3.3.5.4 shows the movement of the M-enantiomer into the 3WJ

This hypothesis is supported by the RMSD data, shown in Figure 3.3.5.5. Initially, the plot shows large fluctuations in the graph indicative of conformational changes, however, at approximately 2000 ns the graph stabilises suggesting that the most energetically favourable configuration has been reached.

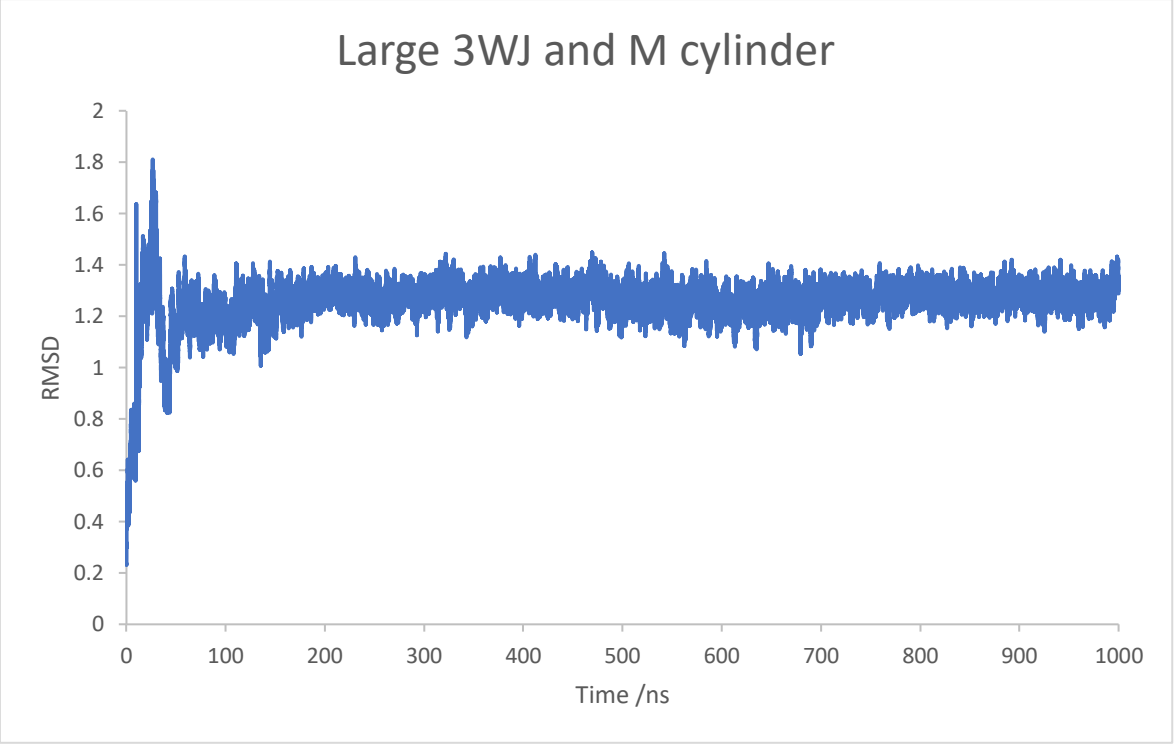


Figure 3.3.5.5 shows the Root Mean Square Deviation (RMSD) plot of how the large 3WJ M-enantiomer move/interact against time (ns)

The *P*-enantiomer behaved similarly to *M*-enantiomer when placed outside the large 3WJ. In two simulations the enantiomer was shown to move to the end of the of the junction strands and interact exclusively with the bases A1-T14. However, in the remaining simulation the enantiomer was observed to move into the junction. Once more the first 80 ns of the simulation was analysed. Unlike, the 1F442M simulation, the *P*-enantiomer approached the junction perpendicular to the DNA rather than at an angle. The *P*-enantiomer then proceeded to slowly thread into the junction until one end of the cylinder was bound with the cavity. While this process involved much less movement of the cylinder, there was substantial movement of DNA junction, stretching in order to allow the cylinder to enter the junction before reforming. The slight resistance by the DNA to the cylinder entering the junction is once more suggestive of the hypothesis that the larger

size of the cylinder causes some volumetric strain on the DNA as the diameter of the junction cavity is not quite large enough to accommodate the cylinder. Nonetheless, binding of the *P*-enantiomer to the 3WJ occurred more easily and rapidly than the *M*-enantiomer, again indicative of the binding preference of the junction

Overall, the simulations indicated that both the *M*- and *P*-enantiomers of the $[Fe_2Py_3]^{4+}$ cylinder showed an affinity for binding DNA 3WJs despite the large size of the cylinder. Interestingly, the simulations further suggested that the 3WJ may have a binding preference for the *P*- enantiomer over *M*-enantiomer, which was unexpected as the literature reported that the *M*-enantiomer of all other triple stranded helicates was bound preferentially.[57–59] Although it is unclear what causes this binding preference, it could be due to the significant size of this novel helicate.

3.3.5.2) 4WJ and $[Fe_2L^1_3]^{4+}$

Unbound DNA 4WJ, shown in Figure 3.3.5.6, was modelled across 3 individual simulations, totalling 3μs of simulation with all of the simulations being generated from the 1XNS PDB file. Each of the free simulations showed the 4WJ in its stacked, closed form with no cavity being present between the central base pairings, A9 – T8 on all DNA strands.

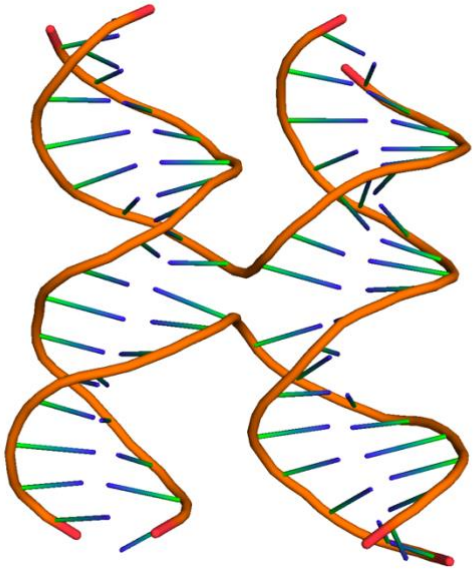


Figure 3.3.5.6 Structure of the unbound 4WJ in its stacked (closed) form

Furthermore, the simulations all showed dynamic movement of the closed conformation, with structure twisting and rotating, indicating its instability in its unbound form. However, it must again be noted that the closing up of the cavity is partially an artefact of the molecular dynamic forcefield used in these simulations and therefore, the observation cannot be fully attributed to the instability of the structure.

With the aim of simulating the interaction between the novel cylinder and the 4WJ both the *M*- and *P*-enantiomers of cylinder were introduced into the molecular dynamic models. Initially, 6 individual simulations totalling 6μs of simulation were generated where both *M*- and *P*-enantiomers were placed at the centre of the junction. A further 6 simulations were generated using the *M*- and *P*-enantiomers, where the cylinder was placed outside of the junction.

The *M*-enantiomer was placed at the centre of the 4WJ. All three simulations showed the enantiomer bound to and remained with the junction throughout the simulation forcing the 4WJ to remain in its open conformation (Figure 3.3.5.7). Further analysis showed that the cylinder bound centrally but at an angle within the junction, enabling both ends of the cylinder to interact with the bases at the cavity of the open form of the 4WJ. The

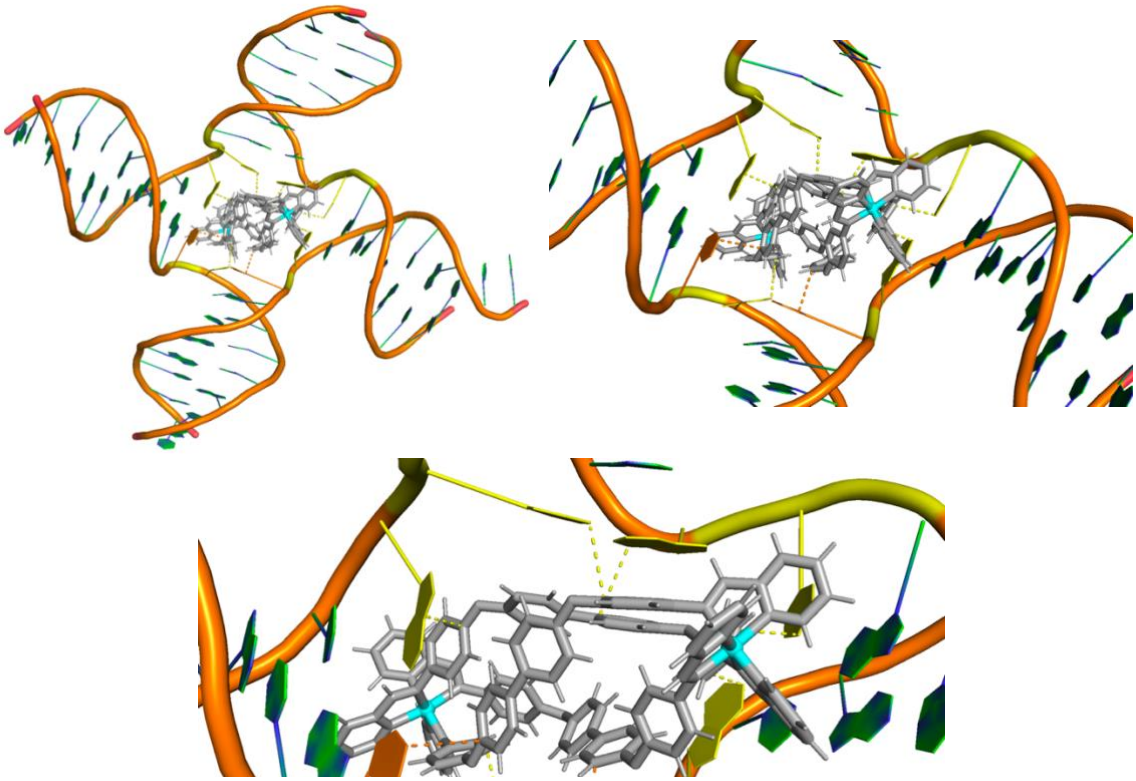


Figure 3.3.5.7 M-enantiomer bound within DNA 4WJ. Top and bottom images show pi-stacking interactions between the bases of the DNA and the phenyl rings in the cylinder backbone. The dotted red lines show the permanent pi-stacking interactions and the dotted yellow lines show the transient pi-stacking interactions

enantiomer was shown to form 8 temporary pi-stacking interactions with the central 4WJ bases, 6 that appeared as consistent interactions and two that were transient. Interestingly, both of the transient pi-stacking interactions were formed between the same end of the cylinder and the DNA suggesting one end of the cylinder is bound more strongly within the cavity. Once more, this is likely due to the significant size of the cylinder, which causes one end to be further away from the DNA bases than is ideal to form more permanent pi-stacking interactions, hence transient ones are only able to form between the two structures due to their natural, fluctuating movement.

In order to compare the interactions of both enantiomers with the 4WJ, simulations where the *P*-enantiomer was placed at the centre of the junction were also generated, as shown in figure 3.3.5.8. Similarly, to the *M*-enantiomer simulations, all 3 simulations showed the enantiomer bound to and remained within the junction throughout the simulation forcing the 4WJ to remain in its open conformation. Further analysis of the simulation showed that like the *M*-enantiomer, the *P*-enantiomer bound at an angle within the cavity of the junction. However, the *P*-enantiomer was able to form 4 permanent pi-stacking interactions and 1 temporary pi-stacking interaction with the DNA bases which were fewer but stronger interactions than the *M*-enantiomer formed (where residency time is acting as a proxy for strength). Potentially, this could indicate that the *P*-enantiomer of the novel cylinder interacts more strongly with 4WJs than the *M*-enantiomer. The *P*-enantiomer did form the majority of its pi-stacking interactions on one side of the cylinder, consistent with the *M*-enantiomer simulations, indicating that the large size of the cylinder meant that it did not entirely fit within the 4WJ cavity and, therefore, one end of the cylinder bound more strongly within the cavity.

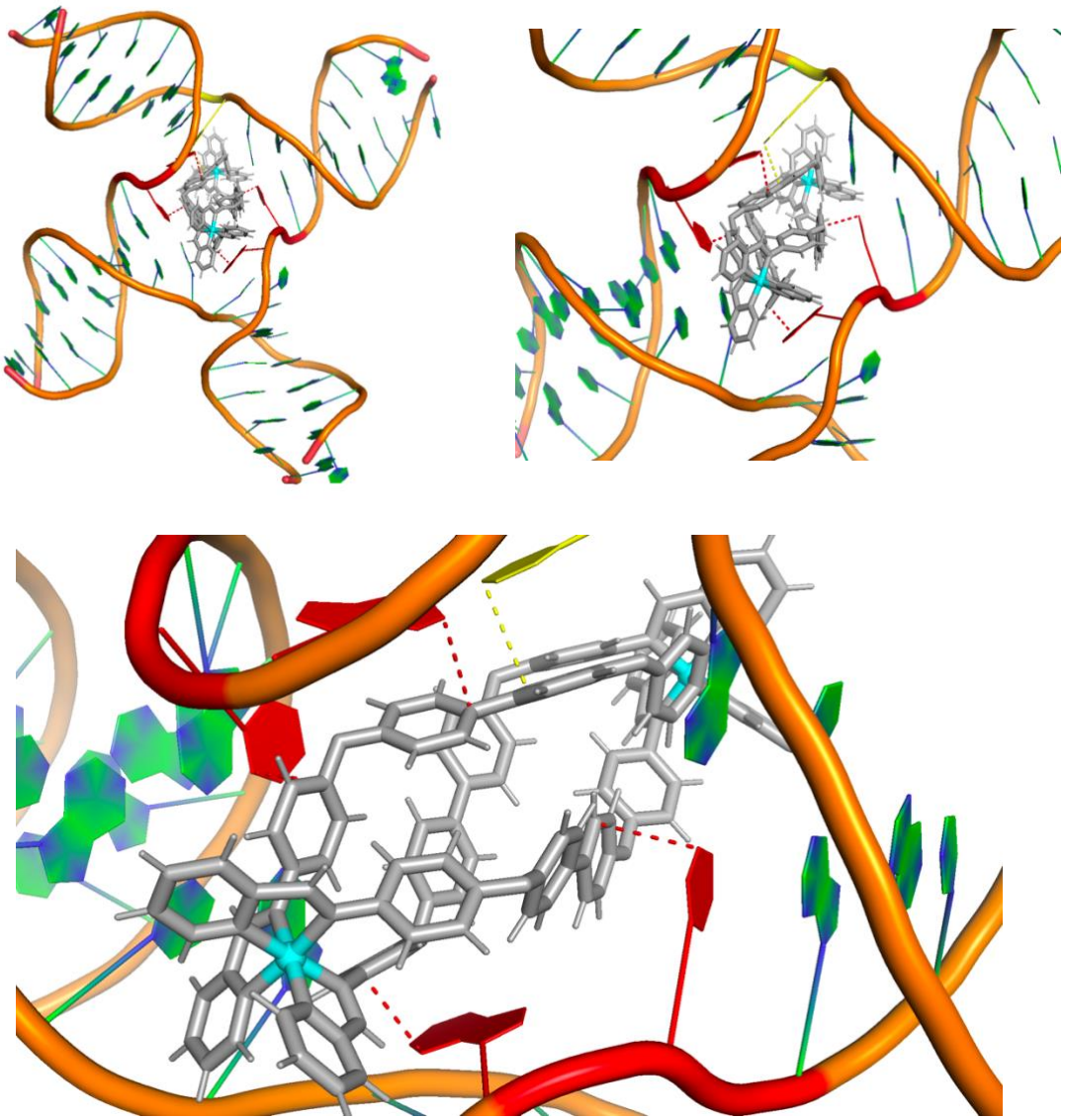


Figure 3.3.5.8 P-enantiomer bound within DNA 4WJ. Top and bottom images show pi-stacking interactions between the bases of the DNA and the phenyl rings in the cylinder backbone. The dotted red lines show the permanent pi-stacking interactions and the do

From the initial central placement of both enantiomers, it was difficult to distinguish whether DNA 4WJs had a binding preference for either the *P* or *M* form of the cylinder. Consequently, further simulations were generated where the enantiomers were placed just outside of the junction. All three of the simulations generated from the *M*-enantiomer

being placed outside of the 4WJ showed the enantiomer thread into the cavity of the junction. Once more, the first 80ns of movement of the cylinder and the 4WJ was investigated in order to analyse the binding mechanism further. In each case, the enantiomer approached the junction at an angle almost parallel to the DNA, which enabled the enantiomer to then thread into the cavity and bind to the 4WJ. Interestingly, the same exact mechanism of interaction was seen with the P enantiomer, however, this was only seen for two of the three simulations as final simulation generated showed no interaction between the P enantiomer and 4WJ at all.

Based on these 4WJ simulations it could be suggested that the 4WJ binds the *M*-enantiomer of the cylinder in slight preference to the *P*. What can be concluded overall, is that a robust model for simulating how the cylinder interacts with and binds to 3WJs and 4WJs has been developed. In comparison to previous molecular dynamic studies, the results of the method show that the novel cylinder interacts with 3WJs in a similar way to the parent cylinder by binding into the junction's cavity and forming pi-stacking interactions with the nucleobases at the branch point, which help substantiate the findings detailed in this thesis.[11] Moreover, the evidence for a difference in binding preference of the enantiomers of the novel helicate to 3WJs suggested by these MD simulations is also reminiscent of previous work completed by Hannon and colleagues, which showed that there was a difference in binding activity between the *M*- and *P*- enantiomers of the parent cylinder.[106] This is not unexpected due to the chiral nature of DNA, however, ultimately in simulations, for both the novel and parent helicate, both enantiomers remain within the junction cavity for the entirety of the simulations, suggesting that the difference between the enantiomers does not affect the stability of the binding.

Chapter 4: CONCLUSIONS

Synthesis of a triple-stranded helicate analogous to the parent cylinder but much more significant in size was achieved through a multistep synthesis. The resulting Fe-based cylinder was shown to have both solubility and stability issues that presented a challenge for further in vitro investigations. These issues were overcome through the development of $[Ni_2L^1_3][Ac]_4$ helicate, which displayed the ability to dissolve in water as well as reasonable stability in the presence of DNA, making it suitable for in vitro investigations. While the studies did not provide evidence that the cylinder had the ability to bind to B-DNA, PAGE displayed the cylinder's ability bind to 3WJs, 4WJs and C-MYC G-quadruplexes. Coupled with the *in-silico* data, it is proposed that the cylinder potentially preferentially bound to the 4WJ over the 3WJ, most likely due to the slightly larger size of the 4WJ which could accommodate the significantly sized cylinder better although tis could also be the result of electrostatics. Finally, the simulations showed that the *P*- and *M*-enantiomeric forms of the helicate interacted with higher-order DNA structures slightly differently, with the 3WJ showing preference for binding the P-enantiomer, which was the opposite result to the analogous parent cylinder.

It is likely that the larger size of this helicate is what contributes to its unusual properties that often differed to the findings of previous research into analogous cylinders. However, there is evidence to suggest that the novel cylinder has potential to act as a therapeutic agent due to its ability to bind to higher-order DNA structures as well as its solubility in biocompatible solvents. Future research into the therapeutic characteristics of this novel cylinder should focus on improving the stability of cylinder through the modification of ligands as was attempted in this research. The potential ability of the cylinder to

rotaxanate with CB10 also presents an interesting potential research route that could be explored.[73] It could enable the selective binding of DNA through controlled de-threading of the cylinder, which in turn would improve the therapeutic potential of the cylinder as it could be used to target specific environments, for example cancerous cells.

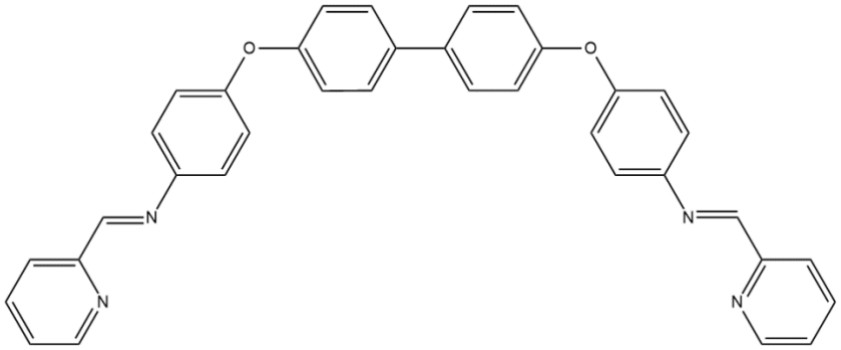
Overall, the creation of this novel cylinder not only adds a new triple-stranded helicate to the current, known portfolio but also presents a metallo-supramolecular helicate that has demonstrated a strong ability to bind to higher-order DNA, suggesting its potential application to therapeutics.

Chapter 5: EXPERIMENTAL

5.1) Materials

All solvents, NMR solvents, chemical reagents and buffer components were purchased from Fischer Scientific, VWR chemicals or Sigma Aldrich and used without further purification. Nickel parent cylinder [Ni2L3]Cl4 was prepared as described previously.[17]

5.1.1) L^1 Ligand

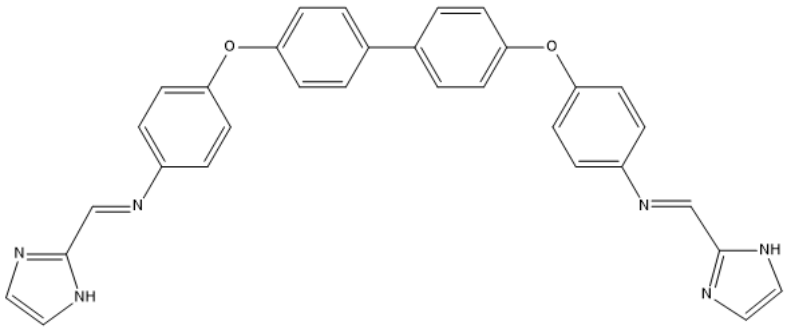


Ligand (L1) (1E,1'E)-N,N'-(([1,1'-biphenyl]-4,4'-diylbis(oxy))bis(4,1-phenylene))bis(1-(pyridin-2-yl)methanimine) was synthesised as follows. 4,4'-(1,1'-biphenyl-4,4'-diyldioxy)dianiline (251 mg, 0.679 mmol) was dissolved in degassed methanol (25 mL) under nitrogen and pyridine-2-carboxaldehyde (0.44 mL, 4.67 mmol, excess) and a couple of drops of acetic acid were added dropwise. The solution was heated under reflux for 24 hours and the yellow precipitate that formed was collected by filtration, washed with cold ethanol and then dried *in vacuo* (300 mg, 80 %)

^{1}H NMR (400 MHz, CD_2Cl_2) δ 8.69 (br d, J = 4.7 Hz, 1H, H_{6py}), 8.63 (s, 1H, H_{imine}), 8.21 (br d, J = 7.8 Hz, 1H, H_{3py}), 7.83 (td, J = 7.6, 1.7 Hz, 1H, H_{4py}), 7.59 (d, J = 8.6 Hz, 2H,

H_{Ph}), 7.38 (ddd, *J* = 7.5, 4.7, 1.2 Hz, 1H, H_{5py}), 7.35 (d, *J* = 8.9 Hz, 2H, H_{Ph}), 7.12 (m comprising two overlapping d, *J*~ 8.8 Hz, 4H, 2x H_{Ph}). MS (ESI+): *m/z* 547 ($[M+H]^+$, 100%), *m/z* 458 ($[C_{30}H_{23}N_3O_2]^+$, 40%).

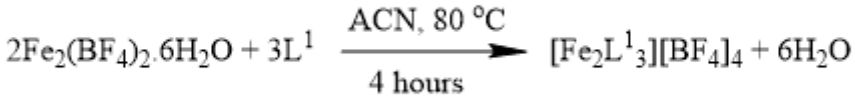
5.1.2) L^2 Ligand



Ligand (L^2) (1E,1'E)-N,N'-(([1,1'-biphenyl]-4,4'-diylbis(oxy))bis(4,1-phenylene))bis(1-(1H-imidazol-2-yl)methanimine) was synthesised as follows. 4,4'-diaminodiphenylmethane (0.595 g, 3 mmol)) was dissolved in methanol (50 mL) and imidazole-2-carboxaldehyde (0.576 g, 6 mmol) also dissolved in methanol (20 mL) was added dropwise. The solution is stirred and refluxed for 24 hours under inert conditions before filtering under vacuo. The resultant off-white powder was washed with methanol before transferring to vial. Yield: 42 %, 544 mg.

^{1}H NMR of L^2 (400 MHz, DMSO-d) δ 8.55 (s, 2H, NH), 8.46 (s, 2H, H imine), 7.71 (s, 2H, H1), 7.69 (s, 2H, H2), 7.41 (s, 2H, Ph1), 7.39 (s, 2H, Ph1), 7.35 (s, 2H, Ph2), 7.19 (s, 2H, Ph2), 7.13 (dd, 8H, Ph3 + Ph4). MS (ESI^+): *m/z* 525 ($[M+H]^+$, 100%), *m/z* 447 ($[C_{28}H_{19}N_3O_2+H]^+$, 10%).

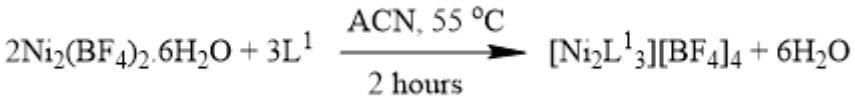
5.1.3) $[Fe_2L^1{}_3][BF_4]_4$



Ligand L1 (30.1 mg, 0.055 mmol) was dissolved in acetonitrile (5 mL) and heated to 50ºC. Iron(II) tetrafluoroborate hexahydrate (12.4 mg, 0.037 mmol) was dissolved in acetonitrile (5 mL) and added to the ligand mixture dropwise with stirring. The reaction mixture was then stirred and heated under reflux for 4 hours. The purple solution was filtered to remove any impurities and the solvent reduced to half volume *in vacuo*. This solution was added into cold diethyl ether (50 mL) and the dark purple solid that formed collected by vacuum filtration (26 mg; 68% yield)

^{1}H NMR (500 MHz, CD_3CN) δ 9.49 (s, 7H, $H_{imine}B$), 9.06 (s, 2H, $H_{imine}C$), 8.71 (d, J = 7.6 Hz, 7H, $H_{3py}B$), 8.58 (d, J = 7.6 Hz, 2H, $H_{3py}C$), 8.38 (m, comprising two overlapping t, J ~ 7.6Hz, 9H, $H_{4py}B$ + $H_{4py}C$), 7.93 (br t, J = 6.6 Hz, 7H, $H_{5py}B$), 7.80 (br t, J = 6.6 Hz, 2H, $H_{5py}C$), 7.74 (br d, J = 5.3 Hz, 7H, $H_{6py}B$), 7.45 (br d, J = 5.5 Hz, 2H, $H_{6py}C$), 7.31 (d, J = 8.6 Hz, 4H, $H_{Ph}C$), 6.98 (m, 42H, 3x $H_{Ph}B$), 6.90 (d, J = 7.6 Hz, 4H, $H_{Ph}C$), 6.77 (br d, J = 8.2 Hz, 4H, $H_{Ph}C$), 5.73 (d, J = 8.3 Hz, 14H, $H_{Ph}B$), 5.46 (br d, J = 8.3 Hz, 4H, $H_{Ph}C$). MS (ESI^+): *m/z* 437 ($[M_2L^1{}_3]^{4+}$, 100%), *m/z* 547 ($C_{36}H_{26}O_2N_4^+$, 10%), *m/z* 274 ($C_{18}H_{13}ON_2^+$, 5%).

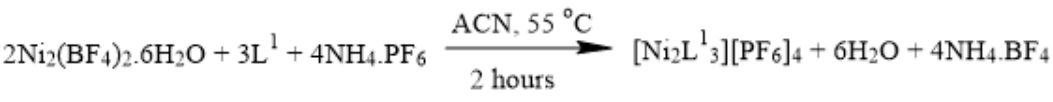
5.1.4) $[Ni_2L^1_3][BF_4]$



Ligand L1 (31 mg, 0.057 mmol) was dissolved in acetonitrile (5 mL) and heated to 50°C. Nickel(II) tetrafluoroborate hexahydrate (9.7 mg, 0.0285 mmol) was dissolved in acetonitrile (5 mL) and added to the ligand mixture dropwise with stirring. The reaction mixture was then warmed to 55°C with stirring for 2 hours. The orange solution was filtered to remove any impurities and the solvent reduced to half volume *in vacuo*. This solution was added into cold diethyl ether (50 mL) and the orange solid that formed was collected by vacuum filtration (28 mg. Yield: 93%)

MS (ESI^+): *m/z* 439 ($[M_2L^1_3]^{4+}$, 100%), *m/z* 403 ($C_{25}H_{18}NO_2K^+$, 5%). Elemental Analysis: Found: C 59.5; H 4.02; N 7.63 %. Calculated for $[Ni_2L1_3][BF_4]_4.5H_2O$: C 59.1; H 4.04; N 7.66 %

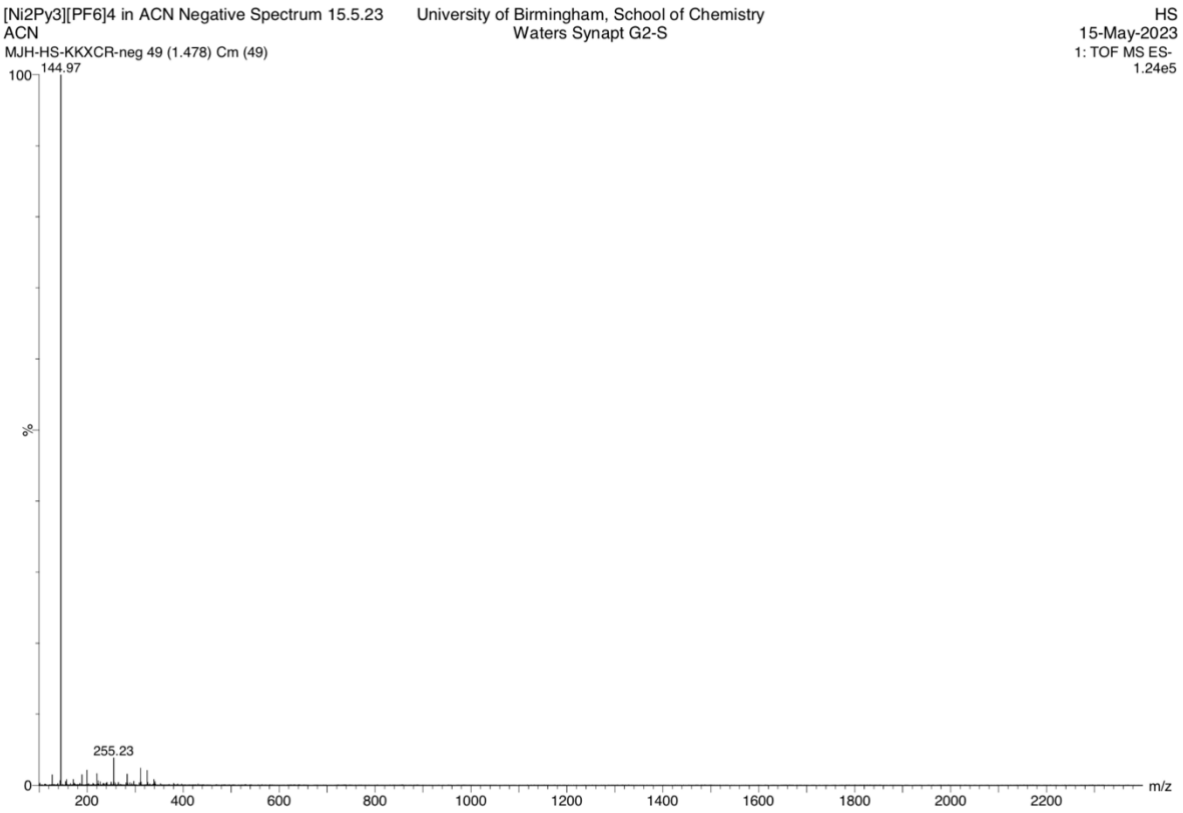
5.1.5) $[Ni_2L^1_3][PF_6]_4$



The pyridine ligand (L^1) (31.40 mg, 0.057 mmol) was dissolved in acetonitrile (5 mL) which was heated to 50ºC. Nickel (II) tetrafluoroborate hexahydrate (9.4 mg, 0.028 mmol) was dissolved acetonitrile (5 mL) and added to the ligand mixture dropwise followed by ammonium hexafluorophosphate (12.2 mg, 0.075 mmol) in acetonitrile (5 mL) with stirring. The reaction mixture was then stirred and heated (55ºC) for 2 hours. The precipitate was collected by vacuum filtration and redissolved in the minimum volume of acetonitrile. Finally, the product was crashed out in cold diethyl ether and isolated by vacuum filtration to give an orange powder. Yield: 84 %, 37 mg

MS (ESI^+): *m/z* 439 ($[M_2L^1_3]^{4+}$, 100%), *m/z* 458 ($[C_{30}H_{23}N_3O_2]^+$, 6%), *m/z* 547 ($[C_{36}H_{26}N_4O_2]^+$, 4%), *m/z* 591 ($[M_2L^1_3F]^{3+}$, 20%)

The negative MS confirming the PF_6 anion is shown below.



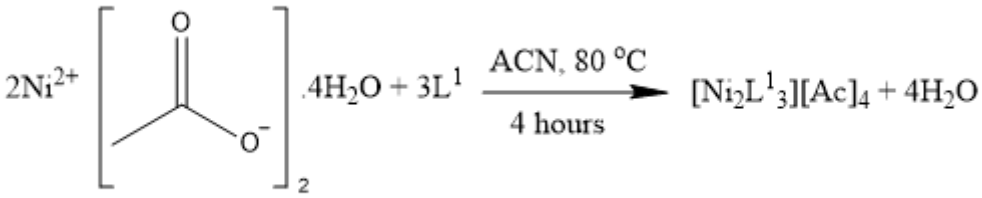
5.1.6) $[Ni_2L^1_3][Cl]_4$

$$2Ni_2(BF_4)_2.6H_2O + 3L^1 + 4NH_4.PF_6 \xrightarrow[2\ hours]{ACN,\ 55\ ^oC} [Ni_2L^1_3][PF_6]_4 + 6H_2O + 4NH_4.BF_4 \xrightarrow[2\ hours]{MeOH,\ Dowex} [Ni_2L^1_3][Cl]_4$$

A column was prepared by packing a Dowex slurry, composed of Dowex and water, into a small glass column and this was then flushed with methanol. $[Ni_2L^1_3][PF_6]_4$ (16.6 mg, 7.110 mmol) was dissolved in the minimum volume of acetonitrile and was added to the top of the column with excess methanol. A small pump was used to aid the passing pf the complex through the column. Finally, the resulting orange powder, $[Ni_2L^1_3][Cl]_4$, was eluted from the bottom in a small pre-weighed vial. Yield: <1 %

MS (ESI$^+$): *m/z* 274 ($C_{18}H_{13}ON_2^+$, 30%), *m/z* 403 ($[M_2L^1{}_2]^{3+}$, 40%), *m/z* 439 ($[M_2L^1{}_3]^{4+}$, 25%), *m/z* 458 ($[C_{30}H_{23}N_3O_2]^+$, 90%), *m/z* 547 ($[C_{36}H_{26}N_4O_2]^+$, 100%),

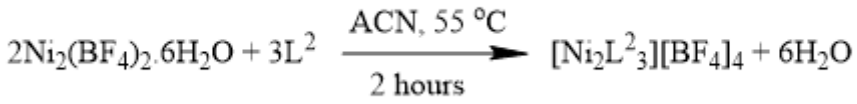
5.1.6) $[Ni_2L^1{}_3][OAc]_4$



Ligand L1 (30 mg, 0.055 mmol) was dissolved in methanol (5 mL) and heated to 50ºC. Nickel(II) acetate tetrahydrate (9.1 mg, 0.037 mmol) was dissolved in methanol (5 mL) and added to the ligand mixture dropwise with stirring. The reaction mixture was then heated under stirring and reflux for 4 hours. The orange solution was filtered to remove any impurities and the solvent reduced to half volume *in vacuo*. This solution was added into cold diethyl ether (50 mL) and the orange solid that formed was collected by vacuum filtration (17 mg; Yield 47%).

MS (ESI$^+$): *m/z* 439 ($[M_2L^1{}_3]^{4+}$, 100), *m/z* 458 ($[C_{30}H_{23}N_3O_2]^+$, 37%), *m/z* 381 ($[C_{25}H_{20}N_2O_2]^+$, 27%), *m/z* 274 ($[C_{18}H_{13}N_2O]^+$, 5%), *m/z* 249 (unknown, 18%).

5.1.7) $[Ni_2L^2_3][BF_4]_4$



The imidazole ligand (L^2) (29.5 mg, 0.056mmol) was dissolved in acetonitrile (5 mL) which was heated to 50ºC. Nickel (II) tetrafluoroborate hexahydrate (9.8 mg, 0.029 mmol) was dissolved acetonitrile (5 mL) and added to the ligand mixture dropwise with stirring. The reaction mixture was then heated under reflux (55ºC) and left stirring for 3 hours. The precipitate was collected by vacuum filtration and redissolved in the minimum volume of acetonitrile. Finally, the product was crashed out in cold diethyl ether (50 mL) and isolated by vacuum filtration as an orange/brown powder. Yield: 75 %, 22 mg

MS (ESI^+): *m/z* 422 ($[M_2L^2_3]^{4+}$, 100%), *m/z* 263 (L^2+2H^+, 95%), *m/z* 582 ($[Ni_2(L^2-H)_2]^{2+}$, 73%), *m/z* 563 ($[Ni_2Im_2(L^2-H)]^{3+}$, 62%), *m/z* 525 (L^2+H^+, 32%), *m/z* 844 ($[Ni_2L^2(L^2-H)_2]^{2+}$, 10%). Where L = $C_{32}H_{24}N_6O_2$.

5.2) Methods

Milli-Q water (18.2 MΩ) was used for the entirety of the biophysical work completed in this thesis. For Circular and Linear dichroism, DNA samples were made up from Calf thymus DNA sodium salt (Sigma Aldrich) by dissolving in milli-Q water (18.2 MΩ). After this, the concentration of the solution was measured by UV-Vis spectroscopy (Cary 5000 NIR spectrometer) by ε (260 nm) = 13,200 mol^{-1} dm^3 cm^{-1} to give a concentration

in DNA base pairs. The stock solution of this was stored at 4 ºC with fresh aliquots being taken out for each experiment as well as fresh buffer was made up before each experiment. Moreover, complexes were dissolved in Milli-Q water (18.2 MΩ) only with fresh solutions being used for each batch of experiments. Electrospray ionisation (ESI) mass spectrometry characterisation was carried out on a Waters SYNAPT-G2-S in positive ion mode. 1H NMR studies were carried out on AVIII 300 (300 MHz), AVANCE NEO400 and AVIII 400 (400MHz) Bruker spectrometers.

5.2.1) Polyacrylamide Gel Electrophoresis (PAGE)

Polyacrylamide gels (12%) were prepared by mixing 20 ml of 37.5:1 acrylamide/bis-acrylamide with 5 mL of 10x Tris-Boric acid buffer (890 mM each, pH 8.3) and 25 ml of deionised water. To this 400 µL of a 10% w/v ammonium persulfate/water solution and 40 µL of TEMED were added to initialise polymerisation. This is then immediately poured between 2 glass plates and a comb inserted at the top this is then left to set for 1 hr. Samples were made up to 30 µl containing 1 µM of each DNA strand, 1xTBN buffer (89 mM Tris, 89 mM Boric acid, 10 mM NaCl, pH 8.3) and the indicated ratio of complex. DNA, water, and buffer were mixed in solution before addition of the stated ratios of complex and additional components. Samples were then incubated at 37 degrees Celsius for 1 hr. Then 7. 5 µl of 50% w/v glycerol was added to each sample and the sample was then centrifuged, mixed, and pipetted into the wells on the gel. The Gel was run at 140 V for 2.5 hours in 1xTB buffer. The gel was then removed from the plates and stained using SYBR™ Gold Nucleic Acid Gel Stain (Thermofisher cientific) in 1xTB buffer for 45 minutes before imaging on a bio-rad ChemiDoc fluorescent imager with 305 nm excitation. DNA structures were at 1 µM in the samples with equal concentrations of each

strand needed for their respective structures, complex concentrations are indicated in ratios to this.

Mini PAGE gels (12%) were prepared similarly to the larger PAGE gels but on a reduced scale. These ran for 35 minutes at 140 V in 1xTB buffer. All of conditions for staining remained the same.

5.2.2) Circular Dichroism (CD)

Circular dichroism (CD): spectra were recorded on a Chirascan+ spectrometer (Applied Photophysics limited). The samples were scanned in a 1 cm path length cuvette between 800 and 200 nm with 3 technical repeats at 1 nm step size and 0.5s dwell time per point. Titrations were carried out at a constant concentration of CT-DNA (200 uM), sodium chloride (40 mM) and TrisHCL buffer (20 mM, pH 7.4) by adding compensating solutions of 2x DNA/Buffer of equal volume to the titre of the complex solution. The concentration of complex in the cuvette was increased step wise by adding set volumes of a stock complex solution. CT-DNA concentration was 1000 μM in DNA base pairs.

5.2.3) Flow Linear Dichroism (LD)

LD was carried out on the same Chirascan+ spectrometer (Applied Photophysics limited) using the LD accessory (Applied Photophysics limited). The LD cell has an angular gap of 0.25 mm giving an overall path length of 0.5 mm. Samples volumes began at 150 μl and stopped at 250 μl. The cell was rotated at 40 revolutions per second to optimise the

DNA signal. The titration series was carried out the same as in the CD studies, with a 3-minute incubation time at a lower revolution speed of 4 revolutions per second. CT-DNA concentration was 100 µM in DNA base pairs.

5.2.4) UV-Vis Absorbance Spectroscopy

UV-Vis absorption spectra were recorded on a Cary 5000 Spectrometer. The samples were made up to a concentration of 1 mM in the relevant buffer (ACN, MeOH, TrisHCl, H_2O) and were scanned in a 1 cm path length cuvette between 800 and 200 nm. Samples were scanned periodically over 24 hours and then at regular intervals across a week.

5.2.5) Molecular Dynamics

Cylinder Parameterisation: The crystal structure of the Li group's cylinder was altered it in Avogadro, converting thiol to pyridine, by collegue Sam Dettmer. Then, the geometry of the parent cylinder optimised using oxygen in the linker over CH2 (as the larger helicate has oxygen conjugated to benzene) in ORCA at ωB97XD/DEF2SVP level of theory. This was fed into MCPB.py to generate the Forcmod files which defined bond lengths, bond angles and dihedral angles. After this, MOL2 files were produced from the larger helicate and the Muliken charges were generated by optimising the larger cylinder in ORCA. These charges were copied from ORCA and pasted into these MOL2 files. Then, the MCPB generated Forecmod file from the modified parent cylinder was used to redefine the atoms in the MOL2 files, with all other atoms being defined by the general

amber force field (GAFF). Amber T-leap was used to generate the Amber topology file from the adapted MOL2 files, which incorporated the Forecmod file from the modified parent cylinder. Finally, the Amber topology file was converted to the GROMACS topology file using the ParmEd function, resulting in the successful parameterisation of the $[Fe_2L^1_3]^{4+}$ cylinder. It should be noted that above methodology details how the *M* -enantiomer of the helicate was optimised. It should be noted that above methodology details how the *M* -enantiomer of the helicate was optimised. In order to generate the parameterised *P*-enantiomer, the chirality of parameterised *M*-enantiomer was inverted in AVOGADRO by collegue, Sam Dettmer.

Parameterisation of DNA: Several DNA structures have already been parameterised using the Hannon group's method [11] including DNA 3WJ (one larger version and one smaller version), DNA 4WJ and duplex DNA and so could be used in this research. The smaller 3WJ was prepared from PDB 2ET0[6], with the compounds removed. The large 3WJ was adapted by colleague Sam Dettmer from PDB 1F44[106] as reported previously. The extra solvent and protein were removed from the 3WJ crystal structure files before the backbone of the structures were joined up using Avogadro to create 3WJs formed of continuous DNA strands. Finally, these 3WJ structures underwent energy minimisation in GROMACS to produce the final structures used in the molecular dynamic (MD) simulations. The 4WJ was parameterised similarly to the 3WJs, where the crystal structure, taken from PDB 1XNS[103], had all water molecules and peptide strands removed before each arm strand was shortened to 9 base pairs as this reduced the modelling cost. Again, this structure was energy minimised in GROMACS to produce the final structure used in the MD simulations. The final DNA structure that was parameterised was B-

DNA, which was generated from 22 adenine bases utilizing the Nucleic Acid Builder (NAB). It is important to note that all these DNA structures were parameterised using parmbsc1, an AMBER forcefield. [104]

Molecular Dynamics (MD) method: The molecular dynamics simulations method used in this research closely follows that detailed by Larry Melidis in his 2021 paper.[105] In the initial simulations, the cylinder was placed in the centre of the cavity in either the 3WJ or 4WJ, while in the later simulations the cylinder was placed directly outside of the junctions. The system (the cylinder and DNA) was then placed in a dodecahedral box of 1.0 nm size and solvated using the TIP3P water model. A 50 mmol concentration of NaCl ions was added to neutralise it. Subsequently the system was minimised such that the maximum force on any one atom was less than 500 kJ mol-1, followed by NVT and NPT equilibration using V-rescale modified Berendsen thermostat at 310 K and for 1000 ps each, in which the cylinder was coupled with the DNA. The simulations were then run using a time-step of 2 fs. This procedure was also used to model the un-complexed (free) DNA by not including the parameters of the cylinder. Finally, the compressed trajectories were processed in GROMACS to remove rotational and translational motion, and to centre the DNA within the box for visualisation and analysis purposes, which were done in PYMOL.

Root Mean Square Deviation (RMSD) Calculations: RMSD calculations are used to show how much a system diverges from a known reference structure and is often used to investigate the stability of a binding interaction. Large conformational changes are expressed as spikes on the RMSD plot and therefore can indicate if an interaction is

unstable. In this research, RMSD plots were used to explore how energetically stable the interactions between the novel cylinder and the higher-order DNA structures were and consequently, evaluate whether the novel cylinder exhibited a binding preference to a particular structure. The plots were calculated in GROMACS using a reference and trajectory file for each system.

It should be noted that the simulations used the Bluebear and Castles HPC facility at the University of Birmingham.[107]

Chapter 6: REFERENCES

1 M. J. Hannon, *Chem. Soc. Rev.*, 2007, **36**, 280–295.
2 M. Santos, P. M. Pereira, A. S. Varanda, J. Carvalho, M. Azevedo, D. D. Mateus, N. Mendes, P. Oliveira, F. Trindade, M. T. Pinto, R. Bordeira-Carriço, F. Carneiro, R. Vitorino, C. Oliveira and M. A. S. Santos, *RNA Biol.*, 2018, **15**, 773–786.
3 A. A. Mercadante and A. Kasi, *Genetics, Cancer Cell Cycle Phases*, StatPearls Publishing, Treasure Island, 2022.
4 G. Guilloux and R. Gibeaux, *Biol. Cell*, 2020, **112**, 369–382.
5 N. Deng, H. Zhou, H. Fan and Y. Yuan, *Oncotarget*, 2017, **8**, 110635-110649.
6 A. Oleksi, A. G. Blanco, R. Boer, I. Usón, J. Aymamí, A. Rodger, M. J. Hannon and M. Coll, *Angew. Chem. Int. Ed.*, 2006, **45**, 1227–1231.
7 P. B. Dervan and B. S. Edelson, *Curr. Opin. Struct. Biol.*, 2003, **13**, 284–299.
8 J. Arpalahti, K. D. Klika, R. Sillanpää and R. Kivekäs, *J. Chem. Soc., Dalton Trans.*, 1998, 1397-1402.
9 Z. J. Guo and P. J. Sadler, *Adv. Inorg. Chem.*, 2000, **49**, 183–306.
10 S. Dasari and P. Bernard Tchounwou, *Eur. J. Pharmacol.*, 2014, **740**, 364–378.
11 J. S. Craig, L. Melidis, H. D. Williams, S. J. Dettmer, A. A. Heidecker, P. J. Altmann, S. Guan, C. Campbell, D. F. Browning, R. K. O. Sigel, S. Johannsen, R. T. Egan, B. Aikman, A. Casini, A. Pöthig and M. J. Hannon, *J. Am. Chem. Soc.*, 2023, **145**, 13570–13580.
12 D. R. Boer, A. Canals and M. Coll, *Dalton Trans.*, 2009, 399–414.
13 S. M. Bhattacharjee, *J. Phys.: Condens. Matter*, 2010, **22**, 155102–155108
14 J. Gómez-González, L. Martínez-Castro, J. Tolosa-Barrilero, A. Alcalde-Ordóñez, S. Learte-Aymamí, J. L. Mascareñas, J. C. García-Martínez, J. Martínez-Costas, J. D. Maréchal, M. Vázquez López and M. E. Vázquez, *Chem. Commun.*, 2022, **58**, 7769–7772.
15 J. Malina, H. Kostrhunova, P. Scott and V. Brabec, *Nucleic Acids Res.*, 2023, **51**, 7174-7183.
16 S. Vuong, L. Stefan, P. Lejault, Y. Rousselin, F. Denat and D. Monchaud, *Biochimie*, 2012, **94**, 442–450.
17 M. J. Hannon, C. L. Painting, A. Jackson, J. Hamblin and W. Errington, *Chem. Commun.*, 1997, **11**, 1807-1808.
18 J. Gómez-Gonzµlez, Y. Pérez, G. iuseppe Sciortino, L. Roldan-Martín, J. Martínez-Costas, J. ean-Didier Maréchal, I. Alfonso, M. Vµzquez López and M. Eugenio Vµzquez, *Angew. Chem. Int. Ed.*, 2021, **133**, 8941–8948.
19 J. Gómez-González, D. Bouzada, L. A. Pérez-Márquez, G. Sciortino, J. D. Maréchal, M. Vázquez López and M. E. Vázquez, *Bioconjugate Chem.*, 2021, **32**, 1564–1569.
20 J. Zhu, C. J. E. Haynes, M. Kieffer, J. L. Greenfield, R. D. Greenhalgh, J. R. Nitschke and U. F. Keyser, *J. Am. Chem. Soc.*, 2019, **141**, 11358–11362.
21 S. Phongtongpasuk, S. Paulus, J. Schnabl, R. K. O. Siegel, B. Spingler, M. J. Hannon and E. Freisinger, *Angew. Chem. Int. Ed.*, 2013, **52**, 11513–11516.
22 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.*, 2010, **49**, 2336–2339.
23 M. Jaroslav, M. J. Hannon and V. Brabec, *Chem. – Eur. J.*, 2007, **13**, 3871–3877.
24 L. Cerasino, M. J. Hannon and E. Sletten, *Inorg. Chem.*, 2007, **46**, 6245–6251.

25 S. A. Barros and D. M. Chenoweth, *Angew. Chem. Int. Ed.*, 2014, **53**, 13746–13750.
26 K. C. Woods, S. S. Martin, V. C. Chu and E. P. Baldwin, *J. Mol. Biol.*, 2001, **313**, 49–69.
27 T. M. Long and T. M. Swager, *J. Am. Chem. Soc.*, 2002, **124**, 3826–3827.
28 E. Ivens, M. M. D. Cominetti and M. Searcey, *Bioorg. Med. Chem.*, 2022, **69,** 116897.
29 V. H. S. van Rixel, A. Busemann, M. F. Wissingh, S. L. Hopkins, B. Siewert, C. van de Griend, M. A. Siegler, T. Marzo, F. Papi, M. Ferraroni, P. Gratteri, C. Bazzicalupi, L. Messori and S. Bonnet, *Angew. Chem. Int. Ed.*, 2019, **131**, 9478–9482.
30 L. A. Howell, Z. A. E. Waller, R. Bowater, M. O'Connell and M. Searcey, *Chem. Comm.*, 2011, **47**, 8262–8264.
31 A. L. Brogden, N. H. Hopcroft, M. Searcey and C. J. Cardin, *Angew. Chem. Int. Ed*, 2007, **46**, 3850–3854.
32 H. Han and L. H. Hurley, *TiPS*, 2000, **21**, 137-141.
33 G. Miglietta, J. Marinello, M. Russo and G. Capranico, Mol Cancer, 2022, **21**, 180.
34 H. Qin, C. Zhao, Y. Sun, J. Ren and X. Qu, *J. Am. Chem. Soc.*, 2017, **139**, 16201–16209.
35 E. Palma, J. Carvalho, C. Cruz and A. Paulo, *Pharm.*, 2021, **14**, 605.
36 J. Zegers, M. Peters and B. Albada, *J. Bio. Inorg. Chem.*, 2023, **28**, 117–138.
37 A. Ghosh, M. Trajkovski, M. P. Teulade-Fichou, V. Gabelica and J. Plavec, *Angew. Chem. Int. Ed.*, 2022, 61.
38 H. Yu, X. Wang, M. Fu, J. Ren and X. Qu, *Nucleic Acids Res.*, 2008, **36**, 5695–5703.
39 K. Duskova, P. Lejault, É. Benchimol, R. Guillot, S. Britton, A. Granzhan and D. Monchaud, *J. Am. Chem. Soc.*, 2020, **142**, 424–435.
40 B. P. Belotserkovskii, S. M. Mirkin and P. C. Hanawalt, *Chem. Rev.*, 2013, **113**, 8620–8637.
41 T. M. Bryan, *Mol.*, 2019, **24**, 3439.
42 N. Puget, K. M. Miller and G. Legube, *DNA Repair*, 2019, **81**, 102661.
43 W. Wang, S. Hu, Y. Gu, Y. Yan, D. B. Stovall, D. Li and G. Sui, *Biochim. Biophys. Acta, Rev. Cancer*, 2020, **1874**, 188410.
44 A. M. Zahler, J. R. Williamson, T. R. Cech and D. M. Prescott, *Nat.*, 1991, **350**, 718–720.
45 S. Neidle, *J. Med. Chem.*, 2016, **59**, 5987–6011.
46 D. Sun, B. Thompson, B. E. Cathers, M. Salazar, S. M. Kerwin, J. O. Trent, T. C. Jenkins, S. Neidle and L. H. Hurley, *J. Med. Chem.*, 1997, **40**, 2113–2116.
47 S. Balasubramanian, L. H. Hurley and S. Neidle, *Nat. Rev. Drug Discovery*, 2011, **10**, 261–275.
48 C. Piguet, G. Bernardinelli and G. Hopfgartner, *Chem. Rev.*, 1997, **97**, 2005–2062.
49 L. G. Warner, N. J. Rose and D. H. Busch, *J. Am. Chem. Soc.*, 1968, **90**, 6938–6946.
50 D. C. Jicha and D. H. Busch, *Inorg. Chem.*, 1962, **1**, 878–893.
51 L. T. Taylor and D. H. Busch, *J. Am. Chem. Soc.*, 1967, **89**, 5372–5376.

52 J.-M. Lehn, A. Rigault, J. Siegel, J. Harrowfield, B. Chevriert and D. Morast, *Proc. Natl. Acad. Sci. USA*, 1987, **84**, 2565–2569.
53 B. R. Serr, K. A. Andersen, C. Michael Elliott and O. P. Anderson, *Inorg. Chem.*, 1988, **27**, 4499-4504.
54 R. Krämer, J. -M Lehn, A. De Cian and J. Fischer, *Angew. Chem. Int. Ed*., 1993, **32**, 703–706.
55 M. J. Hannon, V. Moreno, M. J. Prieto, E. Moldrheim, E. Sletten, I. Meistermann, C. J. Isaac, K. J. Sanders and A. Rodger, *Angew. Chem. Int. Ed*., 2001, **40**, 879–884.
56 L. Cardo and M. J. Hannon, *Inorg. Chim. Acta*, 2009, **362**, 784–792.
57 M. J. Hannon and L. J. Childs, *Supramol. Chem.*, 2004, **16**, 7–22.
58 L. Cardo, V. Sadovnikova, S. Phongtongpasuk, N. J. Hodges and M. J. Hannon, *Chem. Commun.*, 2011, **47**, 6575–6577.
59 C. R. K. Glasson, G. V. Meehan, J. K. Clegg, L. F. Lindoy, J. A. Smith, F. R. Keene and C. Motti, *Chem. – Eur. J.*, 2008, **14**, 10535–10538.
60 J. Malina, H. Kostrhunova, P. Scott and V. Brabec, *Chem. – Eur. J.*, 2021, **27**, 11682–11692.
61 S. V. Kumar, W. K. C. Lo, H. J. L. Brooks and J. D. Crowley, *Inorg. Chim. Acta*, 2015, **425**, 1–6.
62 S. M. McNeill, D. Preston, J. E. M. Lewis, A. Robert, K. Knerr-Rupp, D. O. Graham, J. R. Wright, G. I. Giles and J. D. Crowley, *Dalton Trans.*, 2015, **44**, 11129–11136.
63 R. A. Kaner and P. Scott, *Future Med. Chem.*, 2015, **7**, 1–4.
64 H. Song, N. J. Rogers, S. J. Allison, V. Brabec, H. Bridgewater, H. Kostrhunova, L. Markova, R. M. Phillips, E. C. Pinder, S. L. Shepherd, L. S. Young, J. Zajac and P. Scott, *Chem. Sci.*, 2019, **10**, 8547–8557.
65 O. Domarco, D. Lötsch, J. Schreiber, C. Dinhof, S. Van Schoonhoven, M. D. García, C. Peinador, B. K. Keppler, W. Berger and A. Terenzi, *Dalton Trans.*, 2017, **46**, 329–332.
66 C. Ducani, A. Leczkowska, N. J. Hodges and M. J. Hannon, *Angew. Chem. Int. Ed*., 2010, **49**, 8942–8945.
67 L. Cardo and M. J. Hannon, *Met. Ions in Life Sci.*, 2018, **18**, 29394030.
68 X. Li, M. Han, H. Zhang, F. Liu, Y. Pan, J. Zhu, Z. Liao, X. Chen and B. Zhang, *Biomark. Res.*, 2022, **10**, DOI: 10.1186/s40364-021-00345-1.
69 A. Chen and A. N. Koehler, *Trends Mol. Med.*, 2020, **26**, 508–518.
70 S. Kim and J. Shendure, *Mol. Cell.*, 2019, **76**, 306–319.
71 J. Jen and Y. C. Wang, *J. Biomed. Sci*, 2016, **23,** DOI: 10.1186/s12929-016-0269-9.
72 L. Yang, S. R. Hamilton, A. Sood, T. Kuwai, L. Ellis, A. Sanguino, G. Lopez-Berestein and D. D. Boyd, *Cancer Res.*, 2008, **68**, 4321–4330.
73 C. A. J. Hooper, L. Cardo, J. S. Craig, L. Melidis, A. Garai, R. T. Egan, V. Sadovnikova, F. Burkert, L. Male, N. J. Hodges, D. F. Browning, R. Rosas, F. Liu, F. V. Rocha, M. A. Lima, S. Liu, D. Bardelang and M. J. Hannon, *J. Am. Chem. Soc.*, 2020, **142**, 20651–20660.
74 Z. Huang, W. Song and X. Chen, *Front. Chem.*, 2020, **8**, DOI: 10.3389/fchem.2020.00124
75 J. M. Lehn, *Angew. Chem. Int. Ed*., 1988, **27**, 89–112.
76 X. Ma and Y. Zhao, *Chem. Rev.*, 2015, **115,** 7794–7839.

77 S. Y. Tan, C. Y. Ang and Y. Zhao, in *Comprehensive Supramolecular Chemistry II*, ed. J. L. Atwood, Elsevier, Oxford, 2017, vol. 2, pp. 391–420.
78 X. Yang, F. Liu, Z. Zhao, F. Liang, H. Zhang and S. Liu, *Chin. Chem. Lett.*, 2018, **29**, 1560–1566.
79 S. Liu, P. Y. Zavalij and L. Isaacs, *J. Am. Chem. Soc.*, 2005, **127**, 16798–16799.
80 R. Barat, T. Legigan, I. Tranoy-Opalinski, B. Renoux, E. Péraudeau, J. Clarhaut, P. Poinot, A. E. Fernandes, V. Aucagne, D. A. Leigh and S. Papot, *Chem. Sci.*, 2015, **6**, 2608–2613.
81 H. Bruce Bosmann and T. C. Hall, *Proc. Nat. Acad. Sci. USA*, 1974, **71**, 1833-1837.
82 V. Paradis, N. Youssef, D. Dargère, N. Bâ, F. Bonvoust, J. Deschatrette and P. Bedossa, *Hum. Pathol.*, 2001, **32**, 327–332.
83 R. H. Te Poele, A. L. Okorokov, L. Jardine, J. Cummings and S. P. Joel, *Cancer Res.*, 2002, **62**, 1876-1883.
84 C. Piguet, *J. Inclusion Phenom. Macrocyclic Chem.*, 1999, **34**, 361-391.
85 J. E. M. Lewis and J. D. Crowley, *Chempluschem*, 2020, **85**, 815–827.
86 F. Habib, J. Long, P. H. Lin, I. Korobkov, L. Ungur, W. Wernsdorfer, L. F. Chibotaru and M. Murugesu, *Chem. Sci.*, 2012, **3**, 2158–2164.
87 K. J. Howard-Smith, A. R. Craze, H. Zenno, J. Yagyu, S. Hayami and F. Li, *Chem. Commun.*, 2020, **56**, 8838–8841.
88 K. J. Howard-Smith, M. J. Wallis, J. P. Flood, H. Min, J. C. Tadros, R. Tian, M. M. Bhadbhade, C. E. Marjo and F. Li, *J. Inclusion Phenom. Macrocyclic Chem.*, 2024, 104, 199-207
89 J. McGinley, M. McCann, K. Ni, T. Tallon, K. Kavanagh, M. Devereux, X. Ma and V. McKee, *Polyhedron*, 2013, **55**, 169–178.
90 M. Albrecht, I. Janser, H. Houjou and R. Fröhlich, *Chem. - Eur. J.*, 2004, **10**, 2839–2850.
91 Claire L. Painting, PhD Thesis, University of Warwick, 1999.
92 W. Kandioller, J. Theiner, B. K. Keppler and C. R. Kowol, *Inorg. Chem. Front.*, 2022, **9**, 412–416.
93 R. Saha, A. K. Ghosh, R. N. Samajdar and P. S. Mukherjee, *Inorg. Chem.*, 2018, **57**, 6540–6548.
94 L. Y. Yao, L. Qin, T. Z. Xie, Y. Z. Li and S. Y. Yu, *Inorg. Chem.*, 2011, **50**, 6055–6062.
95 K. K. Patel, PhD Thesis, University of Warwick, 2001.
96 S. Arcoria, *Ann. Chim. (Rome)*, 1967, **57**, 1125–1132.
97 S. Solinski and E. Halasa, *Roczniki Chemii*, 1954, **48**, 1459–1468.
98 K. Yoshida, R. Mitsumori, K. Horii, A. Takashima and I. Nishio, *Colloids and Interfaces*, 2018, **2**.
99 W. G. Van and D. Sluys, *J. Chem. Ed.*, 2001, **78**, 111-115.
100 A. Rodger, G. Dorrington and D. L. Ang, *Analyst*, 2016, **141**, 6490–6498.
101 P. Li and K. M. Merz, *J. Chem. Inf. Model.*, 2016, **56**, 599–604.
102 C. J. Cramer, D. G. Truhlar, *Phys. Chem. Chem. Phys.*, 2009, **11**, 10757-10816.
103 K. Ghosh, K. L. Chi, F. Guo, A. M. Segall and G. D. Van Duyne, *J. Bio. Chem.*, 2005, **280**, 8290–8299.
104 I. Ivani, P. Dans, A. Noy *et al.*, *Nat. Methods*, 2016, **13**, 55-58
105 L. Melidis, I. B. Styles and M. J. Hannon, *Chem. Sci.*, 2021, **12**, 7174–7184.

106 I. Meistermann, V. Moreno, M. J. Prieto, E. Moldrheim, E. Sletten, S. Khalid, P. M. Rodger, J. C. Peberdy, C. J. Isaac, A. Rodger, M. J. Hannon, *Proc Natl Acad Sci*, 2002, **99**, 5069-5074.
107 S. J. Thompson, S. E. M. Thompson, J. B. Cazier, CaStLeS (Compute and Storage for the Life Sciences): A Collection of Compute and Storage Resources for Supporting Research at the University of Birmingham. 2019 DOI: 10.5281/ZENODO.3250616.