Bimetallic Triple Helicates: Nucleic Acid Interaction and Functionalization



UNIVERSITY OF BIRMINGHAM

by

Ross Thomas Egan

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

School of Chemistry
University of Birmingham
College of Engineering and Physical Sciences
May 2022

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

Abstract

This thesis is an exploration into the binding of supramolecular triple helicates to nucleic acids. Ultimately it is hoped that this study can be used to inform further research into controlling this binding.

It begins with a review of biologically available nucleic acid structures and the common characteristics of non-covalent binding drugs. This then follows with a focused review of nucleic acid binding helicates. Chapter 2 introduces a relatively new area of research in the group, namely the rotaxanation of bimetallic helicates. A thiol cleavable rotaxane capping group is applied and then its cleavage is examined with various time dependent methods, showing how helicate DNA binding can be switched off and on again using rotaxanation. Chapters 3 and 4 focus on the binding of helicate enantiomers to nucleic acids. Chapter 3 presents an NMR analysis of a ruthenium cylinder enantiomer to the TAR (Trans Activation Response) RNA hairpin, providing insight into the mechanism and location of binding as well as structural changes to the hairpin on binding. Chapter 4 delves into DNA three-way junction DNA with cylinders, with the aim to understand their binding from the perspective of enantiospecificity. Together they present a picture of the relationship between structure and nucleic acid binding, using enantiospecificity and steric hindrance from macrocycles.

Declaration

The experimental work, observations and recommendations reported in this thesis are the author's unless specifically stated and have not previously been submitted as part of a degree at the University of Birmingham or any other institution.

Ross Egan

May 2022

Acknowledgements

First, my supervisor Professor Michael Hannon deserves acknowledgement for his contribution. Mike has obviously been directly involved throughout, and finally getting me over the finish line.

I want to thank Dr Cécile Le Duff for the time and training offered in the NMR labs. Within the research group I would like to thank Callum for introducing me to the lab and mentoring my first year. James, Cat and Hugo for conversation and companionship in the lab and Tasha for conversation in the office.

There was a noticeable improvement to my research on the arrival into the lab of Dr Aditya Garai with his experience. Together we worked to complete work on a rotaxane structure in which Chapter 2 contains a part. James Craig and Hugo Williams have been involved in the later stages and I expect others may be by time of publication.

The School of Chemistry hosted and funded my scholarship. The Society of Biological Inorganic Chemistry deserves a mention for a grant given to partially cover travel expenses and the conference attendance fee for ICBIC19. Thanks to Rachel Hindle who helped me with my organisation during writeup, with funding from the University of Birmingham through the DSA scheme.

Collaboration was made with Prof. Dr. Roland K. O. Sigel in his research group. With special mention to Dr. Silke Johannsen and Dr. Zenghui Wang for training and running measurements when I could not.

Finally, I would like to thank Ellie whose support has been indispensable over these years. Along with Peggy, Wilbur, Maya, and Eowyn for their thoughtful conversation.

Chapter Contents

- Chapter 1 – Introduction 6
- Chapter 2 – Rotaxanating Cylinders 59
- Chapter 3 – Cylinder Interactions with TAR RNA 153
- Chapter 4 – P Enantiomer with Three-Way Junction DNA 253
- Chapter 5 – Conclusions and Future Work 272

CHAPTER 1 – INTRODUCTION

1 Contents

CHAPTER 1 – INTRODUCTION 6

1 CONTENTS 6

2 GLOSSARY 6

3 INTRODUCTION 8

3.1 Nucleic Acids 8

3.1.1 Deviations From the Duplex 15

3.1.2 Junctions 20

3.1.2.1 Holliday Junction 22

3.1.2.2 Quadruplex 27

3.1.2.2.1 G-Quadruplex 27

3.1.2.2.2 i-motif 37

3.1.3 Concluding nucleic acids 39

3.2 Helicates and Similar Structures 40

3.3 Concluding 45

4 REFERENCES 46

2 Glossary

C – Cytosine

T – Thymine

U – Uracil

A – Adenine

G – Guanine

UTR – Untranslated Region

AUR – Upstream of AUG Region

SLB – Stem Loop B

3WJ – Three-Way Junction

4WJ – Four-Way Junction

G4 – Guanine Quadruplex

M – Left hand twist into helicate (ΛΛ)

P – Right hand twist into helicate (ΔΔ)

HMG – High Mobility Group

3 Introduction

3.1 Nucleic Acids

Nucleic acid (nucleinsäuren) is a term coined by Richard Altmann in 1889,[1] building upon his colleague Friedrich Miescher's discovery of a substance within the nuclei of cells containing a high level of phosphorous.[2] Nucleic acids are formed from nucleotides linked by a phosphate backbone which is where the acidity originates. Near to the backbone, but also present in other positions, are stabilising cations.[3]

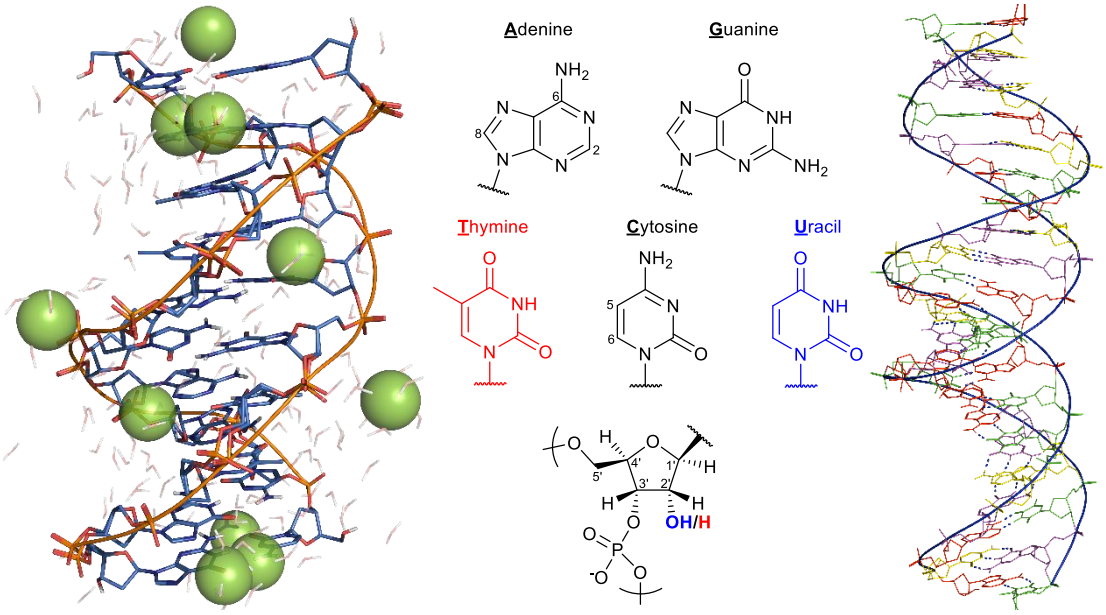


Figure 3.1-1 Left. Crystal structure of B-DNA with water and Mg^{2+} ions. PDB:1EN3.[4] Middle. Basic nucleic acid structure and numbering. Showing the ribose/phosphate monomer, with the nucleobases. Exclusively DNA in red, and RNA in blue. Right. Solution NMR structure of B-DNA, C (yellow), G (magenta), A (red), T (green). PDB:2JYK.[5]

At the 3' and 5' positions ribose units link the phosphates to form the polymer. The sugar differs between DNA and RNA. DNA nucleotides are formed from a 2-deoxy-D-ribose.[6,7] RNA nucleotides are formed from a D-ribose.[8,9] Nucleobases bound to the ribose 1' carbon complete the single stranded structure. The nucleobases fall into two groups of

heterocycles, pyrimidine (Cytosine C, Thymine T, Uracil U) and purine (Adenine A, Guanine G). Where DNA uses T and RNA uses U. (Figure 3.1-1)

The most well-known form of duplex DNA is the helical B-DNA, reported in 1953.[10–12] The bases, A-T and G-C are paired along a phosphate-sugar backbone in all three helical DNAs. The bases are paired through hydrogen bonding between the imine and amine group hydrogens and lone electron pairs. B-DNA experiences a right-handed turn, forming two grooves in the helix.

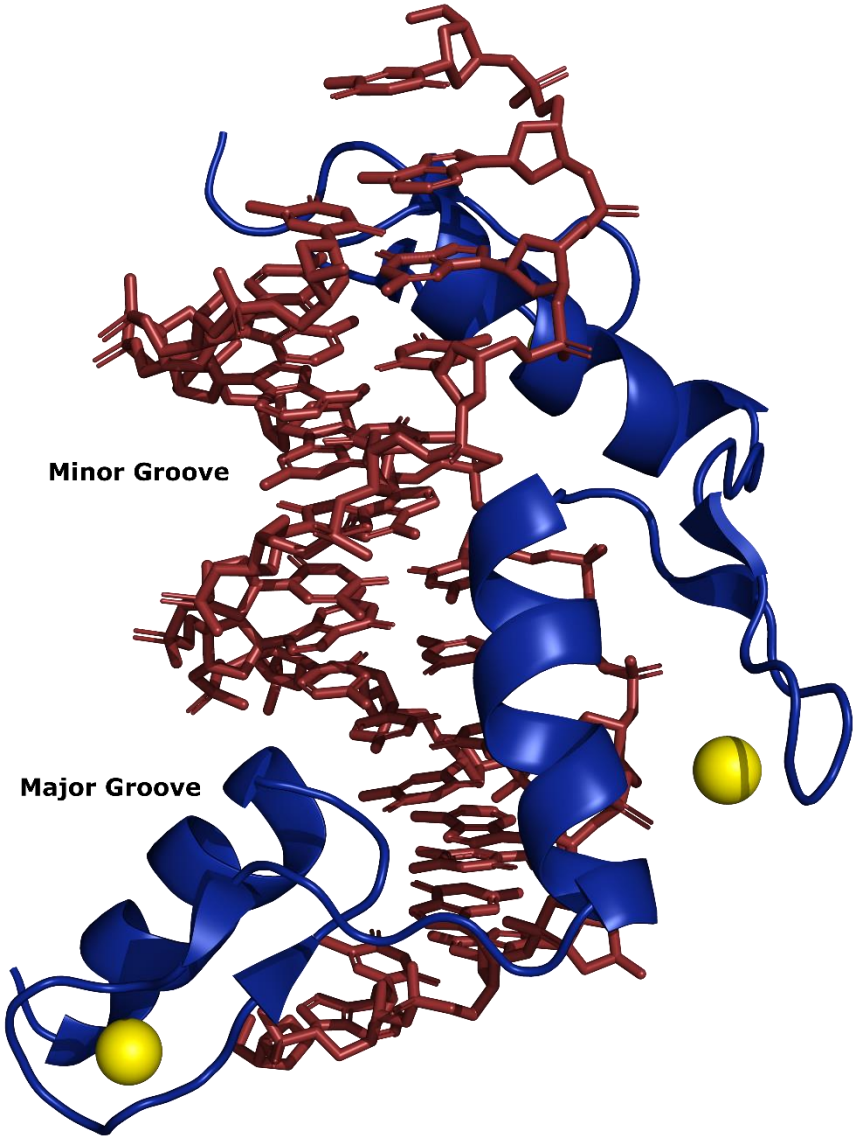


Figure 3.1-2 B-DNA crystal structure showing the two grooves, with a zinc finger protein bound to the major groove. Red: DNA, Blue: Peptide, Yellow: Zinc. Using the crystal structure at PDB:1MEY.[13]

They are the major and the minor grooves and are found on the inside and outside of the base-pairs, walled by the phosphate sugar backbones on either side. The major groove on the base inside of the helix is the larger of the two, consequently it is commonly recognised by proteins such as the zinc finger moiety.[14] (

Figure 3.1-2) B-DNA features 10.4 pairs per full turn of the helix under physiological conditions.[15]

A-DNA also experiences a right-hand turn but has a shorter and wider helix resulting from a lack of stabilising water molecules around the phosphate backbone. It is commonly a dehydrated from of DNA,[16] though some RNA-DNA hybrids can adopt a similar structure.[17] The ribose sugars also adopt an alternative C3' endo conformation. Protein binding is a route to A-DNA where the water can be stripped from the surface, recent work suggests it may be an adaption to extreme environments.[18,19]

Different to the others, Z-DNA experiences a left-hand turn. Though it is not a simple reversal of helicity, the backbone zigzags as the ribosides and phosphates alternate between a diagonal shift and a step up. It is thinner and longer than B-DNA, with approximately 12 base pairs per full turn, and is usually a higher energy state of DNA. Due to the elongated structure, the phosphates on either strand are closer together than in B-DNA. This resulting phosphate-phosphate anion repulsion contributes significantly to the instability of the structure.[20]

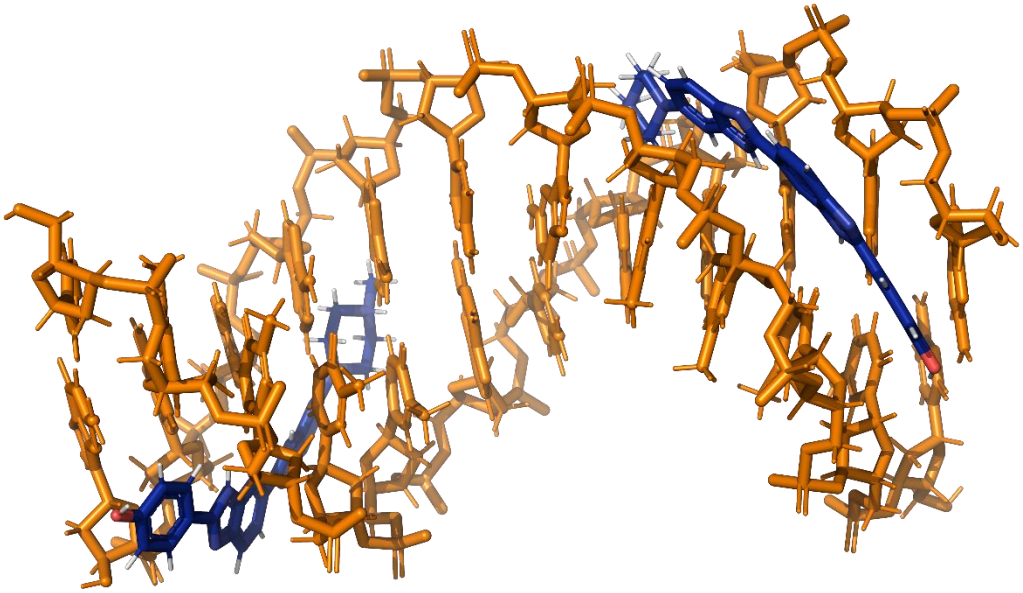


Figure 3.1-3 A-DNA NMR solution structure showing Hoechst 33258 bound to the minor groove of a DNA palindrome. Yellow: DNA, Blue: Hoechst 33258 ligand. PDB:1QSX.[21]

A common nucleic acid stain is minor groove binder Hoechst 33258. Laying parallel to the minor groove its benzimidazole is able to hydrogen-bond with the nucleobases.[22–24] (Figure 3.1-3) Hoechst analogues have been used in a variety of applications, utilizing their DNA binding properties with linkers to achieve more specific goals, such as the inhibition of Escherichia coli RNA and DNA topoisomerase.[25] Similar nucleobase hydrogen bonds can be achieved in the TAR RNA minor groove using a library of amiloride derivatives.[26] By linking two units of binder Hoechst 33258 with a polyether linker it has been demonstrated that the binding affinity to differently distanced A_3T_3 motifs was dependent on the length of the linker. This was shown to be capable of binding across the three-way junction. Though distance limited cooperative binding in this way is more valuable for sequence recognition.[27] TACN conjugates take advantage of various minor groove binders, including Hoechst, to position a metal complex near to the phosphate, achieving DNA cleavage.[28,29]

Intercalators act to insert themselves between nucleic acid base pairs as a form of binding. Like seen for the minor groove binders, intercalators have been used to direct DNA cleavage using linked metal complexes.[30–32]

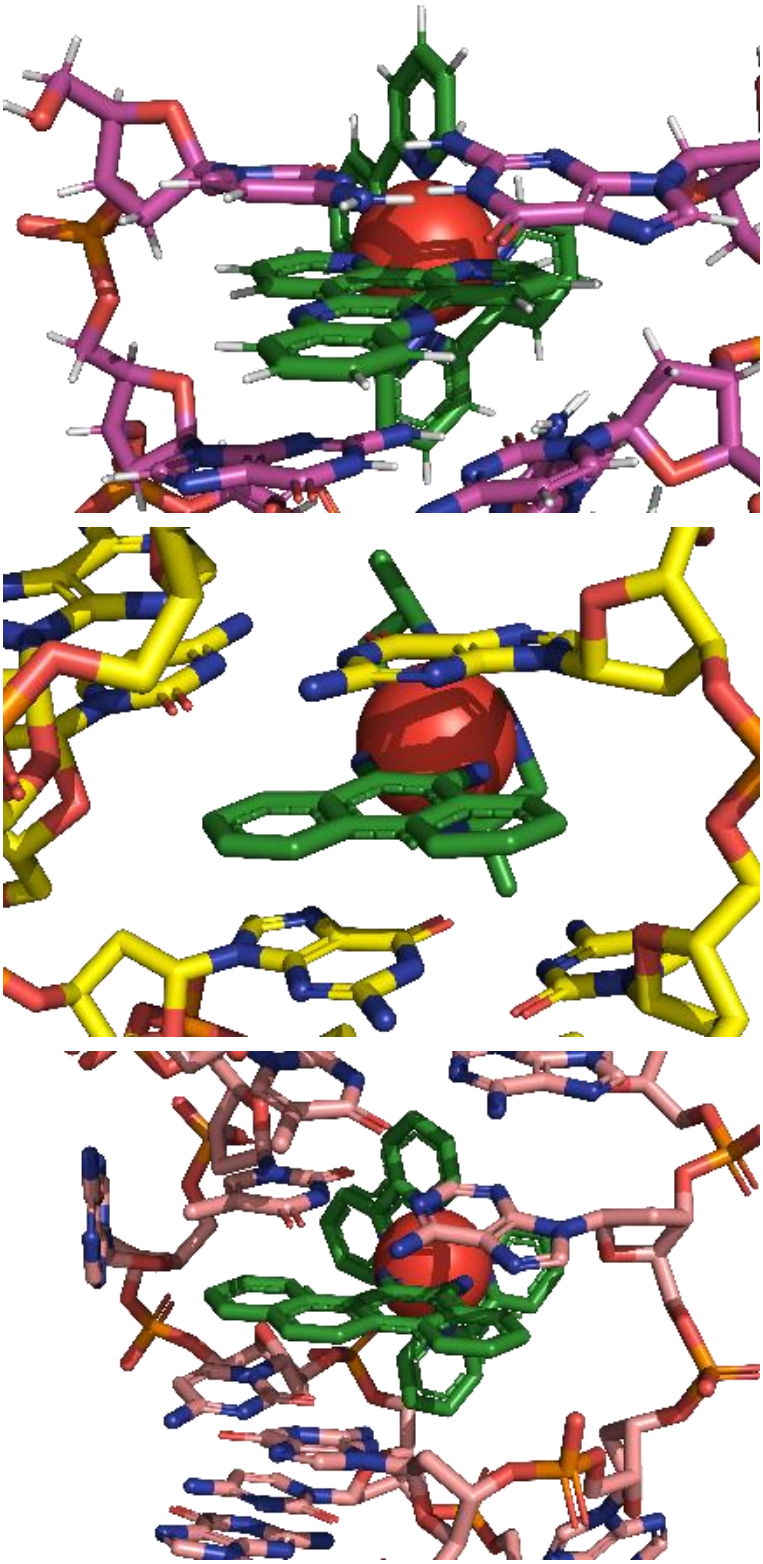


Figure 3.1-4 X-ray diffraction structures of DNA intercalators (green). Top. [Ru(bpy)2 (dppz)]2+. PDB: 4E1U.33 Middle. [Rh(Me2trien) (phi)]3+. PDB:454D.34 Bottom. [Rh(bpy)2 (chrysi)]3+. PDB:3GSK.35

A series of rhodium complex intercalators using different ligands has been built, (Figure 3.1-4) with recent developments in a DNA mismatch targeting 5,6-chrysenequinonediimine (chrysi) ligand.[36] Linking a $[Rh(chrysi)(phen)(PPO)]^{2+}$ complex to a cyanine 3 dye provides a means to image these complexes in cells.[37] A DNA photosensitising 1,4,5,8-tetraazaphenanthren (TAP) ligand can be delivered on a $[Ru(TAP)_2(dppz)]^{+}$ complex, taking advantage of the dipyrido[3,2-a:2',3'-c]phenazine (dppz) intercalator.[38,39]

Another interesting route of research is how the chirality of the threading molecule affects the binding kinetics of a dinuclear intercalator. Rate constants were extracted from NMR exchange peaks, through varied mixing times and temperatures.[40]

3.1.1 Deviations From the Duplex

Many RNAs are single stranded, such as mRNA, tRNA and miRNA. Single stranded nucleic acids are highly flexible and can adopt many structures. However often under certain conditions one structure predominates with various functional elements. Hence by transitioning to alternative conformers binding affinity for structurally specific ligands can be altered. A good example of this is the 5' UTR of mosquito borne flaviviruses such as Dengue virus or West Nile virus, which are capable of genome cyclization from a linear to circular form.[41] Stabilisation of the 5'-UAR flanking stem positioned below stem loop B (SLB) has the effect of impairing cyclization and viral RNA replication.[42]

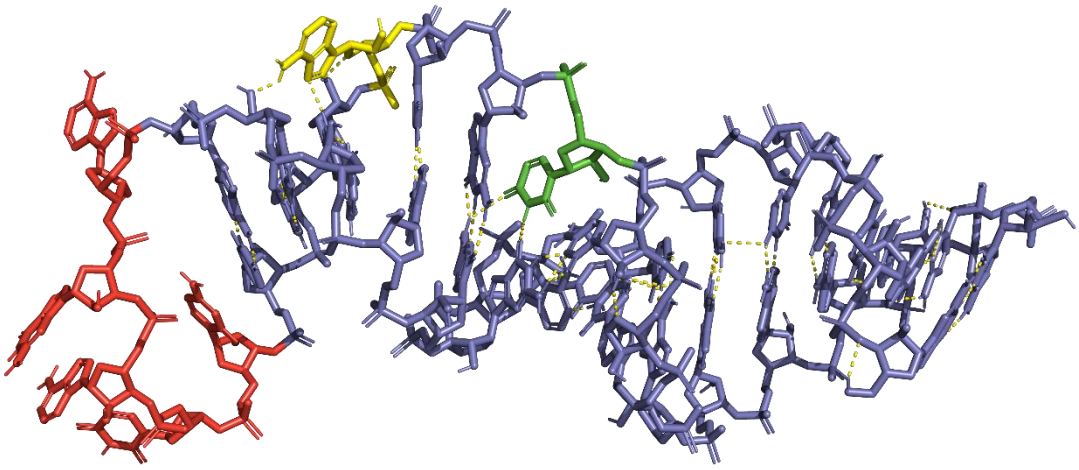


Figure 3.1-5 NMR solution structure of the West Nile virus SLB in its linear form. RNA in blue, bulges in yellow and green, loop in red. PDB:6U79.[43]

SLB contains a nucleotide sequence known as 5'UAR that has a complementary strand in the 3'UTR. SLB differs between viruses, but in the linear form all are held in a stem-loop conformation. For West Nile virus, the structure contains two one base bulges.[42] Helices can form bulges when one or more non-complementary bases loop out from the main helix.[44] The lower bulge base (green) can be seen stacked between the bases above and below, whereas the upper bulge base (yellow) is looped out, interacting with the backbone. Various biological processes require or induce DNA bulges;[45,46] they are commonly found in RNA sequences and are involved in DNA frameshift mutation.[47,48] Interestingly frameshift mutations can be stabilised by small-molecule half-intercalation, leading to the possibility that this can interfere with mismatch repair.[49] At the top, the loop (red) is a region of unpaired bases. In this case and in many others, the loop bases experience greater freedom of movement compared to the duplex.[43] (Figure 3.1-5)

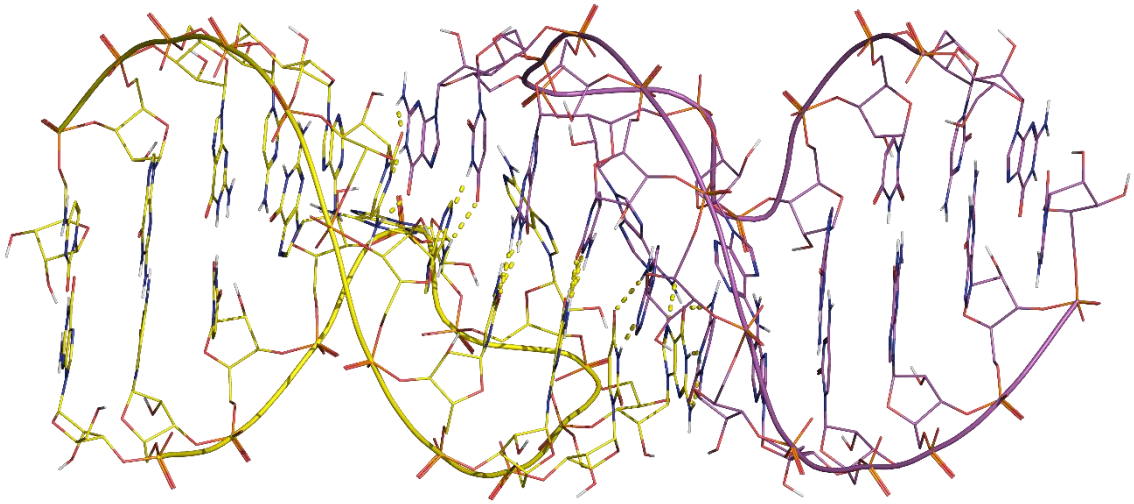


Figure 3.1-6 NMR solution structure of an RNA kissing loop with one strand in yellow and the other strand in purple. Intermolecular hydrogen bonds are shown with yellow dashed lines. HIV-1 dimerization initiation site. PDB: 2D19.[50]

Mobile bases in the free loops are in some cases able to base pair with other complementary loop sequences, these are known as kissing loops. Kissing loops enable distant regions of RNAs to be held in proximity. (Figure 3.1-6) Some attempts at inhibiting the dimerization of HIV-1 have been made by targeting the dimerization initiation site, though no "small-molecule" attempts have been found. Instead, these are solutions designed to interact with the loop nucleotides using peptide nucleic acids and amino glycosides.[51–54]

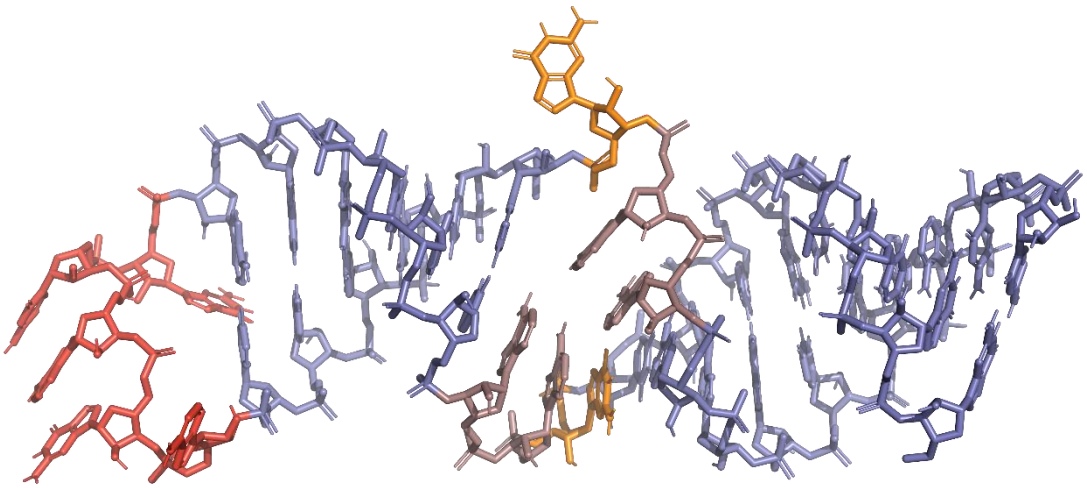


Figure 3.1-7 NMR solution structure of the Dengue virus SLB in its linear form. Internal loop pink (adenine) and orange (guanine). Loop in red. PDB:6W3M.[43]

Another form of loop is the internal loop, similar to a bulge except the non-complementary bases occur on both strands. In this case there are two A-A Watson:Watson cis pairings, holding the loop semi-closed. This pairing is also seen in PDB:2VQE Residues A16-A919. The Gs at the top and bottom of the internal loop are looped out. (Figure 3.1-7)

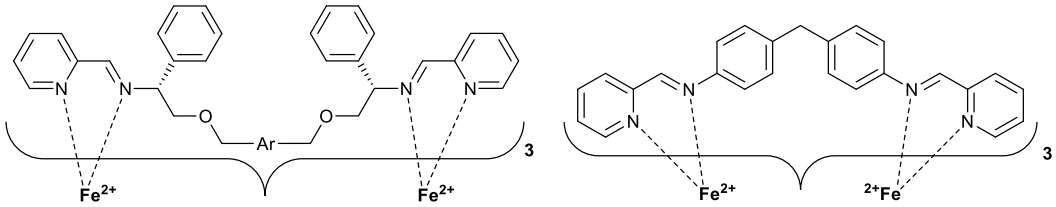


Figure 3.1-8 Helicates with the ability to bind to bulges and loops. Left uses various aromatic spacer units to bind to TAR RNA bulges of varying composition.[55] **Right binds to DNA bulges of varying composition,**[56] **as well as the TAR bulge.**[57–59]

Bulges and internal loops are of particular interest for helicate binding. A three-dimensional drug is aided by ready access to the helix grooves with proximal mobile bases that can interact with the drug further. Recent molecular dynamics simulations suggested that helicates might be capable of binding into the centre of bulges.[57] Helicate

binding is dependent on the bulge size and composition, suggesting a preference in the flexibility and availability of the binding surfaces at the site for certain helicates. Interesting pyrimidine preference has been found for both types of helicate.[55,56] This matches well with the binding seen for the TAR UCU bulge.[57–59] (Figure 3.1-8)

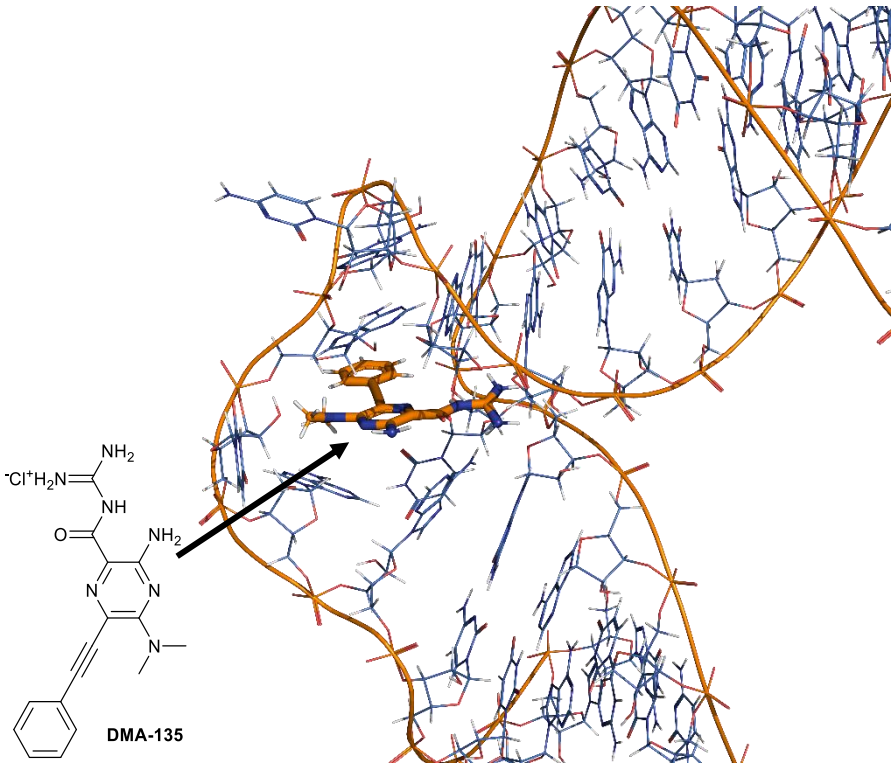


Figure 3.1-9 Left. Structure of DMA-135 bulge binding ligand. Right. NMR solution structure of DMA-135 bound to the bulge of human enterovirus 71 5'UTR stem loop II (residue 128-168). PDB:6XB7.[60]

Screening multiple compounds against a target is a useful method to streamline research. Fluorescence indicator displacement assays identified DMA-135 as a suitable drug for the human enterovirus 71 stem loop II. DMA-135 attenuates translation and replication dependent on the 5'UTR. DMA-135 was then shown to bind to the bulge and induce a significant conformational change.[60] (Figure 3.1-9)

3.1.2 Junctions

Nucleic acids can form junctions. Three-way junctions (3WJ) are one such junction that connects three duplex sections of nucleic acid.[61] There are many configurations and applications of 3WJs, but all form at the convergence of three helices.[62] Each corner of the three-way junction has a branch point where the duplex splits. These branch points are found as a transition between many motifs, such as bulges and hairpins as well as the replication fork, and are associated with many neurodegenerative diseases such as Huntington's and myotonic dystrophy.[63]



Figure 3.1-10 NMR solution three way junction structures. Left. Neurospora VS ribozyme.[64] Right. $PreQ_1$ riboswitch.[65] Backbones are highlighted in orange and pink respectively.

3WJs are commonly found in ribozymes, deoxyribozymes and riboswitches. (Figure 3.1-10) In the case of ribozymes and deoxyribozymes the nucleic acid and metals

cooperate to assume the correct conformation.[66,67] The most common type nucleases, function to cleave nucleic acids with the use of a metal cofactor.[68,69] In the case of the B. subtilis cobalamin riboswitch the conformation is specific to various cobalamin cofactors with the exclusion of cyanocobalamin.[70]

One method of binding to 3WJs is with structures that present three sides, taking advantage of the three branch points/cavity walls. This has been done using calix[3]carbazoles,[71] or azacryptands that position planar biphenyl systems into the junction.[72] Another solution to achieving this arrangement is to use bimetallic helicates.[73–75] These have the advantage of a net positive charge to further interact with the anionic phosphate groups.

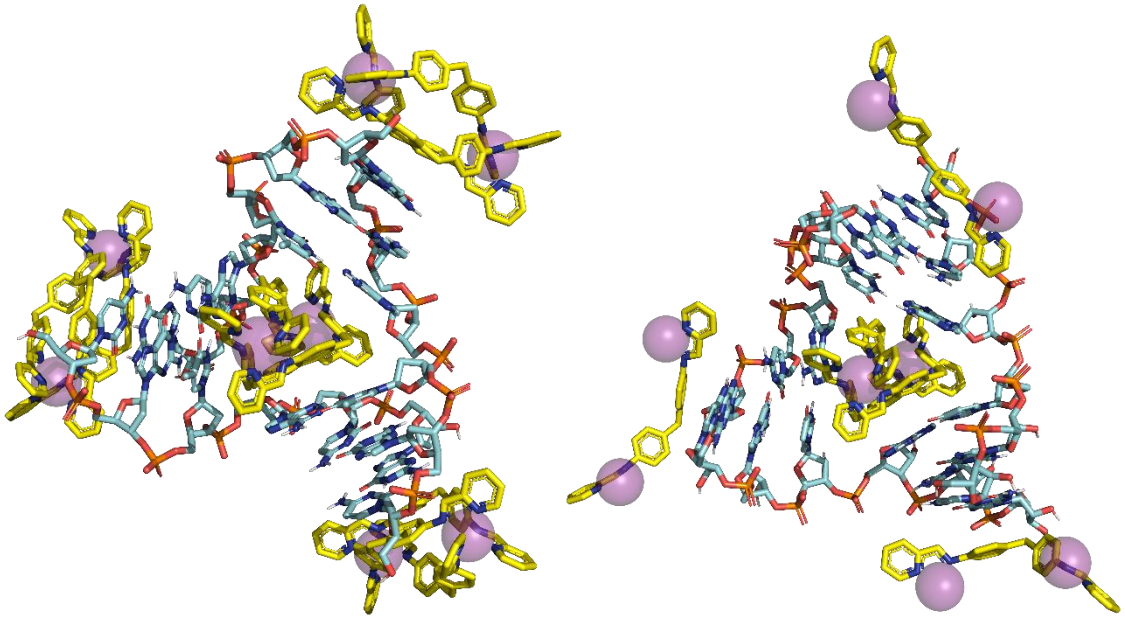


Figure 3.1-11 Crystal structures of bimetallic helicate 3WJs. Left. DNA 3WJ (CGTACG). PDB:2ET0.[74] Right. RNA 3WJ (CGUACG). PDB:4JIY.[76] Iron centers in purple, helicate ligands in yellow, nucleic acids in light blue with heteroatoms highlighted.

The helicate rotatable phenyl groups are well positioned to form a stacking interaction with the 3WJ bases. The cylinder is able to stabilise the 3WJ by binding in the central

cavity. The 4,4'-methylene dianiline structure enables a precise fit to the centre of the junction regardless of substitutions to the bridging CH_2 with O and S,[77] unlike other larger cylinders that do not afford the same stabilization.[78] For the 3WJ, the M enantiomer has a greater stabilising effect, it is proposed that this is a result of the similar helical nature of the complex to the twist at the junctions center.[75] The same cylinder is also able to bind and stabilise an RNA 3WJ to a similar degree as the DNA 3WJ, both in solution and as a crystal.[76] (Figure 3.1-11)

3.1.2.1 Holliday Junction

The four-way junction (4WJ) or Holliday junction, is a cruciform convergence of four duplex helices. It is named after Robin Holliday, who proposed a mechanism for gene recombination that uses the 4WJ,[79]

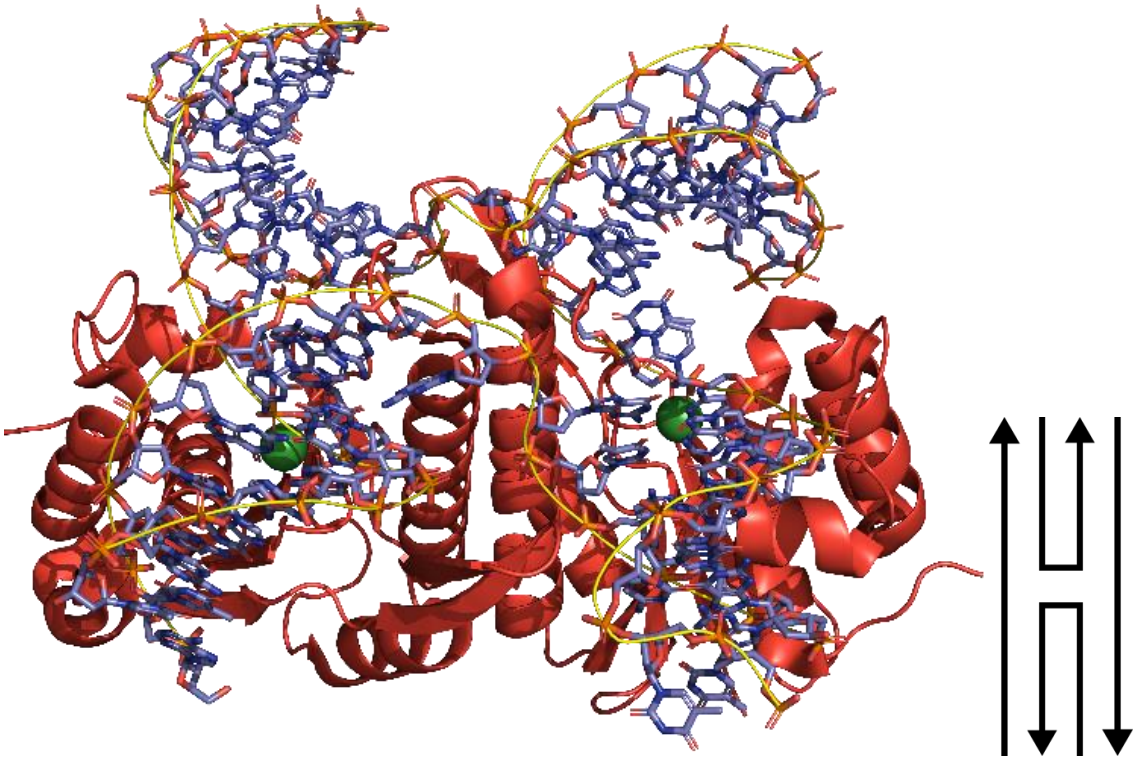


Figure 3.1-12 Left. Crystal structure of a DNA Holiday junction (blue with yellow highlighted backbone) bound to monokaryotic chloroplast 1 HJ resolvase (red). PDB:6IS8.[80] Right. Cartoon of a Holiday junction.

They are found in nature as part of homologous recombination and repair mechanisms; a 4WJ is able to migrate up and down each of its strands by breaking and pairing bases as it moves.[81,82] (Figure 3.1-12)

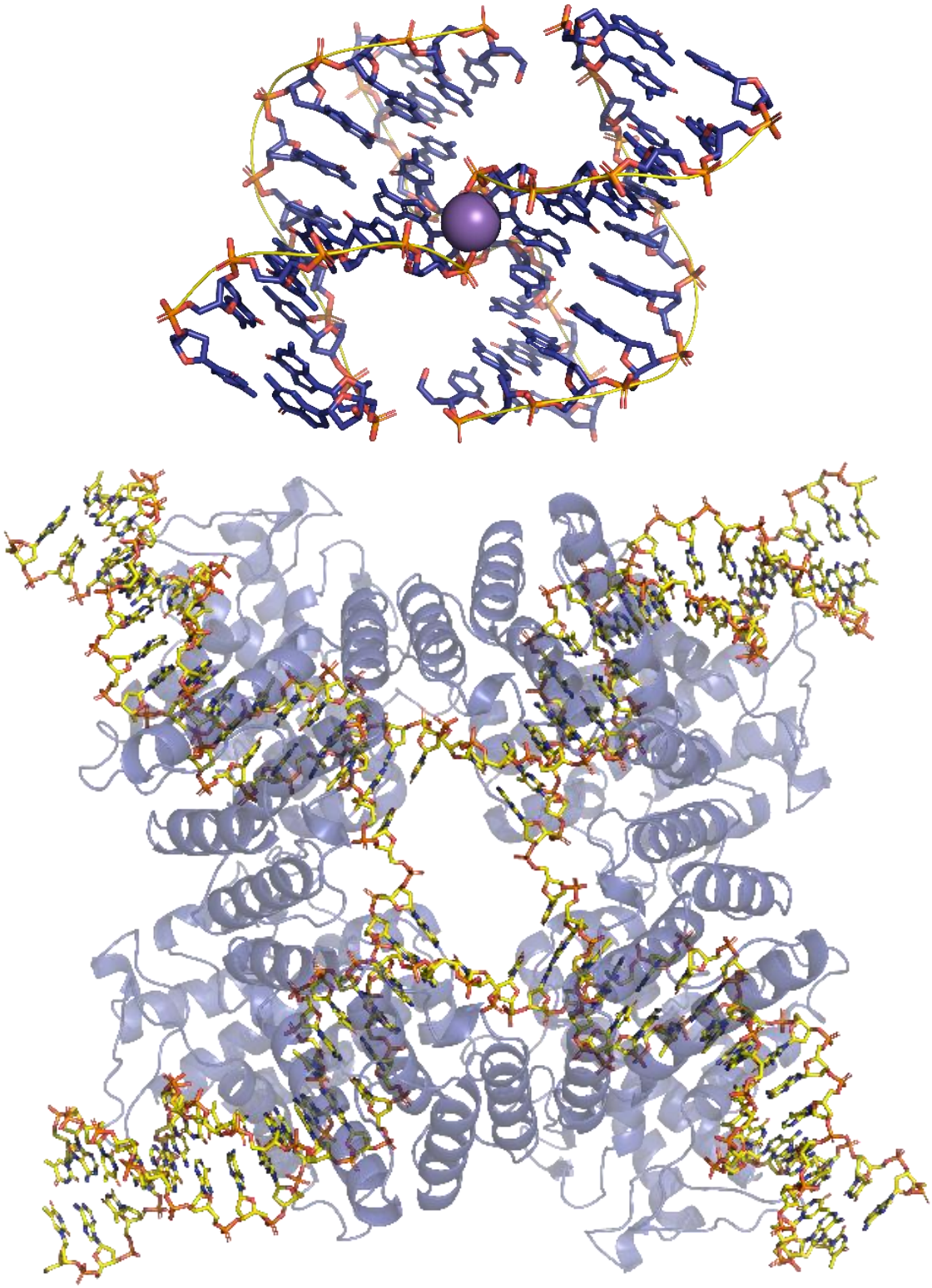


Figure 3.1-13 Top. Crystal structure of a closed form of 4WJ (blue DNA with yellow highlighted backbone, central sodium ion in purple). PDB: 4EZ2.[83]. Bottom. Crystal structure where a 4WJ is complexed with Cre recombinase which induces an opened form (DNA in yellow, CRE recombinase in blue). PDB: 3CRX.[84]

They exist in open and closed states and the positions of the four helices can align in a planar formation or form a pseudo-tetrahedral structure. In order to facilitate branch migration, the 4WJ is moved into the open state, this is done biologically using structures such as Cre recombinase and RuvA, which is one of three proteins in the RuvABC complex. (Figure 3.1-13)

4WJ DNA can be selectively recognised by high mobility group (HMG) protein 1.[85,86] which has been shown to preferentially bind to cisplatin modified DNA.[87] HMG-1 (High Mobility Group protein 1) is a protein with two HMG-box binding domains connected by a 16 base-pair oligonucleotide.[88] Together they have a greater affinity for DNA, as well as a greater ability to bend and supercoil DNA.[89]

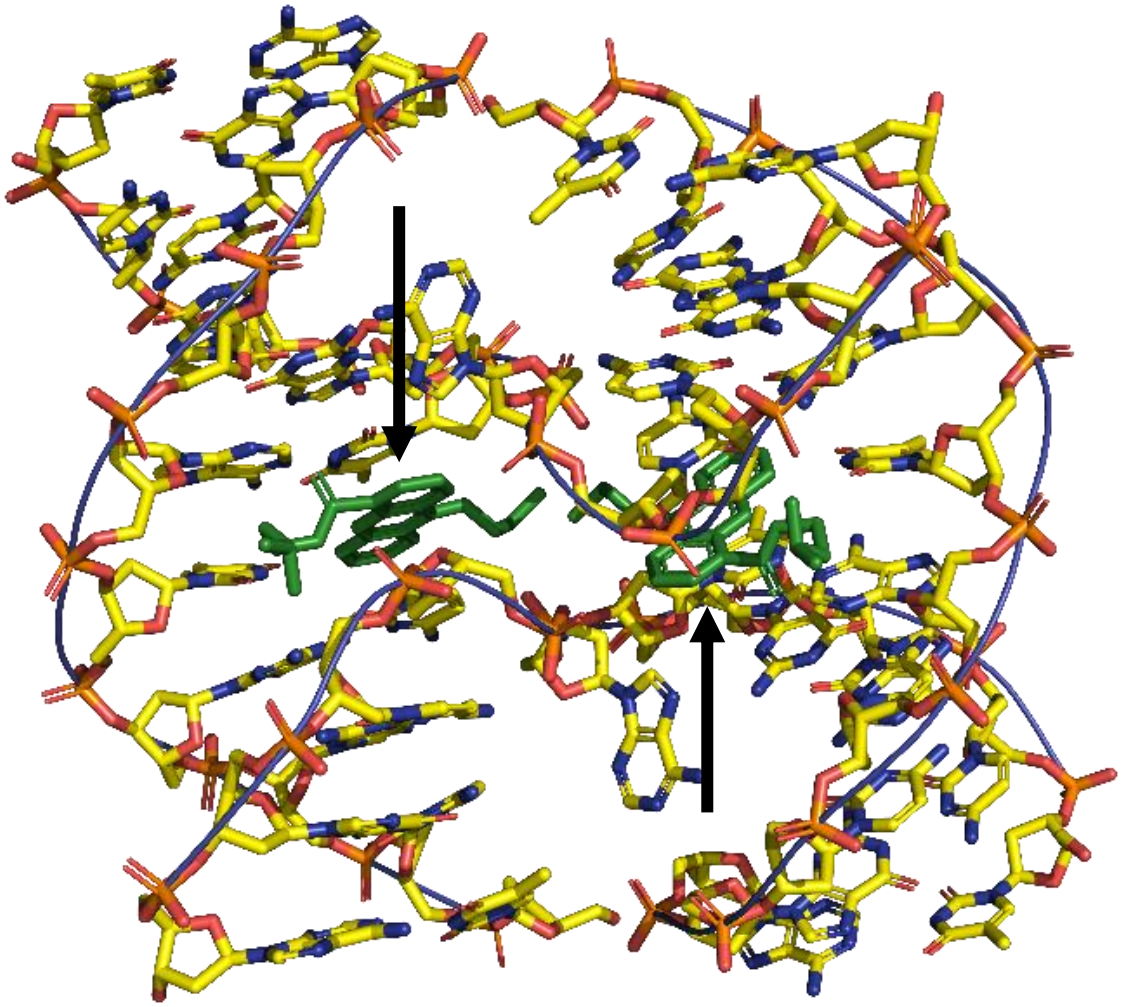


Figure 3.1-14 Crystal structure of a DNA holiday junction (yellow DNA with blue highlighted backbone), bound with two bis-acridine molecules (yellow) indicated with black arrows. PDB:2GWA.[90]

Synthetic 4WJ binders include the bis-acridine molecule, which binds across the centre of the 4WJ. The two acridine moieties stack onto the bases in the junction crossover region, the side chains fitting into the minor grooves of the junction.[90] (Figure 3.1-14)

The interaction between a $Fe(II)_4L_4$ tetrahedron and unpaired DNA bases has been explored, where the four ligands bound to octahedral iron centres are formed from a trianiline centre with three pyridine groups. The tetrahedra are capable of binding to single-stranded, double-stranded, 3WJ, and 4WJ DNA. Using the fluorescence quenching properties of the tetrahedron, the distance from the binding site to the fluorescently

labelled DNA can be gauged. Though able to bind to the 4WJ, the structure showed greater affinity for 3WJ structures.[91]

3.1.2.2 Quadruplex

Quadruplex nucleic acids as the name suggests are formed from four interacting strands.

3.1.2.2.1 G-Quadruplex

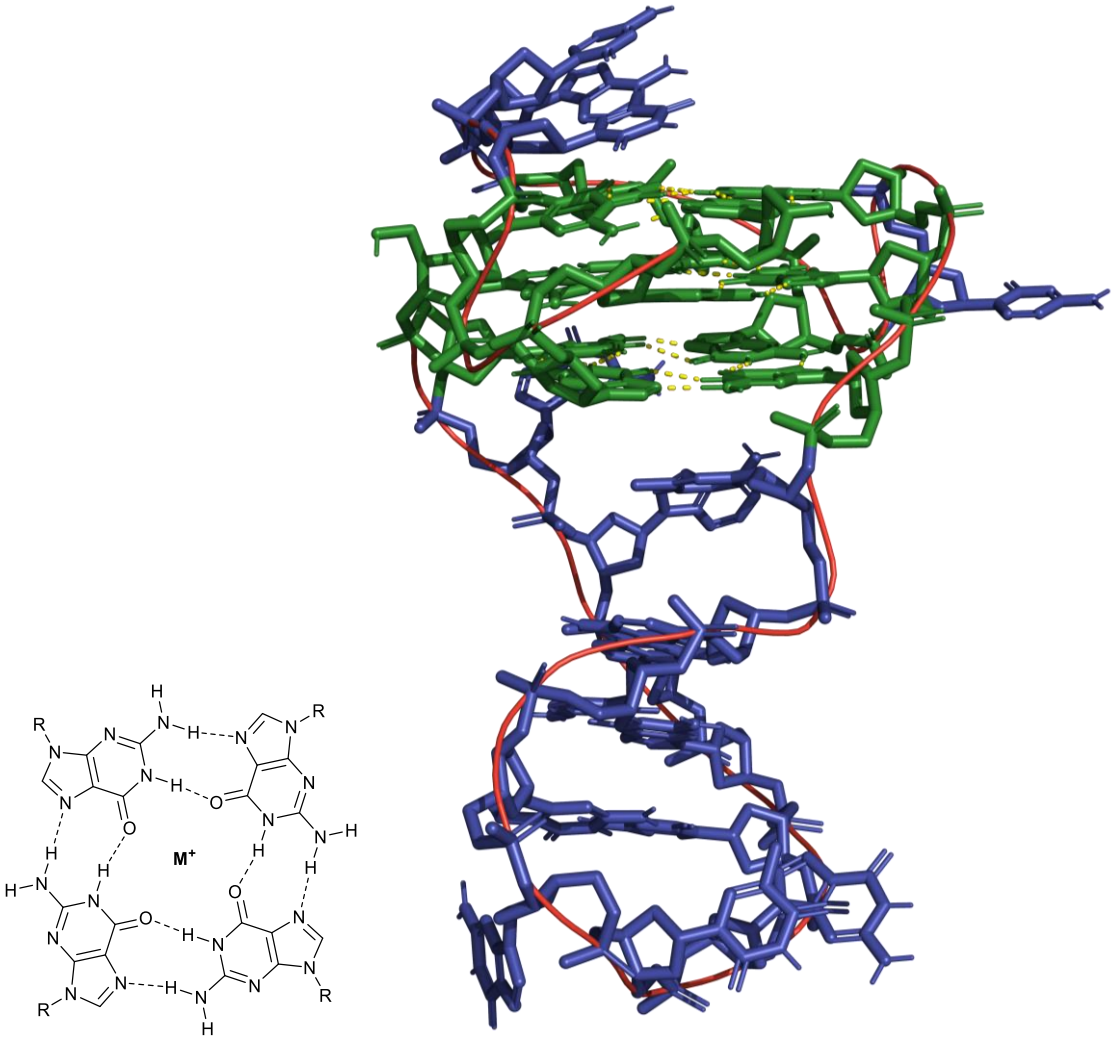


Figure 3.1-15 Left. Structure of a G-quartet/tetrad. Right. Solution structure of a DNA (blue with red highlighted backbone) G-Quadruplex (green) form of HIV-1 LTR-III. PDB:6H1K.[92]

The G-Quadruplex (G4) is a structural motif formed by stacked planar guanine quartets. A guanine quartet consists of four guanine bases hydrogen bonded between their

Hoogsteen and Watson-Crick edges in an ouroboros like end-on-end arrangement. Metal ions such as sodium or potassium stabilise the quartets, these are found in the interstitial space between the planes.[93] G4s are found in guanine rich sequences, such as telomeric nucleic acids and telomerase RNA component.[94] As well as in the RNA 5'-UTRs of the human genome.[95,96] They are present during normal cellular events,[93,97] such as transcription, replication, and translation.[98] As well as a binding site during DNA repair.[99] In addition, G4s are found in the HIV-1 3'UTR,[100,101] and the 5'-UTR,[102,103]. G4s are polymorphic, so can be formed from one (intermolecular) or multiple (intramolecular) strands. There is no limitation on direction of the strand other than what affects the spatial arrangement on the guanine bases. G4s, where all four guanine-rich tracts face the same direction are called parallel, those where one is opposing the other three are hybrid, and antiparallel is where equal numbers face each direction.[104] (Figure 3.1-15)

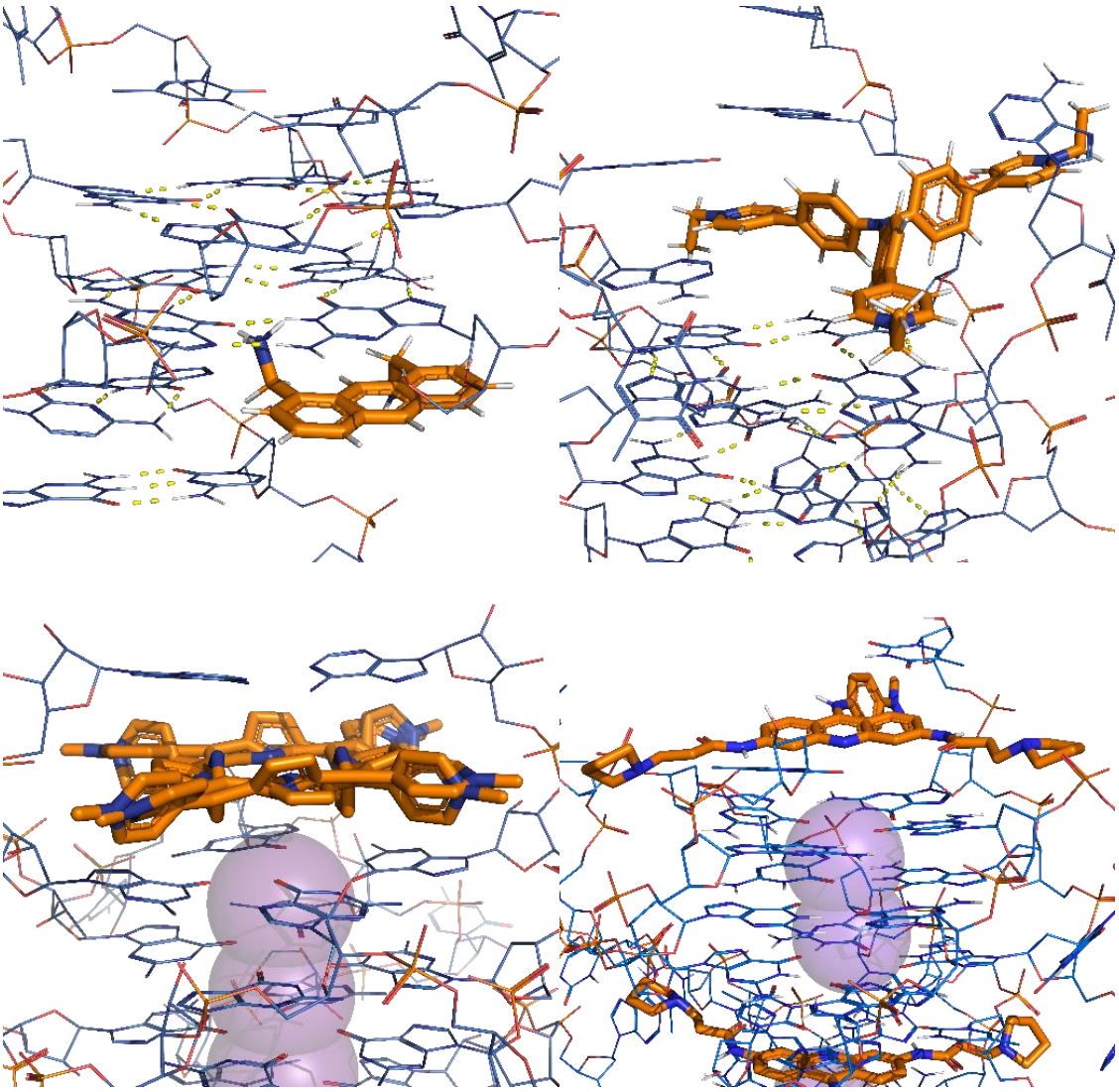


Figure 3.1-16 A collection of G4 binding ligands bound to G4 structures. Where the ligands are orange, the DNA is blue and potassium ions are in purple. Top-Left. Solution structure of a 1,8-bis(aminomethyl)anthracene ligand bound to a guanine on the base of a DNA G4-duplex. PDB:6FC9.[105] Top-Right. Solution structure of a NBTE fluorescent probe bound to the top of a DNA Tel26 G4. PDB:6KFI.[106] Bottom-left. Crystal structure of the top of an RNA G4 bound to TMPyP4 porphyrin (two ligand arrangements overlayed). PDB:6JJI.[107] Bottom-Right. Crystal structure of DNA G4 bound to acridine based BRACO19 ligand. PDB:3CE5.[108]

Initial research for G4 binding has been on small molecules, such as anthraquinones. One of the first ligands to bind and stabilise G4 structures inhibiting telomerase was a modified anthraquinone.[109] With further research into a neomycin-anthraquinone conjugate giving a high binding affinity.[110] Another method has been to use porphyrins, due to their similarities to the G4 in size, shape, and hydrophobicity.[111,112] Metal ions

can be coordinated to the centre to improve their G4 binding.[113] The structure can also be improved by adding hydrophilic groups to reduce self-aggregation and increase water solubility.[114] Interestingly the porphyrin TmPyP4 has been shown to unfold a regulatory all purine G4, increasing the rate of translation.[115]

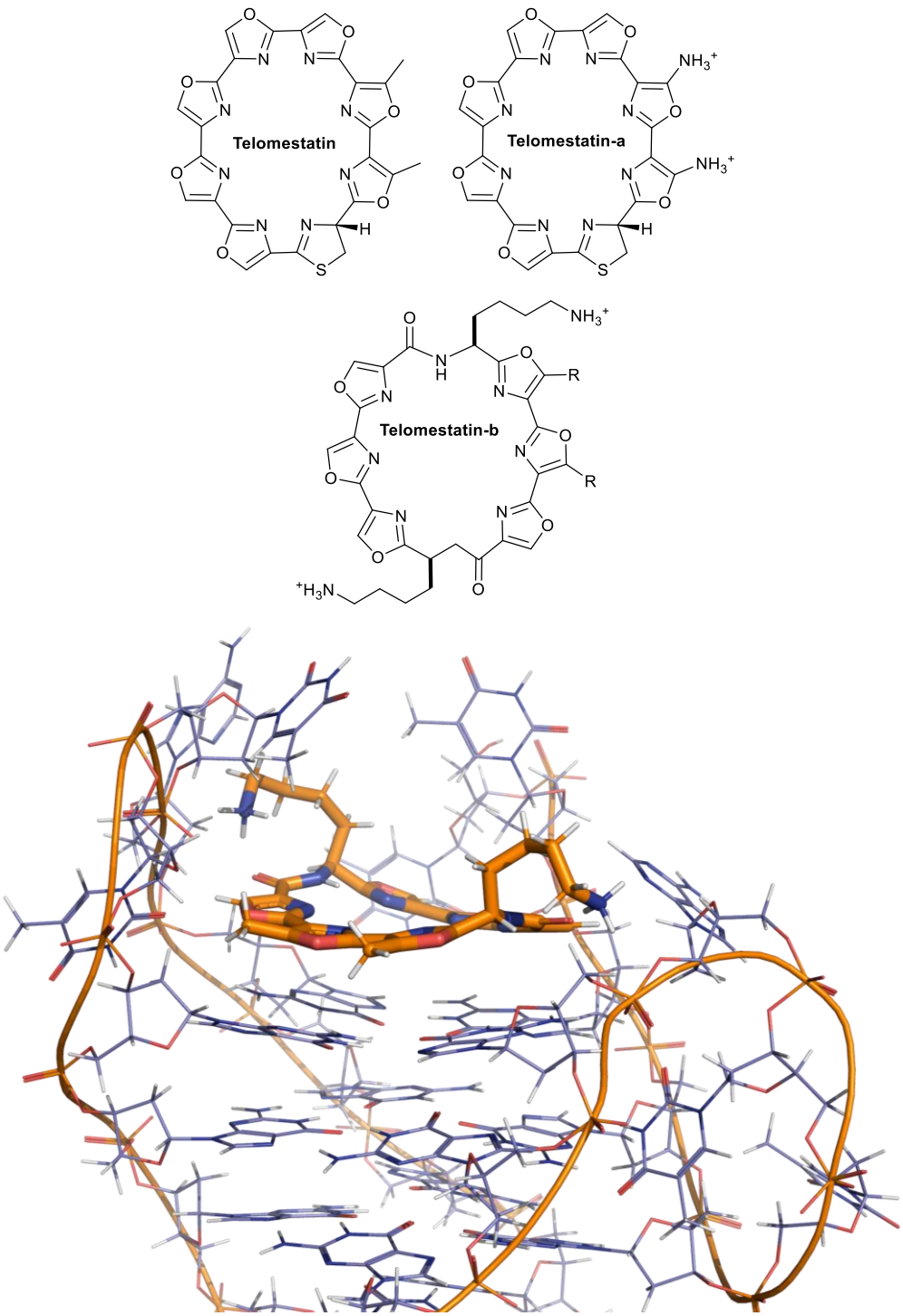


Figure 3.1-17 Top. Telomestatin structures. Telomestatin-a as a proposed improvement.[116] Telomestatin-b as the derivative used in the structure below. Bottom. Solution NMR structure of telomestatin-b derivative bound to a human telomeric G4. DNA in blue with yellow highlighted backbone, telomestatin in orange. PDB:2MB3.[117]

Found in *Streptomyces Anulatus* a soil dwelling bacterium,[118] telomestatin is a telomerase inhibitor which gets its potency from binding to the G-quadruplex.[119] A

series of molecular dynamics experiments on popular G4 binding agents concluded that the addition of two ammonium groups that disrupt the backbone would aid in the binding of the ligand to the G4.[116] This has similarities to another derivative shown to bind to a human telomeric G4. In this study, they noted that it is beneficial to feature flexible planar surfaces that can be angled to better interact with the tetrad, rather than rigid planar surfaces.[117] (Figure 3.1-17)

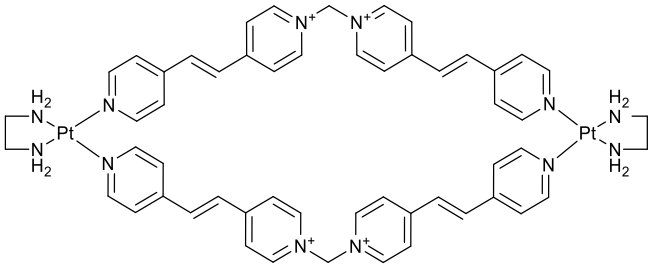


Figure 3.1-18 Self-assembled Pt2L2 box that displayed activity against VM-1 melanoma.

Using a series of self-assembled Pt(II)quadrangles to target G4 DNA showed interesting inhibition of U2OS osteosarcoma cells with the ligand above. On complexing into the Pt-quadrangle an increase in inhibition is seen. Inhibition is also seen for the VM-1 melanoma cell line reported to be cisplatin resistant. Other sized ligands and complexes showed different or lesser binding affinity for various G4 and duplex DNAs, showing the importance of fine-tuning binding agent characteristics.[120]

Mentioned above for use with 3WJs, azacryptands have recently been explored for use with G4 structures, showing that the structure is capable of a dual targeting effect. These can be modified for click chemistry, enabling their use for cellular imaging.[121]

The first molecule reported to selectively induce a parallel quadruplex from unfolded DNA was a modified anthraquinone.[122] Generally, the running theme for these G4

binding ligands is a heteropolyaromatic planar structure to stack onto the quartet, sometimes with a flexible cationic chain to extend into the backbone.[123–125] (Figure 3.1-16)

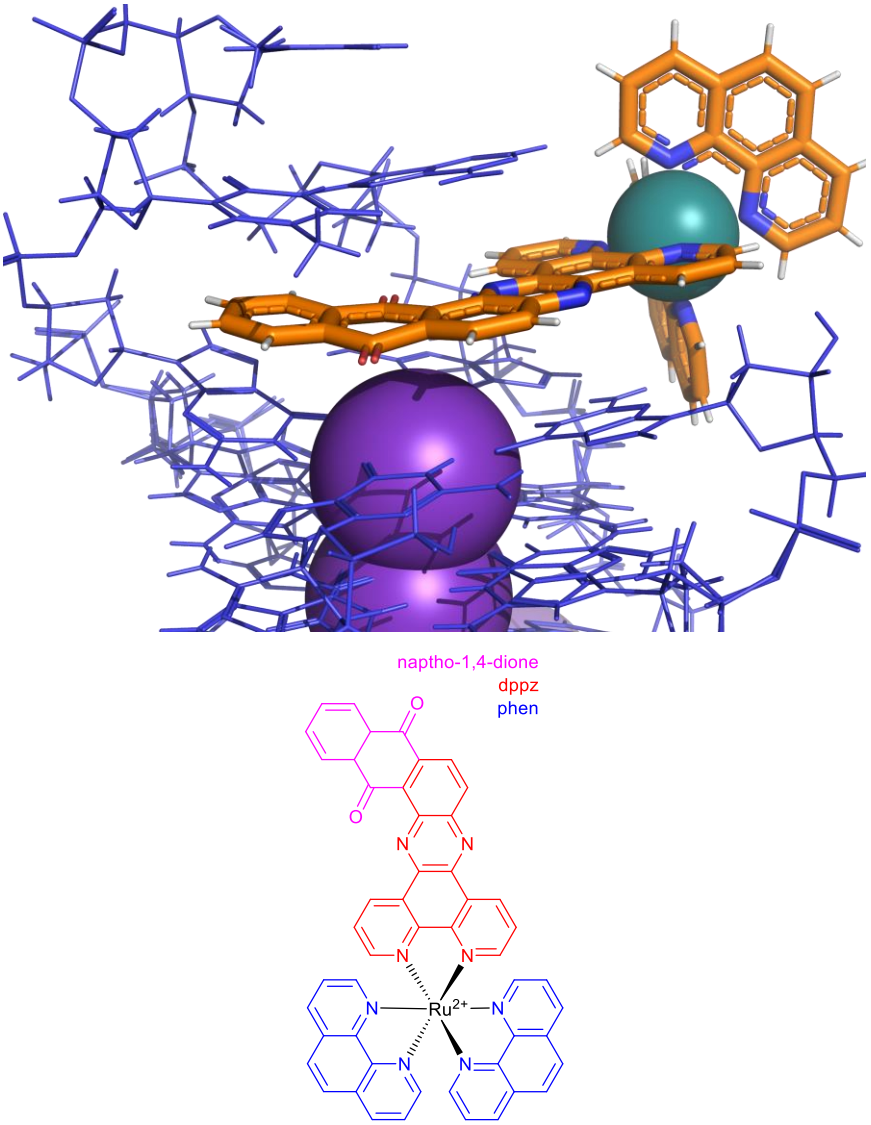


Figure 3.1-19 Left. Crystal structure of a DNA G4 bound to a Λ-[Ru(phen)$_2$(qdppz)]$^{2+}$ ligand. DNA in blue with potassium ion in purple, complex in orange with a teal ruthenium center. PDB:7OTB.[126] Right. Λ-[Ru(phen)$_2$(qdppz)]$^{2+}$ complex.

Similar principles can be achieved using metal complexes, with use of large planar dppz ligands.[127,128] A modified dppz ligand, qdppz is improved for G4 binding with the addition

of a naptho-1,4-dione type moiety. The dione groups are suitably placed to interact with the central G4 cation in a manner reminiscent of anthraquinone.[126] (Figure 3.1-19)

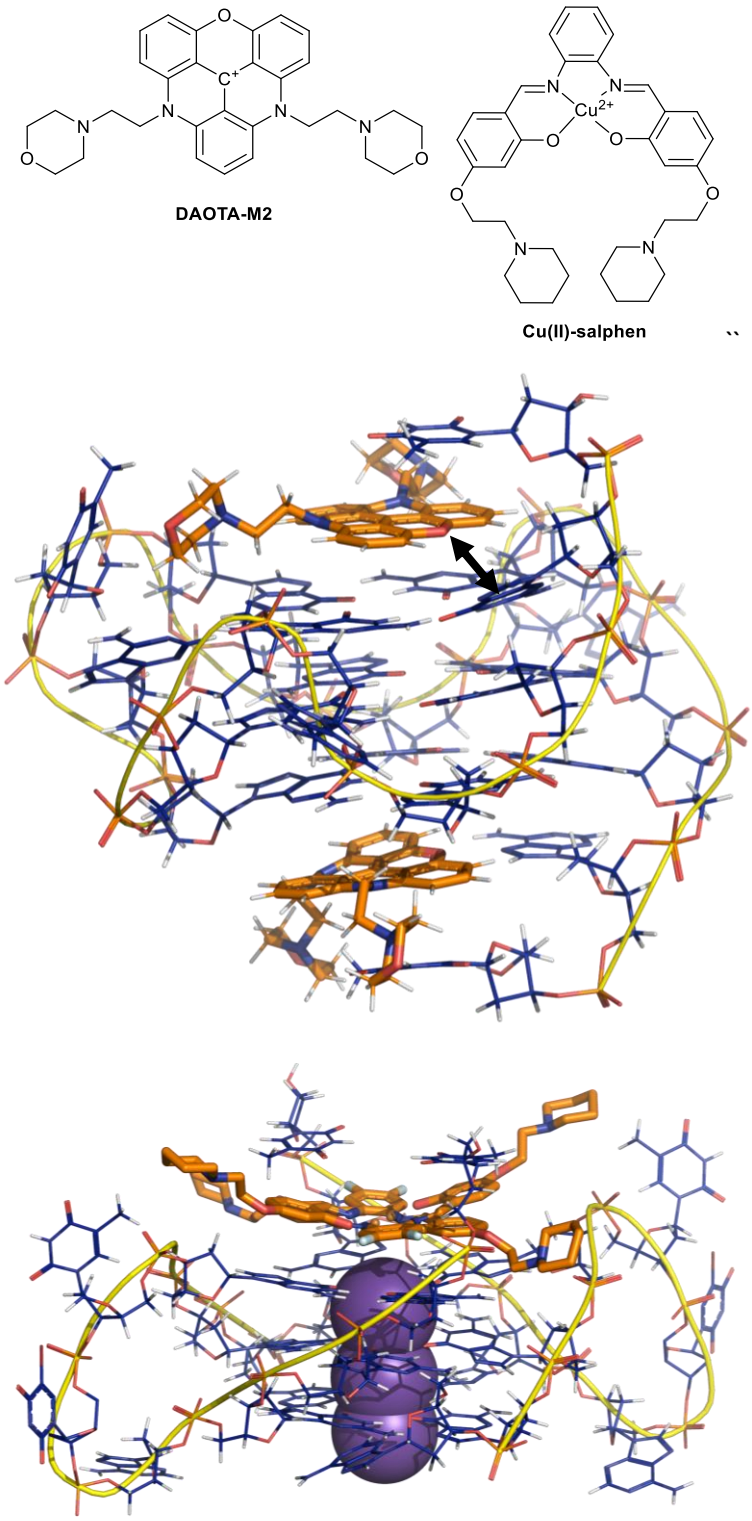


Figure 3.1-20 Top. Structures of the ligands used. Middle. NMR solution structure of DNA G4 (blue with backbone highlighted in yellow) bound to DAOTA-M2 (orange). A tetrad guanine is angled out of plane into the central

DAOTA-M2 alkoxy group shown with a black arrow. PDB:5LIG.[129] Bottom. Crystal structure of DNA G4 (blue with backbone highlighted in yellow) bound to a copper(II)-salphen complex (orange). PDB:3QSC.[130]

The central cation can be part of the binding ligand. This is seen in the fluorescent triangulenium ligands with a central carbocation,[131–133] and can also be achieved using a central metal ion, in the case of salphen ligands.[134,135] This is confirmed when a low G4 binding affinity octahedral Pt(IV) salphen is reduced to remove two z-axis Cl, leaving a square planar Pt(II) complex with a higher G4 binding affinity.[136] Non-planar helicene ligands can also be used to achieve higher selectivity than their DAOTA counterparts.[137] (Figure 3.1-20)

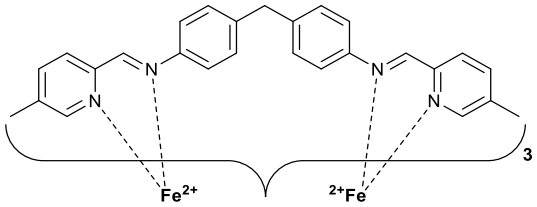


Figure 3.1-21 Modified cylinder found to have the highest chiral selectivity, with 5-methyl pyridine groups.[138]

Chiral metallohelices are able to enantioselectively interact with telomeric G4 DNA, inhibiting telomerase activity.[139] In addition, they have been applied to breast cancer stem cells in vivo showing enantioselective activity against them, with the intended target being telomeric G4s.[140] Using a collection of similar helicates to those mentioned before for junction binding, a helicate with a 5-methyl pyridine was shown to have the greater chiral selectivity for G4 binding. The chiral enantiomers had significantly different binding constants to D-DNA Vs L-DNA G4s. Other helicates with no or differently positioned methyl groups showed strong binding, but lesser selectivity.[138] (Figure 3.1-21) Further research on how exactly this effect occurs could be interesting.

3.1.2.2.2 i-motif

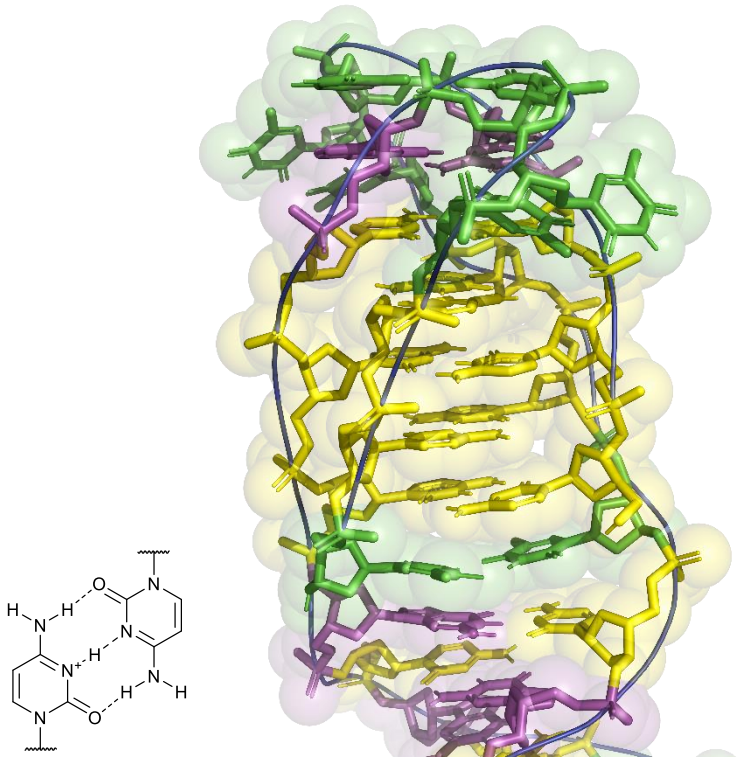


Figure 3.1-22 Left. Hemi-protonated cytosine-cytosine base pair. Right. NMR solution structure of an i-motif/duplex sequence. Yellow = C, Green = T, Purple = G, Red = A. PDB:7O5E.[141]

Believed to be mutually exclusive with respect to G4s,[142] the i-motif is a nucleic acid structure; formed from intermeshed, cross-linked, hemi-protonated cytosine bases.[143] (Figure 3.1-22) Accordingly, it is found in cytosine rich sequences of the DNA, such as the telomeres.[144] The i-motif is thought to modulate replication and promote transcription.[145–147] Up until recently there had been no evidence for the i-motif in vivo as no specific binding agent had been created. Using a method modelled on an earlier study of the G-Quadruplex,[148] it was co-localised with the promoter regions of proto-oncogenes showing its presence in regulatory regions.[149] The i-motif was shown to vary in frequency with the cell cycle, increasing in prevalence in the G1/S boundary and

decreasing in the early S phase; which is noted to be in contrast to the G-Quadruplex, that occurs predominately during the S phase.

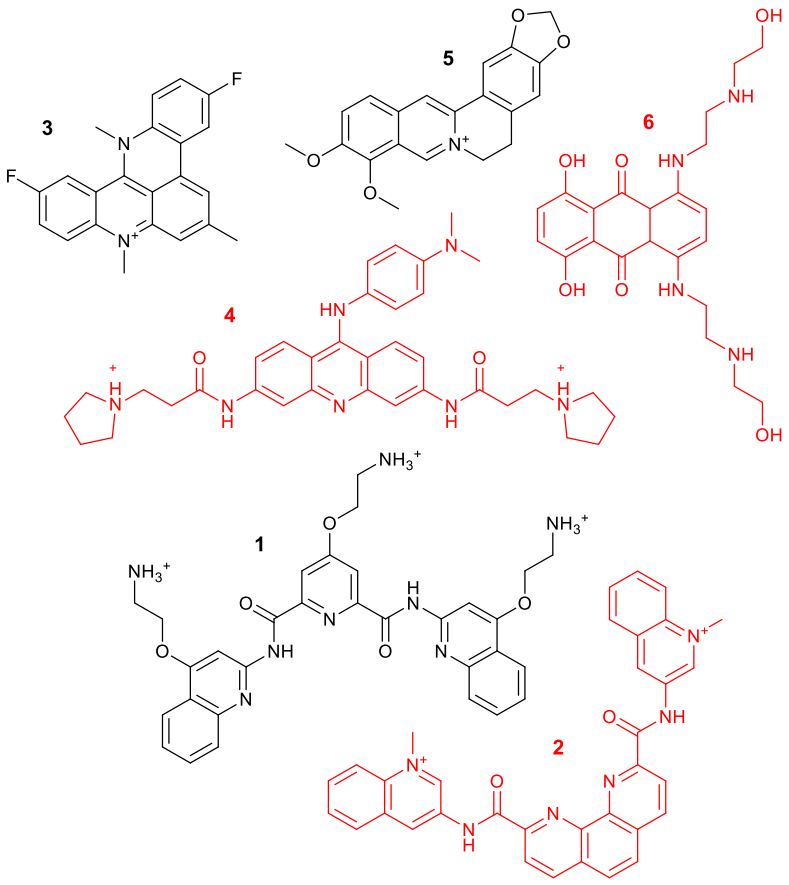


Figure 3.1-23 The G4 binding ligands tested with i-motif DNA. 1. Pyridostatin, 2. Phen-DC3, 3. RHPS4, 4. BRACO-19, 5. Berberine, 6 Mitoxantrone. In red, the compounds that destabilise the i-motif.[150]

In 2018, a paper highlighted a common challenge in nucleic acid binding ligands. A collection of common G4 binding ligands were applied to i-motif forming DNA, all interacted and some of them (BRACO-19,[151] Mitoxantrone,[152] and Phen-DC3,[153]) destabilised the i-motif.[150] (Figure 3.1-23) When the structures of these G4 binding ligands are considered, it stands to reason heteropolyaromatic cations (except neutral mitoxantrone) might interact with non-G4 DNA.

The most interesting interaction however is the destabilisation of the i-motif by some of these compounds, as these are G4 stabilising agents. Destabilising an i-motif that promotes transcription at the same time as stabilising a G4 that inhibits transcription could have a potentially synergistic effect on transcription inhibition.[154,155] It is worth considering that the alternative form for the BCL2 i-motif sequence is a hairpin structure, which could be stabilised.[156,157]

A previously established intercalating complex[158] $[Ru(phen)_2(dppz)]^{2+}$ has been shown to stabilise i-motif structures, with a preference for longer loop lengths.[159,160] Suggesting that the complex may stabilise the structure through interactions with the loops. This perhaps illustrates that i-motif targeting does not have to target the defining characteristic of stacked cytosine base-pairs. But that there are often other targetable or co-targetable features present on quadruplex nucleic acids such as stem loops.[161,162]

3.1.3 Concluding nucleic acids

Generally, agents that target nucleic acids have similar binding mechanisms. The largest portion of RNA binders use stacking interactions (34.8%), closely followed by hydrogen bonding (34.4%). The rest made up of hydrophobic interactions (17.8%) and other (13%). This differs from protein binding which is dominated by hydrophobic interactions (47.2%).[163] It is important to recognise that most small molecule ligands were designed by humans that are susceptible to bias. However, the analysis does match observations made in this review and makes the conclusion that these characteristics can be used to improve screening and scoring libraries of new ligands.

3.2 Helicates and Similar Structures

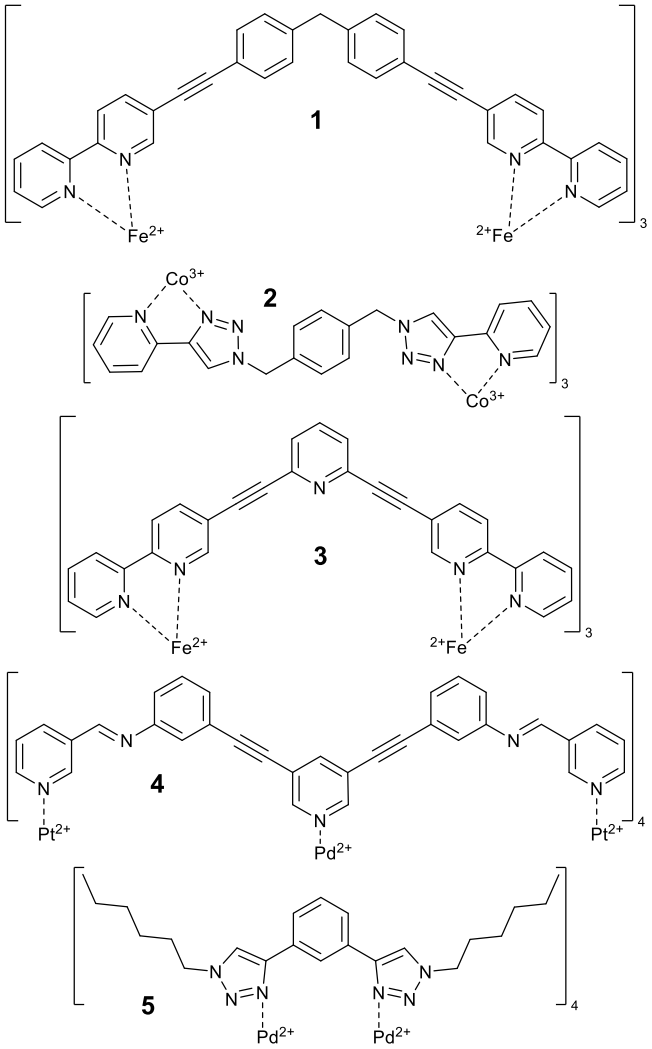


Figure 3.2-1 1-3. Various bimetallic triple helicates. 4. Heterotrimetallic cage. 5. Bimetallic quadruple helicate.

As mentioned before, helicates have promising therapeutic applicability. 2-pyridyl-1,2,3-triazole based ligands with a linking 1,4-xylyl group used in a bi-metallic triple helicate (**2**) are capable of a competitive IC_{50} versus cisplatin in cancerous (A2780, HCT116, MDA-MB-231) cells, with a low activity against non-cancerous (MRC5pd30) cells.[164,165] Bipyridine based helicates with longer 2,6-diethynylpyridine (**3**) and bis(4-ethynylphenyl)methane (**1**) bridge ligands and a central cavity, unfortunately unable to host cytotoxic guests, are capable of cytotoxicity on their own, though not to the same

extent as **2**.[166] A trimetallic cage (**4**) using similar ligands can be used to host cisplatin, two molecules of 5-fluorouracil, or two molecules of 2,6-diaminoanthraquinone.[167] A bistriazole based complex (**5**), with hexyl groups is highly cytotoxic, though poorly specific.[168] Suggesting that the mechanism of action is different to the triple helicate **2**. Later it was shown to act through membrane damage.[169] (Figure 3.2-1)

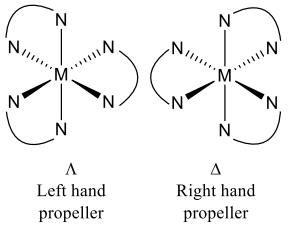


Figure 3.2-2 Bidentate ligands at octahedral centers form either Λ or Δ enantiomers.

When defining triple helicates, it is important to consider there are multiple possible enantiomers. Bidentate ligands arrange around an octahedral centre with either a right hand/clockwise (Δ) or left hand/anticlockwise (Λ) arrangement. Bimetallic helicates formed of two octahedral centres form either the ΛΛ (M enantiomer), ΔΔ (P enantiomer), or ΔΛ (meso-cate). (Figure 3.2-2)

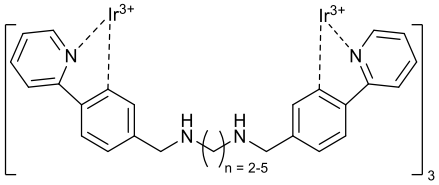


Figure 3.2-3 Bimetallic triple helicate, formed from 2-phenylpyridine based ligands, with different length linker groups.

Helicates with various ligand linking group lengths can be prepared, with intriguing length-controlled chirality. Ligands of even linker lengths were found to form both helicate enantiomers. Odd linker lengths formed the mesocate. The Ir(III) based fluorescent complexes were used to image mitochondria and displayed promising

photocytotoxicity. Where n=2-4, molecular docking using a B-DNA dodecamer predicts binding to the major groove of the helix, with a preference for the mesocate n=3. Further, these experiments show that the longer n=5 linker is able to minor groove bind, owing to the increased flexibility.[170] (Figure 3.2-3)

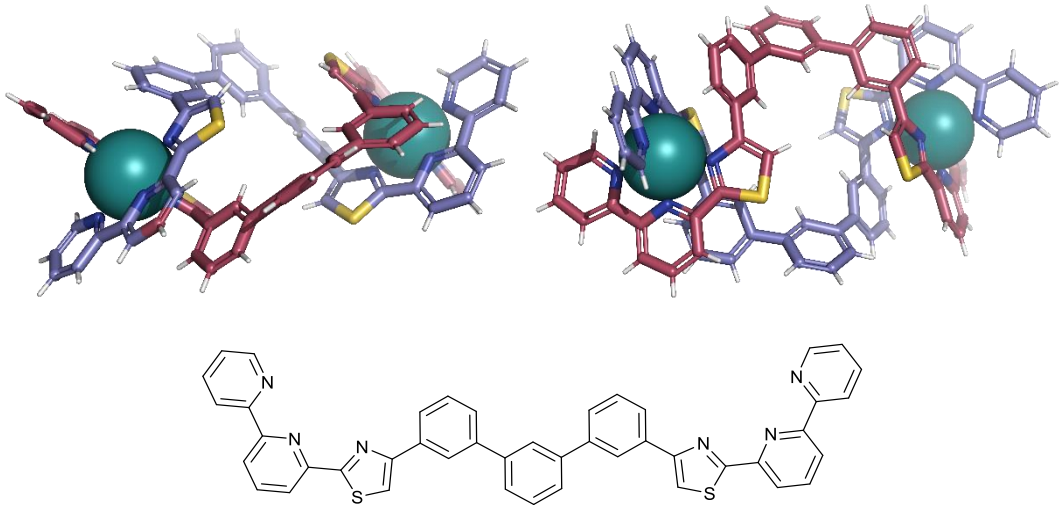


Figure 3.2-4 Structure of a complex $[Ru_2L_2]^{4+}$ with p53-selective HCT116 cytotoxicity. Top-left. Helicate form. CCDC: 1566897. Top-Right. Mesocate form. CCDC: 1566894.[171] Bottom. The ligand used.

Another nucleic acid targeting complex also shows variation between the helicate and mesocate forms. In this case significant variation in selectivity for cytotoxicity, toward the HCT116 colorectal cancer cells without p53 tumour suppressor gene, was seen for the helicate form versus the mesocate.[171]

A fascinating bipyridine based peptide triple helicate has been shown to bind with high affinity to the DNA 3WJ and is capable of binding to DNA 3WJs in cells. This was used as a method for cellular 3WJ staining showing some colocalization with a synthetic fluorescent replication foci targeting protein GFP-PCNAL2.[172]

Ag(I) and Cu(I) $[M_2L_2]$ helicates prepared using quaterpyridine based ligands showed difference in activity, with the Ag helicates showing greater cytotoxicity overall. It is

proposed that this is due to redox properties of the metal ions. The complexes are believed to be groove binding agents, though the slight increase of melting temperatures suggests it may be weak. An experiment was conducted where the complexes were incubated in a HeLa cell extract before being either extracted into chloroform or D_2O to demonstrate their stability. It would be interesting to quantify the extracted complex to identify the proportion still bound to the cellular contents.[173]

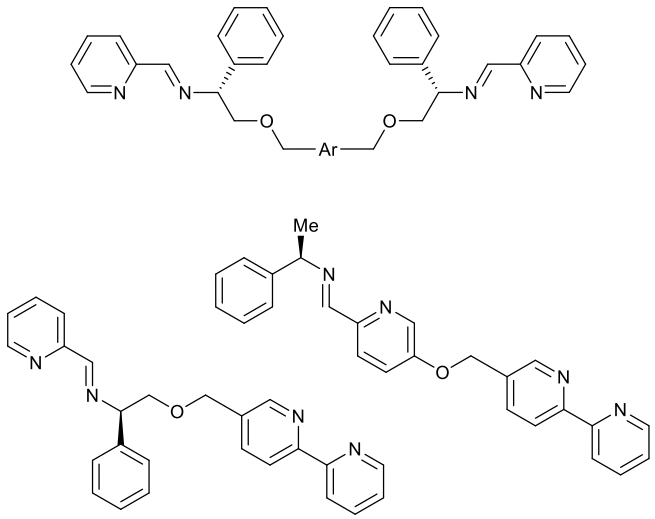


Figure 3.2-5 A collection of ligands used for the creation of flexible helicates. With the basic symmetric flexicate ligand structure at the top, and two asymmetric helicate ligands at the bottom.

Previously mentioned as a bulge binding agent, a series of helicates have been developed with chirality originating from the ligand to create a "pre-programmed" chiral helicate. These complexes exhibited DNA binding and antibiotic activity against MRSA and E. coli.[174] These flexicates have an ability to condense/aggregate DNA, accumulate in HCT116 $p53^{+/+}$ human colon carcinoma cellular nuclei, and inhibit transcription enzymes.[175] Antibiotic activity of these complexes was suggested to occur partially by interacting with DNA at the poles. A particular complex where the group Ar is a 1,3-phenyl and octahedral centres are Fe(II) shows a strong preference for G4 binding.[176]

Developing asymmetric ligands allowed for the creation of helicates with no toxicity toward MRSA or E. coli bacteria, but toxicity against MDA-MB-468 human breast adenocarcinoma cells and high toxicity against HCT116 $p53^{+/+}$ cells.[177]

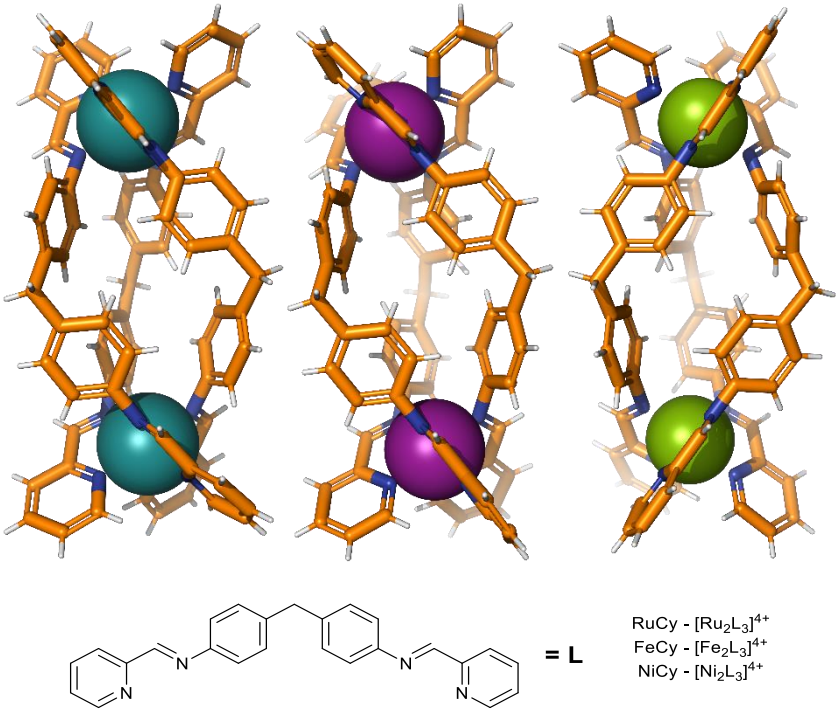


Figure 3.2-6 Top-Left. Crystal structure of RuCy M enantiomer. CCDC: 636261.[178] Top-Middle. Cystal structure of FeCy M enantiomer. CCDC: 622770.[179] Top-Right. Crystal structure of NiCy P enantiomer. CCDC: 722438.[180] Bottom. Structure of the bimetallic triple helicate used in later chapters.

An inexpensive and simple to prepare bimetallic helicate experiences differential binding of the M and P enantiomers to the major and minor groove of B-DNA, respectively.[180,181] The racemic $Fe(II)_2L_3$ complex (FeCy) is shown to have high cytotoxicity causing arrest of cells in the G_0/G_1 phase.[179] The racemic $Ru(II)_2L_3$ complex (RuCy) also shows cytotoxicity slightly greater than the FeCy.[178] The RuCy is also able to inhibit DNA polymerase, blocking PCR at 1.5 µM.[182] The P enantiomers of the FeCy and $Ni(II)_2L_3$ (NiCy) are able to selectively stabilize human telomeric DNA G4s with an

interesting switch from parallel to hybrid form G4. Also interesting is an apparent weakening of the complementary i-motif structure.[183]

3.3 Concluding

Nucleic acids present a range of structural motifs with important, varying but nuanced differences. Nucleic acid binding agents commonly interact, in order of frequency, with stacking, hydrogen bonding or hydrophobic interactions. Many nucleic acid motifs have structural similarities so binding agents often interact with various motifs, requiring fine tuning to achieve specificity. Modifications include the addition of ammonium groups for phosphate interactions, as seen with telomestatin G4 binding. Planar metal complex ligands can be extended or adjusted to reposition polar regions of G4 binders over the tetrad cations, seen with qdppz. Helicates present an interesting solution, presenting a three-dimensional surface to interact at the boundaries of duplex and quadruplex nucleic acids. Recent research has seen them applied to bulges of HIV-1 TAR and the bulges and loops of SARS CoV-2 5'UTR.

Within this thesis is presented research into the binding of supramolecular helicates to three-way junctions and bulge structures. Chapter 2 introduces a method for controlling the binding of a helicate by rotaxanating it with cucurbit[10]uril, using removable capping groups. In Chapter 3 the binding of a helicate with a HIV-1 TAR RNA bulge-stem-loop structure is explored using NMR. Chapter 4 examines the binding of a helicate to DNA three-way junctions using mostly NMR.

4 References

1 R. Altmann, Ueber nucleinsäuren, *Arch. Anat. Physiol.*, 1889, **1**, 524–536.

2 J. F. Miescher, in *Medicinisch-Chemische Untersuchungen*, ed. F. Hoppe-Seyler, Berlin, 1871, pp. 441–460.

3 M. Egli, DNA-Cation Interactions, *Chem. Biol.*, 2002, **9**, 277–286.

4 T. K. Chiu and R. E. Dickerson, 1 Å crystal structures of B-DNA reveal sequence-specific binding and groove-specific bending of DNA by magnesium and calcium, *J. Mol. Biol.*, 2000, **301**, 915–945.

5 G. Masliah, B. René, L. Zargarian, S. Fermandjian and O. Mauffret, Identification of Intrinsic Dynamics in a DNA Sequence Preferentially Cleaved by Topoisomerase II Enzyme, *J. Mol. Biol.*, 2008, **381**, 692–706.

6 P. A. Levene, L. A. Mikeska and T. Mori, on the Carbohydrate of Thymonucleic Acid, *J. Biol. Chem.*, 1930, **85**, 785–787.

7 P. A. Levene and E. S. London, the Structure of Thymonucleic Acid, *J. Biol. Chem.*, 1929, **83**, 793–802.

8 P. A. Levene and W. A. Jacobs, Über Inosinsäure, *Berichte der Dtsch. Chem. Gesellschaft*, 1909, **42**, 1198–1203.

9 P. A. Levene and W. A. Jacobs, Über die Pentose in den Nucleinsäuren, *Berichte der Dtsch. Chem. Gesellschaft*, 1909, **42**, 3247–3251.

10 J. D. Watson and F. H. C. Crick, Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid, *Nature*, 1953, **171**, 737–738.

11 R. E. Franklin and R. G. Gosling, Molecular Configuration in Sodium Thymonucleate, *Nature*, 1953, **171**, 740–741.

12 M. H. F. Wilkins, A. R. Stokes and H. R. Wilson, Molecular Structure of Nucleic Acids: Molecular Structure of Deoxypentose Nucleic Acids, *Nature*, 1953, **171**, 738–740.

13 C. A. Kim and J. M. Berg, A 2.2 Å resolution crystal structure of a designed zinc finger protein bound to DNA, *Nat. Struct. Biol.*, 1996, **3**, 940–945.

14 L. Fairall, J. W. R. Schwabe, L. Chapman, J. T. Finch and D. Rhodes, The crystal structure of a two zinc-finger peptide reveals an extension to the rules for zinc-finger/DNA recognition, *Nature*, 1993, **366**, 483–487.

15 J. C. Wang, Helical repeat of DNA in solution., *Proc. Natl. Acad. Sci.*, 1979, **76**, 200–203.

16 R. E. Franklin and R. G. Gosling, The structure of sodium thymonucleate fibres. I. The influence of water content, *Acta Crystallogr.*, 1953, **6**, 673–677.

17 F. H. Martin and I. Tinoco, DNA-RNA hybrid duplexes containing oligo(dA:rU) sequences are exceptionally unstable and may facilitate termination of transcription, *Nucleic Acids Res.*, 1980, **8**, 2295–2300.

18 F. Wang, Y. Liu, Z. Su, T. Osinski, G. A. P. de Oliveira, J. F. Conway, S. Schouten, M. Krupovic, D. Prangishvili and E. H. Egelman, A packing for A-form DNA in an icosahedral virus, *Proc. Natl. Acad. Sci.*, 2019, **116**, 22591–22597.

19 F. DiMaio, X. Yu, E. Rensen, M. Krupovic, D. Prangishvili and E. H. Egelman, A virus that infects a hyperthermophile encapsidates A-form DNA, *Science (80-.).*, 2015,

348, 914–917.
20 A. Rich, A. Nordheim and A. H. J. Wang, The Chemistry and Biology of Left-Handed Z-DNA, *Annu. Rev. Biochem.*, 1984, **53**, 791–846.
21 E. Gavathiotis, G. J. Sharman and M. S. Searle, Sequence-dependent variation in DNA minor groove width dictates orientational preference of Hoechst 33258 in A-tract recognition: Solution NMR structure of the 2:1 complex with d(CTTTTGCAAAAG)2, *Nucleic Acids Res.*, 2000, **28**, 728–735.
22 N. Spink, D. G. Brown, J. V. Skelly and S. Neidle, Sequence-dependent effects in drug-DNA interaction: The crystal structure of hoechst 33258 bound to the d(CGCAAATTTGCG)2 duplex, *Nucleic Acids Res.*, 1994, **22**, 1607–1612.
23 I. Haq, J. E. Ladbury, B. Z. Chowdhry, T. C. Jenkins and J. B. Chaires, Specific binding of Hoechst 33258 to the d(CGCAAATTTGCG)2 duplex: Calorimetric and spectroscopic studies, *J. Mol. Biol.*, 1997, **271**, 244–257.
24 M. C. Vega, I. García Sáez, J. Aymamí, R. Eritja, G. A. Van der Marel, J. H. Van Boom, A. Rich and M. Coll, Three-dimensional crystal structure of the A-tract DNA dodecamer d(CGCAAATTTGCG) complexed with the minor-groove-binding drug Hoechst 33258., *Eur. J. Biochem.*, 1994, **222**, 721–6.
25 N. Ranjan, S. Story, G. Fulcrand, F. Leng, M. Ahmad, A. King, S. Sur, W. Wang, Y. C. Tse-Dinh and D. P. Arya, Selective Inhibition of Escherichia coli RNA and DNA Topoisomerase i by Hoechst 33258 Derived Mono-and Bisbenzimidazoles, *J. Med. Chem.*, 2017, **60**, 4904–4922.
26 N. N. Patwardhan, L. R. Ganser, G. J. Kapral, C. S. Eubanks, J. Lee, B. Sathyamoorthy, H. M. Al-Hashimi and A. E. Hargrove, Amiloride as a new RNA-binding scaffold with activity against HIV-1 TAR, *Medchemcomm*, 2017, **8**, 1022–1036.
27 M. Tanada, S. Tsujita and S. Sasaki, Design of New Bidentate Ligands Constructed of Two Hoechst 33258 Units for Discrimination of the Length of Two A 3 T 3 Binding Motifs, *J. Org. Chem.*, 2006, **71**, 125–134.
28 M. Tosolini, T. Gianferrara, G. Mion, L. Dovigo, F. Mancin, C. Sissi and P. Tecilla, Interaction with plasmid DNA of Hoechst-TACN conjugates, *Supramol. Chem.*, 2020, **32**, 91–105.
29 C. Sissi, L. Dovigo, M. L. Greco, A. Ciancetta, S. Moro, J. W. Trzciński, F. Mancin, P. Rossi, G. Spalluto and P. Tecilla, Conjugates between minor groove binders and Zn(II)-tach complexes: Synthesis, characterization, and interaction with plasmid DNA, *Tetrahedron*, 2017, **73**, 3014–3024.
30 T. P. Camargo, A. Neves, R. A. Peralta, C. Chaves, E. C. P. Maia, E. H. Lizarazo-Jaimes, D. A. Gomes, T. Bortolotto, D. R. Norberto, H. Terenzi, D. L. Tierney and G. Schenk, Second-Sphere Effects in Dinuclear FeIIIZnII Hydrolase Biomimetics: Tuning Binding and Reactivity Properties, *Inorg. Chem.*, 2018, **57**, 187–203.
31 M. Pellei, F. Del Bello, M. Porchia and C. Santini, Zinc coordination complexes as anticancer agents, *Coord. Chem. Rev.*, 2021, **445**, 214088.
32 E. Boseggia, M. Gatos, L. Lucatello, F. Mancin, S. Moro, M. Palumbo, C. Sissi, P. Tecilla, U. Tonellato and G. Zagotto, Toward Efficient Zn(II)-Based Artificial Nucleases, *J. Am. Chem. Soc.*, 2004, **126**, 4543–4549.
33 H. Song, J. T. Kaiser and J. K. Barton, Crystal structure of Δ-[Ru(bpy) 2 dppz] 2+

bound to mismatched DNA reveals side-by-side metalloinsertion and intercalation, *Nat. Chem.*, 2012, **4**, 615–620.

34 C. L. Kielkopf, K. E. Erkkila, B. P. Hudson, J. K. Barton and D. C. Rees, Structure of a photoactive rhodium complex intercalated into DNA, *Nat. Struct. Biol.*, 2000, **7**, 117–121.

35 B. M. Zeglis, V. C. Pierre, J. T. Kaiser and J. K. Barton, A Bulky Rhodium Complex Bound to an Adenosine-Adenosine DNA Mismatch: General Architecture of the Metalloinsertion Binding Mode, *Biochemistry*, 2009, **48**, 4247–4253.

36 A. Nano, J. Dai, J. M. Bailis and J. K. Barton, Rhodium Complexes Targeting DNA Mismatches as a Basis for New Therapeutics in Cancers Deficient in Mismatch Repair, *Biochemistry*, 2021, **60**, 2055–2063.

37 A. Nano, J. M. Bailis, N. F. Mariano, E. D. Pham, S. D. Threatt and J. K. Barton, Cell-Selective Cytotoxicity of a Fluorescent Rhodium Metalloinsertor Conjugate Results from Irreversible DNA Damage at Base Pair Mismatches, *Biochemistry*, 2020, **59**, 717–726.

38 P. M. Keane, J. Tory, M. Towrie, I. V. Sazanovich, C. J. Cardin, S. J. Quinn, F. Hartl, J. M. Kelly and C. Long, Spectro-electrochemical Studies on [Ru(TAP) 2 (dppz)] 2+ - Insights into the Mechanism of its Photosensitized Oxidation of Oligonucleotides, *Inorg. Chem.*, 2019, **58**, 663–671.

39 P. M. Keane, K. O'Sullivan, F. E. Poynton, B. C. Poulsen, I. V. Sazanovich, M. Towrie, C. J. Cardin, X. Z. Sun, M. W. George, T. Gunnlaugsson, S. J. Quinn and J. M. Kelly, Understanding the factors controlling the photo-oxidation of natural DNA by enantiomerically pure intercalating ruthenium polypyridyl complexes through TA/TRIR studies with polydeoxynucleotides and mixed sequence oligodeoxynucleotides, *Chem. Sci.*, 2020, **11**, 8600–8609.

40 S. D. Fairbanks, C. C. Robertson, F. R. Keene, J. A. Thomas and M. P. Williamson, Structural Investigation into the Threading Intercalation of a Chiral Dinuclear Ruthenium(II) Polypyridyl Complex through a B-DNA Oligonucleotide, *J. Am. Chem. Soc.*, 2019, **141**, 4644–4652.

41 S. M. Villordo, D. E. Alvarez and A. V. Gamarnik, A balance between circular and linear forms of the dengue virus genome is crucial for viral replication, *Rna*, 2010, **16**, 2325–2335.

42 Z. Y. Liu, X. F. Li, T. Jiang, Y. Q. Deng, Q. Ye, H. Zhao, J. Y. Yu and C. F. Qin, Viral RNA switch mediates the dynamic control of flavivirus replicase recruitment by genome cyclization, *Elife*, 2016, **5**, 1–27.

43 S. Sharma and G. Varani, NMR structure of Dengue West Nile viruses stem-loop B: A key cis-acting element for flavivirus replication, *Biochem. Biophys. Res. Commun.*, 2020, **531**, 522–527.

44 I. Tinoco, O. C. Uhlenbeck and M. D. Levine, Estimation of secondary structure in ribonucleic acids, *Nature*, 1971, **230**, 362–367.

45 K. M. Weeks and D. M. Crothers, RNA recognition by Tat-derived peptides: Interaction in the major groove?, *Cell*, 1991, **66**, 577–588.

46 N. Degtyareva, D. Subramanian and J. D. Griffith, Analysis of the Binding of p53 to DNAs Containing Mismatched and Bulged Bases, *J. Biol. Chem.*, 2001, **276**, 8778–8784.

47 G. Streisinger and J. Owen, Mechanisms of spontaneous and induced frameshift mutation in bacteriophage T4, *Genetics*, 1985, **109**, 633–659.
48 W. A. Baase, D. Jose, B. C. Ponedel, P. H. Von Hippel and N. P. Johnson, DNA models of trinucleotide frameshift deletions: The formation of loops and bulges at the primer-template junction, *Nucleic Acids Res.*, 2009, **37**, 1682–1689.
49 A. Kuzminov, Half-Intercalation Stabilizes Slipped Mispairing and Explains Genome Vulnerability to Frameshift Mutagenesis by Endogenous "Molecular Bookmarks", *BioEssays*, 2019, **41**, 1–10.
50 S. Baba, K. I. Takahashi, S. Noguchi, H. Takaku, Y. Koyanagi, N. Yamamoto and G. Kawai, Solution RNA structures of the HIV-1 dimerization initiation site in the kissing-loop and extended-duplex dimers, *J. Biochem.*, 2005, **138**, 583–592.
51 E. Ennifar, J. C. Paillart, A. Bodlenner, P. Walter, J. M. Weibel, A. M. Aubertin, P. Pale, P. Dumas and R. Marquet, Targeting the dimerization initiation site of HIV-1 RNA with aminoglycosides: From crystal to cell, *Nucleic Acids Res.*, 2006, **34**, 2328–2339.
52 B. Parkash, A. Ranjan, V. Tiwari, S. K. Gupta, N. Kaur and V. Tandon, Inhibition of 5'-UTR RNA Conformational Switching in HIV-1 Using Antisense PNAs, *PLoS One*, 2012, **7**, e49310.
53 S. Bernacchi, S. Freisz, C. Maechling, B. Spiess, R. Marquet, P. Dumas and E. Ennifar, Aminoglycoside binding to the HIV-1 RNA dimerization initiation site: Thermodynamics and effect on the kissing-loop to duplex conversion, *Nucleic Acids Res.*, 2007, **35**, 7128–7139.
54 V. K. Tam, D. Kwong and Y. Tor, Fluorescent HIV-1 dimerization initiation site: Design, properties, and use for ligand discovery, *J. Am. Chem. Soc.*, 2007, **129**, 3257–3266.
55 J. Malina, P. Scott and V. Brabec, Shape-Selective Targeting on RNA Bulges by Peptidomimetic Metallohelices, *Chem. - A Eur. J.*, 2020, **26**, 8435–8442.
56 J. Malina, M. J. Hannon and V. Brabec, Recognition of DNA bulges by dinuclear iron(II) metallosupramolecular helicates, *FEBS J.*, 2014, **281**, 987–997.
57 L. Melidis, I. B. Styles and M. J. Hannon, Targeting structural features of viral genomes with a nano-sized supramolecular drug, *Chem. Sci.*, 2021, **12**, 7174–7184.
58 J. Malina, M. J. Hannon and V. Brabec, Iron(II) supramolecular helicates interfere with the HIV-1 Tat-TAR RNA interaction critical for viral replication, *Sci. Rep.*, 2016, **6**, 29674.
59 L. Cardo, I. Nawroth, P. J. Cail, J. A. McKeating and M. J. Hannon, Metallo supramolecular cylinders inhibit HIV-1 TAR-TAT complex formation and viral replication in cellulo, *Sci. Rep.*, 2018, **8**, 2–8.
60 J. Davila-Calderon, N. N. Patwardhan, L.-Y. Chiu, A. Sugarman, Z. Cai, S. R. Penutmutchu, M.-L. Li, G. Brewer, A. E. Hargrove and B. S. Tolbert, IRES-targeting small molecule inhibits enterovirus 71 replication via allosteric stabilization of a ternary complex., *Nat. Commun.*, 2020, **11**, 4775.
61 D. R. Duckett and D. M. Lilley, The three-way DNA junction is a Y-shaped molecule in which there is no helix-helix stacking., *EMBO J.*, 1990, **9**, 1659–64.
62 A. Lescoute and E. Westhof, Topology of three-way junctions in folded RNAs.,

RNA, 2006, **12**, 83–93.
63 R. R. Sinden, Origins of instability, *Nature*, 2001, **411**, 757–758.
64 E. Bonneau, N. Girard, S. Lemieux and P. Legault, The NMR structure of the II-III-VI three-way junction from the Neurospora VS ribozyme reveals a critical tertiary interaction and provides new insights into the global ribozyme structure, *Rna*, 2015, **21**, 1621–1632.
65 M. Kang, C. D. Eichhorn and J. Feigon, Structural determinants for ligand capture by a class II preQ1 riboswitch., *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, E663-71.
66 R. Micura and C. Höbartner, Fundamental studies of functional nucleic acids: aptamers, riboswitches, ribozymes and DNAzymes, *Chem. Soc. Rev.*, 2020, **49**, 7331–7353.
67 L. Ma and J. Liu, Catalytic Nucleic Acids: Biochemistry, Chemical Biology, Biosensors, and Nanotechnology, *iScience*, 2020, **23**, 100815.
68 W. G. Scott, J. T. Finch and A. Klug, The crystal structure of an AII-RNAhammerhead ribozyme: A proposed mechanism for RNA catalytic cleavage, *Cell*, 1995, **81**, 991–1002.
69 T. Lan, K. Furuya and Y. Lu, A highly selective lead sensor based on a classic lead DNAzyme, *Chem. Commun.*, 2010, **46**, 3896.
70 C. W. Chan and A. Mondragón, Crystal structure of an atypical cobalamin riboswitch reveals RNA structural adaptability as basis for promiscuous ligand binding, *Nucleic Acids Res.*, 2020, **48**, 7569–7583.
71 Z. Yang, Y. Chen, G. Li, Z. Tian, L. Zhao, X. Wu, Q. Ma, M. Liu and P. Yang, Supramolecular Recognition of Three Way Junction DNA by a Cationic Calix[3]carbazole., *Chemistry*, 2018, **24**, 6087–6093.
72 K. Duskova, P. Lejault, É. Benchimol, R. Guillot, S. Britton, A. Granzhan and D. Monchaud, DNA Junction Ligands Trigger DNA Damage and Are Synthetic Lethal with DNA Repair Inhibitors in Cancer Cells, *J. Am. Chem. Soc.*, 2020, **142**, 424–435.
73 L. Cerasino, M. J. Hannon and E. Sletten, DNA three-way junction with a dinuclear iron(II) supramolecular helicate at the center: a NMR structural study., *Inorg. Chem.*, 2007, **46**, 6245–51.
74 A. Oleksy, A. Oleksi, A. G. Blanco, R. Boer, I. Usón, J. Aymamí, A. Rodger, M. J. Hannon and M. Coll, Molecular recognition of a three-way DNA junction by a metallosupramolecular helicate., *Angew. Chem. Int. Ed. Engl.*, 2006, **45**, 1227–31.
75 J. Malina, M. J. Hannon and V. Brabec, Recognition of DNA three-way junctions by metallosupramolecular cylinders: Gel electrophoresis studies, *Chem. - A Eur. J.*, 2007, **13**, 3871–3877.
76 S. Phongtongpasuk, S. Paulus, J. Schnabl, R. K. O. Sigel, B. Spingler, M. J. Hannon and E. Freisinger, Binding of a Designed Anti-Cancer Drug to the Central Cavity of an RNA Three-Way Junction, *Angew. Chemie Int. Ed.*, 2013, **52**, 11513–11516.
77 Y. Parajó, J. Malina, I. Meistermann, G. J. Clarkson, M. Pascu, A. Rodger, M. J. Hannon and P. Lincoln, Effect of bridging ligand structure on the thermal stability and DNA binding properties of iron(ii) triple helicates, *Dalt. Trans.*, 2009, 4868–4874.
78 C. Uerpmann, J. Malina, M. Pascu, G. J. Clarkson, V. Moreno, A. Rodger, A. Grandas and M. J. Hannon, Design and DNA binding of an extended triple-

stranded metallo- supramolecular cylinder, *Chem. - A Eur. J.*, 2005, **11**, 1750–1756.
79 R. Holliday, A mechanism for gene conversion in fungi, *Genet. Res.*, 1964, **5**, 282–304.
80 H. Lin, D. Zhang, K. Zuo, C. Yuan, J. Li, M. Huang and Z. Lin, Structural basis of sequence-specific Holliday junction cleavage by MOC1, *Nat. Chem. Biol.*, 2019, **15**, 1241–1248.
81 K. Yamada, M. Ariyoshi and K. Morikawa, Three-dimensional structural views of branch migration and resolution in DNA homologous recombination, *Curr. Opin. Struct. Biol.*, 2004, **14**, 130–137.
82 M. Karymov, D. Daniel, O. F. Sankey and Y. L. Lyubchenko, Holliday junction dynamics and branch migration: Single-molecule analysis, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 8186–8191.
83 A. Chakraborty, P. K. Mandal and N. Gautham, Structure of d(CCGGGACCGG)4 as a four-way junction at 1.6 Å resolution: New insights into solvent interactions, *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.*, 2012, **68**, 1169–1174.
84 C. Biertümpfel, W. Yang and D. Suck, Crystal structure of T4 endonuclease VII resolving a Holliday junction, *Nature*, 2007, **449**, 616–620.
85 M. E. Bianchi, M. Beltrame and G. Paonessa, Specific recognition of cruciform DNA by nuclear protein HMG1, *Science (80-.).*, 1989, **243**, 1056–1059.
86 M. Webb and J. O. Thomas, Structure-specific binding of the two tandem HMG boxes of HMG1 to four-way junction DNA is mediated by the A domain, *J. Mol. Biol.*, 1999, **294**, 373–387.
87 C. S. Chow, C. M. Barnes and S. J. Lippard, A Single HMG Domain in High-Mobility Group 1 Protein Binds to DNAs as Small as 20 Base Pairs Containing the Major Cisplatin Adduce, *Biochemistry*, 1995, **34**, 2956–2964.
88 K. Stott, G. S. F. Tang, K. B. Lee and J. O. Thomas, Structure of a Complex of Tandem HMG Boxes and DNA, *J. Mol. Biol.*, 2006, **360**, 90–104.
89 K. D. Grasser, S. H. Teo, K. B. Lee, R. W. Broadhurst, C. Rees, C. H. Hardman and J. O. Thomas, DNA-binding properties of the tandem HMG boxes of high-mobility-group protein 1 (HMG1), *Eur. J. Biochem.*, 1998, **253**, 787–795.
90 A. L. Brogden, N. H. Hopcroft, M. Searcey and C. J. Cardin, Ligand Bridging of the DNA Holliday Junction: Molecular Recognition of a Stacked-X Four-Way Junction by a Small Molecule, *Angew. Chemie Int. Ed.*, 2007, **46**, 3850–3854.
91 J. Zhu, C. J. E. Haynes, M. Kieffer, J. L. Greenfield, R. D. Greenhalgh, J. R. Nitschke and U. F. Keyser, FeII4L4 Tetrahedron Binds to Nonpaired DNA Bases., *J. Am. Chem. Soc.*, 2019, **141**, 11358–11362.
92 E. Butovskaya, B. Heddi, B. Bakalar, S. N. Richter and A. T. Phan, Major G-Quadruplex Form of HIV-1 LTR Reveals a (3 + 1) Folding Topology Containing a Stem-Loop, *J. Am. Chem. Soc.*, 2018, **140**, 13654–13662.
93 M. L. Bochman, K. Paeschke and V. A. Zakian, DNA secondary structures: Stability and function of G-quadruplex structures, *Nat. Rev. Genet.*, 2012, **13**, 770–780.
94 W. I. Sundquist and A. Klug, Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops, *Nature*, 1989, **342**, 825–829.
95 S. Kumari, A. Bugaut, J. L. Huppert and S. Balasubramanian, An RNA G-quadruplex

in the 5′ UTR of the NRAS proto-oncogene modulates translation, *Nat. Chem. Biol.*, 2007, **3**, 218–221.
96 D. J. Patel, A. T. Phan and V. Kuryavyi, Human telomere, oncogenic promoter and 5′-UTR G-quadruplexes: Diverse higher order DNA and RNA targets for cancer therapeutics, *Nucleic Acids Res.*, 2007, **35**, 7429–7455.
97 A. K. Mukherjee, S. Sharma and S. Chowdhury, Non-duplex G-Quadruplex Structures Emerge as Mediators of Epigenetic Modifications, *Trends Genet.*, 2019, **35**, 129–144.
98 D. Rhodes and H. J. Lipps, Survey and summary G-quadruplexes and their regulatory roles in biology, *Nucleic Acids Res.*, 2015, **43**, 8627–8637.
99 R. Linke, M. Limmer, S. A. Juranek, A. Heine and K. Paeschke, The relevance of g-quadruplexes for dna repair, *Int. J. Mol. Sci.*, 2021, **22**, 1–21.
100 R. Perrone, E. Butovskaya, D. Daelemans, G. Palù, C. Pannecouque and S. N. Richter, Anti-HIV-1 activity of the G-quadruplex ligand BRACO-19, *J. Antimicrob. Chemother.*, 2014, **69**, 3248–3258.
101 E. Butovskaya, P. Soldà, M. Scalabrin, M. Nadai and S. N. Richter, HIV-1 Nucleocapsid Protein Unfolds Stable RNA G-Quadruplexes in the Viral Genome and Is Inhibited by G-Quadruplex Ligands, *ACS Infect. Dis.*, 2019, 2127–2135.
102 W. I. Sundquist and S. Heaphy, Evidence for interstrand quadruplex formation in the dimerization of human immunodeficiency virus 1 genomic RNA, *Proc. Natl. Acad. Sci. U. S. A.*, 1993, **90**, 3393–3397.
103 A. Bugaut and S. Balasubramanian, 5′-UTR RNA G-quadruplexes: Translation regulation and targeting, *Nucleic Acids Res.*, 2012, **40**, 4727–4741.
104 J. T. Grün and H. Schwalbe, Folding dynamics of polymorphic G-quadruplex structures, *Biopolymers*, 2022, **113**, 1–15.
105 L. Díaz-Casado, I. Serrano-Chacón, L. Montalvillo-Jiménez, F. Corzana, A. Bastida, A. G. Santana, C. González and J. L. Asensio, De Novo Design of Selective Quadruplex–Duplex Junction Ligands and Structural Characterisation of Their Binding Mode: Targeting the G4 Hot-Spot, *Chem. - A Eur. J.*, 2020, 6204–6212.
106 L. Y. Liu, W. Liu, K. N. Wang, B. C. Zhu, X. Y. Xia, L. N. Ji and Z. W. Mao, Quantitative Detection of G-Quadruplex DNA in Live Cells Based on Photon Counts and Complex Structure Discrimination, *Angew. Chemie - Int. Ed.*, 2020, **59**, 9719–9726.
107 Y. Zhang, K. El Omari, R. Duman, S. Liu, S. Haider, A. Wagner, G. N. Parkinson and D. Wei, Native de novo structural determinations of non-canonical nucleic acid motifs by X-ray crystallography at long wavelengths, *Nucleic Acids Res.*, 2020, **48**, 9886–9898.
108 N. H. Campbell, G. N. Parkinson, A. P. Reszka and S. Neidle, Structural basis of DNA quadruplex recognition by an acridine drug, *J. Am. Chem. Soc.*, 2008, **130**, 6722–6724.
109 D. Sun, B. Thompson, B. E. Cathers, M. Salazar, S. M. Kerwin, J. O. Trent, T. C. Jenkins, S. Neidle and L. H. Hurley, Inhibition of human telomerase by a G-Quadruplex-Interactive compound, *J. Med. Chem.*, 1997, **40**, 2113–2116.
110 N. Ranjan, E. Davis, L. Xue and D. P. Arya, Dual recognition of the human telomeric G-quadruplex by a neomycin–anthraquinone conjugate, *Chem. Commun.*, 2013, **49**, 5796–5798.

111 T. Rundstadler, E. Mothes, S. Amrane, J.-L. Stigliani, P. Verhaeghe and G. Pratviel, Gold(III) porphyrins: Synthesis and interaction with G-quadruplex DNA., *J. Inorg. Biochem.*, 2021, **223**, 111551.
112 D. Mathew and S. Sujatha, Interactions of porphyrins with DNA: A review focusing recent advances in chemical modifications on porphyrins as artificial nucleases, *J. Inorg. Biochem.*, 2021, **219**, 111434.
113 C. Romera, O. Bombarde, R. Bonnet, D. Gomez, P. Dumy, P. Calsou, J. F. Gwan, J. H. Lin, E. Defrancq and G. Pratviel, Improvement of porphyrins for G-quadruplex DNA targeting, *Biochimie*, 2011, **93**, 1310–1317.
114 L. N. Zhang, R. Zhang, Y. X. Cui, K. K. Liu, D. M. Kong, X. Z. Li and L. N. Zhu, Highly specific G-quadruplex recognition covering physiological pH range by a new water-soluble cationic porphyrin with low self-aggregation tendency, *Dye. Pigment.*, 2017, **145**, 404–417.
115 M. J. Morris, K. L. Wingate, J. Silwal, T. C. Leeper and S. Basu, The porphyrin TmPyP4 unfolds the extremely stable G-quadruplex in MT3-MMP mRNA and alleviates its repressive effect to enhance translation in eukaryotic cells, *Nucleic Acids Res.*, 2012, **40**, 4137–4145.
116 H. J. Sullivan, C. Readmond, C. Radicella, V. Persad, T. J. Fasano and C. Wu, Binding of Telomestatin, TMPyP4, BSU6037, and BRACO19 to a Telomeric G-Quadruplex-Duplex Hybrid Probed by All-Atom Molecular Dynamics Simulations with Explicit Solvent, *ACS Omega*, 2018, **3**, 14788–14806.
117 W. J. Chung, B. Heddi, M. Tera, K. Iida, K. Nagasawa and A. T. Phan, Solution structure of an intramolecular (3 + 1) human telomeric G-quadruplex bound to a telomestatin derivative, *J. Am. Chem. Soc.*, 2013, **135**, 13495–13501.
118 K. Shin-ya, K. Wierzba, K. Matsuo, T. Ohtani, Y. Yamada, K. Furihata, Y. Hayakawa and H. Seto, Telomestatin, a novel telomerase inhibitor from Streptomyces anulatus [17], *J. Am. Chem. Soc.*, 2001, **123**, 1262–1263.
119 M. Y. Kim, H. Vankayalapati, K. Shin-Ya, K. Wierzba and L. H. Hurley, Telomestatin, a potent telomerase inhibitor that interacts quite specifically with the human telomeric intramolecular G-quadruplex, *J. Am. Chem. Soc.*, 2002, **124**, 2098–2099.
120 O. Domarco, D. Lötsch, J. Schreiber, C. Dinhof, S. Van Schoonhoven, M. D. García, C. Peinador, B. K. Keppler, W. Berger and A. Terenzi, Self-assembled Pt2L2 boxes strongly bind G-quadruplex DNA and influence gene expression in cancer cells, *Dalt. Trans.*, 2017, **46**, 329–332.
121 J. Zell, K. Duskova, L. Chouh, M. Bossaert, N. Chéron, A. Granzhan, S. Britton and D. Monchaud, Dual targeting of higher-order DNA structures by azacryptands induces DNA junction-mediated DNA damage in cancer cells., *Nucleic Acids Res.*, 2021, **49**, 10275–10288.
122 R. Rodriguez, G. D. Pantoş, D. P. N. Gonçalves, J. K. M. Sanders and S. Balasubramanian, Ligand-driven G-quadruplex conformational switching by using an unusual mode of interaction, *Angew. Chemie - Int. Ed.*, 2007, **46**, 5405–5407.
123 N. Kosiol, S. Juranek, P. Brossart, A. Heine and K. Paeschke, G-quadruplexes: a promising target for cancer therapy, *Mol. Cancer*, 2021, **20**, 1–18.
124 D. Monchaud and M. P. Teulade-Fichou, A hitchhiker's guide to G-quadruplex ligands, *Org. Biomol. Chem.*, 2008, **6**, 627–636.

125 S. Neidle, Quadruplex Nucleic Acids as Novel Therapeutic Targets, *J. Med. Chem.*, 2016, **59**, 5987–6011.
126 K. T. McQuaid, S. Takahashi, L. Baumgaertner, D. J. Cardin, N. G. Paterson, J. P. Hall, N. Sugimoto and C. J. Cardin, Ruthenium Polypyridyl Complex Bound to a Unimolecular Chair-Form G-Quadruplex, *J. Am. Chem. Soc.*, 2022, **144**, 5956–5964.
127 K. McQuaid, H. Abell, S. P. Gurung, D. R. Allan, G. Winter, T. Sorensen, D. J. Cardin, J. A. Brazier, C. J. Cardin and J. P. Hall, Structural Studies Reveal Enantiospecific Recognition of a DNA G-Quadruplex by a Ruthenium Polypyridyl Complex, *Angew. Chemie - Int. Ed.*, 2019, **58**, 9881–9885.
128 S. J. Devereux, F. E. Poynton, F. R. Baptista, T. Gunnlaugsson, C. J. Cardin, I. V. Sazanovich, M. Towrie, J. M. Kelly and S. J. Quinn, Caught in the Loop: Binding of the [Ru(phen)2(dppz)]2+ Light-Switch Compound to Quadruplex DNA in Solution Informed by Time-Resolved Infrared Spectroscopy, *Chem. - A Eur. J.*, 2020, **26**, 17103–17109.
129 A. Kotar, B. Wang, A. Shivalingam, J. Gonzalez-Garcia, R. Vilar and J. Plavec, NMR Structure of a Triangulenium-Based Long-Lived Fluorescence Probe Bound to a G-Quadruplex, *Angew. Chemie - Int. Ed.*, 2016, **55**, 12696–12699.
130 N. H. Campbell, N. H. A. Karim, G. N. Parkinson, M. Gunaratnam, V. Petrucci, A. K. Todd, R. Vilar and S. Neidle, Molecular basis of structure-activity relationships between salphen metal complexes and human telomeric DNA quadruplexes, *J. Med. Chem.*, 2012, **55**, 209–222.
131 A. Shivalingam, A. Vyšniauskas, T. Albrecht, A. J. P. White, M. K. Kuimova and R. Vilar, Trianguleniums as Optical Probes for G-Quadruplexes: APhotophysical, Electrochemical, and Computational Study, *Chem. - A Eur. J.*, 2016, **22**, 4129–4139.
132 B. W. Lewis, N. Bisballe, M. Santella, P. A. Summers, J. B. Vannier, M. K. Kuimova, B. W. Laursen and R. Vilar, Assessing The Key Photophysical Properties of Triangulenium Dyes for DNA Binding by Alteration of the Fluorescent Core, *Chem. - A Eur. J.*, 2021, **27**, 2523–2536.
133 P. A. Summers, B. W. Lewis, J. Gonzalez-Garcia, R. M. Porreca, A. H. M. Lim, P. Cadinu, N. Martin-Pintado, D. J. Mann, J. B. Edel, J. B. Vannier, M. K. Kuimova and R. Vilar, Visualising G-quadruplex DNA dynamics in live cells by fluorescence lifetime imaging microscopy, *Nat. Commun.*, 2021, **12**, 1–11.
134 C. L. Ruehl, A. H. M. Lim, T. Kench, D. J. Mann and R. Vilar, An Octahedral Cobalt(III) Complex with Axial NH3 Ligands that Templates and Selectively Stabilises G-quadruplex DNA, *Chem. - A Eur. J.*, 2019, **25**, 9691–9700.
135 A. Łęczkowska, J. Gonzalez-Garcia, C. Perez-Arnaiz, B. Garcia, A. J. P. White and R. Vilar, Binding Studies of Metal–Salphen and Metal–Bipyridine Complexes towards G-Quadruplex DNA, *Chem. - A Eur. J.*, 2018, **24**, 11785–11794.
136 S. Bandeira, J. Gonzalez-Garcia, E. Pensa, T. Albrecht and R. Vilar, A Redox-Activated G-Quadruplex DNA Binder Based on a Platinum(IV)–Salphen Complex, *Angew. Chemie - Int. Ed.*, 2018, **57**, 310–313.
137 P. A. Summers, A. P. Thomas, T. Kench, J. Vannier, M. K. Kuimova and R. Vilar, Cationic helicenes as selective G4 DNA binders and optical probes for cellular

imaging, *Chem. Sci.*, 2021, **12**, 14624–14634.
138 C. Zhao, H. Song, P. Scott, A. Zhao, H. Tateishi-Karimata, N. Sugimoto, J. Ren and X. Qu, Mirror-Image Dependence: Targeting Enantiomeric G-Quadruplex DNA Using Triplex Metallohelices, *Angew. Chemie - Int. Ed.*, 2018, **57**, 15723–15727.
139 A. Zhao, S. E. Howson, C. Zhao, J. Ren, P. Scott, W. Chunyu, X. Qu, C. Wang, X. Qu, W. Chunyu, X. Qu, C. Wang, X. Qu, W. Chunyu and X. Qu, Chiral metallohelices enantioselectively target hybrid human telomeric G-quadruplex DNA, *Nucleic Acids Res.*, 2017, **45**, 5026–5035.
140 H. Qin, C. Zhao, Y. Sun, J. Ren and X. Qu, Metallo-supramolecular Complexes Enantioselectively Eradicate Cancer Stem Cells in Vivo, *J. Am. Chem. Soc.*, 2017, **139**, 16201–16209.
141 I. Serrano-Chacón, B. Mir, N. Escaja and C. González, Structure of i-Motif/Duplex Junctions at Neutral pH, *J. Am. Chem. Soc.*, 2021, **143**, 12919–12923.
142 H. A. Assi, M. Garavís, C. González and M. J. Damha, I-motif DNA: Structural features and significance to cell biology, *Nucleic Acids Res.*, 2018, **46**, 8038–8056.
143 K. Gehring, J.-L. Leroy and M. Guéron, A tetrameric DNA structure with protonated cytosine-cytosine base pairs, *Nature*, 1993, **363**, 561–565.
144 S. Dzatko, M. Krafcikova, R. Hänsel-Hertsch, T. Fessl, R. Fiala, T. Loja, D. Krafcik, J. L. Mergny, S. Foldynova-Trantirkova and L. Trantirek, Evaluation of the Stability of DNA i-Motifs in the Nuclei of Living Mammalian Cells, *Angew. Chemie - Int. Ed.*, 2018, **57**, 2165–2169.
145 S. Takahashi, J. A. Brazier and N. Sugimoto, Topological impact of noncanonical DNA structures on Klenow fragment of DNA polymerase, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, 9605–9610.
146 C. E. Kaiser, N. A. Van Ert, P. Agrawal, R. Chawla, D. Yang and L. H. Hurley, Insight into the Complexity of the i-Motif and G-Quadruplex DNA Structures Formed in the KRAS Promoter and Subsequent Drug-Induced Gene Repression, *J. Am. Chem. Soc.*, 2017, **139**, 8522–8536.
147 H. J. Kang, S. Kendrick, S. M. Hecht and L. H. Hurley, The transcriptional complex between the BCL2 i-motif and hnRNP LL is a molecular switch for control of gene expression that can be modulated by small molecules, *J. Am. Chem. Soc.*, 2014, **136**, 4172–4185.
148 G. Biffi, D. Tannahill, J. McCafferty and S. Balasubramanian, Quantitative visualization of DNA G-quadruplex structures in human cells, *Nat. Chem.*, 2013, **5**, 182–186.
149 M. Zeraati, D. B. Langley, P. Schofield, A. L. Moye, R. Rouet, W. E. Hughes, T. M. Bryan, M. E. Dinger and D. Christ, I-motif DNA structures are formed in the nuclei of human cells, *Nat. Chem.*, 2018, **10**, 631–637.
150 A. Pagano, N. Iaccarino, M. A. S. Abdelhamid, D. Brancaccio, E. U. Garzarella, A. Di Porzio, E. Novellino, Z. A. E. Waller, B. Pagano, J. Amato and A. Randazzo, Common G-quadruplex binding agents found to interact with i-motif-forming DNA: Unexpected multi-target-directed compounds, *Front. Chem.*, 2018, **6**, 1–13.
151 S. M. Gowan, J. R. Harrison, L. Patterson, M. Valenti, M. A. Read, S. Neidle and L. R. Kelland, A G-quadruplex-interactive potent small-molecule inhibitor of telomerase exhibiting in vitro and in vivo antitumor activity, *Mol. Pharmacol.*,

2002, **61**, 1154–1162.
152 H. S. Huang, I. B. Chen, K. F. Huang, W. C. Lu, F. Y. Shieh, Y. Y. Huang, F. C. Huang and J. J. Lin, Synthesis and human telomerase inhibition of a series of regioisomeric disubstituted amidoanthraquinones, *Chem. Pharm. Bull.*, 2007, **55**, 284–292.
153 A. De Cian, E. DeLemos, J. L. Mergny, M. P. Teulade-Fichou and D. Monchaud, Highly efficient G-quadruplex recognition by bisquinolinium compounds, *J. Am. Chem. Soc.*, 2007, **129**, 1856–1857.
154 M. I. Onyshchenko, T. I. Gaynutdinov, E. A. Englund, D. H. Appella, R. D. Neumann and I. G. Panyutin, Stabilization of G-quadruplex in the BCL2 promoter region in double-stranded DNA by invading short PNAs, *Nucleic Acids Res.*, 2009, **37**, 7570–7580.
155 M. Debnath, S. Ghosh, A. Chauhan, R. Paul, K. Bhattacharyya and J. Dash, Preferential targeting of i-motifs and G-quadruplexes by small molecules, *Chem. Sci.*, 2017, **8**, 7448–7456.
156 S. Kendrick, H.-J. J. Kang, M. P. Alam, M. M. Madathil, P. Agrawal, V. Gokhale, D. Yang, S. M. Hecht and L. H. Hurley, The dynamic character of the BCL2 promoter i-motif provides a mechanism for modulation of gene expression by compounds that bind selectively to the alternative DNA hairpin structure, *J. Am. Chem. Soc.*, 2014, **136**, 4161–4171.
157 H.-J. J. Kang, S. Kendrick, S. M. Hecht and L. H. Hurley, The transcriptional complex between the BCL2 i-motif and hnRNP LL is a molecular switch for control of gene expression that can be modulated by small molecules, *J. Am. Chem. Soc.*, 2014, **136**, 4172–4185.
158 H. Niyazi, J. P. Hall, K. O'Sullivan, G. Winter, T. Sorensen, J. M. Kelly and C. J. Cardin, Crystal structures of λ-[Ru(phen) 2 dppz] 2+ with oligonucleotides containing TA/TA and AT/AT steps show two intercalation modes, *Nat. Chem.*, 2012, **4**, 621–628.
159 B. J. Pages, S. P. Gurung, K. McQuaid, J. P. Hall, C. J. Cardin and J. A. Brazier, Stabilization of Long-Looped i-Motif DNA by Polypyridyl Ruthenium Complexes, *Front. Chem.*, 2019, **7**, 1–7.
160 F. R. Baptista, S. J. Devereux, S. P. Gurung, J. P. Hall, I. V. Sazanovich, M. Towrie, C. J. Cardin, J. A. Brazier, J. M. Kelly and S. J. Quinn, The influence of loops on the binding of the [Ru(phen)2dppz]2+light-switch compound to i-motif DNA structures revealed by time-resolved spectroscopy, *Chem. Commun.*, 2020, **56**, 9703–9706.
161 E. Ruggiero, S. Lago, P. Šket, M. Nadai, I. Frasson, J. Plavec and S. N. Richter, A dynamic i-motif with a duplex stem-loop in the long terminal repeat promoter of the HIV-1 proviral genome modulates viral transcription, *Nucleic Acids Res.*, 2019, **47**, 11057–11068.
162 C. Harpster, E. Boyle, K. Musier-Forsyth and B. Kankia, HIV-1 genomic RNA U3 region forms a stable quadruplex-hairpin structure, *Biophys. Chem.*, 2021, **272**, 106567.
163 G. Padroni, N. N. Patwardhan, M. Schapira and A. E. Hargrove, Systematic analysis of the interactions driving small molecule-RNA recognition, *RSC Med. Chem.*,

2020, **11**, 802–813.
164 H. Crlikova, J. Malina, V. Novohradsky, H. Kostrhunova, R. A. S. Vasdev, J. D. Crowley, J. Kasparkova and V. Brabec, Antiproliferative Activity and Associated DNA Interactions of [Co2L3]6+Cylinders Derived from Bis(bidentate) 2-Pyridyl-1,2,3-triazole Ligands, *Organometallics*, 2020, **39**, 1448–1455.
165 S. V. Kumar, W. K. C. Lo, H. J. L. Brooks and J. D. Crowley, Synthesis, structure, stability and antimicrobial activity of a ruthenium(II) helicate derived from a bis-bidentate 'click' pyridyl-1,2,3-triazole ligand, *Inorganica Chim. Acta*, 2015, **425**, 1–6.
166 L. S. Lisboa, M. Riisom, R. A. S. Vasdev, S. M. F. Jamieson, L. J. Wright, C. G. Hartinger and J. D. Crowley, Cavity-Containing [Fe2L3]4+ Helicates: An Examination of Host-Guest Chemistry and Cytotoxicity, *Front. Chem.*, 2021, **9**, 1–9.
167 L. S. Lisboa, D. Preston, C. J. McAdam, L. J. Wright, C. G. Hartinger and J. D. Crowley, Heterotrimetallic Double Cavity Cages: Syntheses and Selective Guest Binding, *Angew. Chemie - Int. Ed.*, 2022, **61**, 1–6.
168 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, Biologically active [Pd2L4]4+ quadruply-stranded helicates: stability and cytotoxicity, *Dalt. Trans.*, 2015, **44**, 11129–11136.
169 S. M. Mcneill, N. M. Giles, D. Preston, P. P. Jones, J. D. Crowley and G. I. Giles, Quadruply Stranded Metallo-Supramolecular Helicate [Pd2(hextrz)4]4+Acts as a Molecular Mimic of Cytolytic Peptides, *Chem. Res. Toxicol.*, 2020, **33**, 1822–1834.
170 X. Li, J. Wu, L. Wang, C. He, L. Chen, Y. Jiao and C. Duan, Mitochondrial-DNA-Targeted Ir III -Containing Metallohelices with Tunable Photodynamic Therapy Efficacy in Cancer Cells, *Angew. Chemie Int. Ed.*, 2020, **59**, 6420–6427.
171 S. J. Allison, D. Cooke, F. S. Davidson, P. I. P. Elliott, R. A. Faulkner, H. B. S. Griffiths, O. J. Harper, O. Hussain, P. J. Owen-Lynch, R. M. Phillips, C. R. Rice, S. L. Shepherd and R. T. Wheelhouse, Ruthenium-Containing Linear Helicates and Mesocates with Tuneable p53-Selective Cytotoxicity in Colorectal Cancer Cells, *Angew. Chemie - Int. Ed.*, 2018, **57**, 9799–9804.
172 J. Gómez-González, Y. Pérez, G. Sciortino, L. Roldan-Martín, J. Martínez-Costas, J. Maréchal, I. Alfonso, M. Vázquez López and M. E. Vázquez, Dynamic Stereoselection of Peptide Helicates and Their Selective Labeling of DNA Replication Foci in Cells**, *Angew. Chemie Int. Ed.*, 2021, **60**, 8859–8866.
173 A. Adamski, M. A. Fik, M. Kubicki, Z. Hnatejko, D. Gurda, A. Fedoruk-Wyszomirska, E. Wyszko, D. Kruszka, Z. Dutkiewicz and V. Patroniak, Full characterization and cytotoxic activity of new silver(i) and copper(i) helicates with quaterpyridine, *New J. Chem.*, 2016, **40**, 7943–7957.
174 S. E. Howson, A. Bolhuis, V. Brabec, G. J. Clarkson, J. Malina, A. Rodger and P. Scott, Optically pure, water-stable metallo-helical 'flexicate' assemblies with antibiotic activity., *Nat. Chem.*, 2011, **4**, 31–6.
175 O. Hrabina, J. Malina, H. Kostrhunova, V. Novohradsky, J. Pracharova, N. Rogers, D. H. Simpson, P. Scott and V. Brabec, Optically Pure Metallohelices That Accumulate in Cell Nuclei, Condense/Aggregate DNA, and Inhibit Activities of DNA

Processing Enzymes., *Inorg. Chem.*, 2020, **59**, 3304–3311.
176 D. H. Simpson, A. Hapeshi, N. J. Rogers, V. Brabec, G. J. Clarkson, D. J. Fox, O. Hrabina, G. L. Kay, A. K. King, J. Malina, A. D. Millard, J. Moat, D. I. Roper, H. Song, N. R. Waterfield and P. Scott, Metallohelices that kill Gram-negative pathogens using intracellular antimicrobial peptide pathways, *Chem. Sci.*, 2019, **10**, 9708–9720.
177 A. D. Faulkner, R. A. Kaner, Q. M. A. Abdallah, G. Clarkson, D. J. Fox, P. Gurnani, S. E. Howson, R. M. Phillips, D. I. Roper, D. H. Simpson and P. Scott, Asymmetric triplex metallohelices with high and selective activity against cancer cells., *Nat. Chem.*, 2014, **6**, 797–803.
178 G. I. Pascu, A. C. G. Hotze, C. Sanchez-Cano, B. M. Kariuki and M. J. Hannon, Dinuclear ruthenium(II) triple-stranded helicates: Luminescent supramolecular cylinders that bind and coil DNA and exhibit activity against cancer cell lines, *Angew. Chemie - Int. Ed.*, 2007, **46**, 4374–4378.
179 A. C. G. Hotze, N. J. Hodges, R. E. Hayden, C. Sanchez-Cano, C. Paines, N. Male, M. K. Tse, C. M. Bunce, J. K. Chipman and M. J. Hannon, Supramolecular Iron Cylinder with Unprecedented DNA Binding Is a Potent Cytostatic and Apoptotic Agent without Exhibiting Genotoxicity, *Chem. Biol.*, 2008, **15**, 1258–1267.
180 M. J. Hannon, C. L. Painting, A. Jackson, J. Hamblin and W. Errington, An inexpensive approach to supramolecular architecture, *Chem. Commun.*, 1997, **128**, 1807–1808.
181 I. Meistermann, V. Moreno, M. J. Prieto, E. Moldrheim, E. Sletten, S. Khalid, P. Mark Rodger, J. C. Peberdy, C. J. Isaac, A. Rodger, M. J. Hannon, P. M. Rodger, J. C. Peberdy, C. J. Isaac, A. Rodger and M. J. Hannon, Intramolecular DNA coiling mediated by metallo-supramolecular cylinders: Differential binding of P and M helical enantiomers, *Proc. Natl. Acad. Sci.*, 2002, **99**, 5069–5074.
182 C. Ducani, A. Leczkowska, N. J. Hodges and M. J. Hannon, Noncovalent DNA-binding metallo-supramolecular cylinders prevent DNA transactions in vitro, *Angew. Chemie - Int. Ed.*, 2010, **49**, 8942–8945.
183 H. Yu, X. Wang, M. Fu, J. Ren and X. Qu, Chiral metallo-supramolecular complexes selectively recognize human telomeric G-quadruplex DNA, *Nucleic Acids Res.*, 2008, **36**, 5695–5703.

CHAPTER 2 – ROTAXANATING CYLINDERS

1 Acknowledgements and related work

This work was done as part of a group effort to explore methods of rotaxanating supramolecular cylinders. It began with the work of Dr. Lucia Cardo on pseudorotaxanes and Catherine Hooper on capping rotaxanes with methyl pyridine resulting in the publication, Rotaxanating Metallo-supramolecular Nano-cylinder Helicates to Switch DNA Junction Binding, *J. Am. Chem. Soc.*, 2020, **142**, 20651–20660, which a small part of this work contributed.[1]

The bulk of the chapter was conducted for part of a future joint first authorship publication with Dr. Aditya Garai, on dinitrobenzene capped rotaxanes and their controllable binding to DNA three-way junctions. This will include biological experiments conducted by James Craig and Dr. Nik Hodges that are not included herein. Here, presented as much as possible, is work conducted alone. With exceptions from Dr. Aditya Garai where the experiments were too integral to the project and too entangled with others. Thanks to Dr Aditya Garai, who conducted the synthesis of the 2-((2,4-dinitrophenyl)thio)ethan-1-ol, L^{DNB} with my assistance, DNB_cylinder (ligand method), and assisted with the DNB aldehyde synthesis. He was essential for his positive attitude, experience and to bounce ideas back and forth. Together a mass-spec experimental protocol was developed on a low-resolution instrument to be used and later, using this with the assistance of Dr. Chris Williams, Dr. Aditya conducted a series of high-resolution

mass spectrometry decapping repeats. These serve to better illustrate the effects seen in UV-Vis and NMR and, as the higher resolution versions, are the ones chosen for this thesis.

Finally, I wish to thank Hugo Williams for the acquisition of UV-Vis DNP stability measurements, as well as Dr. Sarah Griffin and Prof. Neil Champness for crystallography.

2 Contents

CHAPTER 2 – ROTAXANATING CYLINDERS 59

1 ACKNOWLEDGEMENTS AND RELATED WORK 59

2 CONTENTS 61

3 GLOSSARY 63

4 INTRODUCTION 64

4.1 Rotaxanation 64

4.1.1 Controlled Delivery Methods 67

4.1.2 Use of GSH 68

5 RESULTS AND DISCUSSION 68

5.1 Pseudorotaxanes and Rotaxanes 68

5.1.1 Assembly 71

5.1.1.1 Slippage formation of dinitrobenzene rotaxane through the synthesis of capped cylinder 74

5.1.1.2 Post-assembly alkylation of the 2-imidazole cylinder and its pseudorotaxane 77

5.1.1.2.1 Crystal Structure 85

5.2 Decapping 87

5.2.1 Dinitrophenol 90

5.2.2 Method of Reaction 91

5.2.2.1 NMR Decapping Experiments 91

5.2.2.1.1 1:6 DNB Rotaxane to Mercaptoethanol NMR Experiment 93

5.2.2.1.2 1:24 DNB Rotaxane to Mercaptoethanol NMR Experiment 96

5.2.2.1.3 DNB Ligand Reaction with Mercaptoethanol. 100

5.2.2.1.4 1:20 DNB Cylinder to Mercaptoethanol NMR Experiment 104

5.2.2.1.5 NMR Decapping Conclusions 106

5.2.3 Releasing an Imidazole Cylinder 108

5.2.3.1 Acetonitrile 110

5.2.3.2 Water 114

5.2.3.3 Stability of the 2-imidazole Cylinder. 120
5.2.3.3.1 Stability in MeCN 124
5.2.3.3.2 Stability of the Pseudorotaxane 124
5.2.3.4 To Yield a DNA 3WJ Forming Species. 128
5.3 Conclusions 130
6 EXPERIMENTAL 132
6.1 Synthesis of 2-((2,4-dinitrophenyl)thio)ethan-1-ol 132
6.2 Synthesis of 1-(2,4-dinitrophenyl)-1H-imidazole-2-carbaldehyde 132
6.3 Synthesis of (1E,1'E)-N,N'-(methylenebis(4,1-phenylene))bis(1-(1-(2,4-dinitrophenyl)-1H-imidazol-2-yl)methanimine) (L^{DNB}) 133
6.4 Synthesis of DNB capped nickel cylinder ($[Ni_2L^{DNB}{}_3]^{4+}$) 134
6.4.1 Ligand first 134
6.4.2 Post-assembly modification of 2-imidazole cylinder 134
6.5 Post-assembly synthesis of DNB capped nickel rotaxane $[Ni_2L^{DNB}{}_3 \bullet CB[10]]\ 4DNP^-$ 135
6.6 Data Acquisition and Analysis 135
6.6.1 NMR Spectroscopy 135
6.6.2 UV-Vis Spectroscopy 136
7 REFERENCES 137
8 APPENDICES 140
8.1 NMR Spectroscopy 140
8.1.1 DNB-aldehyde Assignment and Coupling 140
8.1.2 DNB-ligand assignment 140
8.2 UV-Vis Spectroscopy 141
8.2.1 2-imCy Hydrolysis Products 141
8.2.2 DNB Rotaxane 142
8.2.2.1 In Water 142
8.2.2.2 In Acetonitrile 145
8.2.3 DNB Cylinder 146
8.2.3.1 In Water 146

8.2.3.2 In Acetonitrile 150
8.2.4 2-((2,4-dinitrophenyl)thio)ethan-1-ol 151
8.3 Rotaxane Isotope Pattern 152

3 Glossary

L^{2im} – 2-imidazole carboxaldehyde based ligand

L^{Py} – 2-pyridine carboxaldehyde based ligand

FeCy – Iron Cylinder. $[Fe_2L^{Py}_3]^{4+}$

2imCy – 2-imidazole cylinder $[Ni_2L^{2im}_3]^{4+}$

DNB – 2,4-dinitrobenzene

L^{DNB} – 2,4-dinitrobenzene modified 2-imidazole ligand

DNB-rotaxane (sometimes shortened to DNB-RTXN) – DNB capped 2-imidazole cylinder threaded through a cucurbit[10]uril macrocycle.

DNB-cylinder (sometimes shortened to DNB-Cyl) – DNB capped 2-imidazole cylinder

Slippage – A form of rotaxane synthesis where the completed axle is slipped inside the macrocycle.

Post-assembly modification – Chemical alterations made to a formed supramolecular structure. In this case a pseudo-rotaxane that is capped once formed.

3WJ – Three-way junction

S1 – $^{5'}$CGGAACGGCACTCG, part of a DNA 3WJ

S2 – $^{5'}$CGAGTGCAGCGTGG, part of a DNA 3WJ

S3 – $^{5'}$CCACGCTCGTTCCG, part of a DNA 3WJ

SNAr – Aromatic nucleophilic substitution

4 Introduction

4.1 Rotaxanation

The first reported instance of a rotaxane was in 1967,[2] though the name was first recorded two years later.[3] Rotaxanes have grown steadily in popularity in the last decade, as various approaches to their synthesis have become more widely understood.[4] There are different methods, through capping, clipping, and slipping. Even the mechanical contraction of a macrocycle on palladium coordination has been used.[5]

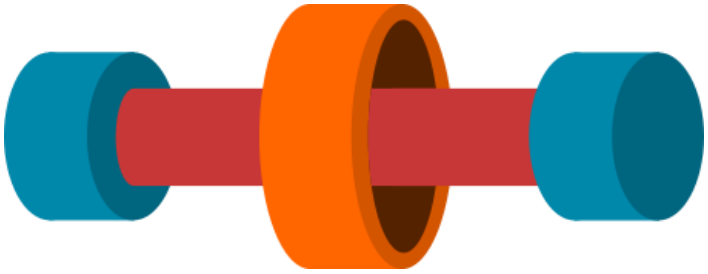


Figure 4.1-1 Generic dumbell illustration of a rotaxane, showing the thread in pink, the stoppers in blue, and the macrocycle in orange.

The uniting design of a rotaxane is the assembly of a structure with a threading molecule onto which stopper groups mechanically hold a threaded macrocycle. (

Figure 4.1-1) Photo-responsive rotaxanes are frequently used for molecular shuttling and switching.[6] Another area of study is ion-recognition as a result of mechanically constrained binding pockets.[7] Macrocycles also serve to protect rotaxanated fluorophores from hydrolytic bleaching in vesicles.[8] One interesting use of rotaxanation, is prevention of duplex DNA assembly. Achieved using a macrocycle threaded onto a

modified DNA sequence, this also has the effect of inhibiting PCR amplification.[9] Perhaps this can be used as a method to create artificial mismatches or direct a structure to a specific conformer. A well-established DNA binding agent is deliverable to cells using a cisplatin derived rotaxane stopper.[10]

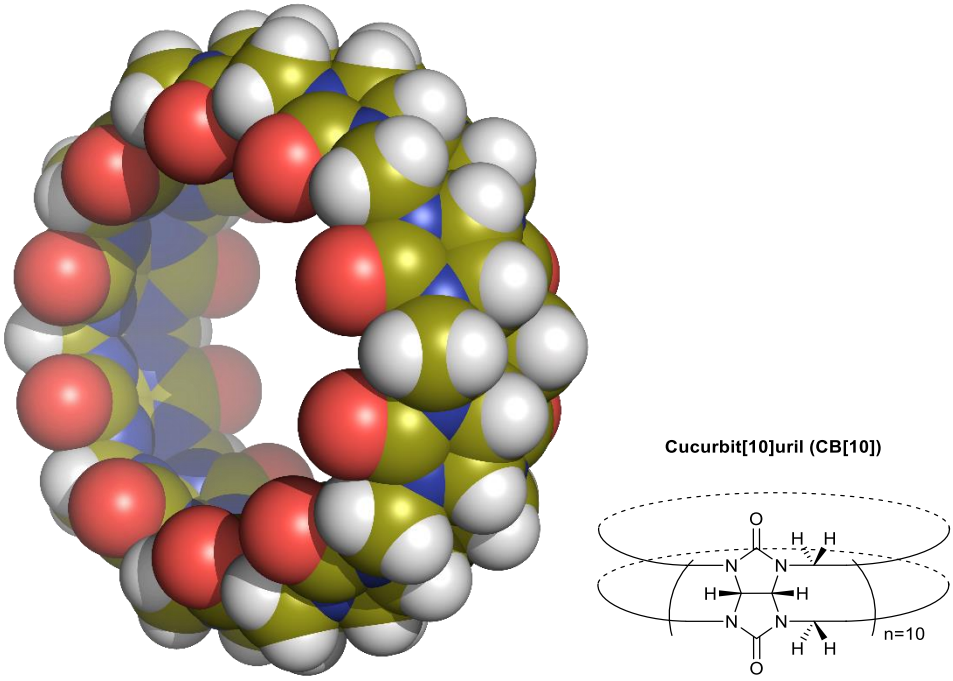


Figure 4.1-2 Left. Crystal structure of cucurbit[10]uril (CCDC no. 1507270). Excluding tetrachloro-cadmum(ii) chloride.[11] Right. Chemical structure of a cucurbit[10]uril.

As a macrocycle cucurbituril (CB) presents a hydrophobic cavity essential in the entropic binding of a threading molecule.[12] Whilst hydrophobic, the cavity has been shown to hold high-energy water molecules that drive guest binding with their release.[13–15] This has been estimated at as many as 22 molecules for cucurbit[10]uril (CB[10]),[16] though the effect is lesser for structures larger than CB[7] as the water is able to form hydrogen bonded networks more similar to bulk water.[17] On either portal are rings of ketones, which provide a highly polarized lipophobic area. This has been demonstrated to direct cations into the portal,[18–20] and has even been used for complexation.[21] (Figure 4.1-2)

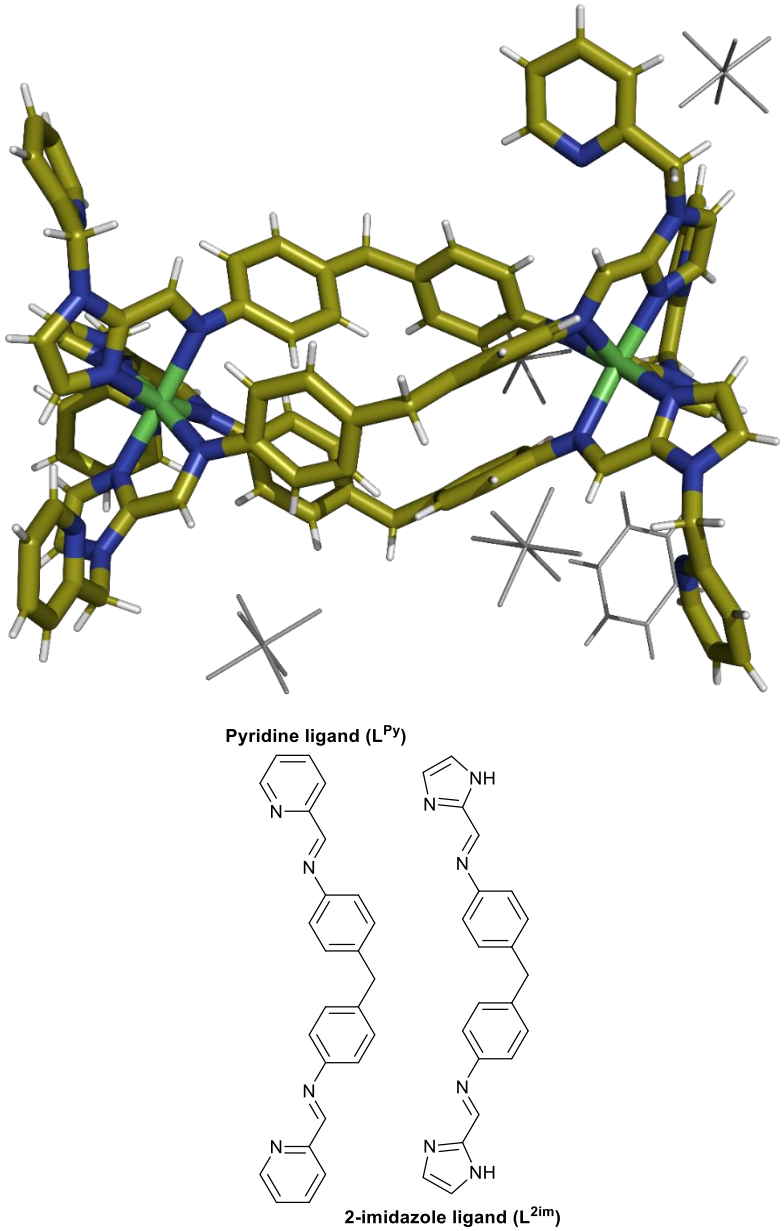


Figure 4.1-3 Left. Crystal structure of a 2-picolyl capped nickel(II) 2-imidazole cylinder $[Ni_2L^{2im}_3]^{4+}$. Co-crystallised with benzene and ammonium hexafluorophosphate.[1] (CCDC no. 1989563) Right. Structures of pyridine and 2-imidazole based ligands.

Using a cucurbit[10]uril we recently showed that a supramolecular helicate is an excellent candidate for binding to the center of the macrocycle to a high degree of affinity.[1] A pseudorotaxane assembled using a L^{Py} based ruthenium(II) or iron(II) helicates is protected from decomposition. Building on this, a L^{2im} based nickel(II) helicate can be capped by alkylating with methyl pyridine to form a rotaxane. It was

demonstrated that rotaxanation deactivated the DNA binding properties of the helicate, whereas the 2-imidazole pseudorotaxane was still able to bind, albeit to a lesser extent than the free helicate.[1] (Figure 4.1-3)

4.1.1 Controlled Delivery Methods

Attempts at controlling the release of bioactive molecules typically use enzyme cleavable covalent bonds, such as the cleavage of benzhydrocodone into hydrocodone,[22] or isavuconazonium into isavuconazole.[23,24] Alternatively use of lipid nanoparticles were recently used to deliver mRNA vaccines, releasing their payload on protonation in the endosome.[25] Using a similar principle macrocycles have been assembled into large cationic polyrotaxanes, that act to create nucleic acid aggregates.[26] These cargo-complexes are intended to deliver their payload intact for gene therapy. Biocleavable macrocycles and stoppers have been developed to trigger their release.[27,28]

Smaller scale, propeptides are shielded in a rotaxane, the stoppers cleavable by galactosidase, triggering spontaneous release.[29] Alternatively a galactosidase trigger on a macrocycle triggers a sequence of reactions ending in the opening of the macrocycle and release of a chemotherapy agent.[30]

DNA binding can be controlled through rotaxanating with a bipyridine-based macrocycle. The macrocycle is held mechanically using cleavable groups triggered enzymatically or via photo-irradiation. On cleavage, the macrocycle de-threads and an aryl ester connecting group hydrolyses, to release a G4 binding ligand.[31]

4.1.2 Use of GSH

Glutathione (GSH) is a tripeptide formed from glutamic acid, cysteine and glycine. It is a commonly found molecule within the body, protecting cells from oxidative stress. Elevated presence is found in bone marrow, breast, ovarian, colon, larynx and lung cancers.[32,33] With general cancer cell nucleus concentrations of GSH between 2-10 mM.[34] Interestingly at the same time, blood glutathione concentration decreases during certain chronic diseases.[35] By cleaving a disulfide bond with GSH, drug release was achieved from a mesoporous silica nanoparticle by removal of a mechanical blocking rotaxane.[36,37] Another small molecule photosensitizer has been shown to be activatable on release by glutathione.[38]

5 Results and discussion

In this chapter, rotaxanation is explored further with the purpose of controlling the DNA binding properties of the cylinders. It begins with investigation into the binding strength of cylinders to CB[10] some of which was featured in publication.[1] Following this the direction changes to the development of a rotaxane with removable capping groups. The rest of the chapter then explores the removal of these capping groups to release an active drug molecule.

5.1 Pseudorotaxanes and Rotaxanes

The following data on pseudorotaxanes comprises data acquired during the development of a publication on rotaxanation of helicates already mentioned in the introduction.[1]

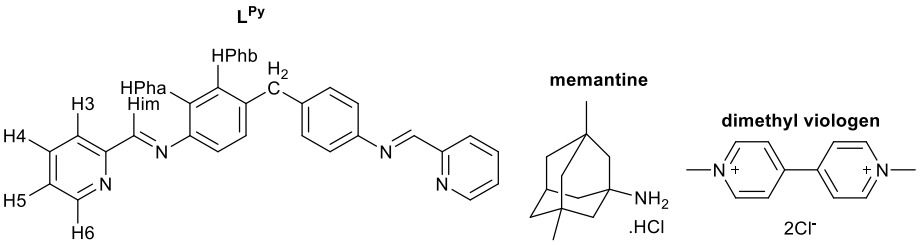


Figure 5.1-1 Structures of L^{Py} (left), memantine (middle), and dimethyl viologen (right).

To ascertain an estimation of binding strength of an iron cylinder (FeCy) bound to CB[10], the iron pseudorotaxane $[Fe_2L^{Py}_3]^{4+}$·CB[10] is used in a series of titrations. These are competition experiments using known cucurbituril binding agents memantine and dimethyl viologen as the competitors.

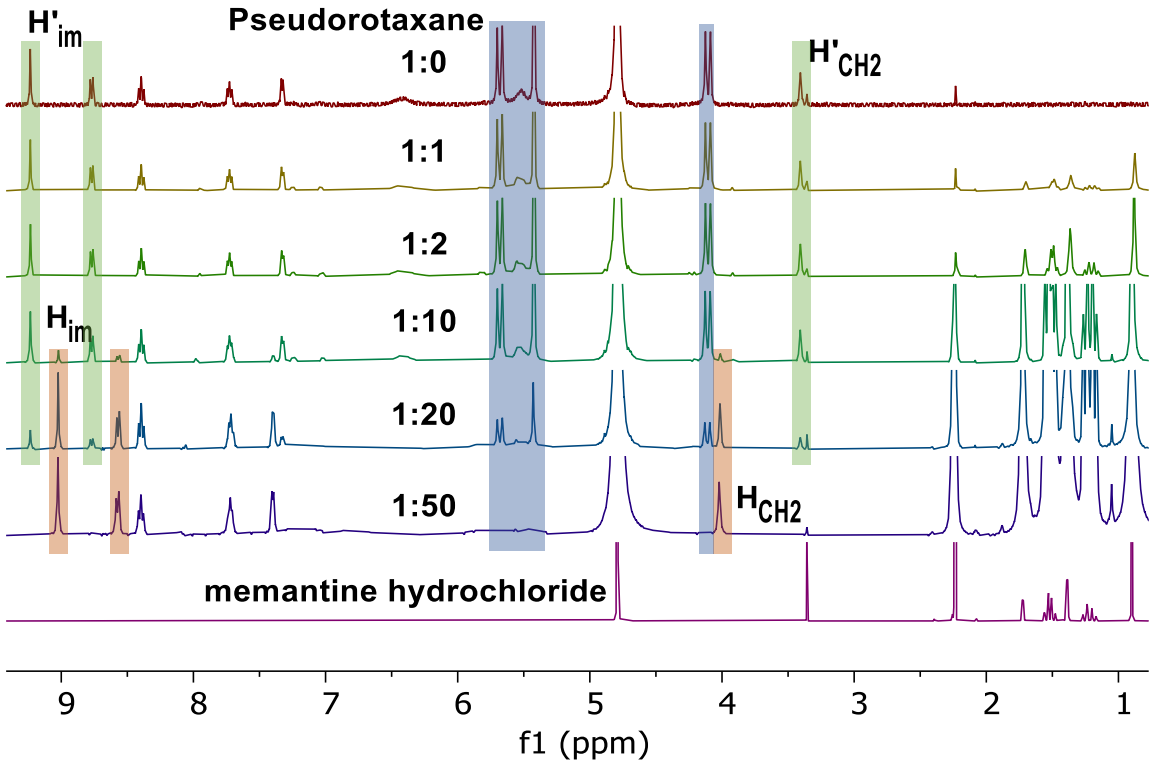


Figure 5.1-2 NMR competition titration of iron(II) pseudorotaxane $[Fe_2L^{Py}_3$·CB[10]]Cl_4 (1 mM) with memantine hydrochloride, in D_2O. Growing cylinder peaks highlighted orange, shrinking pseudorotaxane peaks from complex highlited green, shrinking pseudorotaxane peaks from CB[10] highlighted blue. Published as part of.[1]

On addition of 10 equivalents of memantine to a solution of iron pseudorotaxane a precipitate of (memantine)$_2$·CB[10] begins to form showing a reduction in the

pseudorotaxane peaks (H'_{im} and H'_{CH2}), and the growth of NMR peaks corresponding to free iron cylinder (H_{im} and H_{CH2}). All the peaks corresponding to these complexes, including the CB[10], agree with the labelled peaks. By 50 equivalents the cylinder is completely displaced from the CB[10]. This is evidenced by the complete loss of the pseudorotaxane peaks corresponding to the cylinder, and the reduction with significant broadening of the CB[10] peaks. Interestingly there is broadening of the iron cylinder phenyls, likely as they're interacting with the excess memantine in solution, though how is uncertain.[39]

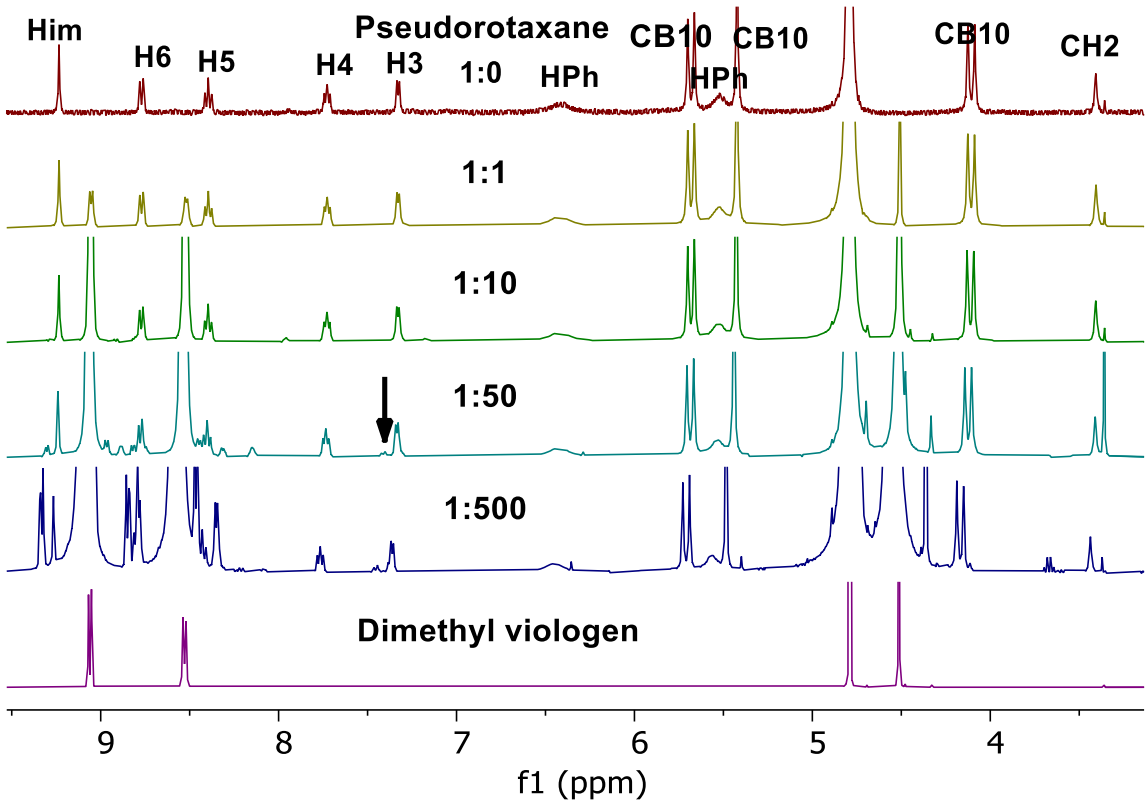


Figure 5.1-3 NMR competition titration of the iron(II)pseudorotaxane $[Fe_2L^{Py}{}_3{\cdot}CB[10]]Cl_4$ (1 mM) with dimethyl viologen chloride in D_2O. Dimethylviologen.CB[10] doublet shown with black arrow. At 1:500 equivalents, the

<1% impurity in the dimethyl viologen becomes more dominant than the pseudo-rotaxane, and is the cause of the new peaks seen here.

Another NMR titration using dimethyl viologen, showed no displacement up to 1:500 equivalents. In this experiment the small quantities of impurities in the dimethyl viologen eventually obscured the pseudorotaxane, it appears that no free cylinder is formed. However, on close inspection there is the beginnings of a characteristic dimethylviologen·CB[10] doublet at approx. 7.45 ppm.[40] There is a slight shift of peaks downfield with respect to the H_2O peak but this may result from short lived transient interactions. (Figure 5.1-3) Viologens are known to bind to cucurbiturils, the positive charge on the nitrogen well positioned for interactions with the portal.[40,41] Recently a viologen-phenylene-benzimidazole was shown to bind at a 3:2 ratio with CB[10].[42] Perhaps the larger cavity CB[10] is too great for 2:1 dimethylviologen/CB[10] to compete with the cylinder. When forming 2:1 dimethyl viologen/CB[10] complexes it was found that an inclusion complex of CB[10]·2methylviologen^{2+}·dihydroxynapthalene was more stable.[40]

5.1.1 Assembly

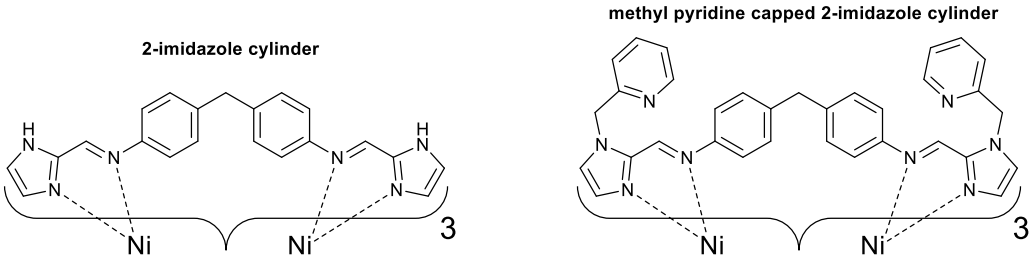


Figure 5.1-4 Structure of the 2-imidazole cylinder and the methyl pyridine capped 2-imidazole cylinder.

Developing upon pseudo-rotaxanes, it was possible to thread an imidazole cylinder and alkylate a nitrogen with methyl pyridine. This was presented in a journal article forming

a rotaxane that had inhibited DNA binding properties. As part of this work, using mass spectroscopy it was shown that mixtures of degrees of capping demonstrated greater activity than fully capped rotaxanes, a first foray into controlling the release of the cylinder from CB[10].[1] (Figure 5.1-4) A limitation was the inability to trigger the release of the 2-imidazole cylinder due to the stability of the methyl C-N bond. Were the complex able to slip from the CB[10] under extreme conditions, the remaining caps would reduce the effectiveness of the junction binding properties of the 2-imidazole cylinder.[1] Therefore, it was decided to explore groups that could be removed with more ease and triggered under specific circumstances. Ideally the capping group should be bulky enough to restrict release, whilst accessible/labile enough to permit the trigger. There are various methods for removable groups such as enzymatic cleavage, photo-activation, and reductive cleavage.[31,43]

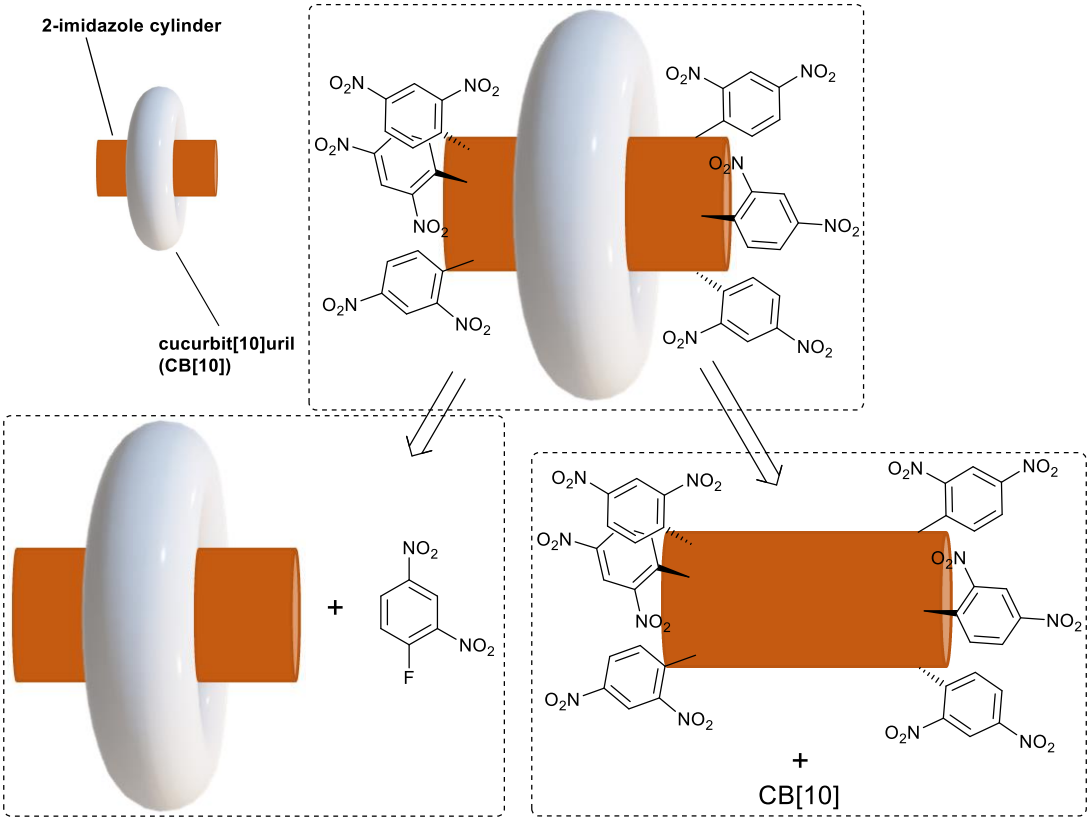


Figure 5.1-5 Illustrating the potential retrosynthetic routes to the rotaxane. Capping, where the pseudo-rotaxane is modified post-assembly; and slippage, where the synthesised capped cylinder is slipped inside the CB[10] to form the rotaxane.

Dinitrobenzene was a simple choice for a first attempt. The highly electron withdrawing nitro groups on 1-fluoro-2,4-dinitrobenzene allow for easy amination as seen with its use as Sanger's reagent.[44] The direct binding to the ring makes this moiety inflexible and a rigid capping species. Moreover, with the time constraints of working during the COVID-19 pandemic a simple design was needed.

Settling on dinitrobenzene for the first attempt at removable capping species, a synthetic route needed to be designed. Not wishing to complicate the synthesis by forming the cylinder core inside the rotaxane, there are two obvious routes to the final product, either the capped cylinder is formed beforehand and then slipped inside under

gentle warming conditions, or the pseudo-rotaxane is formed and the caps are applied as post-assembly modification. (Figure 5.1-5)

5.1.1.1 Slippage formation of dinitrobenzene rotaxane through the synthesis of capped cylinder

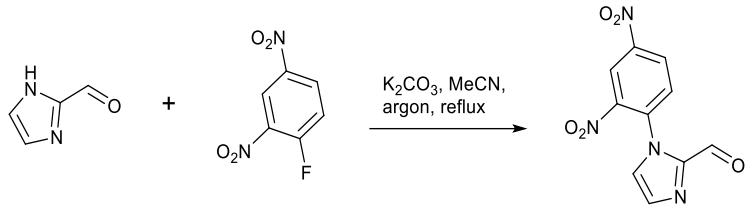


Figure 5.1-6 Scheme for the synthesis of 1-(2,4-dinitrophenyl)-1H-imidazole-2-carbaldehyde.

In order to synthesize the capped cylinder for the slippage method, first the aldehyde needs to be synthesized. This is done using an equivalent quantity of 1-fluoro-2,4-dinitrobenzene in the presence of potassium carbonate base. (Figure 5.1-6) The solvent acetonitrile is chosen for its aprotic nature coupled with likely good solubility of the product. The mixture is stirred under reflux until the starting material is consumed, then filtered to remove potassium carbonate crystals. Following concentration, the product is purified by silica gel flash column chromatography. It is important to isolate the product from other aldehydes present, to prevent irremovable impurities at later stages.

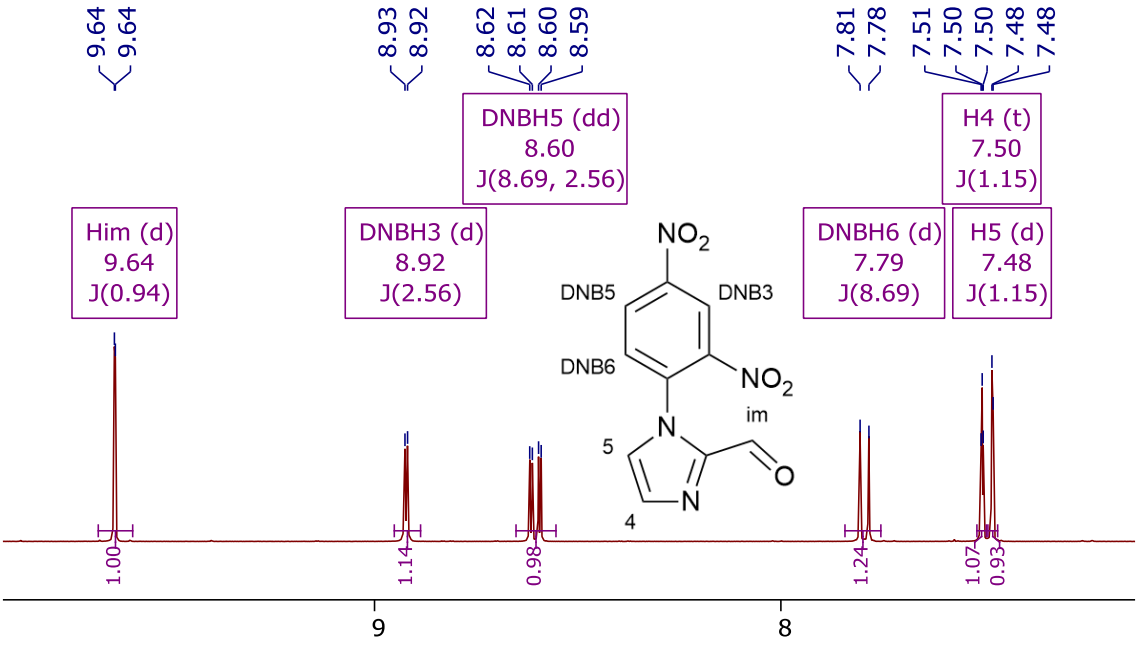


Figure 5.1-7 1H-NMR 400 MHz of DNB-aldehyde in acetonitrile-d3.

The isolated aldehyde has a clear set of coupled DNB peaks, with ortho coupling of 8.69 Hz between DNB5 and DNB6 and a meta coupling of 2.56 Hz between DNB3 and DNB5. (Figure 5.1-7) Interestingly the Him resonance is split into a doublet. There is no other peak split with the same 0.94 Hz constant, though COSY confirms this is a ^{5}J coupling between Him and H4. (Appendix Figure 8.1-1)

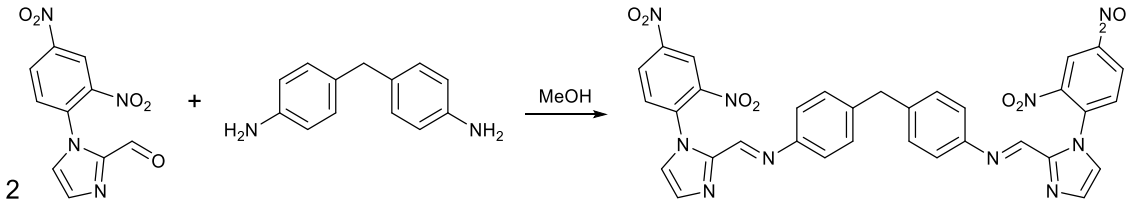


Figure 5.1-8 Scheme for the synthesis of dinitrobenzene modified 2-imidazole ligand (L^{DNB}).

Two equivalents of the aldehyde are then reacted with 4,4'-methylenedianiline in methanol to yield the ligand as a yellow precipitate, which is filtered and washed with cold ethanol. (Figure 5.1-8) Similar J coupling is seen for the ligand, including a 0.22 Hz

^{5}J Him and H4 coupling, though in this case it is more visible in the 1D spectrum. (Figure 8.1-2)

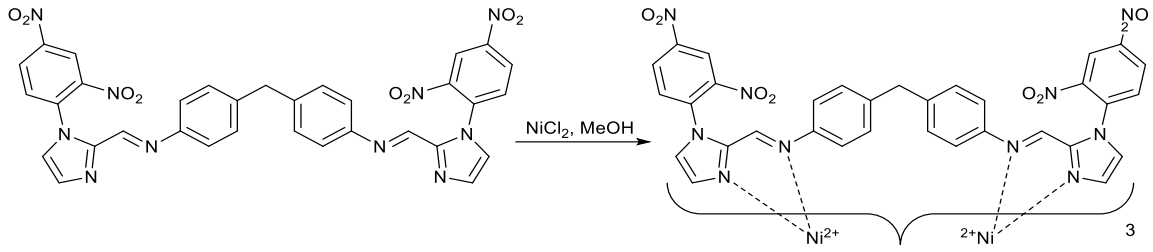


Figure 5.1-9 Scheme of the synthesis of dinitrobenzene modified cylinder from L^{DNB} and nickel chloride in methanol under reflux.

The ligand is then reacted in methanol under reflux with nickel chloride to form the dinitrobenzene modified cylinder as the chloride salt. (Figure 5.1-9)

Previously it has been possible to assemble methyl pyridine capped rotaxanes through slippage, so a similar procedure was attempted. A suspension of CB[10] (8.39 mg, 5.05 umol, 1 eq) and DNB cylinder (12.6 mg, 4.59 umol, 1.1 eq) in water (10 mL) was refluxed for 30 minutes and a portion assessed with mass spectrometry. No rotaxane was formed, so the reflux was continued for another 3 hours. Again, no rotaxane was formed. One reason for this could be poor solubility of the DNB capped cylinder and rotaxane in water as well as the insolubility of un-threaded CB[10]. Building on the understanding that high energy water is trapped within the macrocycle,[13] a suggested driving force for the easy assembly of the pseudo-rotaxane is the greater entropy gained on binding a water soluble threaded CB[10] compared to its crystalline structure. Various attempts have been made to thread the CB[10] with complexes in non-aqueous solvents but these were unsuccessful and adding aqueous solubilized 2-imidazole cylinder or iron parent cylinder pseudo-rotaxane to an excess of methanol proved to be

a suitable way of unthreading the CB[10]. It is likely that for the reaction the threading agent needs to be solubilized first, it was assumed the heat from the reflux would be enough to solubilize a portion of the suspension and so facilitate the steady threading of the rotaxane. Another challenge is the comparable instability of the DNB cylinder to 2-imidazole cylinder in aqueous solution.

A repeat experiment would be wise to try dropping an acetonitrile solution of DNB cylinder into a hot aqueous CB[10] suspension. This would have the combined effect of the raised likelihood of macrocycle/thread encounter, as well as reducing the time the cylinder spent in aqueous solution before threading. It is very apparent that the cylinder body is protected within the CB[10] to hydrolysis in aqueous solution once threaded. (Section 5.2.3.3)

5.1.1.2 Post-assembly alkylation of the 2-imidazole cylinder and its pseudorotaxane

An alternative route to the synthesis of the rotaxane is post-assembly modification. This was less desirable than the slippage method, due to the potential for failure to fully cap all available sites. It was first noticed that with the development of the methylpyridine modified complexes, but at the time there was value in being able to study the effects of different degrees of capping.[1] As such, when beginning the process of post-assembly alkylation with 1-fluoro-2,4-dinitrobenzene it was not expected to achieve complete capping the first try.

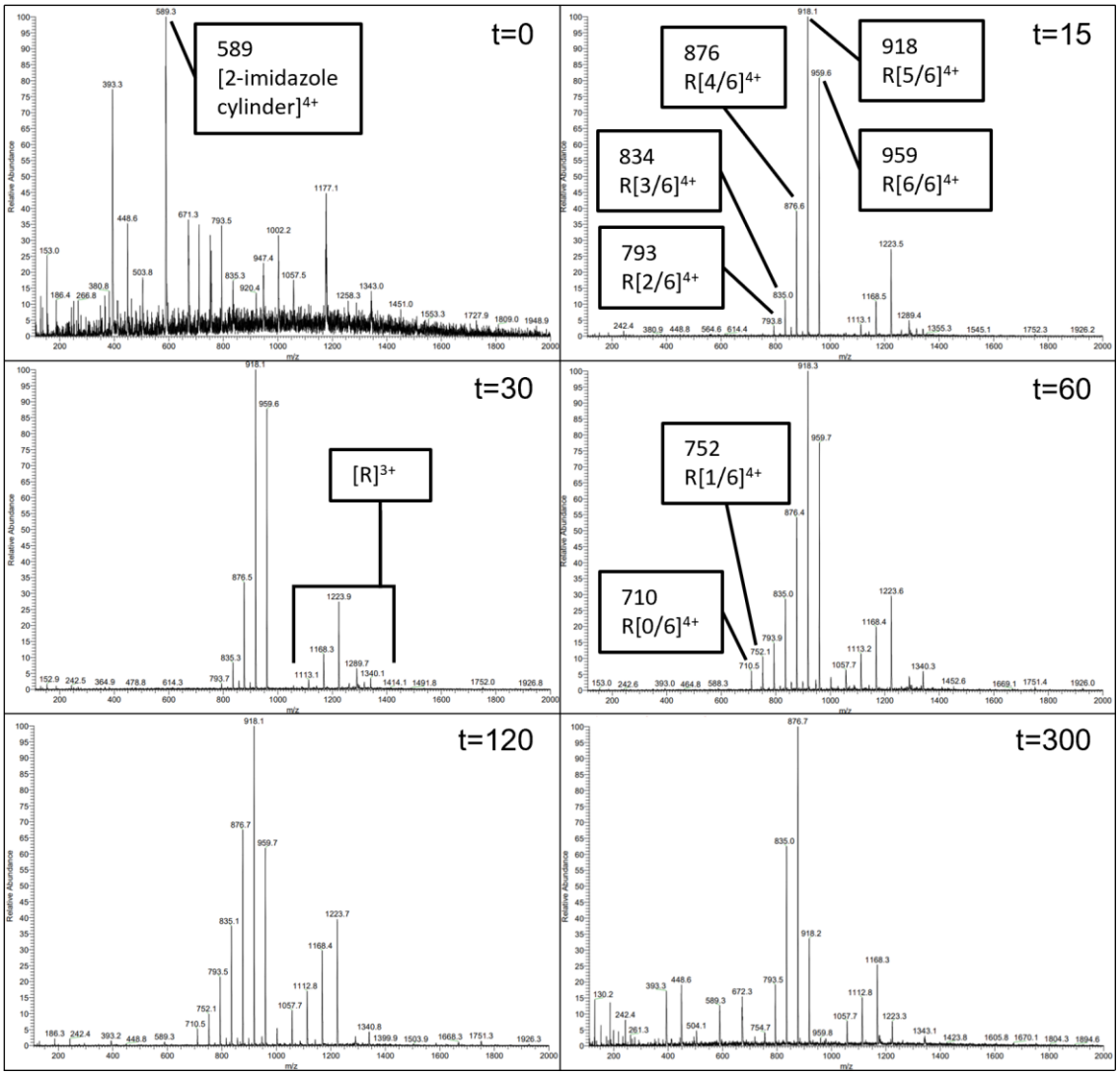


Figure 5.1-10 Low resolution ESI-MS spectra taken at 0, 15, 30, 60, 120, 300 minute time points for a rotaxane capping reaction. 2-imidazole cylinder (2.74 umol), CB[10] (3.01 umol) in water; followed by potassium carbonate (24.6 umol), 1-fluoro-2,4-dinitrobenzene (16.7 umol), stirred at 60 °C. R in labels indicates that the pseudo-rotaxane/rotaxane structure is threaded. Peaks reprentative of different degrees of capping for the rotaxane product, [6/6]4+ (959 m/z), [5/6]4+ (918 m/z), [4/6]4+ (876 m/z), [3/6]4+ (834 m/z), [2/6]4+ (793 m/z), [1/6]4+ (751 m/z), [0/6]4+ (710 m/z), [6/6+Cl]3+ (1291 m/z), [5/6]3+ (1224 m/z), [4/6]3+ (1168 m/z), [3/6]3+ (1113 m/z), [2/6]3+ (1057 m/z), [1/6]3+ (1002 m/z), [0/6]3+ (947 m/z).

To monitor the reaction a quick low resolution mass spectrum was effective, as the variably capped rotaxane species gave a visible distribution that could be tracked. Initial experiments with stoichiometric ratios failed to reach completion. The intensity of the $[6/6]^{4+}$ peak appears to decrease between 30 and 60 minutes at stoichiometric ratios.

(Figure 5.1-10) This is due to a competing reaction decapping the cylinder consuming the 2,4-dinitrobenzene. This is likely the hydrolysis product, 2,4-dinitrophenol which is seen as a counterion as the phenolate. Excess 2,4-dinitrophenol can be washed from the product with water.[45]

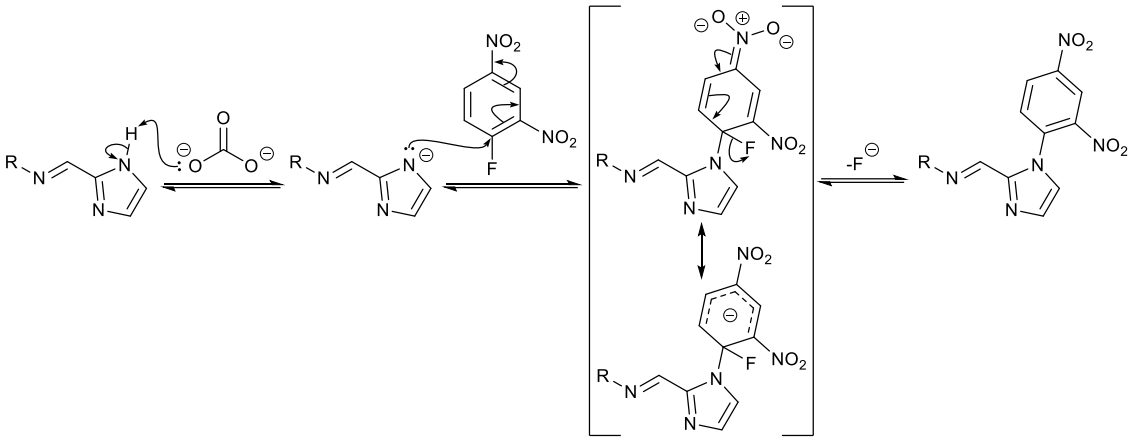


Figure 5.1-11 Proposed mechanism for alkylation of 2-imidazole in the presence of carbonate base with a fluoro-dinitrobenzene

The rate limiting step of the SNAr mechanism is the association of the aromatic species with the nucleophile, the dissociation of the leaving group happening rapidly, meaning the concentration of product was likely less important than reactants. (Figure 5.1-11) It occurred that the caps needed to be applied much faster than they were removed.

The apparent decrease in higher numbers of caps could have been due to an energetic barrier causing an increase in lower numbers of caps toward the end of reaction with no actual degradation, however from the initial increase in caps to form a 4/6, 5/6, 6/6 capped dominant population and the almost symmetry of the distributions it seems that there is no significant increase in energy required for the later caps.

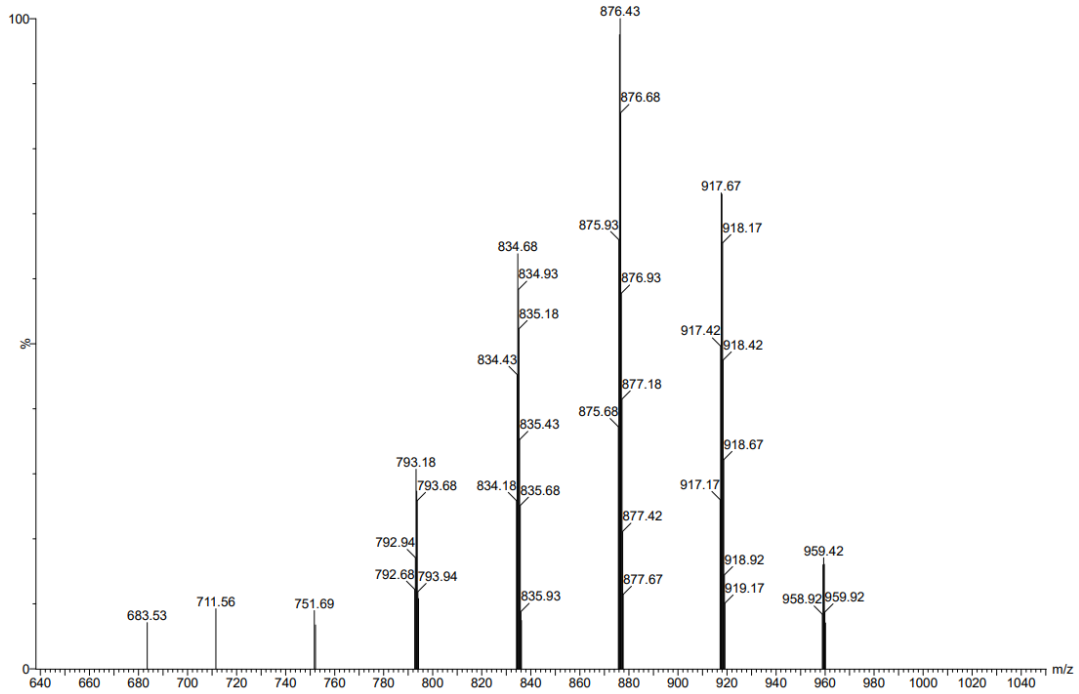


Figure 5.1-12 Mass spectrum of rotaxane synthesis product showing broad range of degrees of capped rotaxane. Potassium carbonate (12 eq), 1-fluoro-2,4-dinitrobeznene (18 eq), 20 degC 16 hours.

To make sure the degradation couldcould not be mitigated by lower temperature or that the reaction was not simply incomplete, repeats were conducted over longer reaction times, including ones with a larger excess of capping agent. The result was a well-formed distribution centered around 4/6 caps. (Figure 5.1-12)

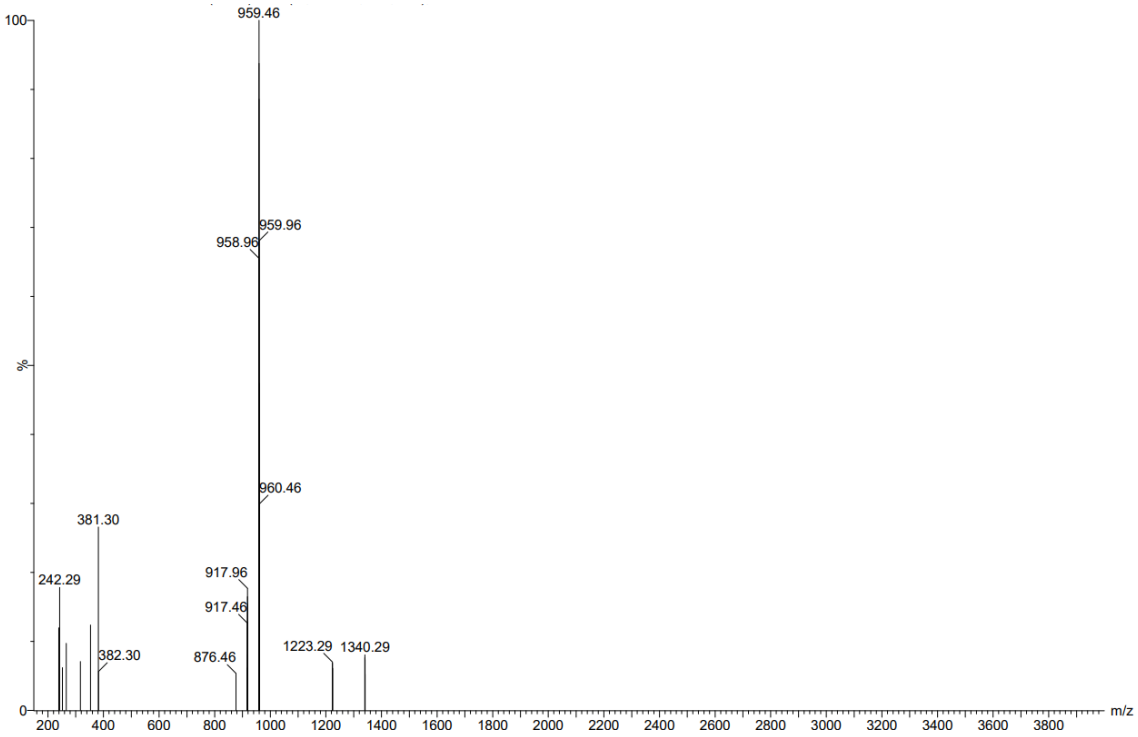


Figure 5.1-13 Mass spectra taken of the product of reacting nickel 2-imidazole cylinder (1 eq) with CB[10] (1.1 eq), 1-fluoro-2,4-dinitrobenzene (18.3 eq) and potassium carbonate (12.2 eq) in water at 80 °C for 50 minutes.

The most effective method found was to react the 2-imidazole cylinder with an excess of over 18 equivalents in deionized and degassed water at 80 °C, the reaction was complete after 50 minutes. (Figure 5.1-13) It is important to consider that the cylinder has 6 available binding sites so 18 equivalents is not excessive, at 3 equivalents per binding site.

The product is isolated as a bright orange precipitate. The synthesized rotaxane was found to be soluble in acetonitrile and DMSO, and poorly/in soluble in methanol and water. Later it was realized that the anion is the phenolate form of 2,4-dinitrophenol. This has the potential to complicate things with another dinitro-aromatic species in solution. However, for decapping purposes this is investigated later. (Section 5.2.1)

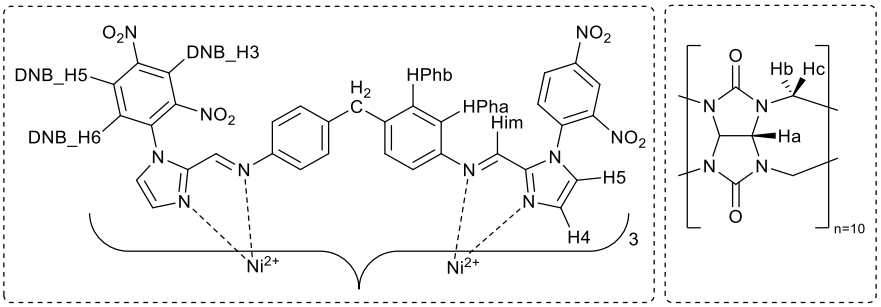


Figure 5.1-14 Labelling of a DNB capped 2-imidazole cylinder cylinder, and cucurbit[10]uril. Labelling applies to non-capped 2-imidazole cylinder too.

The assignment of the corresponding rotaxane NMR signals is complicated by its poor solubility and paramagnetism from the Ni(II). Numbering is shown in (Figure 5.1-14). Though similarities from the 2-imidazole cylinder/pseudorotaxane and capped cylinder help. In this case, the peaks shifted away from the “diamagnetic region”, are similar between spectra. An effect known as pseudocontact shift, caused by dipolar coupling between the magnetic moments of the nucleus and the unpaired electrons on the Ni(II), is seen. H5 and Him are pseudocontact shifted owing to their proximity to the Ni(II), H4 so much that the peak is not visible. It is possible slight humps between 90-130 ppm could be a H4 in the 2-imidazole cylinder and DNB rotaxane spectra, though a hump appears at 200 ppm for the DNB cylinder. When viewing humps in the spectra, it is important to be careful distinguishing from the distorted baseline resulting from loss of the early FID. (Figure 5.1-15)

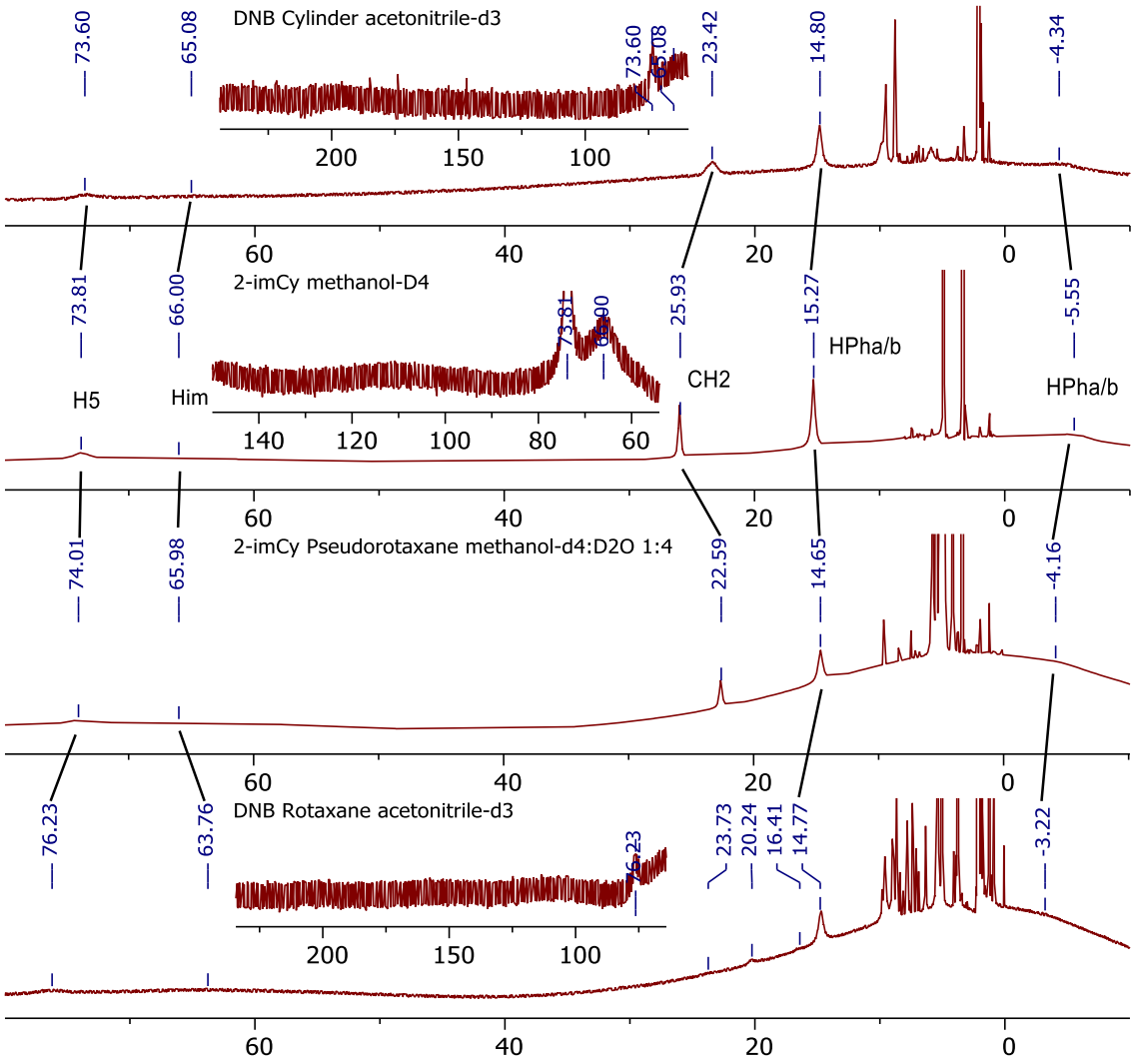


Figure 5.1-15 1H-NMR spectra showing the larger contact and pseudocontact shifts, stacked. Top to bottom. DNB cylinder in acetonitrile-d3 400 MHz, 2-imidazole cylinder in methanol-d4 500 MHz, 2-imidazole pseudorotaxane in methanol-d4:D2O 1:4 500 MHz, DNB rotaxane in acetonitrile-d3 400 MHz. 2-imidazole cylinder peaks are assigned, and dotted lines show the peaks occurring in other spectra.

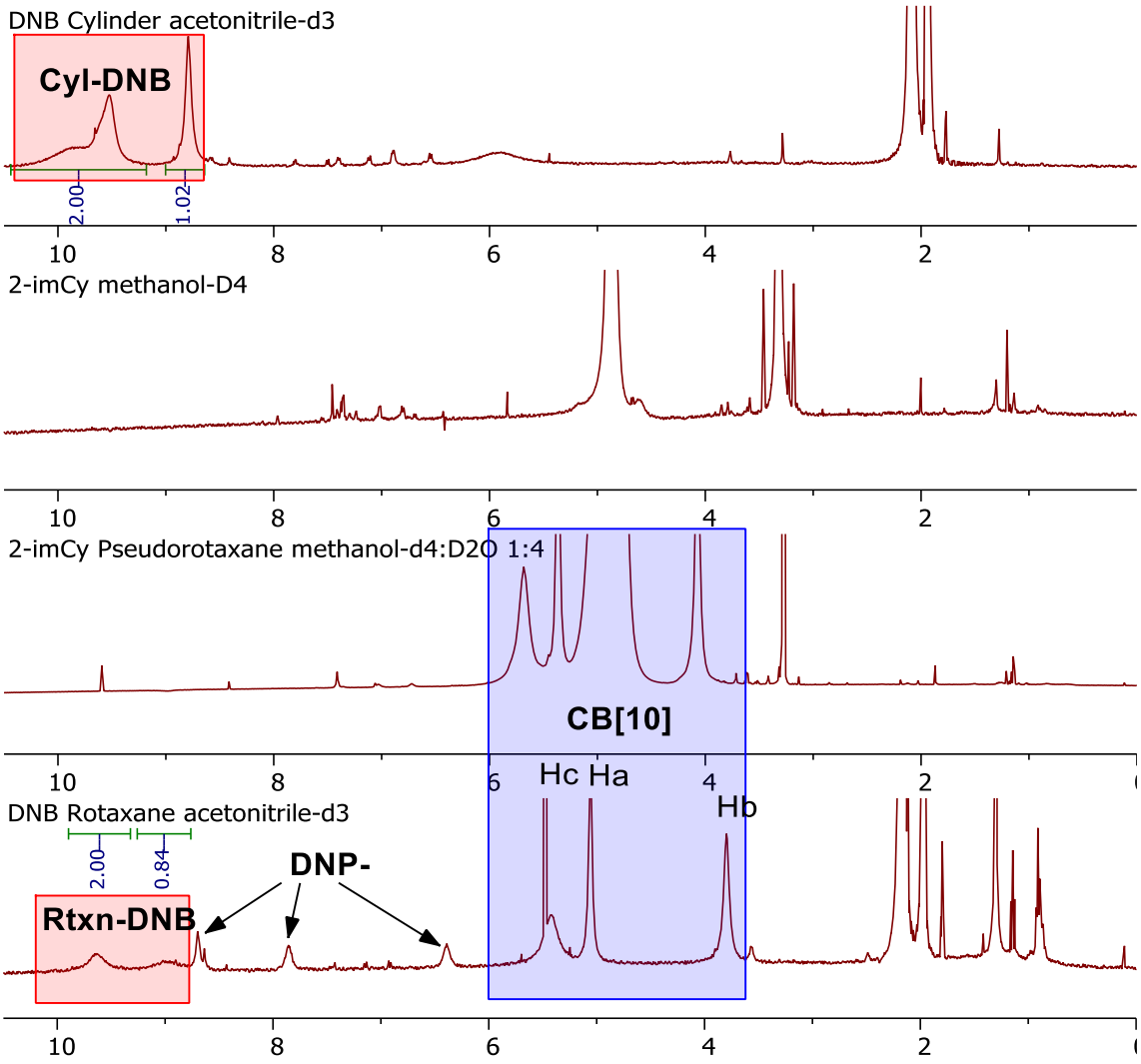


Figure 5.1-16 1H-NMR spectra showing the pseudocontact and diamagnetic shifts, stacked. Top to bottom. DNB cylinder in acetonitrile-d3 400 Mhz, 2-imidazole cylinder in methanol-d4 500 MHz, 2-imidazole pseudorotaxane in methanol-d4:D2O 1:4 500 MHz, DNB rotaxane in acetonitrile-d3 400 MHz.

Out of range for contact shifts, the cylinder DNB peaks appear closer to those of the diamagnetic ligand. Though the peaks still experience pseudocontact shifts, and broadening. The same is true for the rotaxane, though the peaks are significantly more broadened, likely due to proximity of the CB[10]. The DNB also appears as two peaks, though can be integrated to show that they are overlapped. CB[10] from the rotaxane can be compared to the 2-imidazole pseudorotaxane, where the peaks are similar. Also visible is the dinitrophenolate counterion (DNP^-).

5.1.1.2.1 Crystal Structure

Attempts at crystallizing the methyl pyridine capped rotaxane were unsuccessful. This included vapor diffusion of various solvent combinations using benzene, diethyl ether, acetonitrile and dichlorobenzene into methanol solution with the chloride and perchlorate counter ions. A small crystal was obtained from the iodide salt. This was done by dissolving the chloride salt in water, to this a potassium iodide solution was added precipitating the iodide salt. The mixture is then warmed, and water is steadily added until the sample is fully dissolved. The solution is then slowly cooled to permit crystallization. On analysis this was found to be cucurbit[10]uril alone.

Using the DNB rotaxane prepared through post-assembly modification as described in this chapter again attempts were made unsuccessfully with vapor diffusion methods. The sample was soluble in acetonitrile, so attempted crystallization was made through the slow vapor diffusion of benzene, methanol, ethanol, diethyl ether dichlorobenzene into a concentrated acetonitrile solution. These all resulted in precipitation, but no solid product of reasonable quality for single crystal analysis was acquired.

Crystals were obtained through the slow evaporation of acetonitrile from a sample. The complex was almost completely dissolved in acetonitrile, the undissolved remainder was filtered through a cotton plug with the filtrate passed directly into a cleaned vial, further acetonitrile of approximately 5% of the total volume was added. Onto the top of the vial a parafilm lid was added with a hole pricked using a needle to permit diffusion into the atmosphere. Crystal formation took approximately 2 months.

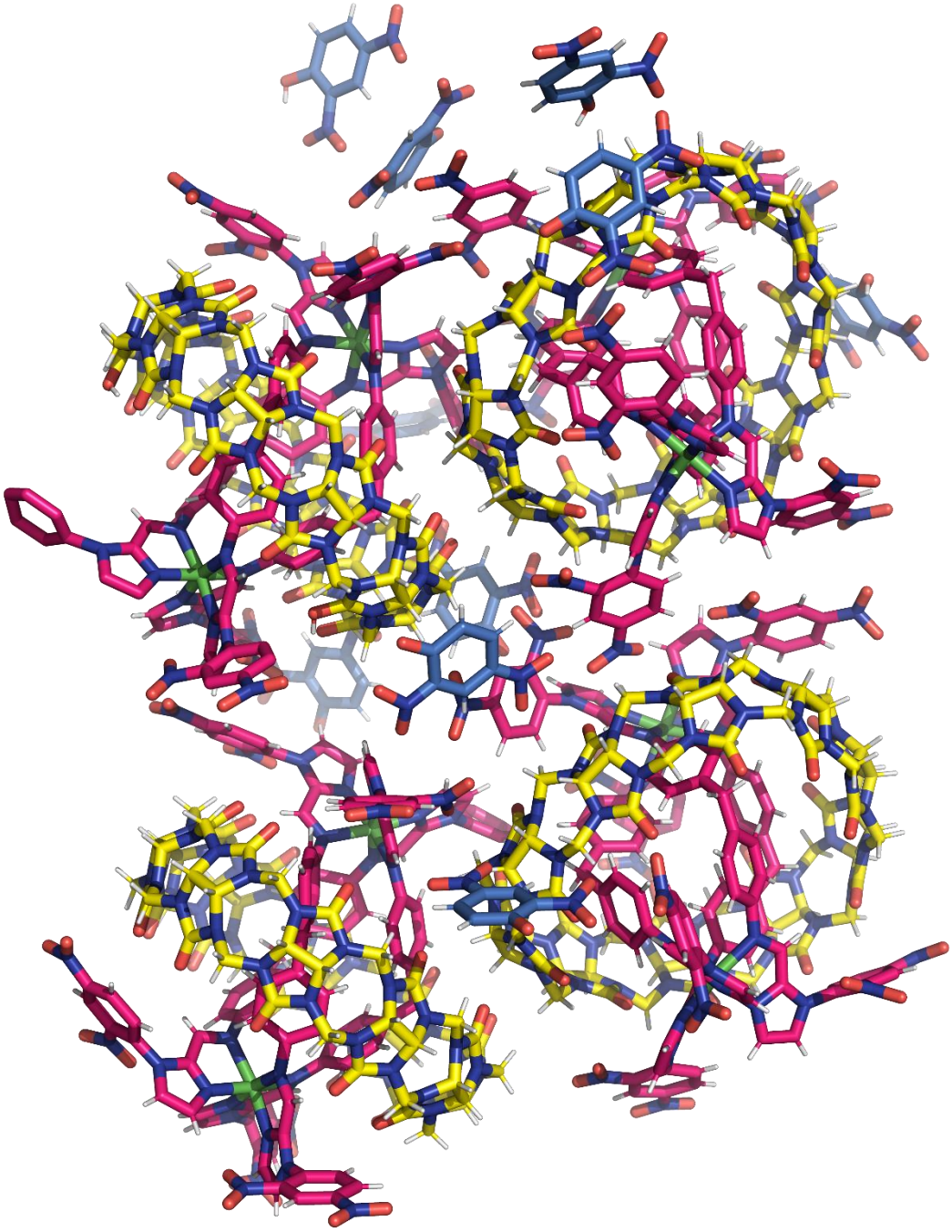


Figure 5.1-17 Partial crystal structure of the DNB rotaxane.

The structure is incomplete, most obvious from the 10 DNP^- present for the 4 $[Ni_2L_3]\cdot CB[10]^{4+}$ in the cell, leaving 4 F^- as likely anions. (Figure 5.1-17) Interestingly in this unit cell all complexes are the M enantiomer, it is possible that in the crystal there

are other P enantiomer containing unit cells. From preliminary testing, this seems to be the case though cannot be fully confirmed at this stage.

5.2 Decapping

Before experiments using DNA, it is important to understand how the rotaxane behaves in isolation; to identify the product of decapping reactions and a grasp on the rates at different concentrations. Initially glutathione was used, though the mass spectra were better simplified by using mercaptoethanol.

On incubating with mercaptoethanol (10 mM) at 37 °C, mass spectra of the solution at time point T=0 min demonstrated a reduction in $[R6/6]^{4+}$ and $[R5/6]^{4+}$ capped rotaxane peaks, with a similar increase in the $[R4,3,2/6]^{4+}$ capped rotaxanes to show an apparent shift in the center of the distribution to R4/6 caps. This continues in later time points with the center of the distribution around R2/6 caps by T=60 mins. The apparent even distribution on either side of $[R4/6]^{4+}$ capped rotaxane peak at T=0 min suggests no preference for the mercaptoethanol to react with the capped species above R1/6 and can be assumed to be similar for R1/6 too. It is harder to judge from the data the proportions of cylinder to rotaxanes in solution given the substantial difference in size and possible changes to ionizability resulting from the loss of the CB[10]. As the rotaxanes lose caps to become fully decapped pseudo-rotaxane or partially capped the cylinder is expected to be able to leave and bind to a competitor nucleic acid. In the mass spectrum an auspicious release of cylinder from CB[10] is seen in the form of $[C0/6]^{2+}$ and $[C1/6]^{2+}$ peaks at 1 hours. Adding iron parent cylinder to the mixture at 21 hours, a clear [FeCy pseudorotaxane]$^{4+}$ peak forms showing that the CB[10] is still intact

upon release. Considering the 3+ peaks, it is hard to compare between R6/6 and R5/6, as the available imidazole nitrogen is useful for the localization of a negative charge. This leaves the need for adducts to stabilize the R6/6, and as a result, there is $[R5/6\text{-}H^{+}]^{3+}$ representing the R5/6, and $[R6/6+NO2+H]^{3+}$, $[R6/6+DNP]^{3+}$ representing R6/6. (

Figure 5.2-1)

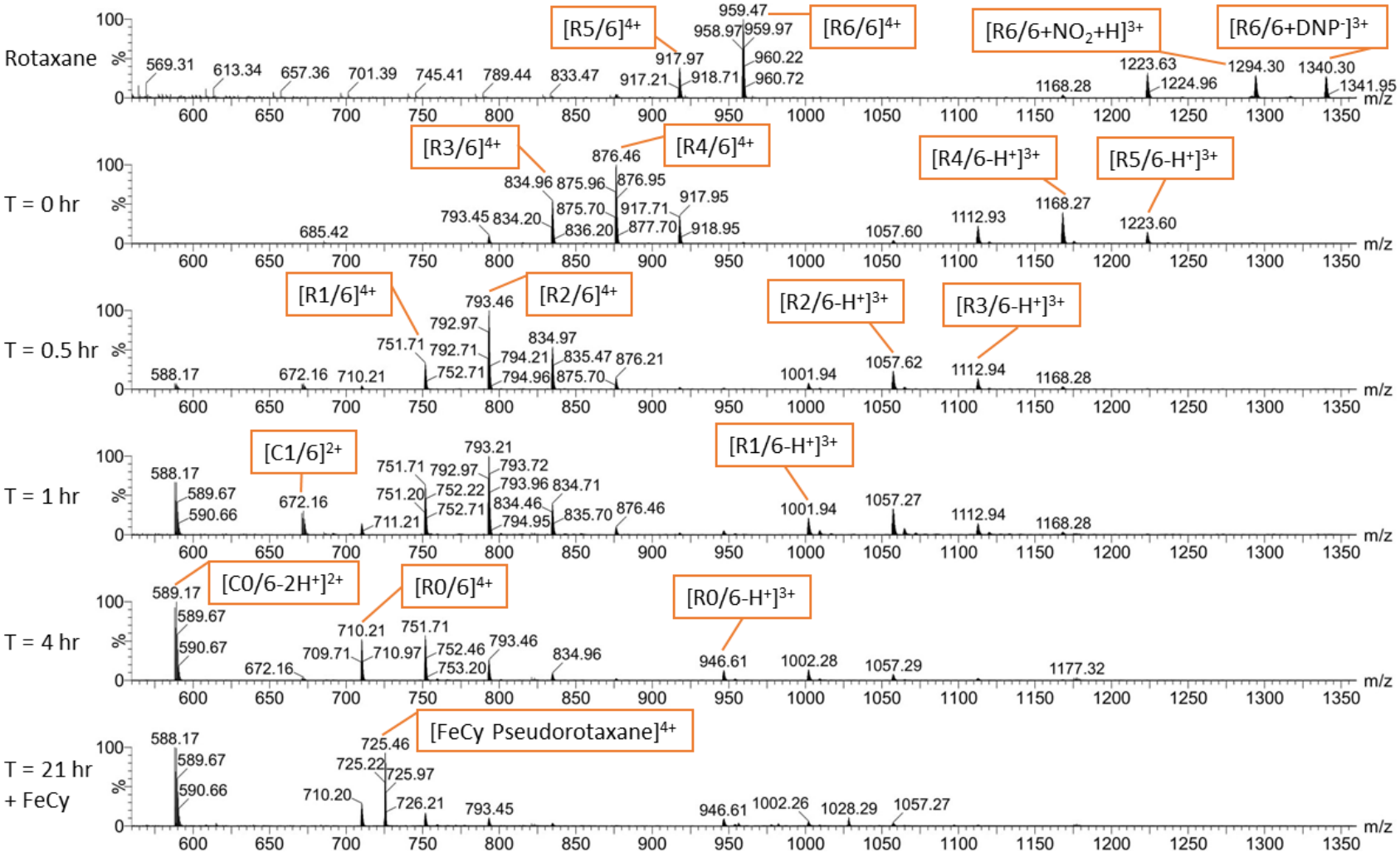


Figure 5.2-1 Mass spectrum of DNB rotaxane in methanol. Taken alone and with 10 mM mercaptoethanol at T = 0, 0.5, 1, and 4 hours and at 21 hours with added FeCy.

5.2.1 Dinitrophenol

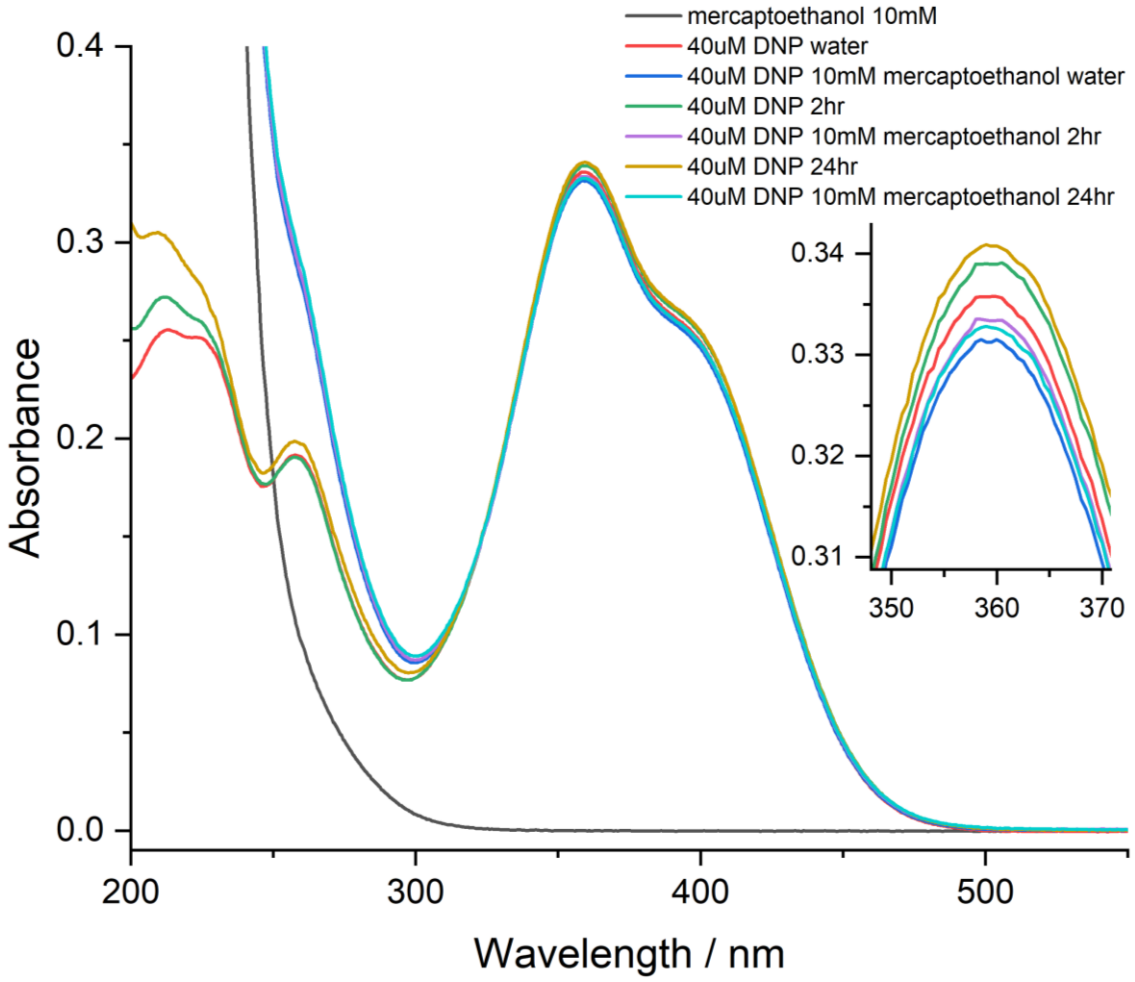


Figure 5.2-2 UV-Vis spectra of DNP (40 µM) in water alone, or water with mercaptoethanol (10 mM). Samples are measured and incubated at 37 °C at 0, 2 and 24 hours.

Incubating DNP over 24 hours, with mercaptoethanol has no changes to the UV-Vis spectra. This shows that the DNP does not react with mercaptoethanol in water. Considering the DNP acts as an acid with a pK_a of 4.1,[46] it is possible the reaction would proceed faster in the presence of a base that can deprotonate the mercaptoethanol. Though there should be no base present in the rotaxane samples to affect the decapping, so it can be assumed that the DNP^- is uninvolved with the decapping process. (Figure 5.2-2)

5.2.2 Method of Reaction.

Considering the history of researchers reacting nucleophiles with dinitrobenzene moieties, it can be theorized that the decapping method would undergo a similar nucleophilic aromatic substitution (SNAr) reaction through a Meisenheimer complex intermediate to form the 2-((2,4-dinitrophenyl)thio)ethan-1-ol (DNB-SEtOH) product. (Figure 5.2-3)

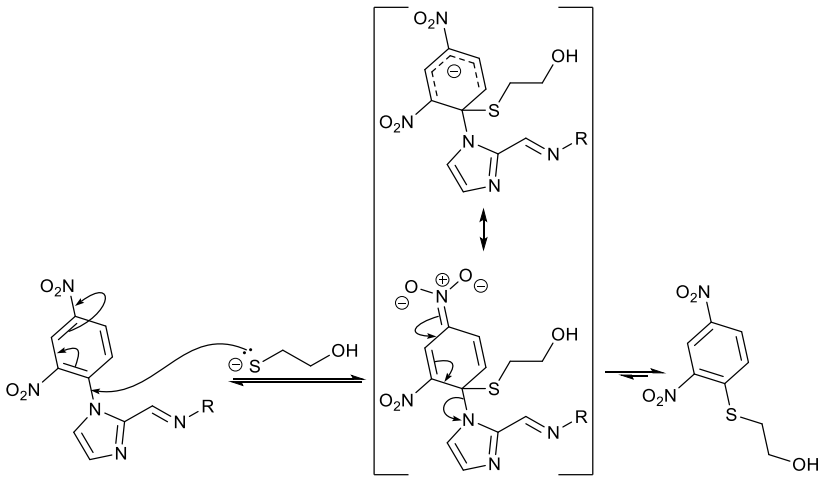


Figure 5.2-3 Proposed SNAr decapping reaction mechanism. The rate limiting step is the formation of the Meisenheimer complex, oweing to its higher energy compared to the starting material and product. The loss of the imidazole should be relatively fast as the aromaticity of the benzene is returned.

It is apparent from the mass spectrometry that the rotaxane reacts in a mixture of mercaptoethanol, with a distribution that moves towards completely decapped pseudorotaxane and 2-imidazole cylinder. However, to determine whether the dominant reaction is the SNAr involving the thiol, as opposed to another nucleophile, and whether the product is the expected DNB-SEtOH, further experiments were needed.

5.2.2.1 NMR Decapping Experiments

The expected product of decapping 2-((2,4-dinitrophenyl)thio)ethan-1-ol (DNB-SEtOH) should be diamagnetic. It was synthesized from equivalent quantities of

mercaptoethanol and 1-fluoro-2,4-dinitrobenzene using excess triethylamine base in dry chloroform. On running the NMR in acetonitrile-d3, an apparent quartet is visible at 3.84 ppm, which was theorized to be a doublet of triplets. This was proven to be the case with a repeat NMR in methanol-d4 where the alcohol was deuterated by the solvent. (Figure 5.2-4)

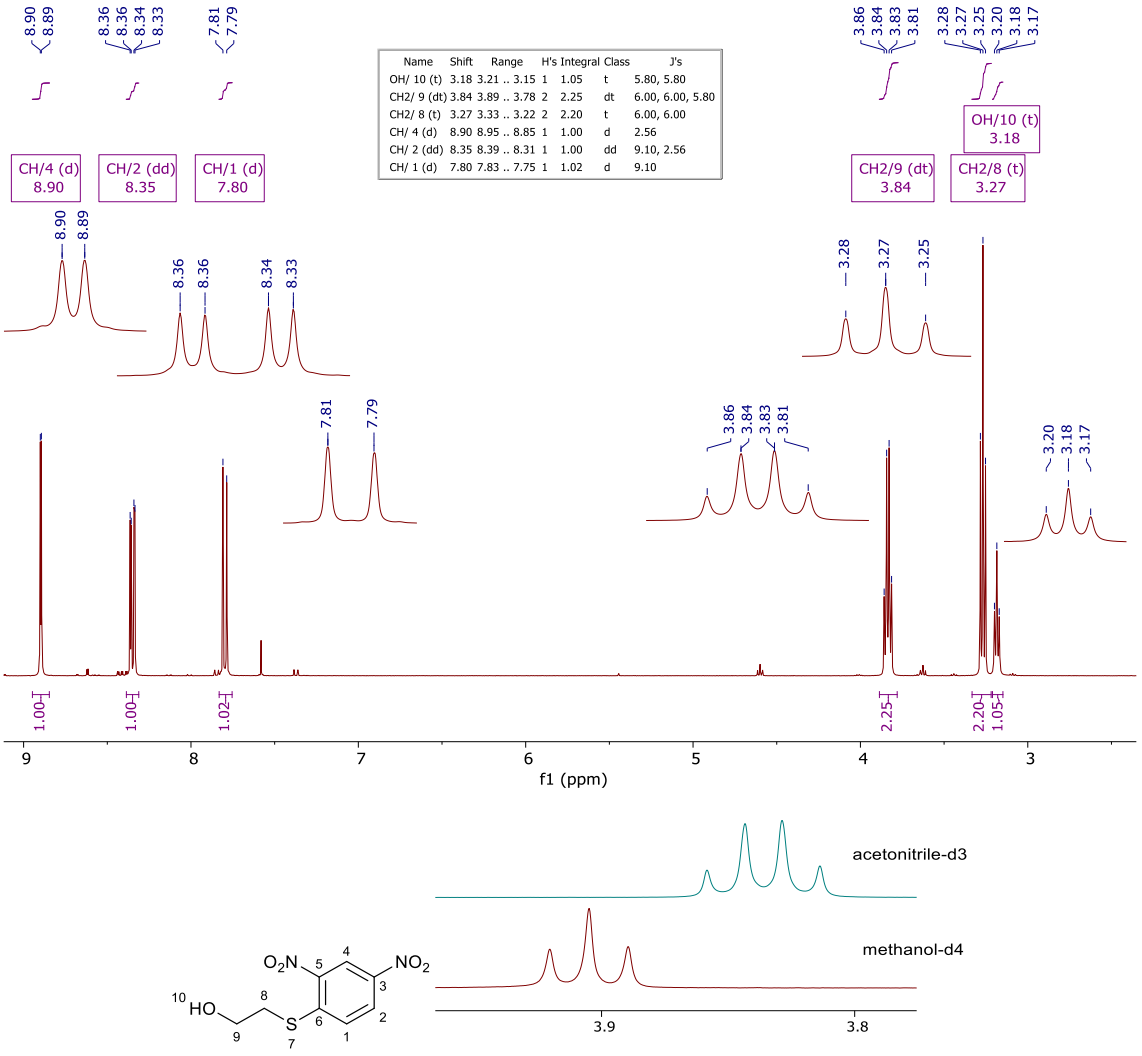


Figure 5.2-4 Top. NMR spectrum 400 MHz of 2-((2,4-dinitrophenyl)thio)ethan-1-ol in acetonitrile-d3. Bottom Left. Numbering scheme used in NMR. Bottom Right. Stacked NMR 400 MHz in acetonitrile-d3 and methanol-d4 to show the change in observed coupling on the CH2/9 nuclei resulting from exchange between H10 and D10.

For the NMR decapping experiment acetonitrile-d_3 proved to be the most suitable. Attempts were made to develop a method of adding DMSO-d_6 solubilized complex to

D_2O, but issues with precipitation prevented this, leaving acetonitrile as the best solvent for most experiments.

5.2.2.1.1 1:6 DNB Rotaxane to Mercaptoethanol NMR Experiment

The DNB rotaxane was decapped under a ratio of 1:6 at 0.43 mM rotaxane to 2.56 mM mercaptoethanol. Stoichiometric ratios of reactants, as there are 6 caps per rotaxane, were of interest to lower the rate of reaction whilst still maintaining the possibility of full decapping. Having a slower rate of reaction meant it was possible to increase the number of scans per spectrum, so these are clearer than the 10 mM measurements later. The experiment was designed with the intention to compare to UV-Vis data, so a lower concentration was also beneficial for UV-Vis to minimize the overshadowing of the other peaks with mercaptoethanol.

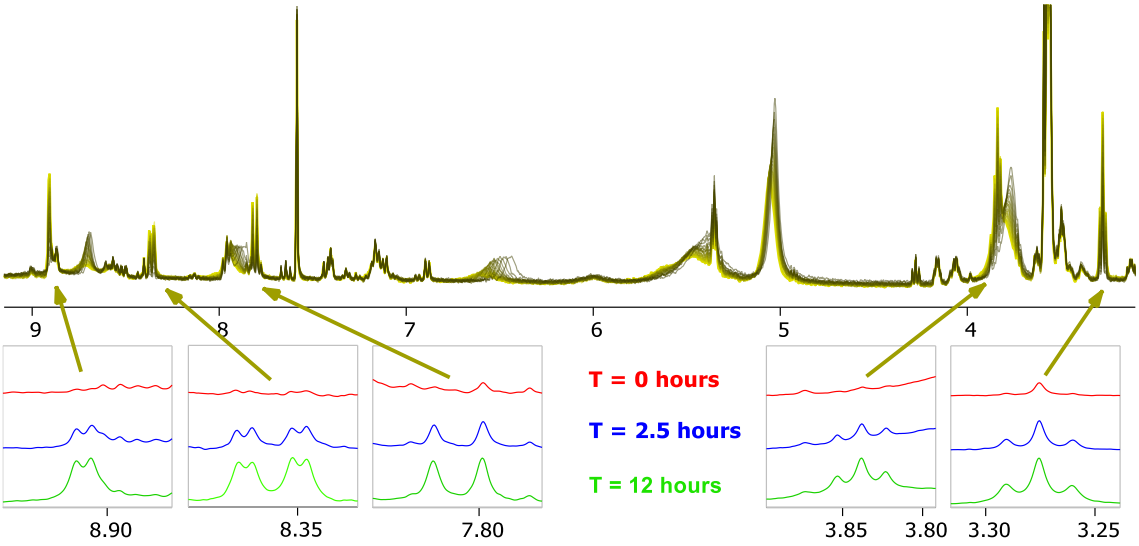


Figure 5.2-5 A series of overlayed 400 MHz NMR spectra of, 1:6 rotaxane (0.43 mM) to mercaptoethanol (2.56 mM) in acetonitrile-D3. Showing the changes in diamagnetic peaks over time, with expansions showing the growing DNB-SEtOH product.

Looking at the overlayed spectra, the only changing diamagnetic peaks are from the DNB-SEtOH increasing and the mercaptoethanol decreasing, demonstrating no release

of L^{DNB}, L^{2im}, their respective aldehydes or 4,4'-methylenedianiline, confirming that the process of release is entirely governed by decapping and that the resultant complexes are produced intact. (Figure 5.2-5)

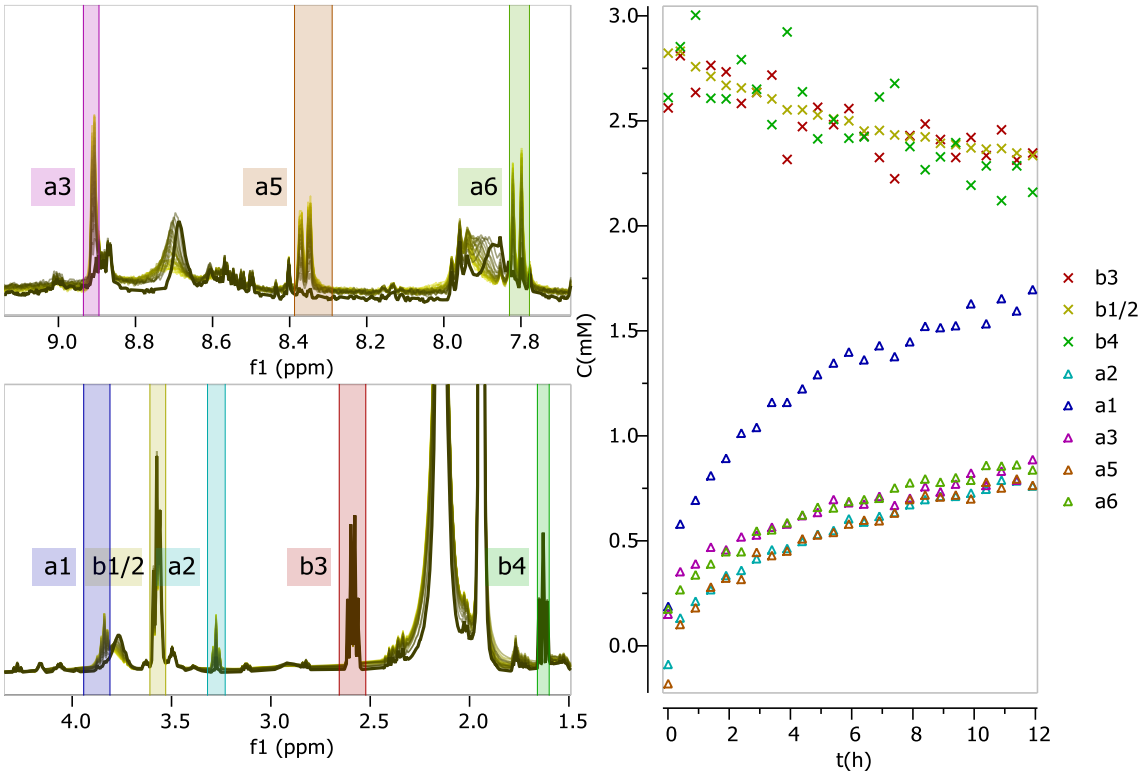


Figure 5.2-6 400 MHz NMR spectra of 1:6 rotaxane (0.43 mM) and mercaptoethanol (2.56 mM) in acetonitrile-D3. Showing the decrease of mercaptoethanol peaks (b1 = OH, b2 = CH2, b3 = CH2, b4 = SH) and the increase of DNB-SEtOH peaks (a1 = CH2, a2 = CH2, a3 = DNBH3, a5 = DNBH5, a6 = DNBH6). Concentration is set using the initial T=0 concentration of mercaptoethanol at 2.56 mM at peak b3. Concentrations are adjusted according to the number of nuclei involved in a peak.

Mercaptoethanol peaks show a shallow exponential decay over 12 hours, the OH and CH_2 are overlapped (b1/2). All three integrations seem to follow a similar curve, the 3 protons on b1/2 gives a smoother curve. At the same time the DNB-SEtOH peaks show an exponential growth over a similar amount. DNB-SEtOH CH_2 (a1) overlaps with a broadening CB[10] peak so appears slightly too intense. Series for the CH_2 (a2) and

DNBH5 (a5) begin at t=0 below 0, this is a result of the baseline appearing marginally above the spectrum. (Figure 5.2-6)

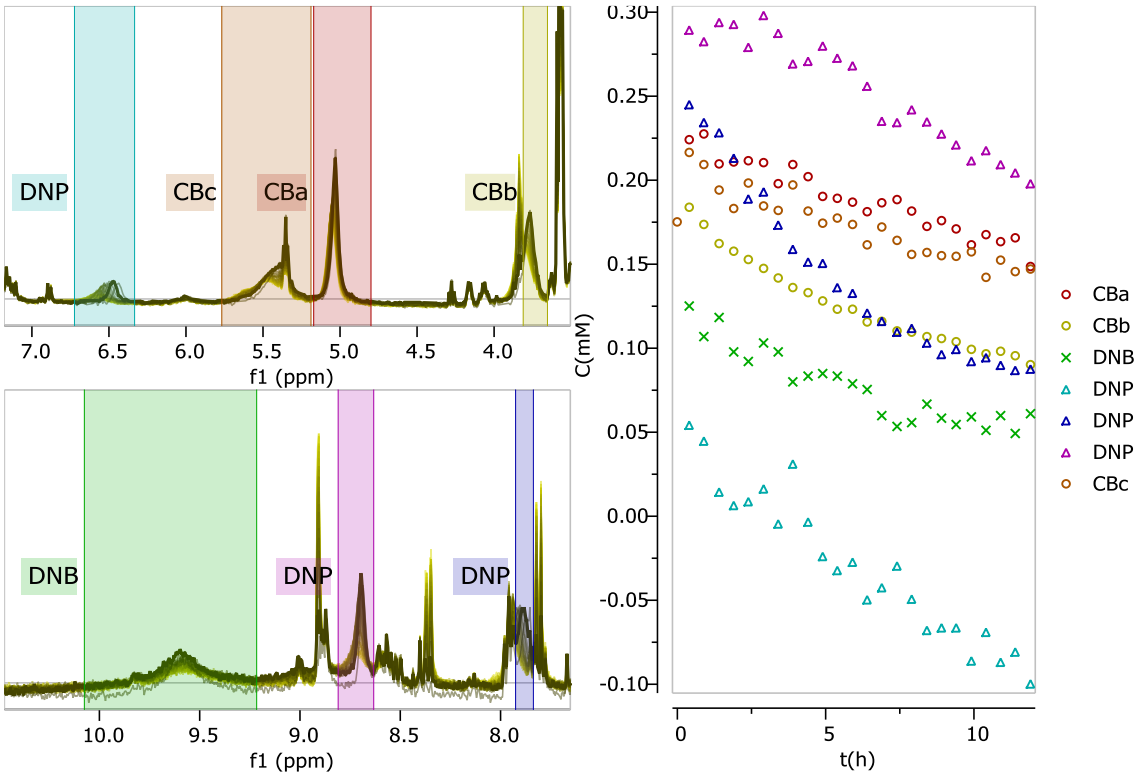


Figure 5.2-7 400 MHz NMR spectra of 1:6 rotaxane (0.43 mM) and mercaptoethanol (2.56 mM) in acetonitrile-D3. Showing the decrease of rotaxane peaks (CBa = CB[10]-Ha, 20H; CBb = CB[10]-Hb, 20H; CBc = CB[10]-Hc, 20H; DNB = DNB-2H, 12 H; DNP = DNP-H, 4H). Concentration is set using the initial T=0 concentration of mercaptoethanol at 2.56 mM at peak b3 (Figure 5.2-6). Concentrations are adjusted according to the number of nuclei involved in a peak.

The rotaxane peaks present a challenge, being paramagnetic they're broadened leading to them often overlapping with other peaks. A good example already mentioned is CB[10]-Hb, which overlaps with a DNB-SEtOH peak. More significant is the DNB peak, representing two of the three protons on the DNB ring. This is heavily broadened, reducing the intensity of the observed peak. The DNP illustrates another challenge with this dataset. The spectrum tends to drop below the baseline in regions of few peaks, leading to integrations lower than the "real" values. For the DNP peak at 6.5 ppm, this

follows the same trend as other DNP peaks, but ends at a negative integration due to broadening below the baseline. At 9.5 ppm the DNB peak also experiences this, but to a lesser extent. The broadening is likely a result of the increase in heterogeneity of capped rotaxanes. (Figure 5.2-7)

5.2.2.1.2 1:24 DNB Rotaxane to Mercaptoethanol NMR Experiment

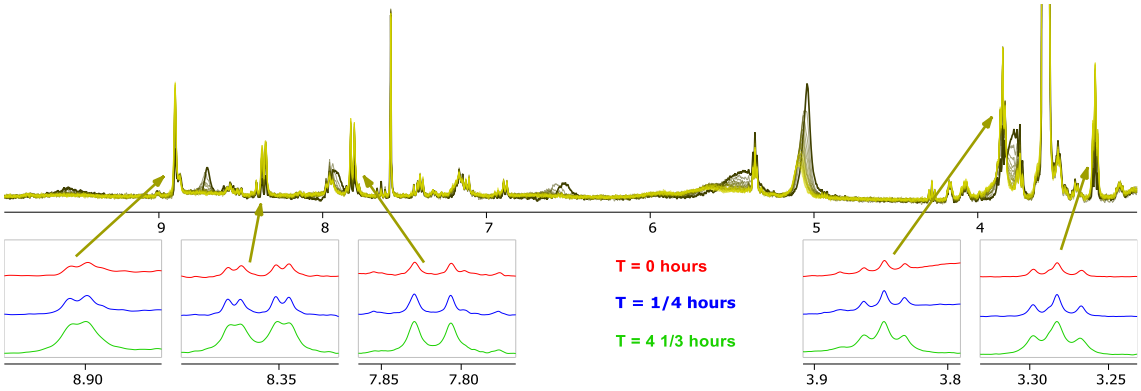


Figure 5.2-8 A series of overlayed 400 MHz NMR spectra of, 1:24 rotaxane (0.41 mM) to mercaptoethanol (10 mM) in acetonitrile-D3. Showing the changes in diamagnetic peaks over time, with expansions showing the growing DNB-SEtOH product.

Repeating the experiment with 0.41 mM rotaxane and 10 mM mercaptoethanol to give a ratio of approx. 1:24 shows the reaction occurring at a greater rate. As a result, the spectra need to be acquired faster, reducing the s/n ratio. However more dramatic changes to the peaks should provide useful information. Like the 1:6 ratio, the same diamagnetic DNB-SEtOH is the product of the reaction. There are also changes to the CB[10] peaks, and the DNB appears to be rapidly consumed. (Figure 5.2-8)

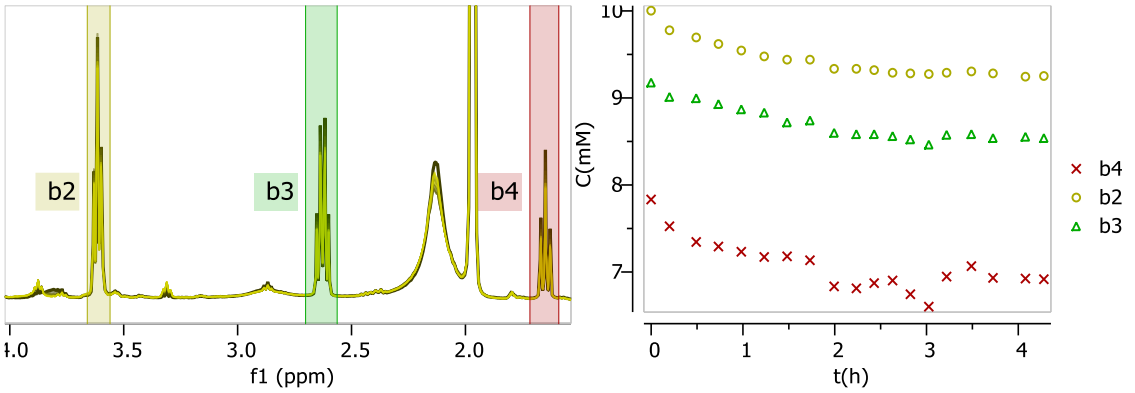


Figure 5.2-9 400 MHz 1H-NMR spectra of 1:24 rotaxane (0.41 mM) to mercaptoethanol (10 mM) in acetonitrile-d3, showing the decrease in mercaptoethanol peaks (b2 = CH2, b3 = CH2, b4 = SH). Concentration is set using the initial T=0 concentration of mercaptoethanol of 10 mM at peak b2. Concentrations are adjusted according to the number of nuclei involved in a peak.

The mercaptoethanol peaks differ in this dataset, because the OH group is not visible, whereas at 1:6 the OH peak was under a CH_2. There is also a greater variability between integrations, especially the SH peak, which appears lower. Possibly due to its proximity to the acetonitrile and water peaks, or perhaps the SH is exchanging more rapidly in the lowered pH caused by the SNAr reaction. All exponential decays appear similar except for the offset. (Figure 5.2-9)

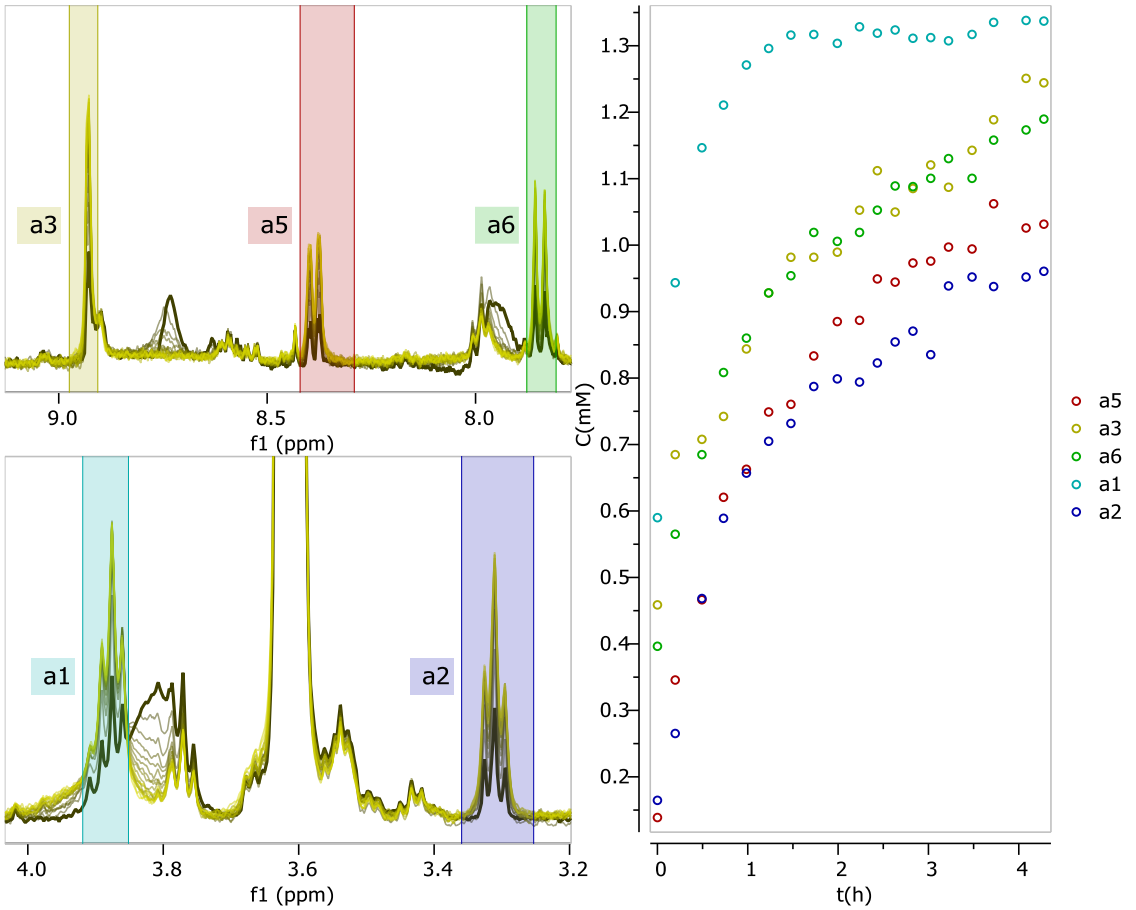


Figure 5.2-10 400 MHz 1H-NMR spectra of 1:24 rotaxane (0.41 mM) to mercaptoethanol (10 mM) in acetonitrile-d3. Showing the increase of DNB-SEtOH peaks (a1 = CH2, a2 = CH2, a3 = DNBH3, a5 = DNBH5, a6 = DNBH6). Concentration is set using the initial T=0 concentration of mercaptoethanol of 10 mM at peak b2. (Figure 5.2-9) Concentrations are adjusted according to the number of nuclei involved in a peak.

Again, the DNB-SEtOH forms as the diamagnetic product, more rapidly this time showing a dependence for its formation on mercaptoethanol concentration which is unsurprising. The CH2 in proximity to the OH (a1) overlaps with a broadening CB[10] peak, resulting in a rapid initial increase in integration as the resonances further overlap. The peak then plateaus as the CB[10] peak reduces in intensity in the region integrated countering the growth of the CH2 peak. (Figure 5.2-10)

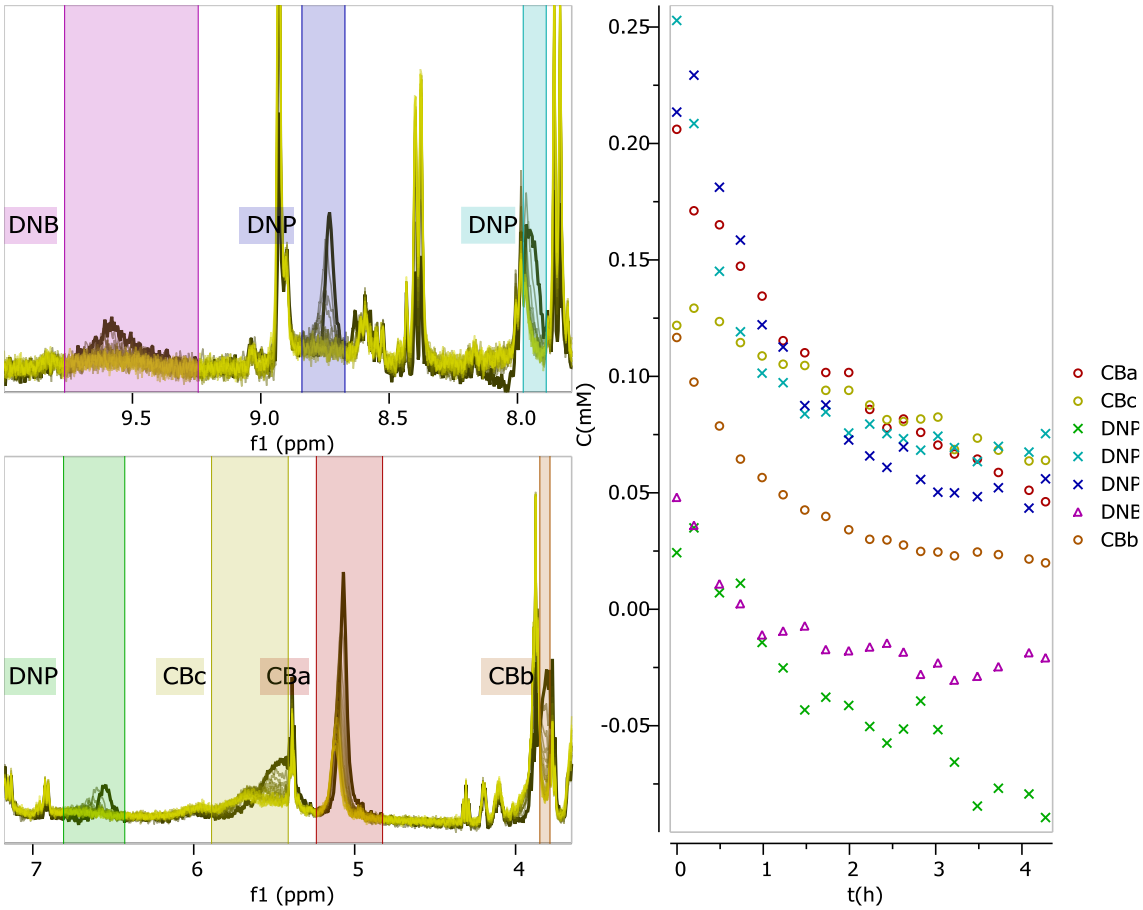


Figure 5.2-11 1H-NMR spectra of 1:24 rotaxane (0.41 mM) to mercaptoethanol (10 mM) in acetonitrile-d3. Showing the decrease of rotaxane peaks (CBa = CB[10]-Ha, 20H; CBb = CB[10]-Hb, 20H; CBc = CB[10]-Hc, 20H; DNB = DNB-2H, 12 H; DNP = DNP-H, 4H). Concentration is set using the initial T=0 concentration of mercaptoethanol of 10 mM at peak b2. (Figure 5.2-9) Concentrations are adjusted according to the number of nuclei involved in a peak.

The rotaxane peaks all show an exponential decay, some shallower than others. The variation is likely a combination of the broadness of the peaks and the limits of integrating regions over large overlapping peaks. The CB[10] peaks show broadening and de-shielding, which might occur from removal of electron dense DNB groups from nearby. Again, the DNP seems to be consumed in the experiment, whether this is due to a reaction with the mercaptoethanol, or another reaction is uncertain. Though the UV-Vis experiment with these in isolation showed no change. (section 5.2.1) Given the peaks seem to shift and broaden this would point to a change in the local environment

for the DNP. One way this could occur would be changes to the rotaxane that render it a non-uniform cation for the DNP^-. (Figure 5.2-11)

From the NMR data whilst decapping the rotaxane it is hard to determine whether it is releasing a cylinder or not as all the rotaxane peaks appear to be decreasing. There is also no apparent growth in capped cylinder, or 2-imidazole pseudorotaxane, though these would not necessarily be the products of decapping. Perhaps the species formed here are a mixture of partially decapped rotaxanes. The mixture of populations would provide peaks too broad to visualize and a transition to these would appear as a decrease in the regions integrated. This considered it is still important to rule out other processes occurring.

5.2.2.1.3 DNB Ligand Reaction with Mercaptoethanol.

Reacting DNB ligand with mercaptoethanol was conducted inside the NMR tube, measured using a series of timed experiments with automated shimming, tuning and matching before each measurement. Measurements were taken every 20 minutes for 24 hours, then at 48 hours another measurement was taken.

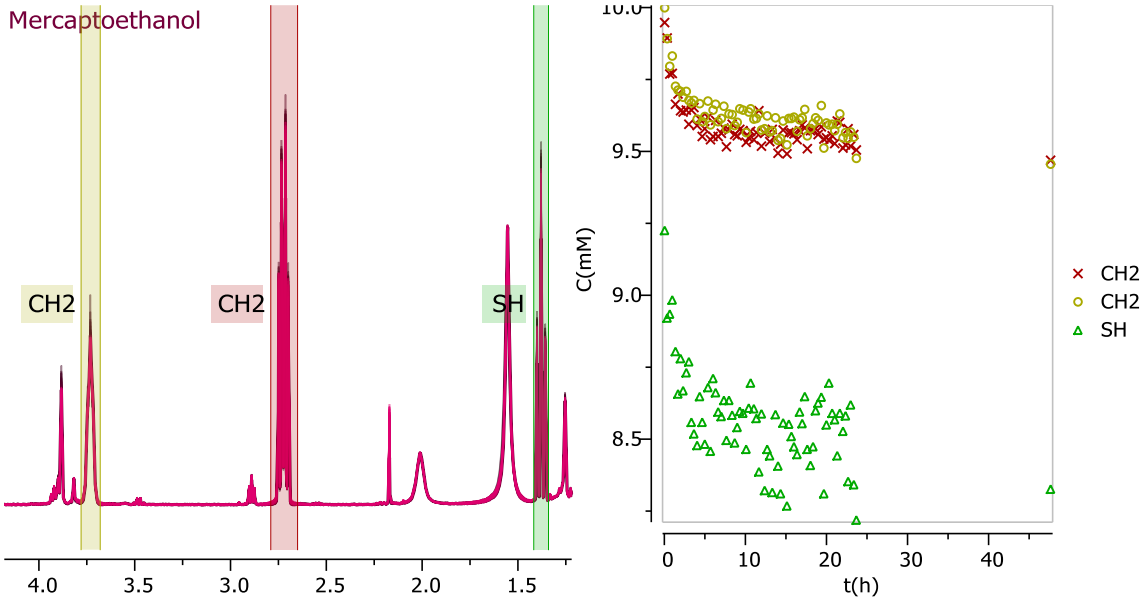


Figure 5.2-12 Time dependent 1H-NMR in chloroform-d, showing the mercaptoethanol peaks (left) and their integrations vs time (right). DNB ligand (ca. 4 mM) with mercaptoethanol (10 mM) at 298K. Concentrations are set to 10 mM with respect to mercaptoethanol region C at T=0.

Looking at the mercaptoethanol peaks, there are three distinct peaks. The integration at T=0 between 2.65-2.79 ppm has been used to derive the integrated concentrations for the rest of the experiment. The peak between 1.34-1.42 ppm is close to a broad water peak resulting in greater variation between measurements and a slightly reduced integration. These three peaks show an exponential decay over 24 hours, with some reaction still occurring between 24 hours and 48 hours. Decreasing approximately 0.5 mM over 48 hours, it is apparent from the rapid initial reaction that the time between mixing and first measurement could contain a significant portion of the reaction. (Figure 5.2-12)

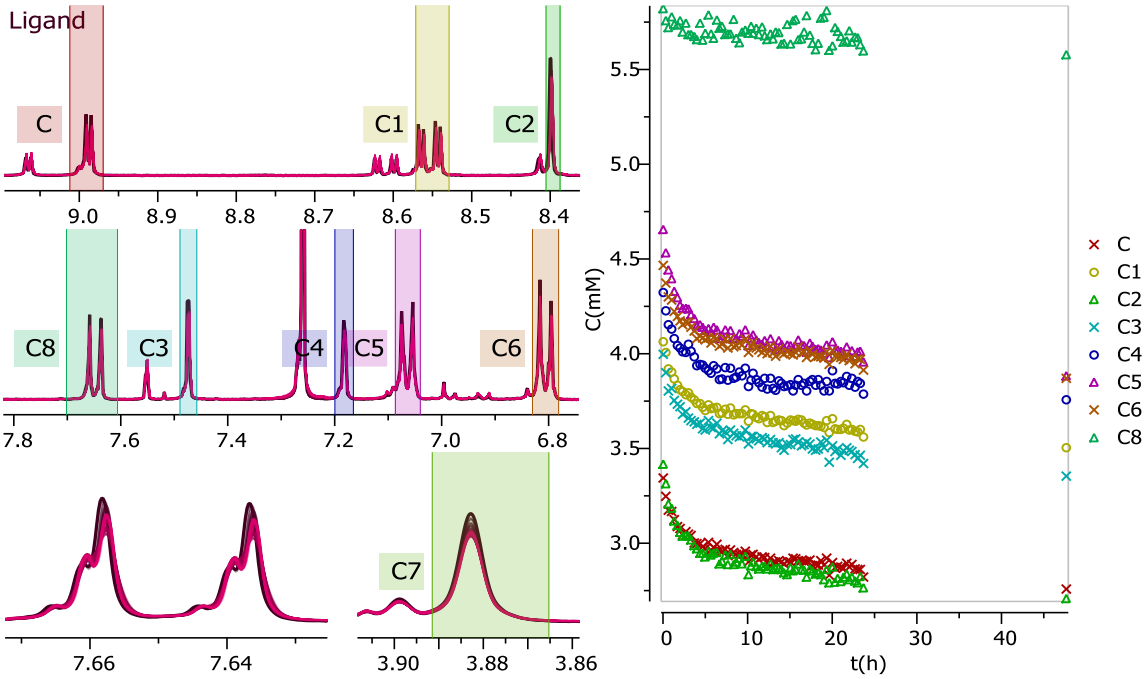


Figure 5.2-13 Time dependent 1H-NMR in chloroform-d, showing the ligand peaks (left) and their integrations vs time (right). DNB ligand (ca. 4 mM) with mercaptoethanol (10 mM) at 298K. Concentration is set to the integration of a mercaptoethanol CH2 peak at 10 mM. The region at 7.65 ppm is two or more peaks overlapped resulting in a relatively unchanging integration, see arrow.

Considering the ligand peaks these are similar to the assigned ligand in DMSO. (Figure 8.1-2) All peaks follow the same exponential decay regardless of which part of the ligand. The decrease over 48 hours, is in line with the mercaptoethanol, with a decrease of approximately 0.5 mM. At 7.65 ppm is a set of peaks resulting from a decreasing DNB peak, and a growing DNB peak from the product. When integrated the two changes are cancelling each other out, and result in an overall relatively stable integration. Though after 48 hours, there does appear to have been a slight decrease of approx. 1-2 mM. This data and that from the mercaptoethanol shows that the two species are reacting at a similar rate, likely with one another. (Figure 5.2-13)

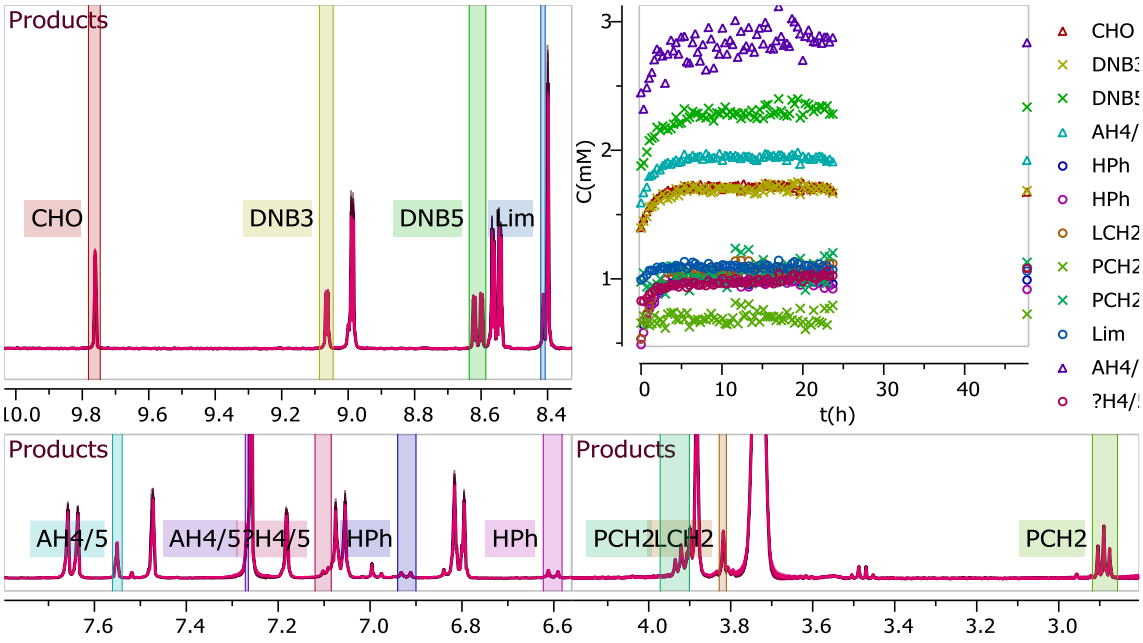


Figure 5.2-14 Time dependent 1H-NMR in chloroform-d, showing the product peaks (left and bottom) and their integrations vs time (top right). DNB ligand (ca. 4 mM) with mercaptoethanol (10 mM) at 298K. Concentration is set to the integration of a mercaptoethanol CH2 peak at 10 mM. Products are represented by crosses (DNB-SEtOH), circles (2-imidazole ligand), triangles (aldehyde).

It is obvious that, like the mercaptoethanol the initial formation of the products has occurred rapidly before the first measurement, as none begin at 0 mM. As mentioned in the paragraph above, one of the DNB-SEtOH DNB peaks is obscured by a DNB-ligand DNB peak, so is not included here. It and the rest are assigned with the aid of COSY and are similar to the spectrum in acetonitrile-d3. Other DNB-SEtOH peaks are the CH2s from the 2-(thio)ethan-1-ol group, which appear near to their respective free mercaptoethanol peaks. Another product of the reaction is the decapped 2-imidazole ligand which appears as an imine, a CH2 and two phenyl peaks, as well as an ambiguous H4/5 peak that overlaps with a DNB-ligand phenyl peak. Oddly another product forms as three coupled peaks, an aldehyde (CHO), with two imidazole (AH4/5) peaks. The asymmetry of the imidazole suggests something could be bound to one N, but whether it is just protonated is uncertain. That a ligand imine is being hydrolyzed, whether it is

capped or uncapped is interesting, though a large water peak at 1.5 ppm could indicate the cause. (Figure 5.2-14)

That this is not present in the rotaxane samples above suggests that the CB[10] is still able to protect the imine after decapping, pointing to formation of decapped rotaxanes tending toward pseudorotaxanes. (Sections 5.2.2.1.1 and 5.2.2.1.2)

5.2.2.1.4 1:20 DNB Cylinder to Mercaptoethanol NMR Experiment

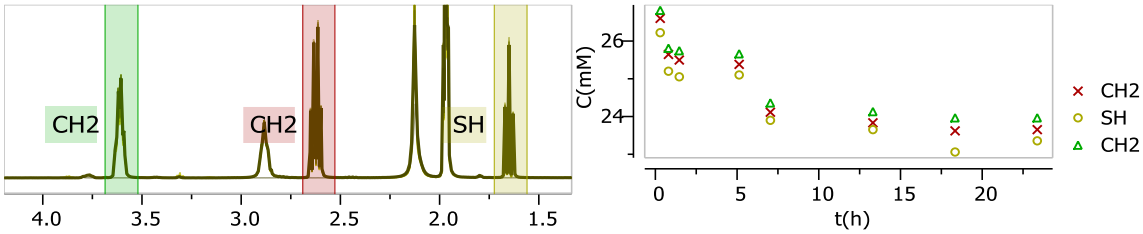


Figure 5.2-15 1H-NMR spectra of 1:20 cylinder (1.33 mM) to mercaptoethanol (26.6 mM) in acetonitrile-D3. Showing the decrease of the mercaptoethanol peaks. Concentration is set using the initial T=0 concentration of mercapteothanol of 26.6 mM at peak CH2 2.6 ppm. Concentrations are adjusted according to the number of nuclei involved in a peak.

Reacting the DNB cylinder (1.33 mM) with mercaptoethanol (26.6 mM) at a ratio of 1:20 or 1:3.3 mercaptoethanol to DNB group, shows a decrease in the mercaptoethanol peaks. The greater concentration of mercaptoethanol increases the signal to noise ratio, compared to the rotaxane data above. Compared to the mercaptoethanol, the water peak is also lesser in this sample reducing the effect of detector saturation. Again at a higher mercaptoethanol concentration, the OH peak is absent. (Figure 5.2-15)

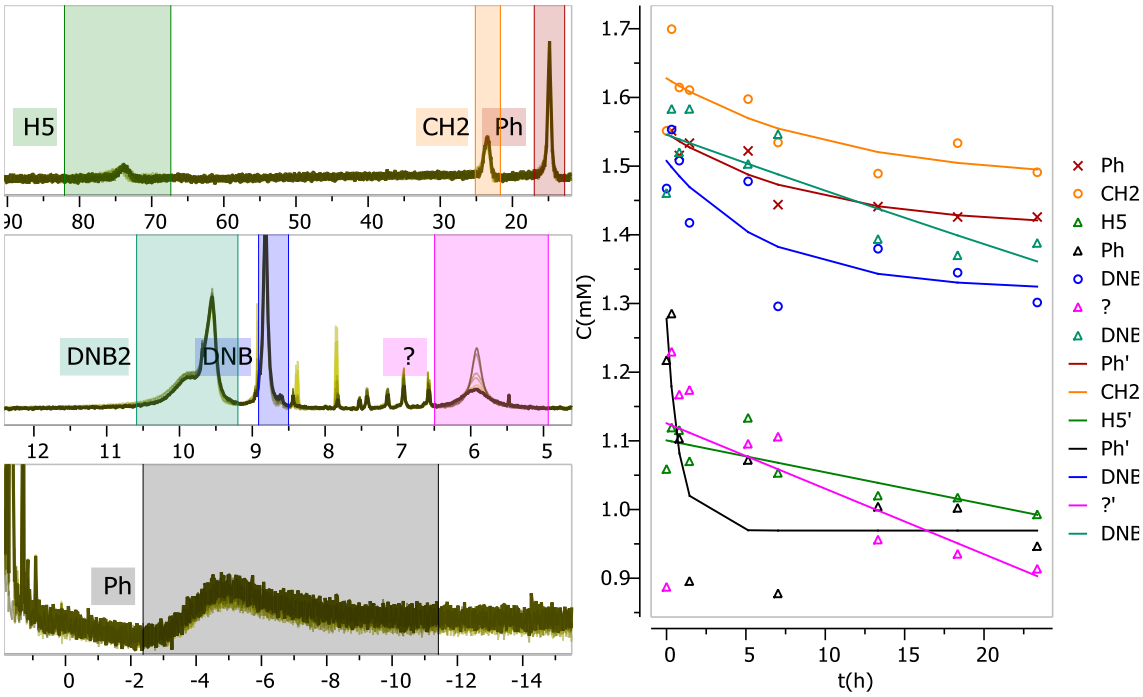


Figure 5.2-16 1H-NMR spectra of 1:20 cylinder (1.33 mM) to mercaptoethanol (26.6 mM) in acetonitrile-D3. Showing the decrease of the cylinder peaks (DNB2 = DNB-2H, 12H; DNB = DNB-H, 6H; DNP = DNP-H, 4H; CH2 = CH_2, 2H; H5 = imidazole H5, 6H; Ph = HPha/b, 12H). Concentration is set using the initial T=0 concentration of mercaptoethanol of 26.6 mM at peak CH2 2.6 ppm. (Figure 5.2-15) Concentrations are adjusted according to the number of nuclei involved in a peak.

The cylinder DNB peaks are simpler to track versus the rotaxane as they are less broad. All cylinder peaks appear to undergo a steady decrease over 24 hours. This is similar to what is seen for the rotaxane experiments. It is likely this is a result of peak broadening caused by a shift to a range of capped populations rather than a homogenous 6/6 capped sample. The peak at approx. -5 ppm for the phenyl a/b is particularly broadened (Figure 5.2-16) Also present is a peak at 6 ppm, which interestingly changes by broadening steadily over the experiment. However, in the measurement before the addition of mercaptoethanol the peak is broader and has a lower integration than the final measurement at T = 24 hours. This behavior is more akin to an exchangeable proton. Within 1 ppm of what is expected for an ammonium peak, and within the range

for an aniline or amide these could be possible. Given the sample is prepared using ammonium hexafluorophosphate this could be the source, though the characteristic 52.4 Hz $^{1}J_{13N\text{-}1H}$ coupling is not present this could be obscured by paramagnetic broadening.

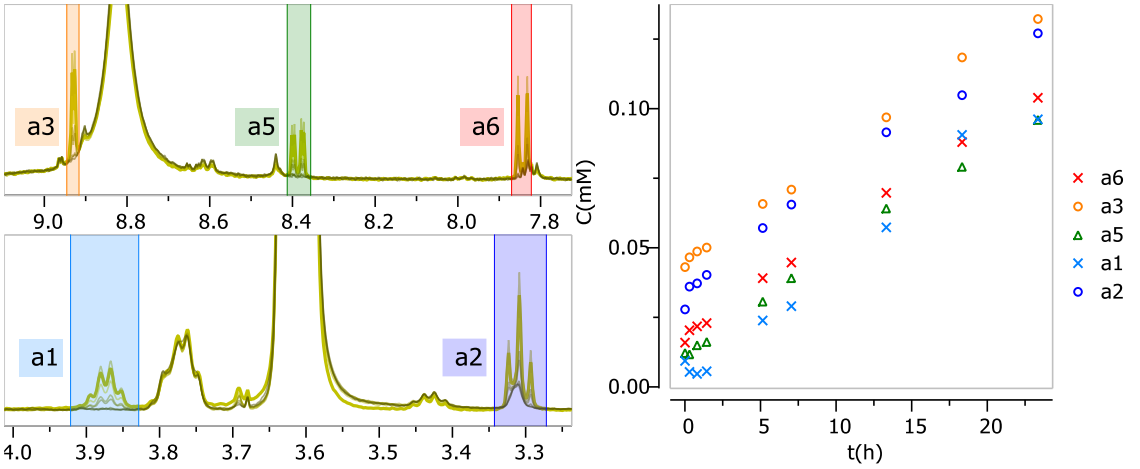


Figure 5.2-17 1H-NMR spectra of 1:20 cylinder (1.33 mM) to mercaptoethanol (26.6 mM) in acetonitrile-D3. Showing the increase of the DNB-SEtOH peaks (a1 = CH2, a2 = CH2, a3 = DNBH3, a5 = DNBH5, a6 = DNBH6). Concentration is set using the initial T=0 concentration of mercaptoethanol of 26.6 mM at peak CH2 2.6 ppm. (Figure 5.2-15) Concentrations are adjusted according to the number of nuclei involved in a peak.

The DNB-SEtOH is not the only diamagnetic species present in this experiment, but by the end of the reaction it is the strongest. The DNB-SEtOH is produced at a steady rate and does not seem to be plateauing, suggesting this reaction will continue for more time if allowed. (Figure 5.2-17) The other species present appear to be a slight DNB-ligand impurity slowly hydrolyzing into methylene dianiline and the capped aldehyde.

5.2.2.1.5 NMR Decapping Conclusions

The presence of a water peak in all the NMR data acquired suggests its origin is the mercaptoethanol used. The ligand when exposed to water readily hydrolyses, though this does not seem to occur in the cylinder, and certainly does not occur in the rotaxane.

When exposed to mercaptoethanol all derivatives with the DNB readily react to form the DNB-SEtOH product pointing to a SNAr mechanism of reaction. However, there was no clear formation of a pseudorotaxane or imidazole cylinder, so this needs to be explored further. It is worth considering that whilst the intensities for the rotaxane and capped cylinder peaks decreased there was no precipitate observed as would be expected for CB[10], so a new CB[10] · Guest species has formed.

5.2.3 Releasing an Imidazole Cylinder

Understanding from mass spectrometry, that it is possible to remove the caps from a rotaxane and that from a population perspective it appears to occur in the form of a shifting distribution eventually releasing a fully decapped 2-imidazole cylinder. From the NMR the main diamagnetic product of the reaction is 2-((2,4-dinitrophenyl)thio)ethan-1-ol, suggesting a SNAr mechanism and unobservable rates of other reactions. Using UV-Vis spectroscopy this section will further explore the decapping process and understand whether the rotaxane is decapped to release an intact imidazole cylinder.

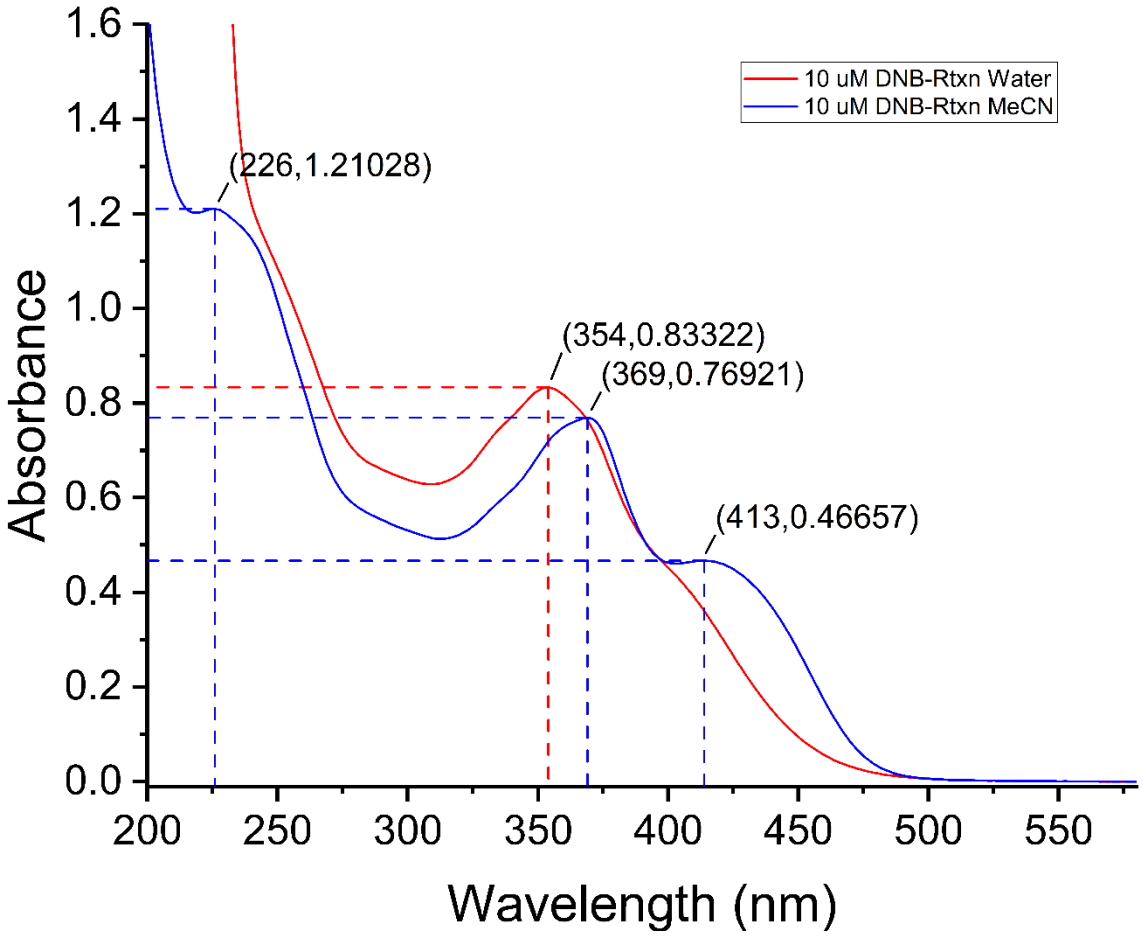


Figure 5.2-18 UV-Vis spectra of 10 uM DNB rotaxane in water and acetonitrile at 37 °C. Showing the location of the peaks.

In water the peak, which would appear around 413 nm in acetonitrile, is broadened and potentially blue shifted making it harder to interpret. The strong peak at 369 nm in acetonitrile is blue shifted to 354 nm in water, a possible cause of the blueshift could be polar solvent effect indicating a CT transition between a polar ground state and an excited state. (Figure 5.2-18)

5.2.3.1 Acetonitrile

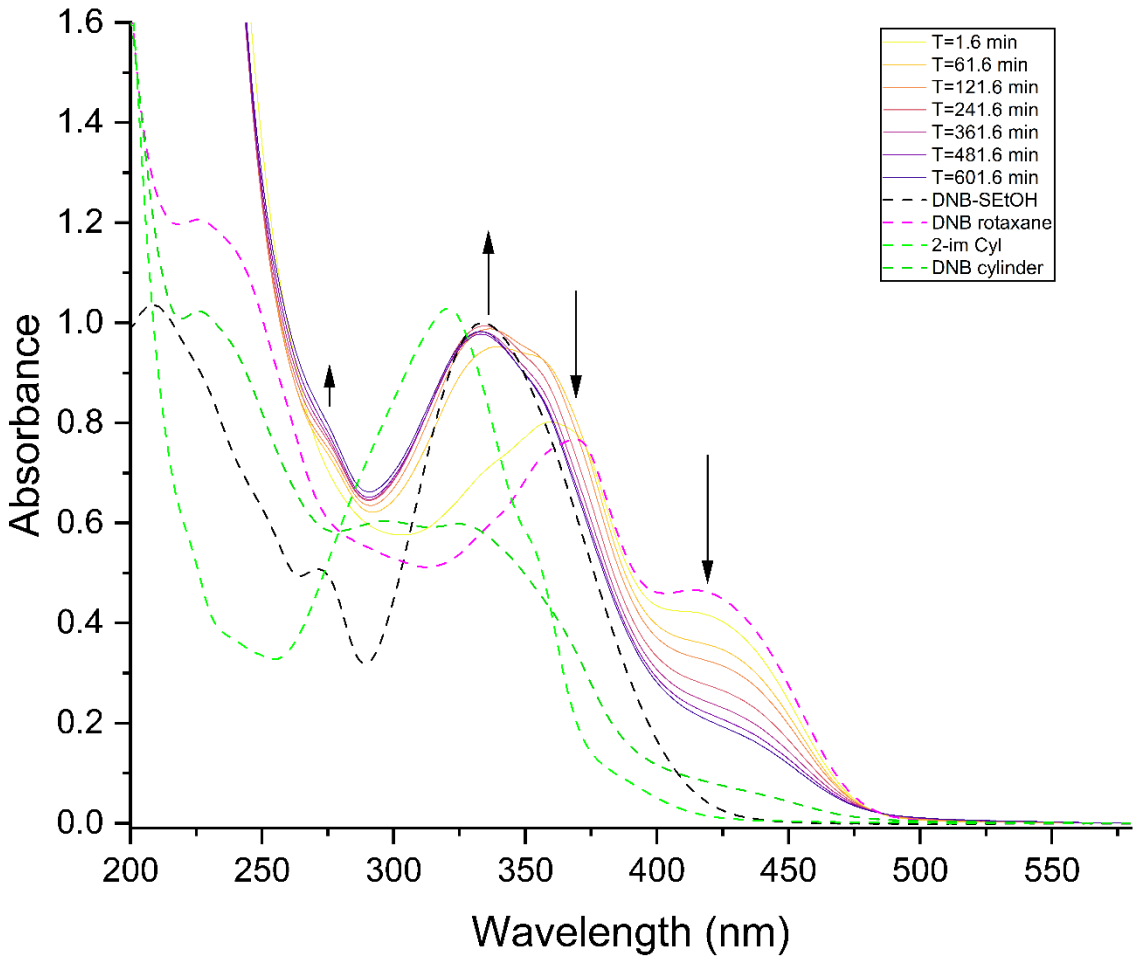


Figure 5.2-19 UV-Vis spectra in MeCN 37 °C. Solid lines. Rotaxane at 10 uM with 10mM mercaptoethanol in MeCN taken at different time points over 10 hours. Dashed lines. 2-((2,4-dinitrophenyl)thio)ethan-1-ol in MeCN, DNB rotaxane at 10 uM in MeCN, 2-imidazole cylinder at 10 uM in MeCN, DNB cylinder at 10 uM in MeCN. Full dataseries at (

Figure 8.2-5)

As a result of the instability of the threaded pseudo-rotaxane in organic solvents, an alternative comparison, the 2-imidazole cylinder, was used. A reaction of DNB rotaxane at 10 uM with mercaptoethanol in acetonitrile produces a UV-Vis spectra that transforms over 10 hours, to be attributable to a solution of DNB-SEtOH, 2-imidazole cylinder pseudorotaxane, and a decreased quantity of DNB rotaxane. A decrease in peak intensity at 369 nm and 413 nm represents consumption of DNB rotaxane, with an

increase in peak intensity at 333 nm and 271 nm. (Figure 5.2-19) This is more apparent when looking closely at the peak forming around 333 nm. (Figure 5.2-20)

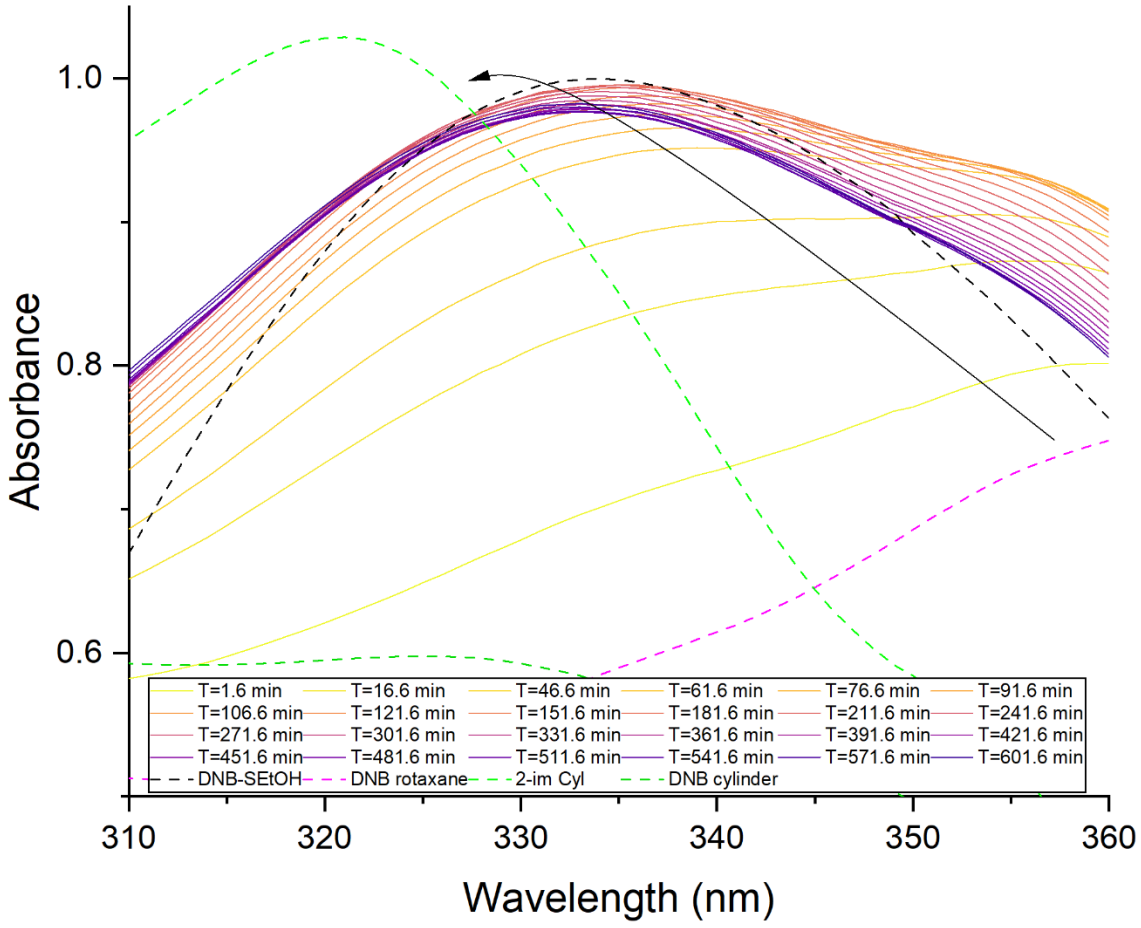


Figure 5.2-20 UV-Vis spectra in MeCN at 37 °C showing the growth and shifting of a peak, as the rotaxane is consumed and 2-imidazole cylinder pseudorotaxane and 2-((2,4-dinitrophenyl)thio)ethan-1-ol are producted. Rotaxane at 10 uM with 10mM mercaptoethanol taken at different time points over 10 hours.

The rotaxane peak at 369 nm shows a shift toward 333 nm, resulting from the growth of 2-imidazole cylinder, and 2-((2,4-dinitrophenyl)thio)ethan-1-ol peaks and a decrease in the rotaxane peak at 369 nm that becomes more apparent toward later time points. It is likely that a varied population of partially decapped rotaxane/cylinder would absorb broadly across 300-350 nm.

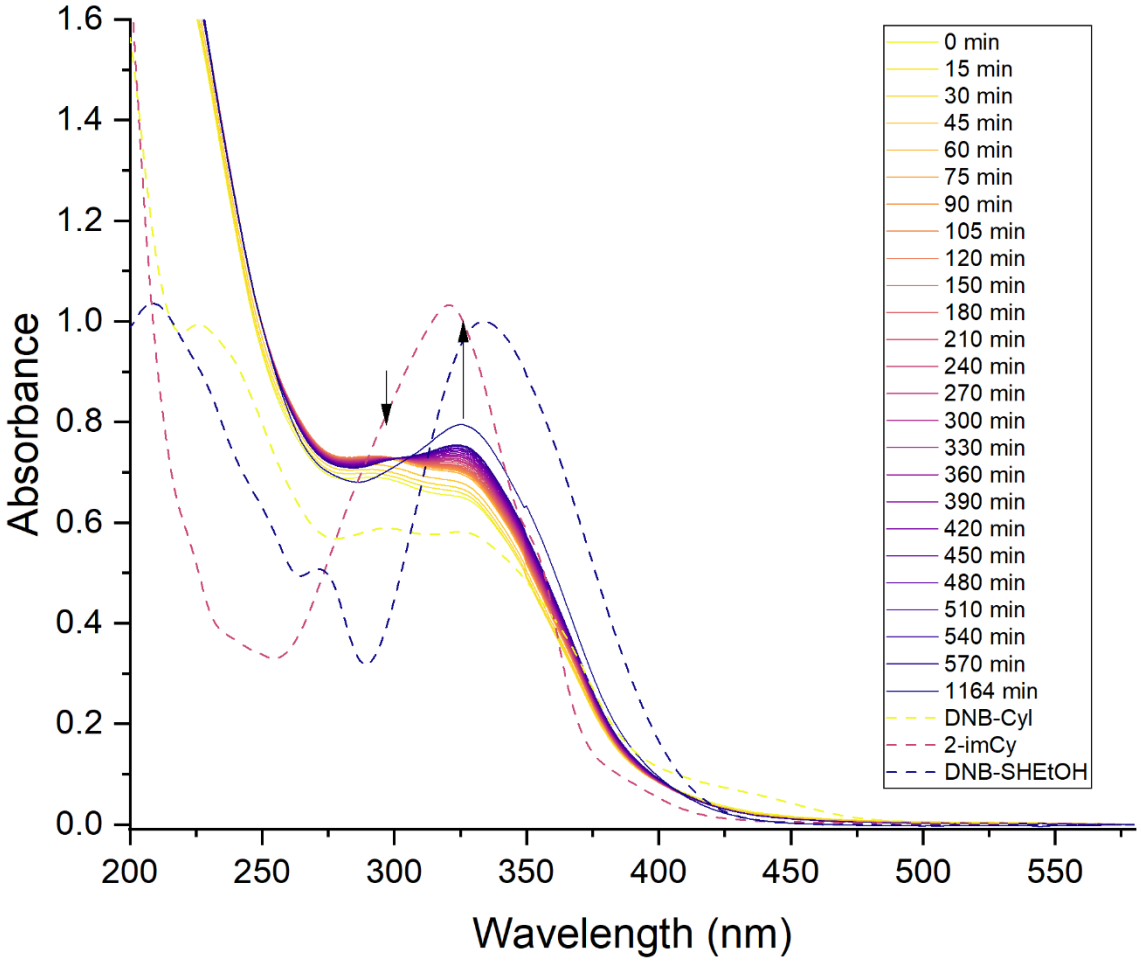


Figure 5.2-21 UV-Vis spectra in MeCN. Solid lines. DNB cylinder at 10 uM with 10 mM mercaptoethanol in MeCN, taken at different time points. Dashed lines. 2-((2,4-dinitrophenyl)thio)ethan-1-ol in MeCN, DNB cylinder at 10 uM in MeCN, 2-imidazole cylinder at 10 uM in MeCN.

The same reaction was performed using 10 uM DNB cylinder with 10 mM mercaptoethanol, and demonstrates a similar decrease in DNB cylinder, with a subsequent increase in a peak representative of a combination of 2-imidazole cylinder and 2-((2,4-dinitrophenyl)thio)ethan-1-ol. (Figure 5.2-21)

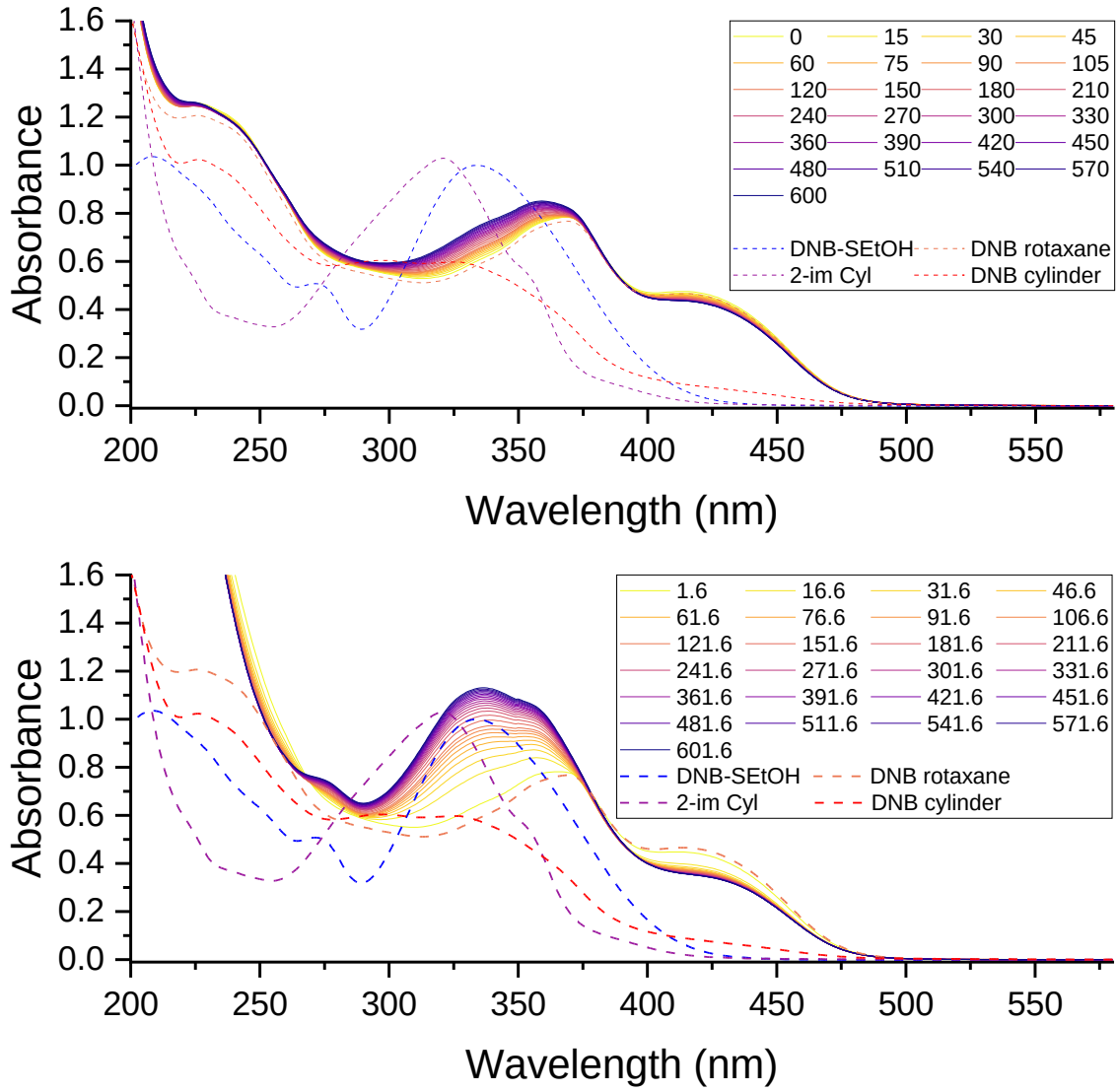


Figure 5.2-22 Top. UV-Vis spectra in MeCN at 37 °C, DNB rotaxane at 10 uM with 60 uM mercaptoethanol taken at different time points. Bottom. UV-Vis spectra in MeCN at 37 °C, DNB rotaxane at 10 uM with 5 mM mercaptoethanol taken at different time points. Top and bottom. Dashed lines. 2-((2,4-dinitrophenyl)thio)ethan-1-ol, DNB cylinder 10 uM, DNB rotaxane 10 uM, 2-imidazole cylinder 10 uM.

At 60 uM mercaptoethanol, the process is expectedly slower, the lesser mercaptoethanol absorbance shows a bit more detail below 300nm. At 5 mM the balance between mercaptoethanol absorption and reaction rate, better shows the formation of the small peak at 271 nm from DNB-SEtOH. (Figure 5.2-22)

5.2.3.2 Water

In water, the pseudorotaxane freely forms at room temperature, so is stable for UV-Vis measurements. This is a better comparison for a decapped product complex.

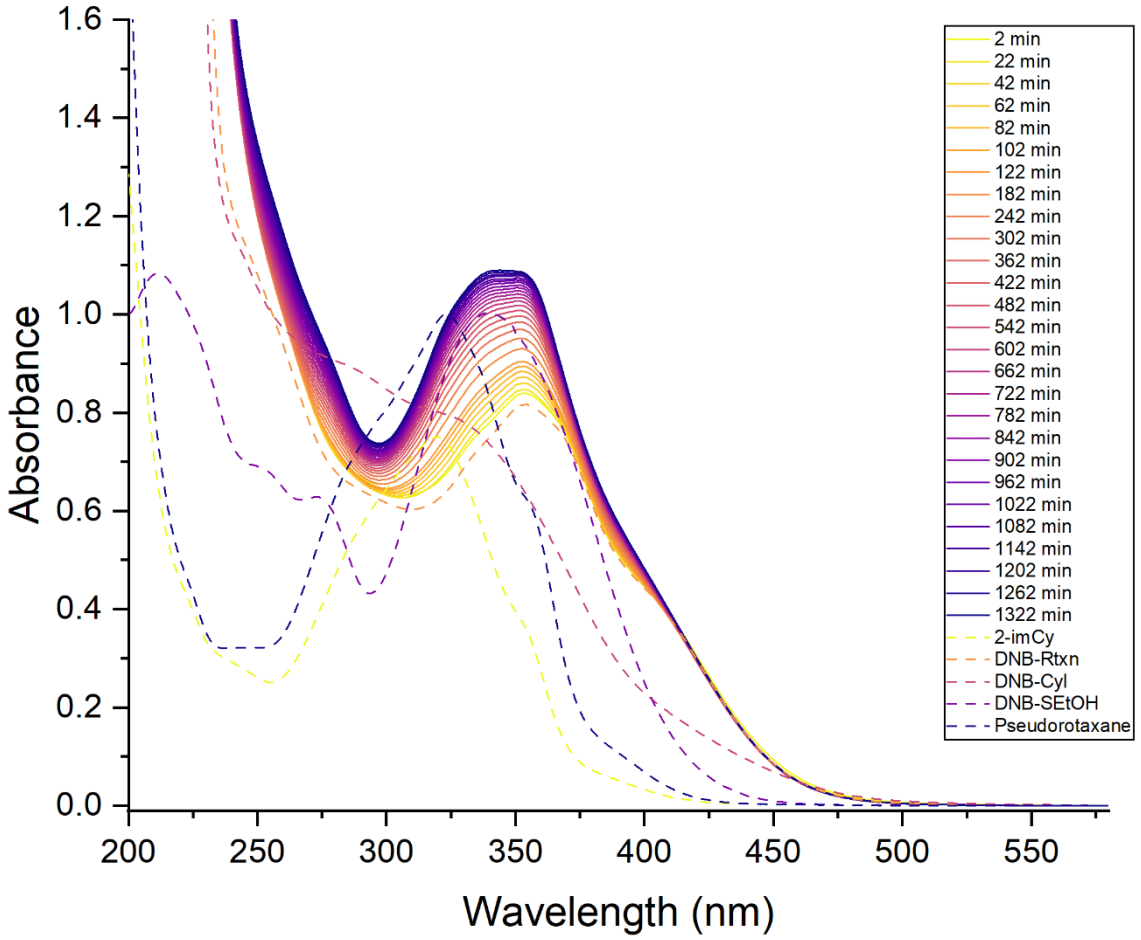


Figure 5.2-23 UV-Vis spectra in H2O. Solid lines. DNB rotaxane at 10 uM with 10 mM mercaptoethanol, taken at different time points. Dashed lines. 2-((2,4-dinitrophenyl)thio)ethan-1-ol, DNB cylinder 10 uM, DNB rotaxane 10 uM, 2-imidazole cylinder 10 uM.

Repeating the experiment in water, the shifted/broadened DNB rotaxane spectra makes interpretation of the consumption of rotaxane less certain. However, there is a similar increase in peaks in the frequency expected for increased concentrations of 2-imidazole pseudorotaxane and DNB-SEtOH in solution. (Figure 5.2-23)

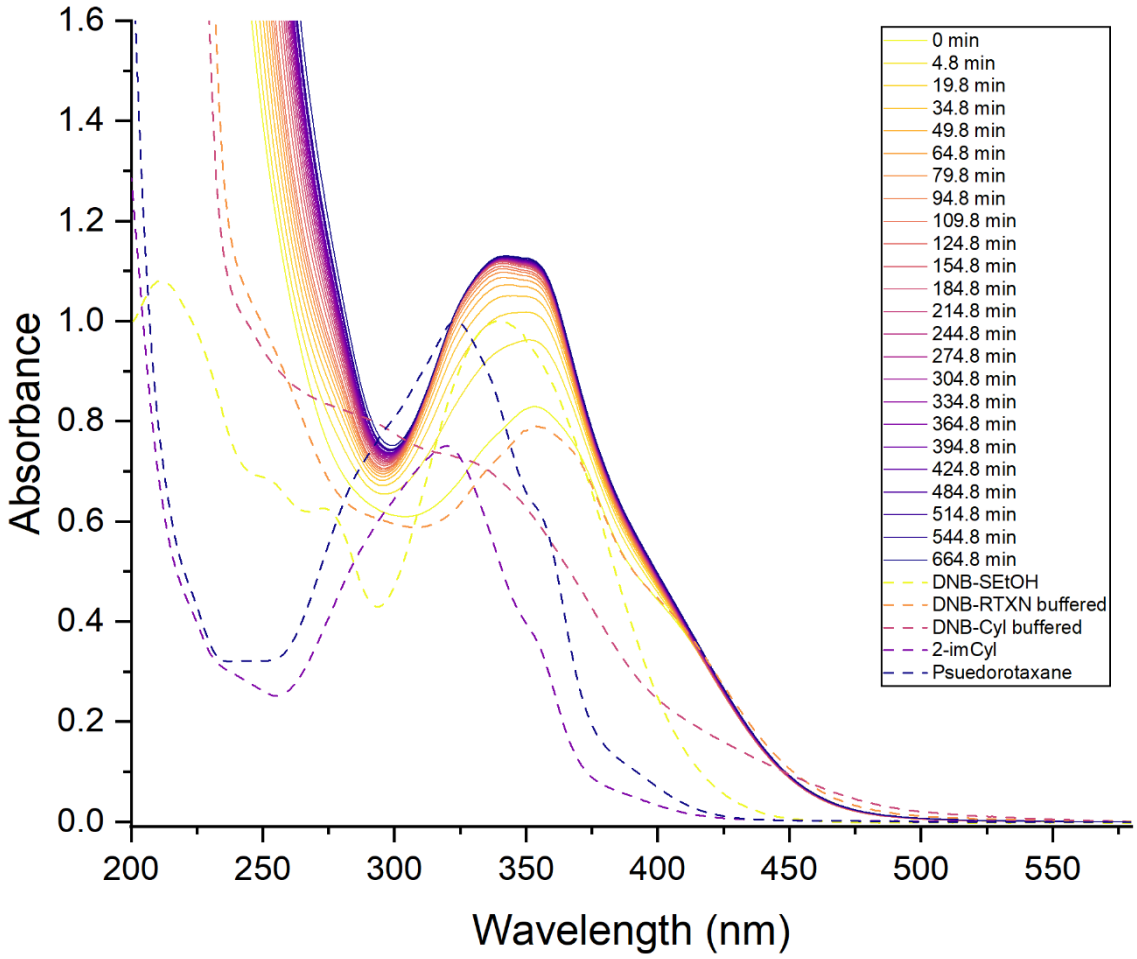


Figure 5.2-24 UV-Vis spectra in H2O. Solid lines. DNB rotaxane at 10 uM with 10 mM mercaptoethanol and 1 mM sodium cacodylate, taken at different time points. Dashed lines. 2-((2,4-dinitrophenyl)thio)ethan-1-ol, DNB cylinder 10 uM, DNB rotaxane 10 uM, 2-imidazole cylinder 10 uM.

Running the experiment in buffered water using 1 mM sodium cacodylate showed little difference in outcome, though an apparent slightly faster rate of reaction. (Figure 5.2-24) Perhaps buffering the acidity of the mercaptoethanol had an effect of increasing the rate of nucleophilic attack. In line with affecting the mercaptoethanol, the absorption attributed to the mercaptoethanol appears to have either red shifted or has intensified. Though this can not be confirmed as the signal saturated the detector at lower wavelengths.

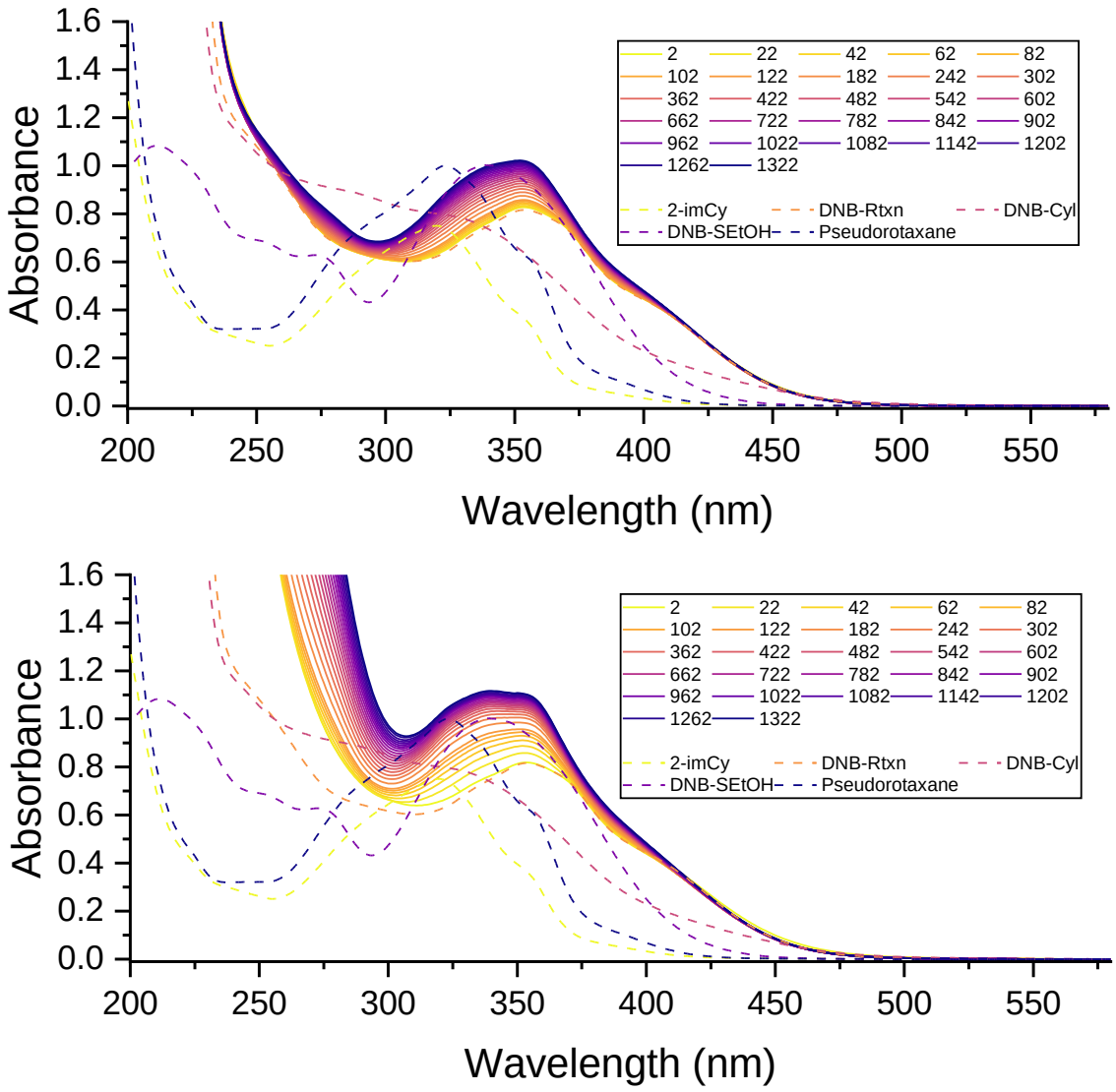


Figure 5.2-25 Top. UV-Vis spectra in H_2O at 37 °C, DNB rotaxane at 10 uM with 5 mM mercaptoethanol taken at different time points. Bottom. UV-Vis spectra in H_2O at 37 °C, DNB rotaxane at 10 uM with 100 mM mercaptoethanol taken at different time points. Top and bottom. Dashed lines. 2-((2,4-dinitrophenyl)thio)ethan-1-ol, DNB cylinder 10 uM, DNB rotaxane 10 uM, 2-imidazole cylinder 10 uM, 2-imidazole cylinder pseudorotaxane.

Changing the concentration of mercaptoethanol to 5 mM reduces the overlap with the mercaptoethanol peak, giving better visibility of the small peak around 273 nm from the DNB-SEtOH, and shows the same though slower growth of the dominant DNB-SEtOH and pseudorotaxane peaks. In contrast, changing the concentration to 100 mM has the

effect of obscuring the smaller DNB-SEtOH peak but increasing the rate of growth observed in the dominant peaks. (Figure 5.2-25)

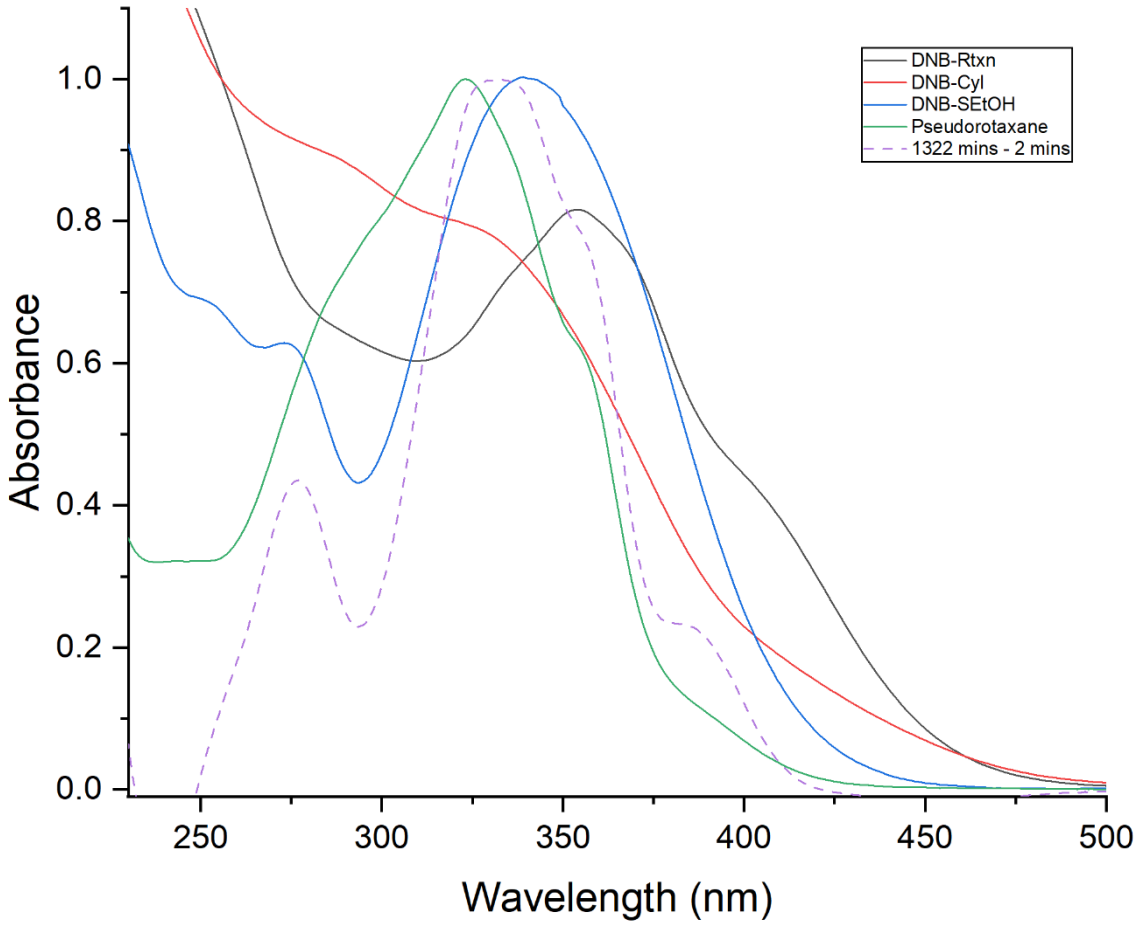


Figure 5.2-26 UV-Vis spectra in H_2O at 37 °C. Dashed. DNB rotaxane at 10 uM with 5 mM mercaptoethanol time point 1322 mins minus time point 2 mins then adjusted so the peak maximum equals 1 for comparison. Solid. 2-((2,4-dinitrophenyl)thio)ethan-1-ol, DNB cylinder 10 uM, DNB rotaxane 10 uM, 2-imidazole cylinder pseudorotaxane.

By subtracting initial time points from later ones, it is simpler to visualize the changes. (Figure 5.2-26) It shows a large increase between the DNB-SEtOH and the pseudorotaxane, importantly this highlights that the forming peak is a combination of the two compounds. There is also a well-defined peak at 273 nm showing clearly the change mentioned before that identifies the DNB-SEtOH.

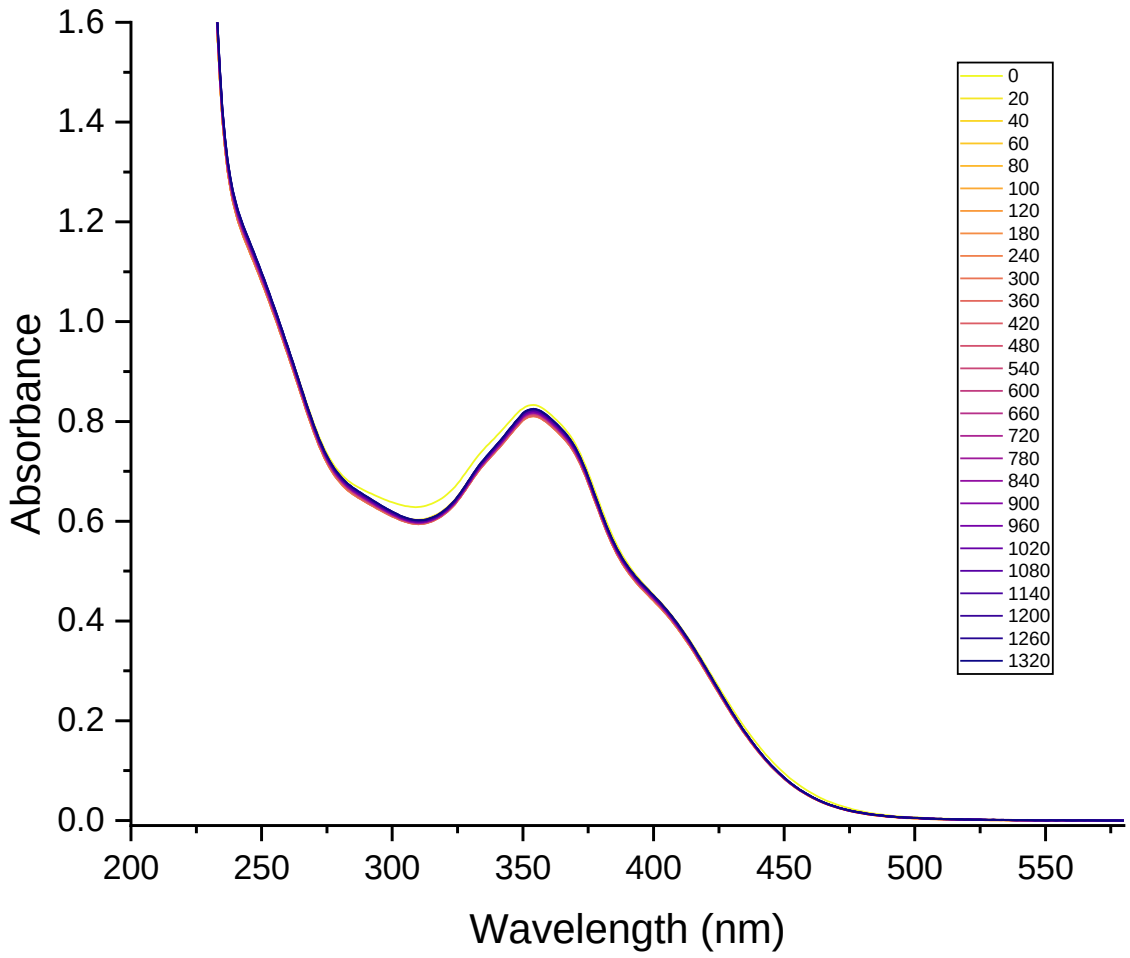


Figure 5.2-27 UV-Vis spectra in H_2O at 37 °C. DNB rotaxane at 10 uM at various time points.

Comparing the spectra of the rotaxane in water it is apparent that the decapping is entirely a result of the thiol in solution. After 22 hours of incubation at 37 °C in water, the appearance of the rotaxane UV-Vis spectra is unchanged. (Figure 5.2-27) The slight difference between 0 and 20 minutes is caused by the acclimatization of the cuvette to the temperature of the Peltier temperature controller.

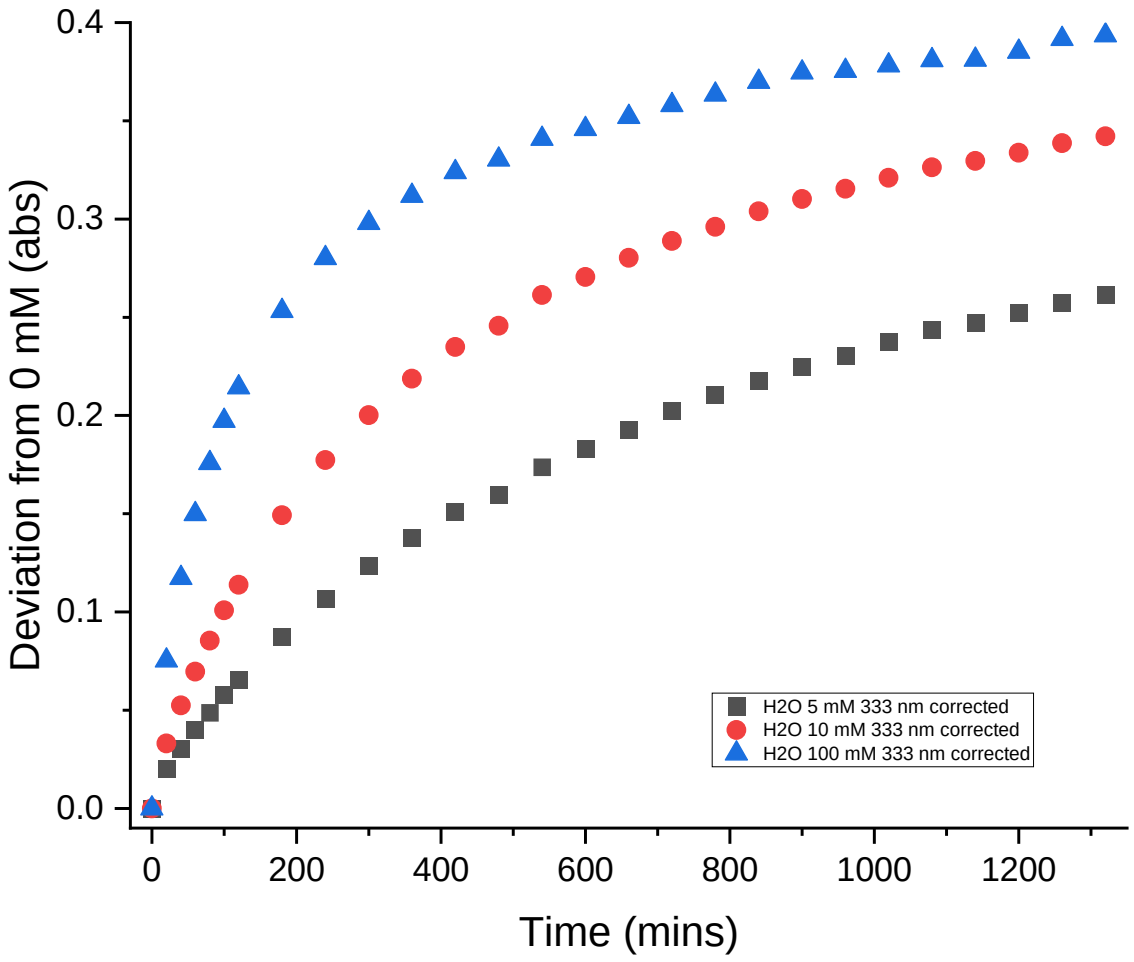


Figure 5.2-28 UV-Vis absorbance values at 333 nm for 10 uM rotaxane with 5 mM, 10 mM, and 100 mM mercaptoethanol subtracted from 0 mM data values at their corresponding time point to yield a deviation from the behavior expected for an aqueous solution. Values are finally zeroed to the initial time point to ignore deviation in time between mixing and the first data aquisition.

Plotting the evolution over time of the absorbance at 333 nm as a deviation from the value at 0 mM, shows a clear dependence on concentration. Though the difference between 5 mM and 10 mM seems to be the same as 10 mM and 100 mM, showing a non-linear dependence on mercaptoethanol concentration. That the increase seems to be lesser at higher concentration suggests a fractional reaction order with respect to mercaptoethanol. This is unsurprising and is evidence that there are multiple steps as expected for the SNAr mechanism. (Section 5.2.2) The slight deviation of the plot at 100

mM shows the presence of more than one contributing peak, with one likely shifting causing the effect. (Figure 5.2-28)

5.2.3.3 Stability of the 2-imidazole Cylinder.

If the rotaxane is to be developed for use in biological systems, it is important to assess the stability of the 2-imidazole cylinder once released. This was best done using UV-Vis spectroscopy owing to the paramagnetic nature of these nickel complexes.

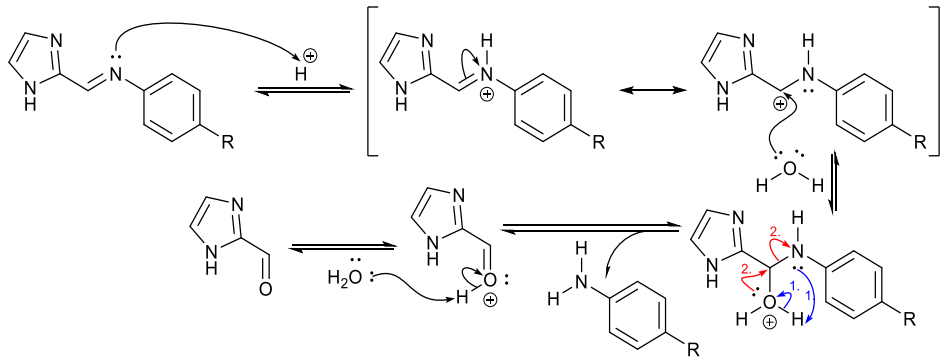


Figure 5.2-29 Generic immine hydrolysis mechanism.

A form of degradation observed with these complexes and ligands is hydrolysis of the immine. (Figure 5.2-29) Hydrolysis prevents use of normal-phase silica gel chromatography and most forms of HPLC to purify these immine based ligands and complexes.

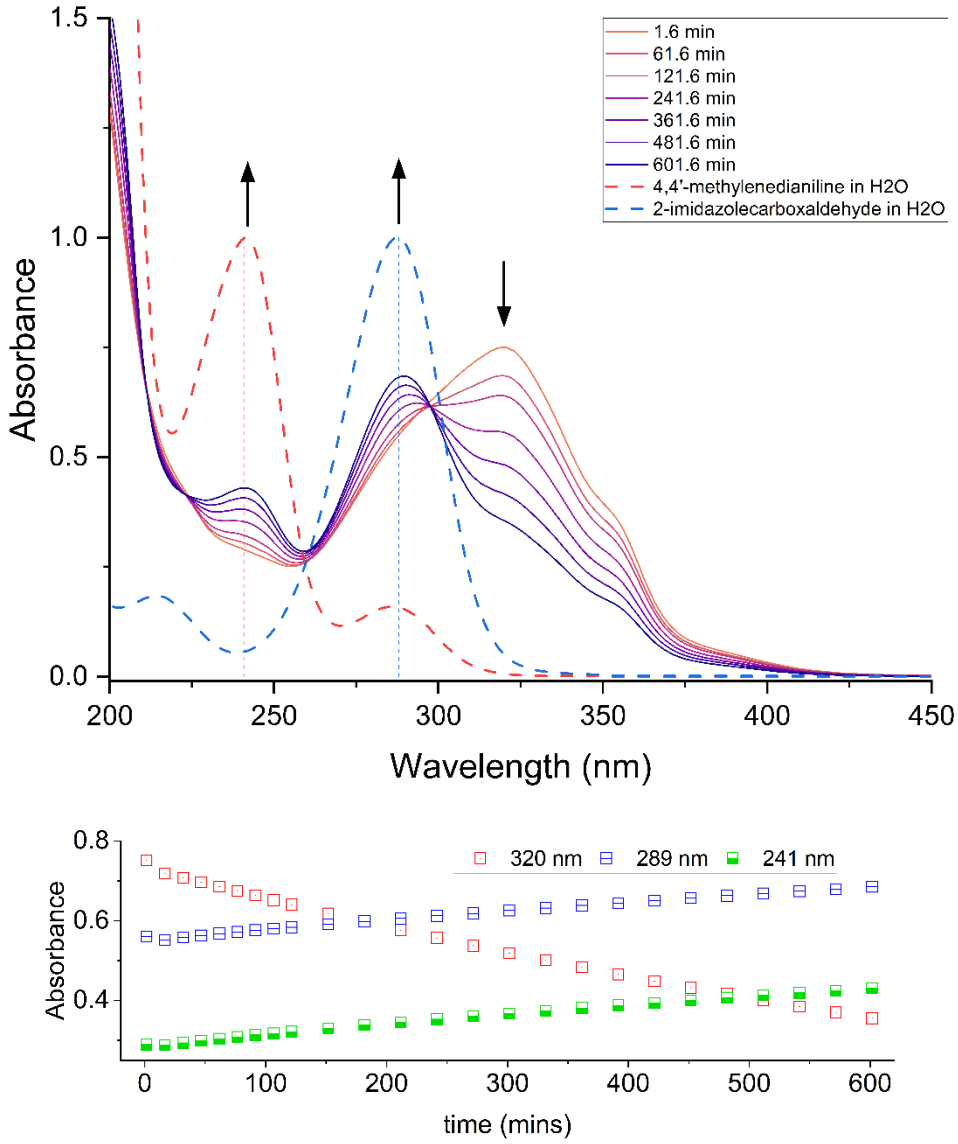


Figure 5.2-30 Top. UV-Vis spectra taken at 37 °C. Solid lines: 2-imidazole cylinder 10 uM in H2O at various time points. Dashed lines: Likely products of the 2-imidazole cylinder hydrolysis, 4,4'-methylenedianiline and 2-imidazolecarboxaldehyde in H2O. Bottom. Increase or decrease of absorbance at 320, 289, and 241 nm Vs time.

On incubating the 2-imidazole cylinder in water, the cylinder peak at 320 nm decreases proportionally to the increase at 241 nm and 289 nm for the 4,4'-methylene dianiline and 2-imidazolecarboxaldehyde respectively, creating an isosbestic point at 297 nm. (Figure 5.2-30) From this, it is apparent the 2-imidazole cylinder is vulnerable to hydrolysis in water.

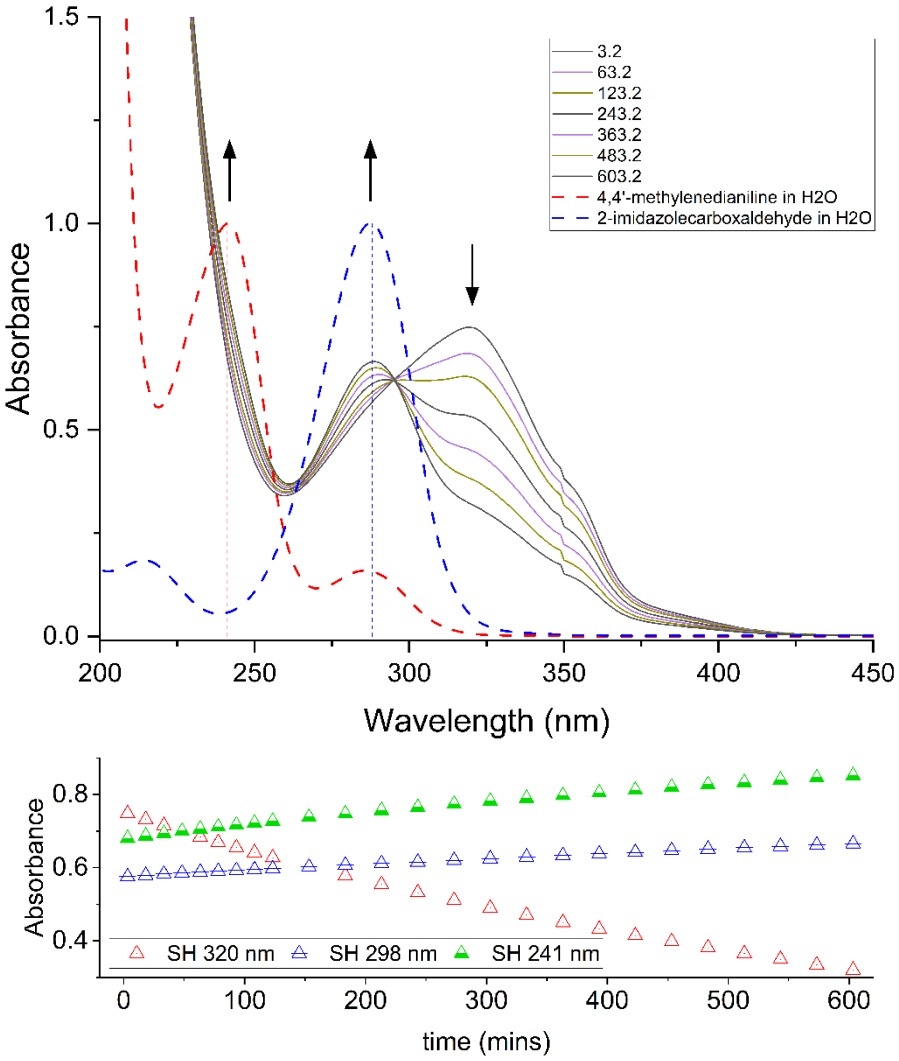


Figure 5.2-31 Top. UV-Vis spectra taken at 37 °C. Solid lines: 2-imidazole cylinder 10 uM with mercaptoethanol 10 mM in H_2O at various time points. Dashed lines: Likely products of the 2-imidazole cylinder hydrolysis, 4,4'-methylenedianiline and 2-imidazolecarboxaldehyde in H_2O. Bottom. Increase or decrease of absorbance at 320, 289, and 241 nm Vs time.

Mercaptoethanol absorbs below about 230 nm, and at 10 mM it gives a strong peak. This obscures the 4,4'-methylenedianiline absorption, however because the concentration of thiol remains the same throughout the experiment it is still possible to measure the change in intensity at 241 nm to a reasonable degree of accuracy. (Figure 5.2-31) Looking at the two graphs, it is apparent that addition of mercaptoethanol has little effect on the 2-imidazole cylinder.

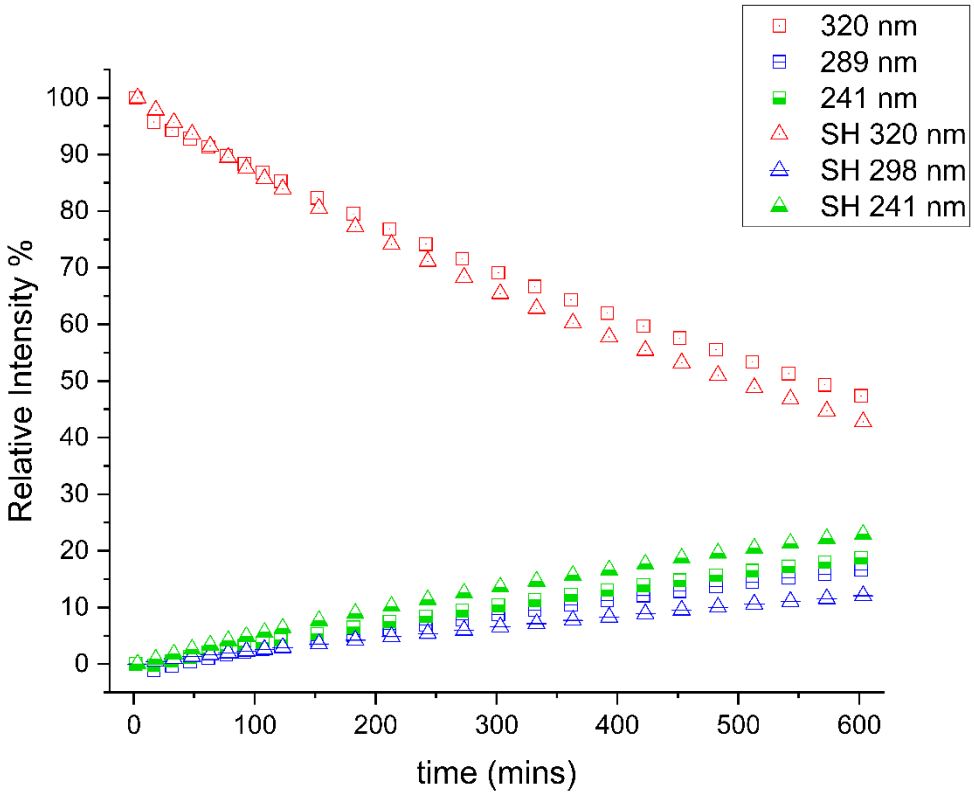


Figure 5.2-32 Intensity of an increasing or decreasing peak as a proportion of percentage, relative to the decrease of the 2-imidaozle cylinder peak at 320 nm in their respestive spectra. Increases adjusted to begin from 0.

Plotting the values of absorption as a function of their intensity relative to the intensity of the 2-imidazole cylinder at 320 nm, shows little difference between the two experiments. The data was prepared in a way to show the decrease of the 2-imidazole cylinder from its initial 100%, in comparison to the increase of the two organic products adjusted so the increase in percentage is proportional in scale to the increase in absorbance relative to the 2-imidazole cylinder. Looking at the decrease of the 2-imidazole cylinders, there is a slight increase in the rate of hydrolysis in the presence of 10 mM mercaptoethanol. Perhaps this is a result of the mercaptoethanol acting as a weak acid in solution, furthering the rate the immine is protonated. (Figure 5.2-32)

Acquiring DNB-cylinder decapping data in water was complicated by its hydrolysis through the same mechanism as seen for the 2-imidazole cylinder. (Figure 8.2-6)

5.2.3.3.1 Stability in MeCN

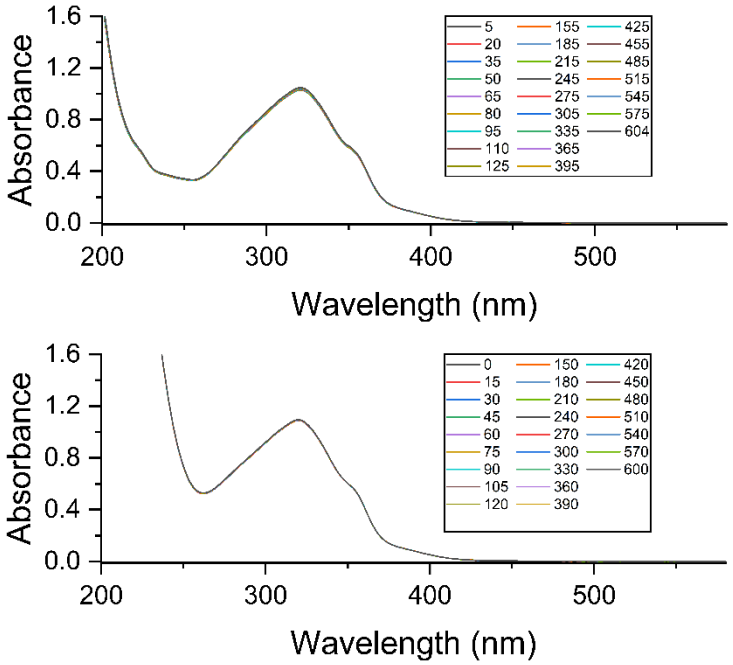


Figure 5.2-33 Top. 10 uM 2-imidazole cylinder in MeCN 37 °C. Bottom. 10 uM 2-imidazole cylinder with 10 mM mercaptoethanol in MeCN 37 °C.

In acetonitrile there is little to no change in the UV-Vis spectrum when incubated with and without mercaptoethanol. (Figure 5.2-33) This has implications that it might be easier to assess the binding ability of the released cylinder if it is released in acetonitrile.

5.2.3.3.2 Stability of the Pseudorotaxane

To achieve concentrations strong enough to make stock solutions, the 2-imidazole cylinder is assisted into water by dissolving in methanol first. The solution is then mixed into a suspension of CB[10] which is subsequently dissolved as it forms an aqueous solution of pseudo-rotaxane.

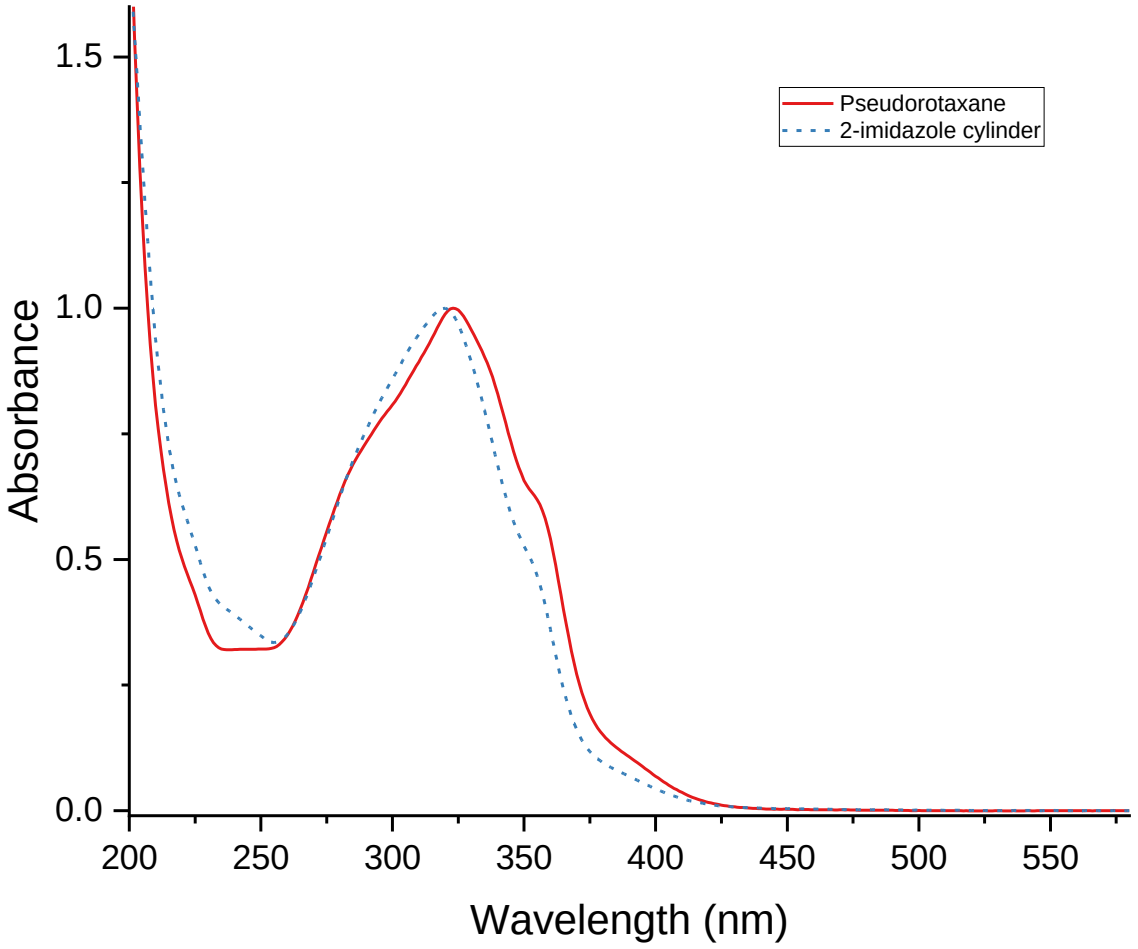


Figure 5.2-34 UV-Vis spectrum in water taken at 37 °C of the 2-imidazole cylinder (dashed blue) and its CB[10] pseudorotaxane (solid red).

Comparing the cylinder and pseudorotaxane, the UV-Vis spectra are very similar with a slight red shift where the peak shifts from 320 nm for the 2-imidazole cylinder to 323 nm for the pseudorotaxane, and increased resolution of the shoulders around 300 nm and 355 nm. (

Figure 5.2-34)

To assess whether the threading affected the hydrolysis of the immine the stability under heating over time was examined.

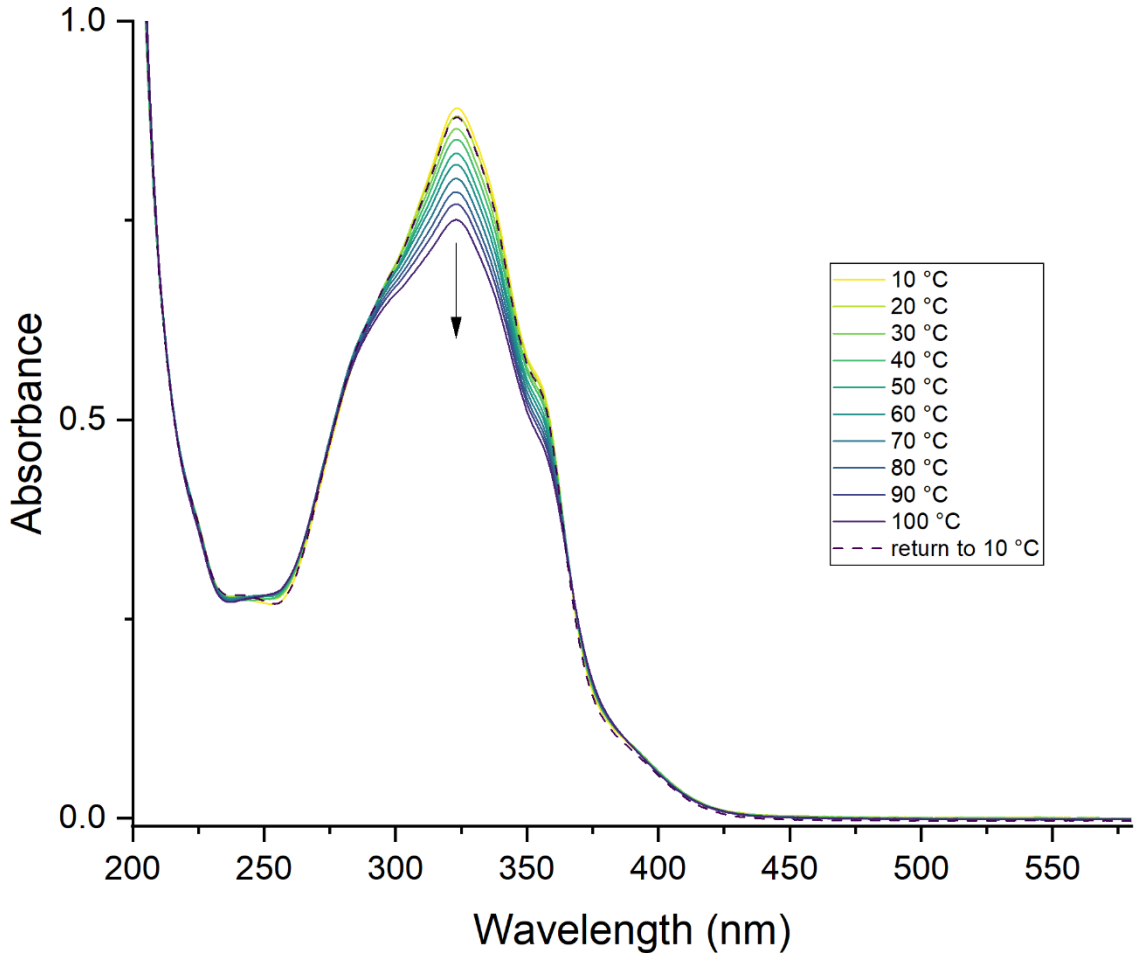


Figure 5.2-35 UV-Vis spectrum of 2-imidazole cylinder pseudorotaxane in <5% merthanol/water. Variable temperature increments of 10 °C from 10 °C to 100 °C returning to 10 °C.

As the temperature of an aqueous solution increases up to 100 °C there is a decrease in absorbance at 323 nm as well as the 355 nm shoulder, though the shoulder at approx. 300 nm appears relatively unaffected and becomes more apparent. On returning the solution to 10 °C there is very little change from the initial measurement. Perhaps a slight formation of a peak at 241 nm discussed in the next section. (Figure 5.2-35)

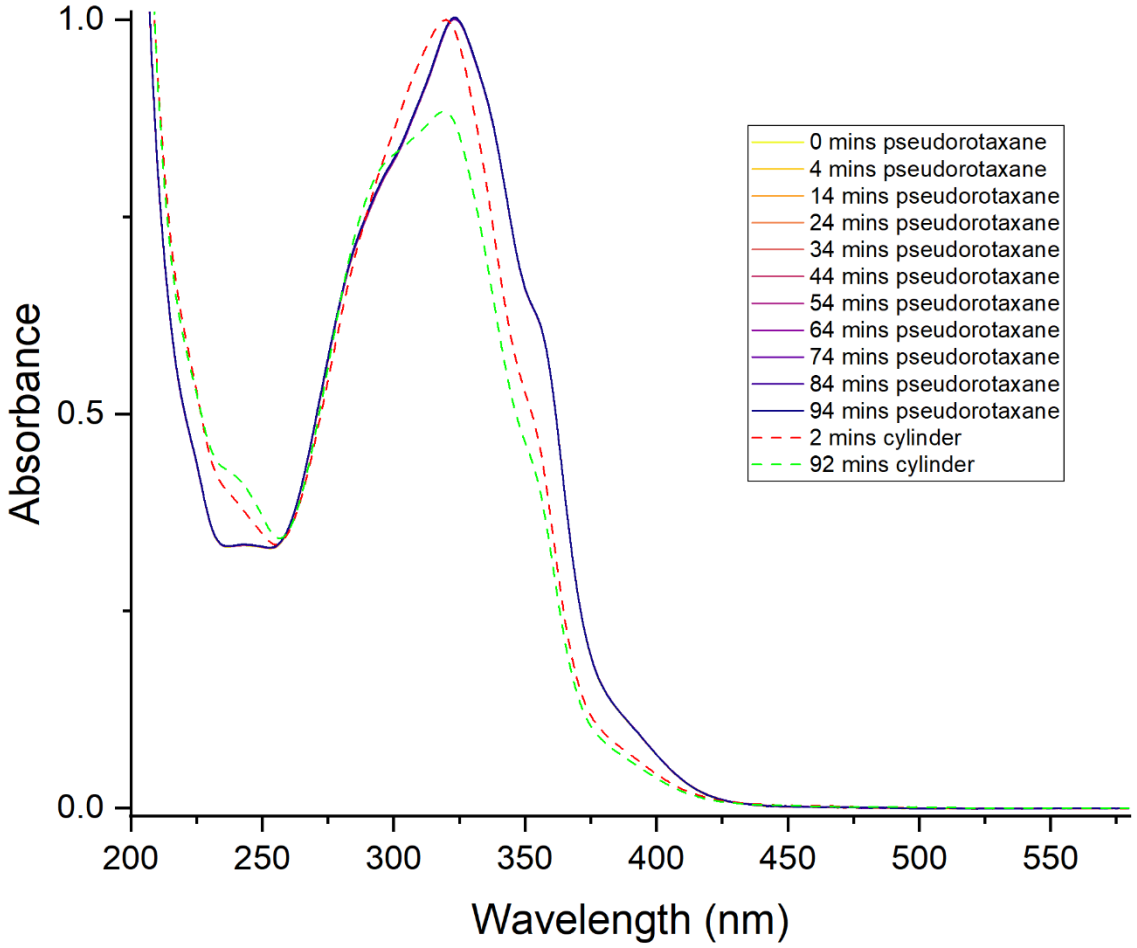


Figure 5.2-36 UV-Vis spectrum of 2-imidazole cylinder pseudorotaxane at 37 °C in <5% methanol/water. Taken at time points from 0 to 94 minutes. Overlayed with similar timepoints from the cylinder in water at 37 °C. Both sets of data are processed independently, and normalised to make the initial peak height 1.

Incubating the same solution of pseudorotaxane in aqueous solution at 37 °C shows no change over 90 minutes. The cylinder in comparison shows much greater change over a similar timeframe. (Figure 5.2-36) There is a marginal peak at 241 nm for 10 °C possibly representative of a small portion of cylinder hydrolysis before its threading. The peak is slightly increased on its return to 10 °C, suggesting either the slight unthreading and subsequent hydrolysis of pseudorotaxane at higher temperatures, or the accelerated hydrolysis of a non-threaded population. (Figure 5.2-35) But the lack of change in the

time dependent experiment suggests that it is a small degree of un-threading at higher temperatures causing this small change. (Figure 5.2-36)

Considering the improved stability resulting from the protection of the immine by the CB[10], it is possible that further research could be done on separating cylinders using HPLC. This is also relevant for the reverse where the cylinder can solubilize the CB[10] to be separated post functionalization using chromatography.

5.2.3.4 To Yield a DNA 3WJ Forming Species.

Decapping the rotaxane results in a mixture of partially capped rotaxanes and pseudo-rotaxane. Selecting reliable decapping conditions, a solution of 10 mM mercaptoethanol in acetonitrile, resulted in a product capable of forming a DNA 3WJ with 14mer DNA. (Figure 5.2-37)

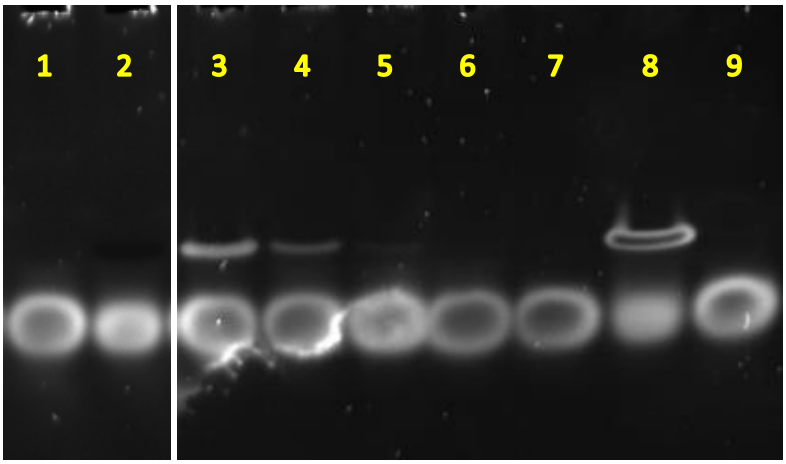


Figure 5.2-37 20% 37.5:1 PAGE image showing the effect of the uncapped rotaxane on 14mer 3WJ DNA (S1+S2+S3) in pH 7.05 TBA buffer. Lanes in order left to right: 1. 14mer, 2. 14mer + FeCy 1:1, 3-6. 14mer + decapped rotaxane 1:0.2 (10 uM DNB_Rtxn incubated with 10 mM mercaptoethanol in MeCN at 37 °C for 3, 2, 1, 0 hours respectively. Then lyophilised and dissolved in 15% methanol/water to form a stock solution), 7. 14mer +

DNB_Rtxn 1:1, 8. 14mer + 2imCy 1:1, 9. 14mer + DNB_Cyl 1:1. All samples are incubated with S1, S2 and S3 at 37 °C in 89 mM pH 7.05 tris boric acid buffer and 10 mM NaCl for 1 hour. Fluorescence imaged using SYBR Gold stain.

To prepare decapped samples the rotaxane was introduced to a solution of 10 mM mercaptoethanol at 10 uM, intentionally comparable to the UV-Vis studies conducted previously and resulting in a 1:167 ratio of caps to decappant. This enabled an observable difference between different incubation times, with the rotaxane added to vials at set time points and the whole batch cooled in liquid nitrogen and lyophilized together. (Figure 5.2-38) The dry decapped product, now in the absence of mercaptoethanol, was then dissolved in 15% methanol/water, to form a stock solution.

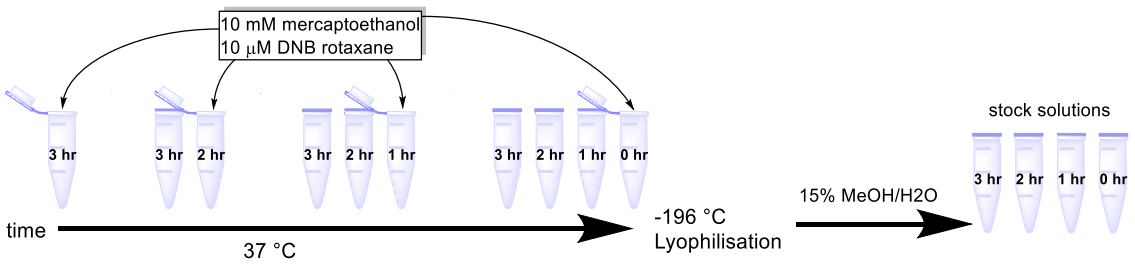


Figure 5.2-38 Illustration of the experimental design, showing the preparation of decapped rotaxane stock solutions.

Preparation of the gel samples begins with the addition of pH 7.05 10x buffer (890 mM tris base, 890 mM boric acid and HCl) and 10x NaCl (100 mM). Followed by S1 (CGGAACGGCACTCG), S2 (CGAGTGCAGCGTGG) and S3 (CCACGCTCGTTCCG), a set of 14mer DNA oligos designed to form either SS-DNA or in the presence of a 3WJ forming agent DNA 3WJ (5 uM). Finally, the compound to be assessed is added (5 uM), and the samples are incubated at 37 °C for 1 hour. After adding glycerol to 10%, the samples are added to the polyacrylamide gel (20 %, 37.5:1, pH 7.05 TB buffer) and a potential of 140 V is applied for 2 hours.

The resulting gel shows that increasing incubation times of the DNB rotaxane with mercaptoethanol results in greater formation of 3WJ. Considering the experimental evidence reported in this chapter suggesting the formation of the 2-imidazole cylinder pseudorotaxane, the agent being released to form the 3WJ is 2-imidazole cylinder. A positive control was conducted with synthesized 2-imidazole cylinder, to demonstrate the formation of 3WJ. Additionally, a control of FeCy was run and shows a negative band, thought to be the result of FRET between the dye and the complex. Negative controls were conducted with the DNB cylinder, and rotaxane as well as 14mer DNA alone. Time point T=0 was included with mercaptoethanol to demonstrate no interference resulting from the incubation and lyophilization sample preparation.

5.3 Conclusions

An assembly that can be reliably synthesized has been produced through post-assembly modification. The synthesis makes use of the high binding strength of the uncapped cylinder to the CB[10]. Once inside, the cylinder is protected from hydrolysis until its release. The assembled rotaxane can then be activated by decapping in the presence of a thiol. The reaction produces two products, the 2-imidazole pseudorotaxane and the DNB-SEtOH. It has a similar release profile across NMR, mass spec and UV-Vis spectroscopy, leading to the release of a DNA binding pseudorotaxane. Quantifying the pseudorotaxane formation remained elusive. Though its formation is certain when considering the presence in mass spectrometry it is apparent presence in UV-Vis and the consistent decapping behavior of other DNB derivatives in NMR spectroscopy.

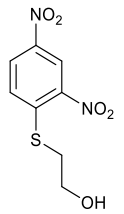
The decapped complex has a preferential affinity for a DNA three-way junction (3WJ) that rivals the CB[10], forming the 3WJ structure in solution. This type of interplay between DNA binding and threaded molecules has been seen when CB[7] has been used to modulate ligand DNA interactions.[47] Though in that case the affinity for the CB[7] was used to reduce aggregation between a ligands on a DNA sequence.

By creating a stimuli responsive molecule with a product that functions through non-covalent interactions has a dual effect. The thiol response creates an increased activity at locations with higher concentrations, with potential targets among various cancers. The non-covalent interactions avoid nucleic acid damage, an alternative to the common DNA damage route used by platinum-based agents, leaving no genotoxicity. Combined this has promise for use in cancer treatments with reduced damage to the surrounding tissues.

It is worth considering though that this release mechanism can be improved upon. Poor solubility made analysis challenging and would have clinical implications too. For release there are other methods, such as light triggered or enzymatic cleavage that are worth exploring.

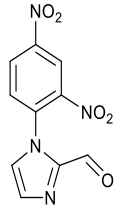
6 Experimental

6.1 Synthesis of 2-((2,4-dinitrophenyl)thio)ethan-1-ol



Mercaptoethanol (5.2 mmol) in 20 mL of dry $CHCl_3$ was slowly added to a solution of 1-fluoro-2,4-dinitrobenzene (1 g, 5.2 mmol) in triethylamine (7 mL) at room temperature. The reaction process was monitored by TLC. The mixture was extracted with HCl (1 M), and then the organic layer was washed twice with water. The product was separated, dried over magnesium sulfate ($MgSO_4$), filtered, and concentrated under vacuum. Crude products were then recrystallized from $CHCl_3$, giving a bright yellow solid. Yield 783 mg (65%). ^{1}H NMR (400 MHz, CD_3CN) δ 8.90 (d, 1H, J = 2.6 Hz), 8.35 (dd, J = 9.1, 2.6 Hz, 1H), 7.80 (d, J = 9.1 Hz, 1H), 3.84(dt, J = 6.0 Hz, 2H), 3.27 (t, J = 6.0 Hz, 2H), 3.10 (t, J = 5.8 Hz, 1H)

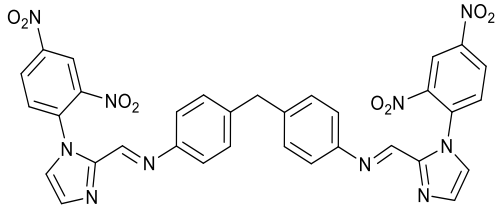
6.2 Synthesis of 1-(2,4-dinitrophenyl)-1H-imidazole-2-carbaldehyde



2-imidazole carboxaldehyde (103 mg, 1.07 mmol) is added to an RBF and dissolved in acetonitrile (20 mL), to this solution is added potassium carbonate (223 mg, 1.61 mmol),

followed by 1-fluoro-2,4-dinitrobenzene (200 mg, 1.07 mmol) dropwise in a solution of acetonitrile (20 mL). The mixture is stirred under reflux for 67 hours in an inert atmosphere and monitored by TLC. The mixture is filtered through cotton to remove the potassium carbonate. Followed by concentrating and separation by column chromatography on silica eluting with 88% chloroform, 10% ethyl acetate, 2% methanol. Yield (196 mg) 75%. ^{1}H NMR (400 MHz, Acetonitrile-d_3) δ 9.68 (d, J = 0.9 Hz, 1H, CHO), 8.96 (dd, J = 2.6, 0.4 Hz, 1H, DNB-H), 8.64 (dd, J = 8.7, 2.6 Hz, 1H, DNB-H), 7.83 (dd, J = 8.7, 0.4 Hz, 1H, DNB-H), 7.54 (t, J = 1.0 Hz, 1H, Imm-H), 7.52 (d, J = 1.1 Hz, 1H, Imm-H). 1H). Mass spectrum (ESI): m/z = 263.04 (C10H7N4O5)$[M+H]^+$. CHN: C 45.95 %, H 2.25 %, N 21.34 % (Predicted C 45.81 %, H 2.31 %, N 21.37 %).

6.3 Synthesis of (1E,1'E)-N,N'-(methylenebis(4,1-phenylene))bis(1-(1-(2,4-dinitrophenyl)-1H-imidazol-2-yl)methanimine) (L^{DNB})

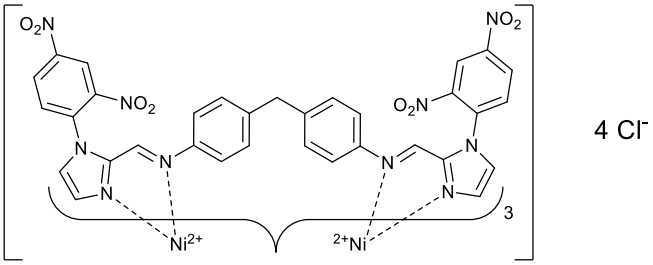


DNB aldehyde (50 mg, 0.19 mmol, 2.1 equivalents) and 4,4' Methylene dianiline(18 mg, 0.09 mmol, 1 equivalent) were taken in a 25 ml round bottom flask in 15 ml ethanol and stirred at room temperature over night to obtain the ligand as a yellow precipitate. The precipitate was filtered and washed with cold ethanol and dried over vacuum. Yield (50 mg, 80%). ^{1}H NMR (400 MHz, DMSO-d6) δ 8.92 (d, J = 2.6 Hz, 1H, DNB-Ha), 8.67 (dd, J =

8.7, 2.6 Hz, 1H, DNB-Hb), 8.47 (d, J = 0.6 Hz, 1H, Imine-H), 7.98 (d, J = 8.6 Hz, 1H, DNB-Hc), 7.75 (dd, J = 1.2, 0.6 Hz, 1H, Imidazole-H), 7.47 (d, J = 1.2 Hz, 1H, Imidazole-H), 7.14 (d, J = 8.4 Hz, 2H, Ph-2H), 6.93 (d, J = 8.4 Hz, 2H, Ph-2H), 3.86(s, 2H, CH_2). ^{13}C NMR (101 MHz, DMSO) δ 119.94, 128.51, 148.06, 131.09, 125.46, 131.04, 129.20, 121.04, 39.83

Mass spectrum (ESI): m/z = 687.3 ($M+H^+$). CHN: C 57.87 %, H 3.34 %, N 20.67 % (Predicted C 57.73 %, H 3.23 %, N 20.40 %).

6.4 Synthesis of DNB capped nickel cylinder ($[Ni_2L^{DNB}_3]^{4+}$)



6.4.1 Ligand first

A suspension of L^{DNB} (20 mg, 0.029 mmol, 3 eq) in dry methanol (5 mL) is stirred under argon, as a solution of $NiCl_2.6H_2O$ (4.62 mg, 0.019 mmol, 2 eq) in methanol (5 mL) is added slowly. The mixture is heated at 50 °C for 1 hour after which the transparent yellow solution is filtered. The filtrate is treated with excess methanolic ammonium hexafluorophosphate. The resultant precipitate is filtered, washed with methanol, and dried under vacuum to yield **$[Ni_2L^{DNB}_3]$ $4PF_6^-$** (18 mg, 80%) Mass spectrum (ESI): m/z = 544.0949 ($[M-4Cl]^{4+}$).

6.4.2 Post-assembly modification of 2-imidazole cylinder

To a solution of nickel 2-imidazole cylinder (14.5 mg, 8.236 umol) and 1-fluoro-2-4-dinitrobenzene (48.9 mg, 0.263 mmol) in acetonitrile (5 mL) was added potassium

carbonate (18.5 mg, 0.134 mmol). The mixture was stirred under reflux for 70 minutes. On cooling, the precipitate was filtered to yield the product (20.6 mg, 7.08 umol, 86 %) as the phenolate. Mass spectrum (ESI): m/z = 544.09 ($[M-4DNP]^{4+}$), 786.45 ($[M-3DNP]^{3+}$), 1015.17 ($[M-3DNP + ligand]^{3+}$), 1271.19 ($[M-2DNP]^{2+}$), 715.63 ($[M+L]^{4+}$).

6.5 Post-assembly synthesis of DNB capped nickel rotaxane $[Ni_2L^{DNB}_3 \bullet CB[10]]\ 4DNP^-$

A suspension of CB[10] (10 mg, 6.0 µmol, 1.1 eq) is stirred in a degassed aqueous (10 mL) solution of $[Ni_2L^{2im}_3]Cl_4$ (7.2 mg, 5.5 mol, 1 eq). The mixture is stirred under argon for 10 minutes. To the mixture is added aqueous (5 mL) potassium carbonate (9.2 mg, 66.8 µmol, 12 eq) followed by aqueous (5 mL) 1-fluoro-2,4-dinitrobenzene (18.6 mg, 100 µmol, 18.3 eq), and the mixture is heated to 80 °C for 60 minutes. The yellow/orange precipitate is filtered and washed with water, methanol, ether, and chloroform and dried under high vacuum to yield **$[Ni_2L^{DNB}_3 \bullet CB[10]]\ 4DNP^-$** (16.5 mg, 76% yield). Mass spectrum (ESI): m/z = 959.4623 ($[M-4DNP]^{4+}$, Predicted = 959.4627). CHN: C 46.82 %, 3.1 %, N 22.75 % (Predicted C 48.09 %, H 3.04 %, N 23.90 %).

6.6 Data Acquisition and Analysis

6.6.1 NMR Spectroscopy

For timed kinetics experiments a solution of complex in acetonitrile-D_3 was prepared and preliminary 1D-^{1}H-NMR spectra were acquired. Then a solution of rotaxane and mercaptoethanol was prepared with a quick inversion to stir, the mixture was then added to the carousel where the sample was automatically inserted and the instrument

locked, tuned, and shimmed. The spectrum is run using a 30° pulse, with an acquisition time of 3.7 s and a relaxation delay of 15 s. Attempts were made to keep the sample at a set temperature within the spectrometer and use the spooler to sequence measurements and repeats, though issues with the VTU rendered this unsuitable. The room temperature was set to a stable 25 °C and the VTU spectrometer temperature control was disabled, preventing the sample experiencing different temperatures inside and outside of the spectrometer.

MestreNova was used to prepare figures for most NMR spectra, however for some analysis Brucker Dynamics Center was used, with pre-processing done in Brucker TopSpin. When using TopSpin, to ensure no bias the spectra are processed the same way in batch. This is done with the “multicmd” command, where a set of spectra have their phasing (phc0, phc1) set to the same value followed by an “abs” automatic baseline correction. An individual dataset cannot be corrected to give a good baseline, the whole series must be considered. This comes from a desire to preserve the fitting shape over the origin, the intercept is independent of the slope and can be corrected if needed.

6.6.2 UV-Vis Spectroscopy

All UV-Vis experiments were performed using a Varian Cary 5000 spectrometer, measurements were run in single beam mode and baseline correction was taken using blank solvents. Measurements were taken using 1 cm pathlength quartz cuvettes and were recorded between 200-800 nm. Timed experiments were conducted with temperatures controlled using the Varian Cary Peltier temperature controller, often in series with measurements staggered between cuvettes.

7 References

1 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, Rotaxanating Metallo-supramolecular Nano-cylinder Helicates to Switch DNA Junction Binding., *J. Am. Chem. Soc.*, 2020, **142**, 20651–20660.

2 P. E. Eaton and C. E. Stubbs, The Synthesis of a Stable Complex of a Macrocycle and a Threaded Chain, 1967, **3789**, 5723–5724.

3 G. Schill and H. Zollenkopf, Rotaxan-Verbindungen, I1), *Justus Liebigs Ann. Chem.*, 1969, **721**, 53–74.

4 A. W. Heard and S. M. Goldup, Simplicity in the Design, Operation, and Applications of Mechanically Interlocked Molecular Machines, *ACS Cent. Sci.*, 2020, **6**, 117–128.

5 I. Yoon, M. Narita, T. Shimizu and M. Asakawa, Threading-Followed-by-Shrinking Protocol for the Synthesis of a [2]Rotaxane Incorporating a Pd(II)−Salophen Moiety, *J. Am. Chem. Soc.*, 2004, **126**, 16740–16741.

6 S. Saha and J. F. Stoddart, Photo-driven molecular devices, *Chem. Soc. Rev.*, 2007, **36**, 77–92.

7 M. J. Langton and P. D. Beer, Rotaxane and catenane host structures for sensing charged guest species, *Acc. Chem. Res.*, 2014, **47**, 1935–1949.

8 J. J. Gassensmith, E. Arunkumar, L. Barr, J. M. Baumes, K. M. DiVittorio, J. R. Johnson, B. C. Noll and B. D. Smith, Self-assembly of fluorescent inclusion complexes in competitive media including the interior of living cells, *J. Am. Chem. Soc.*, 2007, **129**, 15054–15059.

9 A. Acevedo-jake, A. T. Ball, M. Galli, M. Kukwikila, D. G. Singleton, A. Tavassoli, S. M. Goldup, M. Denis, D. G. Singleton, A. Tavassoli and S. M. Goldup, AT-CuAAC Synthesis of Mechanically Interlocked Oligonucleotides, *J. Am. Chem. Soc.*, 2020, **142**, 5985–5990.

10 X. Wang and D. B. Smithrud, Pt-rotaxanes as cytotoxic agents, *Bioorganic Med. Chem. Lett.*, 2011, **21**, 6880–6883.

11 Y.-Q. Yao, Q. Liu, Y. Huang, Q.-J. Zhu, Y.-Q. Zhang, X. Xiao, Z. Tao and G. Wei, Supramolecular Assemblies of Cucurbit[10]uril Based on Outer Surface Interactions, *Aust. J. Chem.*, 2017, **70**, 637.

12 K. I. Assaf and W. M. Nau, Cucurbiturils: From synthesis to high-affinity binding and catalysis, *Chem. Soc. Rev.*, 2015, **44**, 394–418.

13 F. Biedermann, W. M. Nau and H.-J. Schneider, The Hydrophobic Effect Revisited-Studies with Supramolecular Complexes Imply High-Energy Water as a Noncovalent Driving Force, *Angew. Chemie Int. Ed.*, 2014, **53**, 11158–11171.

14 F. Biedermann, M. Vendruscolo, O. A. Scherman, A. De Simone and W. M. Nau, Cucurbit[8]uril and Blue-Box: High-Energy Water Release Overwhelms Electrostatic Interactions, *J. Am. Chem. Soc.*, 2013, **135**, 14879–14888.

15 F. Biedermann, V. D. Uzunova, O. A. Scherman, W. M. Nau and A. De Simone, Release of High-Energy Water as an Essential Driving Force for the High-Affinity Binding of Cucurbit[n]urils, *J. Am. Chem. Soc.*, 2012, **134**, 15318–15323.

16 W. M. Nau, M. Florea and K. I. Assaf, Deep Inside Cucurbiturils: Physical Properties and Volumes of their Inner Cavity Determine the Hydrophobic Driving Force for Host-Guest Complexation, *Isr. J. Chem.*, 2011, **51**, 559–577.

17 S. J. C. Lee, J. W. Lee, H. H. Lee, J. Seo, D. H. Noh, Y. H. Ko, K. Kim and H. I. Kim, Host–Guest Chemistry from Solution to the Gas Phase: An Essential Role of Direct Interaction with Water for High-Affinity Binding of Cucurbit[n]urils, *J. Phys. Chem. B*, 2013, **117**, 8855–8864.

18 W. L. Mock and N. Y. Shih, Structure and selectivity in host-guest complexes of cucurbituril, *J. Org. Chem.*, 1986, **51**, 4440–4446.
19 W. L. Mock and N. Y. Shih, Organic ligand-receptor interactions between cucurbituril and alkylammonium ions, *J. Am. Chem. Soc.*, 1988, **110**, 4706–4710.
20 S. Senler, W. Li, M. H. Tootoonchi, S. Yi and A. E. Kaifer, The cucurbituril 'portal' effect, *Supramol. Chem.*, 2014, **26**, 677–683.
21 Y. Zhang, M. Liu, I. Karatchevtseva, J. R. Price, Z. Tao and G. Wei, Yttrium and lanthanide (Ln = La and Gd) complexes with cucurbit[10]uril: crystals transforming from supramolecular frameworks to coordination nanotubes, *New J. Chem.*, 2020, **44**, 18208–18215.
22 US 8,461,137 B2, 2013, 2.
23 M. H. Miceli and C. A. Kauffman, Isavuconazole: A New Broad-Spectrum Triazole Antifungal Agent, *Clin. Infect. Dis.*, 2015, **61**, 1558–1565.
24 US6300350B1, 2001.
25 World Intellectual Property Organization, WO/2018/081480, 2018.
26 R. Ardeleanu, A. I. Dascalu, A. Neamtu, D. Peptanariu, C. M. Uritu, S. S. Maier, A. Nicolescu, B. C. Simionescu, M. Barboiu and M. Pinteala, Multivalent polyrotaxane vectors as adaptive cargo complexes for gene therapy, *Polym. Chem.*, 2018, **9**, 845–859.
27 T. Ooya, H. S. Choi, A. Yamashita, N. Yui, Y. Sugaya, A. Kano, A. Maruyama, H. Akita, R. Ito, K. Kogure and H. Harashima, Biocleavable polyrotaxane-plasmid DNA polyplex for enhanced gene delivery., *J. Am. Chem. Soc.*, 2006, **128**, 3852–3853.
28 A. Tamura and N. Yui, Rational design of stimuli-cleavable polyrotaxanes for therapeutic applications, *Polym. J.*, 2017, **49**, 527–534.
29 A. Fernandes, A. Viterisi, V. Aucagne, D. A. Leigh and S. Papot, Second generation specific-enzyme-activated rotaxane propeptides., *Chem. Commun. (Camb).*, 2012, **48**, 2083–2085.
30 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, A mechanically interlocked molecular system programmed for the delivery of an anticancer drug, *Chem. Sci.*, 2015, **6**, 2608–2613.
31 T. Kench, P. A. Summers, M. K. Kuimova, J. E. M. M. Lewis and R. Vilar, Rotaxanes as Cages to Control DNA Binding, Cytotoxicity, and Cellular Uptake of a Small Molecule**, *Angew. Chemie Int. Ed.*, 2021, **60**, 10928–10934.
32 G. K. Balendiran, R. Dabur and D. Fraser, The role of glutathione in cancer, *Cell Biochem. Funct.*, 2004, **22**, 343–352.
33 A. H. Abou Ghalia and I. M. Fouad, Glutathione and Its Metabolizing Enzymes in Patients with Different Benign and Malignant Diseases, *Clin. Biochem.*, 2000, **33**, 657–662.
34 M. Gulfam, F. F. Sahle and T. L. Lowe, Design strategies for chemical-stimuli-responsive programmable nanotherapeutics, *Drug Discov. Today*, 2019, **24**, 129–147.
35 C. A. Lang, B. J. Mills, W. Mastropaolo and M. C. Liu, Blood glutathione decreases in chronic diseases, *J. Lab. Clin. Med.*, 2000, **135**, 402–405.
36 Z. Luo, X. Ding, Y. Hu, S. Wu, Y. Xiang, Y. Zeng, B. Zhang, H. Yan, H. Zhang, L. Zhu, J. Liu, J. Li, K. Cai and Y. Zhao, Engineering a hollow nanocontainer platform with multifunctional molecular machines for tumor-targeted therapy in vitro and in vivo, *ACS Nano*, 2013, **7**, 10271–10284.
37 Z. Luo, Y. Hu, K. Cai, X. Ding, Q. Zhang, M. Li, X. Ma, B. Zhang, Y. Zeng, P. Li, J. Li, J. Liu and Y. Zhao, Intracellular redox-activated anticancer drug delivery by functionalized hollow mesoporous silica nanoreservoirs with tumor specificity, *Biomaterials*, 2014, **35**, 7951–7962.
38 R. Wang, X. Xia, Y. Yang, X. Rong, T. Liu, Z. Su, X. Zeng, J. Du, J. Fan, W. Sun and X. Peng,

A Glutathione Activatable Photosensitizer for Combined Photodynamic and Gas Therapy under Red Light Irradiation, *Adv. Healthc. Mater.*, 2022, **11**, 1–10.

39 S. Sinn, E. Spuling, S. Bräse and F. Biedermann, Rational design and implementation of a cucurbit[8]uril-based indicator-displacement assay for application in blood serum, *Chem. Sci.*, 2019, **10**, 6584–6593.

40 W. Gong, X. Yang, P. Y. Zavalij, L. Isaacs, Z. Zhao and S. Liu, From Packed "Sandwich" to "Russian Doll": Assembly by Charge-Transfer Interactions in Cucurbit[10]uril, *Chem. - A Eur. J.*, 2016, **22**, 17612–17618.

41 W. Zhang, S. Gan, A. Vezzoli, R. J. Davidson, D. C. Milan, K. V. Luzyanin, S. J. Higgins, R. J. Nichols, A. Beeby, P. J. Low, B. Li and L. Niu, Single-Molecule Conductance of Viologen–Cucurbit[8]uril Host–Guest Complexes, *ACS Nano*, 2016, **10**, 5212–5220.

42 F. Liu, S. Chowdhury, R. Rosas, V. Monnier, L. Charles, H. Karoui, D. Gigmes, O. Ouari, F. Chevallier, C. Bucher, A. Kermagoret, S. Liu and D. Bardelang, Triple Stack of a Viologen Derivative in a CB[10] Pair, *Org. Lett.*, 2021, **23**, 5283–5287.

43 S. R. Gayam and S.-P. Wu, Redox responsive Pd(ii) templated rotaxane nanovalve capped mesoporous silica nanoparticles: a folic acid mediated biocompatible cancer-targeted drug delivery system, *J. Mater. Chem. B*, 2014, **2**, 7009–7016.

44 F. Sanger, The free amino groups of insulin., *Biochem. J.*, 1945, **39**, 507–15.

45 E. B. De Vargas and R. H. De Rossi, Kinetic study of the reaction of N-(2,4-dinitrophenyl)imidazole with piperidine and n-butylamine, *J. Org. Chem.*, 1984, **49**, 3978–3983.

46 B. S. Ross, P. Ganapati Reddy, H. R. Zhang, S. Rachakonda and M. J. Sofia, Synthesis of diastereomerically pure nucleotide phosphoramidates, *J. Org. Chem.*, 2011, **76**, 8311–8319.

47 E. Y. Chernikova, A. Y. Ruleva, V. B. Tsvetkov, Y. V. Fedorov, V. V. Novikov, T. M. Aliyeu, A. A. Pavlov, N. E. Shepel and O. A. Fedorova, Cucurbit[7]uril-driven modulation of ligand-DNA interactions by ternary assembly, *Org. Biomol. Chem.*, 2020, **18**, 755–766.

8 Appendices

8.1 NMR Spectroscopy

8.1.1 DNB-aldehyde Assignment and Coupling

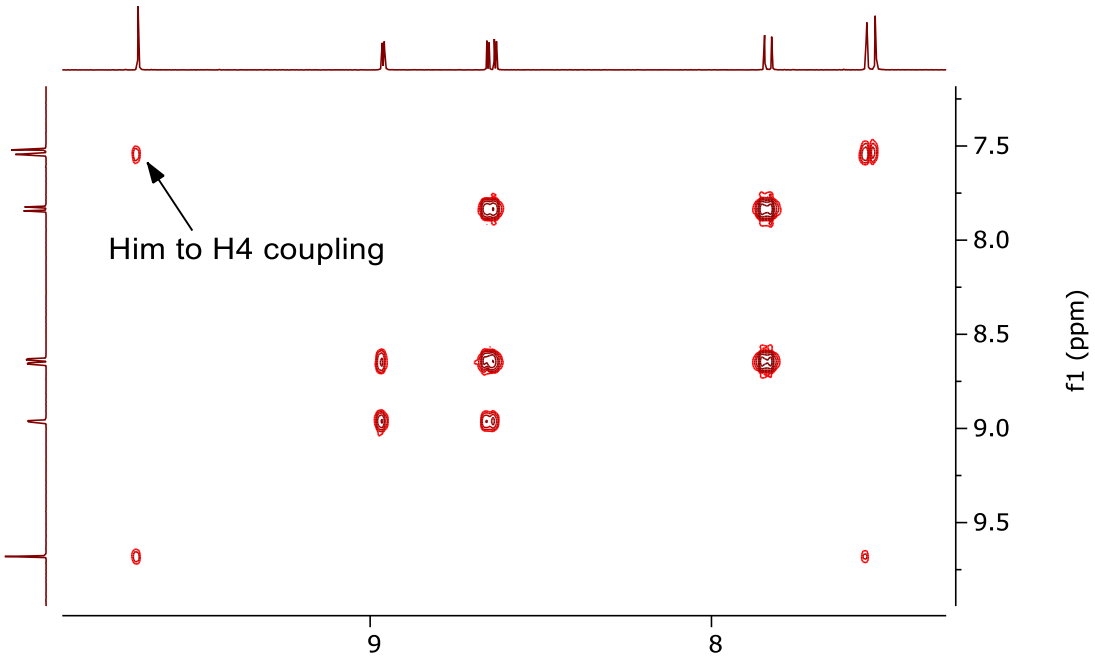


Figure 8.1-1 1H,1H-COSY spectrum of DNB aldehyde in acetonitrile-d3. Showing coupling between DNB peaks as well as interesting 5J coupling between Him and H4.

8.1.2 DNB-ligand assignment

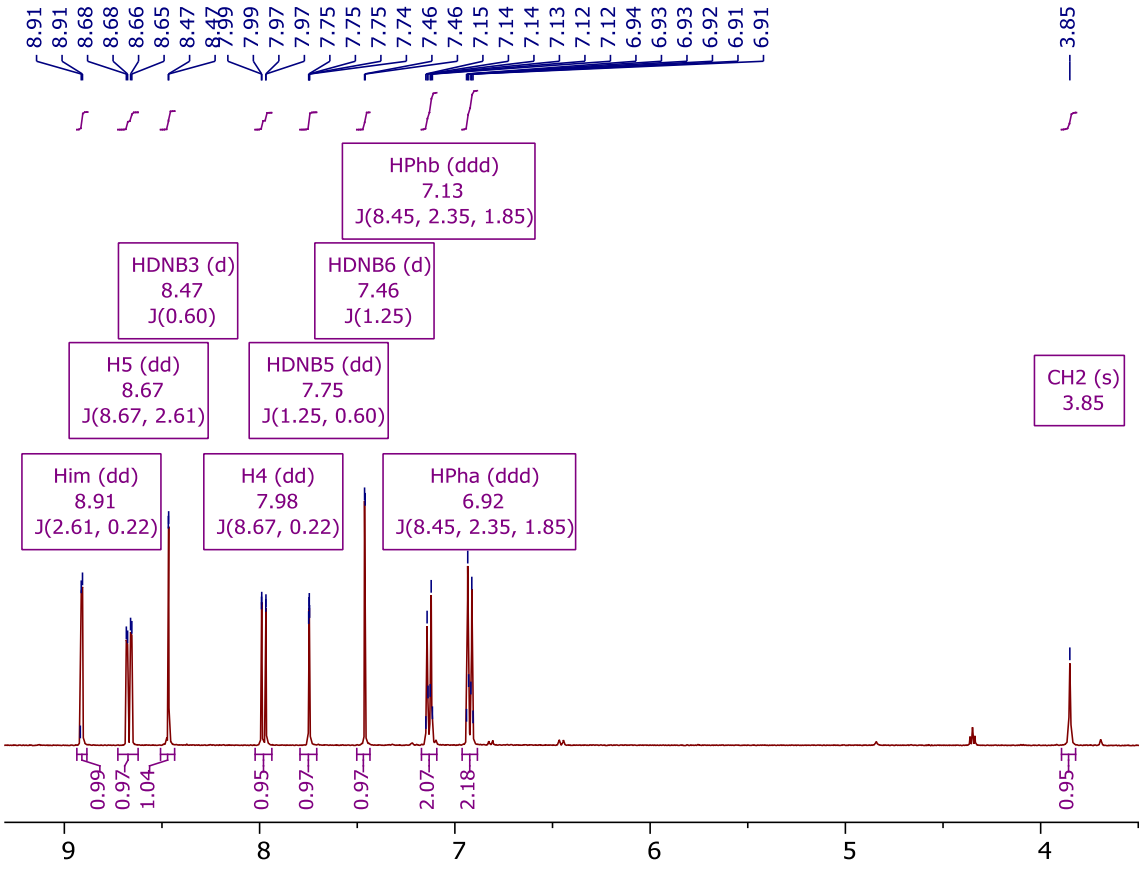


Figure 8.1-2 1H-NMR spectrum of DNB-ligand in DMSO-d6.

8.2 UV-Vis Spectroscopy

8.2.1 2-imCy Hydrolysis Products

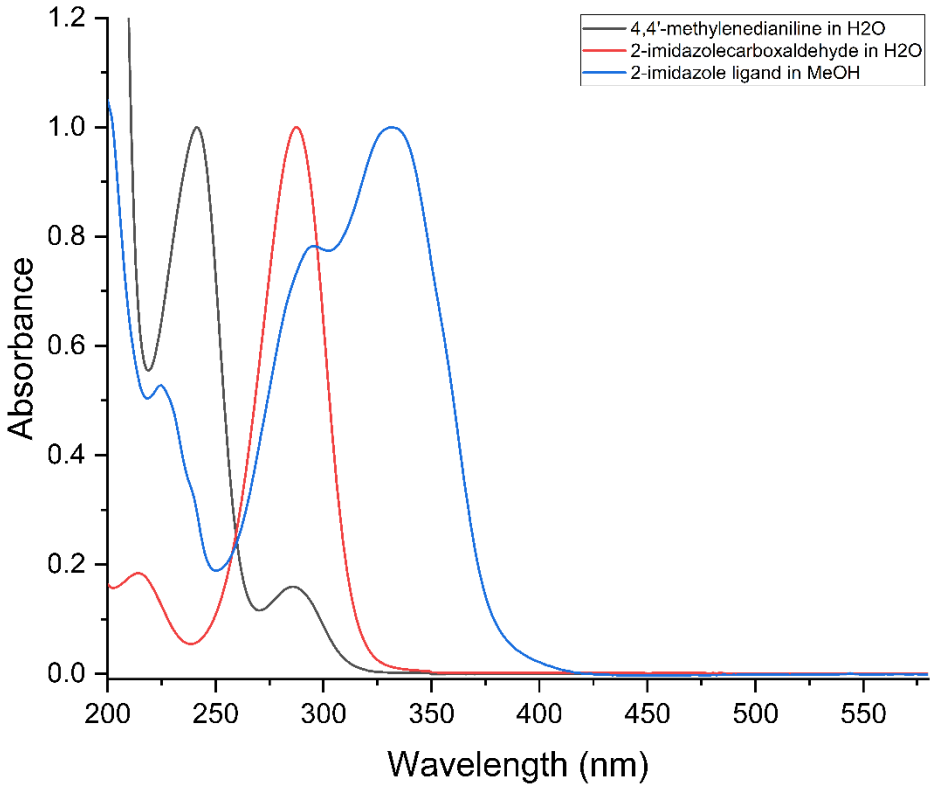


Figure 8.2-1 UV-Vis spectrum of various potential products of the degradation of 2-imidazole cylinder normalised to highest peak.

8.2.2 DNB Rotaxane

8.2.2.1 In Water

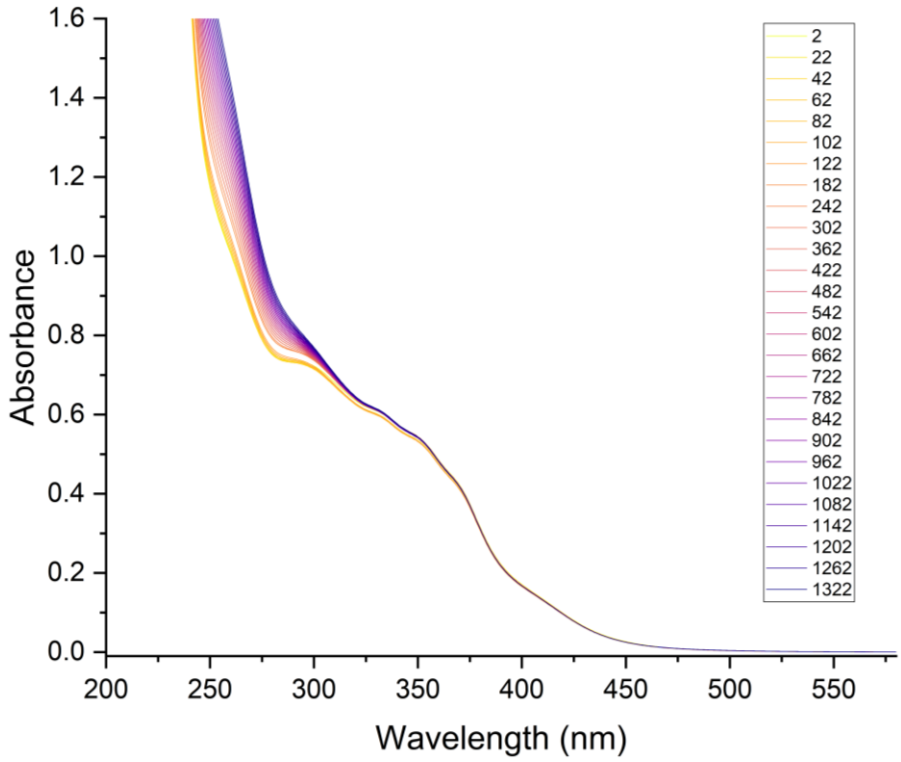


Figure 8.2-2 10 uM DNB rotaxane and 10 mM GSH in H2O 37 °C

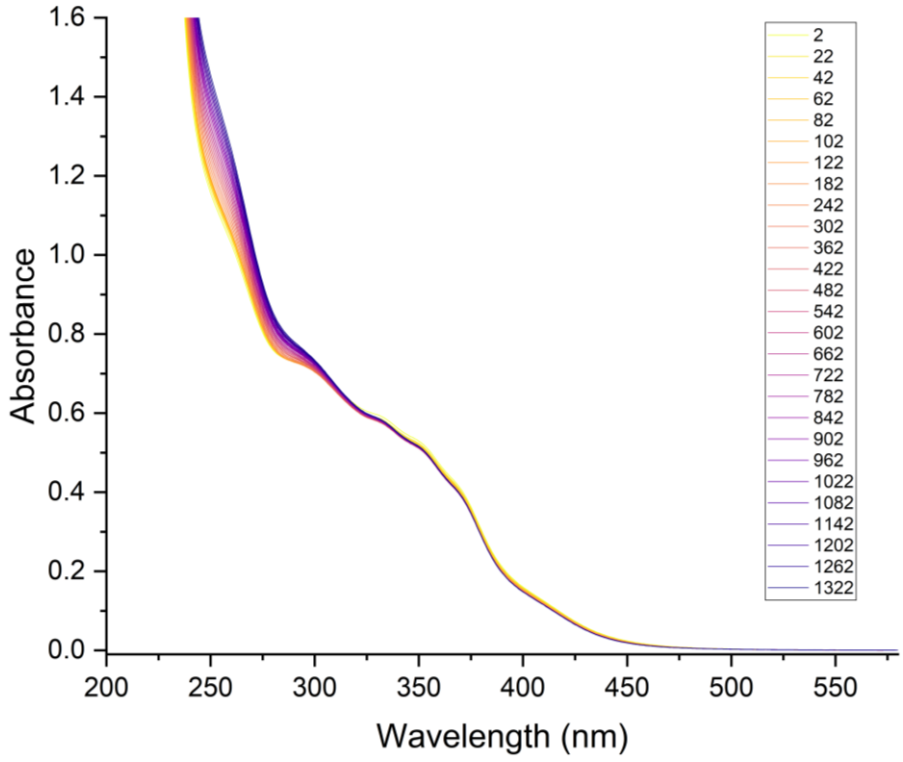


Figure 8.2-3 10 uM DNB rotaxane and 5 mM N-acetyl cysteine in H2O 37 °C

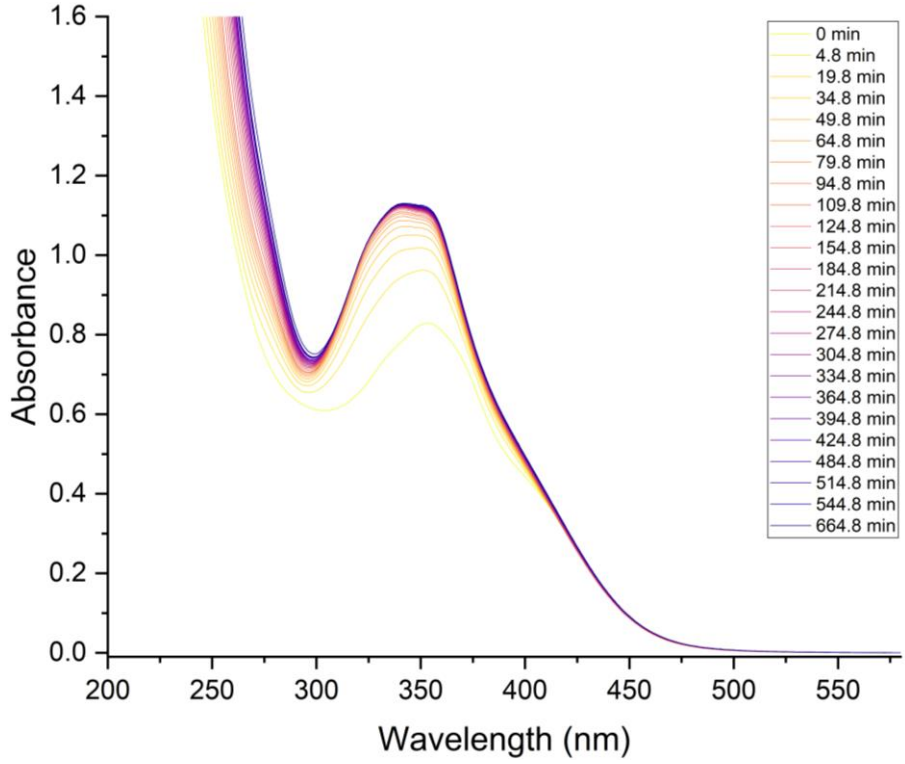


Figure 8.2-4 10 uM DNB rotaxane with 10 mM mercaptoethanol in 1 mM sodium cacodylate buffered H2O 37 °C

8.2.2.2 In Acetonitrile

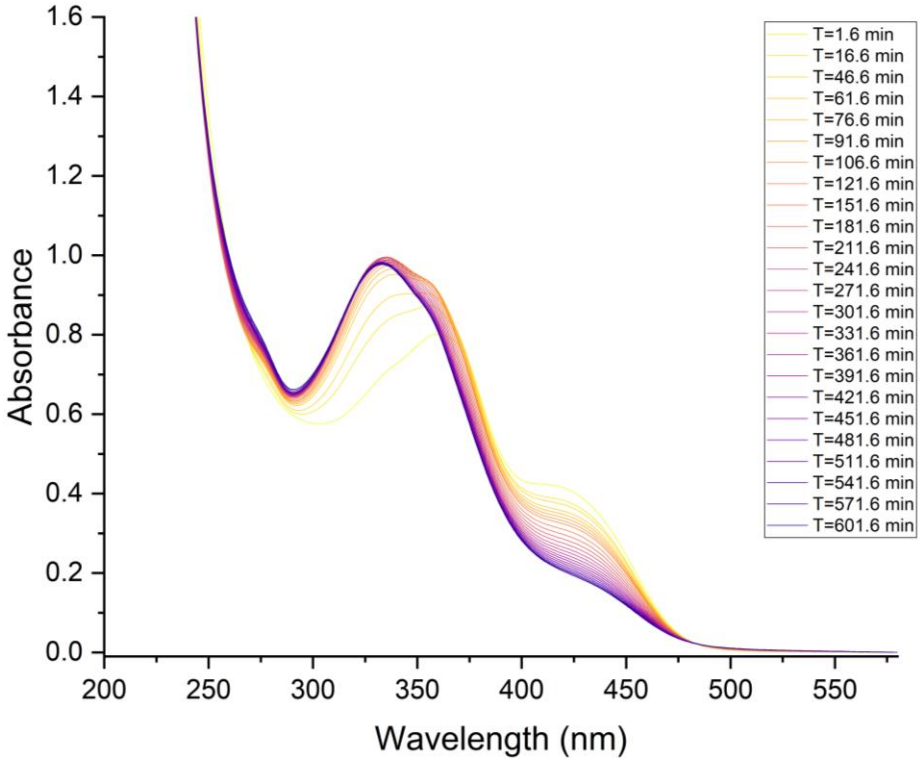


Figure 8.2-5 Figure 3. UV-Vis spectra, showing the reaction of 10 uM DNB rotaxane and 10 mM mercaptoethanol over 10 hours (600 minutes) in MeCN at 37°C.

8.2.3 DNB Cylinder

8.2.3.1 In Water

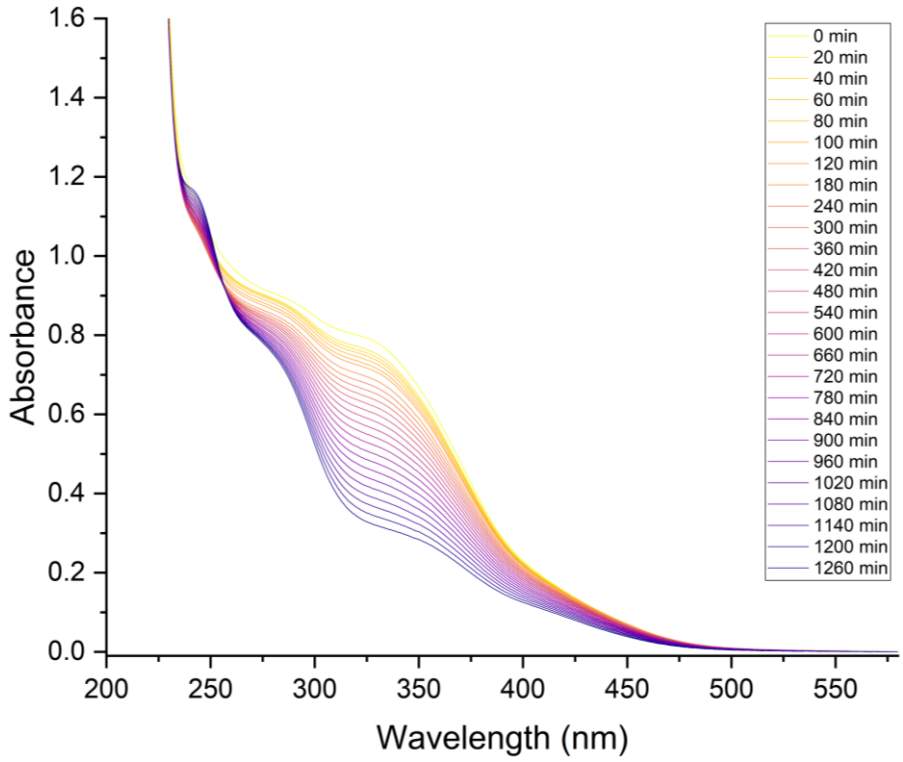


Figure 8.2-6 10 uM DNB cylinder in H2O 37 °C

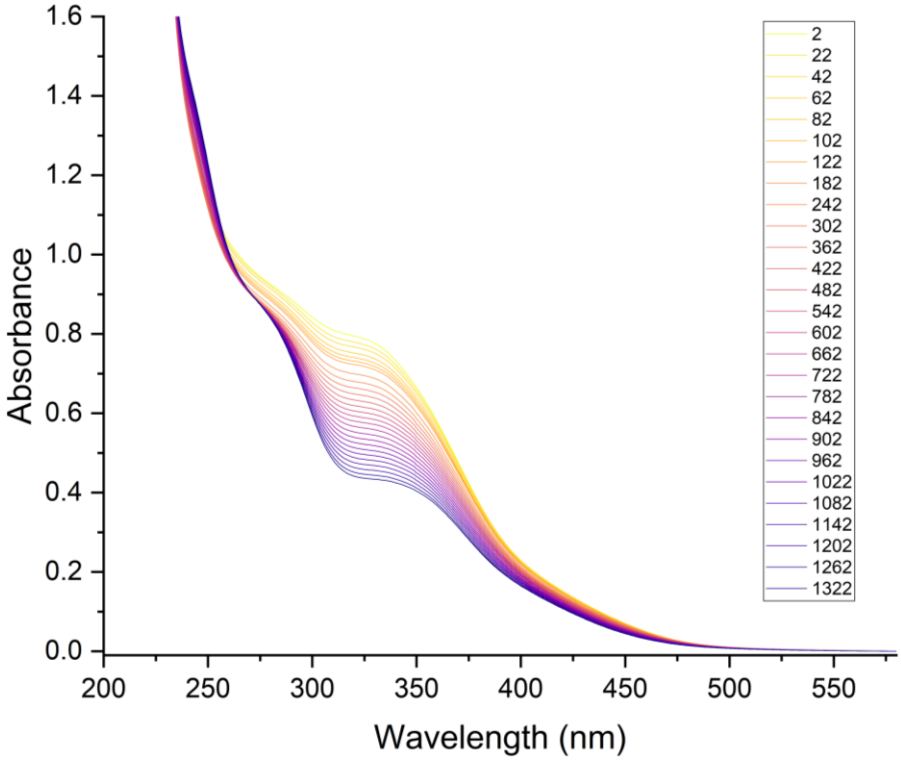


Figure 8.2-7 10 uM DNB cylinder with 5 mM mercaptoethanol in H2O 37 °C

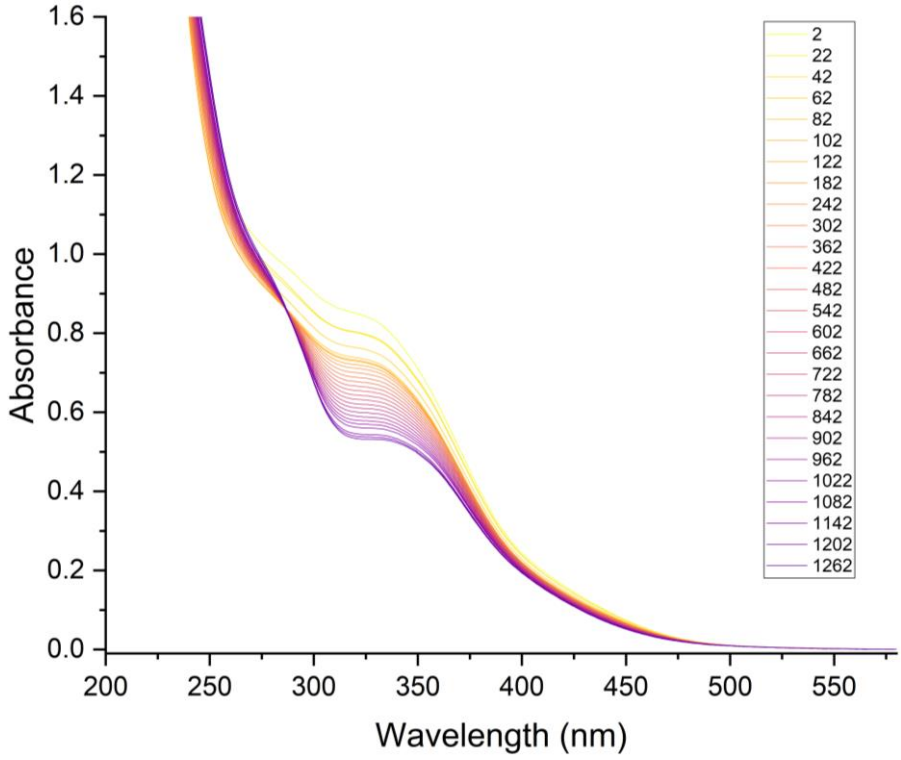


Figure 8.2-8 10 uM DNB cylinder with 10 mM mercaptoethanol in H2O 37 °C

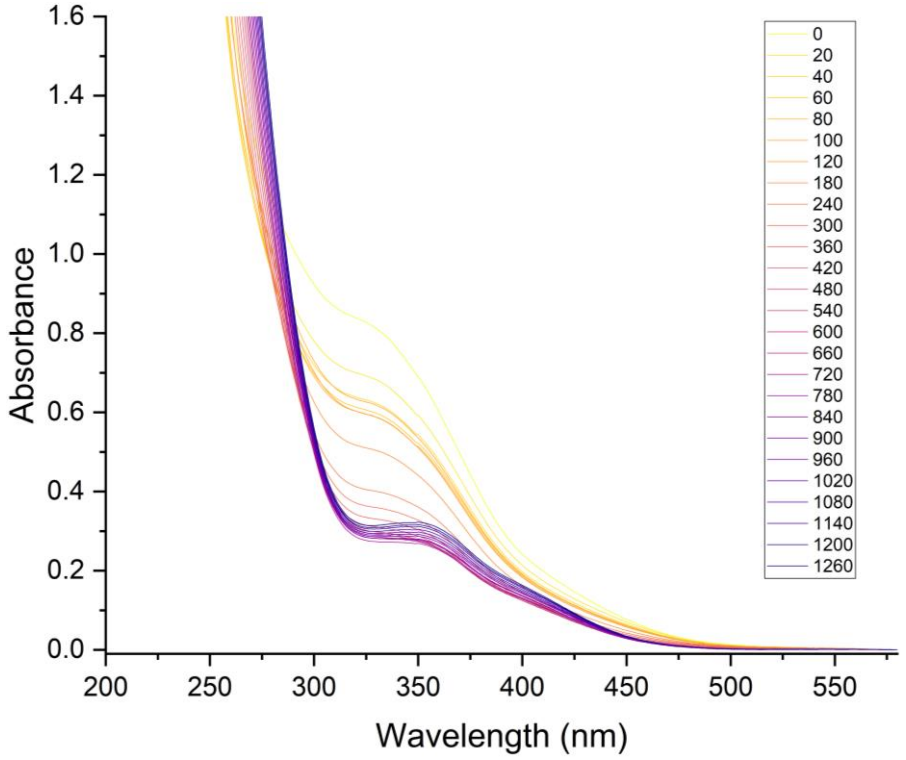


Figure 8.2-9 10 uM DNB cylinder with 100 mM mercaptoethanol in H2O 37 °C

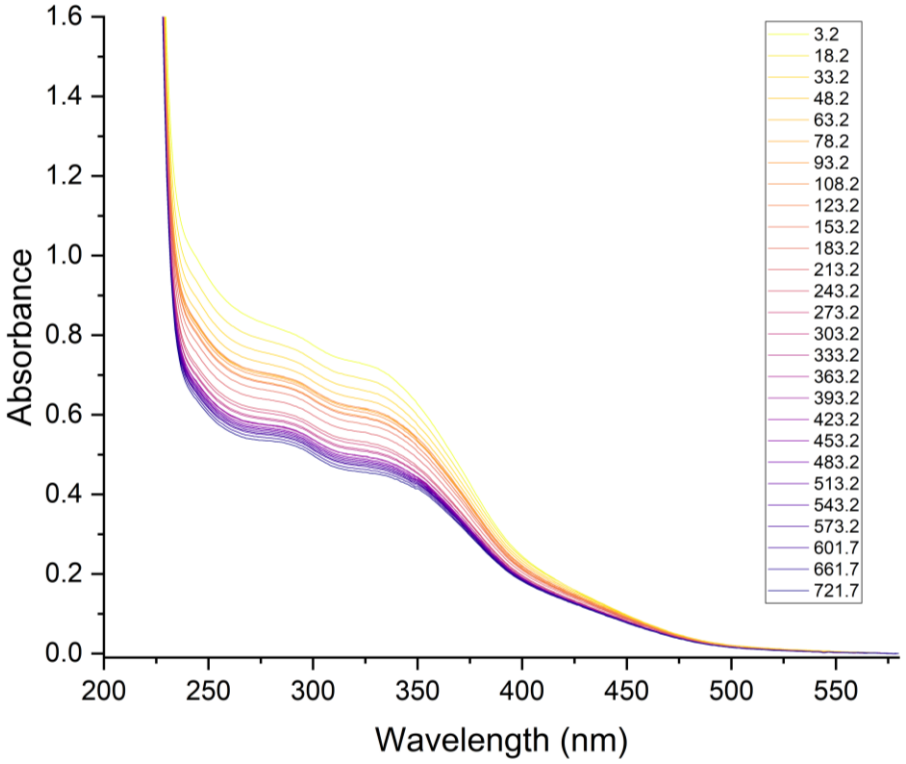


Figure 8.2-10 10 uM DNB cylinder in 1 mM sodium cacodylate buffered H2O 37 °C

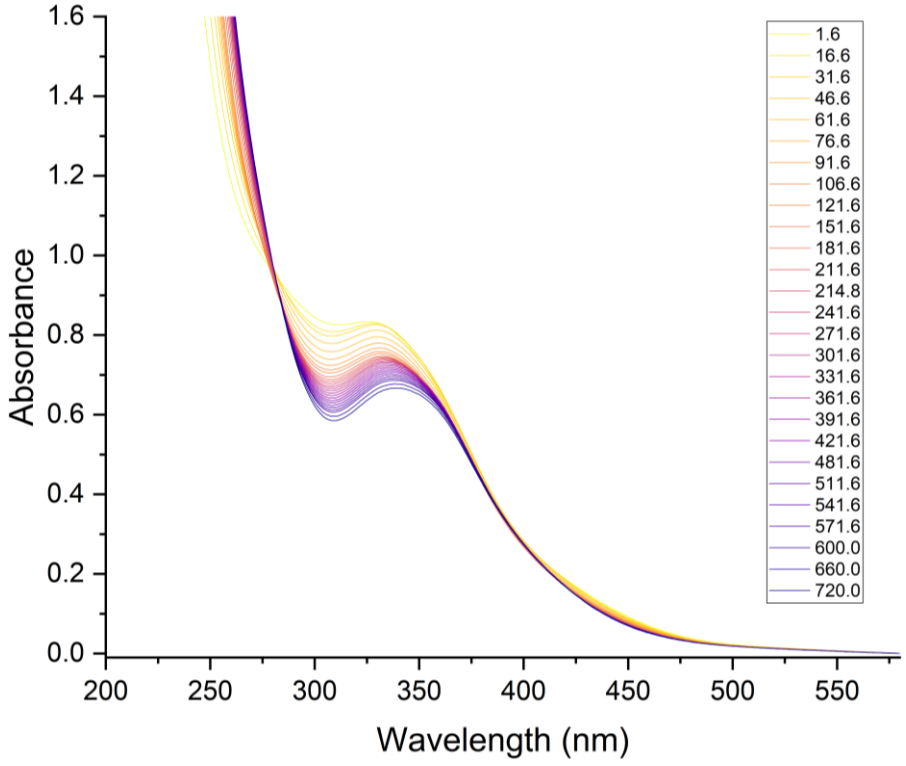


Figure 8.2-11 10 uM DNB cylinder with 10 mM mercaptoethanol in 1 mM sodium cacodylate buffered H2O 37 °C

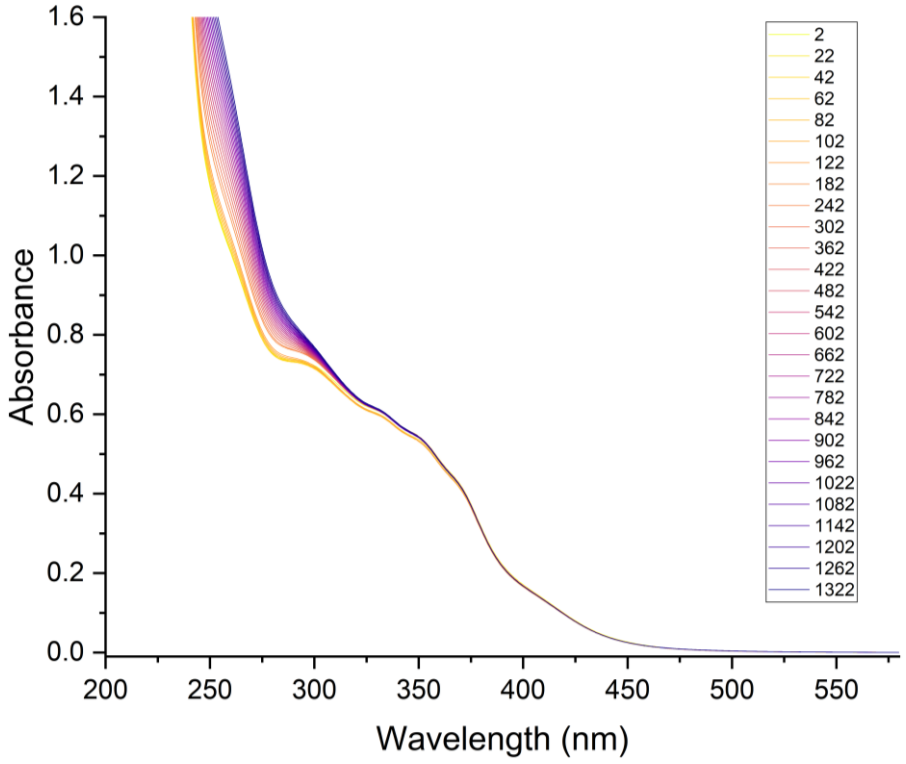


Figure 8.2-12 10 uM DNB cylinder with 10 mM GSH in H2O 37 °C

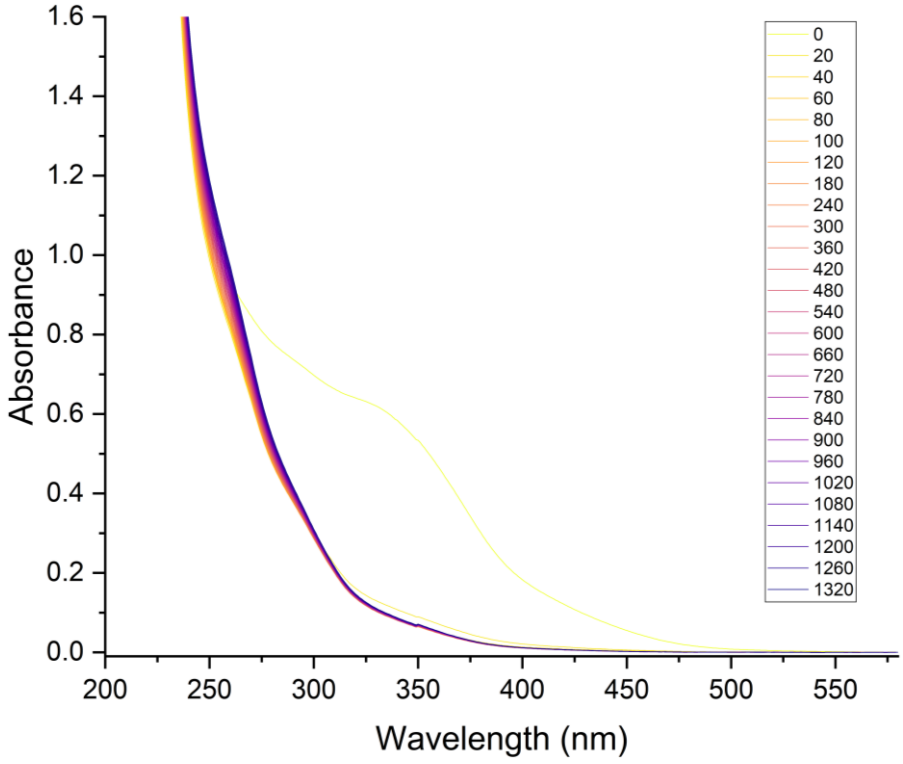


Figure 8.2-13 10 uM DNB cylinder with 5 mM N-acetyl cysteine in H2O 37 °C

8.2.3.2 In Acetonitrile

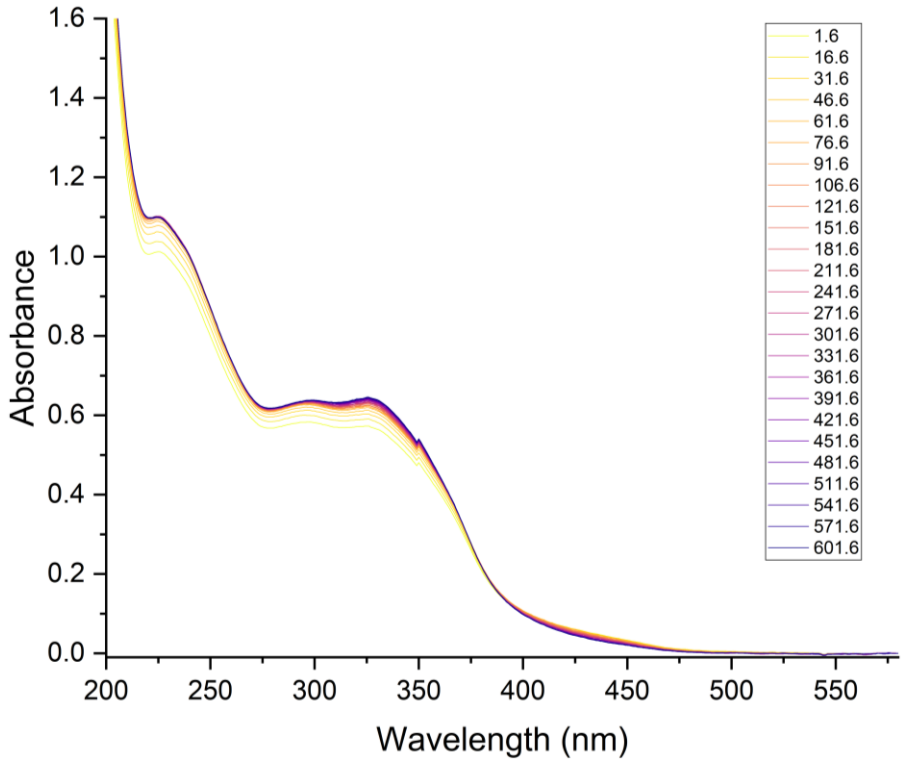


Figure 8.2-14 10 uM DNB cylinder with 0.06 mM mercaptoethanol in MeCN 37 °C

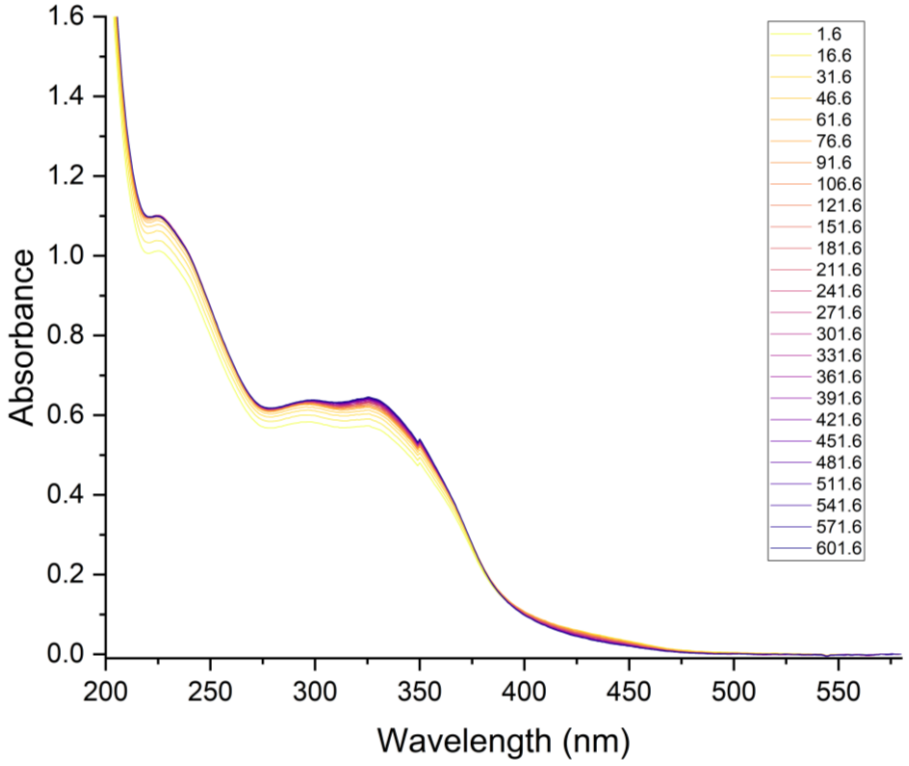


Figure 8.2-15 10 uM DNB cylinder with 5 mM mercaptoethanol in MeCN 37 °C

8.2.4 2-((2,4-dinitrophenyl)thio)ethan-1-ol

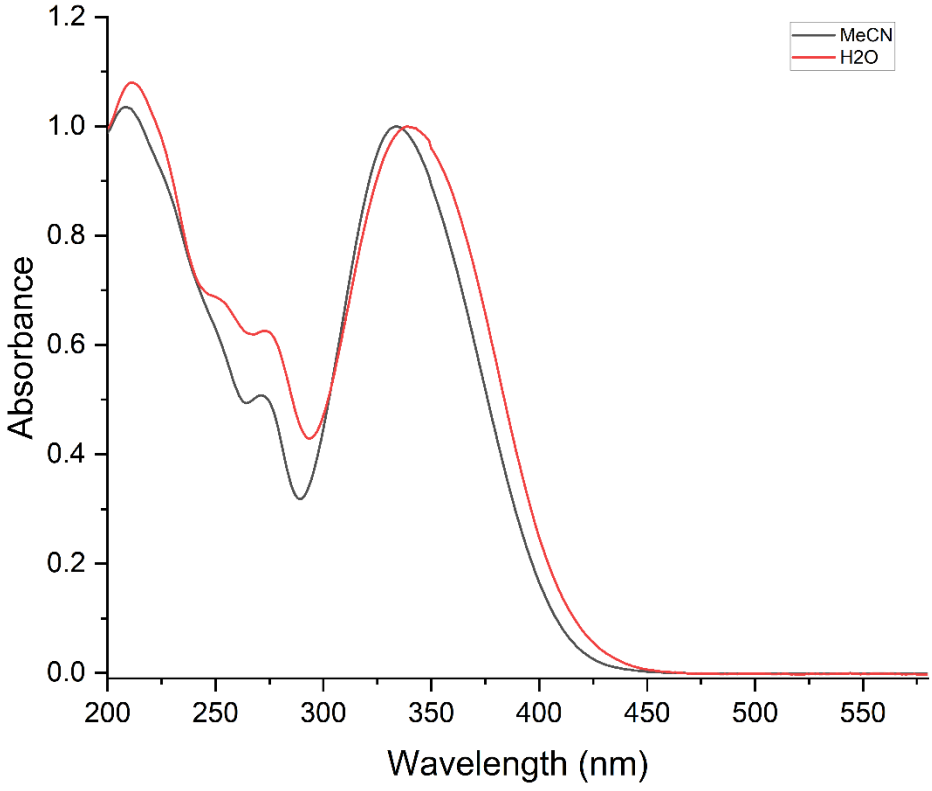


Figure 8.2-16 2-((2,4-dinitrophenyl)thio)ethan-1-ol in MeCN and H2O. Normalized to peak at 350 nm.

8.3 Rotaxane Isotope Pattern

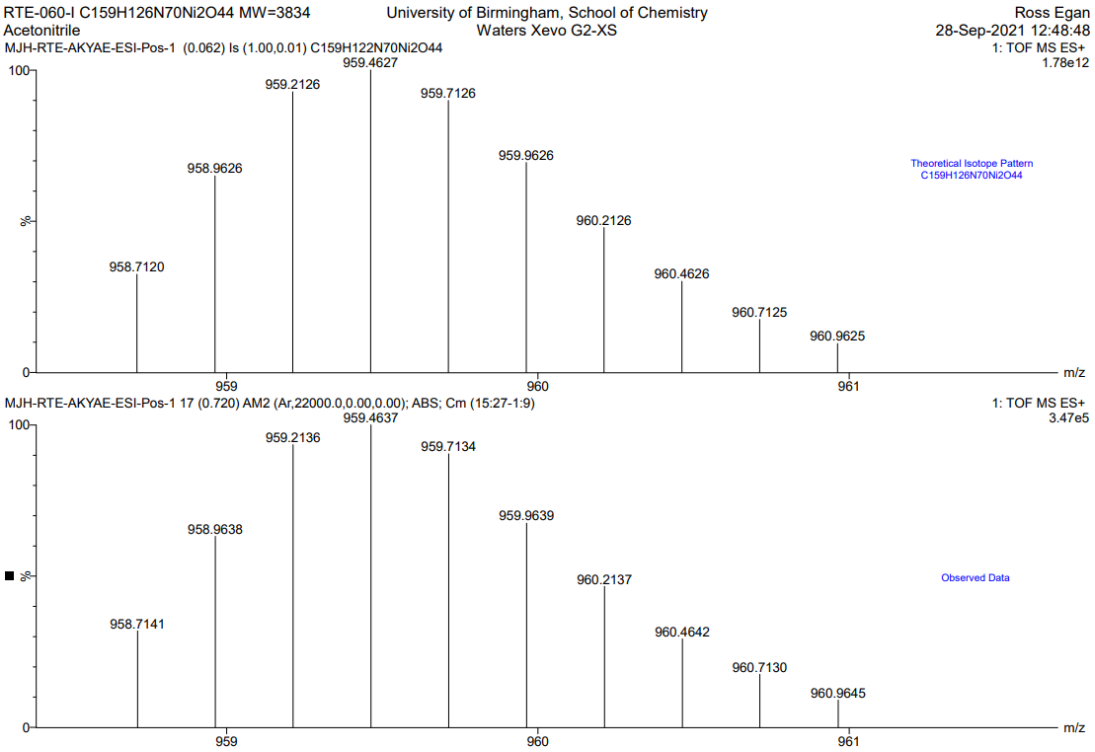


Figure 8.3-1 High resolution isotope pattern for DNB-rotaxane.

CHAPTER 3 – CYLINDER INTERACTIONS WITH TAR RNA

1 Acknowledgements

This work was done in collaboration with Prof. Dr. Roland K. O. Sigel, Dr Silke Johannson and Dr Zenghui Wang.

I wish to thank Dr Silke Johannsen and Dr Zenghui Wang for the continuing effort in acquiring spectra for this project throughout various international lockdowns. Zenghui for his rapid answers to my questions, Silke for her expertise offered and the whole Sigel group for their welcoming attitude to a visiting student with wildly different training to their own.

Dr Aditya Garai was essential posting samples whilst I was in Zurich, but also developing our knowledge of ruthenium cylinder purification further through his work on refining the cellulose column with a celite plug, rapidly improving the speed with which we purified samples from then on with an elegant solution.

AutoSave, for not losing this chapter when my laptop died.

2 Contents

CHAPTER 3 – CYLINDER INTERACTIONS WITH TAR RNA 153

1 ACKNOWLEDGEMENTS 153

2 CONTENTS 153

3 GLOSSARY AND NAMING CONVENTION 156

3.1 NMR Peak Labelling 157

4 INTRODUCTION 158

4.1 RNA 158

4.2 TAR 158

5 RESULTS AND DISCUSSION 163

5.1 Ruthenium Cylinder 163

5.1.1 Synthesis 163

5.1.1.1 Starting Materials 163

5.1.1.2 Ruthenium Cylinder 164

5.1.1.3 Purification 165

5.1.2 Epsilon 166

5.1.3 Enantiomers 171

5.1.3.1 Enantioseparation 171

5.1.3.2 Circular Dichroism (CD) 173

5.1.3.2.1 Finding the Purest Enantiomer. 174

5.2 Free TAR 178

5.2.1 90/10 H_2O to D_2O Spectra 178

5.2.2 100% D_2O Spectra 184

5.3 Bound TAR 190

5.3.1 The Cylinder 190

5.3.2 TAR and RuCy_P 197

5.3.2.1 90% H_2O Spectra 198

5.3.2.2 100% D_2O NOESY 202

5.3.2.3 TOCSY 211

5.3.2.4 HSQC 213

5.3.2.5 DOSY 214

5.4 1:0.5 ratio 215

5.4.1 ^{1}H-NMR ... 215
5.4.2 NOESY ... 218
5.4.3 TOCSY ... 223
5.4.4 NOESY 50 ms ... 225
5.4.5 HSQC ... 228
5.5 Conclusions ... 229
6 EXPERIMENTAL ... 231
6.1 Ruthenium Cylinder ... 231
6.1.1 Parent Ligand **L** ... 231
6.1.2 cis-Tetrakis(dimethylsulphoxide)dichlororuthenium(II) **$Ru(DMSO)_4Cl_2$** ... 232
6.1.3 Ruthenium Cylinder **$[Ru_2(L)_3]Cl_4$** ... 232
6.1.3.1 Celite / Cellulose Plug Purification ... 232
6.1.3.1.1 To further purify as racemic ... 233
6.1.3.1.2 Alternatively, if enantiomers are required ... 234
6.1.3.2 Desalting ... 234
6.2 TAR RNA Preparation (GGCAGAUCUGAGCCUGGGAGCUCUCUGCC) ... 235
6.2.1 Synthesis ... 235
6.2.2 Purification ... 236
6.3 NMR ... 238
6.4 UV-Vis / Circular Dichroism Spectroscopy ... 239
7 REFERENCES ... 239
8 APPENDICES ... 243
8.1 Free TAR ... 243
8.1.1 HSQC ... 243
8.1.2 TOCSY ... 244
8.2 Bound TAR 1:1 ... 245
8.2.1 RNA ... 245
8.2.2 Intermolecular NOEs ... 247
8.3 Bound TAR 1:0.5 ... 250

8.4 Ruthenium Cylinder 250

8.4.1 ^{1}H,^{1}H-NOESY 250

8.4.2 ^{1}H,^{13}C-HSQC/HMBC 251

8.4.3 UV-Vis Spectra Used for Figure 5.1-4. 251

3 Glossary and Naming Convention

L – The ligand used. See Figure 3.1-1 bottom.

Ruthenium cylinder (RuCy) – $[Ru_2(L)_3]^{4+}$ A polypyridyl bimetallic ruthenium(II) helicate.

P enantiomer – (ΔΔ) Right-handed helicate.

M enantiomer – (ΛΛ) Left-handed helicate.

ADP-1 – A peptide containing the active region of the tat protein.[1] (SFTTKALGISYGRKKRRQRRRPPQGSQTHQVSLSKQ)

3.1 NMR Peak Labelling

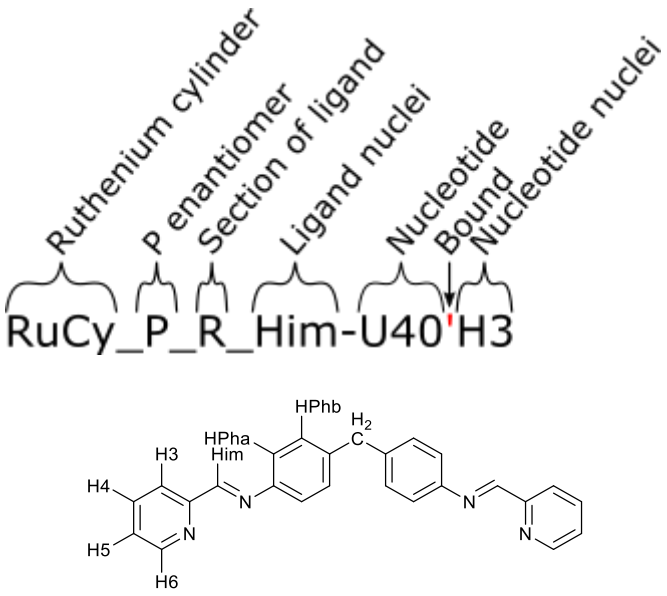


Figure 3.1-1 Top. NMR peak assignment naming convention. The example is a cylinder to RNA NOE. Bottom. The ^{1}H-NMR chemical environments for the free cylinder illustrated on the ligand (L).

There are two molecules identified within the NMR, the complex and the RNA. The complex is identified as RuCy_P meaning that it is the P enantiomer of the ruthenium cylinder. The structure of the complex has 6-fold symmetry, which when broken needs to be described. For NMR assignment these 6 regions are described using the letters R (red), B (blue), T (teal), Y (yellow), P (purple), G (green). There are also different nuclei within these regions: H3-6 (on the pyridine), Him (imine), HPha/b (phenyls), CH2 (CH_2). The RNA is identified first with the nucleotide (U40 – uracil number 40), then ‘ is used to denote if the peak is part of the bound RNA or if absent the free RNA. Finally, the nucleotide nuclei describes the position on the nucleotide (see Figure 4.1-1 for a diagram of the nucleotide). (Figure 3.1-1)

4 Introduction

4.1 RNA

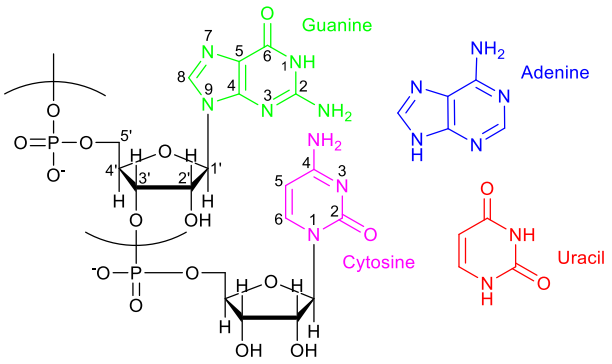


Figure 4.1-1 Molecular structure of RNA nucleotides, showing the numbering of the ribose 1'-5', as well as purine numbering 1-9 with guanine and pyrimidine numbering 1-6 with cytosine.

RNA is comprised of nucleotide units consisting of a nucleobase bound to the ribose C1' forming a nucleoside, and a phosphate group at the 5' carbon. These nucleobases are purines such as adenine and guanine denoted by A and G respectively, or pyrimidines such as uracil and cytosine denoted with U and C respectively. (Figure 4.1-1) AU and GC form Watson-Crick type base pairs which stack onto other base pairs either side to assemble into A-form RNA. RNA can form various motifs other than the A-form duplex. These include hairpin loops, kissing loops, bulges and internal loops, and pseudoknots.[2]

4.2 TAR

The trans-activation response (TAR) element is a non-coding region of the HIV-1 untranslated region (UTR). When bound to the tat protein, the trans-activation of the UTR by tat permits increased transcription of the viral RNA.[3,4] TAR is a stem-bulge-loop structure located from nucleotide position +1 to +59. HIV-1 is also able to release microRNAs derived from the TAR element, which have been found to protect infected

cells from apoptosis by downregulating certain genes.[5,6] The structure is important for dimerization, though does not do this through a kissing loop structure.[7]

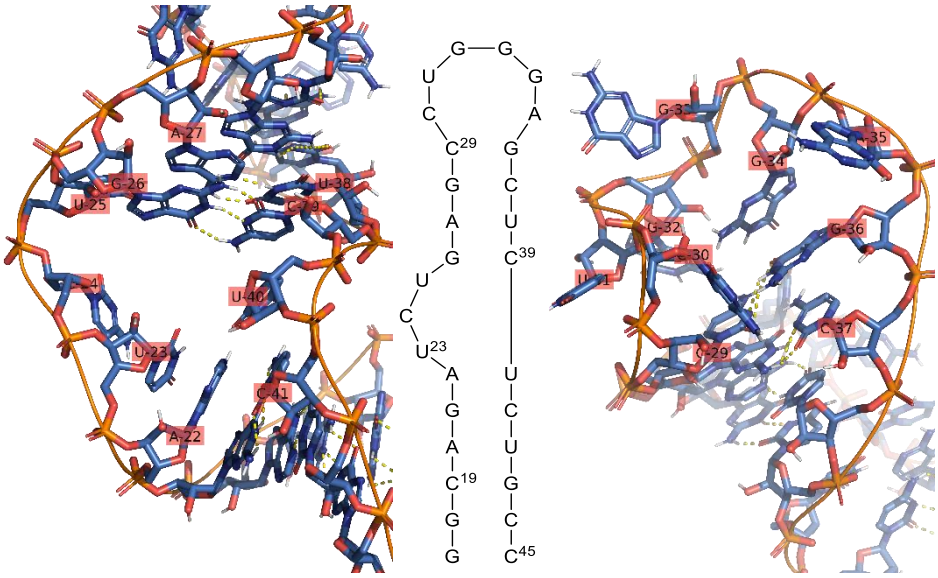


Figure 4.2-1 Left. TAR RNA bulge region. PDB: 1ANR.[8] Middle. Numbered HIV-1 TAR RNA sequence used in our experiments, showing the lower stem, bulge, upper stem, and loop. Right. TAR RNA loop region. PDB: 1ANR.[8]

The sequence used is a shortened one beginning at the 3' end with G17 and ending with C45. It has been demonstrated to assemble correctly with an NMR solution structure,[8] (Figure 4.2-1) and so was used in different experiments previously within the group. During a study into the affinity of an iron(II) cylinder for various DNA bulge sequences,[9] it was shown that the two enantiomers inhibited the ADP-1-TAR interaction (ADP-1 is the binding fragment of the tat protein). The cylinder binding at the bulge was supported with RNase A footprinting.[10] Later it was shown that racemic ruthenium(II) and nickel(II) cylinders can also inhibit the ADP-1-TAR interaction and inhibit HIV-1 infection of 1G5 and TZM-bl cells, with the ruthenium cylinder the most potent (possibly due to the complex's stability).[11] Looking at the interaction of cylinders with RNA structural features, molecular dynamics simulations were used to further understand the binding mechanism. These showed a cylinder migrating to and entering the bulge region to

bind.[12] A dinuclear polypyridyl ruthenium complex has previously been shown to bind around the TAR bulge, with NOEs to the C39H5 and U40H1' in a manner consistent with minor groove binding,[13] however the binding of the cylinder is quite different entering and residing within the bulge.[12]

Unbound TAR is a dynamic structure able to adopt a range of conformers.[14] The binding of TAR to tat is structure specific, involving both the bulge and loop. The bulge is important as it provides a few mobile nucleobases to bind, and also acts as a point of flexibility between the lower and upper stem.[15] Tat and tat derived peptides have been found to stabilise specific TAR conformers;[16–18] other binding ligands also seem to capture these accessible states providing evidence for a conformational capture model.[15,19–21]

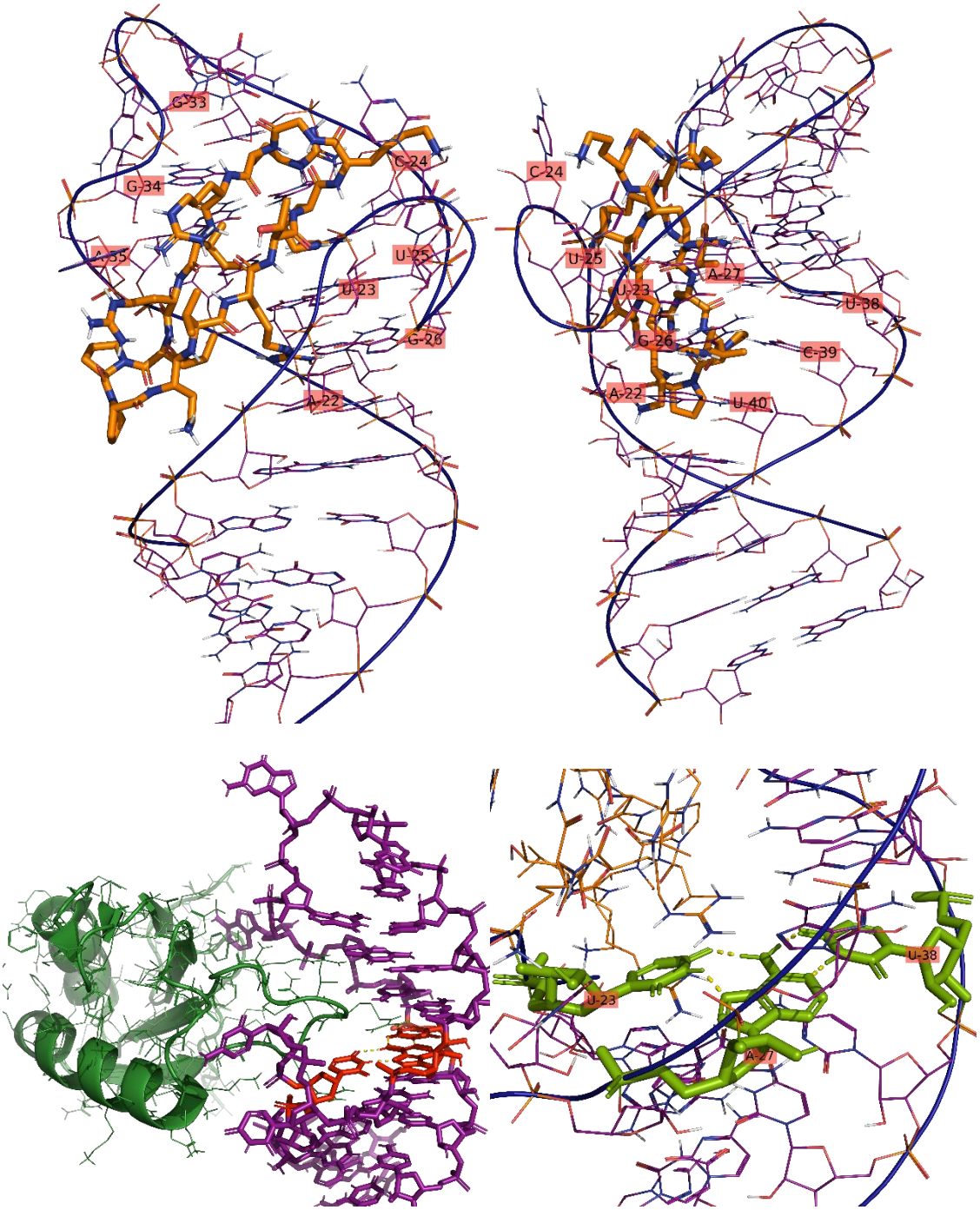


Figure 4.2-2 Bottom-left. Crystal structure of a TAR RNA binding peptide, with TAR (magenta) the base triple (red) and the RRM TBP6.7 peptide (green). PDB: 6CMN.[22] The rest are from a solution NMR structure of a TAR RNA binding peptide. Showing the major groove binding pocket (top-left) the closed bulge from the minor groove side (top-right) and the U23-A27-U38 base-triple (bottom-right). PDB: 6D2U.[23]

Originally suggested to explain NOE contacts between an argininamide ligand and U23, A22, and A27,[24] many have reported indirect evidence indicating a U23-A27-U38 base

triple structure in the bound form.[25,26] Direct evidence was found for a HIV-2 TAR base triple upon argininamide binding, in the form of $^{2}J(N,N)$ scalar couplings.[27] The base triple is reported to contribute to order in the structure,[28] and has been observed in a crystal structure of the TAR bound to a lab evolved protein.[22] Finally in 2019, direct evidence in solution was found in the form of a U23H3 peak at 14.16 ppm. This was achieved using a designed macrocyclic tat peptide mimic.[23] (Figure 4.2-2)

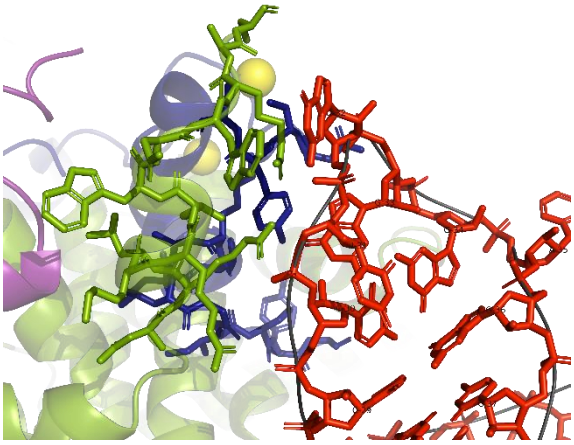


Figure 4.2-3 Crystal structure of the TAR loop in complex with Tat:AFF4:P-TEFb, showing the interactions between cyclin T1 (green) and the TAR loop (red). PDB: 6CYT.[29]

The loop is a very dynamic region required for the binding of the cyclin T1 subunit, an essential part of the positive transcription elongation factor b that complexes with tat.[29,30] The structure is stabilised by a transient cross-loop Watson-Crick base pair between C30 and G34, with G34 stacking onto G36.[31] A35 loops in and out, taking the place of G34 and stacking with G36 when facing in.[32,33] (Figure 4.2-3)

The frequency of HIV-1 TAR publications decreased in the late 2000s, but it remains a valuable target and has had a steady stream of publications in recent years. TAR is also of interest for cylinder binding as a well-researched stem-bulge-loop to study bulge/loop binding agents. This chapter aims to present the binding of a RuCy_P helicate to the TAR

stem-loop, with the intention to identify the binding location and structural changes to the TAR motif. This would provide insight into the mechanism for ADP-1 inhibition, as well as provide information for future RNA target sites this complex may be applicable for.

5 Results and discussion

5.1 Ruthenium Cylinder

The ruthenium cylinder was first synthesised and reported by Pascu et al. in 2007. At the time the procedure was reported as challenging, requiring multiple alumina columns with MeCN, H_2O, and saturated KNO_3 solution at a ratio of 20:1:1 and yielding 1%.[34] The reaction has been attempted using microwave reactor synthesis, and purification on silica gel or Sephadex C-50 resin to improved yield.[35] The cylinder has been shown to interact with DNA sequences, and has the ability to photocleave the DNA.[36]

Here the synthetic procedures tried previously are reviewed,[35,37,38] and further investigations described that lead to a relatively simple method that reproducibly provided pure ruthenium cylinder for the NMR experiments in this thesis.

5.1.1 Synthesis

5.1.1.1 Starting Materials

Both $RuCl_3 \cdot xH_2O$ and cis-$Ru(DMSO)_4Cl_2$ have been used as starting materials previously. Ruthenium chloride is an attractive starting material because it can be purchased and used without extra steps,[35] whereas cis-$Ru(DMSO)_4Cl_2$ requires synthesising from the

ruthenium chloride.[39,40] It was found that whilst ruthenium trichloride could be used for the synthesis, it was inconsistent with its results. Therefore cis-$Ru(DMSO)_4Cl_2$ was used.

The temperamental nature of ruthenium trichloride results from its two known allotropes, the black α-form, and the dark brown β-form. The β-form is converted into the $Ru(DMSO)_4Cl_2$, so in their synthesis of the cis-$Ru(DMSO)_4Cl_2$ Bratsos and Alessio first reflux ruthenium trichloride in ethanol, filtering out insoluble black solid.[39] This was not attempted, but may prove a way in future to maximise yields on ruthenium cylinder when using $RuCl_3 \cdot xH_2O$.

Synthesising cis-dichloridotetrakis(dimethylsulfoxide)ruthenium(II) from ruthenium trichloride is a straightforward reflux in DMSO.[40] The reaction yields yellow crystals that can be aided in precipitation with acetone. Samples paler yellow in colour gave greater yields of ruthenium cylinder, whilst deep yellow crystals less so. The deep yellow crystals are likely a result of partial formation of the trans isomer, which was found to form under the same conditions but at 70 °C.[41]

5.1.1.2 Ruthenium Cylinder

There is value in a reaction that can be done rapidly in the microwave. However, once set up, the standard reflux can be left with minimal interference. By contrast the microwave reaction needed monitoring, as it was susceptible to overpressure due to displaced water from the $RuCl_3 \cdot xH_2O$ and ethylene glycol. Both routes were attempted, for either method 3 equivalents of ligand was reacted with 2 equivalents of Ru. It was found that the most human intensive part with either method was the purification which could be done whilst the synthesis was repeated for multiple batches.

The ruthenium cylinder is synthesised by reacting a 2:3 ratio of cis-$Ru(DMSO)_4Cl_2$ and parent ligand in dry and degassed ethylene glycol. The mixture is refluxed under argon at 200 °C for 7 days, after which the solvent is removed under reduced pressure. Previously NH_4PF_6 has been used to isolate the complexes from the solvent,[34] though given unidentified ruthenium complexes are the main contaminant this step can be avoided. If the PF_6^- salt is required, the exchange can be conducted later.

5.1.1.3 Purification

The distilled solid is scraped from the flask walls and washed with DCM to remove neutrally charged ruthenium and organic impurities. The washings are a deep red to black colour. The solid is filtered off and kept.

A celite cellulose plug is prepared. It was found that using pipettes as mini columns was the easiest way to do this, though in theory a wide carefully assembled plug in a fritted funnel may have a similar effect. First a slurry of cellulose powder in milli-Q water was dropped into the pipette. A pipette bulb is used to provide pressure to the excess water and settle the bed. A clamp can be used to apply pressure to the bulb. The first layer of cellulose is approx. 1.5 cm. Onto the cellulose a slurry of celite is carefully pipetted, and the excess water again eluted to leave a bed of approx. 1 mm.

The solid mixture (ca. 100 mg) containing ruthenium cylinder is dissolved in milli-Q water (1 mL) and a portion (ca. 0.5 mL) pipetted onto the surface of the celite. The complex is then pushed onto the column. When washing with water the celite serves as a filter for small insoluble particulates in the solution. The cellulose method of purification is unknown, though it is known that it has the ability to interact with the cylinder, for it is

used for chiral separation. Cellulose has a chiral molecular structure, though the purification at this stage could also be influenced by the nanostructure, forming fibrils with pores ranging from macro to nano. The cylinder is immobile on cellulose until 0.2 M NaCl solution is used to elute the complex, pointing to a molecular scale effect. The eluted complex is lyophilised to an orange solid.

The cylinder is regularly desalted between purification steps as the excess sodium chloride can affect the initial mobility of the cylinder when added to a column. This is done crudely by triturating with methanol.

The cylinder is further purified using cation exchange resin in the form of CM C-50 Sephadex. The gel is first washed with milli-Q water, then eluted with increasing concentrations of NaCl solution (0.05, 0.1, 0.15, 0.2, 0.3, 0.4 M), discarding 0 - 0.15 M fractions. The complex is eluted by 0.2 - 0.3 M NaCl as an orange solution, leaving brown/red impurities in the resin. The solution is then lyophilised and again crudely desalted by triturating with methanol.

5.1.2 Epsilon

During the course of the NMR experiments described later in this chapter, it was realised that the ruthenium cylinder samples were binding only half of the TAR RNA. This could be explained by a lower concentration of ruthenium cylinder than expected. Which, given the concentration was determined by UV-Vis could only be explained by an incorrect extinction coefficient. These samples had the benefit of providing a dataset of partially bound RNA. (Section 5.4)

The easiest method for establishing an extinction coefficient is to dissolve a known mass in a known volume to a known molarity. A dilution series of UV-Vis spectra are then acquired, and the absorbance is used. However, in preparing a ruthenium cylinder sample, a lot of salts are present. NaCl is often observed when in high quantities lightening the colour of the ruthenium cylinder, and at lower quantities can visually deceive the observer to the size of a portion. It is possible to use the solubility of $[Ru_2(L)_3]Cl_4$ in methanol or other protic solvents to separate salts, however methanol no matter how dry still has solubility for NaCl, and as the salt becomes less soluble in a solvent mixture,[42] so does the complex. Other methods such as G-10 Sephadex can be used for desalting, though again removing 100% of impurities is impossible. In addition faith had been lost in measuring accurate mass, so finding an alternative method was preferred.

A solution of ruthenium cylinder of unknown concentration, with a known concentration of potassium acetate was prepared in D_2O (0.185 mM). A ^{1}H-NMR spectrum was then acquired using long relaxation delays of 15 ms.

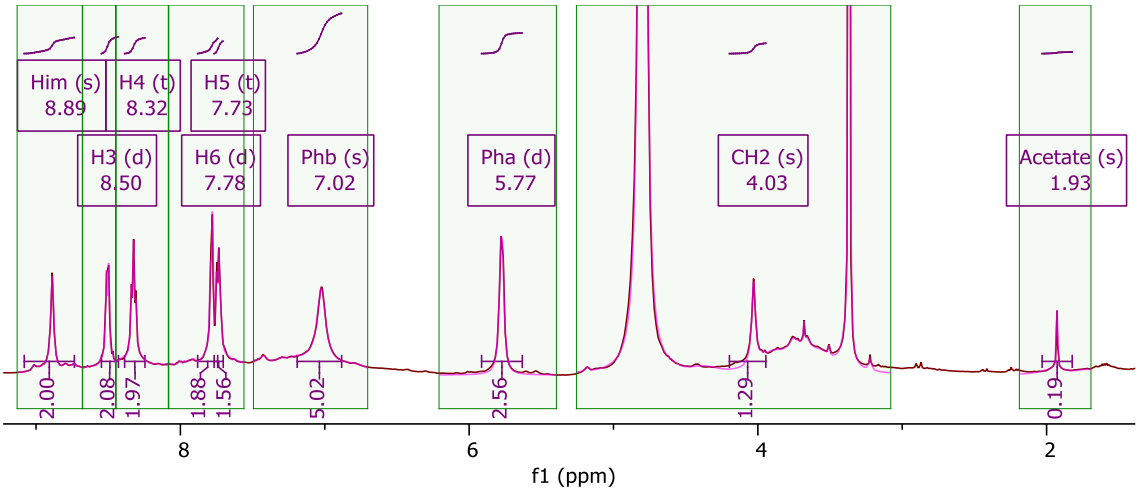


Figure 5.1-1 ^{1}H-NMR spectrum of ruthenium cylinder and potassium acetate (0.185 mM) in D2O. D1 = 15 ms. Black. The original spectrum. Purple. The fitted spectra.

The acquired dataset is phased, and the baseline corrected. Peaks are picked and regions are selected for deconvolution. Multiplets are defined and the integration values are taken. (Figure 5.1-1) Integrals are determined by a mixed Lorentzian-Gaussian equation. This is because the theoretical NMR peak is Lorentzian, but field inhomogeneity and weighting functions can yield partial gaussian line-shapes. The deconvolution tool determines the degree to which the peak is characterised by Gaussian or Lorentzian as well as varying peak width, height, and shift to best fit the acquired data.

Table 5.1.2-1 Ruthenium cylinder peaks and their integration. Whether they are included or excluded from the calculation and various values calculated.

Peak (multiplicity)	Frequency / ppm	integration	(in/ex)clude (reason)
Him	8.89 ppm	2.00	In
H3	8.50 ppm	2.08	In
H4	8.32 ppm	1.97	In
H6	7.78 ppm	1.88	In
H5	7.73 ppm	1.56	In
Phb	7.02 ppm	5.02	Ex (Broad, excessively high integration)
Pha	5.77 ppm	2.56	Ex (Strong proximal water peak, excessively low integration)
CH2	4.03 ppm	1.29	Ex (Large overlap with ethylene, strong proximal water peak)
Mean (Included)		1.90	
Acetate (s)	1.93 ppm	0.19	0.185 mM
(RuCy/6H)/(Acetate/3H)		4.99	
RuCy concentration		0.925 mM	

The peaks are integrated and a ratio of complex to acetate is used to determine the concentration of the ruthenium cylinder in solution, taking into consideration the number of nuclei for each compound. (

Table 5.1.2-1)

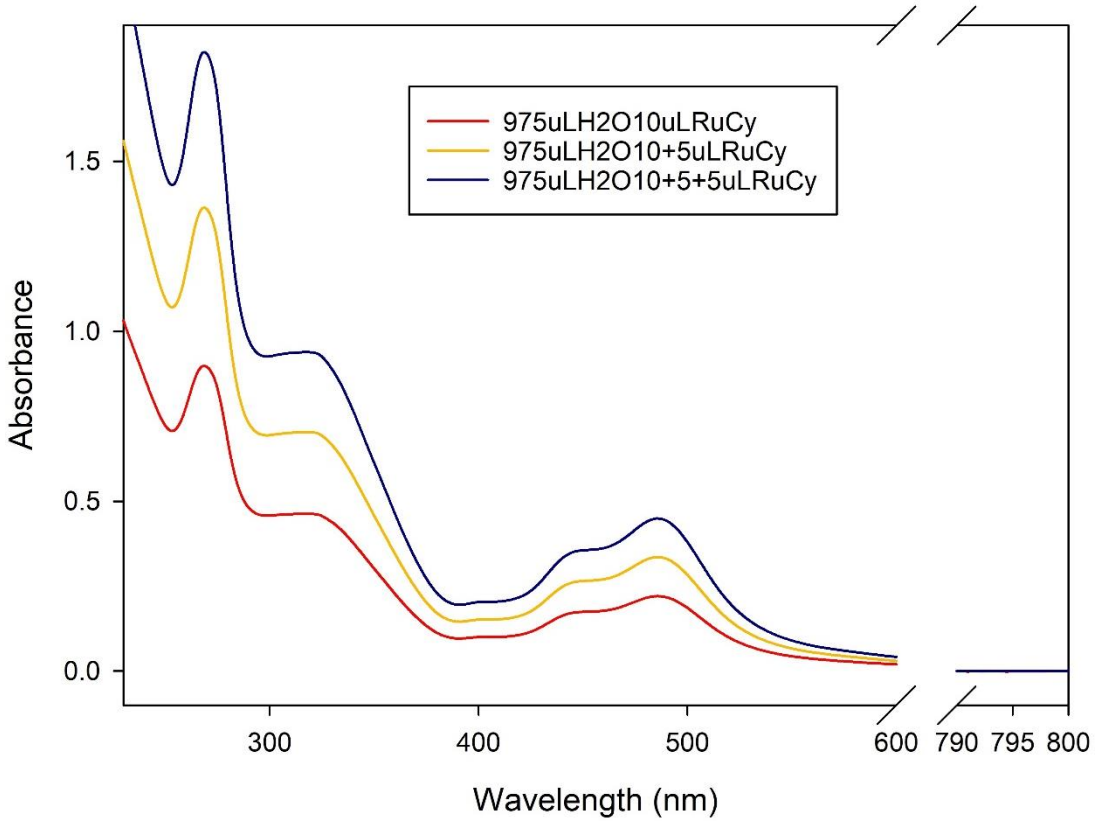


Figure 5.1-2 UV-Vis spectra of portions of ruthenium cylinder (0.63 mM) taken from the NMR experiment in water.

UV-Vis spectra are acquired, and the concentration is used to determine the extinction coefficient at different frequencies. (Figure 5.1-2) Errors are calculated as a combination of estimated experimental error, NMR S/N, NMR standard deviation, and UV-Vis standard deviation. The NMR standard deviation contributes about 98% of the total error, which is understandable considering the variation between values. Calculated extinction coefficients and their respective error are as follows.

ε_{486} = 23871 ± 2589, ε_{449} = 18849 ± 2045, ε_{318} = 50019 ± 5420, ε_{271} = 96982 ± 10506.

Errors are large, ca. ± 11%, so this value has limitations for more concentration dependent experiments. However, determining the difference between 1:1 and 1:0.5 is

within error, so these values can be used to determine the concentration of a RuCy sample for use in this thesis. As expected from the TAR RNA data, the extinction coefficient is greater than previously measured: 485 (16900),[34] and 492 (7939).[38]

5.1.3 Enantiomers

5.1.3.1 Enantioseparation

Enantioseparation is required for applications with complex nucleic acid systems. Each enantiomer has the potential for a different binding mechanism increasing the complexity of an NMR spectrum.

A method for the separation of enantiomers was previously developed within the group using the cellulose in filter paper as the stationary phase.[43] Since then, the method was further refined to use cellulose powder column chromatography.

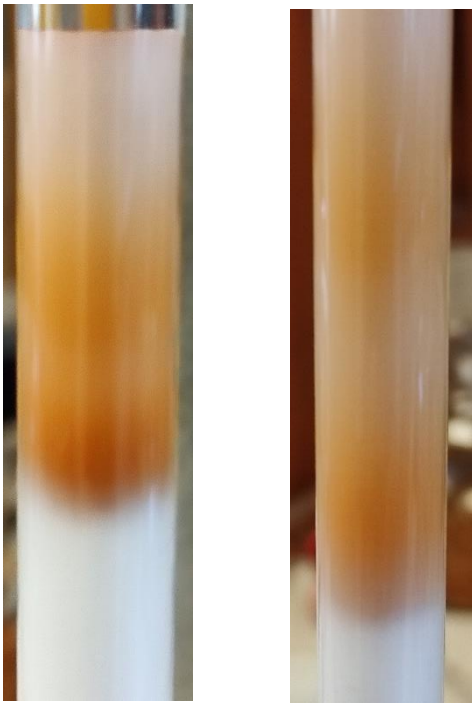


Figure 5.1-3 Images of a cellulose enantioseparation column. Left. First band has travelled approx. 1.5 cm down the column, the bands are beginning to separate. Right. After more time the bands have separated further.

First a cellulose slurry is prepared with 0.2 M NaCl solution. It is likely pore and particle sizes vary between portions of cellulose powder as previously columns have been as long as 25 cm in length.[35] When longer columns were tried, the pressures required and flow rates were prohibitive. It was found that a shorter column was more efficient and easier to pump, and that good separation of the bands was achieved after approx. 5 cm. Further, it was found that using a pipette as a disposable column was simpler and sufficient pressure could be applied through the pipette bulb. Generally, pipette columns of 7 cm were used. The racemic ruthenium cylinder is dissolved in 0.2 M NaCl solution and added to the top of the column where it is carefully pushed onto the column and followed up with a drop of 0.2 M NaCl solution to ensure none remains on the glass. The ruthenium cylinder is eluted using 0.2 M NaCl solution, the first band M>P, the second P>M. About 7 fractions are collected the head, body and tail of each band

and another in between the bands. The fractions are analysed using circular dichroism (CD) to sort into degrees of enantiopurity, then lyophilised.

Attempts were made at developing a chiral HPLC method. Phenomenex Lux Cellulose and Amylose chiral HPLC columns were tried using 0.2 M NaCl solution as the eluent, but these were unsuccessful as the complex was not observed eluting from the column. The detector used was UV-Vis, perhaps this experiment could be repeated using a MS detector.

5.1.3.2 Circular Dichroism (CD)

UV-Vis and CD spectra were acquired for each fraction.

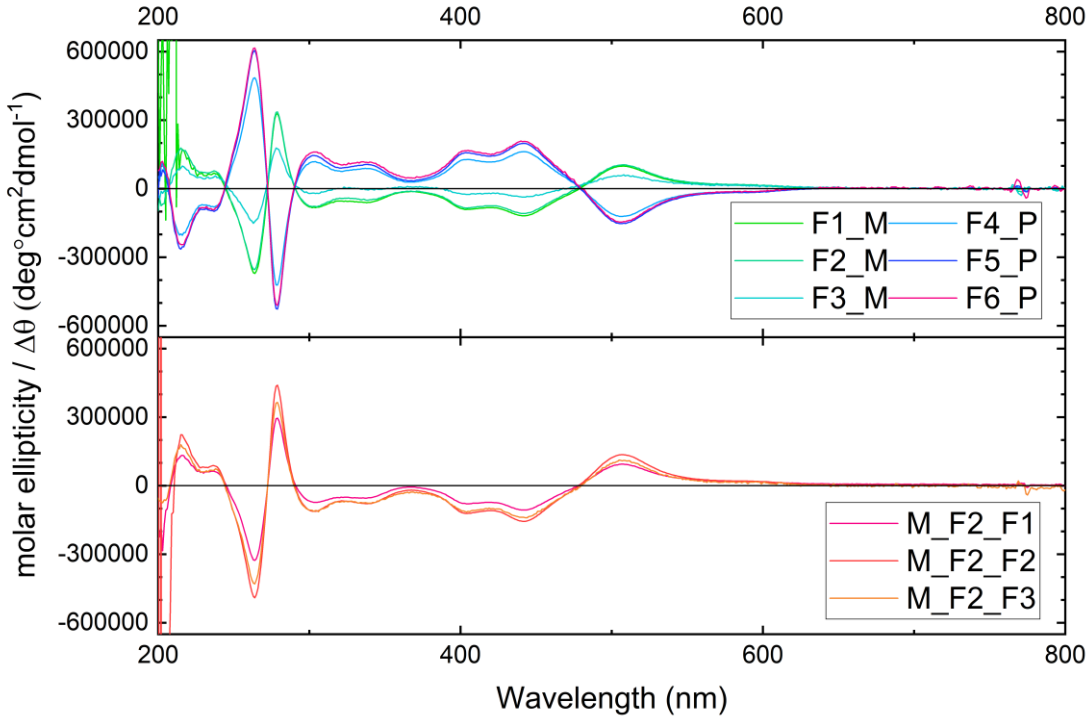


Figure 5.1-4 Top. CD spectra of fractions from an enantioseparation column, Bottom. CD spectra of fractions from a further separation of fraction 2 from the spectra above. Wavelength Vs molar elipticity, calculated from

elipticity in millidegrees and the concentration of the sample calculated from the absorbance spectrum (Figure 8.4-3).

As multiple batches of ruthenium cylinder enantiomers were prepared, a pattern emerged where the P enantiomer consistently provided a stronger signal. When the ruthenium cylinder passes through a cellulose column, the M enantiomer is the first band, the P follows. Often the P enantiomer needs a further column to remove trace M enantiomer from the tail of the first band. It was then theorised that the M enantiomer was, like the P enantiomer, containing small portions of the opposing enantiomer. A strong M fraction was columned again but resulted in little improvement to enantiopurity. M_F2_F1 is not significantly lower in enantiopurity than M_F2_F2 showing that the entire M band is affected, and not just the front of the band. (Figure 5.1-4)

5.1.3.2.1 Finding the Purest Enantiomer.

Realising that there was a discrepancy between CD spectra for the M and P enantiomers it was important to identify why. An experiment was devised using Δ-TRISPHAT, an anionic chiral shift reagent, in the same manner as previously established.[35,37,44] However this time to determine the relationship between peak intensity and circular dichroism and therefore in the process identify a target molar ellipticity for an enantiopure sample. (Figure 5.1-5) Once ^{1}H-NMR is acquired, the samples are exchanged to the chloride salt using Dowex 1x8 Cl^-. These samples are then dissolved in water, for UV-Vis and CD acquisition.

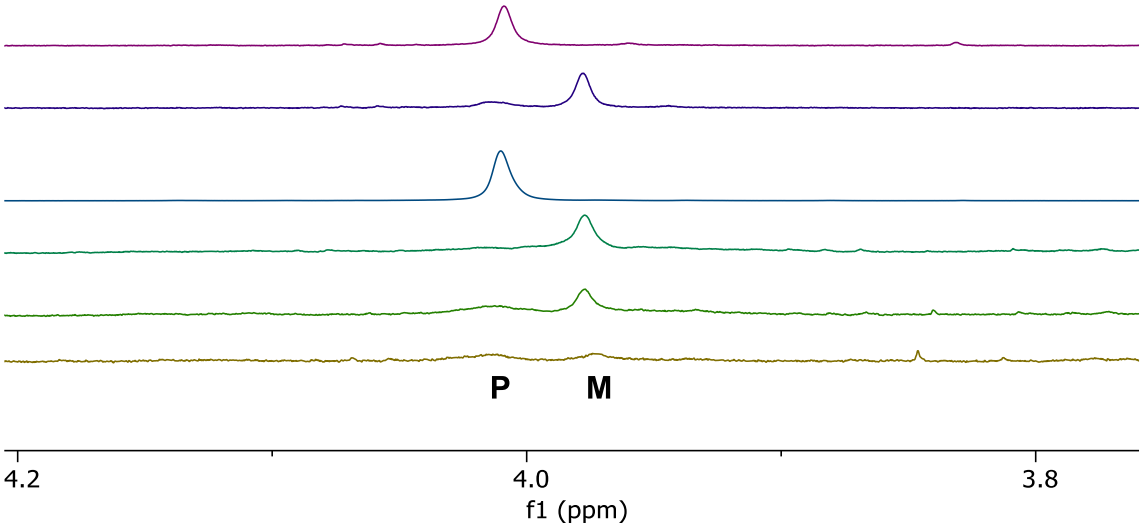


Figure 5.1-5 A selection of ruthenium cylinder spectra of varying enantiopurity in Acetonitrile-d3. Showing the CH2 peak in the presence of Δ-TRISPHAT.

The molar ellipticities at 445 nm and 507 nm were plotted vs the NMR peak integration ratio (P-M) and linear fits were applied. The fits both pass through the axis at the same point, illustrating that the CD spectra are the same for both enantiomer samples. (Figure 5.1-6) This issue is not present in the iron cylinder (see Chapter 4, Appendices, Figure 7.1-1) A few samples were measured in a different institution to reduce both experimental bias and eliminate instrumental issues. The trend is the trend is the same, the M enantiomer samples give a weaker molar ellipticity.

These showed the experimental technique and apparatus were appropriate. Finally, samples were passed through multiple cellulose columns to prepare very enantiopure samples, these are plotted with assumed ratios of 1 or -1 for P and M respectively. For this data, the P enantiomers seem to fit with the linear trends for the NMR data, clustering around the X=1,-1 intercepts. The M enantiomers however are consistently below what would be expected for an enantiopure sample.

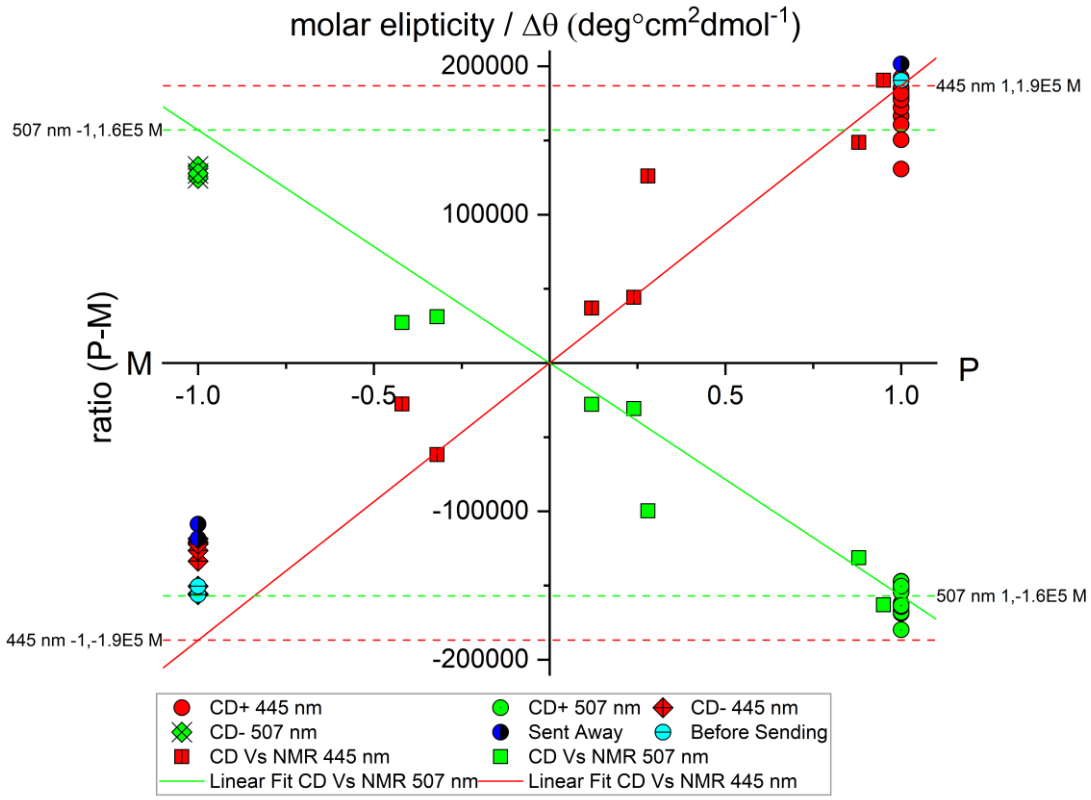


Figure 5.1-6 A graph of molar elipticity (deg°cm²dmol⁻¹) Vs NMR peak ratio (P-M, real or imagined). Datapoints for two frequencies are plotted, 445 nm (red), and 507 (green). Squares show the real NMR molar elipticity datapoints. Linear fits are plotted to only the NMR data, with an intercept of 0. Red and green, circles and diamonds show assumed 100% enantiopure samples. Blue circles are datapoints acquired externally, with the datapoints acquired before sending the samples as turquoise circles. Horizontal dashed bars show the expected molar elipticity at a given frequency for an enantiopure sample.

It is likely that the cause of this discrepancy is some form of impurity in the sample. One that is consistently present in multiple syntheses and enantioseparations. Identifying the impurity would be a valuable step in understanding the preparation of this complex.

In order to travel at the same rate as the M enantiomer through a cellulose column an impurity would likely be similar in structure to the M enantiomer. Further, as the concentration is determined using the MLCT band of the UV-Vis spectrum the impurity would need to have a similar absorption band to affect the determined concentration.

Considering that the P enantiomer interacts with the cellulose more strongly as seen by the slower rate, it would then follow that the M enantiomer interacts less with the cellulose. A species similar in structure to the M enantiomer but unable to interact with the chiral interface to the same degree as the P enantiomer, could potentially be a mesocate structure. The mesocate has not been seen before in triple helicate cylinders with these ligands, however it has been seen in other double and triple helicates.[45–47] The reason the mesocate has not been seen before in other research on the ruthenium cylinder and other triple helicates could be that it simply has not been present. That the same effect has been observed in multiple batches in this study starting with different purchases of $RuCl_3 \cdot xH_2O$ suggests it may have been.

Alternatively, it has not been noticed. Looking through previously published CD spectra, they are reported in mdeg, a concentration dependent unit. It would therefore be simple for this to go unnoticed.[35] One reports the concentration of two samples as the same, however in this case the M enantiomer sample has slightly greater ellipticity.[37] In either case, the quantity of data presented is not enough to draw conclusions.

It cannot be said for certain whether a mesocate has been observed here, but it does seem promising enough to recommend further research. To purify the mesocate from the M enantiomer it would be necessary to identify a stationary phase that can differentiate the two. Given the P enantiomer is slowed by its interactions with the cellulose in a column, it might be worth trying a different chiral polysaccharide. Perhaps amylose or a modified cellulose.

5.2 Free TAR

The free TAR was prepared via transcription with T7 polymerase. The procedure is detailed in section 6.2. Initially it is important to verify the assembly of the TAR structure, following this to develop an idea of conformation and a free TAR dataset to compare to the bound TAR.

5.2.1 90/10 H_2O to D_2O Spectra

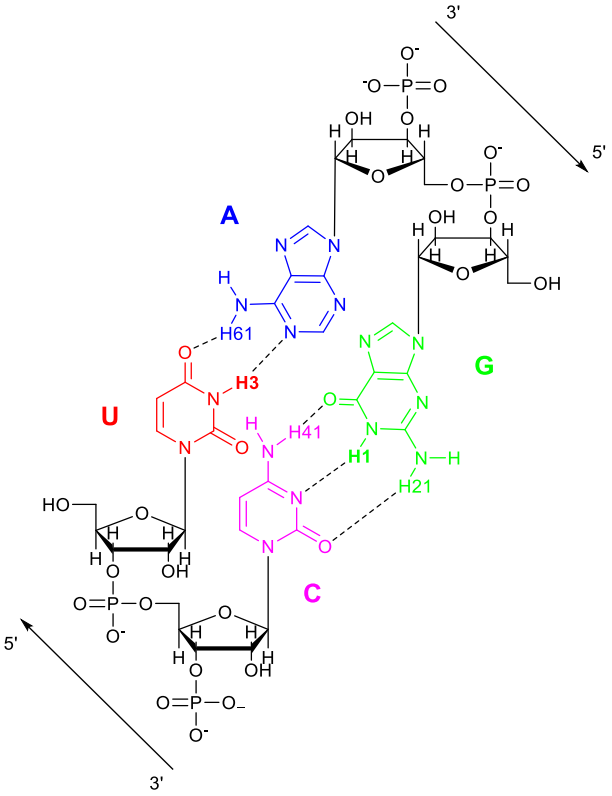


Figure 5.2-1 AU and GC base pairs, showing the exchangable imine and amine hydrogen bonding protons.

The hydrogen bonding between the bases serves to localise hydrogen atoms for visualisation. These hydrogen nuclei are exchangeable and in a protic deuterated solvent will not be visible. Using a solution of 9:1 H2O/D2O we're able to populate these nuclei enough to visualise them. These are the imines and amines. (Figure 5.2-1)

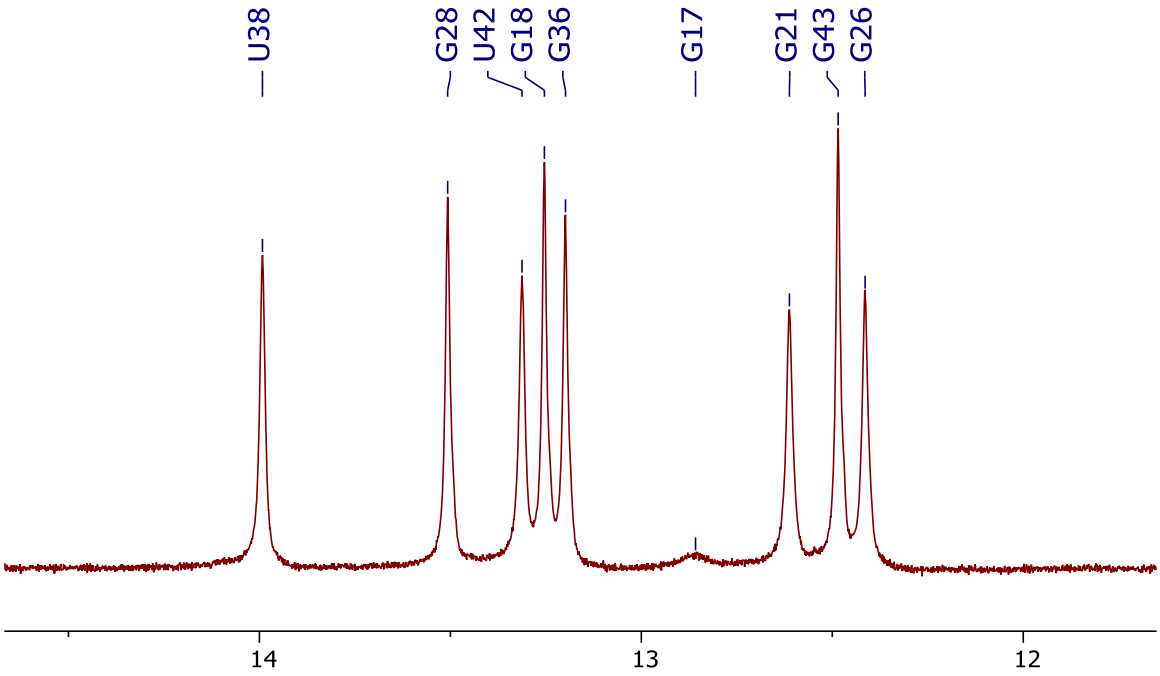


Figure 5.2-2 ^{1}H-NMR spectra of 1.5 mM TAR RNA in 90% H2O. pH 7, 50 mM NaCl, 298K, 700 MHz.

Around 13 ppm, are the hydrogen bonded imine N-H1 and N-H3 peaks between the RNA bases. These give an indication of the base-pairing in the stems. As these are dissociative, their intensity is related to the stability of the hydrogen bond. Here we see strong peaks for G36, G28, U38, G26 showing good assembly of the loop-stem-bulge region, and G21, U42, G43, G18 for the bulge-stem-end. At 298K the G17 H1 hydrogen bond is dissociating as the end frays, giving a weak signal. The absence of a U40 H3 points to poor base-pairing at A22-U40, likely resulting in an opened bulge structure, and contributing to flexibility between the helices. The assignment of these is explained in the following paragraphs. (Figure 5.2-2)

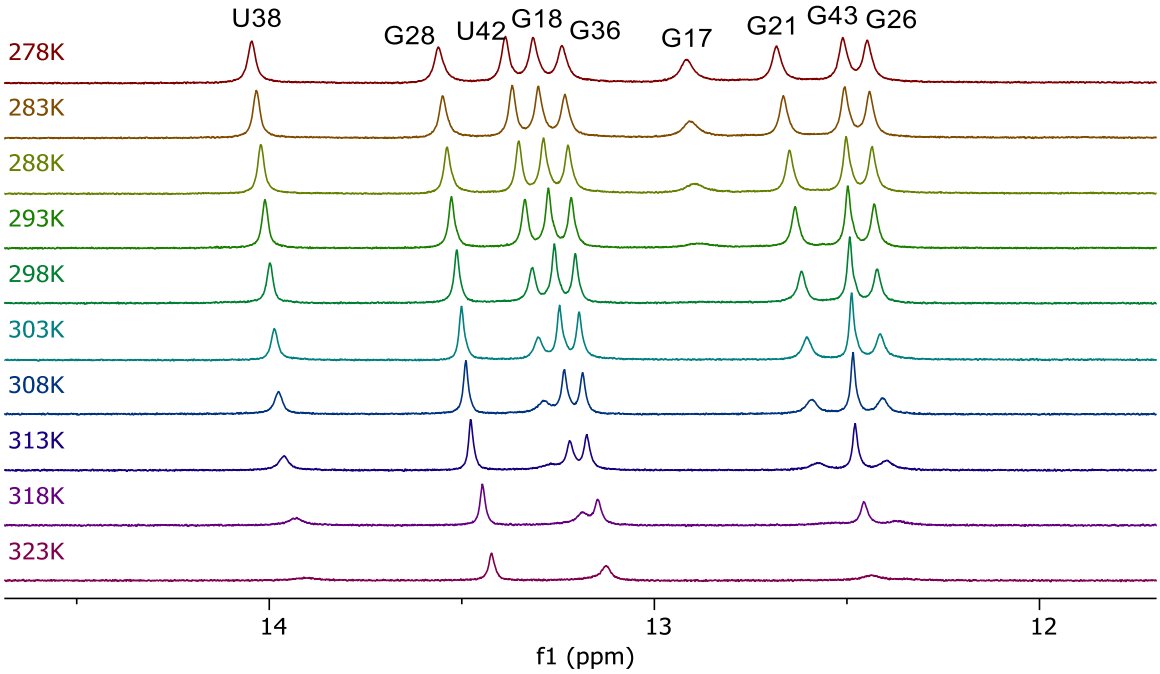


Figure 5.2-3 VT ^{1}H-NMR spectra of 1.5 mM TAR RNA in 90% H2O. pH 7, 50 mM NaCl. Showing the change in the imino region over a range of temperatures between 278K and 323K, 600 MHz.

With increasing temperature, we see a greater degree of fraying, as well as a weakening of the bonds near to bulges and loops, as the RNA tends toward single stranded at higher temperatures. Using a variable temperature titration, we see the first bonds to dissociate are the G17 at the end, and the G21 and U42 nearer the bulge. G21 and U42 are already weakened as a result of the poor A22-U40 pairing. Still visible are the G28, G36, and U38 in a more protected location between the bulge and loop, as well as G43 between the bulge and the frayed end. With so few AU base pairs it is hard to compare their relative stability to the generally more stable GC pair.[48]

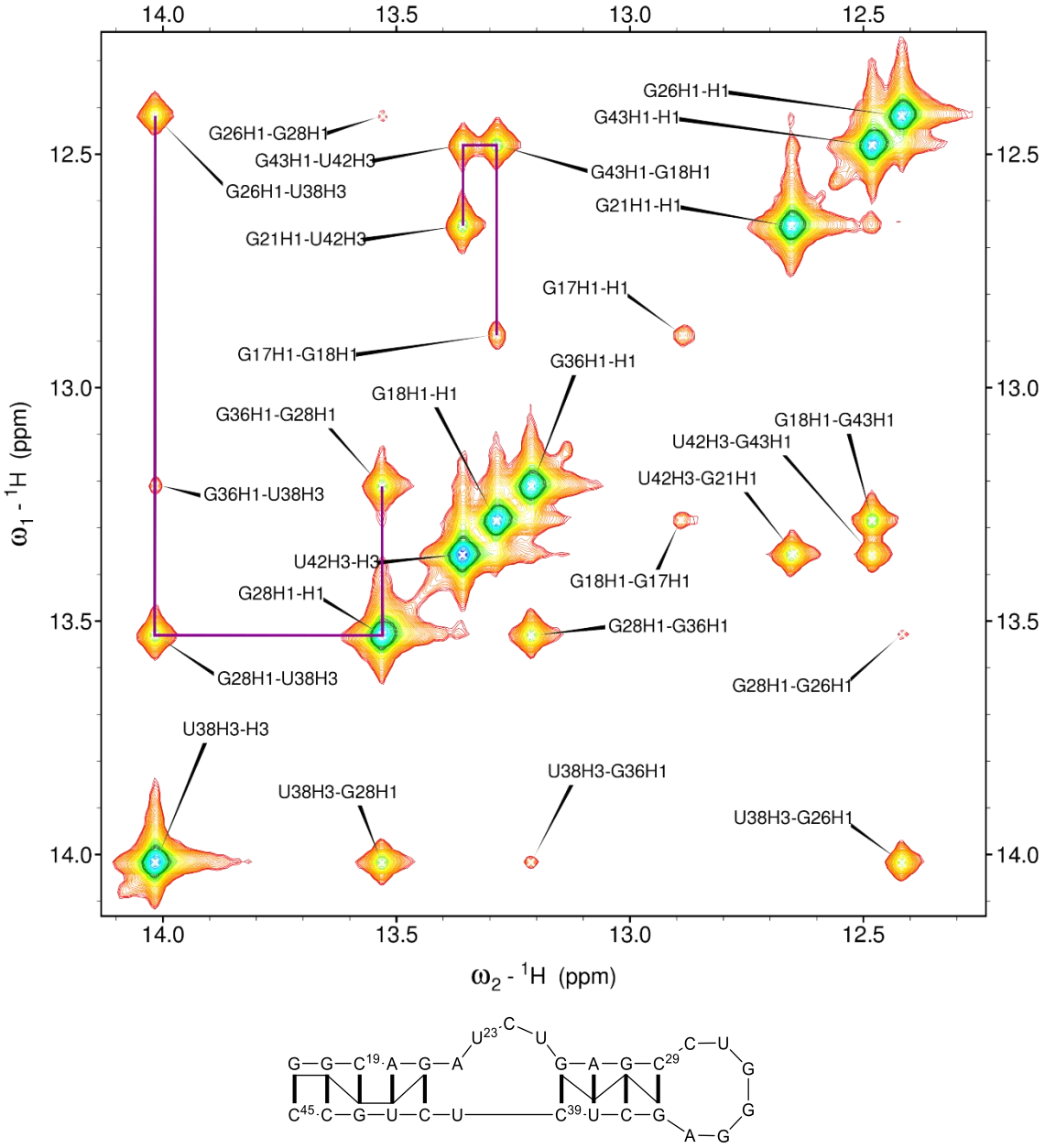


Figure 5.2-4 Top. ^{1}H, ^{1}H-NOESY spectra of 1.5 mM TAR RNA in 90% H2O. pH 7, 50 mM NaCl, 278K, 700 MHz. d8 = 350 ms. Showing the imino region, lines indicate the walk. Bottom. Illustration of free TAR showing the imine walk.

Interconnectivity in the form of NOEs between imines can be used to glean more structural information and following the sequence of NOE crosspeaks aids in their assignment. Beginning with two imine-H at the base of the stem connecting G17-C45N3 to G18-C44; we see the crosspeak G17H1-G18H1, following this sequence

through G43H1-G18H1, G43H1-U42H3 and G21H1-U42H3 we arrive at the G21H1 at the base of the bulge. As the RNA becomes more mobile at the fraying stem, we see lesser intensity in the G17H1 and a weak G18H1-G17H1 NOE. Above the bulge at G26H1 we can follow the G26H1-U38H3, G28H1-U38H3 and G36H1-G28H1 to G36H1 at the base of the loop. Longer range G36H1-U38H3 and G26H1-G28H1 are present, perhaps as a result of strong stacking in the upper stem. This spectrum confirms the identity of the imine peaks, and the correct assembly of the TAR structure. (Figure 5.2-4)

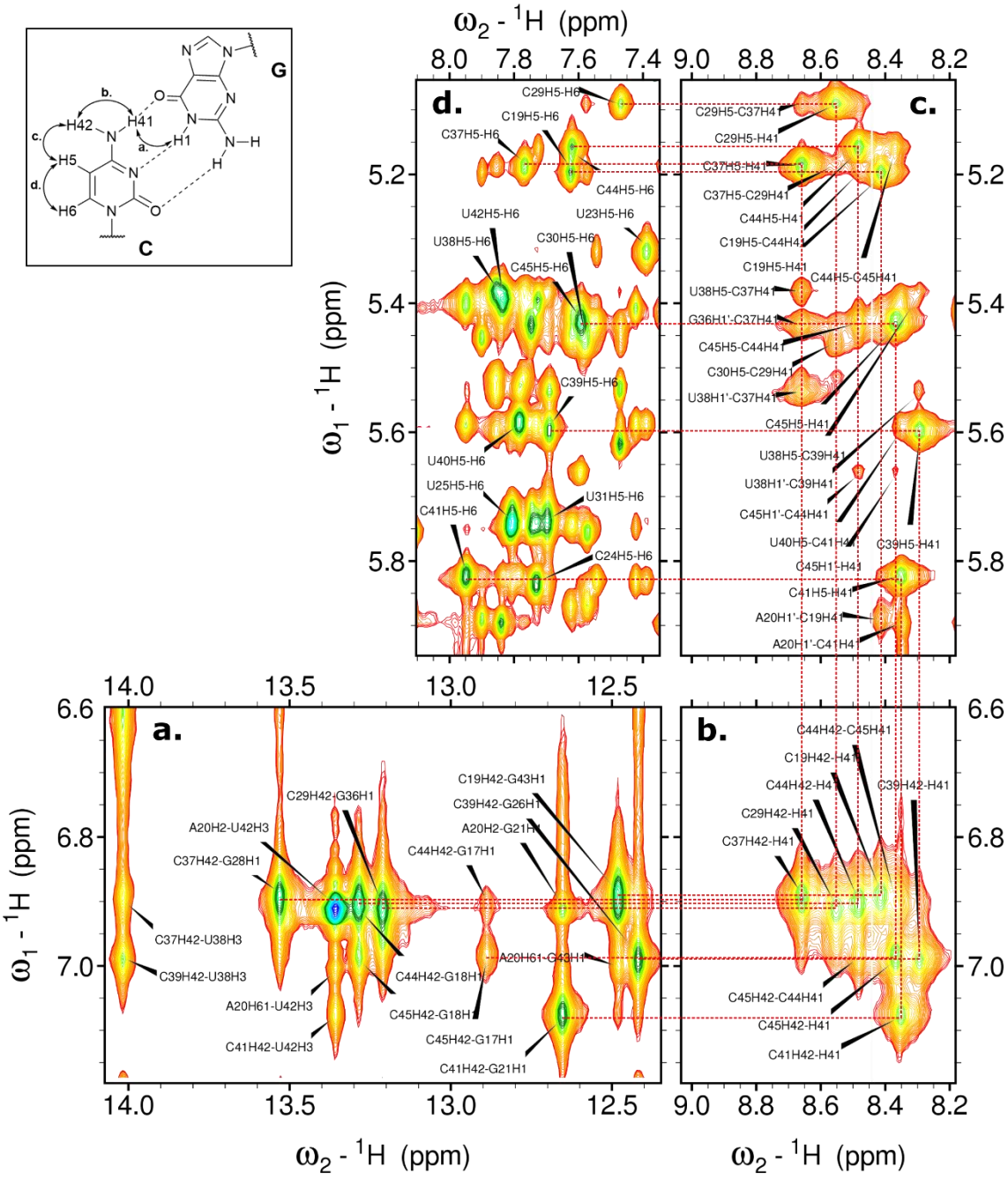


Figure 5.2-5 Top left. Location of the NOEs illustrated, labelled a-d. Main. ^{1}H,^{1}H-NOESY spectra of 1.5 mM TAR RNA in 90% H2O. pH 7, 50 mM NaCl, 278K, 700 MHz. Showing cross-peaks imine-amine (a.), amine-amine (b.), amine-H5, (c.), H5-H6 (d.).

Identifying the H5-H6 peaks is important for assigning the D2O spectra, the CH5-H6 can be assigned using the corresponding CH5-H41 peak. These are identified from the CH41-

H42 peaks in a distinctive cluster at, 8.1-8.8 by 6.7-7.0 ppm which are in-turn identified from the imine-H42 peaks. This process is unique to the cytosine, owing to the proximity of the CH5 to the CH41 as uracil possesses carbonyl groups at the 4 position. (Figure 5.2-5)

5.2.2 100% D_2O Spectra

Transferring these non-exchangeable assignments across from 90:10 H2O/D2O NOESY to 100% D2O samples is then done. For some this is simple and, in this case, the C37H5-H6, C44H5-H6, C19H5-H6 and C29H5-H6 at the top were isolated enough that these could be assigned with certainty. The rest were assigned as part of the H8/6-H1' sequential walk that followed and then confirmed by cross-referencing back to the 90% water NOESY, and data from other experiments such as TOCSY/HSQC. (Figure 5.2-8)

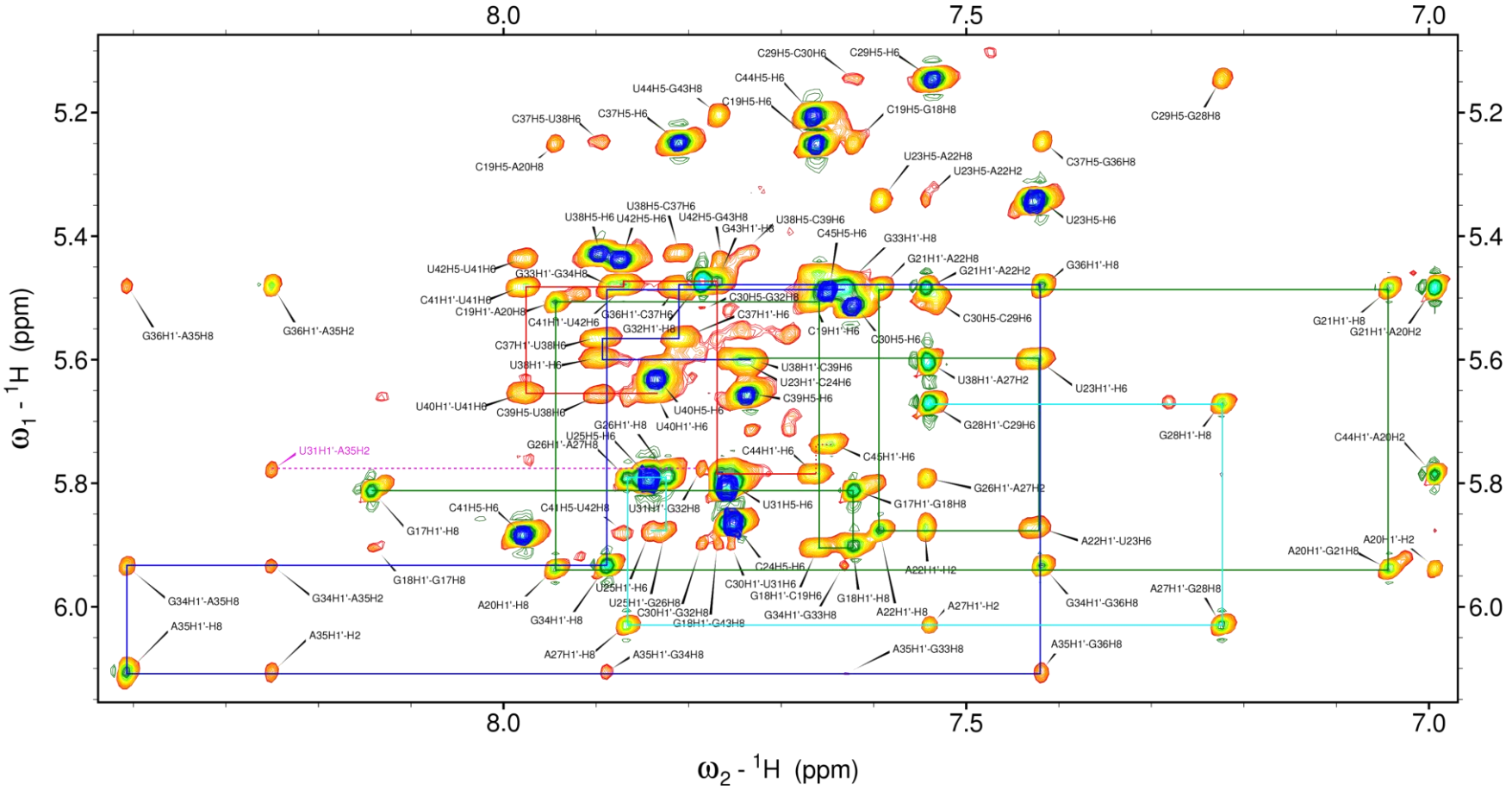


Figure 5.2-6 ^{1}H,^{1}H-NOESY spectrum of 1.5 mM TAR RNA in D2O at 298K, 700 MHz. pH 7, 50 mM NaCl. H8/H6/H2 - H5/H1' peak region. NOESY TAR RNA intensity gradient orange-turquoise overlayed with TOCSY TAR RNA 1.5 mM D2O in blue. Solid lines H1'-H8/6 sequential walk. The purple label indicates a potential cross-loop NOE.

Beginning the sequential walk at the 3' end, the G17H1' has NOE's to the G17 and G18H8. Following the sequence through C19, A20, G21, A22, U23H1'-H6 to U23H1'-C24H6 as there is no C24H1'-H6. This non-connectivity is unsurprising as the bulge is highly flexible and reported as a site of mobility between the stems.[14,49] Slightly different are reports that the sequential walk is interrupted at U25,[8,50] whereas here the loss of connectivity begins at C24 and ends at U25. A lack of inter-base H5-H6 NOE peaks for U23-C24 and C24-U25 suggest limited stacking between these bases. Considering the mobility at the bulge, this is expected. U25H1'-H6 begins the next segment of the sequence through G26, A27, G28 finishing at G28H1'-C29H6 at the beginning of the loop. An NOE across the base pair with U38H1'-A27H2 demonstrates the stability of the upper stem A-form. Resuming with G33H1'-H8 through G34 with both G33H1'-G34H8 and a weak G34H1'-G33H8, followed by A35, G36, C37, U38, and C39 stopping with U38H1'-C39H6 and no connectivity visible between C39 and U40. Given C39 and U40 are unstacked, and the connecting atoms are the axis of mobility between stems, this lack of NOEs is anticipated.

Also present is U31H1'-A35H2, potentially indicative of a cross loop pairing. Previously reported is a C30-G34 base pair, it is possible we'd see this NOE if the A35 was stacked between the G34 and G36 or if there is a triplex structure between the Hoogsteen edge of the A35 and the sugar edges of the G34 and C30. This is not too unreasonable considering the dynamic nature of the loop and that similar stacked conformations have been seen as excited states.[51] Though the G34H1'-G36H8 points to a looped out A35.[52] Alternatively, it is possible there is a U31-A35 pairing, bringing the nuclei close enough for the visible NOE, though this is less likely. (Figure 5.2-6)

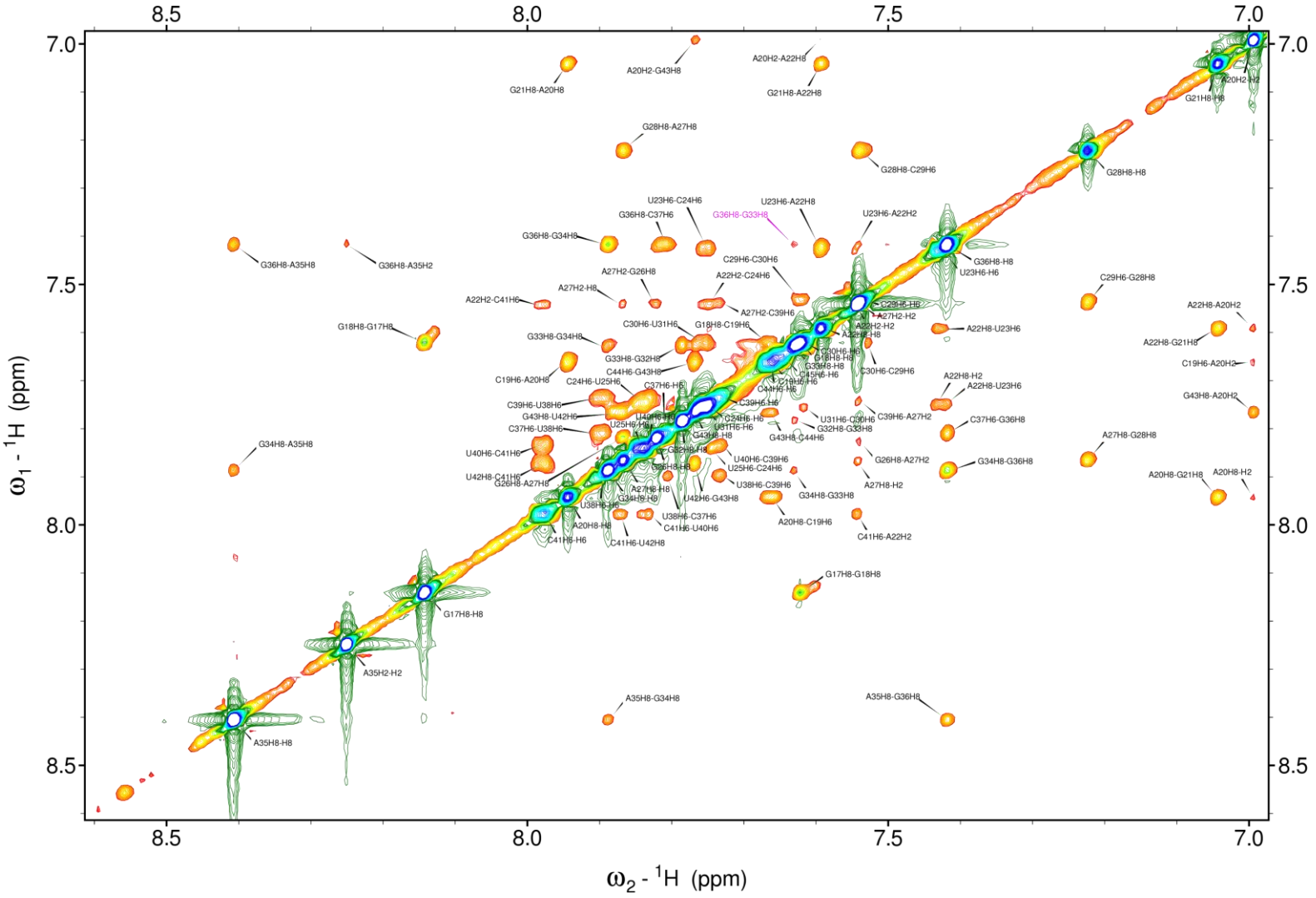


Figure 5.2-7 ^{1}H,^{1}H-NOESY spectrum of 1.5 mM TAR RNA in D2O at 298K, 700 MHz. pH 7, 50 mM NaCl. Aromatic-aromatic peak region. Purple label indicates a cross-loop NOE.

Further cross-loop interconnectivity is found with the G36H8-G33H8, shown in purple. In the ground state G33 has been shown to move into proximity of the G34, which is stacked onto the G36.[33] This is a potential method through which these bases could come into contact. The loop bases show connectivity across G36H8-A35H8, G34H8-A35H8, G33H8-G34H8, G33H8-G32H8 with no connectivity between G32 and U31. This suggests that the U31 is pointed out of the loop. The connectivity resumes with C30H6-U31H6 and C29H6-30H6. Other than the G36H8-A35H2, there are no aromatic NOEs to the A35H2, suggesting it is frequently out of alignment with the other bases. These two bases demonstrating points of reduced stacking align with the possibility of a C30-G34 cross-loop pairing.

This region also helped with the assignment across the bulge, showing aromatic peaks between U23H6-A22H8, U23H6-C24H6, C24H6-U25H6. These connectivity's are also indicative of base stacking in the bulge bases. (Figure 5.2-7)

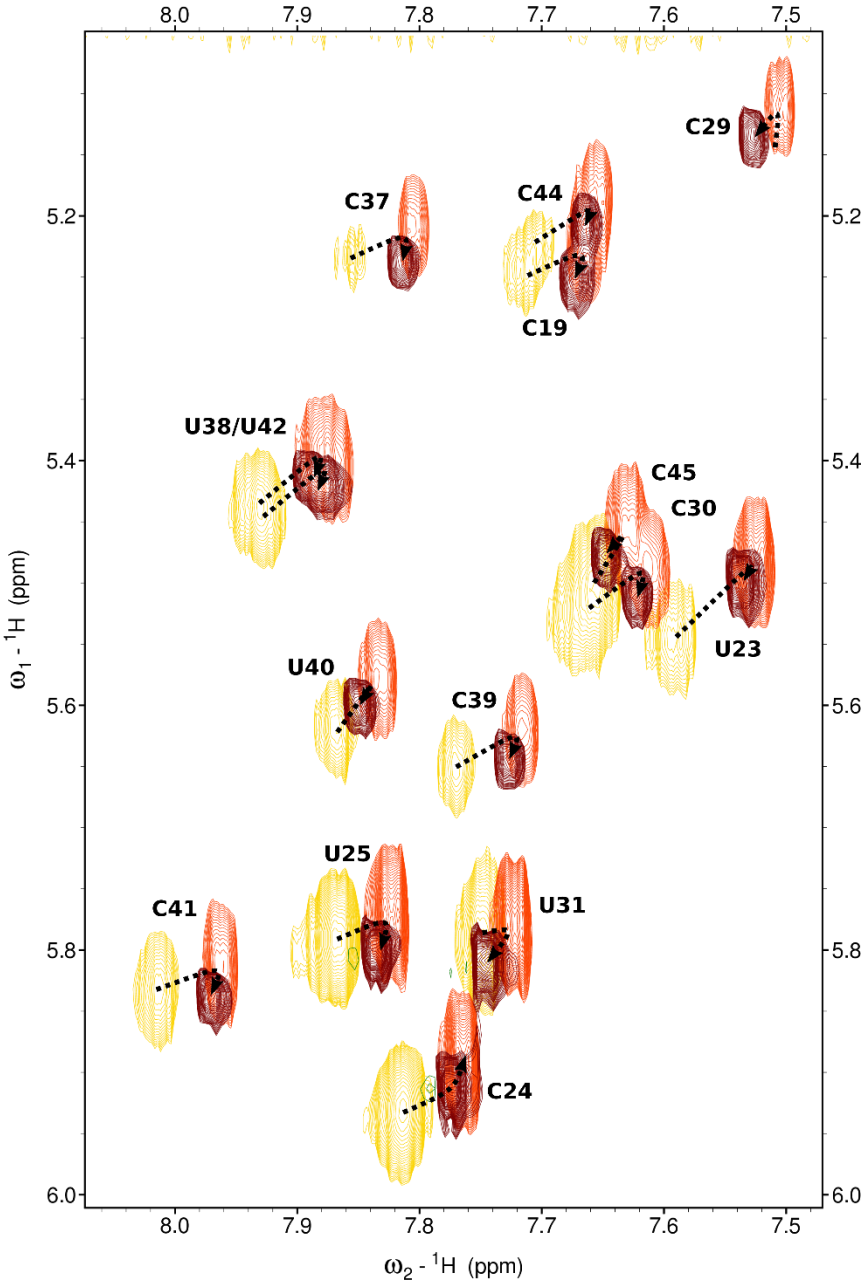


Figure 5.2-8 278K, 293K, 303K ^{1}H, ^{1}H-TOCSY overlay, 700 MHz. Arrows indicate the shift at higher temperatures.

Pyrimidine H5-H6 J coupling is visible using TOCSY. By varying temperatures whilst looking at the H5-H6 coupling we can glean an idea of which bases are most affected by the change. In this case the U23, and C24 are more shifted. The C29 appears to be less visible at low temperatures, possibly due to weakened pairing at the top of the loop, alternatively the peak could be affected by the nearby water. (Figure 5.2-8)

Considering the continuity of the H1'-H8/6 sequential walk, the stem portions are A-form and the locations where it is disrupted are expected, the bulge, loop, and end. The unexpected U31H1'-A35H2 is acceptable considering the freedom of movement available to the loop. The data for the free TAR RNA confirms the structure is assembled correctly, closely matches the previously reported data,[1,8,32,49] and provides a baseline for bound TAR experiments.

5.3 Bound TAR

Binding the TAR RNA is theorised to be a conformational capture model, so it can be expected to assemble a similar structure to other binding molecules. Samples are prepared at a mixture of 1:1 cylinder:TAR and are annealed before measurement. Further details are in section 6.3.

5.3.1 The Cylinder

The ruthenium cylinder has 3 symmetric ligands, resulting in 6 equivalent pyridyl and aniline, and 3 equivalent CH2 groups. Unbound this results in a spectrum of 8 distinct chemical environments.

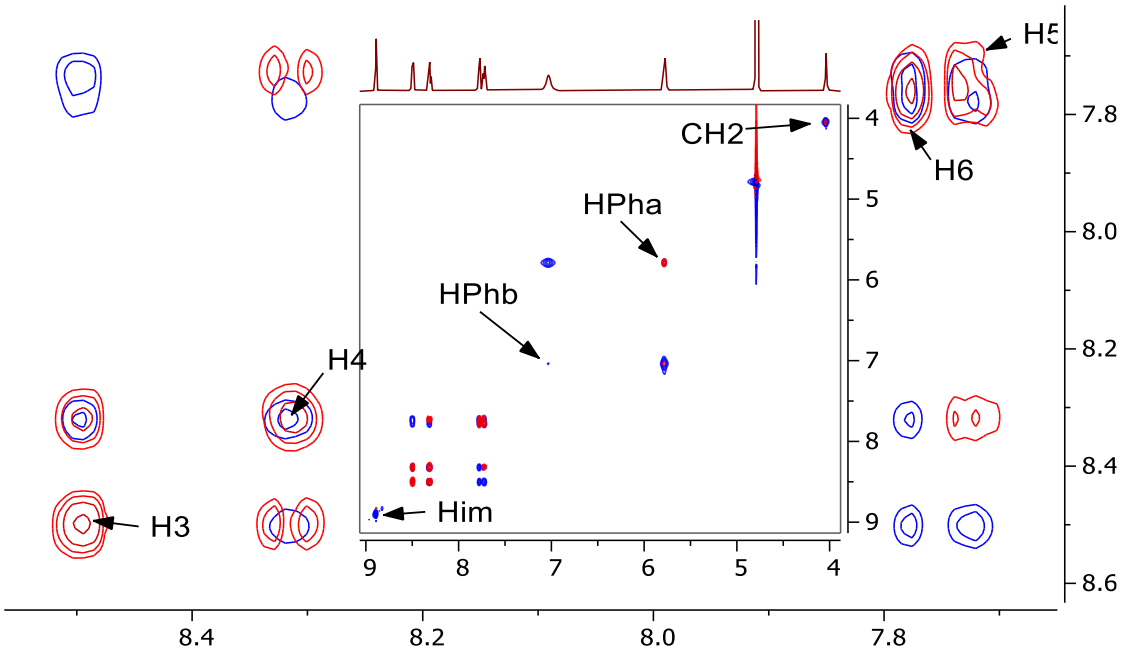


Figure 5.3-1 COSY (red), TOCSY (blue) overlay of ruthenium cylinder in D2O, 298K, 500 MHz, showing a closeup of the pyridine protons. Inset shows the full spectrum. Arrows label the cylinder peaks.

HPha and HPhb are broadened due to the rotatability of the phenyl groups. B is broadened more-so. The H5, is at a higher frequency to the H6, identified through the COSY crosspeak to H4. All peaks can be confirmed with NOESY, (^{1}H,^{13}C)-HSQC, (^{1}H,^{13}C)-HMBC. (Figure 8.4-1, Figure 8.4-2) Visible is the broadening of HPhb believed to be a result of the free rotation experienced by the phenyl group. (Figure 5.3-1)

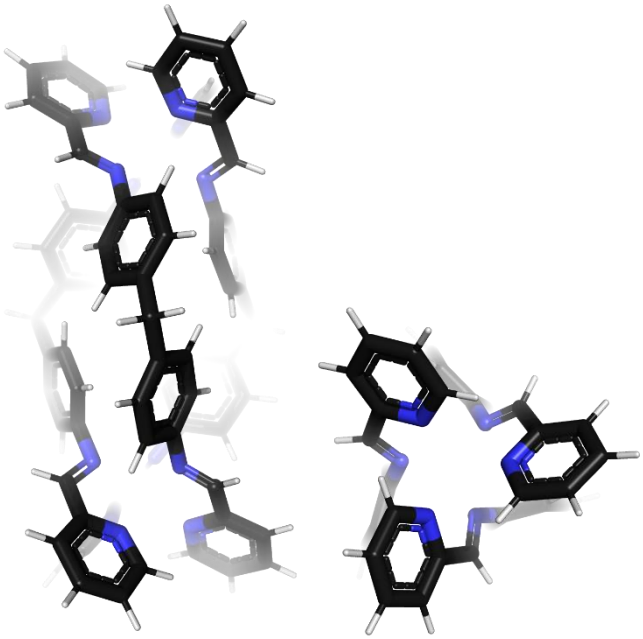


Figure 5.3-2 3D illustration with metals not shown. The two and three fold symmetry of the cylinder, illustrated by a side on and end on view respectively.

As mentioned above, the cylinder experiences a two-fold and three-fold symmetry as a result of the three symmetric ligands, bound octahedrally to two metal centers. (Figure 5.3-2)

Opposite enantiomers will likely have a different response to their environment, and a spectrum is greatly simplified by removing one. Therefore, before NMR experiments are performed on bound TAR it is important to separate the enantiomers. (Section 5.1.3)

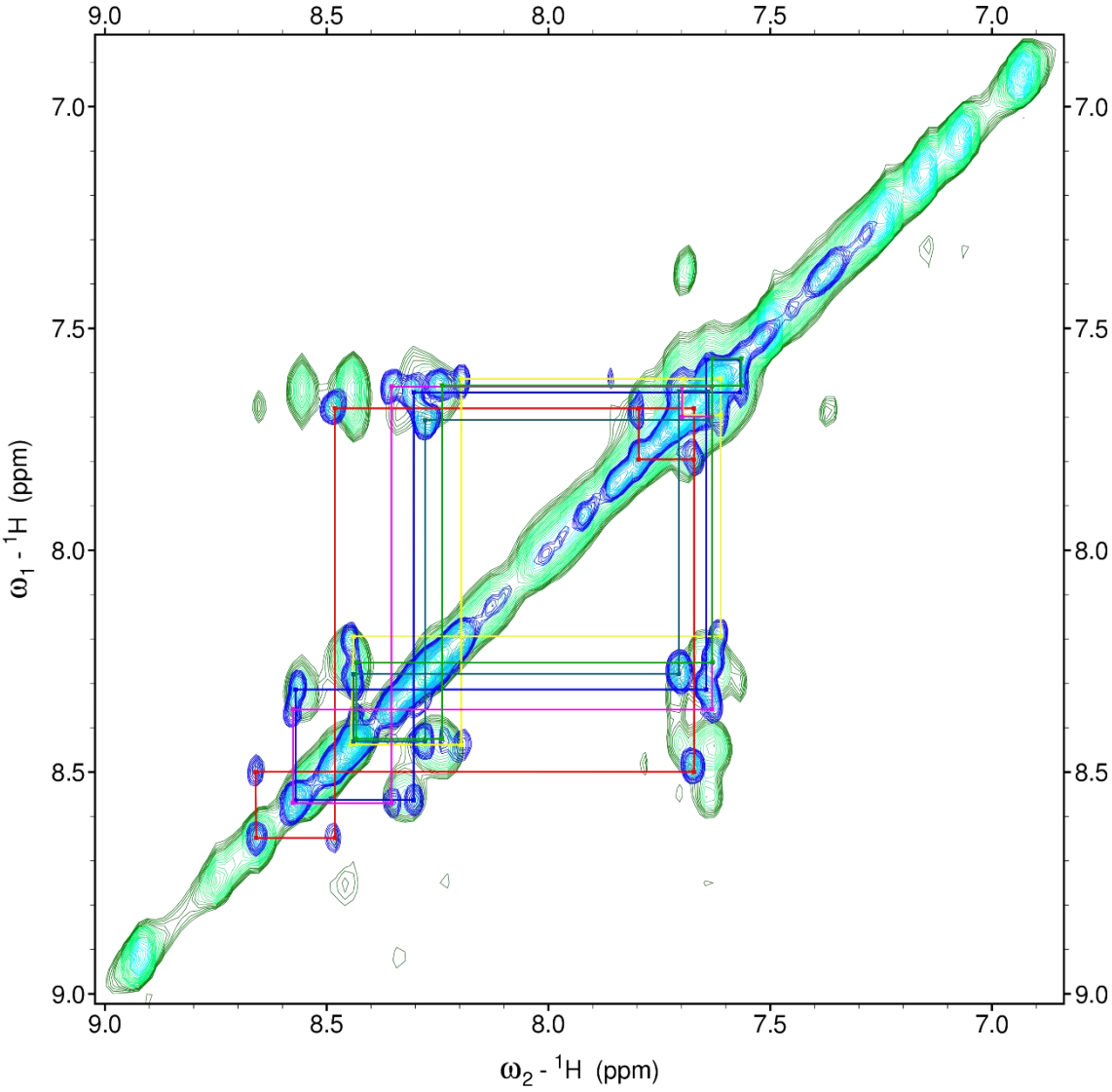


Figure 5.3-3 ^{1}H.^{1}H-COSY (blue) and ^{1}H,^{1}H-TOCSY (green) overlay. Of 1:1 TAR RNA (0.49 mM) with RuCy_P (0.49 mM) in D_2O. 50 mM NaCl, pH 7, 278K, 700 MHz. Showing cylinder pyridine J-coupled peaks. Sets of peaks are indicated with coloured lines to differentiate unique pyridyl environments.

Despite being a pure sample of P enantiomer, on binding to TAR RNA, the pyridyl peaks split into 6 distinct J-coupled sets traceable through COSY cross-peaks. (Figure 5.3-3) This indicates the loss of both the two and three-fold symmetry. The value of this information lies in how the cylinder comes to lose its symmetry. For instance, a nucleus in proximity to a pyridine could be expected to break the symmetry, dependent on the movement of the complex. A complex that can freely move around the RNA would be equally affected

at all pyridines resulting in no difference between the resonances. If it were localised but had no preference for orientation around the metal-metal axis as it rotated, we would see two distinct populations corresponding to either end as the three pyridines contributing to each end would be equivalent. In our case we see 6 distinct populations for each pyridine peak, suggesting the cylinder binds in a specific way to the RNA. There is no J coupling from the imine to pyridine protons, so they are not visible with TOCSY/COSY.

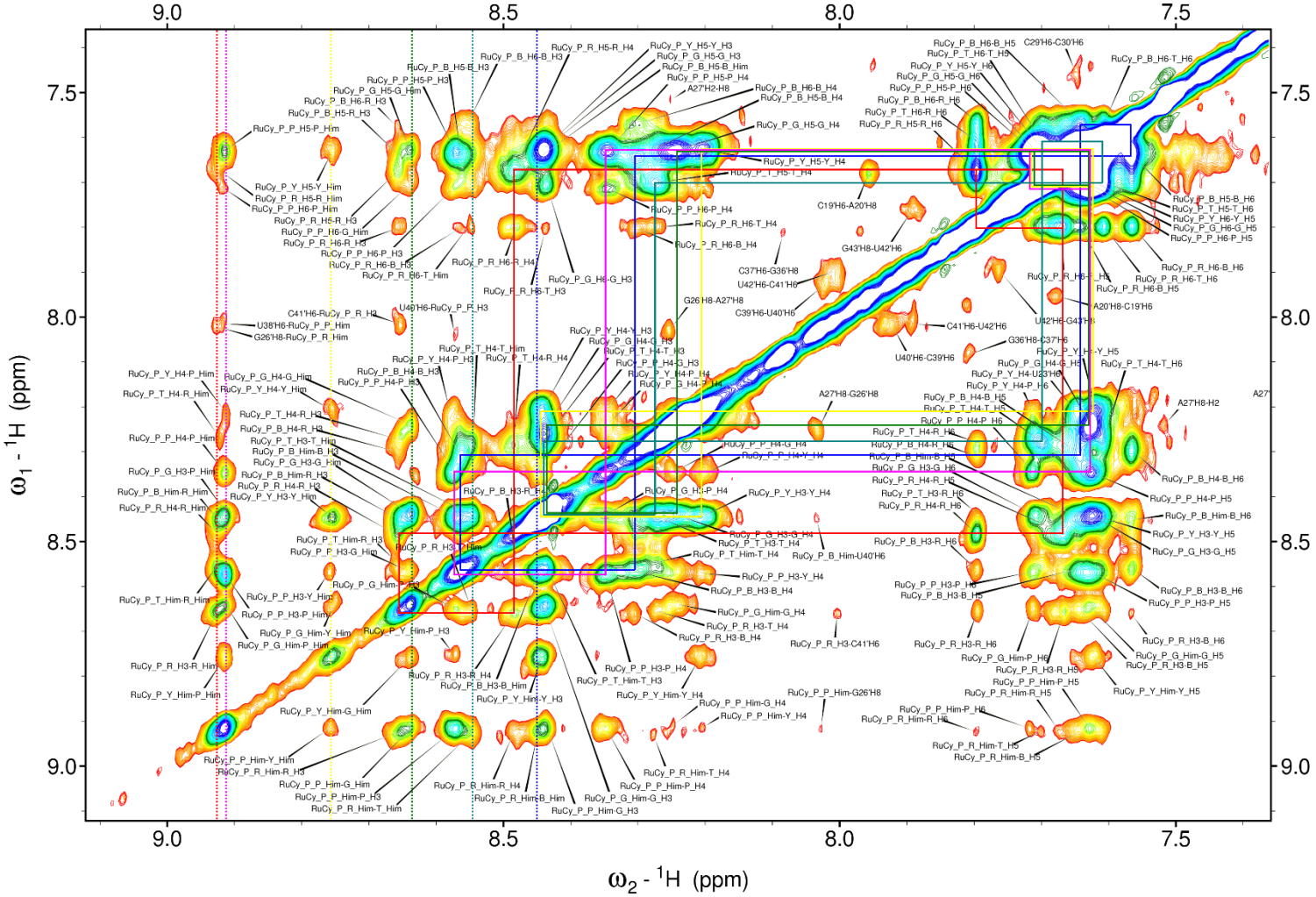


Figure 5.3-4 ^{1}H,^{1}H-NOESY spectrum of 1:1 TAR RNA (0.49 mM) with RuCy_P (0.49 mM) in D_2O. 50 mM NaCl, pH 7, 278K, 700 MHz. Showing the aromatic region, including many of the cylinder pyridine and imine peaks. Peaks are colour coordinated as previously established in Figure 5.3-3. Straight lines indicate the J-coupling sequence H3-6, dashed lines indicate the imine resonances.

By mapping the J coupled pyridine protons onto the NOESY the same pattern can be traced. The imine protons can then be identified through NOEs to the H3, and other nearby protons. Like the pyridines, the imines are also split into 6 distinct resonances, with two groupings: Red, blue, teal (RBT) and yellow, purple, green (YPG). These are linked through inter-ligand NOEs and constitute the two ends. (Figure 5.3-4)

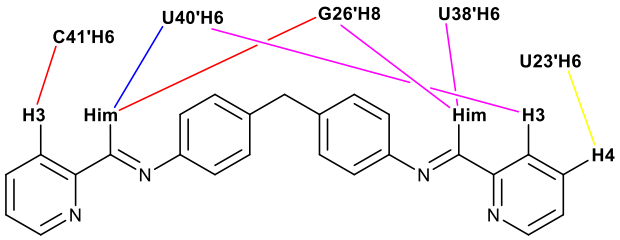


Figure 5.3-5 Diagram of the cylinder ligand showing the NOE contacts from TAR RNA H6 visible in Figure 5.3-4. Lines indicating the contacts are coloured dependent on the colours assigned during the cylinder assignment. Positions of RNA are illustrative.

There are intermolecular NOEs present in this region of the ^{1}H,^{1}H-NOESY spectrum. They are between the cylinder protons and the base H6 protons. Interestingly these seem to follow a pattern where the bases at the bottom of the bulge seem to have more contacts with the RBT end of the cylinder, and the top of the bulge seems to have contacts to the YPG end. These are not exclusive however, as the U40'H6 and G26'H8 appear to have contacts to both ends, perhaps indicating that one side of the cylinder is facing these bases, with the red, blue and purple environments facing toward the bulge. This would be possible with the cylinder placed either parallel to the minor groove, or within the bulge perpendicular to the minor groove. If within the bulge, the blue would be at the bottom of the bulge, the red at the top and at the other end of the helix the purple in the middle. U23'H6 is flexible and being the only base with an NOE to the yellow H4 is reasonable. This model however does not account for the lack of involvement of other

parts of the cylinder with U40'H6 and G26'H8. Parallel to the helix in the minor groove, the cylinder would be positioned with the RBT end at the bottom of the bulge, and the YPG end at the top. The plane that can be drawn between the red, blue and purple imines faces the bulge. For the yellow H4 to U23'H6 NOE, the base would likely be looped over the cylinder. (Figure 5.3-5)

The bound phenyl peaks are no longer strong and broadened as seen in the unbound cylinder spectrum. (Figure 5.3-1) The peaks are instead split into multiple weak resonances. It is unknown whether the phenyl peaks are split further as a result of loss of rotation, there would be 24 potential peaks if so. Though the phenyl rings are the most flexible part of the cylinder and provide a tiltable pi system for face-to-face stacking. This behaviour has been seen in a crystal structure of an RNA 3WJ formed around a cylinder.[53] In NMR DNA 3WJ experiments the bound cylinder experienced phenyl splitting from reduced phenyl rotation, though the symmetry of the 3WJ maintained the three-fold symmetry.[54] Some phenyl protons have weak NOEs to the adjacent imine aiding in their assignment. (Figure 8.2-6, Figure 8.2-8)

5.3.2 TAR and RuCy_P

With the cylinder peaks showing high specificity, focusing now on the RNA, it is important to gain an understanding of the location of the cylinder binding. As well as the effects of cylinder binding on the TAR in and away from the immediate binding site.

5.3.2.1 90% H_2O Spectra

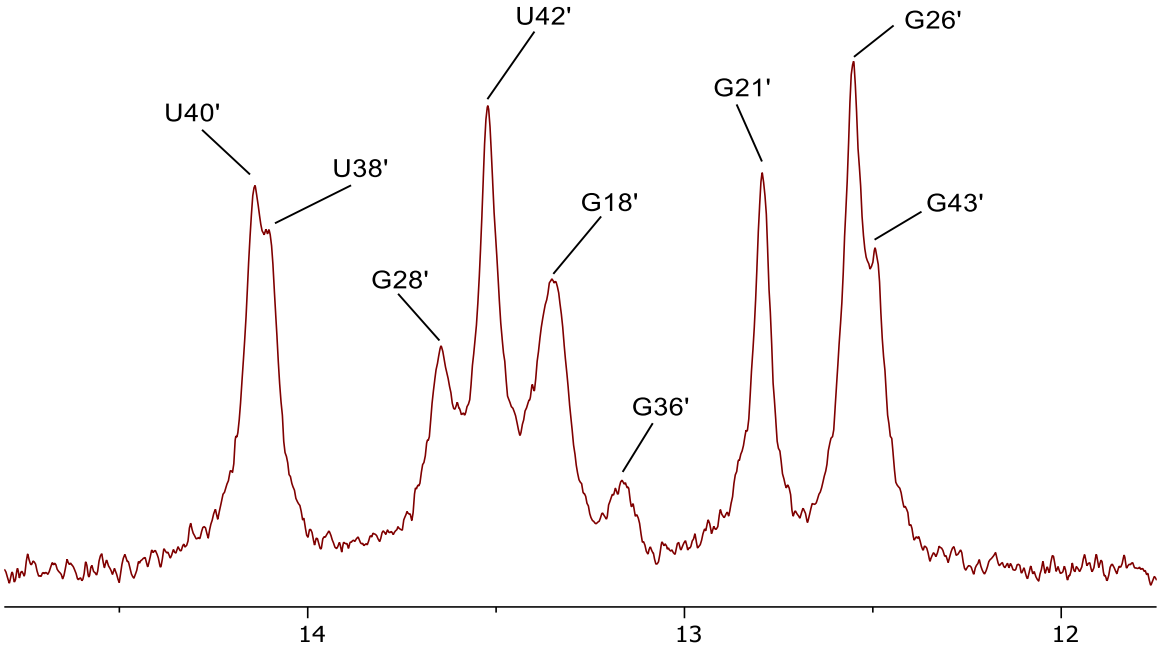


Figure 5.3-6 ^{1}H-NMR spectrum of 1:1 TAR RNA (0.46 mM) and RuCy_P (0.46 mM) in 90% H2O showing the imine peaks, 700 MHz.

The U40'H3 imine peak is present in the bound TAR. This was not present in the free TAR and demonstrates that the A22'-U40' base pair is stabilised by the cylinder binding. This also hints at a stabilisation and loss of flexibility at the bulge. Also observable, is a weakening of the G36'H1 at the top of the loop, and possibly G28'H1 too as a result of the loss of stability from the C29'-G36' pairing. (Figure 5.3-6)

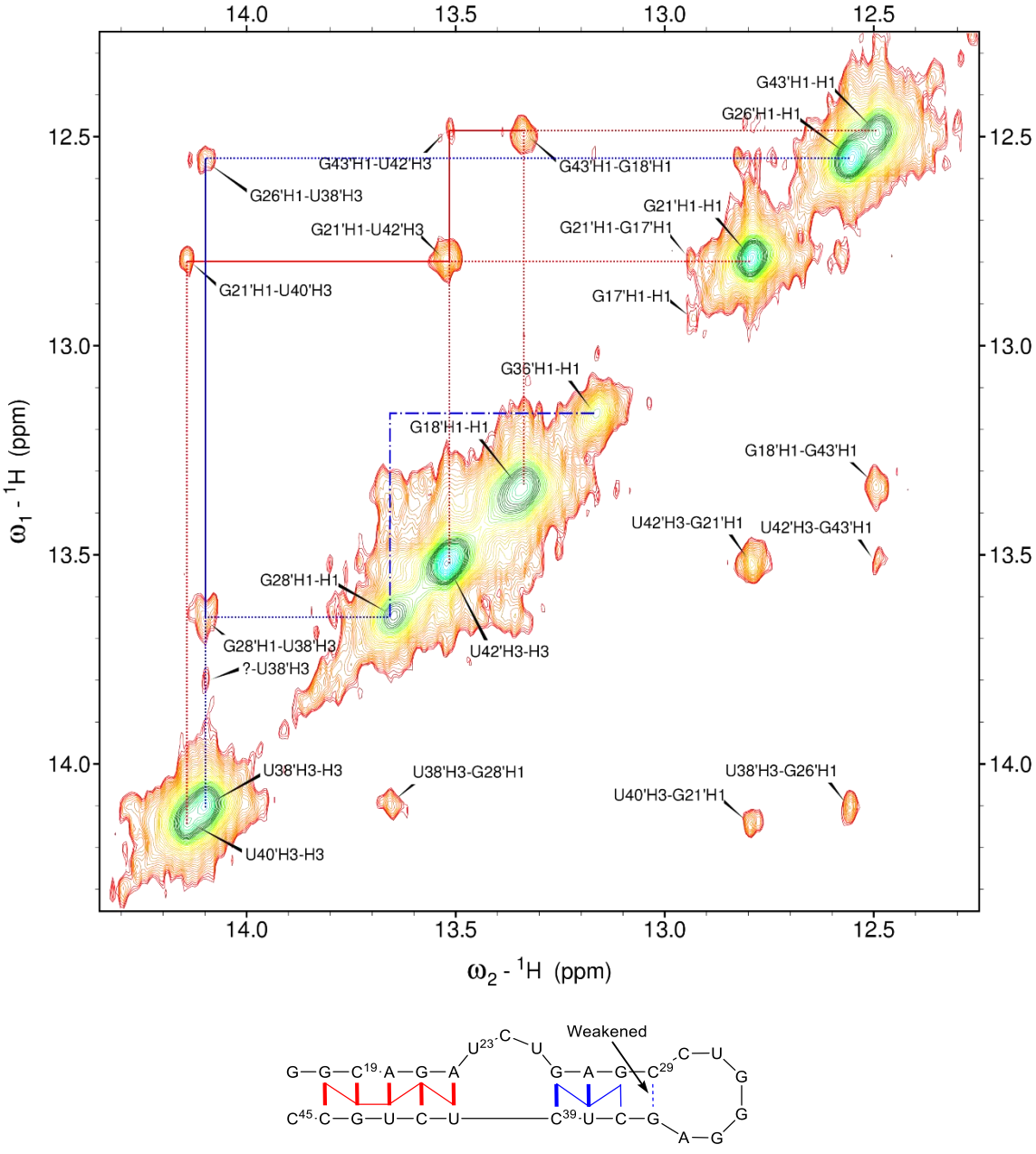


Figure 5.3-7 ^{1}H, ^{1}H-NOESY spectrum of 1:1 TAR RNA (0.46 mM) and RuCy_P (0.46 mM) in 90% H2O showing the imino region, 700 MHz.

As in the 1D spectrum, the G36'H1 is weakened in intensity. It also has no crosspeak to the G28'H1. Both point to a weakening of the C29'-G36' base pair, possibly opening the base of the loop with reduced stacking onto the G28-C37. In MD simulations the upper stem is expected to uncoil on RuCy_P binding, this might coincide with changes in the loop.[12] (Figure 5.3-7)

The G28'H1-U38'H3 crosspeaks are split into multiple resonances that are hard to distinguish, with the U38'H3 showing greater variation. These different populations together with their strong intensity in the 1D spectrum above indicate, not necessarily weakening of the base pair, but a degree of uncertainty in the exact conformation nearby. This could be a result of the cylinder binding close, or effects from the changes to the base of the loop.

A crosspeak from U38 to an unknown resonance high enough frequency to be uracil, hints to the potential for a small population forming a U23'-A27'-U38' base triple, though there is no other evidence to confirm this suggestion. As of the publishing of this thesis there are no reports of directly observed NOEs confirming this structure for HIV-1 TAR. This could be confirmed using ^{2h}J(N,N) scalar couplings.[15,27]

With a strong NOE to the U38'H3, G26'H1 is among the stronger peaks in the 1 and 2D spectra, in comparison to the free TAR data where it is like the other peaks. This relative strength points to an increased stability at the G26-C39 bond.

Across the bulge, U40'H3 is both strong and has a direct NOE to G21'H1, confirming it is assignment. This demonstrates the strengthening of the A22'-U40' base pair on addition of the cylinder. This effect is seen in conjunction with the formation of the base-triple on argininamide binding.[15] Though this may not be a cooperative behaviour, and is not evidence for base-triple formation in this structure.

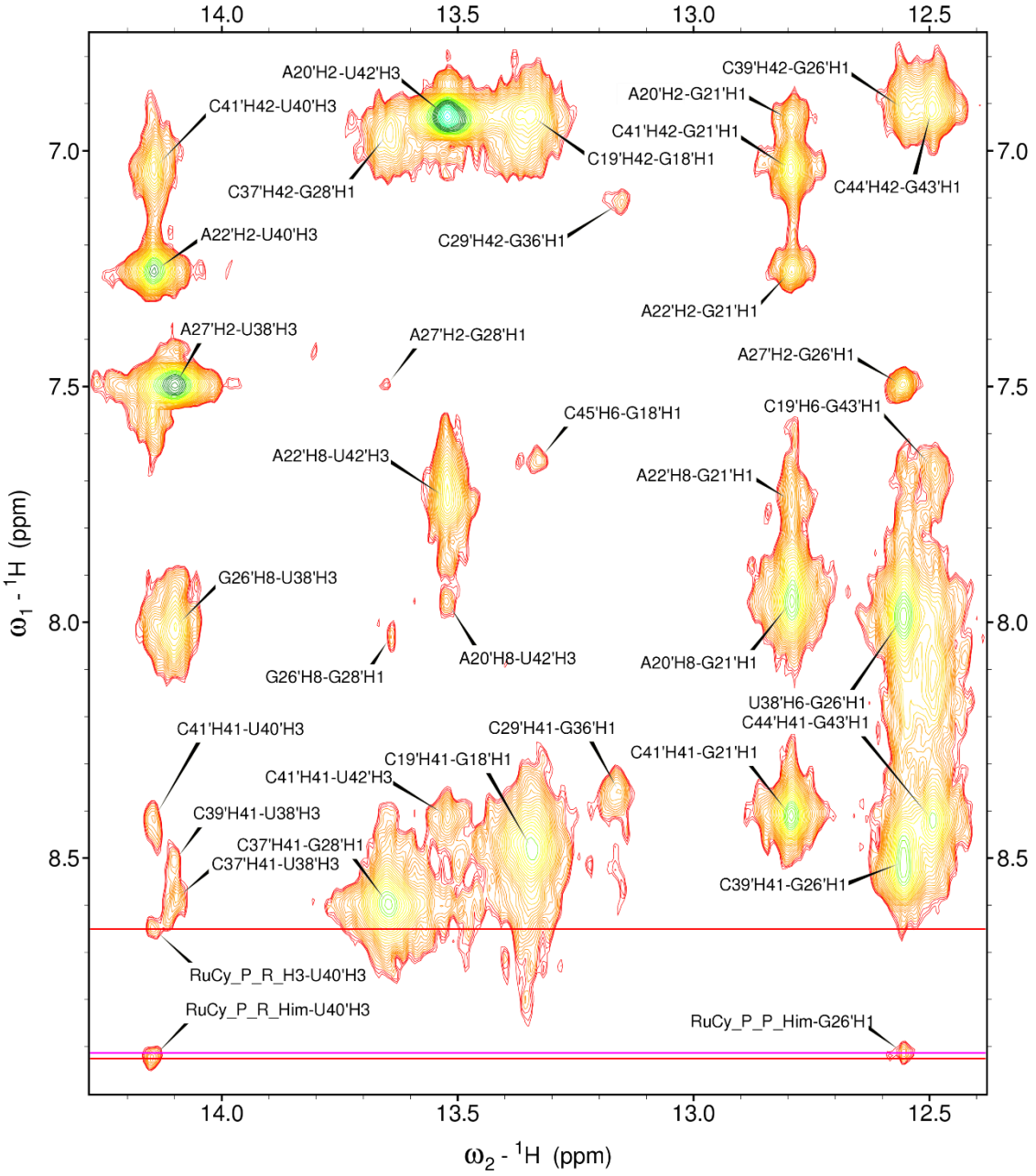


Figure 5.3-8 ^{1}H,^{1}H-NOESY spectrum of 1:1 TAR RNA (0.46 mM) and RuCy_P (0.46 mM) in 90% H2O showing the imino-aromatic/amino cross peaks, 700 MHz. Colours of pyridine groups are detailed in section 3.1.

Direct NOE contacts are visible from the U40'H3 and G26'H1 to cylinder peaks. The U40'H3 at the bottom of the bulge has NOEs to the red H3 and Him, whereas the G26'H1 at the top of the bulge has an NOE to the purple Him, locating the cylinder contiguous to the bulge. The red and purple pyridine/imine peaks are believed to be on opposing ends of the cylinder, assigned from D_2O bound cylinder spectra, though which ligand is

unknown. (Figure 5.3-4) NOEs to different ends of the cylinder at the top and bottom of the bulge could agree with this, though the cylinder is approx. 18 Å lengthways and 7 Å across,[34] so it is likely not aligned along the TAR lengthways.

Strong NOEs from the adenine H2 to the imine protons confirm their assignment in the D_2O spectra. The U40H3-A22H2 NOE, described as evidence for strengthening of the basepair, is present and strong.[1] Also observed are other aromatic-imine peaks that can be crossreferenced to the D_2O spectra.

5.3.2.2 100% D_2O NOESY

The reason for reducing the temperature of the sample to 278K for the 90% H_2O NOESY is less mobility in the bases results in stronger hydrogen bonds, and therefore stronger exchangeable peaks. Considering the D_2O NOESY, the effect of reducing the temperature to 278K in line with the 90% H_2O NOESY is a strengthening in intensity for the lower energy populations, and possibly better connectivity for the sequential walk. (Figure 5.3-9) Confirming the H2 peaks, we can see coupling in the TOCSY spectra from AH2 to the exchangeable UH3. Presumably as ^3J coupling. (

Figure 8.2-2)

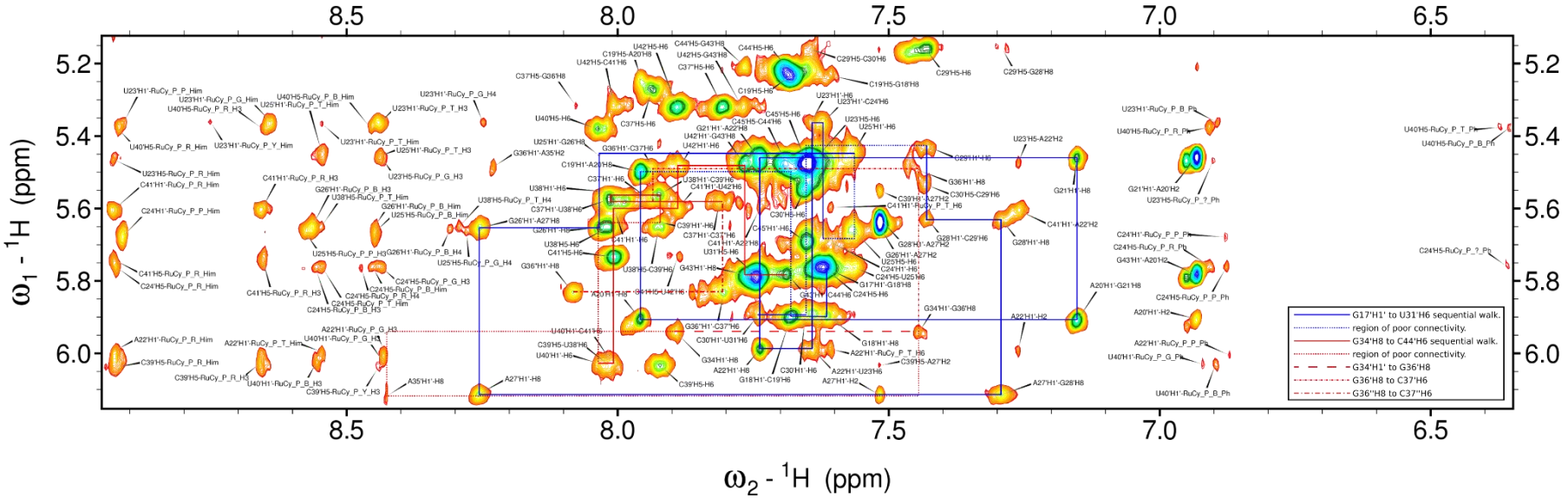


Figure 5.3-9 ^{1}H,^{1}H-NOESY spectrum of 1:1 TAR RNA (0.49 mM) with RuCy_P (0.49 mM), in 100% D2O. 278K, 50 mM NaCl, pH 7. d8 = 250 ms, 700 MHz. Showing the H8/6/2-H1'/5 region as well as intermolecular NOEs from the cylinder to the RNA. On the left side of the spectra are pyridine and imine intermolecular NOEs and on the right are intermolecular NOEs resulting from the phenyl protons. Lines describe the interconectivity of bases. Solid blue shows the G17'H1' to U31'H6 sequential walk. Regions of poor connectivity, with missing C19'H1'-H6 and C29'H1'-C30'H6 are indicated with dotted blue. The walk is continued from G34'H1' to G36'H8 with solid red lines. A series of missing NOEs between G34'H1' and G36'H8 are shown with dotted lines, and the connectivity across G34'H1' and G36'H8 is shown with red dashed lines. The two populations of C37' and C37'', and the surrounding sequential walk are shown in red dot-dash lines of two different lengths.

Beginning at the base of the stem, there is slight fraying seen as unobservable G17'H1'-H8. This is unsurprising considering the trace G17'H1-H1 peak in the imino region. (Figure 5.3-7) A similar behavior is true for C45' on the opposing side, there are no definite NOE peaks connecting C44'H1' to C45'H6, though C45'H1'-H6 is strong. There is then good connectivity until an indistinguishable C19'H1'-H6 likely buried under another peak, resuming all the way to C24'H1'-H6. This is comparably more connectivity at the base of the bulge than the free TAR, though this could be due to temperature so caution must be exercised comparing to the 298K measurement in this thesis. Continuing the opposing side of the sequence, a well resolved walk is obervable from C44'H6 to U40'H1'. In line with good base-pairing in the lower stem G43'H1'-A20'H2 across the strands is visible. Though the well shifted U25'H1'-G26'H8 is weak, a walk can be traced from this around a very strong and also well shifted G26'H1'-A27'H8 around to C29'H6. Differing from the free TAR, the loop appears to be less stacked, with poorer connectivity between C29' and C30' with an absent C29'H1'-C30'H6 peak. There are peaks for C30'H1' to both the C30 and U31 H6's. The walk is then absent until it resumes with G34'H1'-H8. It is worth relating this and what is discussed next to the 90% water sample where G36'H1 is weak and has no NOE contact to the G28'H1 pointing to poor or transient pairing at the base of the loop.

There are two conformations visible for G36 and C37 denoted with (') and ('). G34'H1' has an NOE peak direct to G36'H8, which can be followed to C37'H1'-H8. C37'H1'-C37''H8 is also where the other population converges. G36'' and C37'' are characterised by relatively strong G36''H1'-H8 and G36''H1'-C37''H8 peaks due to either proximity of the nuclei, or proportion of nuclei in this conformation. Considering the C37''H5-H6 peak

is stronger than C37'H5-H6 this is likely a contributing factor, though if as suspected G36'' is facing into the loop both effects could be present. Expanding upon this possibility, G34' and A35' both have connectivity to G36' but not G36''. A peak for G36''H8-C37''-H5 is also present, resulting from close proximity of the bases suggesting a degree of stacking.

Continuing the walk from C37'H1', we can walk through U38' to C39'H1'-H6. Absent is C39'H1'-U40'H6, which is understandable as this is a point of flexibility between the stems. In previous free TAR NMR structures, U38 and C39 stack at the top of the bulge, and at the bottom U40 stacks onto C41.[8]

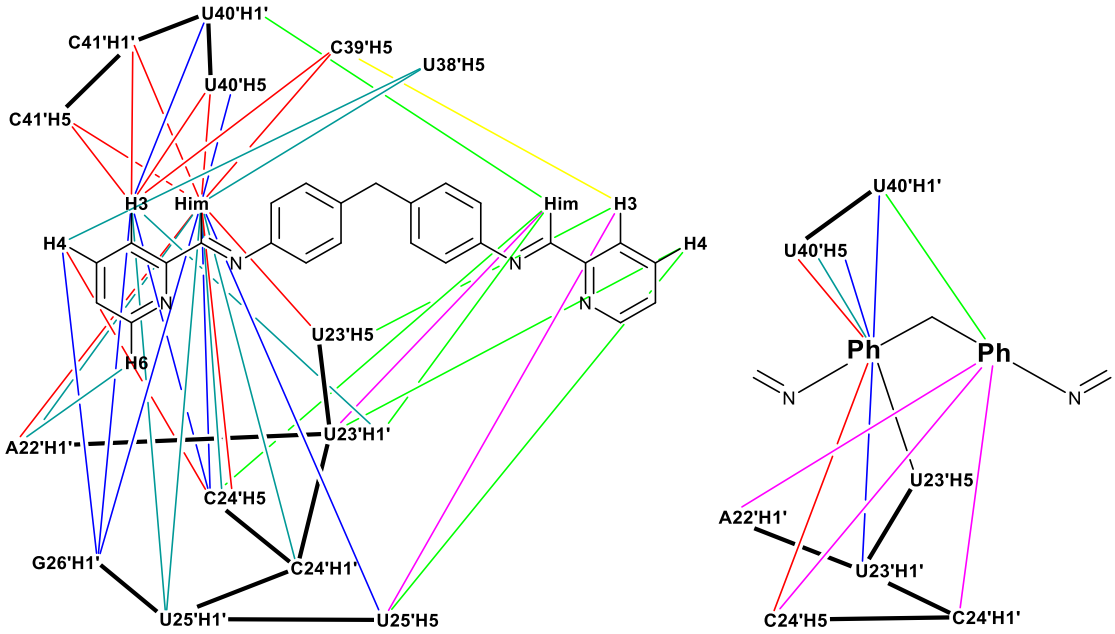


Figure 5.3-10 Diagrams of the cylinder ligand, showing the NOE contacts from TAR RNA H1'/H5. Left. The pyridine and imine peaks. Right. The phenyl peaks. Lines indicating the contacts are coloured dependent on the colours assigned during the cylinder assignment. Positions of RNA are illustrative.

The positions of the RNA were placed to maximise the visibility of connections and simplify viewing only, however it does pose an interesting perspective and is reminiscent of the conformation adopted on argininamide binding reported by J. Puglisi.[24]

Returning to the bulge, there are NOEs from the cylinder to the RNA. Let's start from the lower bases on the 5' strand with C41'. The lowest NOE to the cylinder is C41'H1' with contacts to RuCy_P_R_Him, RuCy_P_R_H3. The same two cylinder peaks also have contacts to C41'H5. These are the same nuclei that have NOE peaks in the 90% H_2O sample to U40'H3. Moving up through U40'H1' and H5, NOEs to red and blue pyridines, but also G_Him are observed. This is the first contact to the other end of the cylinder for this strand. Following this, there is U39'H5-RuCy_P_Y_H3, the only other pyridine NOE to this strand. For the phenyl peaks, U40'H5 and 1' both have blue and red contacts, but the U40'H1' also contacts a green phenyl. All this conforms to expectations for the base of the complex interacting with the minor groove. At the top, G26'H1', a paired base, experiences NOEs exclusively to the blue pyridine/imine, this tells the observer that G26' is interacting with the same end of the cylinder as its partner C39'. That the top and bottom of the bulge are interacting with the same end of the cylinder on both strands suggests that the cylinder might be inserted within the bulge, perhaps with the RBT end facing the minor groove. Below G26', U25'H1' also interacts with the same end, though T_Him and T_H3. Its base, the more mobile of the bulge,[8] has the NOEs U25'H5-RuCy_P_P_H3, U25'H5-RuCy_P_B_Him, and U25'H5-RuCy_P_G_H4. G_H4, being the proton with NOEs to the RNA furthest from the RBT end. C24'H1' contacts T_Him and P_Ph, relatively few contacts. C24'H1'-RuCy_P_T_Him is a strong peak though, so perhaps they are particularly close. C24'H5 comparatively has many contacts, mostly to RBT pyridines, but there is one to G_Him. In addition, C24'H5 has contacts to phenyls on both ends. Considering the NOEs to C24'H1' and H5 it is likely that the C24' sugar is near to the middle of the cylinder, with the base stacking onto the cylinder phenyl, this is

possible within either the minor groove or bulge insertion binding models. Finally, U23' has NOE peaks to both ends, but more to the YPG end. It is also the only base other than U25' to interact with G_H4. Given U23' is n+1 to A22, in the minor groove binding model it will be aligned with the base pointing along the cylinder toward YPG, interacting along the outside. (Figure 5.3-10)

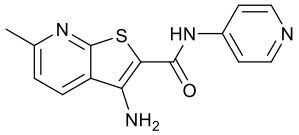


Figure 5.3-11 One of the ligands investigated in the study mentioned below.[55]

Contrasting to previous publications, a small molecule that binds to the TAR bulge also binds to the lower bulge region within the minor groove. TOCSY data shows the C41, U40, C24, and U23 H5-H6 peaks shifting. C39 also shifts significantly on binding, suggesting the ligand induces a conformational change. Where the results differ is in the data presented in this chapter the bulge base U25 is also significantly affected, as well as the minor groove interactions extending further up to U38. This is likely due to the size and cylindrical structure of the helicate, providing a large surface for the bulge bases to interact, capable of interacting in a larger region.[55] (Figure 5.3-11)

In support of the bulge insertion model is the prevalence of YPG NOEs to the bulge bases, and the RBT NOEs to the paired bases. To expand on this, a cylinder within the bulge perpendicular to the helix would have one end closer to the ribose protons in the minor groove, and the other end closer to the bulge bases in the major groove. The base H5 protons would be closer to the cylinder in this model, matching the quantity and intensity of these intermolecular NOEs. Contacts from the cylinder to U38' and C41'

(expected to be further from the cylinder than U40’ and C39’,) would be understood from the perspective that the helical turn of the stem means that the bases are stacked with a slight overhang.

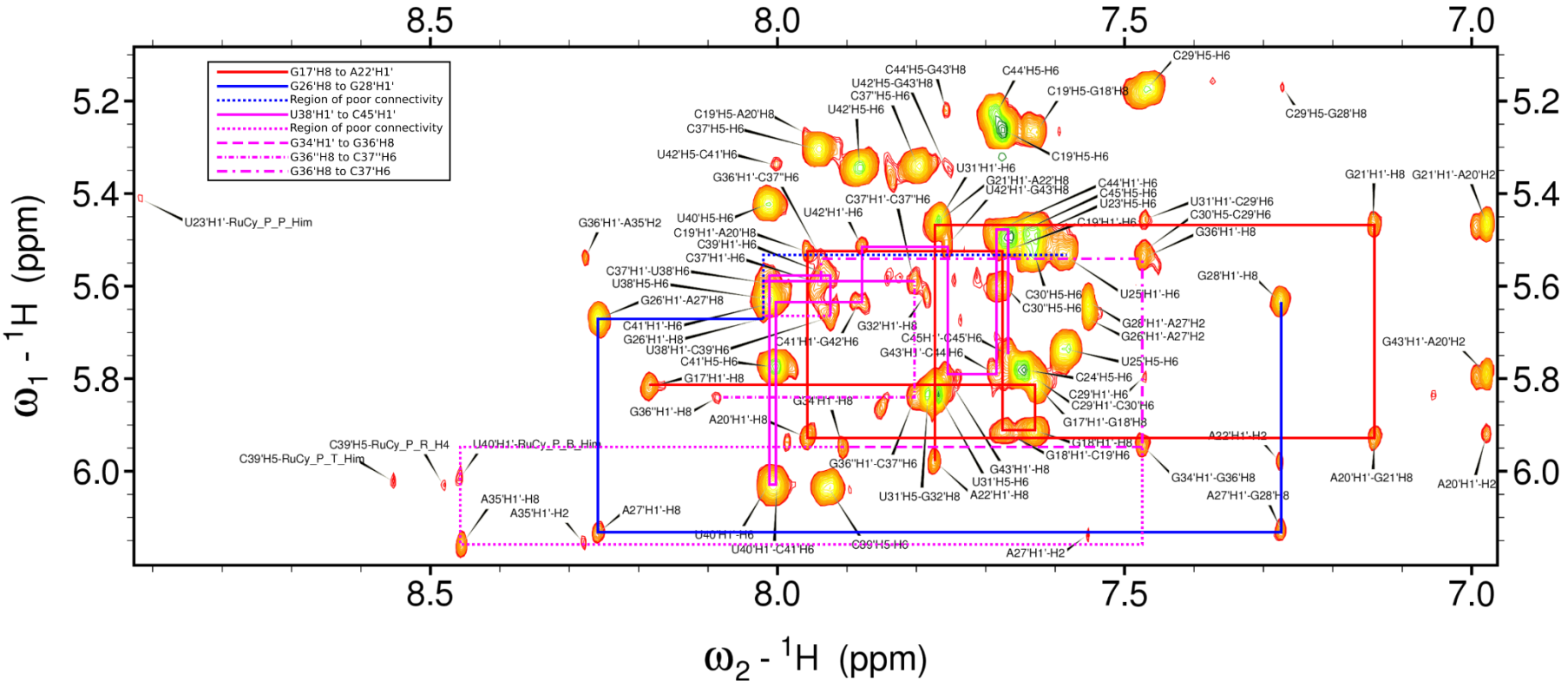


Figure 5.3-12 ^{1}H,^{1}H-NOESY spectrum of 1:1 TAR RNA (0.49 mM) with RuCy_P (0.49 mM), in 100% D2O. 298K, 50 mM NaCl, pH 7. d8 = 250 ms, 700 MHz. Showing the H8/6/2-H1'/5 region as well as intermolecular NOEs from the cylinder to the RNA. On the left side of the spectra are pyridine and imine based NOEs and on the right are NOEs resulting from the phenyl protons. Lines describe the interconectivity of bases. The walk between G17'H8 and A22'H1 is shown in solid red. The walk is continued in solid blue between G26'H8 and G27'H1' with a dotted blue line showing the lack of NOE between U25'H1' and G26'H8. U38'H1' to C45'H1' is shown in solid purple. The absence of G34'H1'-A35'H8 and A35'H1'-G36'H8 are shown with dotted purple lines, and the NOE between G34'H1' and G36'H8 is shown with a dashed purple. The two populations between G36'/G36'' and C37'/C37'' are noted with longer/shorter dot-dash lines.

Range and rate of motions increase with temperature. As a result, at 298K some NOEs are weakened, and some are shifted. Though some are strengthened. Beginning at the lower stem, we now have a strong sequential walk from G17'H1'-H8 to A22'H1'-H8, and a sequential walk from C45'H1'-H6 to C39'H1'-H6. A visible C39'H1'-H6 and C39'H1'-U40'H6 hint to a closing of the inter-helical angle, this is also seen in arginine and tat bound TAR.[24,50] The bulge region has very little connectivity, with the exception of U25'H1'-H6. The walk continues from G26'H1'-H8 to G28'H1'-H8 where the walk is interruped. On the other strand, the walk is disrupted between C39' and U38', continuing at C37'H1'-U38'H6.

Like in the 278K spectra, there are both C37'H6 and C37''H6 present. The spectrum differs though, for the 298K G36''H1'-H8 and G36''H1'-C37''H6 are significantly weakened. A conformer considered to be resulting from further weakening of the C29'-G36' pairing with likely quite significant motion from G36'. Simultaneously in comparison to the 278K spectra, the G34'H1'-G36'H8 is similar in intensity. G34H1'-G36'H8 confirms G34' is stacked onto G36' which identifies this as another comformer, one where the G34' is involved in a transient cross-loop pairing with C30'. The A35', with very weak peaks throughout the spectra, would be looped out, but transiently and able to stack between G34' and G36'.[32]

At the same time as these changes to the C29'-G36' bond, a new peak emerges for C30''H5-H6. To cause a shift in this peak, we're likely seeing a change in the dynamics within the loop. One option is, within the conformer where G34' dynamically pairs with

C30', U31' is able to stack on-top of C30' in a short lived excited state, C30''.[33] Being able to form this excited state for long enough for a strong C30''H5-H6 peak, would potentially disrupt Cyclin T1 and tat binding to the loop.[30] This could be the cause of the broadening of the U31'H5-H6, best seen in the TOCSY spectrum. (Figure 5.3-13) Though there is no C29'H5-C30''H6 or vice-versa bond to point to base stacking.

Also present, though in reduced quantities are intermolecular NOEs to the cylinder. The reduction is likely a result of increased mobility of the cylinder and more dynamic base interactions. Seen here are C39'H5-RuCy_P_T_Him and C39'H5-RuCy_P_R_H4 at the top of the bulge, U40'H1'-RuCy_P_B_Him at the bottom of the bulge, and U23'H1'-RuCy_P_P_Him. Again, like at 278K U40' and C39' are in contact with the RBT end, whilst the bulge strand has contact with the YPG end. There were no intermolecular phenyl NOEs at this temperature, this can be attributed to the same mobility that reduces the quantity in peaks to the pyridine and imine. The free cylinder phenyls freely rotate leading to broadened signals, this effect is likely continued in the bound complex. That the 278K sample was able to visualise these peaks is indicative of the extent of the interactions with the phenyls. (Figure 5.3-12)

5.3.2.3 TOCSY

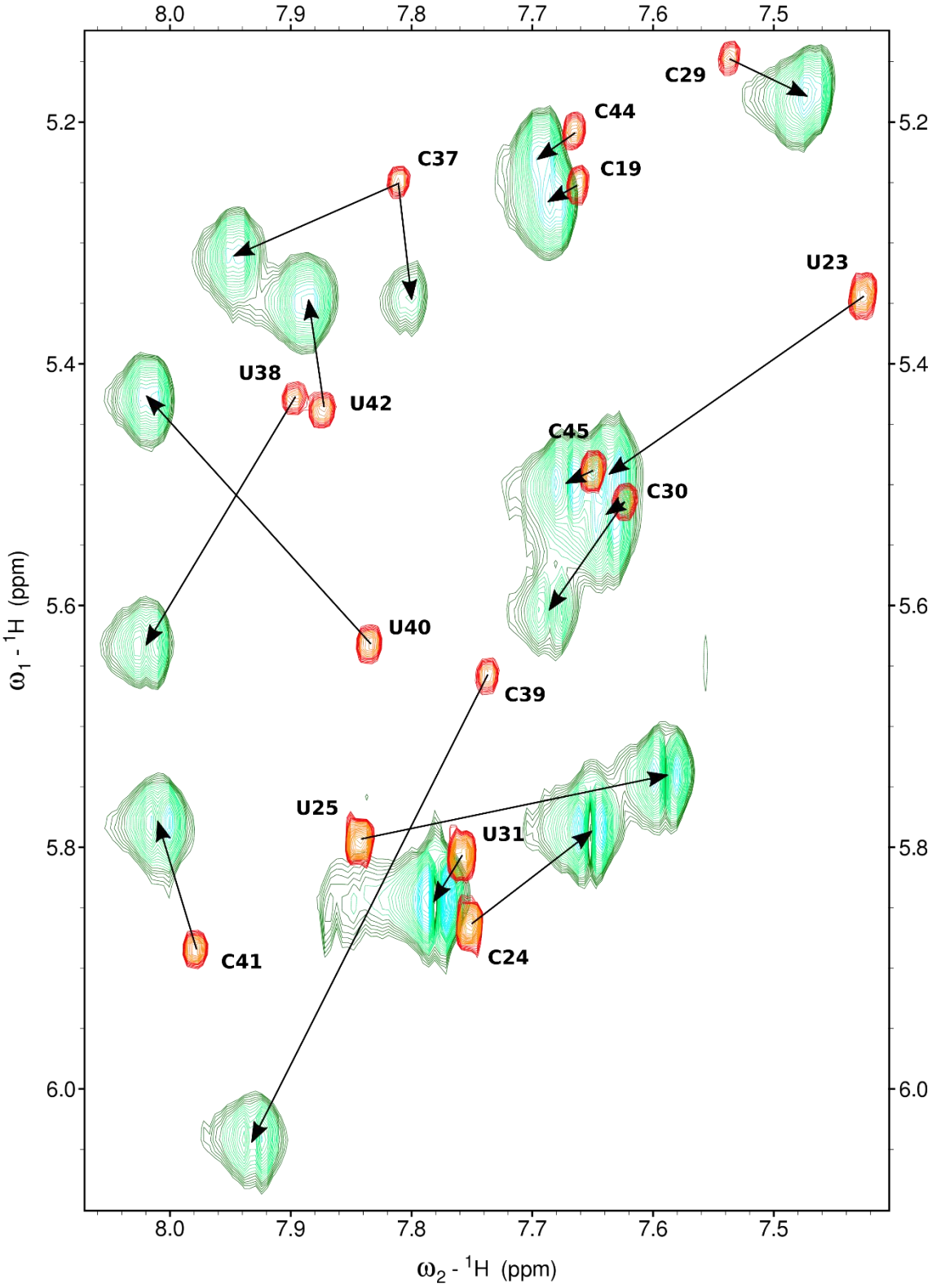


Figure 5.3-13 ^{1}H,^{1}H-TOCSY spectrum overlay of Green. 1:1 TAR RNA (0.49 mM) with RuCy_P (0.49 mM) and Orange. TAR RNA (1.5 mM) in D2O. 50 mM NaCl, pH 7, 298K, 700 MHz. Showing the shifts of H5-H6 peaks marked

with arrows. Green peaks appear broader than the orange as the spectra is zoomed in the vertical dimension. Assignment in (Figure 8.2-1)

The changes to the bulge region are stark when H5-H6 peaks are viewed. When combined with RuCy_P, the lower stem peaks C44,C19,C45 are relatively unaffected. Moving up toward the bulge U42 and C41 have slight shifts. U40,U23,C24,U25, U38, and especially C39 are shifted more so. C29, slightly shifted at the top of the stem. At the base of the loop, two populations are present for C37 and C30, likely due to two potential structures around the loop junction, including a cross loop pairing between C30, and G34.[31] The potential for concerted dynamic motions swapping the G34, and A35 in and out of the loop has been previously explored.[32] The instability around the loop in contrast to the bulge further nullifies the possibility that the cylinder is located in the loop. U31 finally is largely unaffected by cylinder binding, though experiences a "smeared" broadening effect. (Figure 5.3-13)

5.3.2.4 HSQC

HSQC was measured, however in this 1:1 ratio sample the spectra were weak and poorly resolved. Measurements were attempted in 100% D_2O and 90% H_2O at 278K and 298K, with no success.

5.3.2.5 DOSY

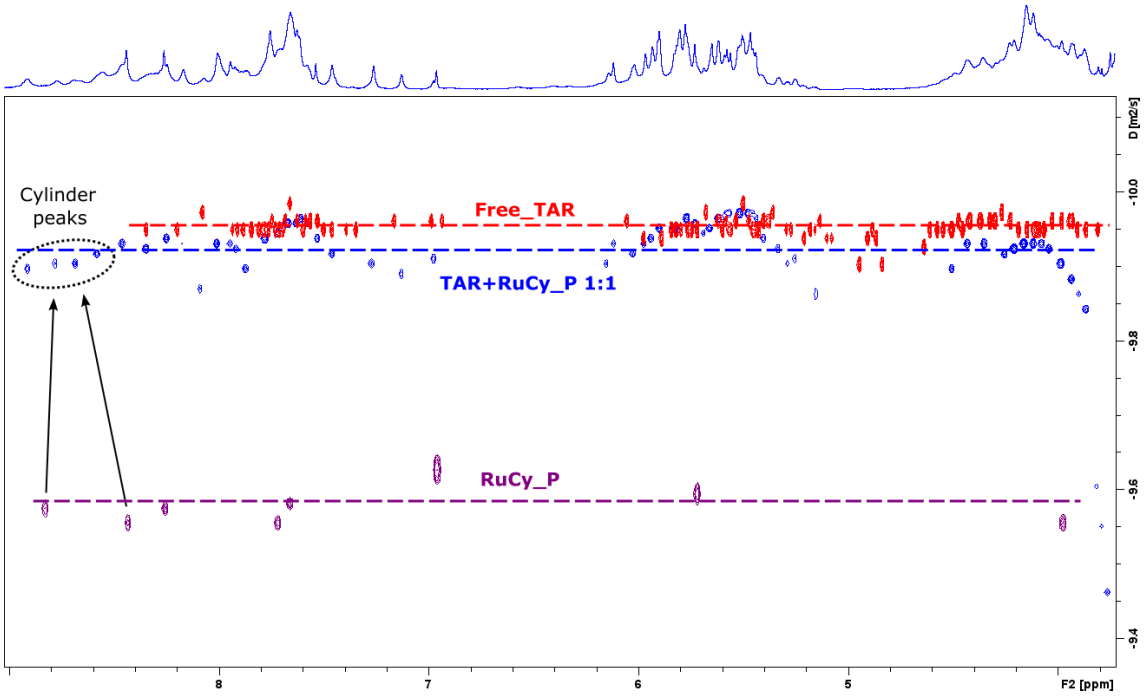


Figure 5.3-14 DOSY overlay 700 MHz. Red. Free_TAR RNA (D2O, 1.5 mM TAR RNA, 50 mM NaCl, 298K), Blue. Bound_TAR with RuCy_P (D2O, 0.49 mM TAR RNA, 0.49 mM RuCy_P, 50 mM NaCl, 298K), Purple. RuCy_P (D2O, 298K). Shown circled are cylinder peaks, corresponding to the Him, and H3 (arrows) as part of the TAR+RuCy_P agglomeration. Dashed lines to illustrate populations.

A DOSY spectrum, of the TAR with RuCy_P mixture has similar diffusion coefficients to the spectrum of the free TAR. Whilst the bound TAR is apparently diffusing slightly faster than the free, given the similarity and overlap between the two populations it is tempting to say that they are the same. However, given the many factors that can affect the transport of the complex and TAR RNA in a tube, it cannot be said whether they are the same or different. Though the RuCy_P in comparison is certainly a unique population. That the peaks attributable to the cylinder in the bound mixture, most obviously Him and H3, are diffusing at a similar rate to the TAR and not the RuCy_P is direct evidence for the association of the two assemblies. In other words, on binding to TAR RNA the cylinder's rate of diffusion is decreased. (Figure 5.3-14)

5.4 1:0.5 ratio

A series of data was acquired at a ratio of 1:0.5 TAR:cylinder. They provide a good set of data for a mixture of bound and unbound TAR in solution, with visible interconversion in some cases.

5.4.1 ^{1}H-NMR

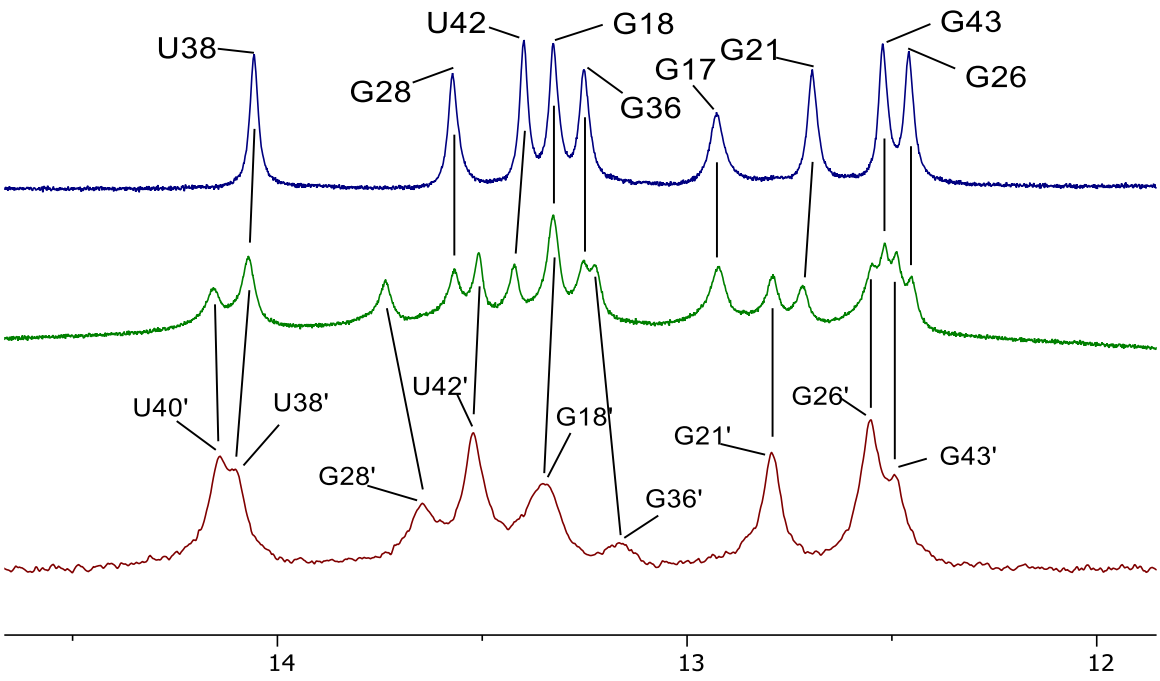


Figure 5.4-1 Stack with 1D NMR in 90% H2O at 278K. Top. Free TAR RNA 600 MHz, Middle. 1:0.5 mixture of RNA to ruthenium cylinder 700 MHz, Bottom. 1:1 mixture of RNA to ruthenium cylinder 700 MHz.

Comparing the mixture of 1:0.5 TAR RNA to cylinder, to the 1:1 and free TAR solutions, many similarities can be seen. The imino region assignments can be mapped across with little variation. Worth mentioning is the G36' is slightly stronger in the mixture, perhaps the rapid exchange with the free TAR structure has a stabilising effect. G18/G18' and U38/U38' have unnoticeable shifts overlapping in the mixture, in contrast to the G28' and G36' which vary between 1:1 and 1:0.5 solutions. Perhaps this difference between

the bulge and loop ends of the stem are a result of relative regions of structural stability and instability respectively. In agreement is the presence of the U40’ on the bottom of the bulge, that is present only in the bound form of both the 1:1 and 1:0.5 mixtures. Looking at the imino region the 1:0.5 mixture seems to have two populations present, the free TAR structure, and a structure with little deviation from the 1:1 mixture. (Figure 5.4-1)

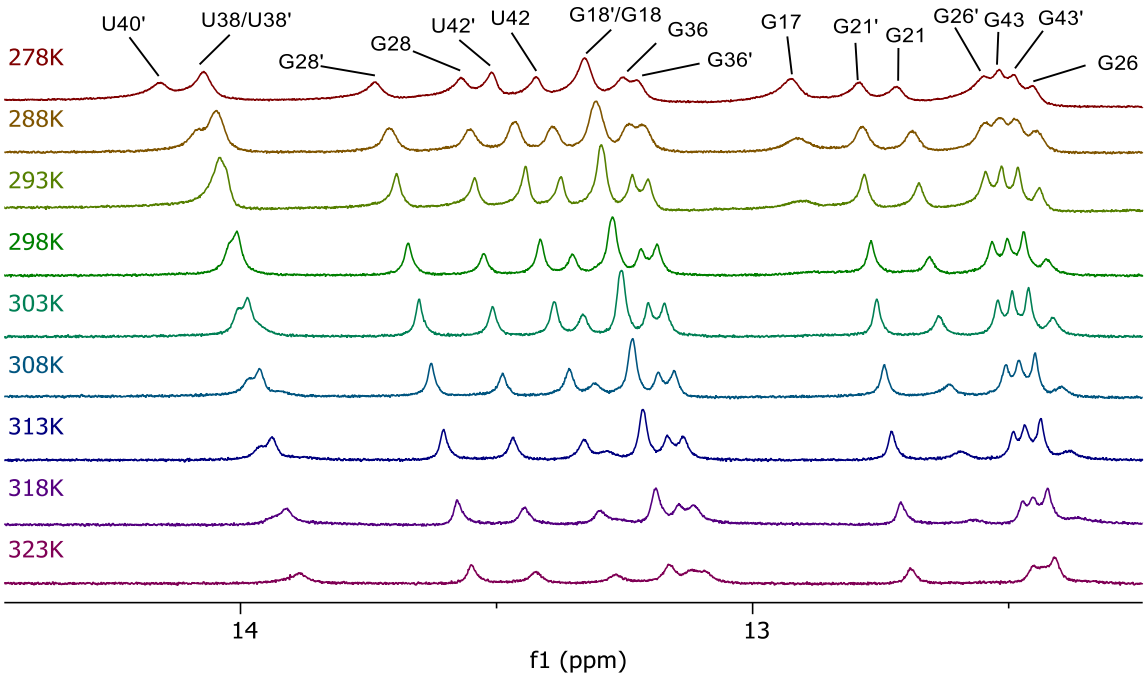


Figure 5.4-2 Stack with 1D NMR in 90% H2O, 700 Mhz. Variable temperature experiment between 278K and 323K.

On warming the mixture to 288K and 293K, there is some sharpening for the majority of peaks, but little change otherwise. U40’ however shifts upfield under the U38/U38’ peak, at higher temperatures it appears again as the U38/U38’ decreases but U40’ gone at 318K leaving the U38/U38’ as one of the peaks remaining at the final temperature of 323K. Increasing to 298K, G17 is almost indistinguishable from the baseline. The other peaks are affected more gradually.

Comparing the less stable unbound forms of G21/G26/U42 to the bound G21'/G26'/U42' which are present at 323K suggests that the cylinder bound structure has a greater stability at higher temperatures. However, the similar intensity of G28/G28' at 323K points to equal presence of both bound and unbound forms in solution. This demonstration that the ratio of bound/unbound is not dependent on temperature shows that the mixture is not dominated by low affinity constants, but a product of the ratio of cylinder to RNA. Perhaps the difference in intensity for the bound form of the G21'/G26'/U42' points to a region of structural stability around the bulge, compared to the relatively unaffected G28', closer to the loop. This is likely a result of the presence of the cylinder near to the bulge.

G36', which in the 1:1 cylinder to TAR spectra is noticeably weaker, (Figure 5.4-1) is seen to sharpen on increased temperature and is present alongside G36 at 323K. (Figure 5.4-2) This might indicate that the effect causing the weakened hydrogen bonding in the 1:1 mixture is not a "normal fraying" effect. But is perhaps due to a distortion of the bound structure that can be reduced by reducing cylinder interaction.

5.4.2 NOESY

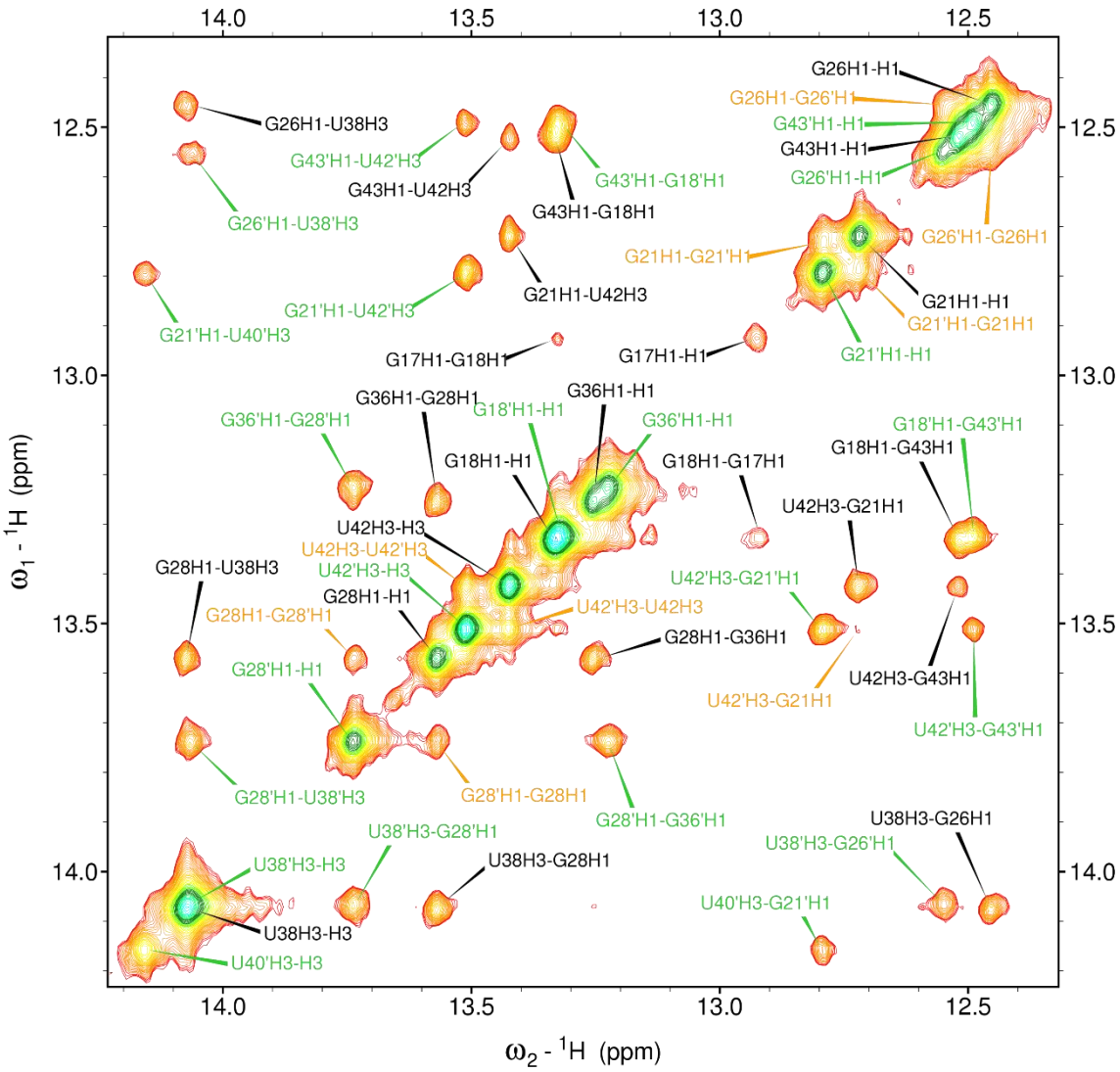


Figure 5.4-3 ^{1}H,^{1}H-NOESY spectrum of 1:0.5 TAR RNA (0.46 mM) with RuCy_P (0.22 mM) Black. free TAR, Green. Bound TAR, Orange. Exchange peaks. 700 MHz.

Using the 90% H2O NOESY spectrum, the connectivity between base-pairs for two distinct populations can be seen, the free and bound TAR indicated in black and green respectively. Here is evidence for the identity of U40'H3 with the G21'H1-U40'H3 crosspeak. Also shown, several exchange peaks visible between peaks with greater shift differences corresponding to the same nuclei of differing populations. As well as a small exchange peak at U42'H3-G21H1. These are indicated in orange. (Figure 5.4-3)

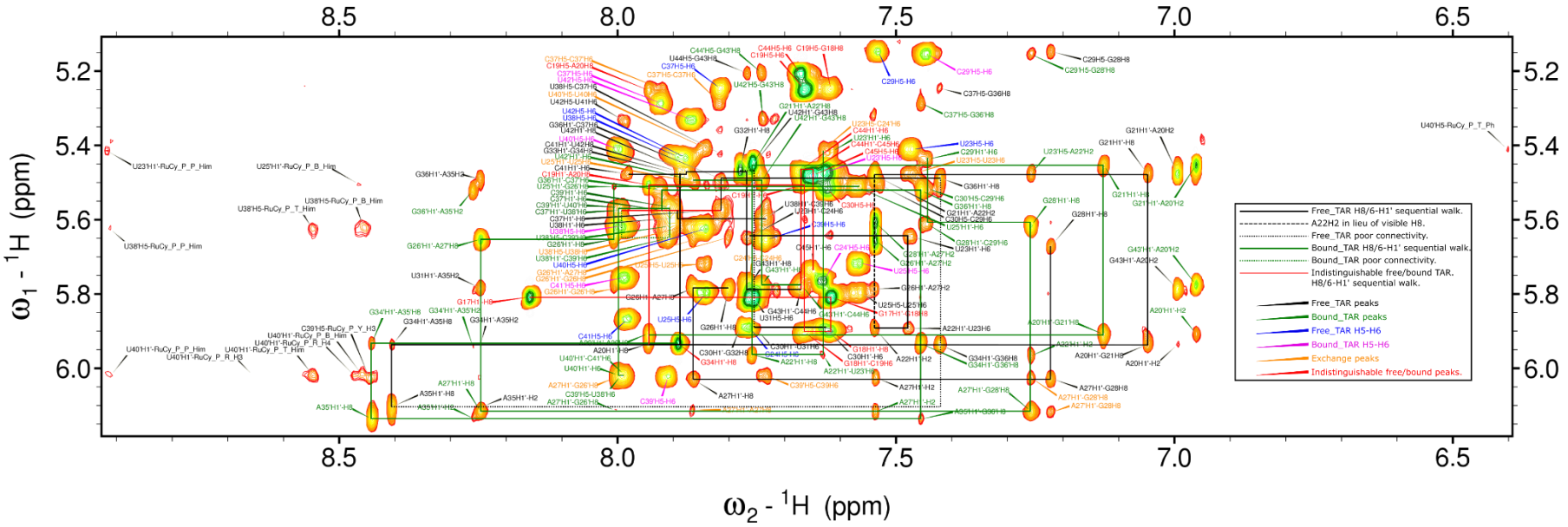


Figure 5.4-4 ^{1}H, ^{1}H-NOESY spectrum showing the H8/6-H1' sequential walks, 700 MHz. In addition, various intermolecular NOEs. 1:0.5 TAR RNA (0.46 mM) with RuCy_P (0.22 mM). Free TAR peaks and walk are in black, a section is indicated with dashed lines as A22H8 is likely buried under other peaks. Bound TAR peaks and walk are in green. Peaks and lines that are the same for both populations are red. Regions of poor connectivity are indicated with dotted lines. Exchange peaks are in orange.

As ascertained by the 90% H_2O NOESY spectra, in 100% D_2O it becomes apparent that a 1:0.5 mixture of TAR to RuCy_P forms both bound and free TAR conformers that exchange in solution. Looking at the H8/6-H1' sequential walk region, certain peaks are the same for free and bound. These can be assumed to be unaffected by RuCy_P binding and are shown in red for the figure. The walk can be traced from G17H1'-H8 to C19H1'-A20H8, and on the opposing side C45H1'-H6 to C44H1'-H6. From A20H8 the bound and unbound populations diverge. The unbound walk is unobserved for A22H8, but as it is present in the free TAR samples, it is likely hidden under other peaks. This considered, on the bulge strand there is a sequential walk to the bulge in both samples ending with U23H1'-C24H6 and a link earlier for the bound sample with U23'H1'-H6. Perhaps this is caused by the flipping out of C24'. On the opposing strand the two populations diverge at G43H1' and G43'H1'. The unbound conformer ending at C41H1'-H6 and the bound at U42'H1'-H6, perhaps as a result of distortion from the cylinder binding. Through the axis of the bulge at C39/U40 we see no connectivity for the free, but for the bound there is a walk from U40'H1'-C41'H6 to U38'H1'-C39'H6 suggesting a reduction of the inter stem angle and a closing of the gap between the A22'-U40' and G26'-C39' pairings.

There are exchange peaks for U23'H5-U23H6, U23H5-C24'H6, C24'H5-C24H6, U25'H5-U25H6, and U25'H1'-U25H6 pointing to a rapid transition between bound and unbound in the bulge. Interestingly many, and those between the same nuclei, are bound-unbound suggesting that the transition to unbound might be faster than the transition to bound. Continuing the walk the bound population begins earlier at U25'H1'-H6 following through a weak U25'H1'-G26'H8 all the way to C29'H1'-H6 at the base of the loop. Unbound continues from G26H1'-H8 through to G28H1'-H8.

There are other exchange peaks present, all are around the bulge area, many are transitions from bound to unbound similar to the bulge bases. On the opposing strand is U40'H5-U40H6 and C39'H5-C39H6. At the top of the bulge there are more exchange peaks, with G26'H1'-G26H8, G26'H1'-A27H8, G26H1'-G26'H8 and A27H1'-G26'H8 and A27'H1'-A27H8. In both directions A27H1'-G28'H8 and A27'H1'-G28H8 are present, almost the furthest up the strand. On the opposing strand is U38'H5-U38H5, and the closest to the loop C37H5-C37'H6.

For both populations the loop, as expected, has regions of poor connectivity. C30H5-H6 and U31H5-H6 are the same for both populations. For the bound TAR there are no observed loop peaks. For the free TAR C30H1'-H6, C30H1'-U31H6, and G32H1'-H8 are present. The walk resumes with G33H1'-G34H8 connected to G34H1'-H8, which is the same for both populations indicated by the red label. As the two conformers are very similar in this region, with little changes it can be assumed that the beginning of the loop is affected little in a 1:0.5 solution. This can be extrapolated to say that the cylinder is certainly interacting very little with the loop. Continuing with the walk at G34', there are again two conformers. G34'H1'-A35'H8 and G34H1'-A35H8 are both present, as are A35'H1'-H8 and A35H1'-H8. However, there is only A35'H1'-G36'H8, suggesting that the A35' is more frequently stacked between G34' and G36'. For the bound and unbound sample there is G34'H1'-G36'H8 and G34H1'-G36H8 suggesting that the A35 and A35' are both mostly looped out. Unlike the TAR and RuCy_P 1:1 spectra, there is only one bound G36'/C37' conformer. With the TAR transitioning rapidly between bound and unbound, alternative conformers are at very low concentrations.

Looking at intermolecular NOEs there is only one weak peak to the phenyl U40'H5-RuCy_P_T_Ph. There are many more pyridine and imine peaks to U40' with weak U40'H1'-RuCy_P_P_Him and U40'H1'-RuCy_P_R_H3; stronger U40'H1'-RuCy_P_T_Him, U40'H1'-RuCy_P_R_H4 and U40'H1'-RuCy_P_B_Him. Most to the bottom of the bulge are the RBT end of the cylinder. At the top of the bulge C39'H5-RuCy_P_Y_H3, and above that medium strength peaks to U38'H5-RuCy_P_T_Him, U38'H5-RuCy_P_B_Him and weak to U38'H5-RuCy_P_P_Him. It is unexpected for the U38' to be in proximity to the same protons as U40', though not if the cylinder is inserted within the bulge with these bases arranged around it. Alternatively, if the RNA exchange peaks are resulting from the motion of the cylinder it would make sense that the cylinder could exchange on the same timescale. Therefore, it would be possible that these peaks could originate from chemical exchange. In the bulge a weak U25'H1'-RuCy_P_B_Him and a medium U23'H1'-RuCy_P_P_Him are present too.

5.4.3 TOCSY

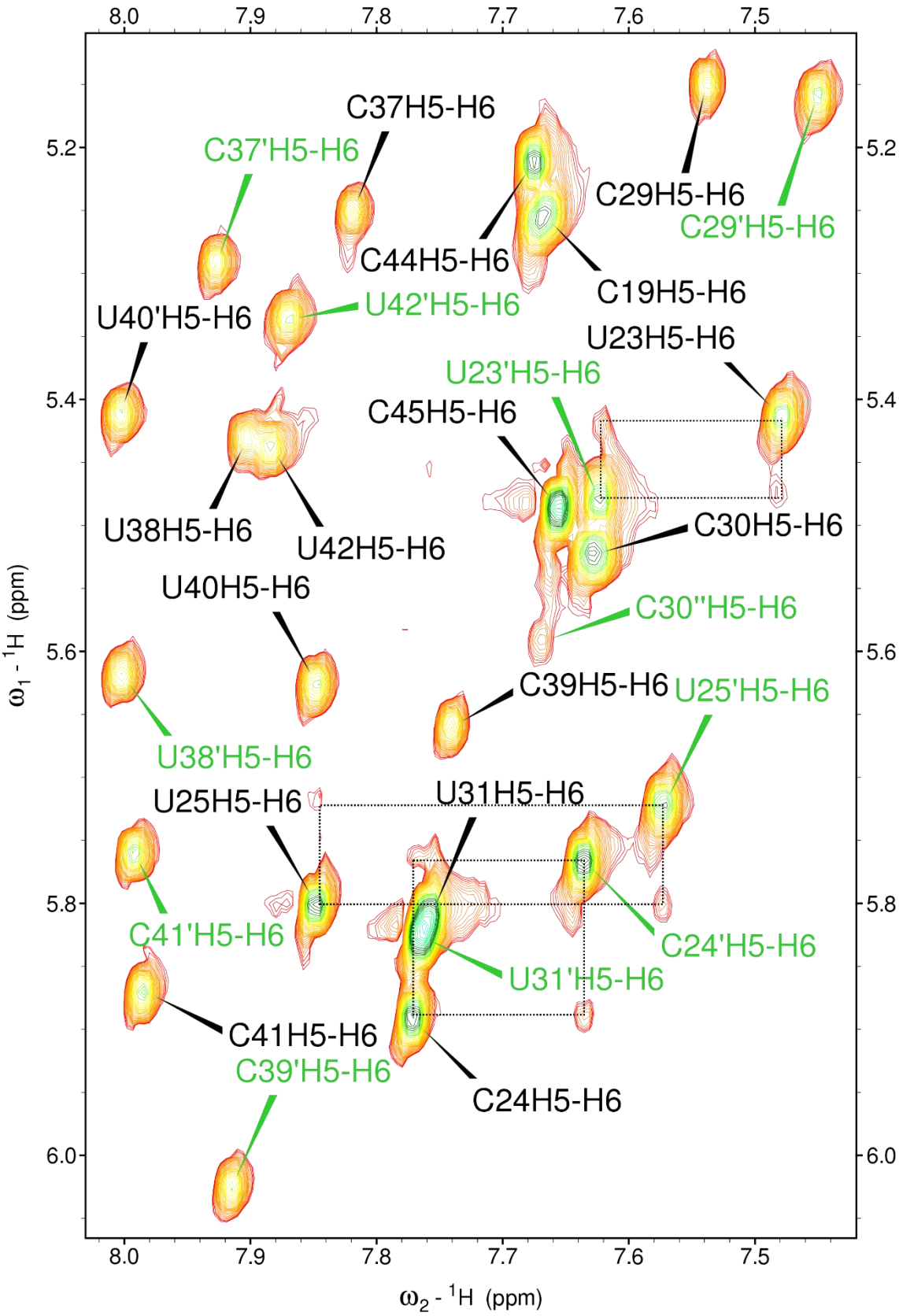

8.0
7.9
7.8
7.7
7.6
7.5
5.2
5.4
5.6
5.8
6.0
C37H5-H6
C37'H5-H6
C29H5-H6
C29'H5-H6
C44H5-H6
C19H5-H6
U40'H5-H6
U42'H5-H6
U23H5-H6
U23'H5-H6
C45H5-H6
C30H5-H6
U38H5-H6
U42H5-H6
U40H5-H6
C30''H5-H6
C39H5-H6
U25'H5-H6
U38'H5-H6
U25H5-H6
U31H5-H6
C41'H5-H6
C24'H5-H6
U31'H5-H6
C41H5-H6
C39'H5-H6
C24H5-H6
ω1 - 1H (ppm)
ω2 - 1H (ppm)

Figure 5.4-5 $^{1}H,^{1}H$-TOCSY spectrum of 1:0.5 TAR RNA (0.46 mM) with RuCy_P (0.22 mM), 700 MHz. Showing RNA H5-H6 peaks. Black. Free TAR, Green. Bound TAR. Black dotted lines indicate exchange peaks.

A TOCSY spectrum of the mixture shows the two distinct sets of H5-H6 peaks. Some bound form peaks are almost indistinguishable from the free TAR. These "immobile" peaks include the C44, C19, and C45, showing little variation in the bottom of the stem. U31 is only slightly shifted and needed to be cross referenced from the NOESY spectra, indicating little variation at the top of the loop stem. C30 is an interesting case, similar to the recently listed peaks it has little shift. However, there is a smaller population that forms another peak labelled C30'', to differentiate from the others. This is discussed further regarding the 1:1 NOESY spectra. (Section 5.3.2.2) But is interesting, as it is not visible in the 1:0.5 NOESY. In stark contrast to the relatively immobile peaks at the ends of the stems, the bulge bases U23, C24, and in particular U25 show large shifts. At the top and bottom of the bulge U40 and C39 show a comparably strong shift. This contributes further evidence pointing to the location of the cylinder near the bulge, engaging U25, U40, and C39. Nearby bases, C41, U42 and U38 show smaller shifts similar to that of C29, pointing to deviation radiating out from the bulge region along the stems.

Exchange peaks are also visible in the TOCSY for U23, C24, and U25 resulting from an exchange between states in this region faster than the rest of the structure. Suggesting the cylinder binding exchanges faster than the RNA conformers.

5.4.4 NOESY 50 ms

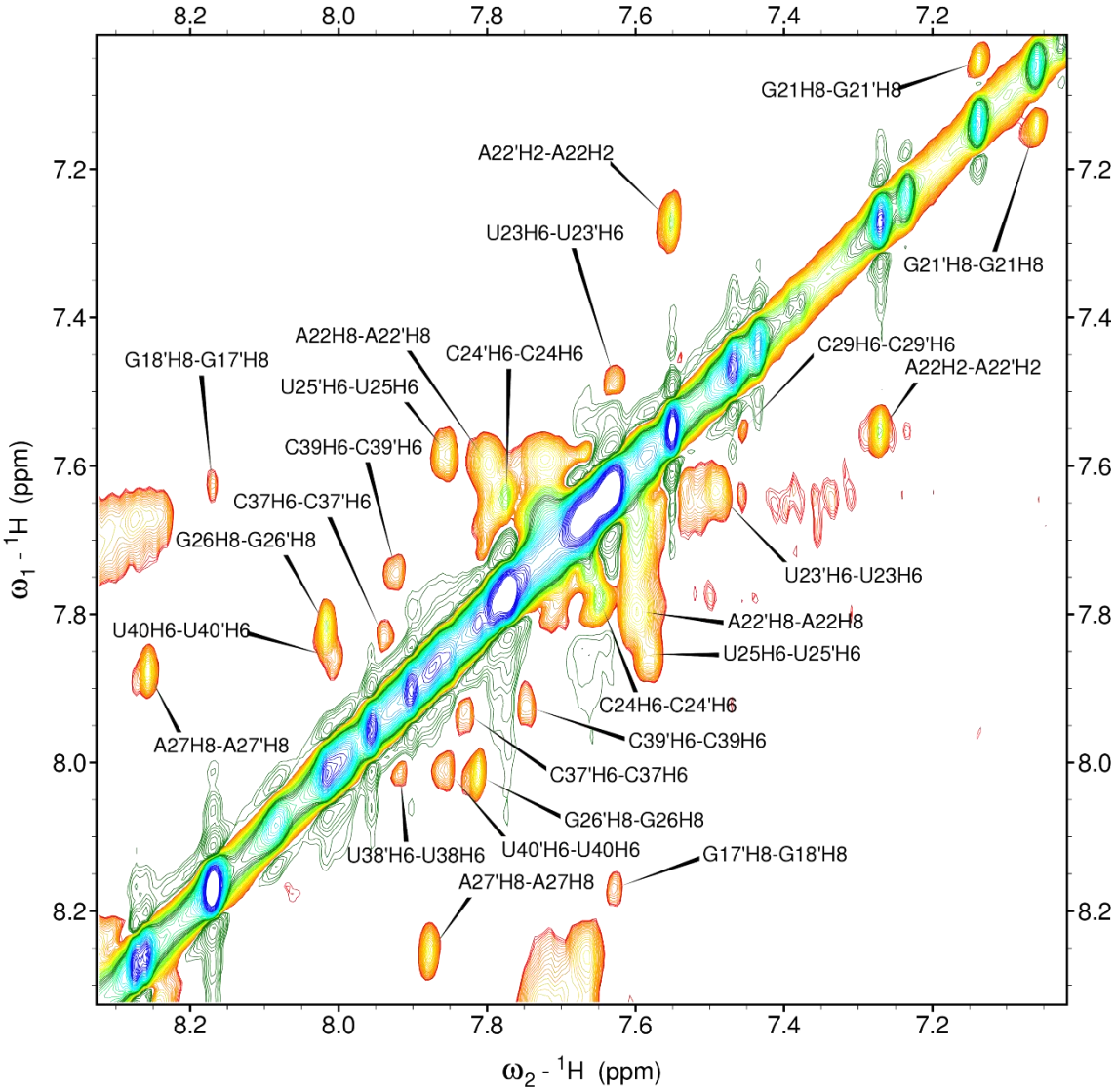


Figure 5.4-6 ^{1}H,^{1}H-NOESY spectrum of 1:0.5 TAR RNA (0.46 mM) with RuCy_P (0.22 mM). Showing aromatic-aromatic short range NOEs and exchange peaks. d8 = 50 ms, 700 MHz.

By reducing the mixing time, medium to long range NOEs have little time to evolve, however exchange peaks that are a result of nuclei exchanging states between excitation and response are visible. So, by doing NOESY with a mixing time of 50 ms, in the region showing aromatic peaks the exchanging H2/H6/H8 peaks and short-range NOEs are observed. (Figure 5.4-6) Notably the peaks visible here are those close to the bulge, with the exception of a short-range NOE for G17'H8-G18'H8. Particularly strong

are all four A22'-A22 and A22-A22', H8 and H2 peaks at the base of the bulge. A27H8-A27'H78, A27'H8-A27H8, G26H8-G26'H8, and G26'H8-G26H8 at the top of the bulge. U25H6-U25'H6, U25'H6-U25H6, C24'H6-C24H6, C24H6-C24'H6 in the bulge. Interestingly the weaker peak intensity of U23H6-U23'H6/U23'H6-U23H6 in comparison to the other bulge exchange peaks matches the lower number of intermolecular NOEs to U23'. Also interesting is there seems to be a trend where peaks from the bulge strand have stronger exchange peaks, and peaks on the opposing strand have weaker exchange peaks. Perhaps this is simply a result of the mobility at the bulge. (Figure 5.4-6)

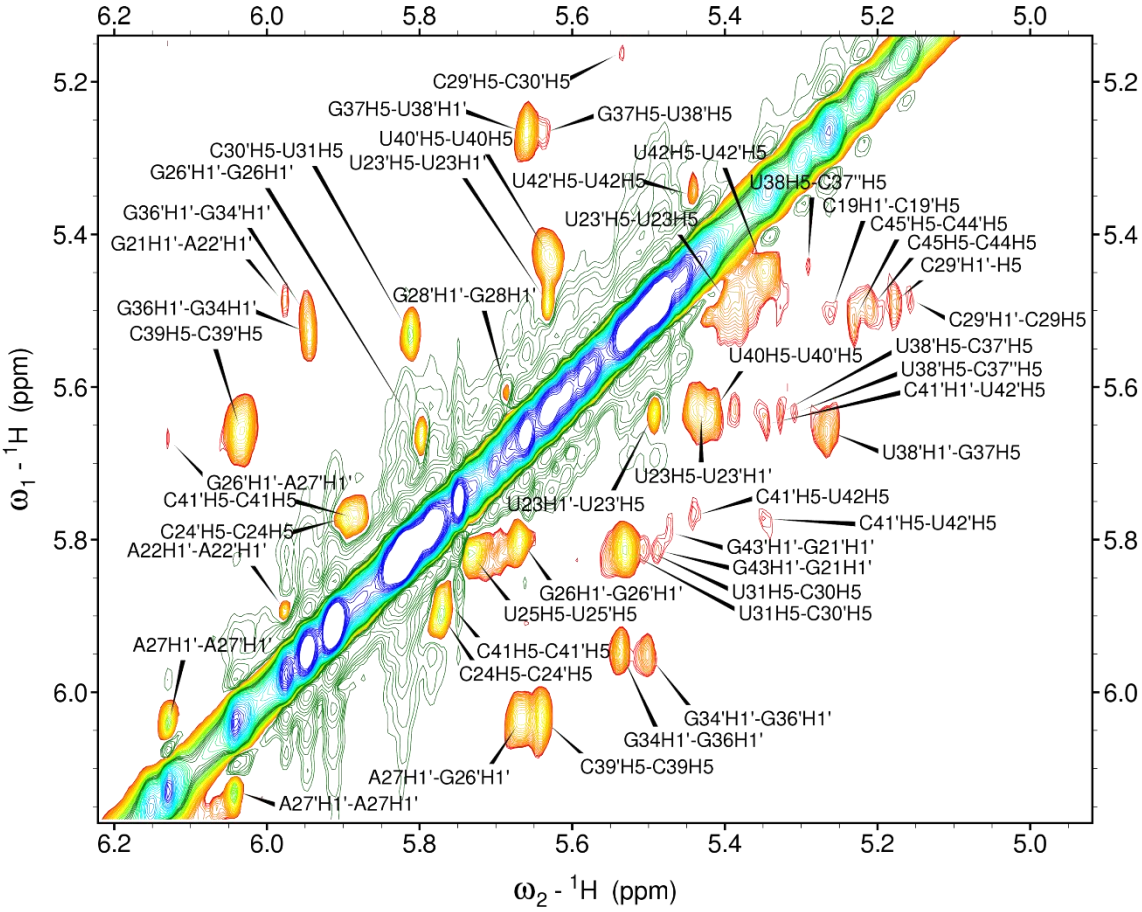


Figure 5.4-7 ^{1}H,^{1}H-NOESY spectrum of 1:0.5 TAR RNA (0.46 mM) with RuCy_P (0.22 mM). Showing H1'/H5 short range NOEs and exchange peaks. d8 = 50 ms, 700 MHz.

Most strong peaks in the H1' region at this mixing time are the result of exchange, though it is possible for H5 and H1' to be close enough for NOE peaks. In this case, many are exchange peaks resulting from nuclei in the bulge region. For instance, strong C39H5-C39'H5, C39'H5-C39H5 and U40'H5-U40H5 peaks. For the loop nucleotides, U25H5-U25'H5 alone could point to a discrepancy between the rate of transition between bound and unbound U25. Interestingly as shown with the TOCSY, (Figure 5.4-5) we again see a degree of mobility at the loop junction, with strong exchange peaks for U31H5-C30'H5, C30'H5-U31H5. There are also strong short-range NOEs between G34H1'-G36H1' and G34'H1'-G36'H1', indicating that G36 is stacked onto G34 in both free and bound conformers. (Figure 5.4-7)

5.4.5 HSQC

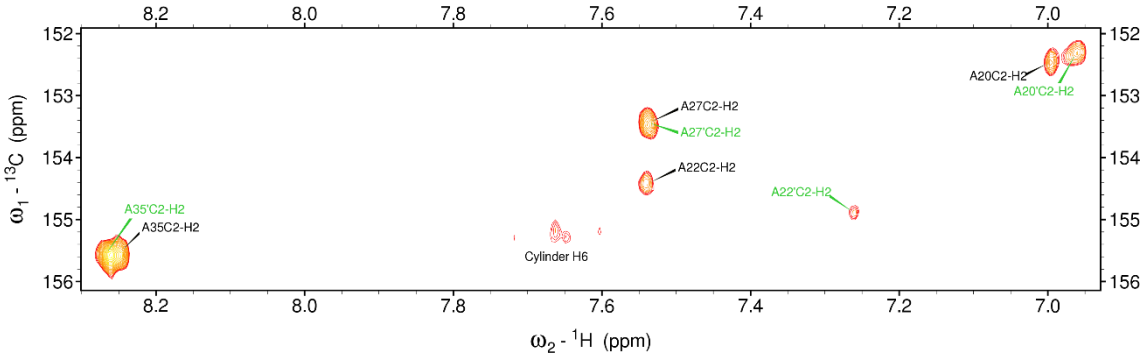


Figure 5.4-8 $^{1}H,^{13}C$-HSQC of a 1:0.5 mixture of TAR to cylinder in D2O, 700 MHz. Showing the two populations present in solution at approximately equal intensitiy, the free TAR with black labels, and the bound TAR in green with the ' suffix. Also present is a collection of cylinder peaks corresponding to the pyridine H6s.

^{1}H-^{13}C HSQC.

In 2006 Q. Zhang and co-workers developed an extended form of the TAR using NMR invisible poly AU and GC segments with the intention of reducing tumbling, demonstrating that there was little effect on the TAR NMR. On binding with argininamide, the A22H2 experiences a large shift, visible in the HSQC.[14] The HSQC reported is very similar to the one seen on binding the cylinder here. A35 and A27 are seemingly unaffected in the bound state, whereas A20 and to a much greater extent A22 are shifted upfield. This pattern would be indicative of a species binding to the bulge, with the bases further away affected to a lesser extent. It is interesting that again there is behaviour similar to argininamide binding, though it could be simply a result of a common binding site rather than common binding mode. (Figure 5.4-8)

In many HSQC spectra, cylinder peaks are weak or absent. This spectrum provided a unique visible set of split peaks, labelled here as "cylinder H6". H6 is visible here as it is the most "protected" from interactions with the RNA. Often ligand binding is described

from the perspective of an RNA binding site, at the 1:0.5 ratio it is worth considering the perspective of the complex. Perhaps these peaks are uniquely visible in the 1:0.5 ratio as there is an excess of TAR pushing the equilibrium toward bound state cylinder.

5.5 Conclusions

Using the ruthenium cylinder, whilst challenging to prepare, is notably easier than other metal centers when applied to NMR. Once prepared a sample can be used without concerns for racemisation, which is valuable for multiple long acquisitions. The cylinder has been previously proven to inhibit tat binding, further research demonstrated the inhibition of HIV-1 infection, and molecular dynamics derived a potential mechanism ending in bulge binding. The results in this chapter show that the cylinder binds to the bulge region. The splitting of the cylinder peaks points to a specific binding conformation and multiple intermolecular NOEs affirms the location, though the specific orientation of the cylinder is uncertain. Two possibilities are proposed, a bulge insertion model and a minor groove binding model. For the minor groove binding model, intermolecular NOEs in the region from C41' to U38' span a distance that matches a parallel orientation of the cylinder to the minor groove. In this model the bulge is closed, with the bulge bases looped over the top of the cylinder to interact with it further. Greater connectivity through the bulge in the bound sample's points to reduction in inter-helical angles in a similar manner to argininamide or tat peptide. For a bulge insertion model, the large range of intermolecular NOEs are explained through the consideration that the open bulge has a high degree of interhelical flexibility allowing the top and bottom of the bulge to be closer than stacked bases. However, there is a strengthening of sequential

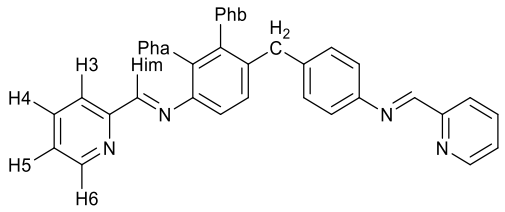
walk peaks between C39' and U40', that would indicate a closing of the bulge. This is a significant difference between these two models that would determine whether the cylinder is inserted within the bulge. NMR residue guided structural calculations may give an indication of the RNA structure onto which the cylinder is bound. Looking at the 1:0.5 mixture we can see that the bound and free forms rapidly exchange, with exchange peaks localised around the bulge. The bulge and its flexibility are essential for tat binding, by locating the cylinder at the bulge it is highly likely this is the mode by which it acts as an inhibitor.

Whilst more NMR experimentation can be done using the TAR and this cylinder, it seems that this avenue of research is unnecessary for now as the question of binding location has been answered. Though further measurements using ^{15}N labelled RNA such as NHN scalar couplings across hydrogen bonds could potentially evidence dynamic pairings across the loop, that is outside of the scope of this thesis. There are, however, other cylinders for which this structure could be applicable and a suitable method to compare them. In addition, there are many more RNA bulges that this cylinder may be appropriate for binding to. In this future research there may be further value in using residual dipolar couplings (RDCs) to learn about the effect of cylinder binding on interhelical angles and mobility. Moving into other techniques it could be valuable to design a modified TAR with bulge binding blocked, to test whether tat binding is still inhibited. Though doing this without also blocking tat binding would be challenging.

6 Experimental

6.1 Ruthenium Cylinder

6.1.1 Parent Ligand L



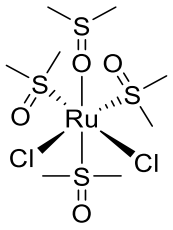
In a 100 mL RBF, 4,4'-methylenedianiline (0.80 g, 4.02 mmol) is dissolved in ethanol (15 mL) with the assistance of sonication. Into this 2-pyridinecarboxaldehyde (0.85 g, 7.95 mmol) in ethanol (10 mL) is added, washed in with ethanol (10 mL). The mixture is stirred for 24 hours. On completion the precipitate is filtered under vacuum and washed with chilled ethanol. The precipitate is then re-crystallised from ethanol and again filtered. The product is dried under high vacuum to yield a pale-yellow powder. (1.34 g, 3.56 mmol, 89% yield).

^{1}H NMR (400 MHz, Chloroform-d) δ 8.74 (dd, J = 4.8, 1.6 Hz, 2H, H6), 8.66 (s, 2H, Him), 8.23 (dd, J = 7.9, 1.2 Hz, 2H, H3), 7.84 (td, J = 7.9, 7.5, 1.6 Hz, 2H, H4), 7.39 (ddd, J = 7.5, 4.8, 1.2 Hz, 2H, H5), 7.30 (m, 8H, HPha and HPhb), 4.08 (s, 2H, CH_2).

Mass spectrum (TOS MS ES+): m/z 337.1766

6.1.2 cis-Tetrakis(dimethylsulphoxide)dichlororuthenium(II)

$Ru(DMSO)_4Cl_2$



In a 25 mL RBF, $RuCl_3 \cdot xH_2O$ (1.00 g, 4.83 mmol) is dissolved in DMSO (10 mL) and heated at 150 °C for 10 minutes. After cooling, half of the solvent is removed under reduced pressure before chilled acetone (20 mL) is added. The precipitate is filtered, washed with chilled acetone (25 mL) and dried overnight under high vacuum. The product is obtained as a yellow solid (0.99 g, 42% yield).

6.1.3 Ruthenium Cylinder $[Ru_2(L)_3]Cl_4$

Dry ethylene glycol is bubbled with argon for 1 hour using vigorous stirring to degas. Parent ligand (0.064 g, 0.170 mmol) and $Ru(DMSO)_4Cl_2$ (0.055 g, 0.113 mmol) were added to a 25 mL RBF, before the ethylene glycol (10 mL) was added. The mixture is heated to 200 °C for 7 days with stirring. The mixture is then kept hot, and a condenser is mounted horizontally with an argon flow and the solvent distilled until dry.

The solid is washed with copious DCM and scraping until the washings are no longer discoloured to yield “crude ruthenium cylinder”.

6.1.3.1 Celite / Cellulose Plug Purification

A pipette is prepared vertically with a cotton plug in the neck. A suspension of cellulose powder in milli-Q H_2O is dropped into the pipette. The eluent is pushed through using a

pipette bulb, taking care not to run dry, until a bed of approx. 1.5 cm of cellulose remains with 1 mm of water above. Onto this a suspension of celite in milli-Q H_2O is pipetted, again pushing the eluent through until this time a bed of 1 mm celite is created. The eluent is then pushed through until the meniscus is in contact with the celite bed.

Crude ruthenium cylinder (ca. 100 mg) is dissolved minimal milli-Q H_2O (1 mL). An aliquot (approx. 0.5 mL) is added onto the surface of the celite, and the eluent pushed through taking care not to run dry. The plug is then washed through with milli-Q H_2O (2 mL), the eluent discarded. The ruthenium cylinder is then eluted with 0.2 M NaCl solution, yielding a bright orange solution. A fresh plug is prepared for each portion of the crude ruthenium cylinder solution and the process is repeated. The collected portions are then lyophilised to yield a mixture of NaCl and ruthenium cylinder powder. The mixture is triturated with methanol and centrifuged keeping the supernatant which is then dried as the partially desalted ruthenium cylinder.

6.1.3.1.1 To further purify as racemic

A gel bed is prepared by hydrating CM C-50 Sephadex resin in milli-Q water for several hours before adding the slurry to a 1.5 cm diameter column, approx. 7 cm high. The gel is then washed with milli-Q water using gravity. The complex is dissolved in milli-Q water and added to a prepared gel bed. The solid ruthenium cylinder is then dissolved in milli-Q H_2O and added to a prepared CM C-50 Sephadex column (approx. 12 cm^2). The column is washed with milli-Q H_2O, then incrementally increasing the concentration of NaCl solution added until the complex is eluted between 0.2 and 0.3 M NaCl.

The solution is then lyophilised, and the solid triturated with methanol. The mixture is then centrifuged, and the supernatant kept and dried.

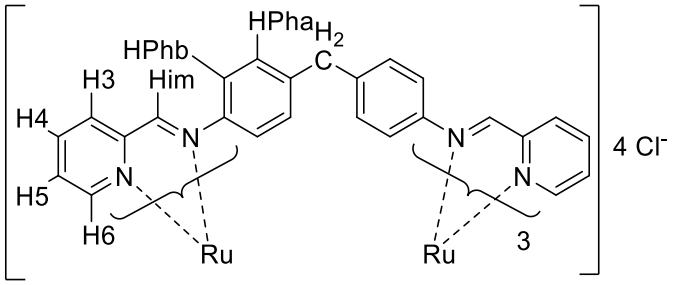
6.1.3.1.2 Alternatively, if enantiomers are required

The dried solid is dissolved in minimal milli-Q H_2O. A pipette is mounted vertically with a cotton plug in the neck. A suspension of cellulose in 0.2 M NaCl solution is dropped into the pipette, the eluent pushed through using a pipette bulb taking care not to run dry. The pipette is topped up with cellulose suspension until a column of approx. 8 cm in length is formed. To the top of this the ruthenium cylinder concentrate (2 mm) is carefully applied, pushing the cylinder into the column with a pipette bulb, followed by a drop of 0.2 M NaCl solution pushed onto the column, and another drop again pushed on. The loaded column is then eluted with 0.2 M NaCl. Commonly the bands are overlapped so first the M is captured in fractions 1-3, the mixture in fraction 4, and the P enantiomer captured in fractions 5-7.

The fractions are then analysed for enantiopurity using circular dichroism and lyophilised.

6.1.3.2 Desalting

The solid is dissolved in milli-Q water before placing onto a prepared G10 Sephadex column. The column is eluted with milli-Q water, discarding the water until orange colouration. The lone orange band is then collected. The water lyophilised to yield the desalted ruthenium cylinder **$[Ru_2(L)_3]4Cl$**.



^{1}H NMR (500 MHz, Deuterium Oxide) δ 8.89 (s, 1H, Him), 8.49 (d, J = 7.9 Hz, 1H, H3), 8.31 (dd, J = 7,8, 7.9 Hz, 1H, H4), 7.78 (d, J = 5.6 Hz, 1H, H6), 7.72 (dd, J = 7.8, 5.6 Hz, 1H, H5), 7.03 (s, 2H, HPhb), 5.78 (d, J = 7.9 Hz, 2H, HPha), 4.04 (s, 1H, CH_2).

^{13}C NMR (126 MHz, Deuterium Oxide) δ 169.84 (s, Cim), 130.67 (s, C3), 138.79 (s, C4), 152.71 (s, C6), 129.41 (s, C5), 129.93 (s, CPhb), 121.09 (s, CPha), 39.00 (s, CH_2)

Mass spectrum (TOF MS EI+): m/z = 333.08 $[Ru_2(L)_3]^{4+}$

UV-Vis (H_2O): (ε / dm^3 mol^{-1} cm^{-1}) 486 (24000), 449 (19000), 318 (50000), 271 (97000)

6.2 TAR RNA Preparation

(GGCAGAUCUGAGCCUGGGAGCUCUCUGCC)

The sequence was prepared using a modified version of the procedure outlined within the paper "The structural stabilization of the κ three-way junction by Mg(II) represents the first step in the folding of a group II intron."[56]

6.2.1 Synthesis

5 = 2'-O-methyl Guanine

NTP equivalents: 10xG, 9xC, 4xA, 6xU

OT = 5' 55CAGAGAGCTCCCAGGCTCAGATCTGCC-TATAGTGAGTCGTATTAATTTC 3'

TS = 5’ GAAATTAATACGACTCACTATAGG 3’

All reagents are carefully defrosted and kept chilled with ice. The parts were combined in a 15 mL centrifuge tube in order: H2O (5880 µL), 5x TXN buffer (2000 µL, 200 mM Tris-HCl pH 7.5, 200 mM DTT, 10 mM spermidine), GTP (229 mM, 301 µL, 69 umol), ATP (219 mM, 126 µL, 27.6 umol), CTP (199 mM, 312 µL, 62.1 umol), UTP (153 mM, 271 µL, 41.4 umol), Triton X-100 (10%, 10 µL), $MgCl_2$ (1 M, 500 µL, 0.5 mmol), OT (100 uM, 50 µL, 5 nmol), TS (100 uM, 50 µL, 5 nmol), mixing between additions.

Finally, T7 polymerase (5’ -> 3‘ synthesis of RNA, on DNA downstream from promoter) chilled in ice is added quickly to the mixture, which is then divided into 3 mL Eppendorf tubes in (10 x 1 mL portions). T7 is manufactured in house and was kindly provided for my use.

The portions are incubated in a thermomixer (37 °C, 300 rpm, 1 hour) forming a precipitate of mainly magnesium pyrophosphate. On completion, the precipitate is span down on a low powered centrifuge. The supernatant is combined into two robust 15 mL centrifuge tubes, and ethanol is added to form a final ratio of 3:1 ethanol/water. Finally, NaCl solution (5 M) is added to make a final concentration of 0.25 M. The mixture is left at -20 °C overnight to precipitate.

6.2.2 Purification

The samples are centrifuged (9500 RCF, 30 min, 4 °C), and the supernatant is poured off. The pellets are then dried under vacuum.

A 20% denaturing poly acrylamide gel is prepared with TBE buffer. Due to its size (approx. 30 x 45 cm), the gel is pre-heated by applying the running potential. The samples are dissolved in TBE buffer, combined, and a loading dye is added. The mixture is then carefully applied to the top of the gel. The gel is run at 10-20W until the loading dye has moved sufficiently (for its distance recommendation as per the sequence length.) On completion the gel is visualised under UV light, and the TAR band is cut out. The gel pieces are frozen when not immediately extracted.

Extracting the RNA from the gel requires electroelution. The Whatman Elutrap is used. A potential of 200V is applied across the apparatus to migrate the RNA through a filter into a chamber where it is trapped between BT1 and BT2 membranes.

On completing the elution, the RNA is extracted in solution from a well between the membranes, precipitated with 3:1 ethanol/water, and centrifuged to collect the RNA. (9500 RCF, 30 min, 4 °C) The supernatant is poured away. To prep the Vivaspin 2 mL centrifuge device it is washed twice with 70% ethanol solution (2 mL), followed by ultrapure water (2 x 2 mL), pouring away the supernatants. The RNA is dissolved in ultrapure water and added to the Vivaspin tubes where it is concentrated by centrifugation (approx. 1000-4000 RCF, 30-60 minutes, 4 °C) and the supernatant is poured away. The RNA is washed with KCl solution to displace the TRIS (2 x 2 mL, 1 M) Followed by ultrapure water (4 x 2 mL); it is centrifuged as above after each addition. The RNA is resuspended in ultrapure water, transferred to a centrifuge tube and frozen.

The RNA concentration is ascertained from UV-Vis using a calculated extinction coefficient (ε_{258} = 268900 $M^{-1}cm^{-1}$). Using this and the volume the molar yield of RNA is

known. A small portion is taken for MALDI-MS [$M+H^+$] = 9.53 kDa, and the solution is then lyophilised.

6.3 NMR

RNA stock solutions prepared above are diluted to a final RNA concentration between 0.5 and 1.5 mM, at a final volume of 220 µL, with 50 mM NaCl at pH 7. The sample is annealed at 95 °C for 5 minutes, followed by a slow cool to room temperature. For 100% D_2O samples, the 90/10 H_2O/D_2O sample is lyophilised. The sample is then dissolved in 100% D2O of the same total volume as was removed. It is again annealed.

RNA is measured using 5 mm Shigemi tubes on either Brucker AV600 or AV700 with a CP-TCI z-axis pulsed field gradient CryoProbe®. Spectra are initially processed using Bruker TopSpin. Assignments and figures are made using Sparky.[57] RNA and complex are assigned separately, avoiding consideration for the implications of intermolecular NOEs as much as possible until assignment was complete, this is to avoid bias.

DOSY spectra are processed using Bruker Dynamics center. Peaks are assigned as manually selected regions for integration.

Other NMR measurements are conducted using either a Bruker AVANCE NEO 400 fitted with a 5 mm BBFO SmartProbe or a Bruker AVANCE NEO 500 fitted with a 5 mm nitrogen-cooled Prodigy BBFO CryoProbe®. Assignments and figures are made using MestreNova.

Pulse programs used in this study are:

- ^{1}H-NMR (zgesgp/zgpr)

- $^{1}H,^{1}H$-NOESY 90% H_2O (rs_noesywatg)
- $^{1}H,^{13}C$-HSQC 100% D_2O (hsqcetgpsisp2.2/hsqcetgpsi/hsqcetgpsi2)
- $^{1}H,^{1}H$-TOCSY 100% D_2O (mlevphpr.2/mlevesgpph)
- $^{1}H,^{1}H$-NOESY 100% D_2O (noesyesgpph)
- $^{1}H,^{1}H$-COSY 100% D_2O (cosydfphpr/cosydfgpph19)
- ^{1}H-DOSY 100% D_2O Or 90% H_2O (ledbpgp2s/ledbpgp2s_es)

6.4 UV-Vis / Circular Dichroism Spectroscopy

UV-Vis spectroscopy is acquired using a Cary Varian 5000. Initial measurements are between 200 and 800 nm.

CD spectroscopy is acquired using an Applied Photophysics Chirascan Plus Spectrometer.

7 References

1 F. Aboul-ela, J. Karn and G. Varani, The Structure of the Human Immunodeficiency Virus Type-1 TAR RNA Reveals Principles of RNA Recognition by Tat Protein, *J. Mol. Biol.*, 1995, **253**, 313–332.

2 D. K. Hendrix, S. E. Brenner and S. R. Holbrook, RNA structural motifs: Building blocks of a modular biomolecule, *Q. Rev. Biophys.*, 2005, **38**, 221–243.

3 S. Bannwarth and A. Gatignol, HIV-1 TAR RNA: The Target of Molecular Interactions Between the Virus and its Host, *Curr. HIV Res.*, 2005, **3**, 61–71.

4 C. Brigati, M. Giacca, D. M. Noonan and A. Albini, HIV Tat, its TARgets and the control of viral gene expression, *FEMS Microbiol. Lett.*, 2003, **220**, 57–65.

5 D. L. Ouellet, I. Plante, P. Landry, C. Barat, M. È. Janelle, L. Flamand, M. J. Tremblay and P. Provost, Identification of functional microRNAs released through asymmetrical processing of HIV-1 TAR element, *Nucleic Acids Res.*, 2008, **36**, 2353–2365.

6 Z. Klase, R. Winograd, J. Davis, L. Carpio, R. Hildreth, M. Heydarian, S. Fu, T. McCaffrey, E. Meiri, M. Ayash-Rashkovsky, S. Gilad, Z. Bentwich and F. Kashanchi, HIV-1 TAR miRNA protects against apoptosis by altering cellular gene expression, *Retrovirology*, 2009, **6**, 1–17.

7 M. Jalalirad, J. Saadatmand and M. Laughrea, Dominant role of the 5′ TAR bulge in dimerization of HIV-1 genomic RNA, but no evidence of TAR-TAR kissing during in vivo virus assembly, *Biochemistry*, 2012, **51**, 3744–3758.

8 F. Aboul-ela, Structure of HIV-1 TAR RNA in the absence of ligands reveals a novel

conformation of the trinucleotide bulge, *Nucleic Acids Res.*, 1996, **24**, 3974–3981.
9 J. Malina, M. J. Hannon and V. Brabec, Recognition of DNA bulges by dinuclear iron(II) metallosupramolecular helicates, *FEBS J.*, 2014, **281**, 987–997.
10 J. Malina, M. J. Hannon and V. Brabec, Iron(II) supramolecular helicates interfere with the HIV-1 Tat-TAR RNA interaction critical for viral replication, *Sci. Rep.*, 2016, **6**, 29674.
11 L. Cardo, I. Nawroth, P. J. Cail, J. A. McKeating and M. J. Hannon, Metallo supramolecular cylinders inhibit HIV-1 TAR-TAT complex formation and viral replication in cellulo, *Sci. Rep.*, 2018, **8**, 13342.
12 L. Melidis, I. B. Styles and M. J. Hannon, Targeting structural features of viral genomes with a nano-sized supramolecular drug, *Chem. Sci.*, 2021, **12**, 7174–7184.
13 D. P. Buck, C. B. Spillane, J. G. Collins and F. R. Keene, Binding of a dinuclear ruthenium(ii) complex to the TAR region of the HIV-AIDS viral RNA, *Mol. Biosyst.*, 2008, **4**, 851.
14 Q. Zhang, X. Sun, E. D. Watt and H. M. Al-Hashimi, Resolving the Motional Modes That Code for RNA Adaptation, *Science (80-.).*, 2006, **311**, 653–656.
15 S. W. Pitt, A. Majumdar, A. Serganov, D. J. Patel and H. M. Al-Hashimi, Argininamide binding arrests global motions in HIV-1 TAR RNA: Comparison with Mg2+-induced conformational stabilization, *J. Mol. Biol.*, 2004, **338**, 7–16.
16 M. Zacharias and P. J. Hagerman, The bend in RNA created by the trans-activation response element bulge of human immunodeficiency virus is straightened by arginine and by Tat-derived peptide, *Proc. Natl. Acad. Sci. U. S. A.*, 1995, **92**, 6052–6056.
17 N. Leulliot and G. Varani, Current topics in RNA-protein recognition: Control of specificity and biological function through induced fit and conformational capture, *Biochemistry*, 2001, **40**, 7947–7956.
18 G. L. Olsen, M. F. Bardaro, D. C. Echodu, G. P. Drobny and G. Varani, Intermediate Rate Atomic Trajectories of RNA by Solid-State NMR Spectroscopy, *J. Am. Chem. Soc.*, 2010, **132**, 303–308.
19 M. H. Bailor, X. Sun and H. M. Al-Hashimi, Topology Links RNA Secondary Structure with Global Conformation, Dynamics, and Adaptation, *Science (80-.).*, 2010, **327**, 202–206.
20 Q. Zhang, A. C. Stelzer, C. K. Fisher and H. M. Al-Hashimi, Visualizing spatially correlated dynamics that directs RNA conformational transitions, *Nature*, 2007, **450**, 1263–1267.
21 H. M. Al-Hashimi, S. W. Pitt, A. Majumdar, W. Xu and D. J. Patel, Mg2+-induced variations in the conformation and dynamics of HIV-1 TAR RNA probed using NMR residual dipolar couplings, *J. Mol. Biol.*, 2003, **329**, 867–873.
22 I. A. Belashov, D. W. Crawford, C. E. Cavender, P. Dai, P. C. Beardslee, D. H. Mathews, B. L. Pentelute, B. R. McNaughton and J. E. Wedekind, Structure of HIV TAR in complex with a Lab-Evolved RRM provides insight into duplex RNA recognition and synthesis of a constrained peptide that impairs transcription, *Nucleic Acids Res.*, 2018, **46**, 6401–6415.
23 M. D. Shortridge, P. T. Wille, A. N. Jones, A. Davidson, J. Bogdanovic, E. Arts, J. Karn, J. A. Robinson and G. Varani, An ultra-high affinity ligand of HIV-1 TAR reveals the RNA structure recognized by P-TEFb, *Nucleic Acids Res.*, 2019, **47**, 1523–1531.
24 J. D. Puglisi, R. Tan, B. J. Calnan, A. D. Frankel and J. R. Williamson, Conformation of the TAR RNA-Arginine Complex by NMR Spectroscopy, *Science (80-.).*, 1992, **257**, 76–80.
25 J. D. Puglisi, L. Chen, A. D. Frankel and J. R. Williamson, Role of RNA structure in arginine recognition of TAR RNA, *Proc. Natl. Acad. Sci. U. S. A.*, 1993, **90**, 3680–3684.
26 J. Tao, L. Chen and A. D. Frankel, Dissection of the proposed base triple in human immunodeficiency virus TAR RNA indicates the importance of the Hoogsteen interaction, *Biochemistry*, 1997, **36**, 3491–3495.
27 M. Hennig and J. R. Williamson, Detection of N-H ··· N hydrogen bonding in RNA via scalar couplings in the absence of observable imino proton resonances, *Nucleic Acids Res.*, 2000, **28**, 1585–1593.

28 R. Nifosì, C. M. Reyes and P. A. Kollman, Molecular dynamics studies of the HIV-1 TAR and its complex with argininamide, *Nucleic Acids Res.*, 2000, **28**, 4944–4955.
29 U. Schulze-Gahmen and J. H. Hurley, Structural mechanism for HIV-1 TAR loop recognition by Tat and the super elongation complex, *Proc. Natl. Acad. Sci.*, 2018, **115**, 12973–12978.
30 S. Richter, H. Cao and T. M. Rana, Specific HIV-1 TAR RNA loop sequence and functional groups are required for human cyclin T1-Tat-TAR ternary complex formation., *Biochemistry*, 2002, **41**, 6391–7.
31 T. Kulinski, M. Olejniczak, H. Huthoff, L. Bielecki, K. Pachulska-Wieczorek, A. T. Das, B. Berkhout and R. W. Adamiak, The apical loop of the HIV-1 TAR RNA hairpin is stabilized by a cross-loop base pair, *J. Biol. Chem.*, 2003, **278**, 38892–38901.
32 E. A. Dethoff, A. L. Hansen, C. Musselman, E. D. Watt, I. Andricioaei and H. M. Al-Hashimi, Characterizing complex dynamics in the transactivation response element apical loop and motional correlations with the bulge by NMR, molecular dynamics, and mutagenesis, *Biophys. J.*, 2008, **95**, 3906–3915.
33 E. A. Dethoff, K. Petzold, J. Chugh, A. Casiano-Negroni and H. M. Al-Hashimi, Visualizing transient low-populated structures of RNA, *Nature*, 2012, **491**, 724–728.
34 G. I. Pascu, A. C. G. Hotze, C. Sanchez-Cano, B. M. Kariuki and M. J. Hannon, Dinuclear ruthenium(II) triple-stranded helicates: Luminescent supramolecular cylinders that bind and coil DNA and exhibit activity against cancer cell lines, *Angew. Chemie - Int. Ed.*, 2007, **46**, 4374–4378.
35 A. Łęczkowska, PhD Thesis, University of Birmingham, 2011.
36 J. Malina, M. J. Hannon and V. Brabec, Interaction of dinuclear ruthenium(II) supramolecular cylinders with DNA: Sequence-specific binding, unwinding, and photocleavage, *Chem. - A Eur. J.*, 2008, **14**, 10408–10414.
37 P. J. Cail, PhD Thesis, University of Birmingham, 2017.
38 J. C. White, PhD Thesis, University of Birmingham, 2015.
39 I. Bratsos and E. Alessio, eds. T. B. Rauchfuss and M. E. Ringenberg, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2010, pp. 148–163.
40 B. R. James, E. Ochiai and G. L. Rampel, Ruthenium(II) halide dimethylsulphoxide complexes from hydrogenation reactions, *Inorg. Nucl. Chem. Lett.*, 1971, **7**, 781–784.
41 J. S. Jaswal, S. J. Rettig and B. R. James, Ruthenium(III) complexes containing dimethylsulfoxide or dimethylsulfide ligands, and a new route to trans - dichlorotetrakis(dimethylsulfoxide)ruthenium(II), *Can. J. Chem.*, 1990, **68**, 1808–1817.
42 S. P. Pinho and E. A. Macedo, Solubility of NaCl, NaBr, and KCl in water, methanol, ethanol, and their mixed solvents, *J. Chem. Eng. Data*, 2005, **50**, 29–32.
43 M. J. Hannon, I. Meistermann, C. J. Isaac, C. Blomme, A. Rodger and J. R. Aldrich-Wright, Paper: a cheap yet effective chiral stationary phase for chromatographic resolution of metallo-supramolecular helicates, *Chem. Commun.*, 2001, 1078–1079.
44 J. A. Norman, PhD Thesis, University of Birmingham, 2012.
45 K. L. Flint, J. G. Collins, S. J. Bradley, T. A. Smith, C. J. Sumby and F. R. Keene, Synthesis and Characterisation of Helicate and Mesocate Forms of a Double-Stranded Diruthenium(II) Complex of a Di(terpyridine) Ligand, *Aust. J. Chem.*, 2019, **72**, 762–768.
46 S. J. Allison, D. Cooke, F. S. Davidson, P. I. P. Elliott, R. A. Faulkner, H. B. S. Griffiths, O. J. Harper, O. Hussain, P. J. Owen-Lynch, R. M. Phillips, C. R. Rice, S. L. Shepherd and R. T. Wheelhouse, Ruthenium-Containing Linear Helicates and Mesocates with Tuneable p53-Selective Cytotoxicity in Colorectal Cancer Cells, *Angew. Chemie - Int. Ed.*, 2018, **57**, 9799–9804.
47 Z. Zhang and D. Dolphin, A triple-stranded helicate and mesocate from the same metal and ligand, *Chem. Commun.*, 2009, 6931–6933.

48 L. F. Sukhodub, Interactions and hydration of nucleic acid bases in a vacuum. Experimental study, *Chem. Rev.*, 1987, **87**, 589–606.
49 D. K. Merriman, Y. Xue, S. Yang, I. J. Kimsey, A. Shakya, M. Clay and H. M. Al-Hashimi, Shortening the HIV-1 TAR RNA Bulge by a Single Nucleotide Preserves Motional Modes over a Broad Range of Time Scales, *Biochemistry*, 2016, **55**, 4445–4456.
50 K. S. Long and D. M. Crothers, Characterization of the Solution Conformations of Unbound and Tat Peptide-Bound Forms of HIV-1 TAR RNA, *Biochemistry*, 1999, **38**, 10059–10069.
51 J. Lee, E. A. Dethoff and H. M. Al-Hashimi, Invisible RNA state dynamically couples distant motifs, *Proc. Natl. Acad. Sci.*, 2014, **111**, 9485–9490.
52 J. A. Jaeger and I. Tinoco, An NMR Study of the HIV-1 TAR Element Hairpin, *Biochemistry*, 1993, **32**, 12522–12530.
53 S. Phongtongpasuk, S. Paulus, J. Schnabl, R. K. O. Sigel, B. Spingler, M. J. Hannon and E. Freisinger, Binding of a Designed Anti-Cancer Drug to the Central Cavity of an RNA Three-Way Junction, *Angew. Chemie Int. Ed.*, 2013, **52**, 11513–11516.
54 L. Cerasino, M. J. Hannon and E. Sletten, DNA three-way junction with a dinuclear iron(II) supramolecular helicate at the center: a NMR structural study., *Inorg. Chem.*, 2007, **46**, 6245–51.
55 F. A. Abulwerdi, M. D. Shortridge, J. Sztuba-Solinska, R. Wilson, S. F. J. Le Grice, G. Varani and J. S. Schneekloth, Development of Small Molecules with a Noncanonical Binding Mode to HIV-1 Trans Activation Response (TAR) RNA., *J. Med. Chem.*, 2016, **59**, 11148–11160.
56 D. Donghi, M. Pechlaner, C. Finazzo, B. Knobloch and R. K. O. Sigel, The structural stabilization of the i three-way junction by Mg(II) represents the first step in the folding of a group II intron, *Nucleic Acids Res.*, 2013, **41**, 2489–2504.
57 W. Lee, M. Tonelli and J. L. Markley, NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy, *Bioinformatics*, 2015, **31**, 1325–1327.

8 Appendices

8.1 Free TAR

8.1.1 HSQC

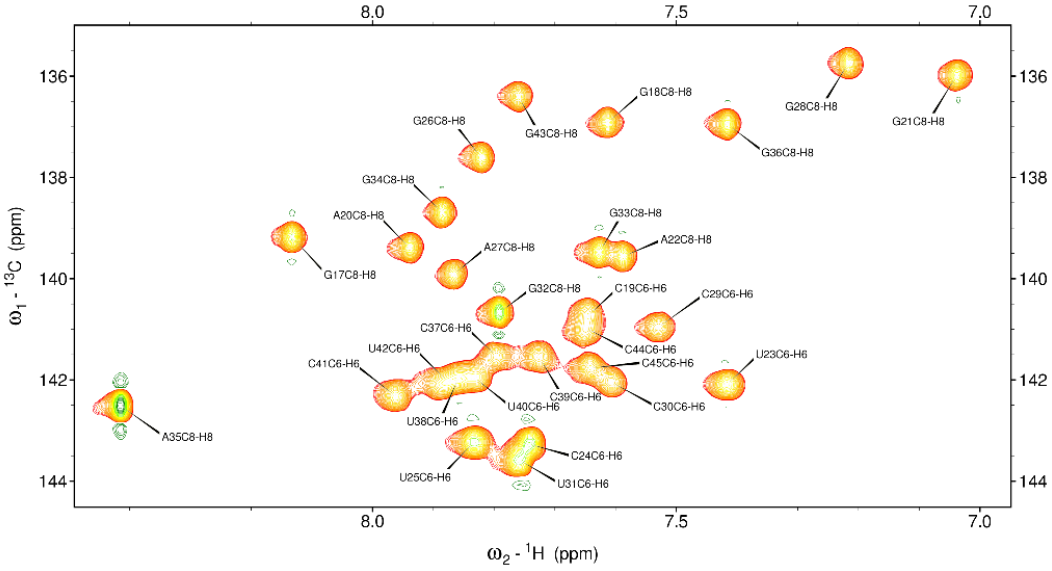


Figure 8.1-1 $^{1}H,^{13}C$-HSQC of 1.5 mM TAR RNA in D2O at 298K, 700 MHz. pH 7, 50 mM NaCl. Showing the C8-H8/C6-H6 peaks.

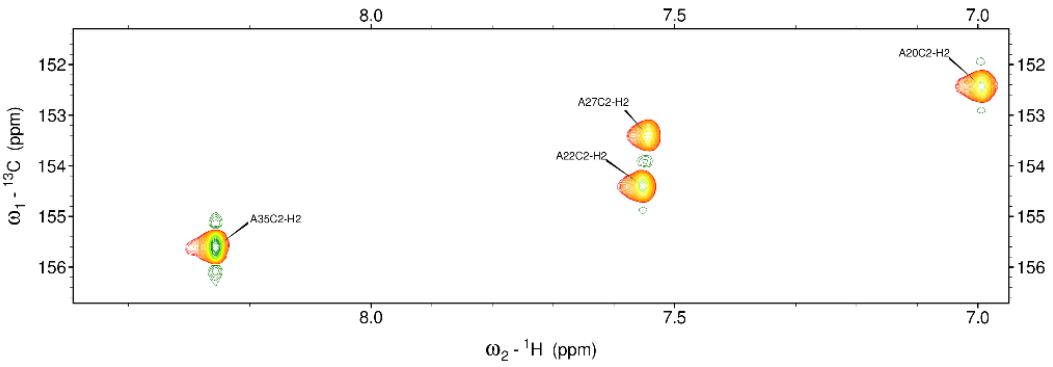


Figure 8.1-2 $^{1}H,^{13}C$-HSQC of 1.5 mM TAR RNA in D2O at 298K, 700 MHz. pH 7, 50 mM NaCl. Showing the C2-H2 peaks.

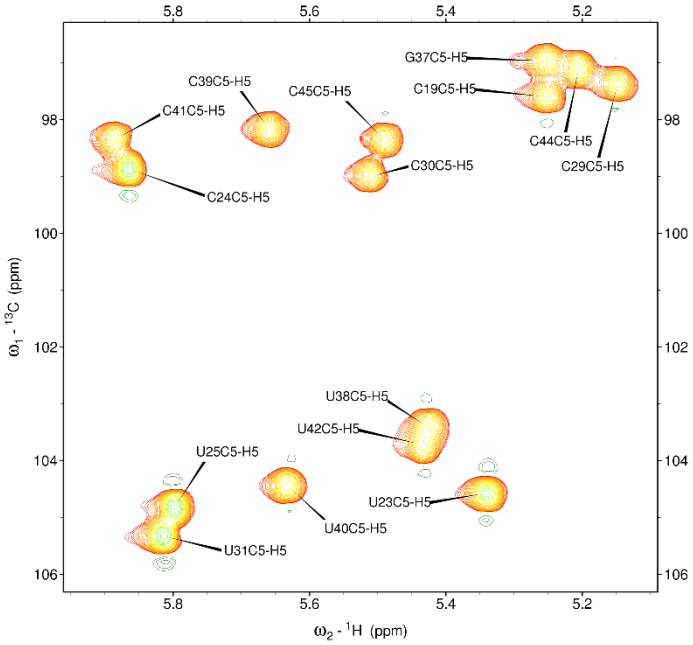


Figure 8.1-3 $^{1}H,^{13}C$-HSQC of 1.5 mM TAR RNA in D2O at 298K, 700 MHz. pH 7, 50 mM NaCl. Showing the C5-H5 peaks.

8.1.2 TOCSY

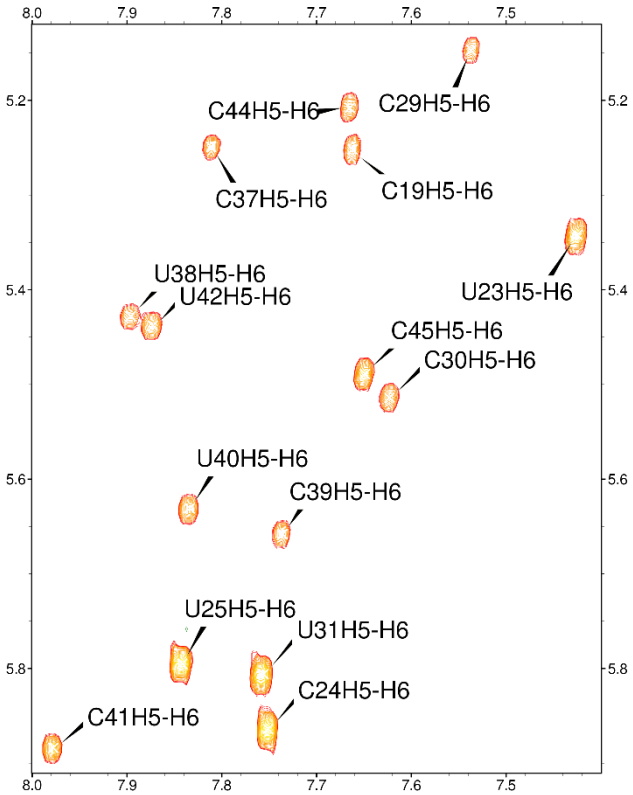


Figure 8.1-4 $^{1}H,^{1}H$-TOCSY spectrum of 1.5 mM TAR RNA in D2O at 298K, 700 MHz. pH 7, 50 mM NaCl. Showing the H5-H6 peaks.

8.2 Bound TAR 1:1

8.2.1 RNA

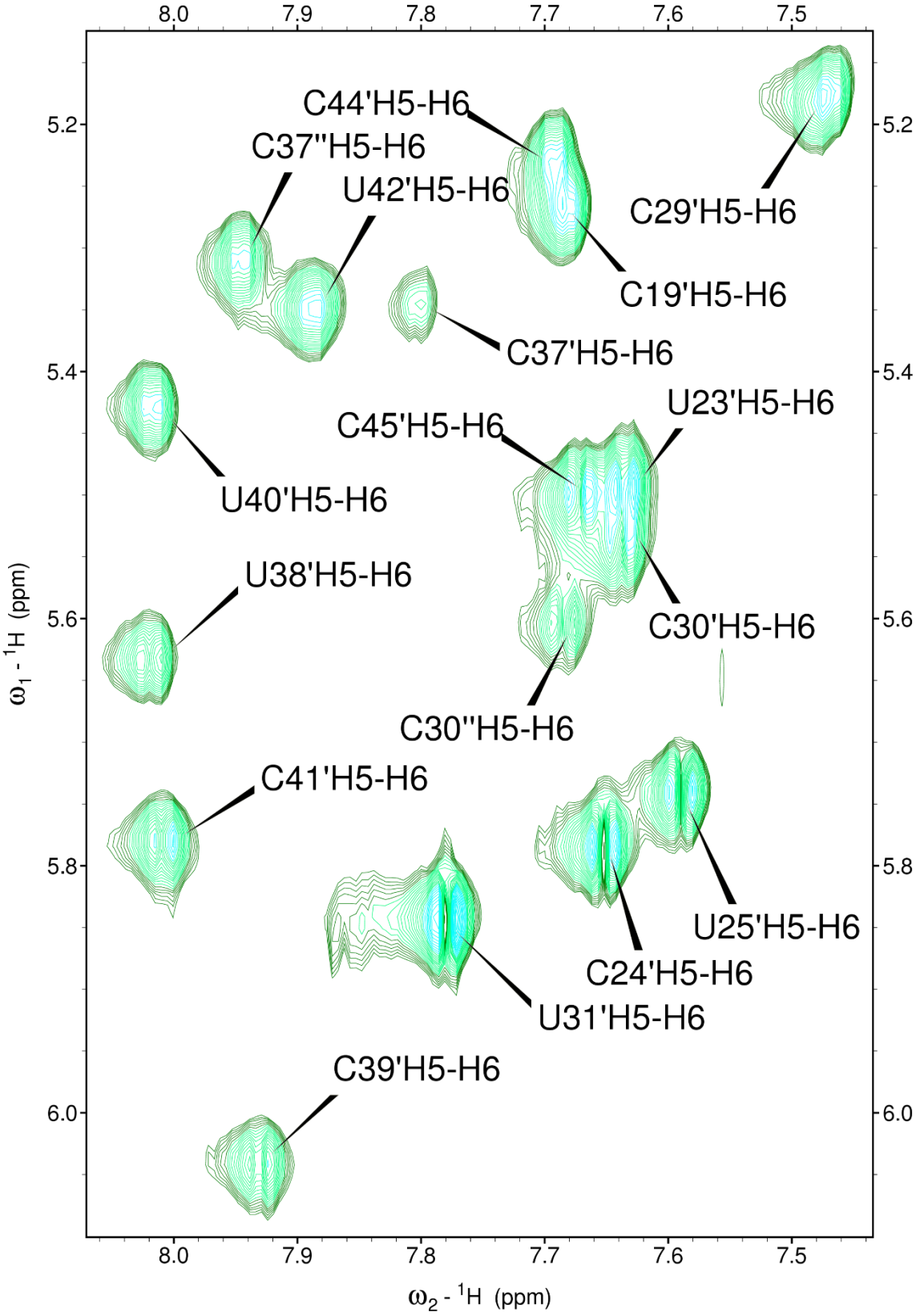


Figure 8.2-1 ^{1}H,^{1}H-TOCSY spectrum of 1:1 TAR RNA (0.49 mM) with RuCy_P (0.49 mM) in D2O, 700 MHz. 50 mM NaCl, pH 7, 298K.

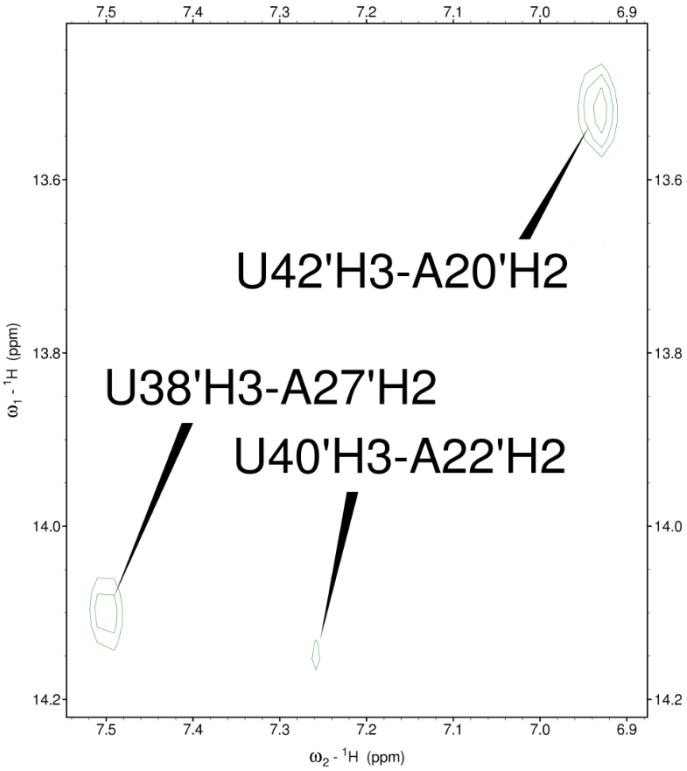


Figure 8.2-2 ^{1}H, ^{1}H-TOCSY spectra of TAR RNA with RuCy_P 1to1 278K 90% H2O, 700 MHz. Showing the H2 to imine peaks.

8.2.2 Intermolecular NOEs

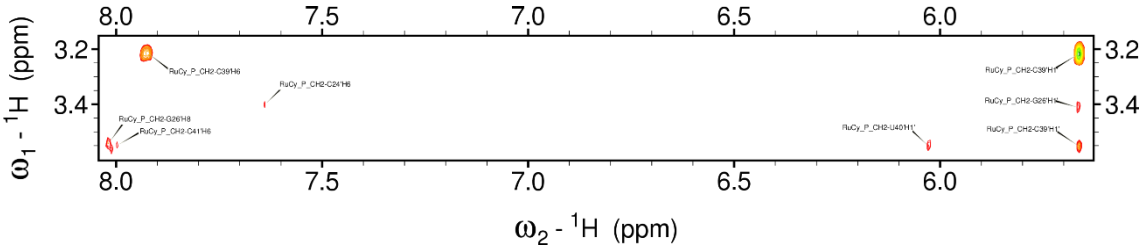


Figure 8.2-3 ^{1}H, ^{1}H-NOESY spectrum of TAR RNA (0.49 mM) with RuCy_P (0.49 mM) in D2O. 50 mM NaCl, pH 7, 298K, 700 MHz. Showing the intermolecular RuCy_P_CH2 to RNA peaks.

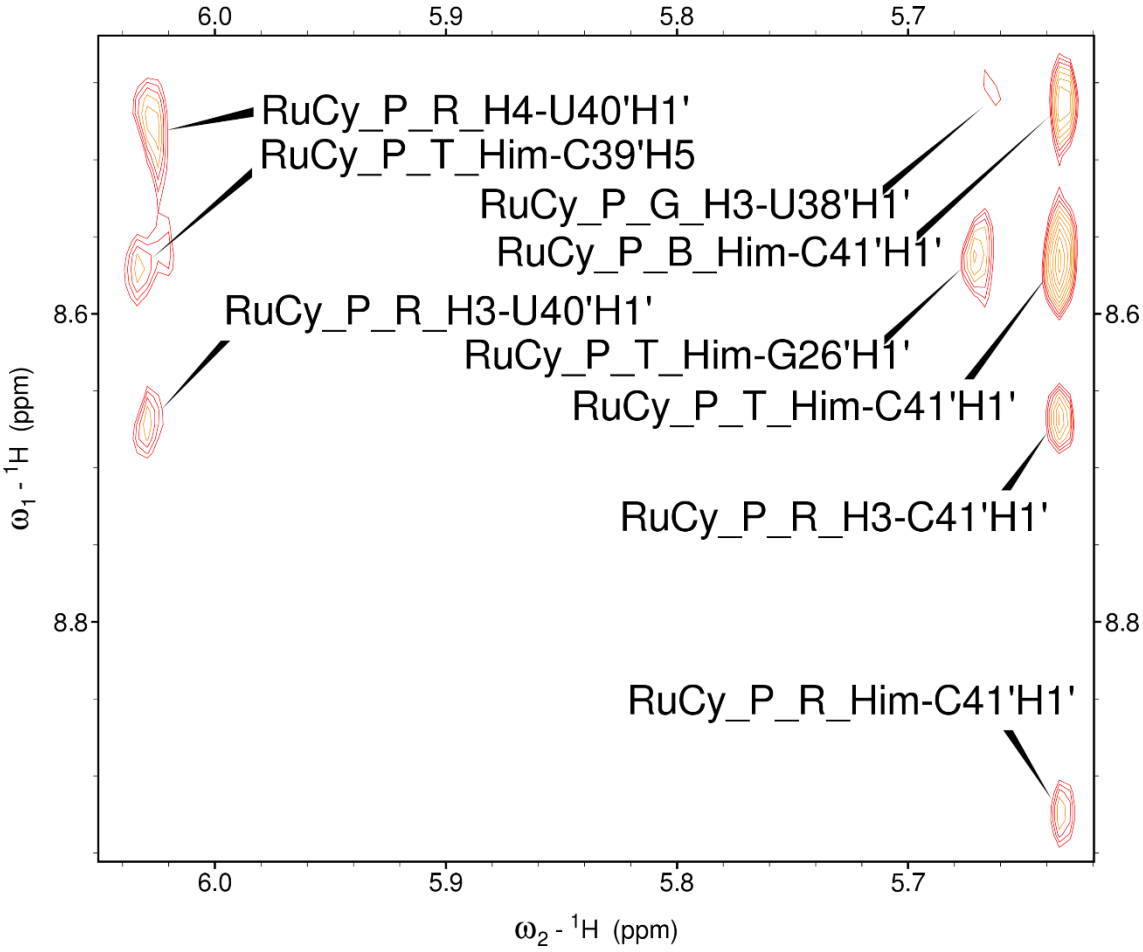


Figure 8.2-4 ^{1}H, ^{1}H-NOESY spectrum of TAR RNA (0.49 mM) with RuCy_P (0.49 mM) in D2O. 50 mM NaCl, pH 7, 298K, 700 MHz. Showing the intermolecular RuCy_P pyridine and imine to RNA peaks.

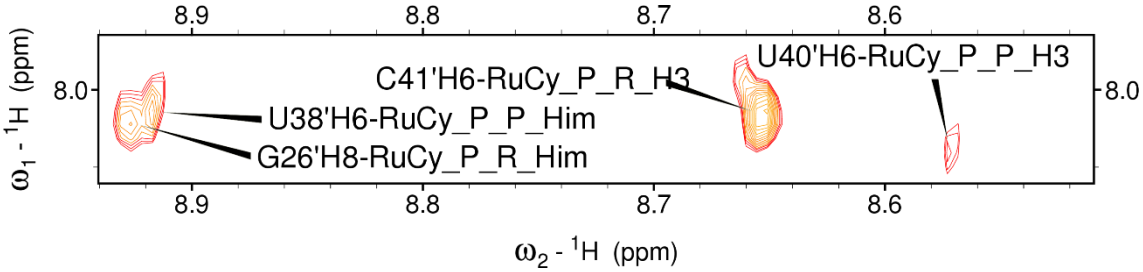


Figure 8.2-5 ^{1}H, ^{1}H-NOESY spectrum of TAR RNA (0.49 mM) with RuCy_P (0.49 mM) in D2O. 50 mM NaCl, pH 7, 278K, 700 MHz. Showing the intermolecular RNA to RuCy_P pyridine and imine peaks.

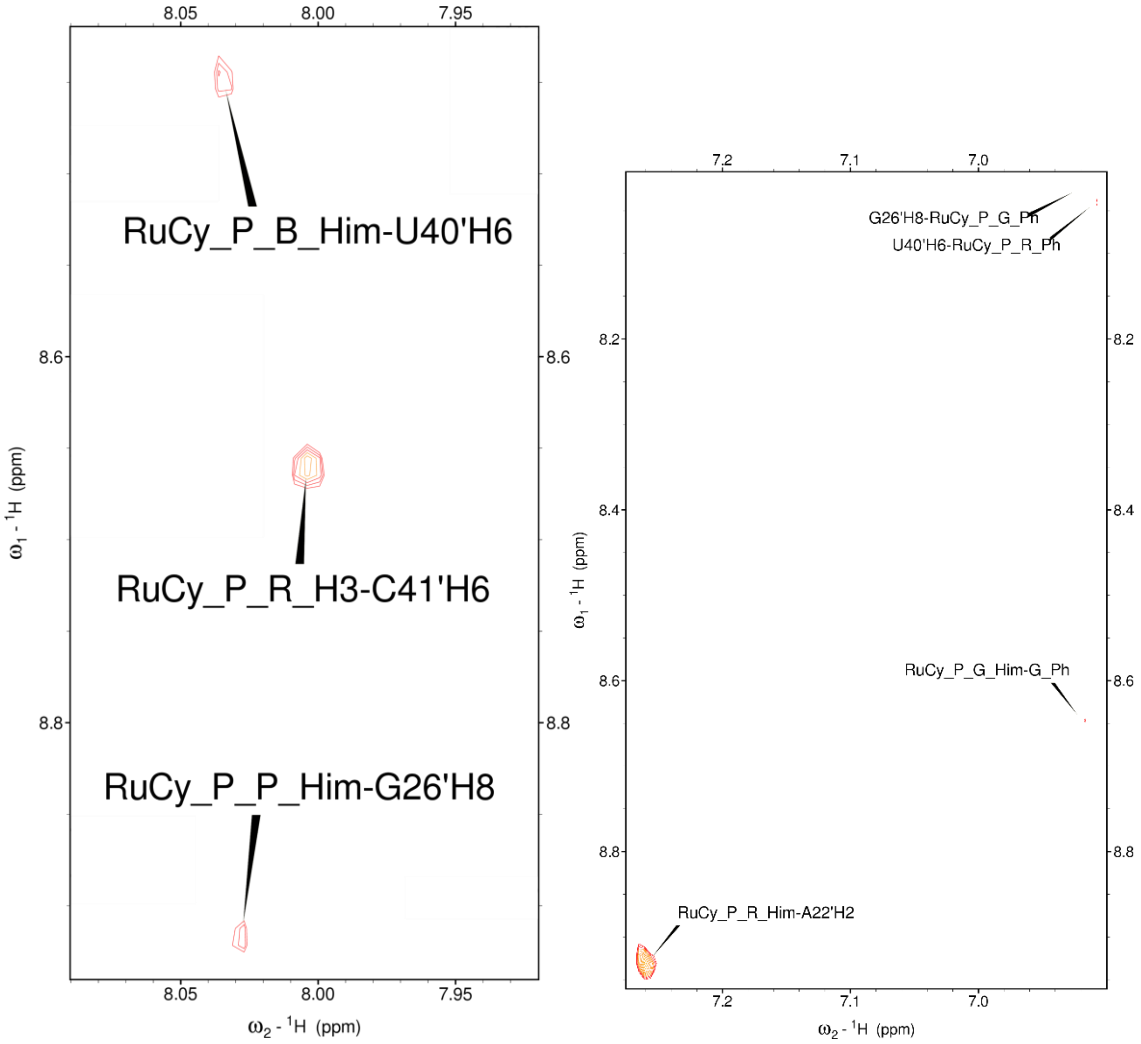


Figure 8.2-6 ^{1}H, ^{1}H-NOESY spectrum of TAR RNA (0.49 mM) with RuCy_P (0.49 mM) in D2O. 50 mM NaCl, pH 7, 278K, 700 MHz. Left. Showing the intermolecular RuCy_P pyridine and imine to RNA peaks. Right. Showing the intermolecular RNA to RuCy_P phenyl and imine peaks, as well as in intramolecular imine to phenyl peak.

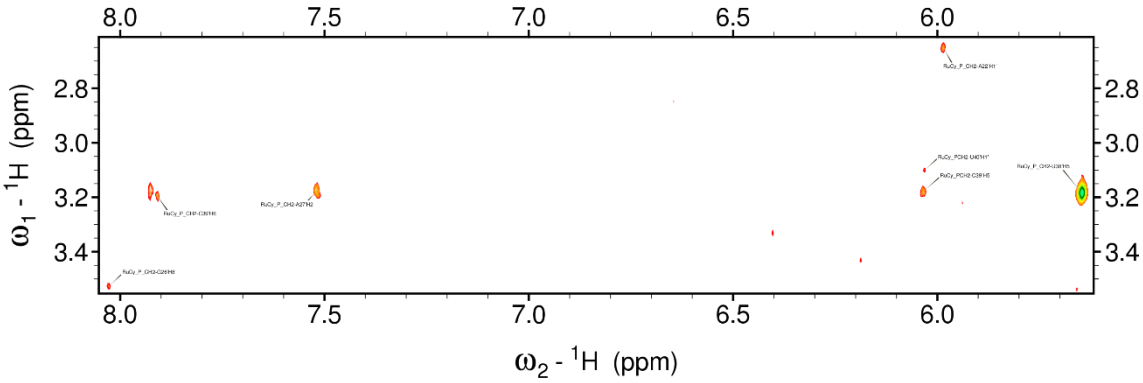


Figure 8.2-7 ^{1}H, ^{1}H-NOESY spectrum of TAR RNA (0.49 mM) with RuCy_P (0.49 mM) in D2O. 50 mM NaCl, pH 7, 278K, 700 MHz. Showing the intermolecular RuCy_P_CH2 to RNA peaks.

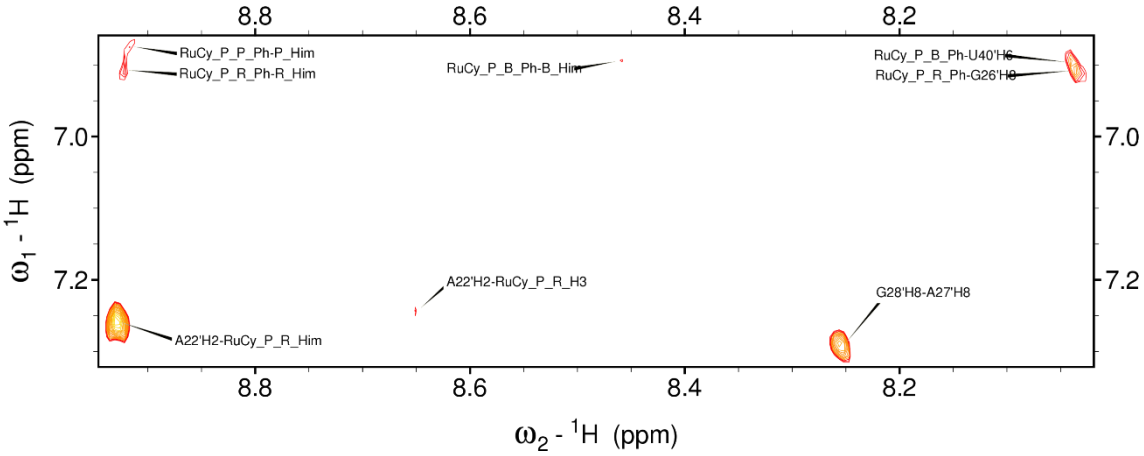


Figure 8.2-8 ^{1}H, ^{1}H-NOESY spectrum of TAR RNA (0.49 mM) with RuCy_P (0.49 mM) in D2O. 50 mM NaCl, pH 7, 278K, 700 MHz. Showing the intramolecular RuCy_P phenyl to imine and RNA inter-base peaks, as well as various intermolecular peaks.

8.3 Bound TAR 1:0.5

8.4 Ruthenium Cylinder

8.4.1 ^{1}H,^{1}H-NOESY

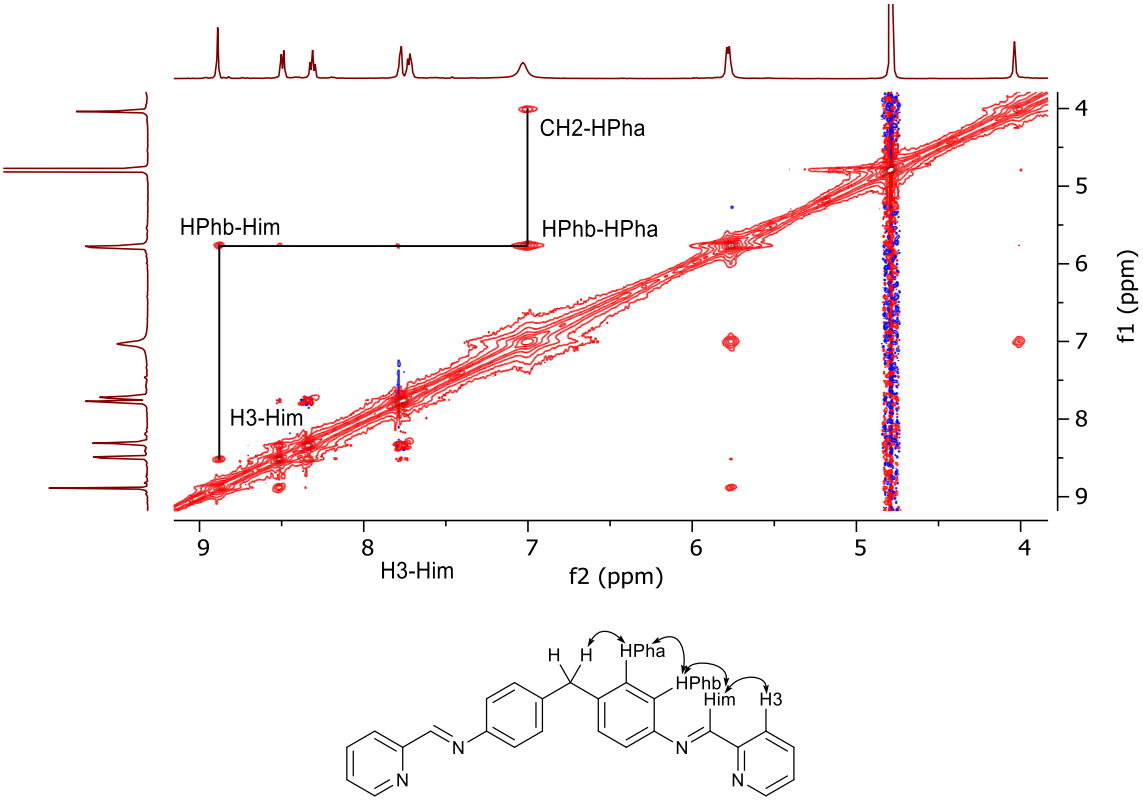


Figure 8.4-1 $^{1}H,^{1}H$-NOESY spectrum at 500 MHz of ruthenium cylinder in D_2O showing the assignment of HPhb and HPha.

8.4.2 $^{1}H,^{13}C$-HSQC/HMBC

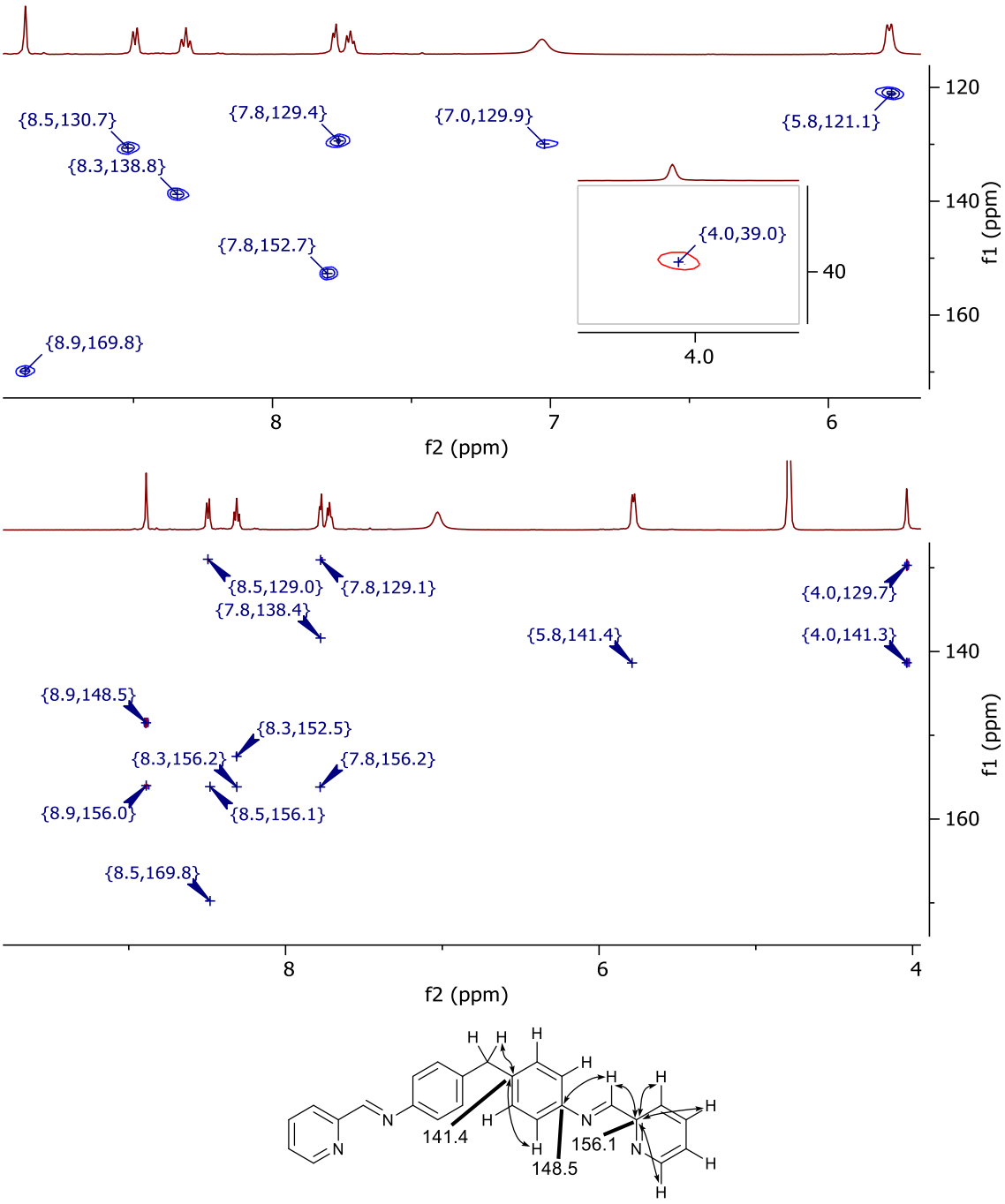


Figure 8.4-2 Ruthenum cylinder ^{13}C assignment in D_2O, 500 MHz, using $^{1}H,^{13}C$-HSQC (top) and $^{1}H,^{13}C$-HMBC (bottom).

8.4.3 UV-Vis Spectra Used for Figure 5.1-4.

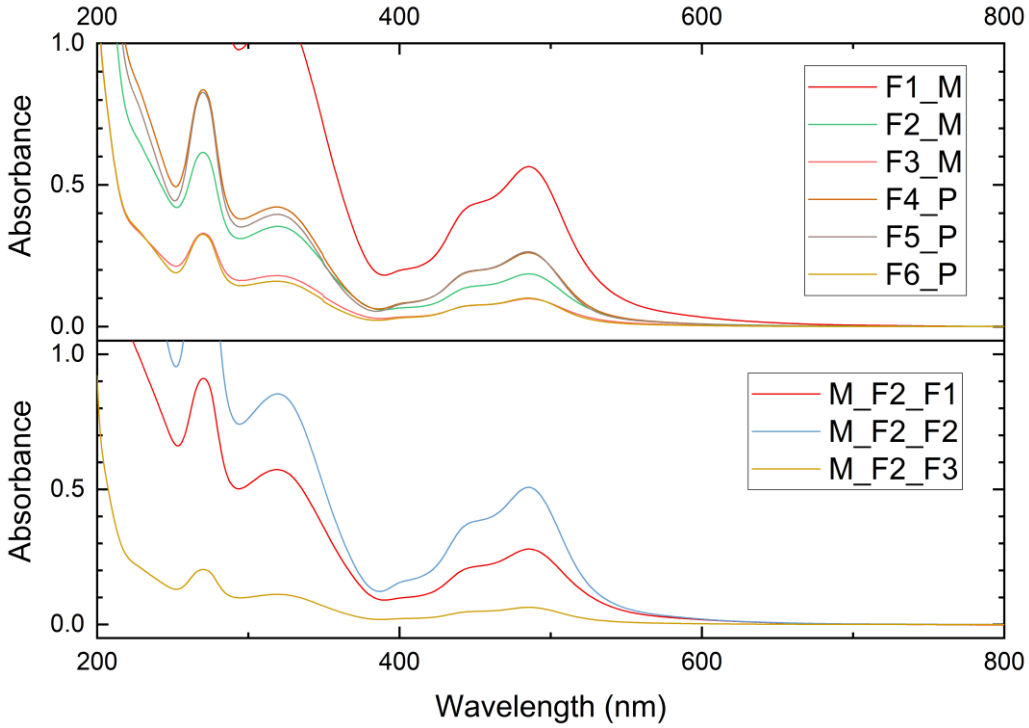


Figure 8.4-3 UV-Vis spectra used to calculate the concentration of enantiomer fractions.

CHAPTER 4 – P ENANTIOMER WITH THREE-WAY JUNCTION DNA

1 Acknowledgements

Thank you to Dr Silke Johannsen and Dr Zenghui Wang for discussion on this project, as well as for running measurements during the SARS-CoV-2 pandemic. Thank you to Prof. Dr. Roland K.O. Sigel for the use of laboratory space and funding. Thank you to Dr. Aditya Garai who like for Chapter 3 was involved in the preparation of RuCy samples and sending them whilst I was in Zurich.

2 Contents

CHAPTER 4 – P ENANTIOMER WITH THREE-WAY JUNCTION DNA 253

1 ACKNOWLEDGEMENTS 253

2 CONTENTS 253

3 GLOSSARY 254

4 INTRODUCTION 254

5 RESULTS AND DISCUSSION 256

5.1 Conclusions 265

6 EXPERIMENTAL 266

6.1 Enantioseparation of Iron Cylinder 266

6.2 Enantiopurity of Fractions 267

6.3 NMR experiments 267

6.4 PAGE experiments 268

7 REFERENCES 268

8 APPENDICES 270

8.1 Circular Dichroism of Iron Cylinder Enantiomers. 270

8.2 ^{1}H,^{1}H-NOESY Spectra of 12mer DNA with RuCy_P 270

3 Glossary

3WJ – Three-Way Junction.

12mer – A 12 nucleotide DNA sequence, $^{5'}$TATGGTACCATA.

6mer – A 6 nucleotide DNA sequence, $^{5'}$CGTACG.

RuCy – Ruthenium cylinder, $[Ru_2L_3]^{4+}$.

FeCy – Iron cylinder, $[Fe_2L_3]^{4+}$.

M – M enantiomer (ΛΛ)

P – P enantiomer (ΔΔ).

4 Introduction

Three-way junctions (3WJ) are found in nature connecting three helix sections for various reasons. These include 5'UTRs,[1] ribozymes,[2] and riboswitches.[3] Replication forks can be considered to have similar Y shaped junction structure to the 3WJ.[4,5]

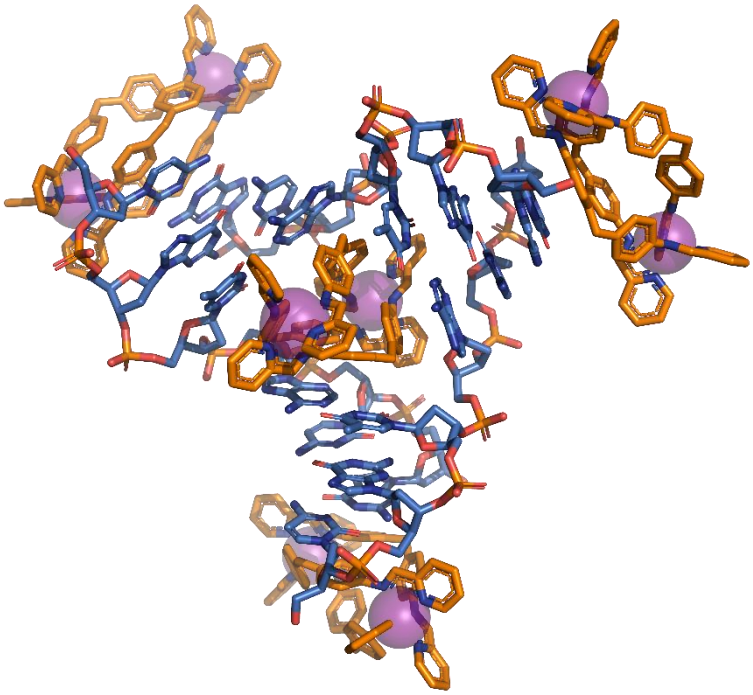


Figure 5.1-1 Crystal structure of FeCy_M (Orange) bound to the center and arms of a 6mer 3WJ (Blue). PDB:2ET0.[6]

Previously one research focus the Hannon group was engaged in, was the investigation of DNA binding properties of supramolecular helicates with duplex DNA for the purpose of triggering DNA coiling.[7,8] Part of this involved investigating the mode of binding of the cylinder to a 10mer palindromic DNA, concluding that a racemic iron helicate (FeCy) is a major groove binder with four helicate populations present in the NMR experiment.[9] In the following years the understanding of short DNA palindrome binding properties expanded. A 6mer DNA sequence was found to form a 3WJ structure present in crystal form,[6] and in solution.[10] The 3WJ does not just form with palindromes, but with three distinct 3WJ forming single stranded sequences, as well as 3WJs with unpaired nucleotides.[4] (Figure 5.1-1) In addition an RNA 3WJ can form for crystallography.[11]

The key publication for this chapter is the NMR of a DNA 3WJ bound to an FeCy M enantiomer (FeCy_M). This was concluded with intermolecular cross-peaks in addition

to comparing the chemical shifts of the DNA sequences, showing that the central bases had greater shifts. The three-fold symmetry of the helicate and loss of two-fold symmetry points to a position at the center of the 3WJ.[10]

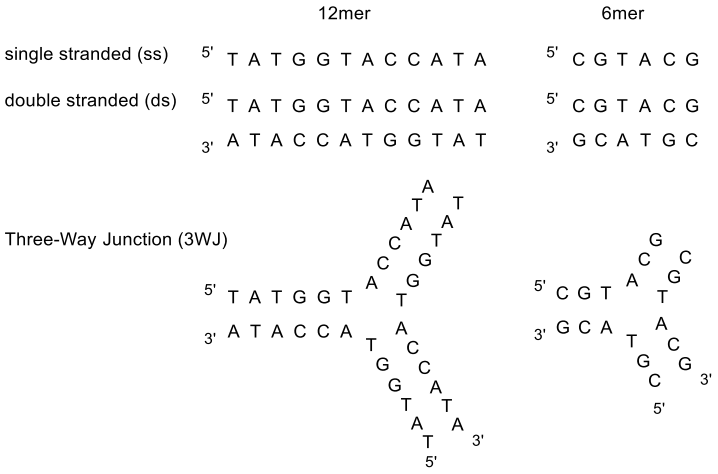


Figure 5.1-2 The DNA structures used within this chapter.

This chapter aims to use this established method to investigate whether the formation and binding of the 3WJ is enantiospecific and whether the results can be repeated using the P enantiomer. The experiments detailed in this chapter use two DNA palindromes, the same used previously.[10] The 12mer (TATGGTACCATA) and the 6mer (CGTACG). They are capable of forming single stranded (ss), double stranded (ds), and 3WJ structures. (Figure 5.1-2)

5 Results and Discussion

The synthesis of FeCy has been exhaustively examined since its initial publication.[12] The enantiomers are separated using cellulose chromatography in the same style as previously reported,[13] using principles from a paper chromatography method.[14] The ruthenium cylinder (RuCy) preparation and enantioseparation is detailed in Chapter 3,

so is not repeated in this chapter. NMR samples are prepared using ratios of cylinder to 3WJ DNA, where 1 3WJ = 1.5 ds DNA = 3 ss DNA.

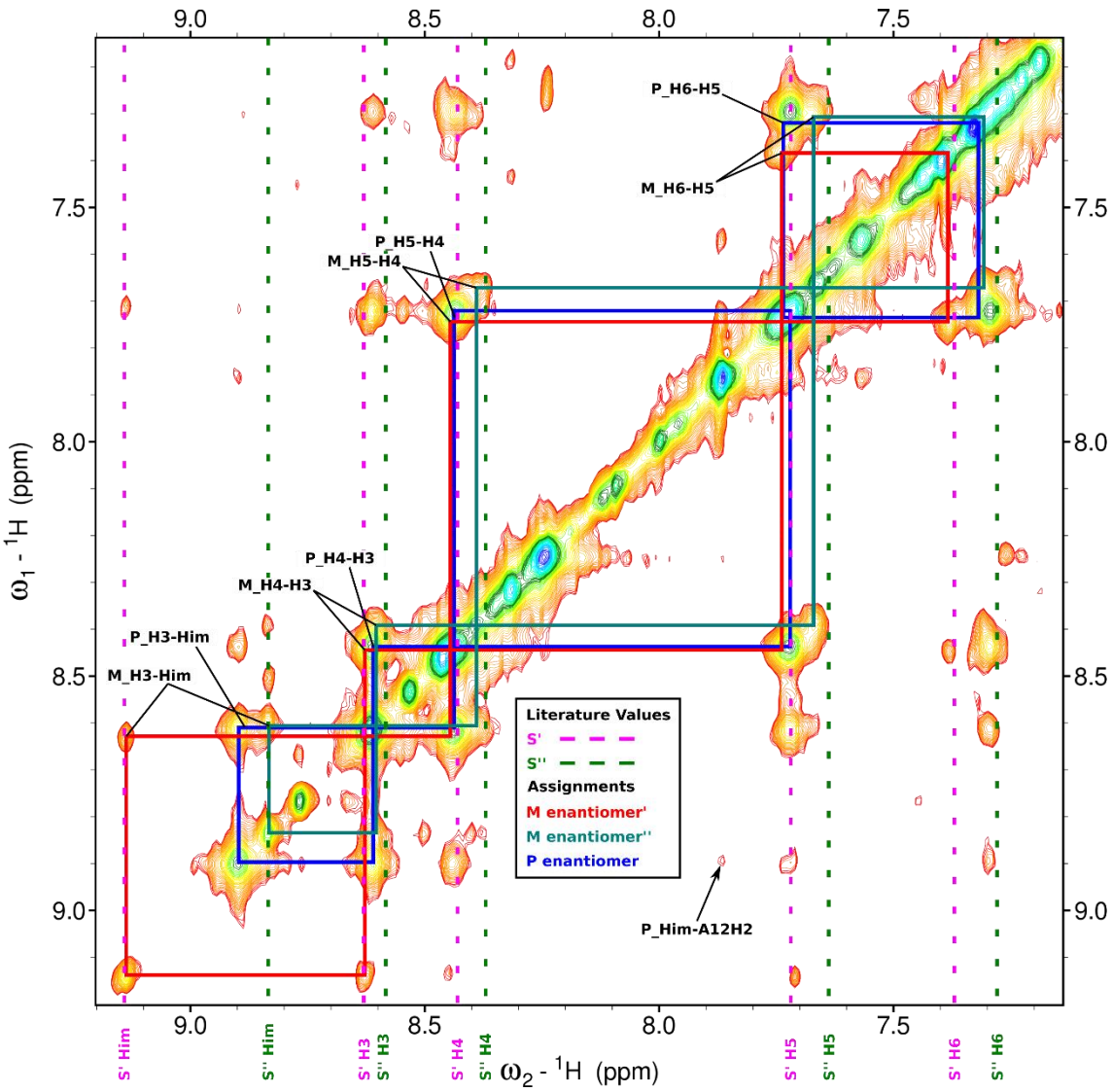


Figure 5.1-1 ^{1}H,^{1}H-NOESY spectrum of 1:1 FeCy_P (0.15 mM) to 12mer 3WJ (0.46 mM ss) in D_2O, 50 mM NaCl, 298K, 600 MHz. Showing the region of aromatic DNA peaks, and cylinder pyridine and imine peaks. Literature values (S', S'') are shown with dashed lines.[10] Assignments are shown with red/teal (M) and blue (P) labels. Literature values adjusted by 0.21 ppm to align S'Him and M_Him.

On mixing a 1:1 FeCy_P to 12mer 3WJ DNA (1:3 ss DNA) a collection of peaks emerges corresponding to the cylinder pyridyl and imine protons. The sample has partially racemised during preparation and acquisitions leading to a smaller population of M

enantiomer. Like with these NMR experiments racemisation of the FeCy has been observed in solution previously,[13] and previously explored at depth in acetonitrile and water.[15] On binding the cylinder has split into two sets of resonances, as noted by L. Cerasino et al.[10] Adjusting the shifts reported previously aligns the S' and S'' with the M enantiomer' and M enantiomer'' values respectively. In their publication they concluded this was due to the cylinder rotating in the center of the 3WJ, with the ends pointing into different sides of the central plane.[10] In this spectrum the major population is the previously unexplored P enantiomer. This is seen forming a single set of peaks, suggesting lower specificity versus the M enantiomer. Finally, a weak NOE contact is seen for FeCy_P_Him-A12H2, pointing to interactions with the ends of the duplex. (Figure 5.1-1)

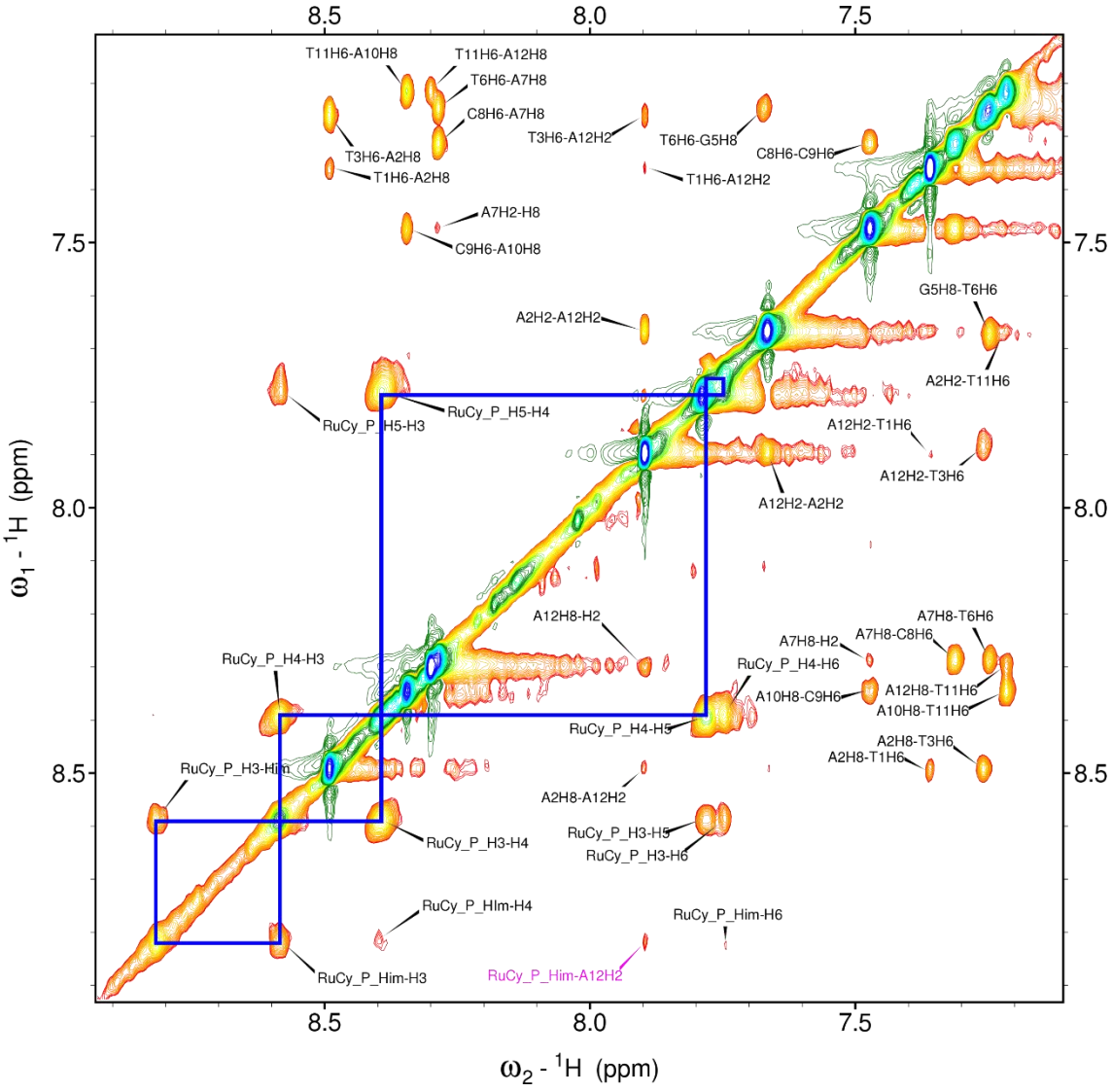


Figure 5.1-2 ^{1}H,^{1}H-NOESY spectrum of 0.5:1 RuCy_P to 12mer 3WJ in D2O, 50 mM NaCl, 298K, d8 = 350 ms, 700 MHz.

The RuCy, unlike the FeCy is less susceptible to racemisation so can be used to provide a spectrum with the isolated P enantiomer. Again, like with the FeCy_P, a 0.5:1 mixture of RuCy_P produces a single set of peaks corresponding to the imine and phenyl protons. There is also a crosspeak for RuCy_P_Him_A12H2 matching the interactions at the end of the duplex seen for the FeCy_P. Together this indicates similarities in the binding of the FeCy_P and RuCy_P showing less dependence on metal center versus helicity. (Figure 5.1-2)

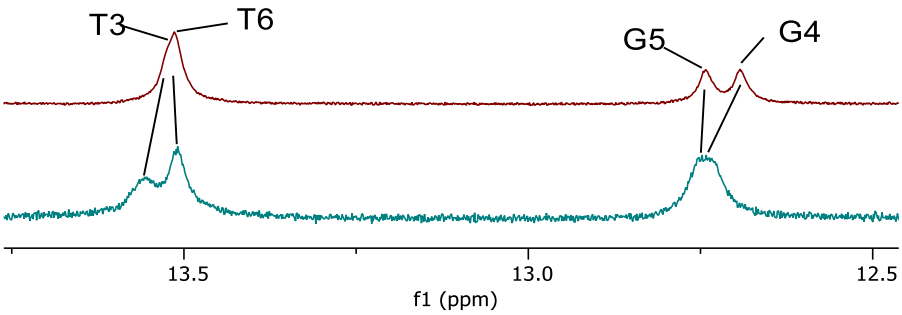


Figure 5.1-3 Stacked ^{1}H-NMR spectrum showing the imino regions. Top. Free 12mer DNA, 600 MHz. Bottom. 0.5:1 RuCy_P to 12mer DNA (3WJ), 700 MHz.

Looking at the imino region of the ^{1}H-NMR spectra at a ratio of 0.5:1 RuCy_P to 3WJ 12mer DNA, it is apparent that little change is made to the base-pairing. Though there is a slight shift for G4 and T3 closer to the ends. Assignments are made using ^{1}H,^{1}H-NOESY. (Figure 8.2-1) Interestingly there does not seem to be simultaneously a bound and an unbound form present in solution, observable through the shift of the G4 peak. This suggests that the binding is either a weak transient form, or a conformation in solution making use of a ratio of 1:6 RuCy_P to ss DNA.

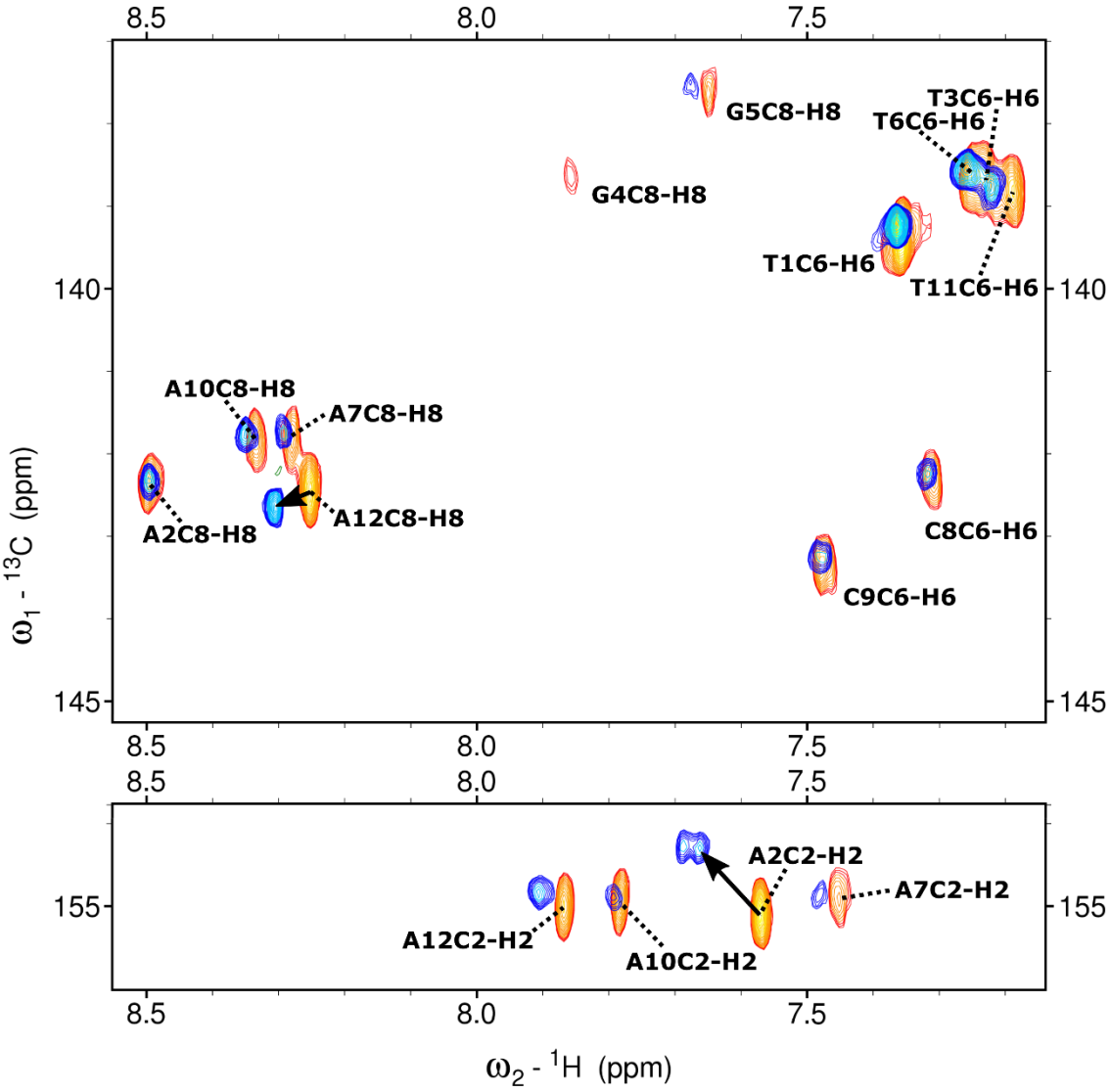


Figure 5.1-4 $^{1}H,^{13}C$-HSQC spectrum overlay at 700 MHz of free (Orange) 12mer DNA (0.47 mM ss) and bound (Blue) 1:0.5 3WJ 12mer DNA (0.42 mM ss) with RuCy_P (0.07 mM). Both in D2O, 50 mM NaCl, 298K, pH 7.00. Showing the C-H2/6/8 peaks. Free peaks are labelled, with dotted lines to help where peaks are close. Bound peaks are not labelled, but overlap with free except where their shift is shown with an arrow. Missing bound G4C8-H8 is circled.

$^{1}H,^{13}C$-HSQC data shows the peaks towards the ends of the duplex experiencing shifts. Using 0.5:1 RuCy_P to 3WJ DNA changes are seen in A12C8-H8, A2C2-H2, T11C6-H6 and A12C2-H2 which is consistent with disruption to the ends of the duplex. Also seen are slight shifts to A7C2-H2 and G5C8-H8 toward the center of the sequence. Most striking is the loss of the G4C8-H8 suggesting mobility at this base. (Figure 5.1-4)

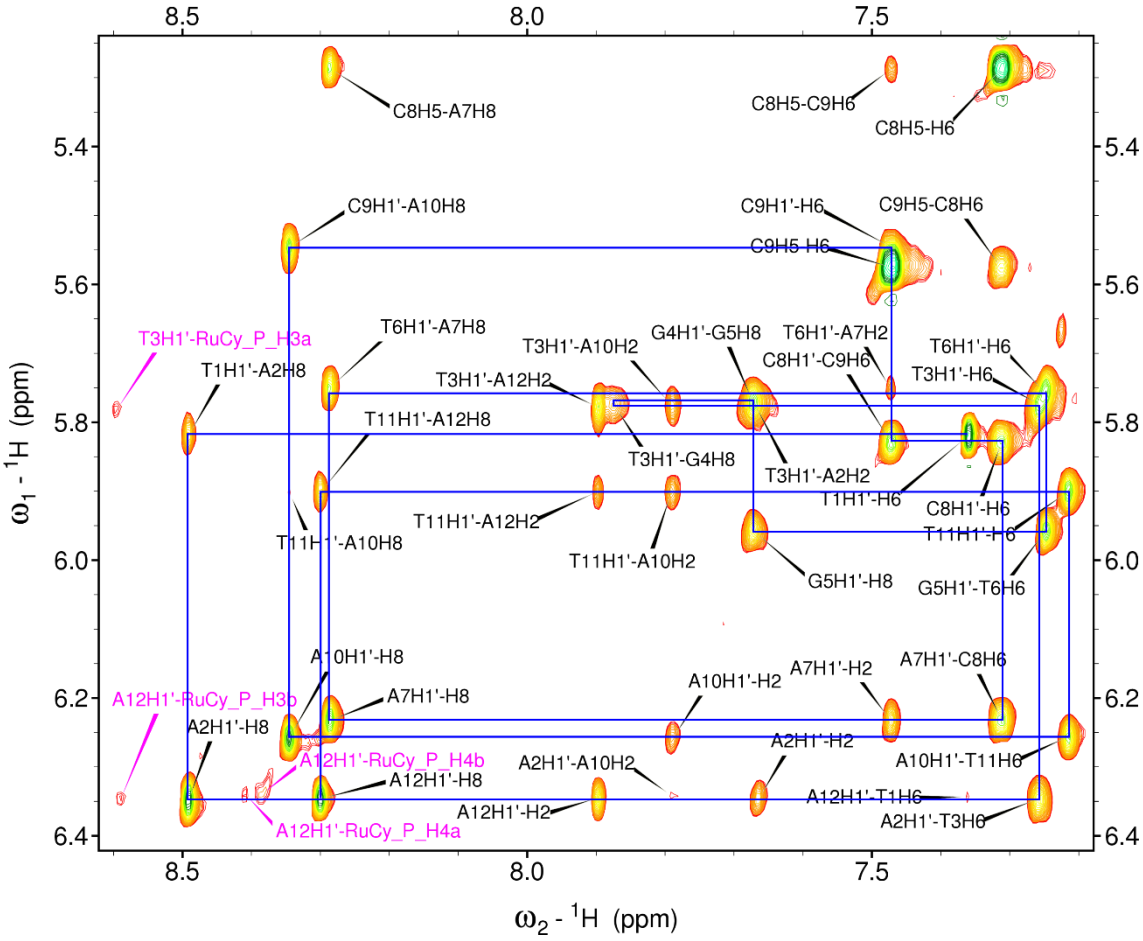


Figure 5.1-5 ^{1}H,^{1}H-NOESY spectrum of 1:0.5 12mer DNA (0.42 mM ss) with RuCy_P (0.07 mM) in D_2O, 50 mM NaCl, 298K, pH 7.00. d8 = 350 ms, 700 MHz. H1'-H8 sequential walk shown with blue. Intermolecular NOEs in purple.

A mixture of 1:0.5 3WJ to RuCy_P produces a ^{1}H,^{1}H-NOESY spectrum with 1 set of DNA peaks consistent with the imino region. Invisible with the intramolecular cylinder NOEs, to the DNA, two populations of intermolecular NOEs are seen. These are labelled a and b though there is no observed relation between H3a/b and H4a/b respectively. As observed in the ^{1}H,^{13}C-HSQC, the G4 appears to be disrupted with weakened NOEs. The neighbouring T3 ribose experiences NOE contacts with T3H1'-RuCy_P_H3a suggesting the cylinder is held deeper within the forked duplex. This would be permitted by fraying up to base-pair T3-A10 which is seen in the imino region. (Figure 5.1-3) Also present are

A12H1' NOEs to RuCy_P_H3b/H4b/H4a. Interestingly T3H1'-A12H2 is present suggesting that perhaps A12 is positioned pointing back toward the minor groove. (Figure 5.1-5)

There are limits to the information that can be gained using a 0.5:1 ratio of RuCy_P to 3WJ, it is possible the 3WJ might form at 1:1. This was attempted by slowly adding RuCy_P in D_2O to a solution of the 12mer/6mer DNA, however the samples prepared repeatedly precipitated and could not be dissolved in a reasonable quantity of solvent. L. Cerasino et al. achieved near 1:1 by adding the cylinder dissolved in methanol-d4 until incipient precipitation. The samples were then lyophilised and dissolved in D_2O. It is possible the experiment conducted here could have achieved a similar result using the same approach with methanol-d4.

Considering the FeCy_M binds to the center of the 3WJ,[10] whereas the FeCy_P and RuCy_P are not seen to do so in NMR. PAGE experiments were then conducted to see whether smaller populations were present.

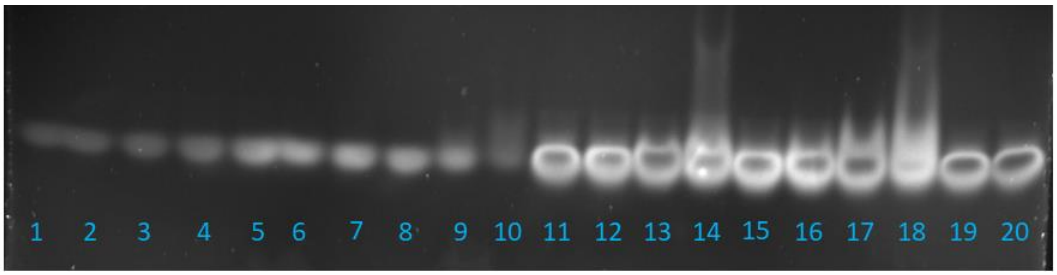


Figure 5.1-6 20% 37.5:1 PAGE image, ran in pH 7.05 TBA buffer, showing the effect of cylinders on 6mer and 12mer DNA. Lanes in order left to right: 1. 6mer+FeCy 1:1, 2. 6mer, 3-6. 6mer+RuCy_P (1:0.5, 1:1, 1:2, 1:4), 7-10. 6mer+RuCy_M (1:0.5, 1:1, 1:2, 1:4), 11-14. 12mer+RuCy_P (1:0.5, 1:1, 1:2, 1:4), 15-18. 12mer+RuCy_M (1:0.5, 1:1, 1:2, 1:4), 19. 12mer, 20. 12mer+FeCy 1:1. All samples are incubated at 37 °C in 89 mM pH 7.05 tris boric acid buffer and 10 mM NaCl for 1 hour. Fluorescence imaged using SYBR Gold stain.

A PAGE experiment of ruthenium cylinder enantiomers with 6mer and 12mer DNA shows no formation of 3WJ at a ratio of up to 1:4 3WJ to cylinder. This is important

because, in agreement with the NMR data, it shows that the RuCy does not induce the three-way junction suggesting the metal center has an impact on the binding. Interestingly beginning at 1:2 and obvious at 1:4, both enantiomers cause some streaking of the 12mer DNA. The M enantiomer appears to be more effective in this, creating a stronger streak in lanes 17 (1:2) and 18 (1:4) for the 12mer, and lanes 9 (1:2) and 10 (1:4) for the 6mer. The P enantiomer does so for the 12mer, seen in lanes 13 (1:2) and 14 (1:4), but appears to have little effect with the 6mer. A streak suggests no definitive structure, which would agree with the NMR analysis that shows poor binding between the RuCy_P and DNA. However, the M enantiomer was expected to form the 3WJ as seen from previous PAGE experiments with FeCy.[16] (Figure 5.1-6) To investigate whether the 3WJ could be formed and visualised using this procedure, the experiment was repeated using racemic FeCy at higher concentrations.

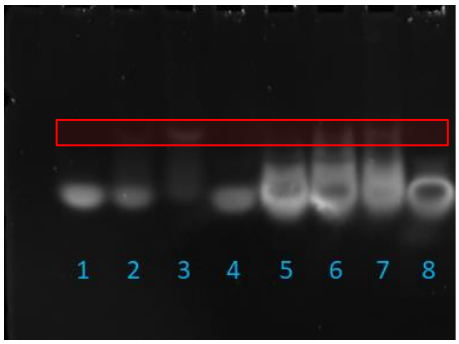


Figure 5.1-7 20% 37.5:1 PAGE image, ran in pH 7.05 TBA buffer, showing the effect of iron cylinder on 6mer and 12mer DNA. Lanes in order left to right: 1-4. 6mer+FeCy (1:2, 1:4, 1:8, 1:0), 4-8. 12mer+FeCy (1:2, 1:4, 1:8, 1:0). All samples are incubated at 37 °C in 89 mM pH 7.05 tris boric acid buffer and 10 mM NaCl for 1 hour. Fluorescence imaged using SYBR Gold stain. 3WJ structure indicated with red rectangle.

Mixing at 1:4 DNA to FeCy, the 3WJ begins to form for both 6mer and 12mer. At 1:8 the strongest 6mer band is the 3WJ. This is visible for the 12mer, though the bands are slightly obscured by streaking. This shows the procedure can form the 3WJ in the same

manner as previously shown using FeCy_M/P/racemic with the same sequence. Interestingly, in this previous work, the FeCy_P formed the 3WJ. Though much weaker than the FeCy_M or racemic, and the bulk of DNA in all cases was either in ss or ds form.[16]

5.1 Conclusions

Whether a repeat using a ratio of 1:1 or even excess RuCy_P would produce NMR visible 3WJ is uncertain. What is apparent is the P enantiomer does not induce the 3WJ as readily as the FeCy_M.[10] For FeCy, the M enantiomer preferentially binds to the center of the 3WJ, with the P enantiomer interacting with the ends of the three arms. This shows the P enantiomer does not compete with the M for binding at the center and likely only interacts transiently with the DNA. Interestingly this arrangement has been seen before in crystal structures where it is referred to as "discontinuous three-way junctions".[17]

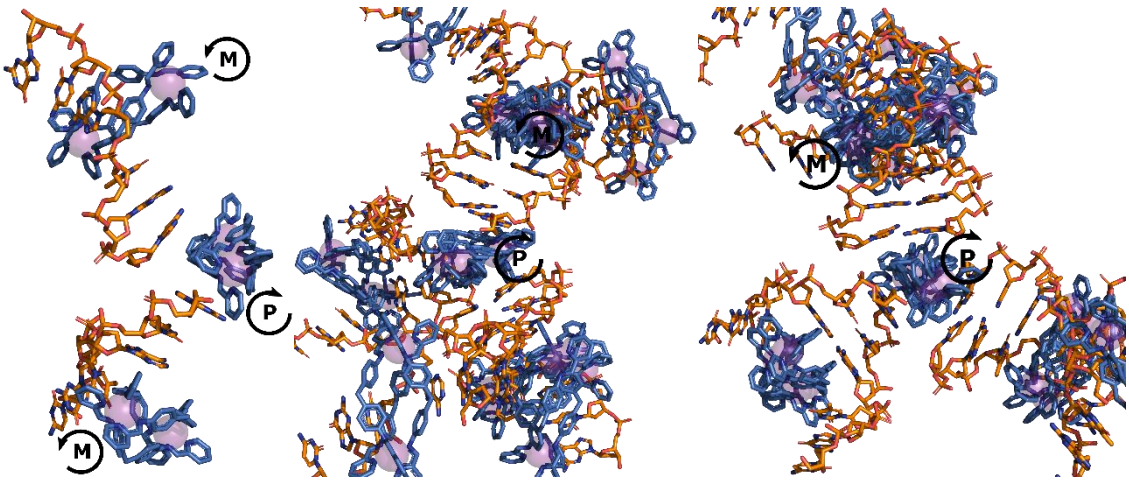


Figure 5.1-1 Crustal structure portions showing the M and P enantiomers arranged around 3WJ forming 6mer DNA. Left. Structure model showing the M and P enantiomers stacking onto the middle and end of the sequences respectively. Center. Selected symmetry mates to show the formation of the 3WJ around the M enantiomer.

Right. Selected symmetry mates to show the formation of discontinuous 3WJ around the P enantiomer. PDB:3I1D.[17]

The RuCy_P with 12mer 3WJ at a ratio of 0.5:1 interacts exclusively with the ends of the duplex DNA in an analogous manner to the FeCy_P, suggesting the failure to bind to a 3WJ could be dependent on helicity. One way to satisfy the formation of one population at a 0.5:1 3WJ ratio is to consider that this corresponds to 1:6 single stranded or 1:3 double stranded implying that the 3WJ hasn't formed. In a crystal structure featuring the M and P enantiomer together. The P enantiomers were bound to the ends of the duplex as a "matrix" holding the 3WJs together. Its highly likely the RuCy_P is transiently binding to the DNA, however its possible these discontinuous 3WJs could be present. The effects of chirality will be important if these complexes or similar helicates are used as drugs, whether they have differing targets or reduced effects.

6 Experimental

Ruthenium cylinder P enantiomer is prepared using the same method as presented in Chapter 3. Racemic iron cylinder is prepared according to previously published procedures.[12]

6.1 Enantioseparation of Iron Cylinder

A slurry of cellulose powder in NaCl solution (0.2 M) is added to a column to generate a cellulose bed (2 cm diameter, 10 cm height). Iron cylinder (0.1 g) is dissolved in minimum NaCl solution (0.2 mM, ca. 2 mL), which is then added to the surface of the cellulose bed to a depth of 0.5 cm. The column is then eluted under pressure with 0.2 NaCl solution. Fractions are taken directly into dry ice chilled vials which are then lyophilised. The dried

fractions are then desalted by triturating the cylinder with minimal methanol, leaving a small portion of undissolved cylinder with the NaCl. The samples are centrifuged (7500 RCF) and the supernatant kept. The desalted complex is then dropped into excess chilled diethyl ether and filtered, then dried under vacuum to yield the desalted fractions.

6.2 Enantiopurity of Fractions

The same cuvette, containing milli-Q water (1 cm), is used to measure a baseline for UV-Vis absorption and circular dichroism. Portions (ca. 5 mg) are dissolved in milli-Q water (1 mL) to create a stock solution, from which aliquots are taken and added into the cuvette until the UV-Vis absorption reads between 0.2 and 0.8, this is used to calculate the concentration using the extinction coefficient in H_2O λ_{max} 584 (16900 $dm^3mol^{-1}cm^{-1}$) nm.[13] The cuvette is then transferred to the CD spectrometer for a measurement between 800-200 nm (3 repeats averaged, 0.3 s acquisition per 1 nm measurement, 25 °C).

6.3 NMR experiments

DNA sequences were sourced from Microsynth. DNA stock solutions prepared above are diluted to a final single stranded concentration between 0.4 and 0.6 mM, at a final volume of 220 µL, with 50 mM NaCl at pH 7. The sample is annealed at 95 °C for 5 minutes, followed by a slow cool to room temperature. For 100% D2O samples, the 90/10 H2O/D2O sample is lyophilised. The sample is then dissolved in 100% D2O of the same total volume as was removed. It is again annealed.

DNA is measured using 5 mm Shigemi tubes on either Brucker AV600 or AV700 with a CP-TCI z-axis pulsed field gradient CryoProbe®. Spectra are initially processed using Bruker TopSpin. Assignments and figures are made using Sparky.

Pulse programs used in this study are:

- ^{1}H-NMR (zgesgp/zgpr)
- ^{1}H,^{1}H-TOCSY 100% D_2O (mlevphpr.2)
- ^{1}H,^{1}H-NOESY 90% H_2O (rs_noesywatg)
- ^{1}H,^{13}C-HSQC 100% D_2O (hsqcetgpsisp2.2)
- ^{1}H,^{1}H-NOESY 100% D_2O (noesyesgpph)

6.4 PAGE experiments

DNA sequences were sourced from Merck. Preparation of the gel samples begins with the addition of pH 7.05 10x buffer (890 mM tris base, 890 mM boric acid and HCl) and 10x NaCl (100 mM). Followed by 12mer (TATGGTACCATA) or 6mer (CGTACG) DNA. Finally, the compound to be assessed is added (5 uM), and the samples are incubated at 37 °C for 1 hour. After adding glycerol to 10%, the samples are added to the polyacrylamide gel (20 %, 37.5:1, pH 7.05 TB buffer) and a potential of 140 V is applied for 2 hours.

7 References

1 J. Ouellet, S. Melcher, A. Iqbal, Y. Ding and D. M. J. Lilley, Structure of the three-way helical junction of the hepatitis C virus IRES element, *Rna*, 2010, **16**, 1597–1609.

2 E. Bonneau, N. Girard, S. Lemieux and P. Legault, The NMR structure of the II-III-VI three-way junction from the Neurospora VS ribozyme reveals a critical tertiary interaction and provides new insights into the global ribozyme structure, *Rna*, 2015, **21**, 1621–1632.

3 R. Micura and C. Höbartner, Fundamental studies of functional nucleic acids: aptamers,

riboswitches, ribozymes and DNAzymes, *Chem. Soc. Rev.*, 2020, **49**, 7331–7353.
4 J. Malina, M. J. Hannon and V. Brabec, Recognition of DNA three-way junctions by metallosupramolecular cylinders: Gel electrophoresis studies, *Chem. - A Eur. J.*, 2007, **13**, 3871–3877.
5 H. Liao, F. Ji, T. Helleday and S. Ying, Mechanisms for stalled replication fork stabilization: new targets for synthetic lethality strategies in cancer treatments, *EMBO Rep.*, 2018, **19**, e2102017.
6 A. Oleksi, A. G. Blanco, R. Boer, I. Usón, J. Aymamí, A. Rodger, M. J. Hannon and M. Coll, Molecular Recognition of a Three-Way DNA Junction by a Metallosupramolecular Helicate, *Angew. Chemie Int. Ed.*, 2006, **45**, 1227–1231.
7 I. Meistermann, V. Moreno, M. J. Prieto, E. Moldrheim, E. Sletten, S. Khalid, P. Mark Rodger, J. C. Peberdy, C. J. Isaac, A. Rodger, M. J. Hannon, P. M. Rodger, J. C. Peberdy, C. J. Isaac, A. Rodger and M. J. Hannon, Intramolecular DNA coiling mediated by metallo-supramolecular cylinders: Differential binding of P and M helical enantiomers, *Proc. Natl. Acad. Sci.*, 2002, **99**, 5069–5074.
8 M. J. Hannon, V. Moreno, M. J. Prieto, E. Moldrheim, E. Sletten, I. Meistermann, C. J. Isaac, K. J. Sanders and A. Rodger, Intramolecular DNA Coiling Mediated by a Metallo-Supramolecular Cylinder, *Angew. Chemie Int. Ed.*, 2001, **40**, 879–884.
9 E. Moldrheim, M. J. Hannon, I. Meistermann, A. Rodger and E. Sletten, Interaction between a DNA oligonucleotide and a dinuclear iron(II) supramolecular cylinder; an NMR and molecular dynamics study, *JBIC J. Biol. Inorg. Chem.*, 2002, **7**, 770–780.
10 L. Cerasino, M. J. Hannon and E. Sletten, DNA Three-Way Junction with a Dinuclear Iron(II) Supramolecular Helicate at the Center: A NMR Structural Study, *Inorg. Chem.*, 2007, **46**, 6245–6251.
11 S. Phongtongpasuk, S. Paulus, J. Schnabl, R. K. O. Sigel, B. Spingler, M. J. Hannon and E. Freisinger, Binding of a Designed Anti-Cancer Drug to the Central Cavity of an RNA Three-Way Junction, *Angew. Chemie Int. Ed.*, 2013, **52**, 11513–11516.
12 M. J. Hannon, C. L. Painting, A. Jackson, J. Hamblin and W. Errington, An inexpensive approach to supramolecular architecture, *Chem. Commun.*, 1997, **128**, 1807–1808.
13 P. J. Cail, PhD Thesis, University of Birmingham, 2017.
14 M. J. Hannon, I. Meistermann, C. J. Isaac, C. Blomme, A. Rodger and J. R. Aldrich-Wright, Paper: a cheap yet effective chiral stationary phase for chromatographic resolution of metallo-supramolecular helicates, *Chem. Commun.*, 2001, 1078–1079.
15 L. Cardo, PhD Thesis, University of Birmingham, 2009.
16 J. Malina, M. J. Hannon and V. Brabec, Recognition of DNA three-way junctions by metallosupramolecular cylinders: Gel electrophoresis studies, *Chem. - A Eur. J.*, 2007, **13**, 3871–3877.
17 D. R. Boer, J. M. C. A. Kerckhoffs, Y. Parajo, M. Pascu, I. Usón, P. Lincoln, M. J. Hannon and M. Coll, Self-Assembly of Functionalizable Two-Component 3D DNA Arrays through the Induced Formation of DNA Three-Way-Junction Branch Points by Supramolecular Cylinders, *Angew. Chemie Int. Ed.*, 2010, **49**, 2336–2339.

8 Appendices

8.1 Circular Dichroism of Iron Cylinder Enantiomers.

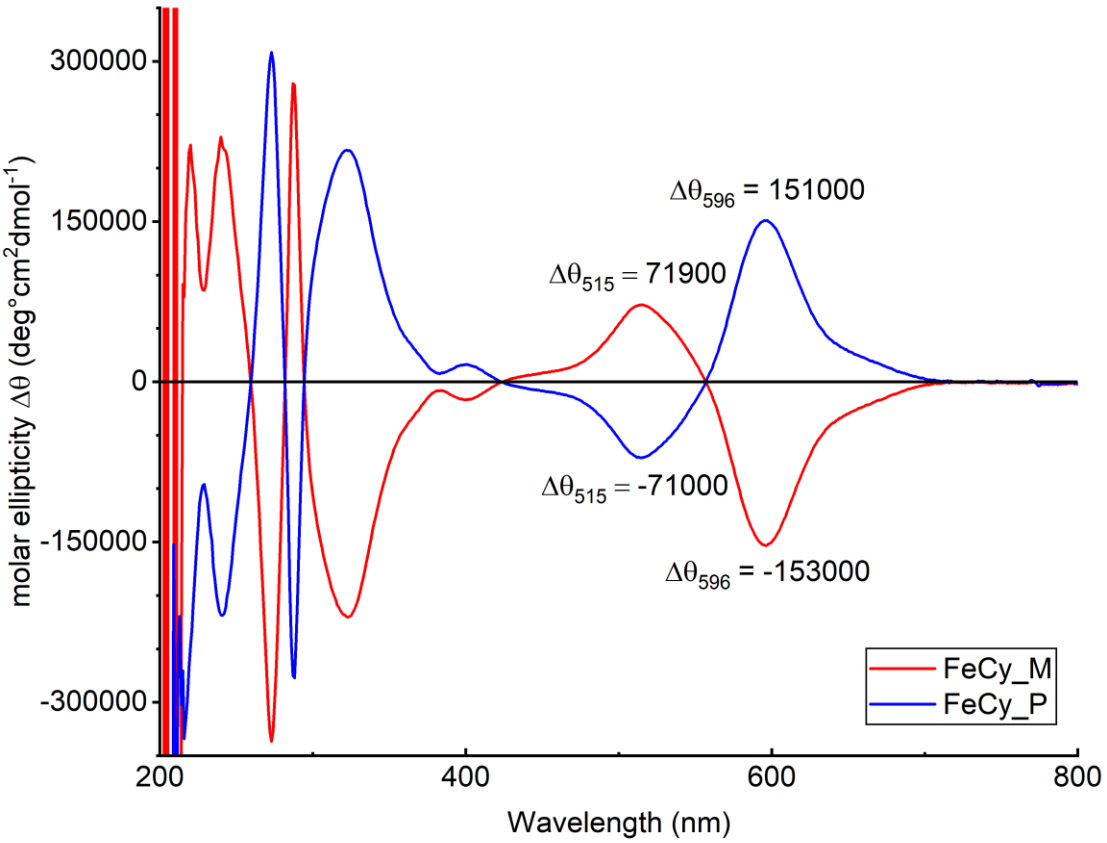


Figure 8.1-1 Circular dichroism spectra of iron cylinder M (red) and P (blue) enantiomers in water. Data adjusted from elipticity (mdeg) using molar extinction coefficient for iron cylinder.[13] Peaks at λmax labelled.

8.2 ^{1}H,^{1}H-NOESY Spectra of 12mer DNA with RuCy_P

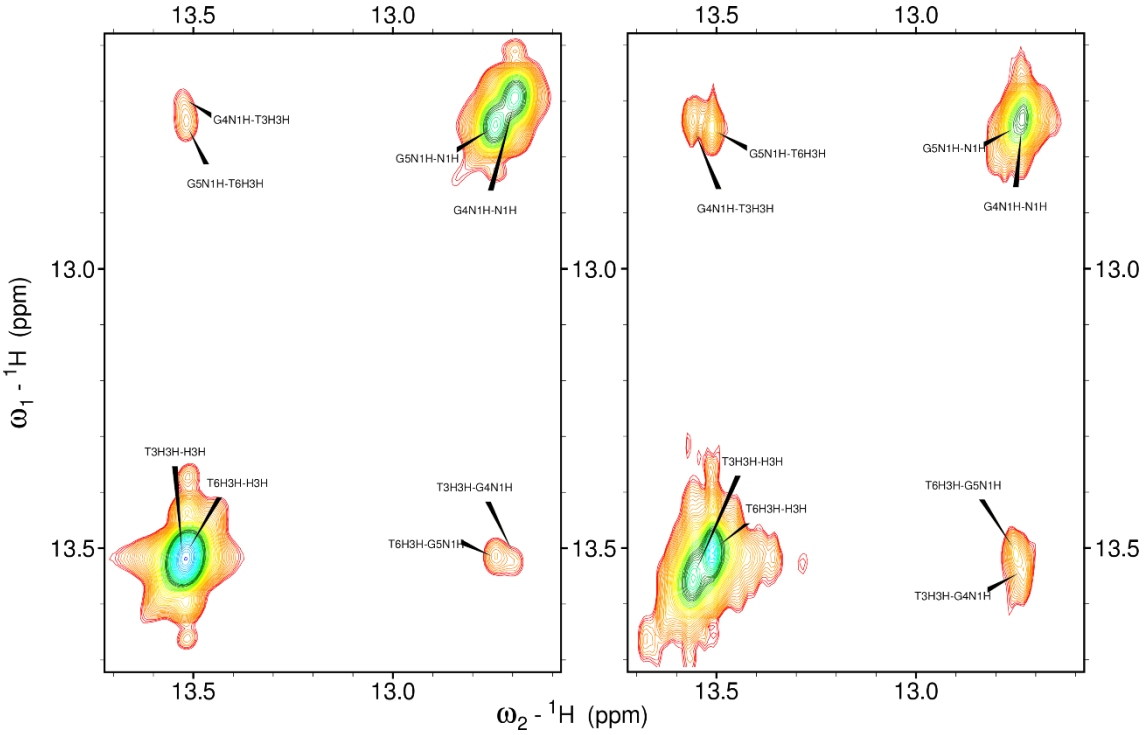


Figure 8.2-1 ^{1}H,^{1}H-NOESY spectra in 90% H_2O showing the imino regions. Left. Free 12mer DNA. Right. 0.5:1 RuCy_P to 3WJ 12mer DNA. 700 MHz.

CHAPTER 5 – CONCLUSIONS AND FUTURE WORK

1 Conclusions

In Chapter 2, a method to control the binding of a 2-imidazole based supramolecular helicate was investigated. This was done through its rotaxanation with CB[10] using thiol cleavable DNB groups to cap. An important part of this chapter has been the development of a reliable method to cap the pseudorotaxane post-assembly. Previously when using flexible methyl pyridine caps slipping the ready-made thread inside was more suitable. The DNB is particularly suited to the post-assembly approach as the labile 1-fluoro position rapidly alkylates onto the imidazole. Once rotaxanated it is unable to form the DNA 3WJ. The rotaxane can then be decapped in a controlled manner with a rate dependence on thiol concentration. On decapping, the rotaxane released a 3WJ forming helicate. The main importance of this is the beginning of research into removable rotaxane caps for supramolecular helicates. DNB has its limits as a capping species, whilst glutathione is found at higher concentrations in certain cancer cells, there are other thiols present in cells in the form of cysteine.

A ruthenium triple helicate presented in Chapter 3, specifically the P enantiomer has been shown to bind to the HIV-1 TAR RNA hairpin and change the conformation. There are two potential models for how the cylinder is bound at the bulge area, one is minor

groove binding, the other is insertion within the bulge. Part of the minor groove binding model is closing the bulge in a manner reminiscent of a coaxially aligned formation, and reducing the interhelical angle such as seen with other stabilised conformations including tat binding. Interestingly the conformational capture of this coaxially aligned structure has been proposed as a mechanism for the binding of the tat protein. If this is the case, then considering previous work demonstrating the inhibition of tat binding using helicates, the cylinder acts not by capturing an inactive conformer of the TAR hairpin. But works instead by competing for the binding of the tat protein. It's possible that normally the correct conformer can be captured by tat, but then the binding requires further alterations to the structure induced by the initial binding. This could then be inhibited by the cylinder, not by preventing the initial tat binding to the major groove, but by engaging the bulge bases to an extent where the tat protein is not binding appropriately. Effectively an indirect competition. For the bulge insertion model, the bulge is held open with a greater interhelical angle. In this conformation the TAR bulge bases would be engaged with the cylinder, preventing their use for tat binding. The bulge unable to close would be incapable of forming the coaxially aligned TAR that is common in bound forms. To understand the mechanism of cylinder TAR inhibition further, it would be valuable to identify the structure at the bulge. Regardless of binding mode this has promising potential as an anti-HIV drug as tat-TAR binding is critical for the assembly of the transcription complex.

Chapter 4 focuses on the binding of ruthenium and iron helicate P enantiomers to DNA 3WJs. The P enantiomer prefers to interact with the ends of the junction arms, on its own, and in the presence of the M enantiomer. Interesting is its similarity to the TAR

bulge binding, where it takes advantage of more flexible unpaired bases. Perhaps this information can be used to direct the binding of the M enantiomer toward 3WJ structures.

2 Future Work

For Chapter 2, a valuable next direction to take this would be applying alternative capping agents to the 2-imidazole pseudorotaxane. Enzyme triggered response could be promising as many of the potential linking groups could aid solubility.[1] Perhaps a peptide Gly-Phe-Leu-Gly or Val-Cit, cleavable by cathepsin B could be used as a capping group to target a variety of cancers.[2] Though whether Val-Cit would need an additional bulky group to properly cap the rotaxane, which could inhibit the cathepsin B, would need to be determined. Considering how this translates to the work in Chapter 3, it's hard to envisage a triggered release using specifically HIV-1 encoded enzymes. An alternative approach might be to attach a slow release group with a glycoprotein targeting antibody, creating areas of higher cylinder concentrations around the HIV particles whilst inhibiting their binding.[3]

To expand on the research in Chapter 3, it would be valuable to identify the structure of the TAR that is adopted on cylinder binding. This could be done with NMR restraint directed structural calculations, or the mobility of the stems could be ascertained with RDCs.

A three base purine GAA bulge occurs in the 5'UTR of Kedougou Virus (KEDV) a mosquito borne flavivirus.[4] This could prove an interesting target for the cylinders presented in Chapter 3, though there is evidence of antibodies in the human population there has

been no reported disease in humans.[5] However this presents an interesting comparison to the pyrimidine UCU HIV-1 TAR bulge.

To expand further on the research presented in Chapter 4, it might be valuable to explore the differential preference for various nucleic acid structures dependent on cylinder enantiomer. For these experiments it would be better to use a different technique to NMR, for the nickel(II) cylinder would be a good alternative to the time-consuming ruthenium cylinder synthesis. Using this the M enantiomer could be fully investigated. With potentially enantiomeric direction toward various targets. In this thread, further research is needed in the area of ruthenium cylinder enantiomers, to identify the cause of the discrepancy in apparent enantiopurity. This would enable reliable continuation of research into the RuCy_M enantiomer.

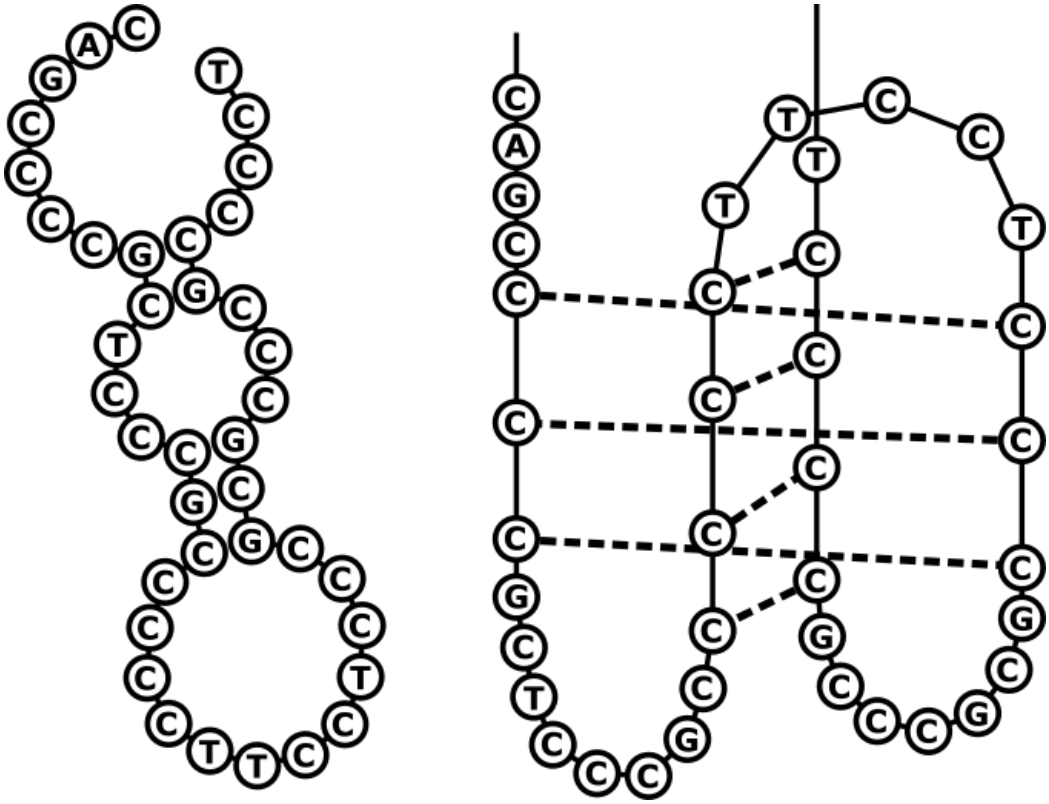


Figure 2-1 Secondary structure of the BCL2 i-motif forming sequence, showing the hairpin form on the left and the i-motif on the right.

A BCL2 promoter i-motif is the active form of the sequence, where their hairpin conformation represses transcription.[6] Stabilisation of the hairpin through loop or bulge binding could be achieved using the helicates presented in Chapters 3 and 4. This should be investigable through CD titrations. (Figure 2-1)

3 References

1 M. Gulfam, F. F. Sahle and T. L. Lowe, Design strategies for chemical-stimuli-responsive programmable nanotherapeutics, *Drug Discov. Today*, 2019, **24**, 129–147.

2 M. Shahriari, M. Zahiri, K. Abnous, S. M. Taghdisi, M. Ramezani and M. Alibolandi, Enzyme responsive drug delivery systems in cancer treatment, *J. Control. Release*, 2019, **308**, 172–189.

3 L. N. Ramana, A. R. Anand, S. Sethuraman and U. M. Krishnan, Targeting strategies for delivery of anti-HIV drugs, *J. Control. Release*, 2014, **192**, 271–283.

4 Z. Y. Liu, X. F. Li, T. Jiang, Y. Q. Deng, Q. Ye, H. Zhao, J. Y. Yu and C. F. Qin, Viral RNA switch mediates the dynamic control of flavivirus replicase recruitment by

genome cyclization, *Elife*, 2016, **5**, 1–27.

5 P. Jansen van Vuren, R. Parry, A. A. Khromykh and J. T. Paweska, A 1958 Isolate of Kedougou Virus (KEDV) from Ndumu, South Africa, Expands the Geographic and Temporal Range of KEDV in Africa., *Viruses*, 2021, **13**, 1368.

6 S. Kendrick, H.-J. Kang, M. P. Alam, M. M. Madathil, P. Agrawal, V. Gokhale, D. Yang, S. M. Hecht and L. H. Hurley, The dynamic character of the BCL2 promoter i-motif provides a mechanism for modulation of gene expression by compounds that bind selectively to the alternative DNA hairpin structure., *J. Am. Chem. Soc.*, 2014, **136**, 4161–71.