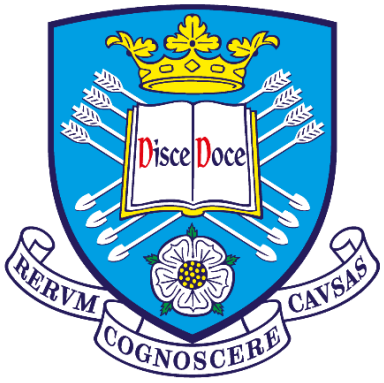


The University Of Sheffield.

A Structural Study into the Binding of a Chiral Dinuclear Ruthenium (II) Polypyridyl Complex to a B-DNA Oligonucleotide

Simon David Fairbanks

Thesis Submitted for the Degree of Doctor of Philosophy

The University of Sheffield

September 2018

Abstract

Herein the synthesis and resolution of the stereoisomers of the dinuclear DNA light switch compound $[\{(bipy)_2Ru\}_2(tpphz)]^{4+}$ (tpphz = tetrapyrido[3,2-a:2',3'-c:3",2"-h:2"',3"'-j]phenazine) is reported. The complexes were resolved by a method of ion exchange chromatography which has been optimised to effectively separate all three stereoisomers. The resolved compounds have been characterised by X- ray crystallography with the absolute structure of each isomer reported. The interaction of these compounds with a DNA octanucleotide d(GCATATCG).(CGATATGC) has been studied using dynamic and structural NMR techniques. Selective deuteration of the bipyridyl ligands of these compounds provided sufficient spectral resolution to characterise the DNA bound structures.

NMR-based models for these complexes have been produced using a restrained molecular dynamics approach through experimentally observed nuclear Overhauser enhancements. A threading intercalation binding mode is proposed with slow and chirality-dependent rates. The Λ,Λ and Λ,Δ stereoisomers show a strong preference for a threaded conformation with inspection of the molecular models revealing the best binding fit for a Λ Ru^{II} centre. Interestingly, the Δ,Δ isomer displays a dynamic exchange process on the NMR timescale translating between threaded and minor groove bound forms. This phenomenon was studied by EXSY NMR and the activation energy of threading calculated to be 116.7 kJ mol^{-1}.

This represents the first solution structure of a threading intercalated bis-ruthenium ligand, outlining that binding selectivity is heavily dependent on the chirality of the intercalating unit.

Acknowledgments

I would like to thank my supervisors Prof. Mike Williamson and Prof. Jim Thomas for their enthusiasm and interest maintained throughout my research. They have both given me invaluable insight and the opportunity to go on many conferences, including the glamourous Pacifichem in which the lovely Hawaiian beaches and many mai tais were one of the main highlights of my PhD.

I would also like to thank all of the people who helped me throughout my research; Andrea for her help with working the NMR spectrometers, Simon down in mass spec, Nick and Pete in stores, Denise and Louise down in accounts and Craig for solving my X ray crystal structures. Also a thank you goes to Prof. Richard Keene for his help with setting up the chiral column chromatography.

As I have been working across two groups I would like to thank all the members past and present of the NMR group in MBB, and the Thomas group in chemistry and to name every person briefly would not do them justice. Thanks for putting up with whatever puns have been thrown at you and realising that conversation may or may not be a joke (depending on my mood). It has been a blast.

To name but a few, I would like to thank the chemistry office mainstays. Sree for his assumed cricket knowledge to which I will always nod, and of course his famous back rubs; Stuart for the dark humour and big moustache; Sam for his ever magnificent hair and blunt calling out of my nonsense; Smasha for the bants and being a founding member of SKS; Caz for the "amazing column fun times" shared; Tom for being half man, half bike and half sharp; Kristy "moth the ripper" for inventing the adjective to be Kirstied, another key member of the SKS; Charlotte for the levels of eccentric energy and optimism - keep up the bee keeping honey; Goaty the Christmas goat; Shahryar for his fantastic use of columns along with his ever impressive memories of birthdays; Paul Jar-nan for his love of the pun; Hawaii teawa Hiwa for just being a lovely guy, and Mike for his desk.

I would like to say a massive thank you to Fliss for brightening up the office and being so amazing and fantastic. I wouldn't have been able to get through this without you.

Declaration

Except where specific references have been made to other sources, the work in this thesis is the original work of the author. It has not been submitted, in whole or in part for any other degree.

Simon D. Fairbanks

September 2018

Abbreviations

A	Adenine
Bipy	2,2'-bipyridine
C	Cytosine
CD	Circular Dichroism
CNS	Crystallography and NMR System
COSY	Correlation Spectroscopy
CT	Calf Thymus
d	Doublet
dd	Doublet of Doublets
DCM	Dichloromethane
DNA	Deoxyribonucleic Acid
DPPZ	Dipyrido[3,2-a:2',3'-c]phenazine
DPQ	1,10-phenanthroline-5,6-dione
DQF	Double Quantum Filtered
ESI	Electrospray Ionisation
G	Guanine
HPLC	High Performance Liquid Chromatography
HTS	Human Telomere Sequence
LD	Linear Dichroism
m	Multiplet
Me	Methyl
MD	Molecular Dynamics
MLCT	Metal-to-ligand Charge Transfer
MTF	Molecular Topology File

NDB	Nucleic Acids Database
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Enhancement
NOESY	Nuclear Overhauser Effect Spectroscopy
ORTEP	Oakridge Thermal Ellipsoid Plot
PAR	Parameter File
PBS	Phosphate Buffered Saline
PDB	Protein Data Bank
Phen	Phenanthroline
ppm	Parts Per Million
q	Quartet
RNA	Ribonucleic Acid
RPMI	Cell Media
s	Singlet
T	Thymine
t	Triplet
TOCSY	Total Correlation Spectroscopy
TPPHZ	Tetrapyrido[3,2-a:2',3'-c:3'',2''-h:2''',3'''-j]phenazine
TSP	Trimethylsilyl propanoic acid (d_4)
UV	Ultraviolet
vdw	Van der Waals

Table of Contents

Chapter 1. Introduction

1.1. The Molecule of Life

The macromolecule deoxyribonucleic acid (DNA) is the primary chemical repository for genetic information inside the cell. It is fundamental to life as it provides the code, in the form of genes, to synthesise proteins through an RNA mediated process of transcription and translation.[1] This pathway forms the central dogma of biology shown in figure 1.1. The information encoded within genes is hereditary and is conserved throughout replication. Thus, the study of DNA is crucial for understanding the basis of cellular processes at the molecular level, but also for developing diagnostics and treatment of genetic diseases. Molecules that bind to DNA can be employed either as biomarkers or to modulate gene function as therapeutics.[2]

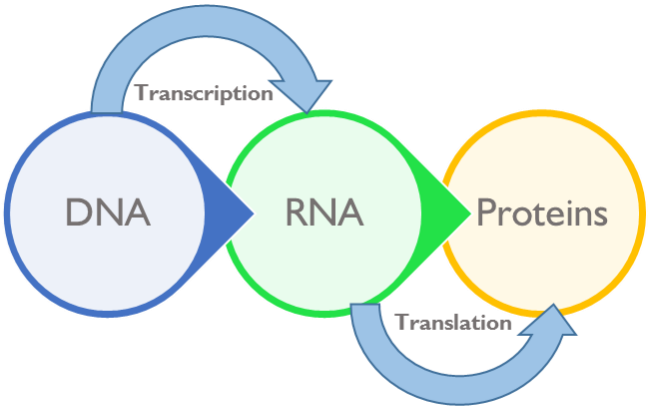


Figure 1.1 The central dogma of biology.

1.2 The Structure of DNA

A strand of DNA can be defined as a polynucleotide chain composed of four types of nucleotide subunit. Each nucleotide consists of deoxyribose sugar, a negatively charged phosphate group and a heteroaromatic nitrogenous base. These nucleotides polymerise to form polynucleotides which are more commonly

known as nucleic acids. The deoxyribose sugars of neighbouring nucleotides are connected by phosphodiester bonds from the 5' position on one nucleotide to the 3' on the next. This phosphodiester bond forms from a condensation reaction between the 5' phosphate group and the 3' hydroxyl group. It is this bond formation that ultimately results in the creation of an alternating sugar-phosphate polyanionic backbone (figure 1.2). The bases are linked to the sugar by a glycosidic bond to the 1' carbon and fall in to two types; the purine structures adenine (A) and guanine (G), and the pyrimidine structures cytosine (C) and thymine (T).[3]

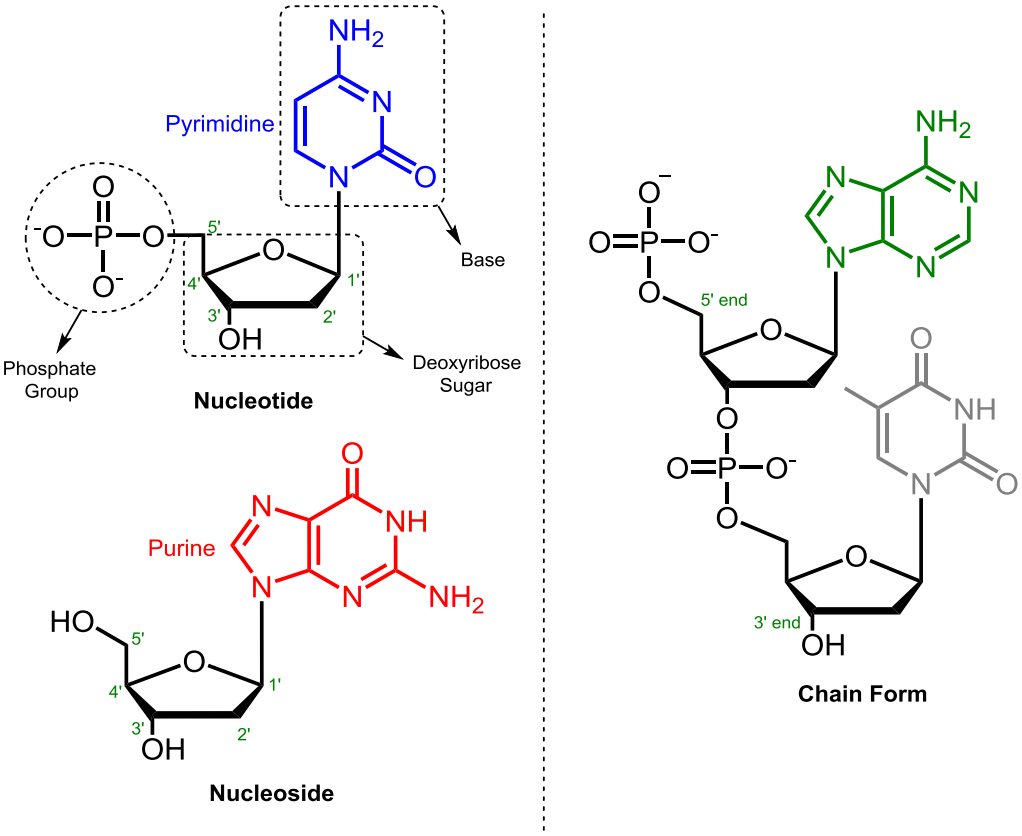


Figure 1.2. Left: The structure of a DNA nucleotide and DNA nucleoside. Right: Nucleotides in chain form linked by phosphodiester bonds from the 5' end to 3' end. The bases are indicated by colours: Cytosine (Blue); Guanine (Red); Adenine (Green) and Thymine (Grey).

In early 1953, Watson and Crick, with the help of X-ray crystallography data from Wilkins and Franklin, published the three dimensional, double helical structure of B-DNA.[4] This discovery was a major breakthrough for molecular biology and for which Watson, Crick and Wilkins earned a share in the Nobel Prize. They disregarded Pauling and Corey's previous scientific model of three intertwined DNA chains with phosphates located at the fibre axis and bases

towards the exterior. It was not evident how the chains were held together and it seemed that the negatively charged phosphates would repel each other.[5] Watson and Crick instead suggested that the bases are internally located and the sugar-phosphate backbone forms the outside of the assembly. This proposed structure consists of a right handed double helix formed from two antiparallel DNA strands to allow the formation of base pairs (figure 1.3). The strands are held together by strong hydrogen bonding interactions between purine and pyrimidine bases giving the complementary base pairs (A,T) and (G,C). Ultimately, this means that one chain of DNA determines the sequence of the opposing strand of DNA. Thus Watson and Crick envisaged that one strand would serve as a template during the process of DNA replication (section 1.2.2).

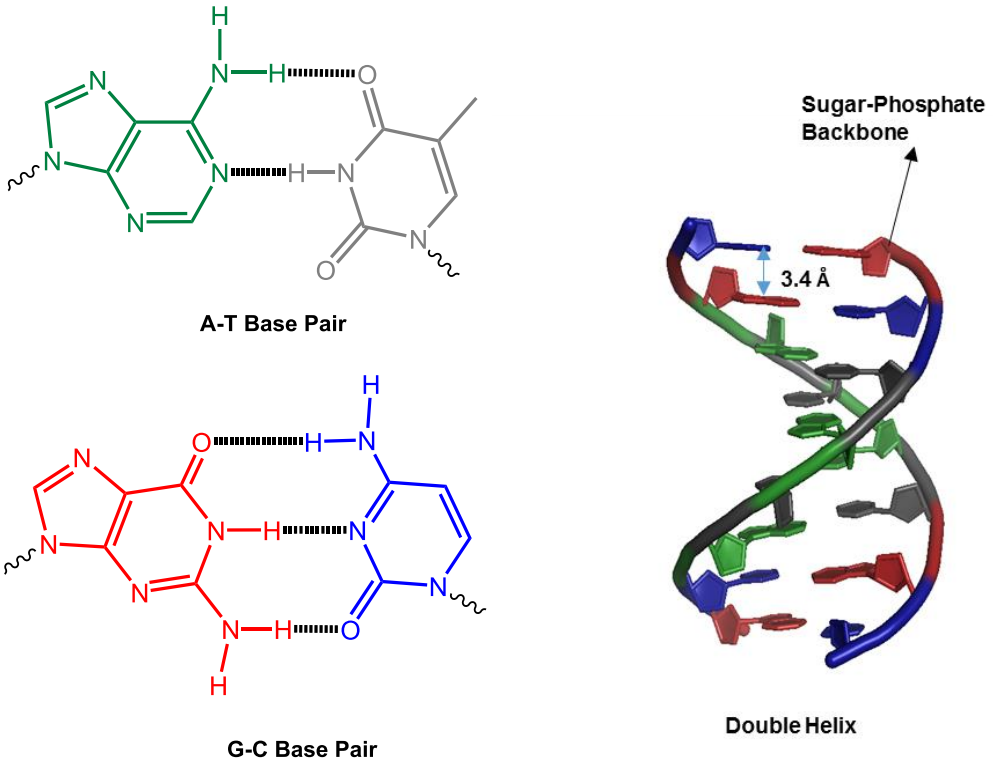


Figure 1.3. Left: Watson and Crick base pairing. Right: B-form double helical structure of DNA illustrating the base pairs and twisted sugar-phosphate backbone.

The helix is 20 Å in diameter, and each residue is 3.4 Å apart with 10.5 base pairs per helical turn.[6] This is the most favourable and efficient arrangement of DNA and also gives rise to stabilising π-π stacking interactions between adjacent base pairs. The base pairing also leads to two types of groove within the double helix, a narrow minor groove and a wider major groove.

1.2.1 The Major and Minor Grooves

The nature of the double helix gives rise to a wide major groove, where the backbones of the two strands are far apart and a narrow minor groove, in which the strands are closer together (figure 1.4). Each of these grooves has different properties due to steric effects and the availability of differing hydrogen bonding sites.

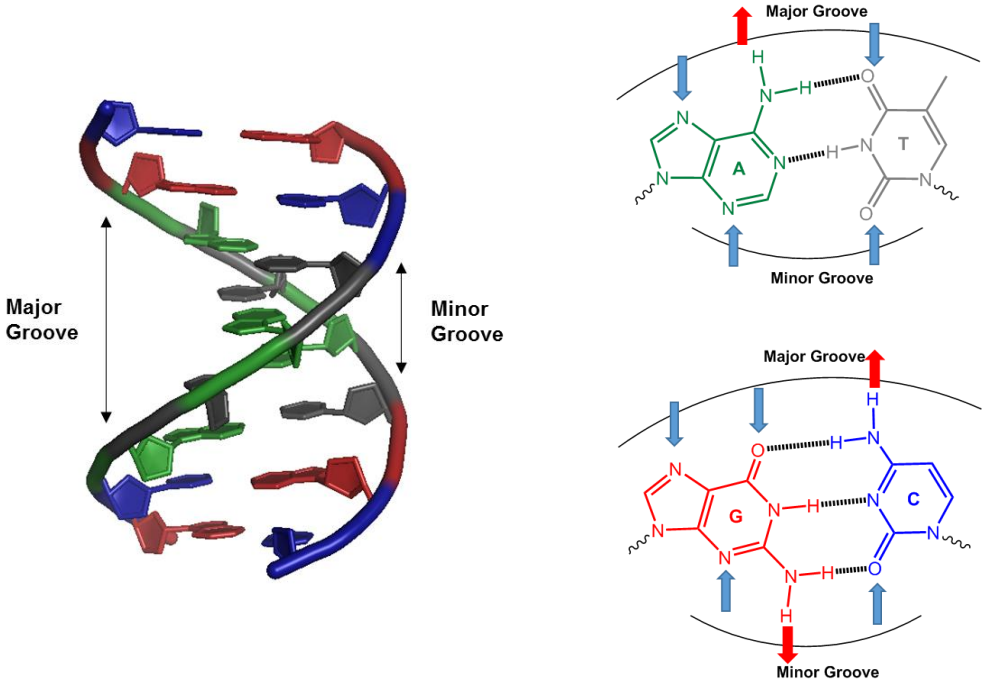


Figure 1.4. Diagram of major and minor grooves of DNA. The major and minor grooves of the A-T and G-C base pairs. The hydrogen bond acceptors are indicated by blue arrows and the donors indicated by red arrows.

The minor groove of the A-T base pair has a hydrogen bond acceptor positioned at N3 of adenine and O2 on thymine. In the G-C base pair the O2 carbonyl can also act as an acceptor and the amino group present on guanine also acts as a hydrogen bond donor site. However, the amino group in guanine does deleteriously affect the availability of N3 in guanine and O2 of cytosine. The major grooves are also lined with different hydrogen bonding sites. In the A-T base pair, the N7 of adenine and the O4 of thymine act as acceptors with the amino group on adenine acting as the hydrogen bond donor. For the G-C base pair, the N7 and the O6 carbonyl act as acceptors and the amino group of cytosine acts as a hydrogen bond donor. This array of different donor and acceptor sites provides a pattern by which small molecules and proteins can recognise sequences of DNA without

disturbing the double helix structure.[7] The pattern is clearly visualised in figure 1.5.

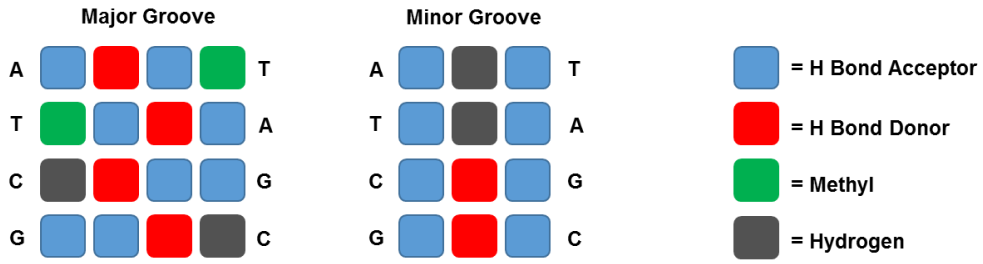


Figure 1.5. The hydrogen bonding patterns of the major and minor grooves according to differing base pairs. Adapted from reference 7.

1.2.2 DNA Replication

As DNA is the storage system for genetic information, it is important that it is duplicated accurately to conserve genes and produce identical daughter cells. Therefore, the process of DNA replication must occur before cell division, and incorporate mechanisms to recognise and correct damaged or mutated DNA. As previously explained, cellular DNA exists in the form of two complementary polynucleotide strands. Therefore, each strand acts as a template for the synthesis of a new complementary strand. This intricate system means that both strands of the DNA can be accurately copied, producing two new double helices from one molecule of DNA.

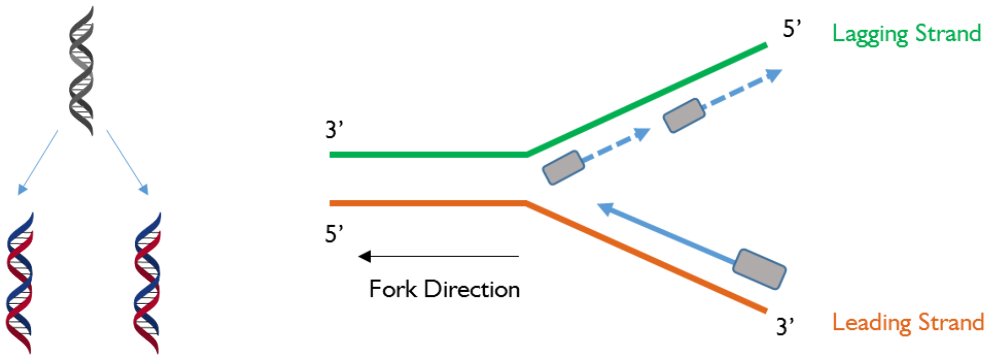


Figure 1.6. a) Two new daughter helices. b) DNA replication fork showing the lagging and leading strands.

Replication takes place at the site of a replication fork (figure 1.6), a DNA junction generated by DNA helicase. Central to the DNA replication apparatus is

the enzyme DNA polymerase, which is responsible for the addition of nucleotides complementary to the template strand. However, DNA polymerase possesses a limitation in that nucleotides can only be added in the 5' to 3' direction. As the DNA strands are antiparallel, this is problematic at the replication fork as each templated strand runs in opposite directions, 5' to 3' and 3' to 5'. The leading strand runs in the 3' to 5' direction towards the replication fork origin and thus can be synthesised continuously by addition of nucleotides from 5' to 3'. The directional problem with the opposing strand, known as the lagging strand, is solved by employing a discontinuous method of replication. This means that the DNA is synthesised in short sections, better known as Okazaki fragments.[8] DNA replication is vital for cell proliferation and thus is often a target for anticancer drugs.

1.3 Duplex DNA Conformations

Although DNA commonly adopts the form of B-DNA under normal physiological conditions, other forms of DNA have also been found.[9] Duplex DNA has the ability to form three possible structures; as well as the classical B-form, structures known as A-DNA and Z-DNA also exist (figure 1.7).

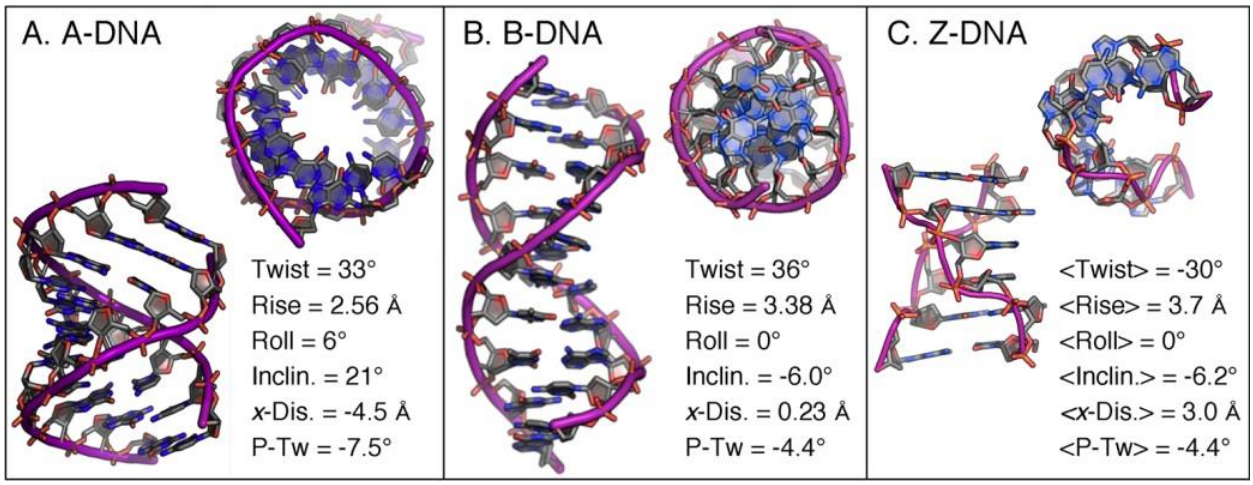


Figure 1.7. Representative images of three possible double helical structures of DNA OF A, B and Z-DNA (left to right). Abbreviations as follows: Incl. = inclination, x-Disp. = x-displacement, P-Tw = propeller twist. For Z-DNA, the helical parameters are given as averages of the alternating dinucleotide steps.[10] Figure from reference 10.

1.3.1 A-DNA

A-DNA is similar in structure to B-DNA in that it is a right handed double helix held together by Watson and Crick base pairing. The main differences arise from different sugar puckering, where the B-DNA deoxyribose is C2'-endo puckered and A-DNA is C3'-endo puckered (figure 1.8). This alternative puckering results in a more compressed structure with base pairs tilted with respect to the helical axis as opposed to being perpendicular. Consequently, this forms a 9 Å cavity down the centre of the helix illustrated in figure 1.7. B-DNA is commonly formed under cellular conditions where DNA is densely hydrated, whereas A-DNA is only stable under conditions with reduced humidity. However, in biological systems A-DNA can sometimes be formed upon binding of certain proteins.[11] RNA duplexes often form A-type structures as the 2'-OH group present on the ribose sugar inhibits the C2'-endo puckering in the B-conformer.[12]

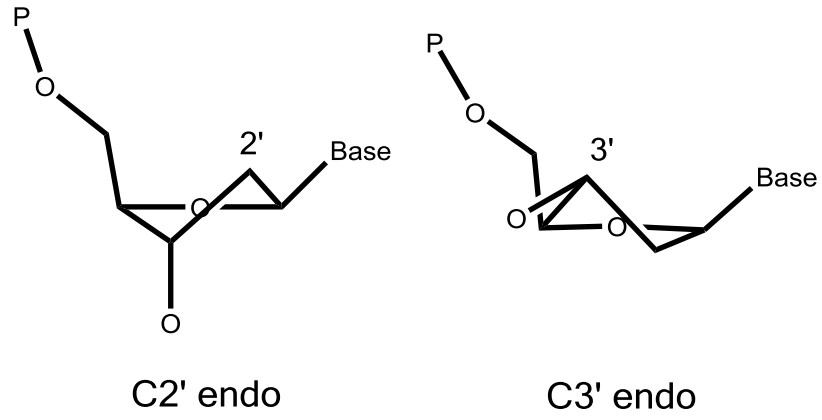


Figure 1.8. The pucker conformations of a pentose sugar in ideal B-DNA (C2'-endo) ideal A-DNA is (3'-endo).

1.3.2 Z-DNA

In contrast to the other 2 major duplex structures, Z-DNA is a left handed helix and has mostly been found in dinucleotide repeating purine pyrimidine sequences of (CG)n or (TG)n.[13] This is where the glycosidic torsion angles and sugar pucker alternate, with the purines adopting a C2'-endo anti conformation and the pyrimidines adopting a C3'-endo syn conformation (figure 1.9). This produces a structure with grooves of equal depth and width analogous to the minor groove of B-DNA but in this case no major groove exists. The conformer name Z-DNA is derived from the zigzag structure of the backbone arising from the alternating bases which differs from the smooth continuous coil present in other conformations.[14] The phosphate groups in Z-DNA are closer together than in B-DNA and under physiological conditions the electrostatic repulsion of the phosphate forces the B conformer. Under high salt concentrations, the electrostatic repulsion is shielded by the electrolyte and Z-DNA can form. In biological systems Z-DNA is formed during transcription under areas of negative torsional strain.[15] The Z-DNA can be found behind the moving RNA-polymerase as the DNA is unwound and subjected to negative torsional strain.

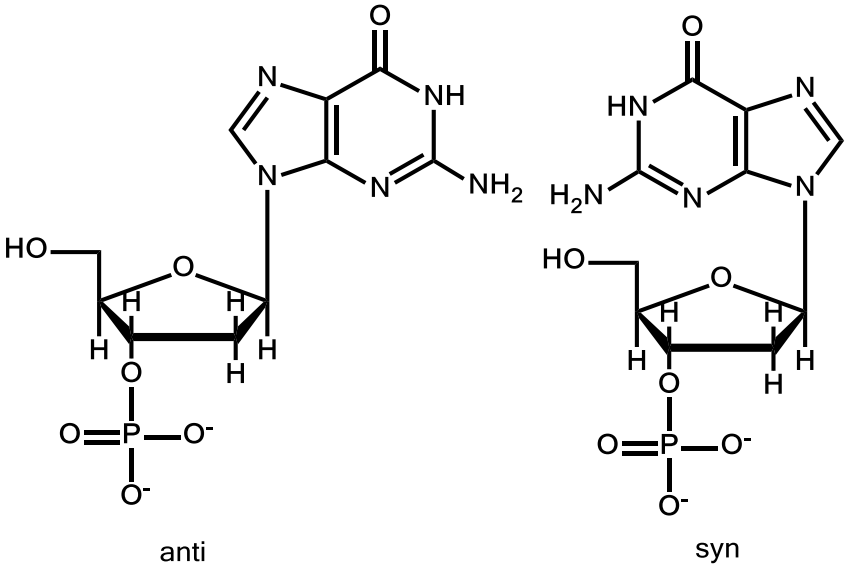


Figure 1.9. A representation of a guanosine nucleotide showing the possible stable conformations adopted through rotation about the glycosidic bond.

1.4 G-Quadruplex DNA

As well as duplex DNA, which consists of two strands of DNA, higher order structures can be formed by interactions between a greater number of strands of DNA.[16] A G-quadruplex is a four stranded DNA secondary structure formed from one or more DNA molecules binding by intermolecular or intramolecular interactions at guanosine residues. The first structure reported by Gellert, *et al.* in 1962, consisted of a square planar arrangement of four guanine residues coordinated through Hoogsteen hydrogen bonding. It is this structure which forms the basic unit of a G-quadruplex and is termed the G-tetrad (figure 1.10).[17] A central monovalent cation stabilises the tetrad through coordination to O6 of the guanine residues. Generally, under physiological conditions this stabilising cation is K^+ but can also be Na^+ or NH_4^+.[18] As the G-tetrads consist of large planar π-systems, they are able to form favourable stacks in solution with a slight helical twist forming the G-quadruplex. X-ray analysis revealed this to be the case with a hole in the centre accommodating a water molecule.

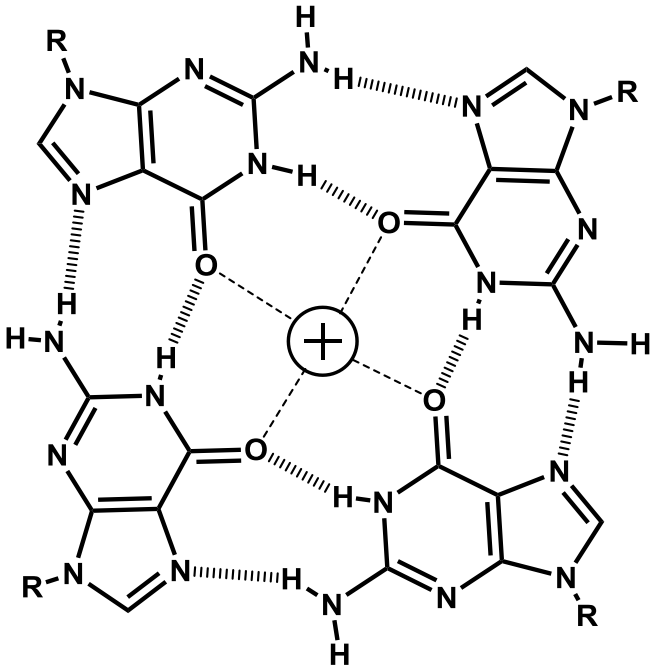


Figure 1.10. A G-quadruplex tetrad composed of four guanosine residues held together by Hoogsteen bonding and stabilised by a central monovalent cation such as K^+.

1.4.1 G-Quadruplex Conformations

G-quadruplexes can fold and be formed in a variety of ways with large amounts of structural polymorphism giving rise to different structures. The same primary sequence can fold into a different topology depending on the conditions (figure 1.11).[19] There are four main factors to be taken into consideration when determining a G-quadruplex structure. The first is stoichiometry which is dependent upon how many separate strands of DNA form the quadruplex. There are several possible combinations of DNA which result in differing structures. A quadruplex can be formed by one strand only which relies on intramolecular interactions formed by the strand folding back on itself. However, there are also intermolecular possibilities whereby multiple strands interact to produce the G-quadruplex.

The primary sequence of the DNA can also affect the structure, as differing nucleotide sequences result in various interactions. For example, longer loop sections can change G-quadruplex folding and strands orientation due to steric or other contributing factors.[20] The third contributing factor is strand orientation itself. As previously mentioned, DNA has a directionality of 5'-3' which means that strands can be aligned in either a parallel or antiparallel fashion affecting the overall structure of the quadruplex. Finally, the type of cation occupying the centre of the tetrads can have a marked effect on the structure, with certain cations stabilising a particular topology. [20]

These contributing factors lead to a number of different possible structures (figure 1.11 c-f). There are four types of unimolecular structures with three observed in the human telomeric repeat sequence. These consist of a 'propeller' type parallel structure with three chains having reversal loops, a 'basket' type antiparallel structure and a mixed structure containing one parallel and three antiparallel strands. The other unimolecular type is known as the 'chair' with alternating parallel strands. As well as this, bimolecular and tetramolecular structures also exist (figure 1.11 a and b).

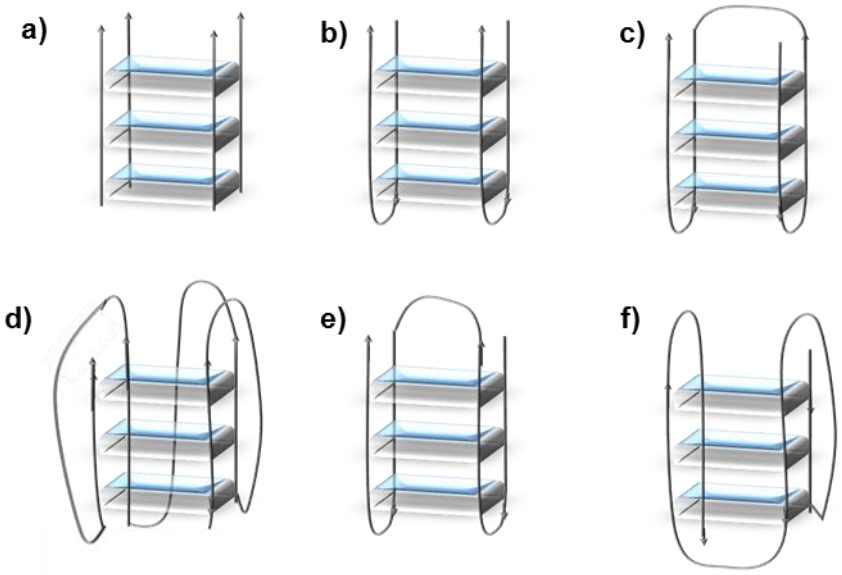


Figure 1.11. Conformations formed by a G3 quadruplex. A) Tetramolecular structure with all strands parallel; (B) Bimolecular antiparallel structure with adjacent parallel strands; (C) Unimolecular antiparallel chair structure; (D) Unimolecular propeller parallel structure; (E) Unimolecular basket antiparallel structure; (F) Unimolecular mixed structure with three parallel and one antiparallel strands.[19]

Many quadruplex forming sequences have been found to occur *in vivo* and are implicated in the down regulation of genes. For example, quadruplex formation around promotor regions causes inhibition of transcription.[21] One such quadruplex forming sequence is found in the promoter region of the c-Myc proto-oncogene.[22] This particular oncogene is linked to the potentiation of cellular proliferation and this has been associated with several malignancies including cancers. Therefore, small molecules that target this quadruplex forming region have the potential to repress c-Myc transcription.[23,24]

1.4.2 The Human Telomere

Another biological example of a G-quadruplex forming region is the telomere. The telomere is a specialised region of linear DNA at the 3' end of chromosomes which in humans consists of noncoding repeat sequences of $(TTAGGG)_n$. The function of the telomere was first elucidated to be a way of distinguishing the ends of chromosomes from broken DNA and therefore protecting from exonuclease degradation, recombination and end to end fusion.[25] The other critical function of the telomere is to eliminate the problem of 'end replication' during DNA replication.[26] This problem arises when the replication fork reaches the end of a chromosome where there is no DNA left to place an RNA

primer. This means that there would be a gap in DNA at the 3' end which would ultimately lead to a loss in genetic information over time through cell division. [27] In order to overcome this, a repetitive telomeric sequence of DNA is positioned at the end of chromosomes (figure 1.12). Each time cell division occurs, the telomeric sequence is shortened which is thought to be a governor for the lifetime of a cell. Once these telomeres become too short the cell enters replicative senescence where no replication occurs and apoptosis takes place (programmed cell death).

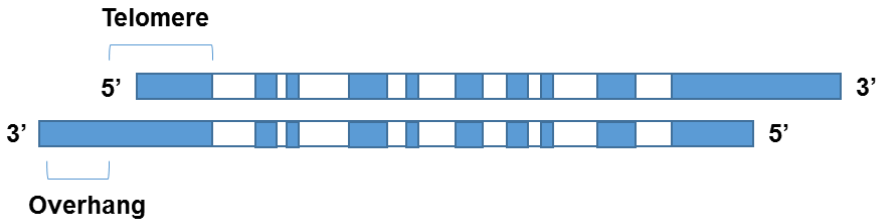


Figure 1.12. Diagram illustrating the position of telomeres in chromosomes.

1.4.2.1 Biological Relevance

One reason for the immortality of cancerous cells is due to the over expression of telomerase which is expressed in 80-85% of proliferating cancer cells.[28] Telomerase is required to synthesise telomeric DNA and works on the basis that there is single stranded DNA as the template. Thus, studies suggest that the stabilisation of telomeres in G-quartet structures can inhibit the activity of telomerase.[29] Therefore, the design of small molecules that bind to and stabilise G-quadruplex structures is an encouraging avenue of research for cancer therapeutics.

The conformation of the human telomere structure *in vivo* has been the subject of debate: a comprehensive study in 1993 using NMR in Na^+ solution showed it formed an anti-parallel basket (E in figure 1.11). However, a crystal structure with K^+ ions produced in 2002 reported a parallel propeller like structure (D in figure 1.11). Although, the X-ray structure uses K^+ ions, which is more likely *in vivo*, it is subject to crystal packing forces, whereas the NMR study uses Na^+ ions but is in solution. Due to this polymorphism it is conceivable that both structures exist *in vivo*.[30] However, a more recent NMR study using K^+ ions

proposes that a hybrid parallel/antiparallel structure is the most likely form in solution.[31,32]

1.5 Cancer

Study into cancer forms one of the largest fields of scientific research, as it presents so many variations no single treatment works for all. Cancer cells are defined by their ability to proliferate uncontrollably and form a group of abnormal cells classed as a tumour.[33] These cancerous cells can then metastasize and invade other areas of the body. It is these metastases that make cancer so difficult to treat. Additionally, resistance can be intrinsic or develop during a course of treatment. Apart from developing new treatment regimes, further research into new diagnostic techniques is required. As previously stated, telomerase is very active in cancer cells and thus it is perceived that the stabilisation of quadruplexes in telomeric DNA is a promising therapeutic lead as telomerase inhibitors.[34] This means that the G-quadruplex DNA is also an attractive target for bioimaging.

1.6 B-DNA Binding Modes

Due to its chemical composition and shape, small molecules can bind to duplex DNA. These systems offer potential as chemotherapeutic drugs and possible treatments for other genetic based diseases, as well as biomarkers for use in diagnostics. There are various ways in which molecules can interact with DNA both irreversibly and reversibly.[35]

1.6.1 Irreversible Binding Modes

Molecules can bind irreversibly to DNA by forming strong covalent bonds which is the essential mode of action behind classical chemotherapy. Cisplatin, $PtCl_2(NH_3)_2$, is the most famous example of a drug that binds irreversibly to DNA and was a breakthrough for chemotherapy, increasing the recovery rate for testicular cancer from 20% to around 90%.[36] The complex forms coordination bonds from the Pt(II) metal centre to the N7 nitrogen on purine bases, most commonly guanine, as this nitrogen is a stronger Lewis base. Most cisplatin-DNA

adducts are formed through intrastrand crosslinks which results in the formation of large bends in the DNA duplex, and an opening and flattening of the minor groove. Ultimately, this change in shape of the DNA helix is thought to block transcription, as well as disrupt the cellular repair mechanism within cells, thus causing cell death.[37] The structure was characterized in 1995 using x-ray crystallography by the Lippard group (figure 1.13).[38]

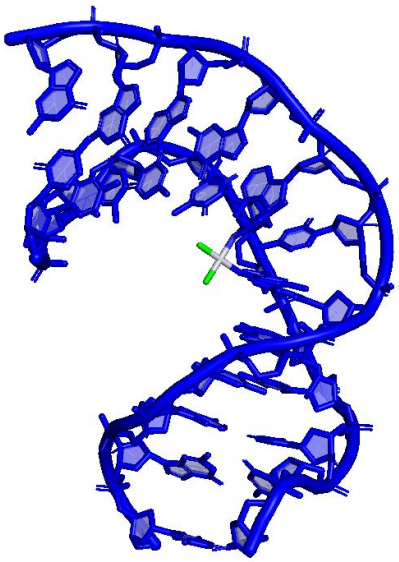


Figure 1.13. An NMR structure of a cisplatin-DNA adduct showing an N7-intrastrand crosslink of adjacent guanines. NDB:1A84[39]

Although cisplatin is a widely employed drug in cancer treatment, the drug itself is associated with severe side effects including nephrotoxicity which puts a limit on dosage.[40] Not only this, but some cancer cells develop resistance against cisplatin meaning that some cell lines are becoming less responsive to treatment. Therefore, alternative drugs with different binding modes are being investigated in an attempt to overcome drug resistance and reduce side effects.

1.6.2 Reversible Binding Modes

Other types of molecules are able to bind to DNA by reversible interactions. This type of binding falls into three types which are discussed in this section.

1.6.2.1 Electrostatic Binding

In physiological conditions DNA exists as a polyanion with a large overall negative charge due to the presence of phosphate groups in the backbone. Thus, the stability of DNA in solution is dependent on the association of cations. This means that positively charged molecules experience an attractive force towards DNA and can become electrostatically associated. However, this type of interaction is not sequence specific and is often relatively weak with low binding affinities.[35] It is this kind of binding that is observed with simple cations such as Mg^{2+}. Spermidine and spermine, which are protonated and thus positively charged at physiological pH, are an example of natural polyamines that bind to DNA through electrostatic interactions to ensure that DNA remains compact (figure 1.14).[41]

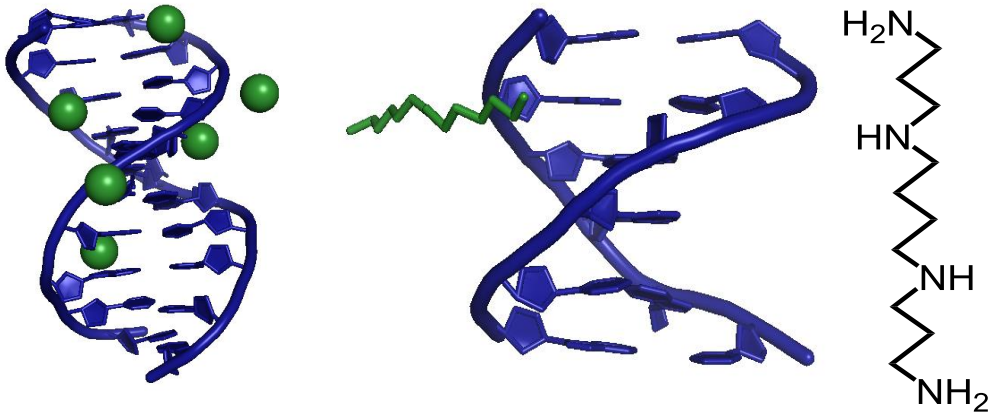


Figure 1.14. X-ray crystallographic structures indicating electrostatic DNA binding of left, magnesium (II) ions (PDB entry 4R4D); middle, spermine (PDB entry 206D) and right, a schematic of the polyamine molecule spermine[42,43]

1.6.2.2 Groove Binding

The reversible mode that most DNA-binding proteins employ is groove binding. As described earlier, duplex grooves exhibit variation in properties such as sterics, hydration, electrostatic potential and hydrogen bonding sites. Thus, their nature means that molecules can display selective and highly specific binding interactions to either the major or minor grooves of the helix, when the molecule approaches within the van der Waals contact distances of the groove walls.[35]

Hydrophobic interactions, hydrogen bonding and electrostatic forces generally play a key role in binding. A typical minor groove binder is cationic and consists of aromatic rings connected by flexible bonds, which enable the molecule to adopt a crescent-shape complementary to the helical structure of DNA. Most molecules that bind this way generally bind in AT rich regions which are more electronegative and narrower, maximising electrostatic and van der Waals forces. Hydrogen bonds that form between the groove binder and DNA enable site-specific interactions. An example of a model groove binder is the antibiotic netropsin which binds to AT rich residues in the minor groove of B-DNA with high affinity (figure 1.15).[44]

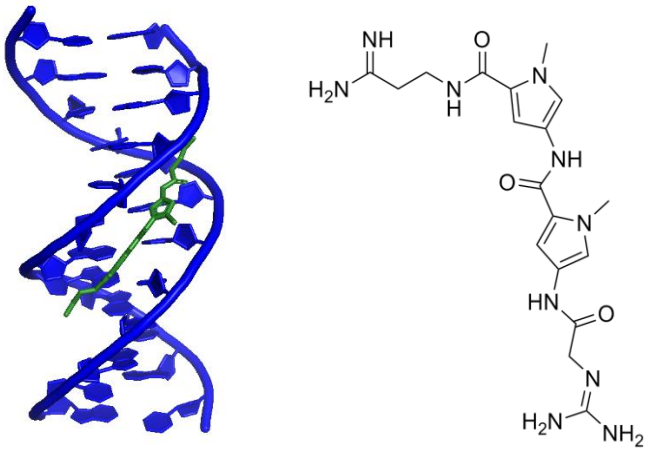


Figure 1.15. The X-ray crystallographic structure of the minor groove binder netropsin-DNA complex PDB entry: 101D.[45]

The mechanism of groove binding can be separated into two steps shown in figure 1.16. Firstly, the molecule must undertake a hydrophobic transfer process from the bulk solution into the minor groove of DNA. The cationic charge on the

groove binder is often responsible for the simultaneous release of condensed counterions which is entropically favourable in terms of the driving forces for binding. Once the molecule resides in the minor groove, the molecular interactions such as hydrogen bonding can be made.[46]

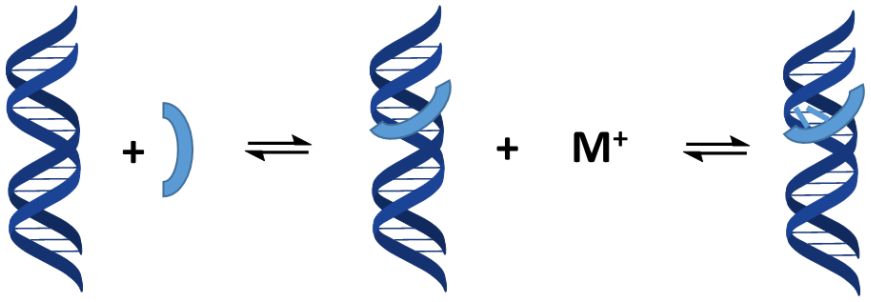


Figure 1.16. Representation of the steps that take place during DNA groove binding.

1.6.2.3 Intercalation

Another method of reversible binding, intercalation, was first reported in 1961 by Lerman.[47] Intercalation occurs when compounds are inserted between the base pairs of the DNA double helix and are accommodated by the unwinding and elongation of the backbone. The structure of such compounds generally consists of polycyclic aromatic rings, which are planar and able to efficiently stack between the base pairs. The binding is stabilised by considerable π-overlap between the base pairs and the planar molecule as well as contributions from electrostatic, hydrophobic and van der Waals interactions. Intercalation causes the stiffening, lengthening and untwisting of the DNA double helix due to the steric strain applied upon insertion, giving them the potential to be therapeutics. Due to the polyaromatic nature of intercalators, they are often chromophores and thus can also be utilised as fluorescent biomarkers for DNA.[48] The commonly used nucleic acid stain, ethidium bromide, is an example of a well-studied intercalator (figure 1.17).[49]

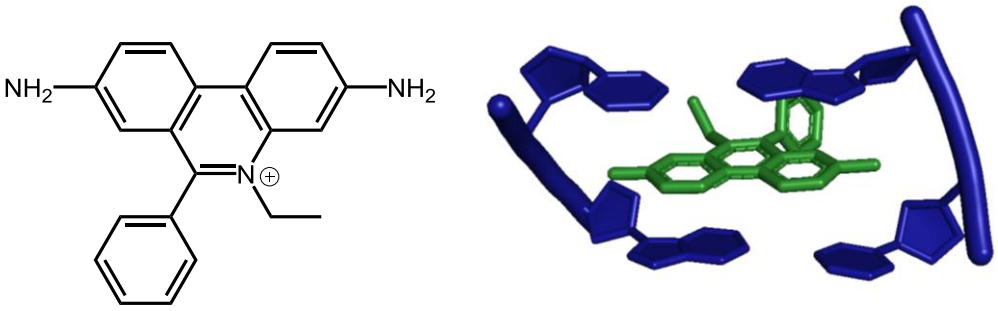


Figure 1.17. The structure of the classical intercalator ethidium (left) and an X-ray crystallographic structure showing a section of an ethidium bromide-DNA complex. NDB entry: DRB006.[50]

Whereas groove binding is an entropically driven process, intercalation is generally more enthalpically driven.[51] Unlike groove binding, intercalation is considered to be divided into three steps. Firstly, the DNA has to produce an intercalation site by undertaking a conformational change whereby the base pairs undergo separation. The intercalator then moves from the bulk solution to occupy the newly formed gap in the base pairs. Once inserted between the base pairs, favourable interactions can form, such as π-π stacking and hydrogen bonding. Although the first step of base pair separation is relatively energetically unfavourable, the release of counterions when the DNA unwinds drives intercalation forward. The insertion of the intercalator itself is also favourable as it is considered to be a hydrophobic transfer process. The principle of this overall model is demonstrated in figure 1.18.

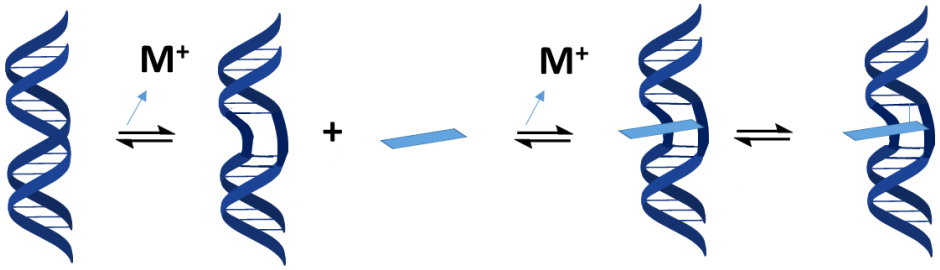


Figure 1.18. Model for intercalation showing DNA unwinding to create an intercalation site.

Metallo-intercalators offer an attractive array of properties in the field of intercalation chemistry and are introduced in the following section.

1.6.2.4 Threading Intercalation

Threading intercalation is a class of intercalation in which the aromatic intercalative unit is flanked by bulky substituents, making classical intercalation impossible until the bulky units are threaded through the base stack. This initially requires severe distortion of the DNA duplex so that the bulky moieties can finally reside in opposite grooves. As a result, this method of binding gives very stable intercalated species with very slow dissociation rates. The binding itself is stabilised by π-π stacking interactions from the planar aromatic intercalated unit but also by favourable groove binding interactions from the bulky substituents. This mode of intercalation has been observed for organic molecules such as the natural product and antibiotic molecule nogalamycin shown in figure 1.19.[52,53]

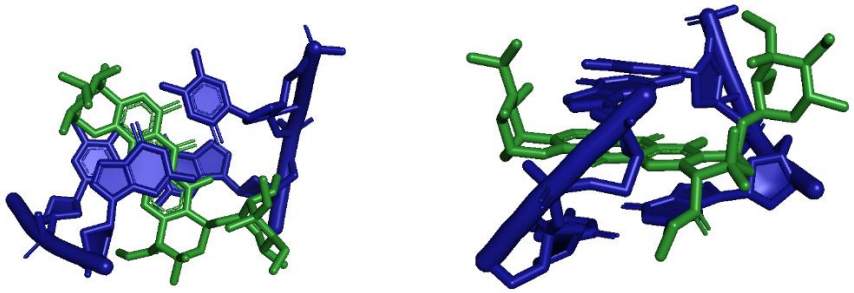


Figure 1.19. X-ray crystallographic structure showing a section of a nogalamycin-DNA complex outlining the threading intercalation mode of binding. Views parallel and perpendicular to the helical axis PDB entry 1D21.[54]

1.7 Transition Metal Complexes as DNA Binding Agents

Since the introduction of cisplatin, research into inorganic transition metal complexes as potential biophysical or therapeutic agents has evolved.[55] Many of these inorganic therapeutics have been based on cisplatin analogues, but there has also been extensive research into other transition metals.[56] A particularly interesting group of this kind are octahedral ruthenium (II) complexes. An example of an archetypical ruthenium complex is shown below in figure 1.20 consisting of a metal centre and multi-dentate pyridine based ligands, such as phenanthroline (phen) and bipyridine (bipy).

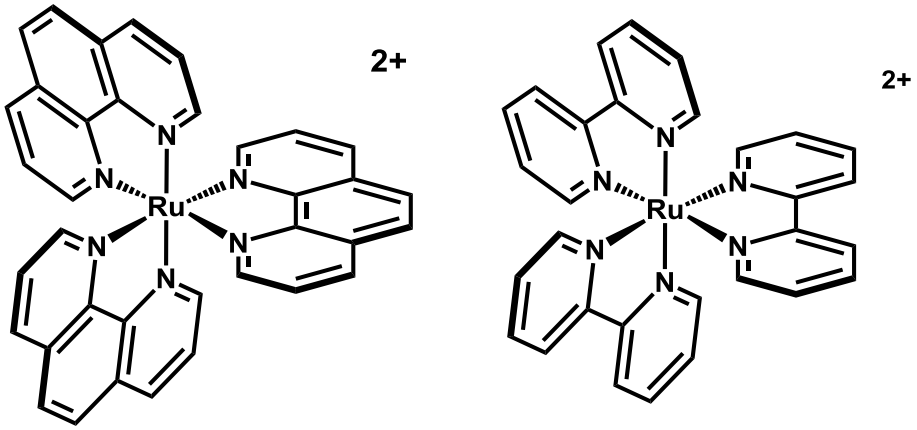


Figure 1.20. Chemical structures of two simple ruthenium polypyridyl complexes $[Ru(phen)_3]^{2+}$ (left) and $[Ru(bipy)_3]^{2+}$ (right).

1.7.1 Early Phenanthroline Complexes

The complex, $[Ru(phen)_3]^{2+}$, was first investigated by Barton, *et al.* during the 1980s.[57] It was reported that the complex binds to DNA via an intercalative method. The octahedral nature of the molecule results in chirality and thus the complexes can exist as left or right handed (Λ or Δ) enantiomers shown in figure 1.21. Barton noted that as DNA is also a chiral molecule, the Λ and Δ forms exhibit different binding affinities and reported that the Δ enantiomer displays a stronger affinity for the right handed DNA helix. This chiral discrimination was suggested to be due to the sterics involved during intercalation into the DNA. The intercalation of one phenanthroline ligand would leave the two remaining ancillary ligands to fit along the helical groove. This occurs with little steric impact for the Δ enantiomer, as the ligands run parallel to the groove. Thus, this study introduced the interest in ruthenium complexes as probes to study the secondary structure of DNA.[58]

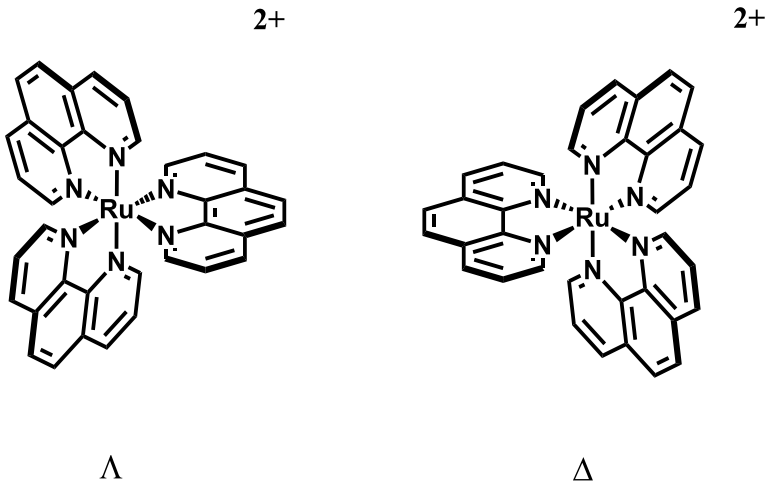


Figure 1.21. The Λ (left) and Δ (right) forms of $[Ru(phen)_3]^{2+}$.

After further study, Barton released a paper in 1986 detailing that surface binding was also contributing to the DNA binding of $[Ru(phen)_3]^{2+}$.[59] It was not until 1990, that a different research group elucidated that intercalation was, in fact, not taking place.[60] Instead the complexes were found to be located in the major groove. Although, the binding mode was revealed to be different, it was still deduced to be sterics causing the difference between the two, and neither enantiomer were shown to intercalate.[61] As discussed in section 1.6.2.3, ethidium bromide is an established intercalator and is known to lengthen DNA, thus increasing viscosity. However, groove binders such as Hoescht 33258 demonstrate no change in viscosity upon the addition of DNA and Chaires showed that neither enantiomer of $[Ru(phen)_3]^{2+}$ induced a change in viscosity.

1.8 Metallo-Intercalators

Despite this initial failure, ruthenium(II) polypyridyl complexes certainly provide a three dimensional scaffold to introduce an intercalating ligand and potentially engineer intercalators with targeted physical and photophysical properties. Thus, these systems display advantages over organic systems and are attractive starting points for the design of DNA interactive systems. Their syntheses are often more facile than organic compounds, with less need for protecting groups and fewer intermediate steps. As they are frequently cationic this improves binding affinity to the polyanionic structure of DNA and enhances water solubility, which is a necessity for efficient drug delivery. Ruthenium(II) complexes are also very stable, due to a low spin d^6 electron configuration, rendering them kinetically inert. This is one of the most important properties when designing a new therapeutic complex as it is essential that the molecule is stable within the cell. This also means that most complexes exhibit an octahedral geometry, resulting in a rigid 3D structure which can facilitate recognition of the DNA target.

As the non-intercalative ancillary ligands are also often planar and aromatic, they can provide additional sites for hydrophobic contacts to be made during intercalation into the DNA. As the ligands can be easily modified, they can be used to vary the steric and electronic properties and enhance cellular uptake or DNA binding specificity. The combination of the ancillary ligand with the metal centre equips the complexes with unique photophysical characteristics which will be explained in the following section. These luminescence properties often have high stokes shifts for application as microscopy staining agents.[62] Some complexes have also been found to photochemically interact with long lived triplet excited states capable of generating singlet oxygen. This gives potential application as photosensitisers for photodynamic therapy.[63]

1.8.1 The DNA Light Switch Effect

In 1990, a novel ruthenium complex utilising the intercalating ligand dipyridophenazine $[Ru(bipy)(dppz)]^{2+}$ (figure 1.22) was reported by Barton et al.[64] The dppz ligand has an extended planar surface in comparison to the parent phen ligand, which can intercalate into B-DNA, resulting in a complex with a heightened binding affinity (K_b) of ~10^7 M^{-1}. In the presence of DNA, an interesting photophysical phenomenon now commonly known as 'the molecular light switch effect' is observed. It was reported that when bound to DNA or in aprotic solvents the metal complex displayed luminescence.

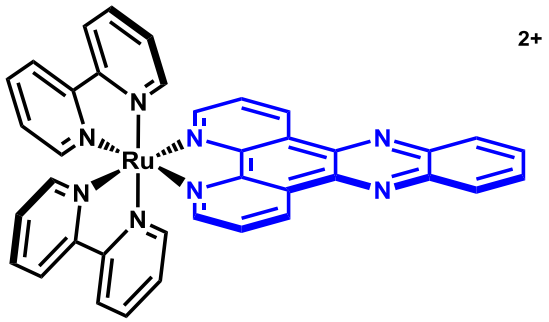


Figure 1.22. Chemical structure of [Ru(bipy)2dppz]2+ with the intercalating dppz ligand shown in blue.

The molecule transitions from virtually zero emission in aqueous solvents when unbound to DNA, to a strong luminescent emission when bound to DNA. It is this distinct change in property, enabling the distinction of two different environments that accounts for the 'switch' term. Overall, the effect is useful in the field of biomarkers for DNA as there is no background emission from any unbound complexes. This enables the complex to function as a probe by 'switching on' when bound to DNA and 'switching off' in the absence of DNA as demonstrated in figure 1.23.

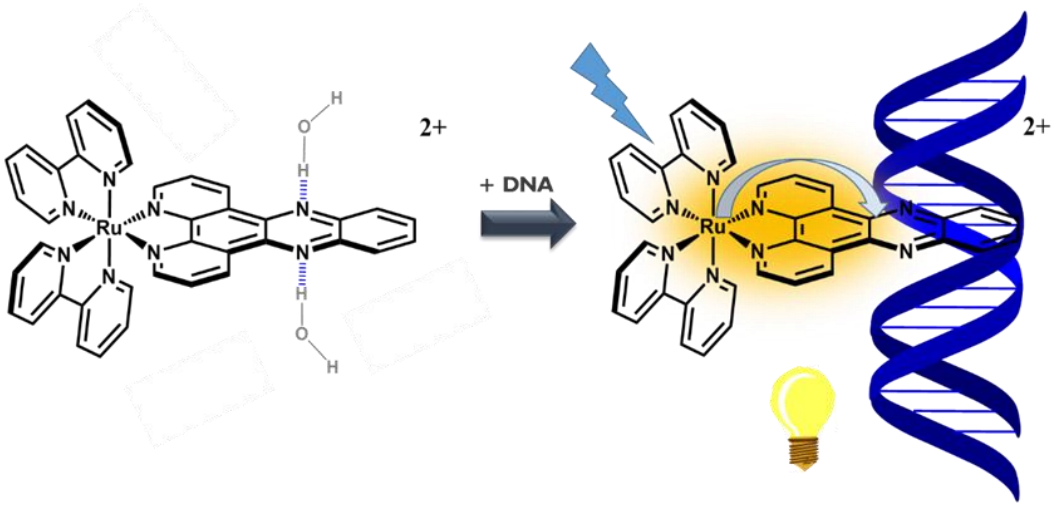


Figure 1.23. A schematic representation of the light switch effect.

Study into the photophysics of the complex has revealed that luminescence occurs as a result of the formation of an excited triplet state formed via a metal-to-ligand charge transfer (MLCT) shown in figure 1.24. This charge transfer was recognised to be a transition from the metal centre to the phenazine moiety of the dppz ligand.[65] There are then possible deactivation pathways including both emissive and non-emissive. Quenching of the excited state occurs in aqueous solvents due to the formation of hydrogen bonds to the free nitrogens on the dppz ligand resulting in a non-emissive deactivation.[66] The consequent luminescence enhancement upon binding to DNA is attributed to the intercalated species being shielded from water molecules. This means that water can no longer access the intercalating ligand and emission is observed. Other derivatives of the original complex have been developed to take advantage of the DNA binding and photophysical properties for utilisation as therapeutics and DNA probes.[67]

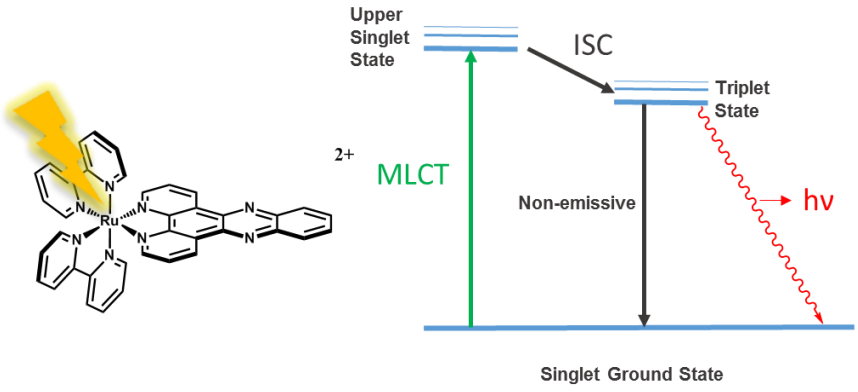


Figure 1.24. Schematic represenation of the decay pathways for the MLCT.

As mentioned previously, two enantiomers exist for this class of complex. Lincoln, *et al.* demonstrated that although they bind with similar binding affinities, the quantum yield of the related Δ-$[Ru(phen)_2dppz]^{2+}$ is much higher than that of the Λ-$[Ru(phen)_2dppz]^{2+}$.[68] Therefore, the emission observed for the racemic mixture is mostly due to the Δ enantiomer. The Barton group reported that the difference in luminescence was due to two different orientations upon binding to the DNA through intercalation in the major groove.[69] Later groups suggested the binding mode to B-DNA was from the minor grooves with controversy arising as to which was the correct orientation.[70] More recent crystallographic studies have confirmed these metal complexes bind by intercalation in the minor groove.[71,72] A crystallographic study of the binding of Λ-$[Ru(phen)_2dppz]^{2+}$ to oligonucleotide sequences $d(CCGGTACCGG)_2$ demonstrated that the complex bound to the central T|A base step by a symmetrical 'head on' intercalation, but also to the CC|GG base step by an angled mode. However, crystals of the Λ-$[Ru(phen)_2dppz]^{2+}$ bound to the alternative DNA sequence $d(CCGGATCCGG)_2$ showed no binding at the central A|T base step with only the canted mode at the CC|GG step (figure 1.25). This effect was attributed to the lower flexibility of the A|T base step over a T|A step and an unfavourable steric interaction with the adenine base in the A|T site.

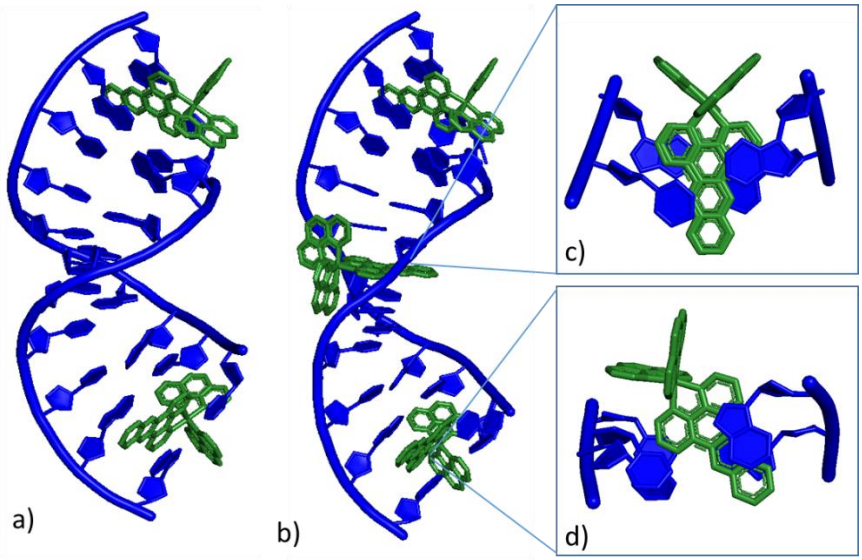


Figure 1.25. The crystallographic structures indicating the intercalation Λ-$[Ru(phen)_2dppz]^{2+}$ into the minor groove of the DNA duplex a) $d(CCGGATCCGG)_2$ and b) $d(CCGGTACCGG)_2$. Expansions of the views perpendicular to the base pair axis show c) the head on intercalation at the T|A base step and d) the canted intercalation at the CC|GG step. PDB entries 3U38 and 4E7Y.[71]

1.8.2 Tpphz

The expansion of the surface area of intercalating ligands is a possible route to improve the binding affinity towards DNA. The ditopic bridging ligand, tpphz (figure 1.26), was first synthesised by Bolger, *et al.* as a component for dinuclear Ru^{II}-based "molecular wires" [73] as it closely resembles dppz. The ligand behaves in a similar manner to dppz and when bound to a metal centre it functions as a DNA intercalator.

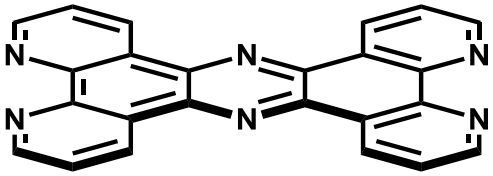


Figure 1.26. The structure of the tpphz ligand

The tpphz ligand presents two possible metal ion coordination sites at either end of the molecule and thus metal ions can bind at either end of the molecule. Therefore, mononuclear or dinuclear (both homo and heteronuclear) complexes can be formed using this ligand.[74]

1.8.2.1 Mononuclear Ruthenium (II) Complexes of tpphz

The mononuclear complexes $[Ru(bipy)_2tpphz]^{2+}$ and $[Ru(phen)_2tpphz]^{2+}$ displayed in figure 1.27 bind to DNA through intercalation and display the light switch effect in cell free conditions.[75] Given these properties, the cellular uptake was monitored by Gill, *et al.* and the light switch effect was observed in MCF-7 cancer cells. As both complexes were taken into cells it was feasible that intercalation could interfere with DNA processes such as replication. Therefore, the cytotoxicity was analysed in several cancer cell lines and both complexes displayed IC_{50} values in the same order of magnitude to that of cisplatin. More recently, Gill, *et al.* have shown that $[Ru(phen)_2tpphz]^{2+}$ acts as a 'three-in-one-bullet' for oesophageal cancer cells by blocking replication, interfering with chromosomal spindle attachment, and radio-sensitising hard-to-treat oesophageal cancer cells.[76]

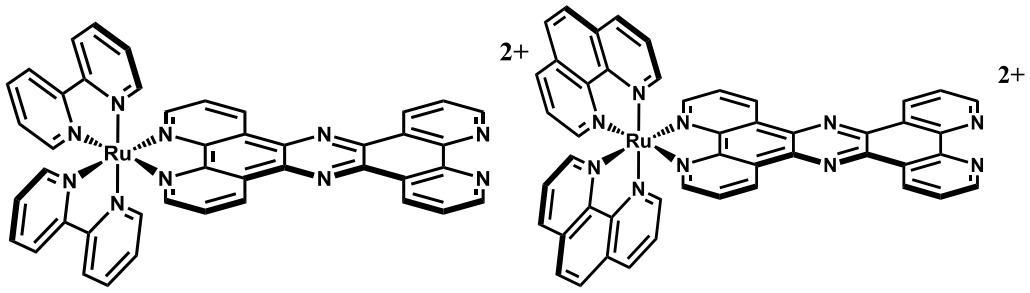


Figure 1.27. Chemical structures of $[Ru(bipy)_2tpphz]^{2+}$ and $[Ru(phen)_2tpphz]^{2+}$

Due the luminescent properties and their anticancer activity, the complexes can be considered as theranostic cellular DNA imaging compounds. Subcellular imaging of organelles and biomolecules such as DNA is not only of increasing importance in basic research and diagnostics, but also aids in understanding the dynamics of DNA structure itself within cells. Most currently employed DNA-imaging probes require ultraviolet (UV) light, which can cause structural damage to DNA,[77] and possess low water solubility, both factors which can affect their application in live cell imaging. However, as demonstrated above, DNA metallo-intercalators offer a promising range of photophysical properties to overcome these limitations.

1.9 Dinuclear Ruthenium (II) Complexes

The complexes discussed so far only incorporate one metal centre. However, it is established that dinuclear ruthenium complexes can enhance DNA binding in comparison to the equivalent mononuclear complex. It is also possible to increase chiral discrimination, as the introduction of another ruthenium centre gives three possible stereoisomers, the forms of which are given below in figure 1.28.

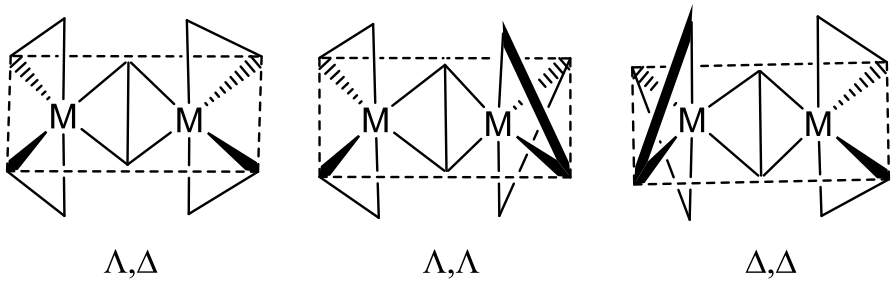


Figure 1.28. The possible enantiomers of a dinuclear ruthenium polypyridyl complex.

1.9.1. Dinuclear Threading Complexes

As discussed in section 1.6.2.4, planar molecules with bulky terminal substituents can intercalate through a threading mechanism. Consequently, dinuclear ruthenium complexes that exhibit this type of structure will have the potential to bind to DNA through threading.

Dinuclear ruthenium complexes that incorporate the intercalating light switch moiety dppz have been developed by Lincoln, *et al.*[78–80] These structures were designed to increase binding affinities over mononuclear complexes by employing two Ru^{II} centres for increased electrostatic attraction to the polyanionic DNA duplex. The structure of $[Ru(phen)_2)_2bidppz]^{4+}$ is shown in figure 1.29.[81] Lincoln, *et al.* reported that these complexes display extremely tight binding to DNA, with K_b values in the region of 10^{12} M^{-1}.

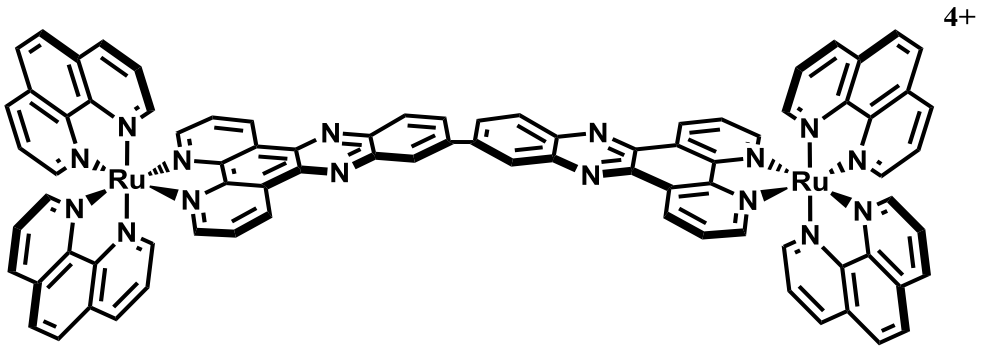


Figure 1.29. Chemical structure of the dinuclear ruthenium complex $[Ru(phen)_2)_2bidppz]^{4+}$

They then employed linear dichroism (LD) to investigate the binding of the Δ,Δ complex. Long polymers such as DNA <1000 base pairs can be aligned by the viscous drag caused by applying a hydrodynamic flow.[82] In this oriented form the electronic transitions are anisotropic and will have a different absorbance of light, parallel ($A_{\parallel}$) or perpendicular ($A_{\perp}$) to the long axis of the DNA based on the direction of the electric transition moment. This can therefore be used to deduce the binding orientations of small molecules to DNA.[83,84] LD measures the absorbance of linearly polarised light at two angles 90° from each other, it is a differential absorption spectroscopy which is defined as $LD = A_{\parallel} - A_{\perp}$. A positive LD shows an electronic transition more aligned with the long axis of the DNA and a negative LD is more perpendicular as shown in figure 1.30.

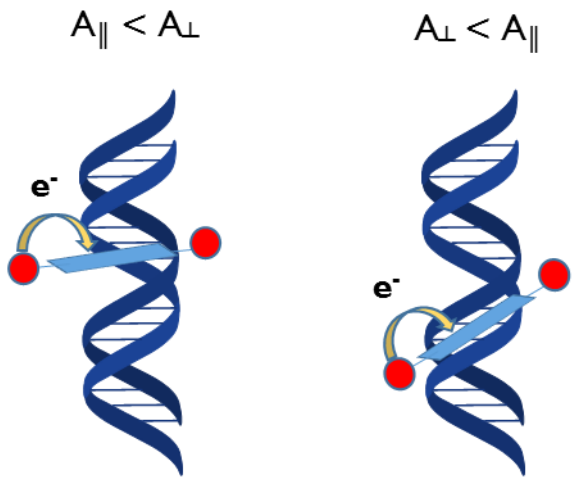


Figure 1.30. A schematic illustration of the different orientations of dinuclear complexes bound to B-DNA, intercalated (left) and groove bound (right). The arrows indicate the electronic polarisation upon MLCT excitation perpendicular to the long axis of the DNA for an intercalated complex and more parallel in the groove bound form.

It was initially reported that $[Ru(phen)_2)_2bidppz]^{4+}$ bound in the groove of calf thymus (CT) DNA, with flow linear dichroism studies giving a positive signal for the MLCT and dppz based π-π* bands. This is indicative of a transition moment parallel to the long axis of the DNA. However, in a serendipitous discovery a sample that was left for two weeks at room temperature showed a shift in the LD signal from positive to negative, indicating that it had threaded through the base pairs, as this causes the intercalating unit to be aligned with the base pair stack perpendicular to the long axis of the DNA.[80] The $[Ru(phen)_2)_2bidppz]^{4+}$ complex has since been used as a model compound for various studies in order to more deeply understand threading mechanisms.[78,85–87]

It was found that the threading event for this structure typically shows extremely slow association and dissociation kinetics,[88–90] as a large deformation of the DNA at the binding site is required for threading of the bulky ruthenium centre to occur, including the transient opening of at least one base pair for some species.[91,92] It was also found that threading was sequence specific with a heavy preference for binding at sites containing multiple AT residues with the threading of Λ,Λ-$[Ru(phen)_2)_2bidppz]^{4+}$ requiring more than six consecutive AT base pairs.[93–95] This preference for AT residues was ascribed to the higher flexibility of the AT base pair containing only two hydrogen bonding units compared to the GC base

pairs containing three. Therefore, the opening of these base pairs is energetically less demanding.

It has been shown that binding first occurs by rapid association of a metal complex in the DNA minor groove at AT rich sequences prior to slow threading.[85] This initial groove bound form of metal complex has been characterized by an NMR study by looking at the binding of Λ,Λ-$[bidppz(bipy)_4Ru_2]^{4+}$ to the DNA duplex $d(CGCGAATTCGCG)_2$ in which the lack of long AT tracts was chosen to inhibit the threading event. ^1H Nuclear Overhauser effect spectroscopy (NOESY) is an NMR experiment that probes the dipolar coupling of ^1H atoms through space. Hydrogens closer than 5 Å will display a cross peak in a NOESY spectrum. the observation of several nuclear Overhauser enhancements NOEs to protons in the minor groove of the DNA allowed for structural determination of initial groove bound complex through molecular dynamics simulations (figure 1.31).[96] The simulation was in agreement with the NOE close contacts and also identified two possible hydrogen bonding sites.

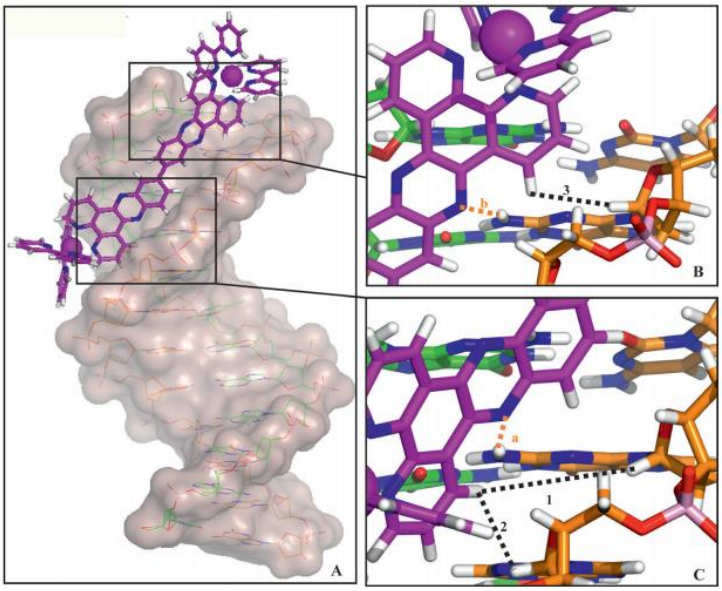


Figure 1.31. The resultant model from molecular dynamics simulations indicating the groove binding orientation of a DNA:Λ,Λ-$[bidppz(bipy)_4Ru_2]^{4+}$ complex. A) The side view of the DNA complex B,C) expansions of regions of the structure showing the short distances from intermolecular NOEs dotted lines indicate NOEs (black, numbered 1–3) and potential hydrogen bonds (orange, labelled a and b). Figure from reference 96. [96]

Lincoln, *et al.* have also shown that stereochemistry of the bulky ruthenium centres is an important factor in threading.[97] $[Ru(phen)_2)_2bidppz]^{4+}$ threading is easier for the Δ,Δ enantiomer than the Λ,Λ. However, by shortening

the threading moeity from bidppz to dppzip (figure 1.32) this effect was reversed. The effect of shortening the bridge also inhibited threading into CT DNA with threading only occuring in more flexible poly d(AT) DNA.

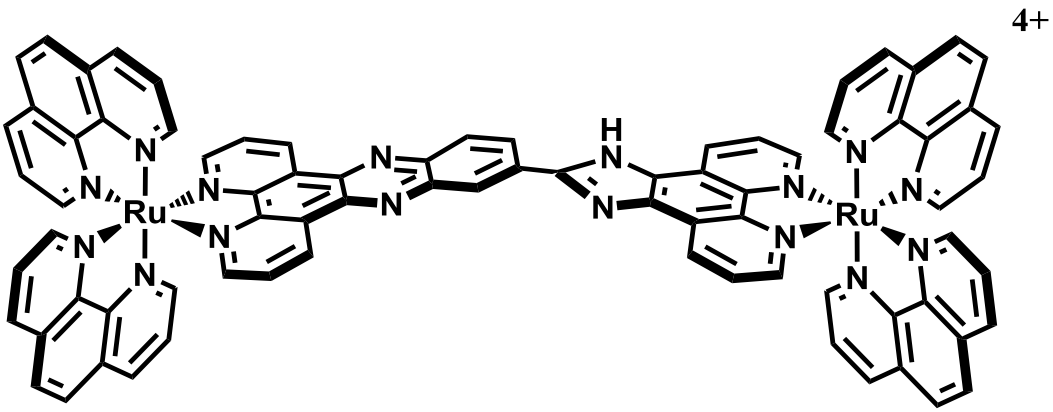


Figure 1.32. Chemical structure of $[Ru_2dppzip(phen)_4]^{4+}$

1.9.2. Bridging tpphz Complexes

The dinuclear ruthenium complexes $[(Ru(phen)_2)_2tpphz]^{4+}$ [2] and $[(Ru(bipy)_2)_2tpphz]^{4+}$ [1] which incorporate the bridging ligand tetrapyrido[3,2-a:2',3'- c:3",2"-h:2",3"'-j]phenazine (tpphz) have been developed as dinuclear derivatives of the dppz complexes (figure 1.33).[98] These compounds exhibit a high binding affinity to duplex DNA of $K_b > 10^6\ M^{-1}$ which is approximately two orders of magnitude higher than the mononuclear counterpart. Not only this, but the complexes also retain the characteristic "light switch effect" of the original dppz complexes.[99]

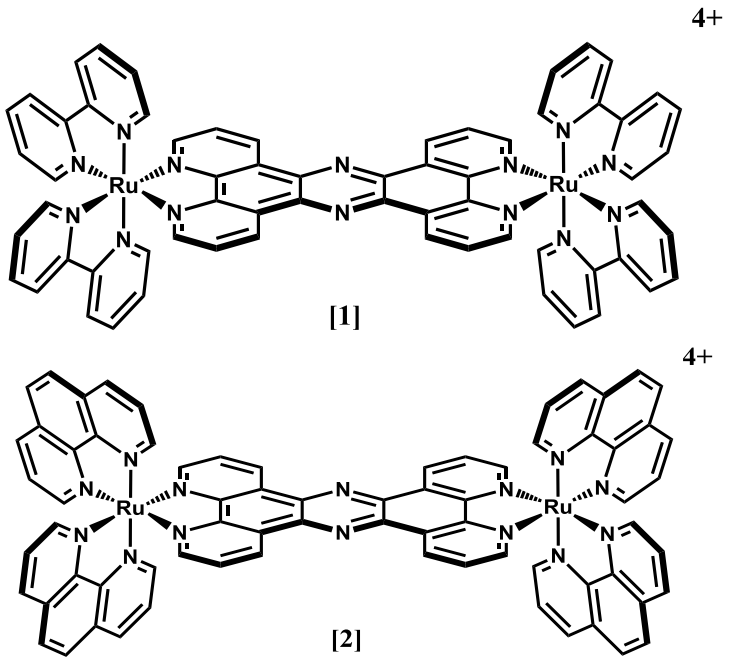


Figure 1.33. The chemical structures of $[(Ru(bipy)_2)_2tpphz]^{4+}$ [1] and $[(Ru(phen)_2)_2tpphz]^{4+}$ [2].

The molecules were also found to bind to quadruplex DNA with a similar affinity of $K_b > 10^6$ M^{-1}. Interestingly, upon binding to quadruplex DNA, the complexes display the light switch effect with blue shifted emission. The binding preferences of the complexes were assessed by screening against several quadruplex-forming sequences.[100] They were determined to bind with the highest affinity to the antiparallel form of the human telomere sequence. This provides the basis for a DNA imaging agent that can discriminate between quadruplex and duplex DNA. When bound to quadruplexes, the complex produces emission at 630 nm but when bound to duplex DNA the emission shifts to 680 nm as illustrated in figure 1.34.

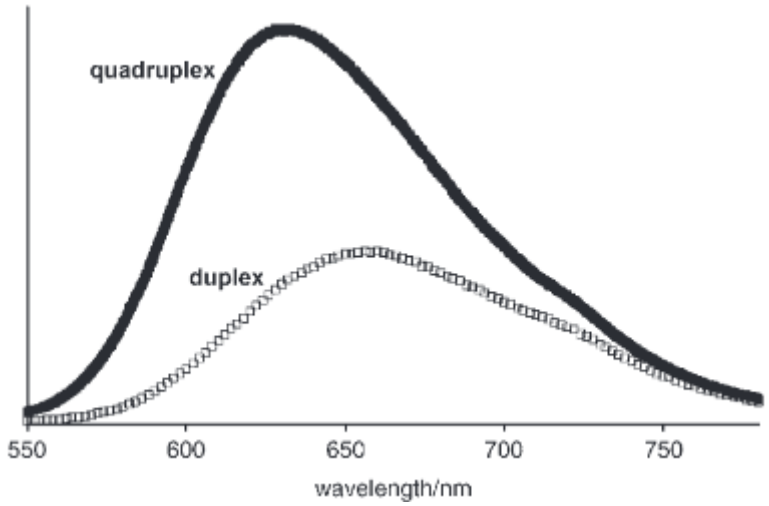


Figure 1.34. From reference 99. A comparison of emission observed for a 5 mM solution of $[(Ru(phen)_2)_2tpphz]^{4+}$ in the presence of G3- quadruplex and calf-thymus duplex DNA. λEx= 450 nm.[99]

Given the interesting potential applications of the compounds, cellular studies have been undertaken to examine whether the complexes are viable, *in cellulo*. The difference in ancillary ligands seems to play an important role in cellular uptake as previous work by Gill *et al* reveals that whilst the phen complex is readily taken up into cells, the bipy complex is not.[77] Thus, the dinuclear phen complex acts as a stain for confocal microscopy images enabling the visualisation of structural changes of DNA during the cell cycle (Figure 1.35).

Figure 1.35. From reference 77. Images of mitotic cells stained by [2] enabling the visualisation of chromosome aggregation throughout the progression of mitosis. [77]

As previously explained, the stabilisation of quadruplex DNA in telomeres is considered to be a potential approach to targeting cancer. As discussed above the complex is capable of distinguishing quadruplex from duplex DNA by a characteristic blue-shift in the 'light switch' emission. Therefore, the application of the phen complex as a structural DNA probe was investigated in the L5178Y-R cell line which contains long regions of G-telomeric DNA. In comparison with MCF-7 cancer cells, there is a significant increase in emission within the 630-640

nm range (figure 1.36). This indicates that [2] can function as a biomarker for quadruplex DNA.

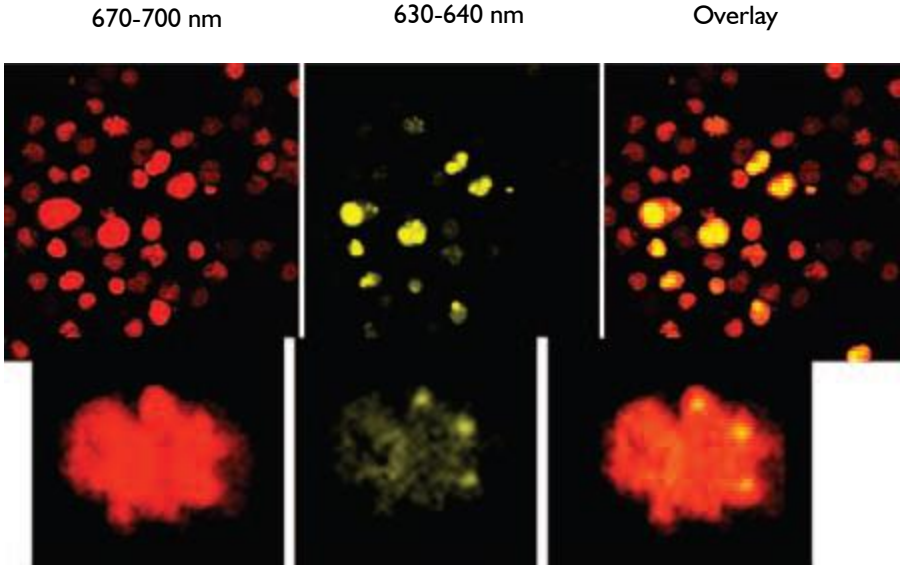


Figure 1.36. Confocal microscopy images of L5178Y-R cells (G-rich telomeric DNA) stained with 1 (200 μM, 30 min) taken using two separate detection channels at 670-700 nm for duplex DNA (red) as well as quadruplex DNA at 630–640 nm emission (green). [77]

Binding studies were undertaken to assess the strength of binding between [2] and calf-thymus (B-DNA) as well as the human telomere sequence (HTS) (quadruplex-DNA).[101] As previously described, dinuclear ruthenium complexes have two chiral centres and therefore exist as three possible stereoisomers, the racemic Δ,Δ and Λ,Λ forms and the meso Δ,Λ diastereoisomer (figure 1.28). In order to study binding affinity, the enantiomerically pure Δ,Δ and Λ,Λ forms of [2] were synthesised. The binding study works on the principle of luminescence enhancement of each stereoisomer upon addition of DNA. The study shows both forms exhibit similar binding affinities to B-DNA as the racemic mixture. However, when bound to quadruplex DNA the Λ,Λ showed two orders of magnitude higher affinity than the Δ,Δ with K_b of 1.16 x 10^7 M^{-1} compared to 2.95 x 10^5 M^{-1}. The luminescent intensity for the HTS saturated Λ,Λ isomer was also 6 fold higher than that of the Δ,Δ isomer. This data indicates that the Λ,Λ isomer binding configuration must have the phenazine nitrogens more shielded from the quenching water molecules as transfer to a hydrophobic environment produces the luminescence enhancement for [2].

Structural data was obtained using NOESY NMR to study the binding of the unresolved mixture of stereoisomers of [1] to the HTS in its antiparallel basket conformation. The obtained NOEs from [1] to the DNA revealed two possible binding sites: end stacking under the diagonal loop or at the lateral loops. Molecular dynamics calculations were used to optimise a structure and it was found that whilst the Λ,Λ molecule could fit under the diagonal loop, the Δ,Δ form could not and bound more weakly at the lateral loops (figure 1.37). This result is consistent with the luminescence enhancement based study on [2] as the Λ,Λ molecule is much less solvent accessible under the diagonal loop.[101]

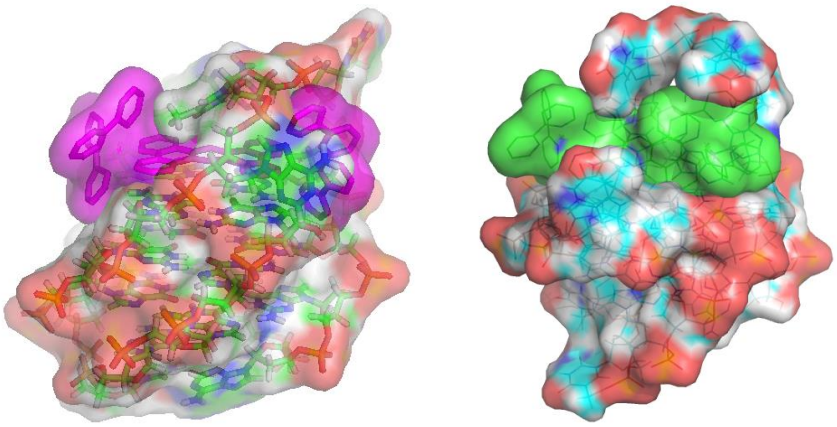


Figure 1.37. Λ,Λ-[1] bound under the diagonal loop (left) and Δ,Δ-[1] bound at the lateral loops (right) images produced by XPLOR. Figure from reference 101. [101]

1.9.2.1 Binding Study to B-DNA

Structural studies on the binding of these complexes to B-DNA are less complete with initial binding studies carried out by viscosity measurements by Rajput, *et al.* It was suggested that the luminescent enhancement of [1] upon binding to calf thymus (CT) DNA was caused by threading intercalation.[99] With addition to calf thymus DNA resulting in a slight increase in viscosity. Later a study by Turro, *et al.* indicated that the observed luminescent enhancement was actually caused by a groove binding event, in which binding into the hydrophobic pocket of a DNA groove was responsible for the light switch effect.[102] They suggested that the viscosity profile shown previously was actually caused by an impurity of the mononuclear complex $[Ru(bipy)_2(tpphz)]^{2+}$. When extensively purified by reverse phase HPLC, the addition of the dinuclear complex to herring sperm DNA showed no increase in viscosity (figure 1.38). However, the Turro

group did not manage to thread the complex, even by annealing CT DNA at 90 °C in the presence of [1]. In the same study, Turro, *et al.* managed to synthesise a threaded form of the dinuclear complex by first intercalating the mononuclear [1] followed by the photochemical coordination of the other ruthenium centre.

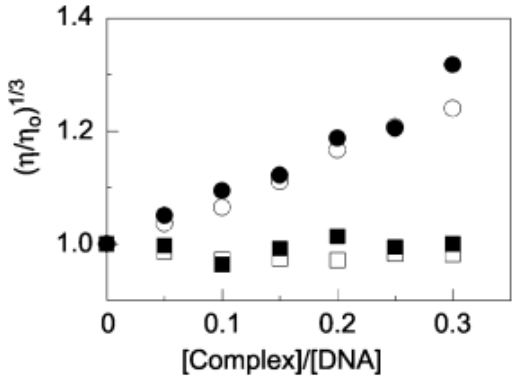


Figure 1.38. Figure from reference 100. A plot of relative viscosity η/η_o at increasing complex to DNA ratios with [1] (white squares), known groove binder hoescht 33258 (black squares), known intercalator ethidium bromide (black circles) and known metallo-intercalator $[Ru(bipy)_2dppz]^{2+}$ (white circles).[102]

1.10 Research Aims

The aim of this research project reported herein, is to delineate a suitable method to resolve all three stereoisomers of each of the nucleic acid binding luminescent ruthenium complexes $[(Ru(bipy)_2)_2tpphz]^{4+}$ [1] and $[(Ru(phen)_2)_2tpphz]^{4+}$ [2] in order to study their interactions with DNA motifs. Although there have been more in depth studies on the binding of the dinuclear ruthenium complexes to quadruplex DNA, there is scope for further study into binding to B-DNA. Thus, another aim of this project is to utilise NMR to provide further structural data on the binding of these complexes to B-DNA in solution with the long-term intention of developing specific probes for DNA motifs within cells.

1.11 References

1 A. McLennan, P. Turner, M. White and A. Bates, *Molecular Biology*, Taylor and Francis, New York, 4th edn., 2005.

2 R. Palchaudhuri and P. J. Hergenrother, *Curr. Opin. Biotechnol.*, 2007, **18**, 497–503.

3 B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter, *Molecular Biology of the Cell*, Garland Science, 2002.

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

5 L. Pauling and R. B. Corey, *Proc. Natl. Acad. Sci. U. S. A.*, 1953, **39**, 84–97.

6 J. M. Vargason, K. Henderson and P. S. Ho, *Proc. Natl. Acad. Sci.*, 2001, **98**, 7265–7270.

7 E.-A. Raiber, R. Hardisty, P. van Delft and S. Balasubramanian, *Nat. Rev. Chem.*, 2017, **1**, 85.

8 B. Alberts, D. Bray, K. Hopkin, A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter, *Essential Cell Biology*, 2009.

9 T. Ohyama, *DNA Conformation and Transcription*, Springer US, Boston, MA, 2005.

10 H. Seligmann, *DNA Replication-Current Advances*, InTech, 2011.

11 Y. Timsit, *J. Mol. Biol.*, 1999, **293**, 835–853.

12 R. E. Dickerson, H. R. Drew, B. N. Conner, R. M. Wing, A. V Fratini and M. L. Kopka, *Science*, 1982, **216**, 475–85.

13 A. Herbert and A. Rich, *Genetica*, 1999, **106**, 1–19.

14 A. Rich and S. Zhang, *Nat. Rev. Genet.*, 2003, **4**, 566–572.

15 L. F. Liu and J. C. Wang, *Proc. Natl. Acad. Sci.*, 1987, **84**, 7024–7027.

16 P. P. Chan and P. M. Glazer, *J. Mol. Med. (Berl).*, 1997, **75**, 267–82.

17 M. Gellert, M. N. Lipsett and D. R. Davies, *Proc. Natl. Acad. Sci.*, 1962, **48**, 2013–2018.

18 S. Neidle and S. Balasubramanian, *Quadruplex Nucleic Acids*, Royal Society of Chemistry, Cambridge, 2006.

19 J. L. Huppert, *Chem. Soc. Rev.*, 2008, **37**, 1375.

20 V. M. Marathias and P. H. Bolton, *Biochemistry*, 1999, **38**, 4355–64.

21 S. Balasubramanian, L. H. Hurley and S. Neidle, *Nat. Rev. Drug Discov.*, 2011, **10**, 261–275.

22 A. Siddiqui-Jain, C. L. Grand, D. J. Bearss and L. H. Hurley, *Proc. Natl. Acad. Sci.*, 2002, **99**, 11593–11598.

23 T. Lemarteleur, D. Gomez, R. Paterski, E. Mandine, P. Mailliet and J.-F. Riou, *Biochem. Biophys. Res. Commun.*, 2004, **323**, 802–808.

24 T.-M. Ou, Y.-J. Lu, C. Zhang, Z.-S. Huang, X.-D. Wang, J.-H. Tan, Y. Chen, D.-L. Ma, K.-Y.

Wong, J. C.-O. Tang, A. S.-C. Chan and L.-Q. Gu, *J. Med. Chem.*, 2007, **50**, 1465–1474.

25 S. Neidle and G. N. Parkinson, *Curr. Opin. Struct. Biol.*, 2003, **13**, 275–283.

26 E. H. Blackburn, *Nature*, 1991, **350**, 569–573.

27 D. Wynford-Thomas and D. Kipling, *Nature*, 1997, **389**, 551.

28 N. W. Kim, M. A. Piatyszek, K. R. Prowse, C. B. Harley, M. D. West, P. L. Ho, G. M. Coviello, W. E. Wright, S. L. Weinrich and J. W. Shay, *Science*, 1994, **266**, 2011–5.

29 A. M. Zahler, J. R. Williamson, T. R. Cech and D. M. Prescott, *Nature*, 1991, **350**, 718–720.

30 J. Li, *Nucleic Acids Res.*, 2005, **33**, 4649–4659.

31 K. N. Luu, A. T. Phan, V. Kuryavyi, L. Lacroix and D. J. Patel, *J. Am. Chem. Soc.*, 2006, **128**, 9963–9970.

32 S. Neidle, *Nat. Rev. Chem.*, 2017, **1**, 41.

33 R. A. Weinberg, *The Biology of Cancer*, 2014.

34 S. Neidle and G. Parkinson, *Nat. Rev. Drug Discov.*, 2002, **1**, 383–393.

35 M. R. Gill and J. A. Thomas, *Chem. Soc. Rev.*, 2012, **41**, 3179.

36 G. J. Bosl and R. J. Motzer, *N. Engl. J. Med.*, 1997, **337**, 242–254.

37 E. R. Jamieson and S. J. Lippard, *Chem. Rev.*, 1999, **99**, 2467–2498.

38 P. M. Takahara, A. Rosenzweig, C., C. Frederick, A., S. Lippard and J., *Nature*, 1995, **377**, 649–652.

39 A. Gelasco and S. J. Lippard, *Biochemistry*, 1998, **37**, 9230–9239.

40 X. Yao, K. Panichpisal, N. Kurtzman and K. Nugent, *Am. J. Med. Sci.*, 2007, **334**, 115–124.

41 H. Deng, *Nucleic Acids Res.*, 2000, **28**, 3379–3385.

42 P. K. Mandal, G. W. Collie, B. Kauffmann and I. Huc, *Angew.Chem.Int.Ed.Engl.*, 2014, **53**, 14424–14427.

43 L. W. Tari and A. S. Secco, *Nucleic Acids Res.*, 1995, **23**, 2065–2073.

44 H. M. Berman, S. Neidle, C. Zimmer and H. Thrum, *Biochim. Biophys. Acta*, 1979, **561**, 124–31.

45 D. S. Goodsell, M. L. Kopka and R. E. Dickerson, *Biochemistry*, 1995, **34**, 4983–4993.

46 C. Bailly and J. B. Chaires, *Bioconjug. Chem.*, 1998, **9**, 513–538.

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

48 R. T. Ranasinghe and T. Brown, *Chem. Commun.*, 2005, 5487.

49 J. B. LePecq and C. Paoletti, *J. Mol. Biol.*, 1967, **27**, 87–106.

50 S. C. Jain and H. M. Sobell, *J. Biomol. Struct. Dyn.*, 1984, **1**, 1179–1194.

51 J. B. Chaires, *Arch. Biochem. Biophys.*, 2006, **453**, 24–29.

52 L. D. Williams, M. Egli, G. Qi, P. Bash, G. A. van der Marel, J. H. van Boom, A. Rich and C. A. Frederick, *Proc. Natl. Acad. Sci. U. S. A.*, 1990, **87**, 2225–9.

53 K. R. Fox and M. J. Waring, *Biochim. Biophys. Acta*, 1984, **802**, 162–8.

54 Y. G. Gao, Y. C. Liaw, H. Robinson and A. H. Wang, *Biochemistry*, 1990, **29**, 10307–10316.

55 D. B. Zamble and S. J. Lippard, *Trends Biochem. Sci.*, 1995, **20**, 435–9.

56 E. Wong and C. M. Giandomenico, *Chem. Rev.*, 1999, **99**, 2451–66.

57 J. K. Barton, A. T. Danishefsky and J. M. Goldberg, *J. Am. Chem. Soc.*, 1984, **106**, 2172–2176.

58 C. Hiort, P. Lincoln and B. Norden, *J. Am. Chem. Soc.*, 1993, **115**, 3448–3454.

59 J. K. Barton, J. M. Goldberg, C. V. Kumar and N. J. Turro, *J. Am. Chem. Soc.*, 1986, **108**, 2081–2088.

60 C. Hiort, B. Norden and A. Rodger, *J. Am. Chem. Soc.*, 1990, **112**, 1971–1982.

61 S. Satyanarayana, J. C. Dabrowiak and J. B. Chaires, *Biochemistry*, 1992, **31**, 9319–9324.

62 R. B. P. Elmes, M. Erby, S. A. Bright, D. C. Williams and T. Gunnlaugsson, *Chem. Commun.*, 2012, **48**, 2588.

63 A. Frei, R. Rubbiani, S. Tubafard, O. Blacque, P. Anstaett, A. Felgenträger, T. Maisch, L. Spiccia and G. Gasser, *J. Med. Chem.*, 2014, **57**, 7280–7292.

64 A. E. Friedman, J. C. Chambron, J. P. Sauvage, N. J. Turro and J. K. Barton, *J. Am. Chem. Soc.*, 1990, **112**, 4960–4962.

65 A. E. Friedman, J. K. Barton, J. C. Chambron, J. P. Sauvage, N. J. Turro and J. K. Barton, *J. Am. Chem. Soc.*, 1990, **112**, 4960–4962.

66 E. R. Batista and R. L. Martin, *J. Phys. Chem. A*, 2005, **109**, 3128–33.

67 C. Mari, V. Pierroz, R. Rubbiani, M. Patra, J. Hess, B. Spingler, L. Oehninger, J. Schur, I. Ott, L. Salassa, S. Ferrari and G. Gasser, *Chem. - A Eur. J.*, 2014, **20**, 14421–14436.

68 C. Hiort, P. Lincoln and B. Norden, *J. Am. Chem. Soc.*, 1993, **115**, 3448–3454.

69 Y. Jenkins, J. K. Barton, A. E. Friedman and N. J. Turro, *Biochemistry*, 1992, **31**, 10809–10816.

70 E. Tuite, P. Lincoln and B. Nordén, *J. Am. Chem. Soc.*, 1997, **119**, 239–240.

71 H. Niyazi, J. P. Hall, K. O'Sullivan, G. Winter, T. Sorensen, J. M. Kelly and C. J. Cardin, *Nat. Chem.*, 2012, **4**, 621–628.

72 J. P. Hall, D. Cook, S. R. Morte, P. McIntyre, K. Buchner, H. Beer, D. J. Cardin, J. A. Brazier, G. Winter, J. M. Kelly and C. J. Cardin, *J. Am. Chem. Soc.*, 2013, **135**, 12652–12659.

73 J. Bolger, A. Gourdon, E. Ishow and J.-P. Launay, *Inorg. Chem.*, 1996, **35**, 2937–2944.

74 S. A. Tysoe, R. Kopelman and D. Schelzig, *Inorg. Chem.*, 1999, **38**, 5196–5197.

75 M. R. Gill, H. Derrat, C. G. W. Smythe, G. Battaglia and J. A. Thomas, *ChemBioChem*, 2011, **12**, 877–880.

76 M. R. Gill, P. J. Jarman, S. Halder, M. G. Walker, H. K. Saeed, J. A. Thomas, C. Smythe, K. Ramadan and K. A. Vallis, *Chem. Sci.*, 2018, **9**, 841–849.

77 M. R. Gill, J. Garcia-Lara, S. J. Foster, C. Smythe, G. Battaglia and J. A. Thomas, *Nat. Chem.*, 2009, **1**, 662–667.

78 D. R. Boer, L. Wu, P. Lincoln and M. Coll, *Angew. Chemie Int. Ed.*, 2014, **53**, 1949–1952.

79 B. Önfelt, P. Lincoln and B. Nordén, *J. Am. Chem. Soc.*, 1999, **121**, 10846–10847.

80 L. M. Wilhelmsson, F. Westerlund, P. Lincoln and B. Nordén, *J. Am. Chem. Soc.*, 2002, **124**, 12092–12093.

81 P. Lincoln and B. Nordén, *Chem. Commun.*, 1996, **0**, 2145–2146.

82 Rodger A. and B. Norden, *Circular Dichroism and Linear Dichroism*, Oxford University Press, 1997.

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

84 B. Nordén and T. Kurucsev, *J. Mol. Recognit.*, 1994, **7**, 141–155.

85 P. Nordell and P. Lincoln, *J. Am. Chem. Soc.*, 2005, **127**, 9670–9671.

86 F. Westerlund, M. P. Eng, M. U. Winters and P. Lincoln, *J. Phys. Chem. B*, 2007, **111**, 310–317.

87 L. M. Wilhelmsson, E. K. Esborner, F. Westerlund, B. Norden and P. Lincoln, *J. Phys. Chem. B*, 2003, **107**, 11784–11793.

88 A. A. Almaqwashi, J. Andersson, P. Lincoln, I. Rouzina, F. Westerlund and M. C. Williams, *Biophys. J.*, 2016, **110**, 1255–1263.

89 F. Westerlund, P. Nordell, J. Blechinger, T. M. Santos, B. Nordén and P. Lincoln, *J. Phys. Chem. B*, 2008, **112**, 6688–6694.

90 C. M. Dupureur and J. K. Barton, *Inorg. Chem.*, 1997, **36**, 33–43.

91 A. A. Almaqwashi, T. Paramanathan, P. Lincoln, I. Rouzina, F. Westerlund and M. C. Williams, *Nucleic Acids Res.*, 2014, **42**, 11634–11641.

92 T. Paramanathan, F. Westerlund, M. J. McCauley, I. Rouzina, P. Lincoln and M. C. Williams, *J. Am. Chem. Soc.*, 2008, **130**, 3752–3753.

93 J. Andersson, M. Li and P. Lincoln, *Chem. - A Eur. J.*, 2010, **16**, 11037–11046.

94 P. Nordell, F. Westerlund, A. Reymer, A. H. El-Sagheer, T. Brown, B. Nordén and P. Lincoln, *J. Am. Chem. Soc.*, 2008, **130**, 14651–14658.

95 F. Westerlund and P. Lincoln, *Biophys. Chem.*, 2007, **129**, 11–17.

96 L. Wu, A. Reymer, C. Persson, K. Kazimierczuk, T. Brown, P. Lincoln, B. Nordén and M. Billeter, *Chem. - A Eur. J.*, 2013, **19**, 5401–5410.

97 J. Andersson and P. Lincoln, *J. Phys. Chem. B*, 2011, **115**, 14768–14775.

98 J. Bolger, A. Gourdon, E. Ishow and J.-P. Launay, *J. Chem. Soc., Chem. Commun.*, 1995, 1799–1800.

99 C. Rajput, R. Rutkaite, L. Swanson, I. Haq and J. A. Thomas, *Chem. - A Eur. J.*, 2006, **12**, 4611–4619.

100 T. Wilson, M. P. Williamson and J. A. Thomas, *Org. Biomol. Chem.*, 2010, **8**, 2617.

101 T. Wilson, P. J. Costa, V. Félix, M. P. Williamson and J. A. Thomas, *J. Med. Chem.*, 2013, **56**, 8674–8683.

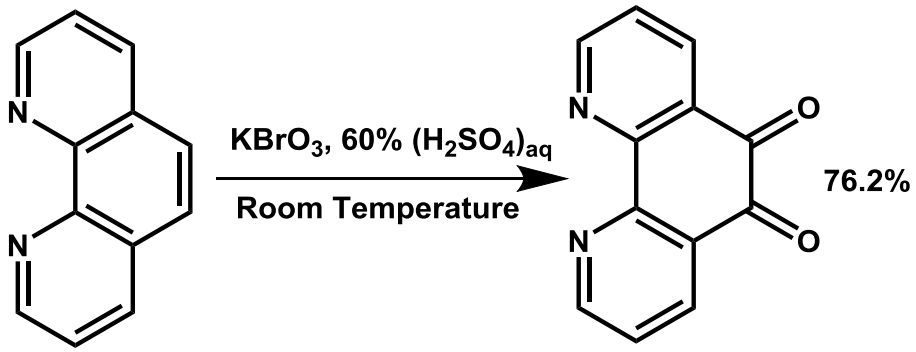
102 D. A. Lutterman, A. Chouai, Y. Liu, Y. Sun, C. D. Stewart, K. R. Dunbar and C. Turro, *J. Am. Chem. Soc.*, 2008, **130**, 1163–1170.

Chapter 2. Synthesis

To study the DNA binding properties of the optical isomers of $[(Ru(bipy)_2)_2tpphz][Cl]_4$ [1] and $[(Ru(phen)_2)_2tpphz][Cl]_4$ [2], the complexes needed to be synthesised and completely resolved. Ancillary bipy and phen ligands were purchased from suppliers and the general method for the synthesis of [1] and [2] was that employed by Bolger *et al.*[1]

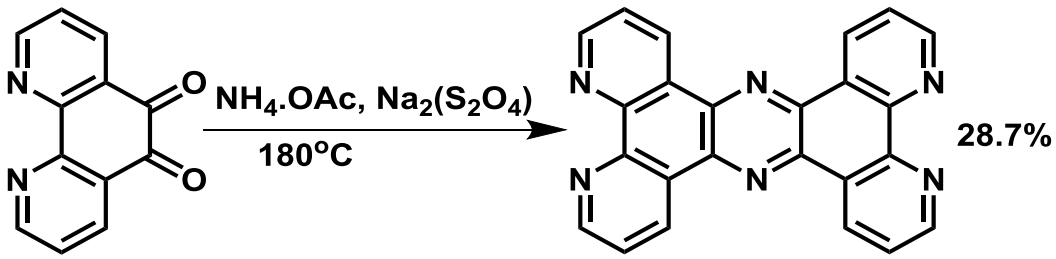
2.1 Precursor Synthesis

Synthesis of the bridging tpphz ligand was carried out in two steps. Firstly the 1,10-phenanthroline (phen) was oxidised to 1,10-phenthroline-5,6-dione (dpq) using the method outlined by Zheng, *et al.* employing potassium bromate and sulphuric acid (60%) as an oxidant as shown in scheme 2.1.[2]



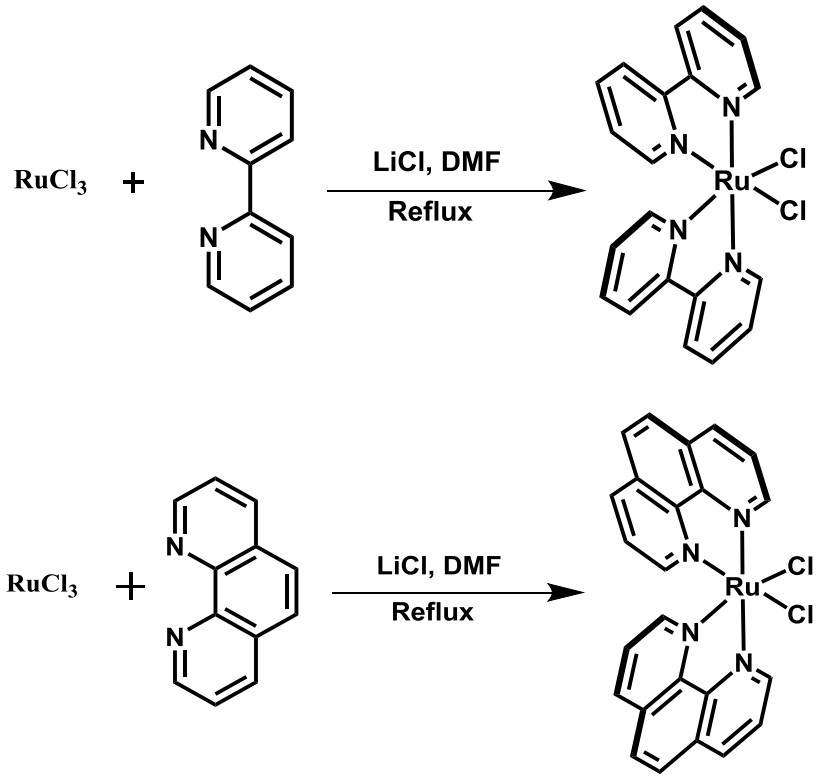
Scheme 2.1. The synthesis of dpq.

The quinone was then converted into tpphz by reacting with ammonium acetate by a method described by Bolger, *et al.* illustrated by scheme 2.2.[1]



Scheme 2.2. The synthesis of tpphz from dpq.

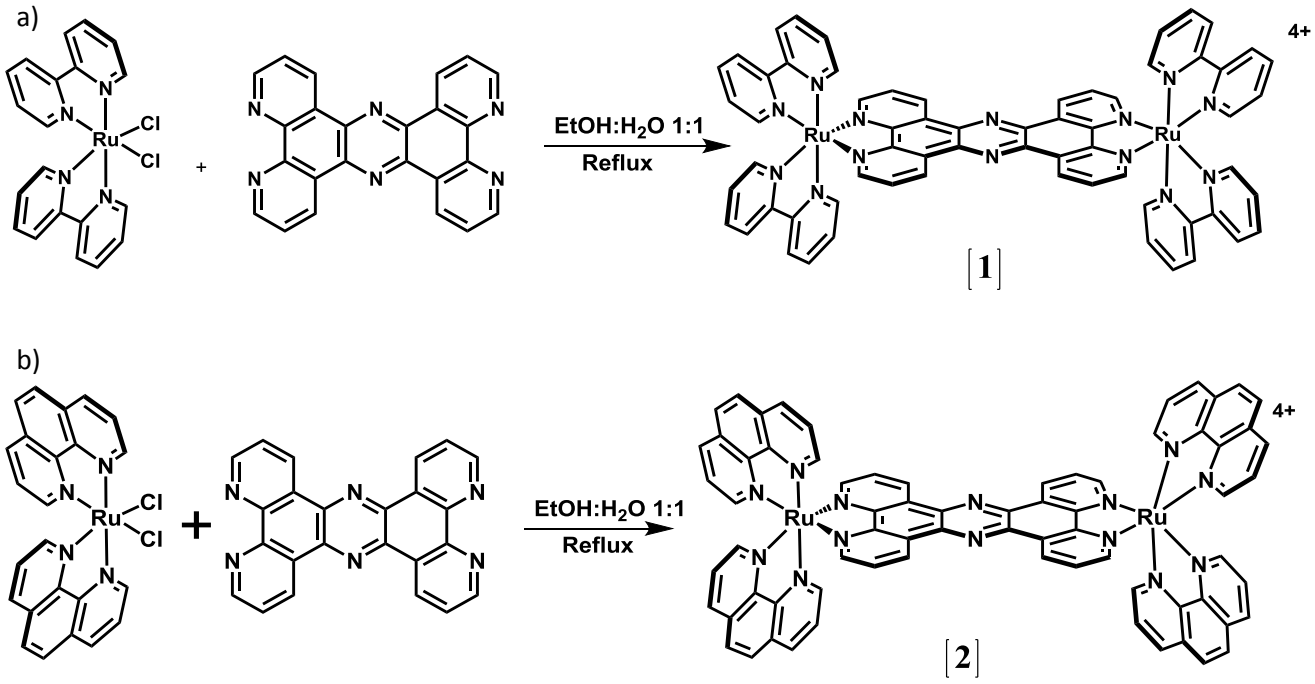
For the ruthenium dimers [1] and [2] to be synthesised, the precursor complexes $Ru(bipy)_2Cl_2.2H_2O$ and $Ru(phen)_2Cl_2.2H_2O$ were first synthesised by the method of Sullivan, *et al.*[3] The coordination of bipy and phen to $RuCl_3.3H_2O$ was carried out in the presence of LiCl to avoid the formation of $Ru(bipy)_3$ and $Ru(phen)_3$ and to leave chloride ions coordinated for eventual substitution with tpphz. The precursor synthesis is shown in scheme 2.3.



Scheme 2.3. The synthesis of mononuclear ruthenium precursor complexes $Ru(bipy)_2Cl_2.2H_2O$ and $Ru(phen)_2Cl_2.2H_2O$.

2.2 Synthesis of Mixed [1] and [2]

The coordination of the two ruthenium centres was successfully achieved using the method described by Bolger, *et al.* via the coordination of the ditopic tpphz ligand.[1] The reaction conditions are detailed below in scheme 2.4.



Scheme 2.4. The non-stereospecific synthesis of dinuclear ruthenium complexes a) $[(Ru(bipy)_2)_2tpphz]^{4+}$ [1] and b) $[(Ru(phen)_2)_2tpphz]^{4+}$ [2].

As previously discussed the octahedral ruthenium complexes with bidentate ligands exist as two stereoisomers; Λ or Δ. The reaction detailed in scheme 2.4 is not chirally selective and therefore yielded a statistical mixture, 2:1:1, of the three stereoisomers Λ,Δ: Λ,Λ: Δ,Δ. Therefore, a method of separation of the stereoisomers is required.

2.2.1 Resolution of Optical Isomers by Crystallisation

In order to probe the binding of chiral molecules to DNA, resolution of stereoisomers was required. Conventional methods of the resolution of ruthenium polypyridyl complexes are that of co-crystallisation with chiral tartrate salts. In these systems, salts of enantiomerically pure d or l tartrate selectively crystallize out of aqueous solution with the corresponding enantiomer of the ruthenium polypyridyl complex by ion pairing. This chiral recognition is due to favourable crystal packing in the diastereomeric salt as demonstrated in figure 2.1.[4]

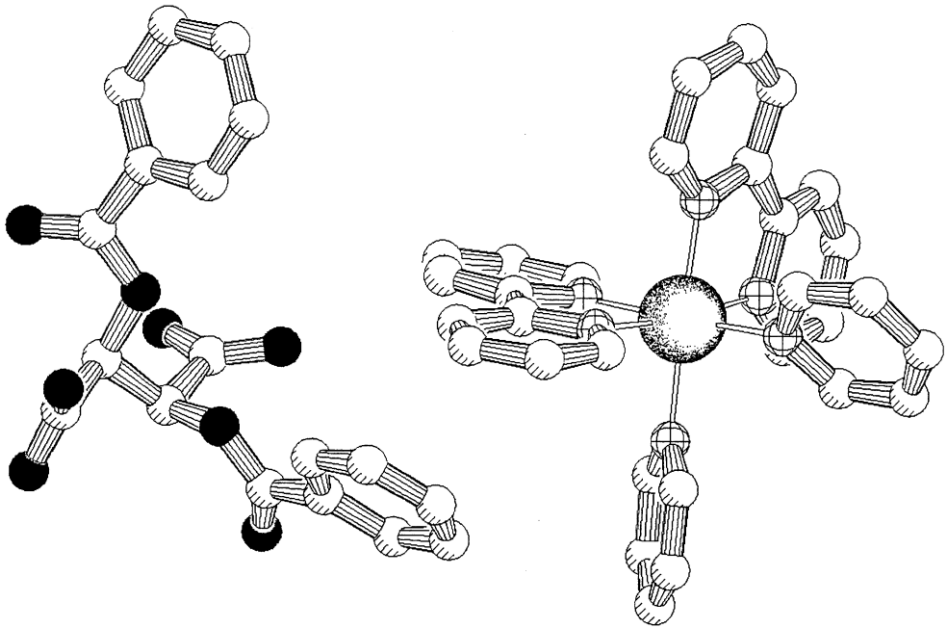
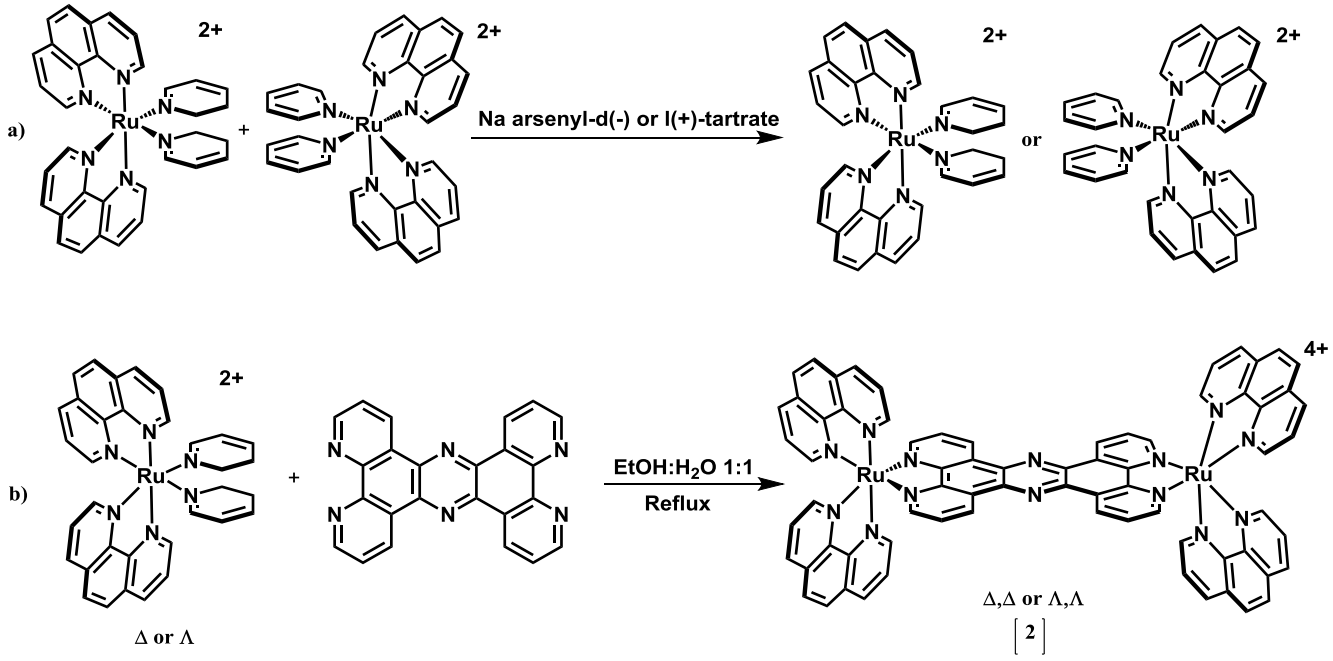


Figure 2.1. X-ray crystallographic structure of a Δ-[Ru(bipy)$_2$(py)$_2$][(+)-O,O'-dibenzoyl-d-tartrate] ion pair.[4]

Previously, this method of resolution has been effectively applied to mononuclear ruthenium complexes containing one chiral centre.[5,6] However, attempts to resolve complexes containing multiple optical isomers has been less successful, and as a result enantiomerically pure dinuclear ruthenium complexes must be synthesised from chirally pure mononuclear precursors. In order to carry out this resolution, a suitable mononuclear precursor is needed, as Ru(bipy)$_2$Cl$_2$ or Ru(phen)$_2$Cl$_2$ cannot be resolved due to the fluxional re-racemisation of complexes. It has been previously shown that the bis-pyridine ruthenium complex

$[Ru(phen)_2(py)_2]^{2+}$ can be resolved and, as long as it is not exposed to light, retains its stereochemistry.[6] Therefore, using this method a stereoselective synthesis was proposed, whereby the $[Ru(phen)_2(py)_2]^{2+}$ would be resolved and used as a chiral building block, with the pyridine ligands substituted by tpphz in a symmetry retained reaction; a methodology first described by Hua and Von Zelewsky.[7] (Scheme 2.5).



Scheme 2.5. A two-step stereospecific synthesis of $[(Ru(phen)_2)_2tpphz]^{4+}$ [**2**].

After several co-crystallisations from aqueous solution with sodium arsenyl-d(-)-tartrate the Δ-$[Ru(phen)_2(py)_2]^{2+}$ was considered to have been suitably enriched as monitored by circular dichroism (CD). Circular dichroism monitors the differential absorption of left handed and right handed circularly polarized light by chiral molecules. Therefore, if a metal complex has been resolved it will show a CD signal as seen in (figure 2.2). Although CD provides a good means of monitoring resolution, without a reference sample it is impossible to know the enantiopurity of the resolved sample.[8] Other methods of determining enantiopurity include the use of NMR chiral lanthanide shift reagents.[9] Chiral lanthanide shift reagents are used in most applications for neutral organic molecules and are less appropriate in determining enantiomeric excess of chiral cations.[10] However, chiral lanthanide shift reagents were not employed in this

study as they have proven to be both expensive and complicated, and sometimes unreliable for the method of chiral discrimination.[11]

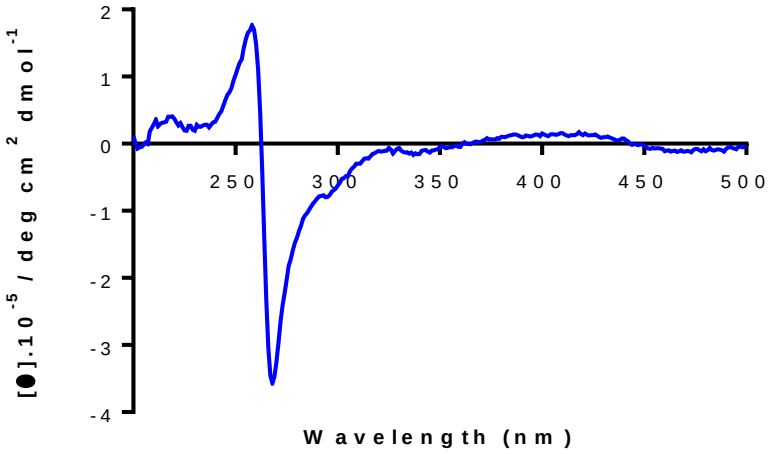


Figure 2.2. Normalised CD Spectrum of Δ-[Ru(phen)$_2$ (Py)$_2$].2[Cl] in H_2O enriched by co-crystalisation with sodium aresenyl-(d)-tartate

Upon the coordination reaction to form Δ,Δ-[2] it was seen that conservation of the Δ configuration was not complete as the CD signal was lost. Although great care was taken to keep the reaction under reduced light, the long and harsh reaction conditions, 4 days in refluxing water:ethanol 1:1, needed for substitution of the pyridine units meant that it was possible that some photo induced racemisation had occurred. $[Ru(phen)_2(py)_2]^{2+}$ has been shown to be very photo-sensitive, with pyridine dissociation occurring in 10 minutes under sunlight.[6] This photo-induced pyridine disassociation has been utilised in the synthesis of photoactivated anti-cancer agents, as dissociation and substitution with a solvent molecule can provide vacant coordination sites for binding to DNA bases.[12]

With this method of stereoselective synthesis producing sub-optimal results, an alternative method of chiral resolution was proposed using chromatographic procedures.

2.3 Resolution of Optical Isomers by Column Chromatography

In studies by Keene, *et al.* chromatographic methods have been employed in the separation of polypyridyl complexes containing multiple ruthenium centres.[13–15] Utilisation of this method is advantageous as compounds can be synthesised in a non-stereospecific manner prior to separation into the optical isomers. Unlike the co-crystallisation method, this uses mild conditions and it also provides a method to synthesise and isolate the Λ,Δ stereoisomers which cannot be produced via the stereoselective synthesis, due to the reaction of both Λ and Δ forms of $[Ru(phen)_2(py)_2]^{2+}$ yielding the statistical mixture of optical isomers shown below.

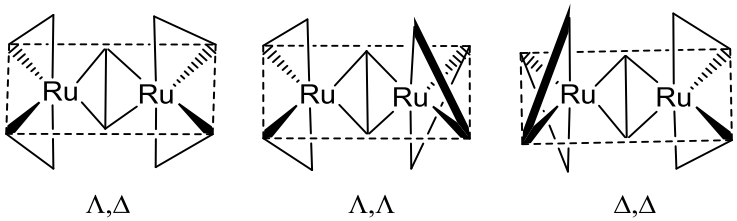


Figure 2.3. Simplified chemical structures of dinuclear ruthenium complexes demonstrating the difference in isomers.

The technique employs the use of a SP-sephadex C-25 cation exchange resin as the stationary phase. SP-sephadex consists of negative propylsulphonate functional groups attached to a cross linked dextran support. Adsorbed cationic ruthenium polypyridyl complexes can be separated based on the differential equilibria between cations in the eluent, such as sodium, with lower charged complexes eluting more rapidly. Keene, *et al.* discovered that although the charge of the cation in the eluent mixture plays an important role in the dissociation of the complexes, the anions also have a significant effect on elution.[16] It was noted that this could be exploited in the separation of ruthenium polypyridyl complexes.[16] They postulated that if an eluent anion associates more strongly to one diastereoisomer or enantiomer of the ruthenium complex, it could effectively lower the charge of that species and aid in its separation. Further study into the separation of the isomers of the ruthenium complex, $[Ru(dmbpy)_2(bipym)]$, concluded that the separation occurs as a result of a cation-anion 'host-guest' type interaction. This interaction was probed through ^1H NMR and the change in

chemical shifts of specific protons were monitored upon the addition of the guest anion to the host ruthenium complex. Using this method it was established that the binding of organic anions was close to the central hydrophobic cavity. The accessibility of the central cavity is dependent on the stereoisomer present as shown in figure 2.4. Therefore, this means that the anions can distinguish between meso and rac stereoisomers.

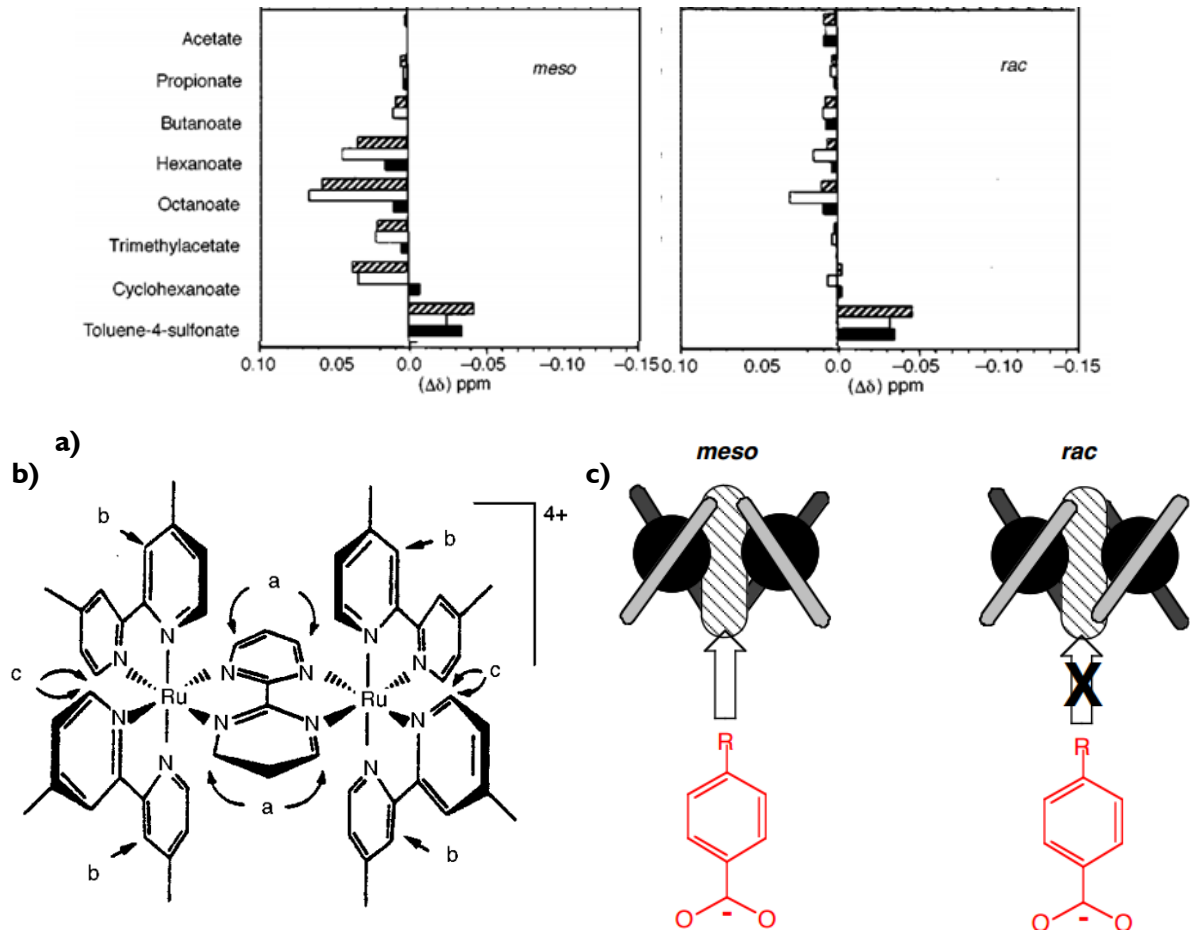


Figure 2.4. **a)** The difference in ^{1}H-NMR chemical shift of specific protons of [Ru(dmbpy)$_2$(bipym)] dmbpy (4,4'-dimethyl-2,2'-bipyridine) bipym (2,2'-bipyrimidine) upon addition of several anion sodium salts to ruthenium complex.[14] Proton a (hashed), proton b (open), proton c (filled). **b)** Chemical structure of [Ru(dmbpy)$_2$(bipym)]. **c**) Illustration of the meso and racemic complexes in association with an eluent anion. Adapted from reference 14. [14]

This effect has been attributed to the chiral anion second sphere association coupled with the inherent chirality of the dextran matrix. This has been used as a powerful tool in the resolution of several ruthenium dimeric species.[13]

2.3.1 Separation of the Diastereoisomers of [1]

Separation of mixed [1] (p43 scheme 2.4) into the corresponding optical isomers was carried out by the method outlined above. Several different eluent mixtures, including sodium toluene-4-sulphonate, were used in an attempt to resolve the complexes and eventually sodium octanoate eluent mixture provided an effective separation. The chromatographic technique itself was found to be very sensitive to flow rate, solution pH and eluent concentration. The optimisation of the conditions proved an arduous task. After many attempted columns, a method was achieved giving an effective separation, which was clearly distinguishable as a result of the red colour.

The compound [1] was loaded onto a metre-long column containing SP-sephadex and eluted using an aqueous solution of 0.15 M sodium octanoate at pH 10.5. The column was set up to recirculate at 0.75 ml min^{-1} rendering the column length effectively limitless. After four passes of the column a separation into two distinct bands was observed. The first eluted band was found to consist of the meso complex (Λ,Δ), followed by the racemic mixture (rac) in the second band consisting of the chemically identical Λ,Λ and Δ,Δ complexes (figure 2.5). This was later confirmed by X-ray crystallography.

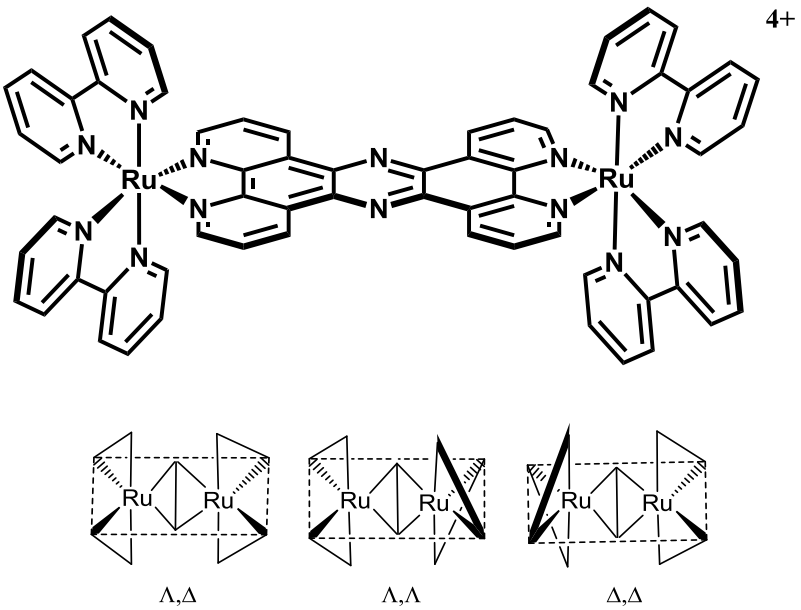


Figure 2.5. Chemical structure of $[(Ru(bipy)_2)_2tpphz]^{4+}$ [1].

The meso complex eluted first due to less steric hinderance upon association of the aliphatic octanoate eluent ion into the cavity through both favourable electrostatic and hydrophobic effects. As the anion associates more closely it lowers the overall charge of the complex enabling the meso isomer to elute through the column more quickly. Interestingly, even though chemically different, the meso and racemic fractions showed identical NMR spectra. This differs from the complex resolved by Keene, *et al.* in which the NMR spectra of the two systems differed. The bipyrimidine bridging ligand employed in that study is considerably shorter than the tpphz ligand. Therefore, in [1] the two ruthenium centres are sufficiently separated so as to exclude any influences of ring current shifts from the opposing bipyridine ligands.

2.3.2 Resolution of Racemic [1]

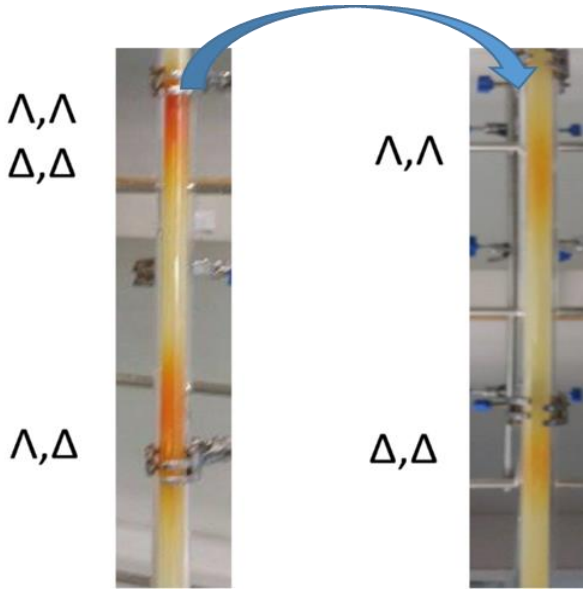


Figure 2.6. The chromatic separation of [1] showing the two SP sephadex C-25 columns with the left column showing the separation of the racemic and meso bands using sodium octanoate eluent and the right column the separation of the Λ,Λ and Δ,Δ fractions using sodium dibenzoyl-L-tartrate eluent mixture.

The second eluted fraction consisting of the racemic mixture of Λ,Λ and Δ,Δ then needed to be separated. Thus, the fraction was returned to the column and separated using a chiral eluent mixture consisting of 0.05 M sodium dibenzoyl-(l)-tartrate at pH 10.5. After several passes of the column the two enantiomers separated, with the Δ,Δ isomer comprising the first fraction, and the Λ,Λ isomer as the second. CD was used as a confirmation of the resolution, as each ruthenium

centre should rotate a plane of polarized light by an equal and opposite amount. This means that the Δ,Δ and Λ,Λ isomers have opposite CD spectra and the Λ,Δ has a CD of zero as the Λ and Δ signals cancel each other out. The CD spectra of the three resolved optical isomers is shown in figure 2.7.

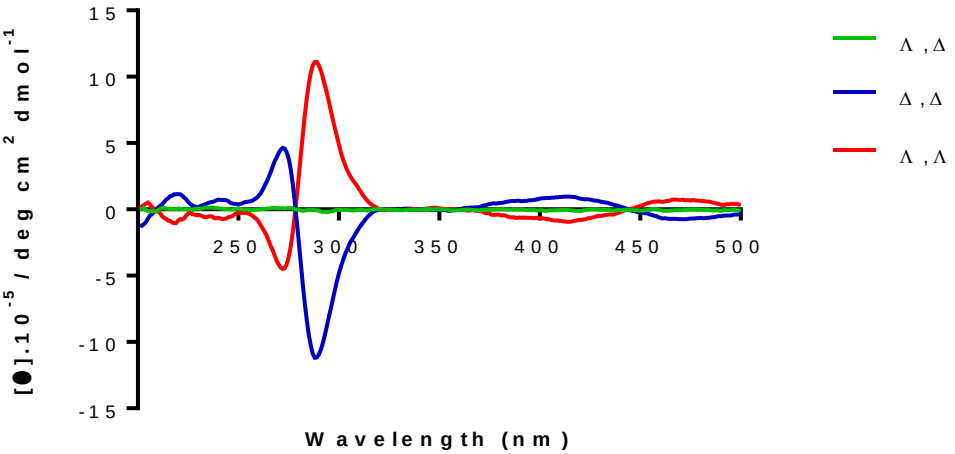


Figure 2.7. Normalised CD Spectrum of Δ,Δ-[1] (blue) Λ,Λ-[1] (red) and Λ,Δ-[1] (green) in H_2O.

2.3.3 Resolution of Racemic [2]

After the successful resolution of [1] had been carried out, the same method was applied to the resolve the mixture of the related phenanthroline complex [2] (p43 scheme 2.4) under the same conditions and separation was successfully accomplished. However, it was noted that the racemic band showed significant spreading. Therefore, this band was left to run on the column for a longer period of time with only the initial non chiral 0.15 M sodium octanoate eluent added. This indeed resulted in separation of the racemic band into two fractions without the addition of any chiral eluent and was confirmed by CD. This surprising feature of chiral separation on sephadex has been reported previously in the resolutions performed by Keene, *et al.* and is due to the selectivity of the chiral sephadex support alone. However, the separation was improved by the addition of a chiral eluent mixture consisting of 0.05 M sodium dibenzoyl-(l)-tartrate at pH 10.5. The column separation conditions were eventually optimised so that all three

stereoisomers of [2] could be separated on one column by employing 0.05 M sodium dibenzoyl-(l)-tartrate at pH 10.5 as an eluent mixture. The Δ,Δ was the first to be eluted followed by Λ,Λ and the meso stereoisomer Λ,Δ was the final eluted complex. The CD spectrum of the three fractions is shown below in (figure 2.8).

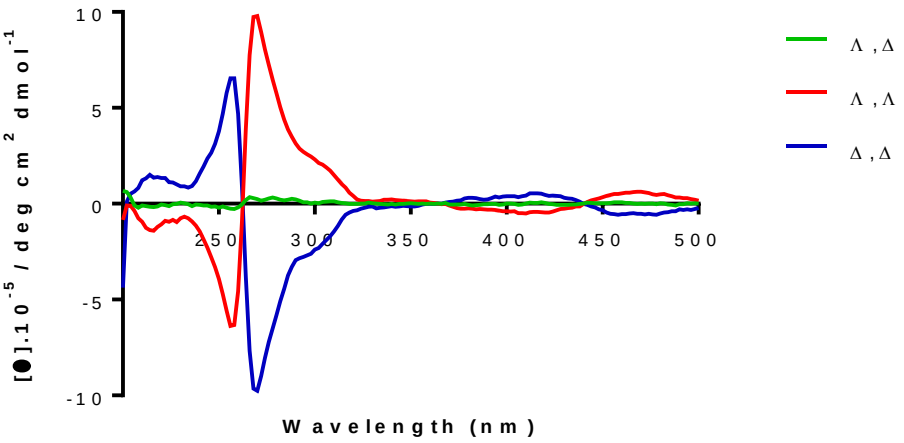


Figure 2.8. Normalised CD Spectrum of Δ,Δ-[2](blue), Λ,Λ-[2](red) and Λ,Δ-[2](green) in H_2O

2.4 X-Ray Crystallography

As previously discussed, the synthesis of [1] was originally reported by Bolger, *et al.* This report also described an X-ray crystallographic structure grown from the unresolved mixture of optical isomers as their nitrate salts (figure 2.9). The resulting crystals solely contained the Δ,Δ and Λ,Λ forms co-crystallized together.[17] It was speculated that this occurred because Λ,Δ-[1] was more soluble than the other two forms.

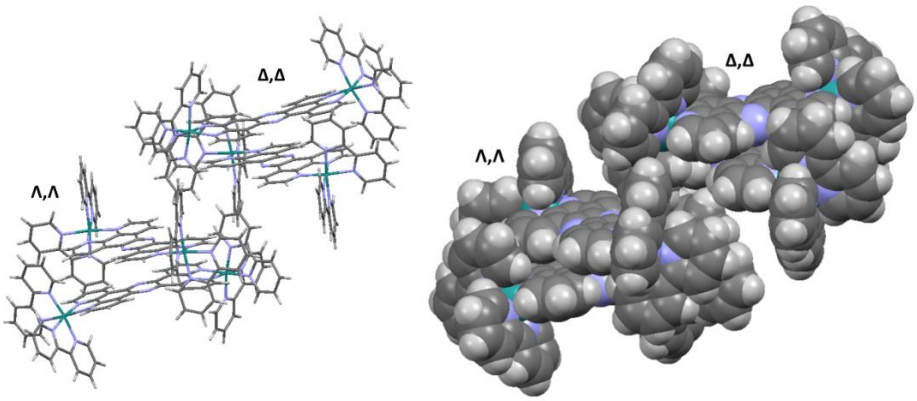


Figure 2.9. The crystal structure of co-crystallised Λ,Λ-[1] and Δ,Δ-[1](right) space filling form.[17] CCDC Deposition Number: 1310877

As X-ray quality crystals of metal complexes are typically grown from an unresolved mixture, only a few examples of crystal growth from pre-resolved complexes exist.[15] Thus, crystal growth from the pre-resolved metal complexes was attempted. Crystal structures listed in this section were solved by Dr. Craig Robertson at the Department of Chemistry X-ray Crystallography Service.

2.4.1 X-Ray Crystallographic Studies of Λ,Δ-[1]

Crystals of Λ,Δ-[1] suitable for X-ray structure determination were successfully obtained by vapour diffusion of acetone into a methanolic solution of the chloride salt. A summary of the crystallographic data can be found in the appendix. The crystal structure is shown in figure 2.10 below and is the first reported for this stereoisomer.

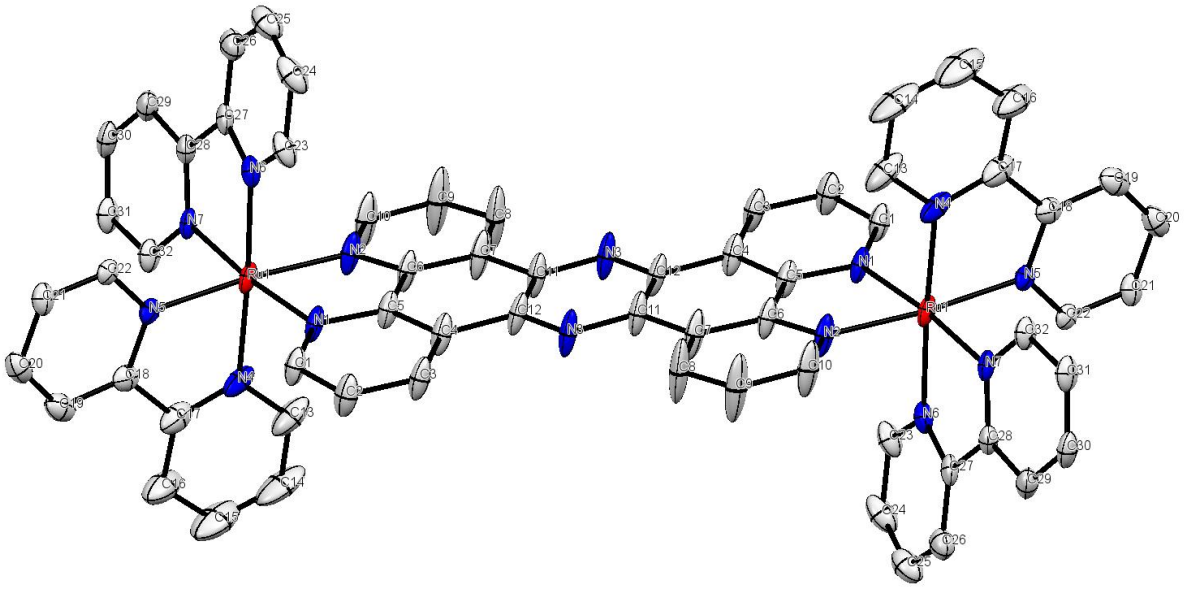


Figure 2.10. ORTEP plot of the X-ray crystallographic structure of Λ,Δ-[1] with thermal ellipsoids indicating 50% probability.

The asymmetric unit consists of half of the molecule along the bridging tpphz moiety, applying inversion generated the other half of the molecule in the full unit cell as shown in figure 2.11.

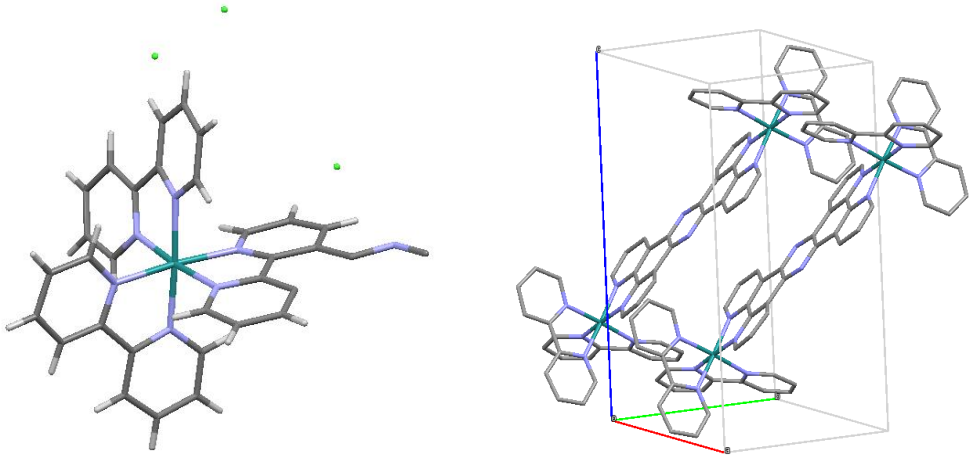


Figure 2.11. Left: The asymmetric unit and right: the full unit cell packing diagram of the triclinic Λ,Δ-[1] with anions and hydrogen atoms omitted for clarity.

The crystals form a triclinic system with the space group P1. The geometry of the $(bipy)_2Ru(tpphz)$ stereocentre is very similar to that found in $[Ru(bipy)_3]^{2+}$.[18] The ruthenium-nitrogen bond lengths are 2.045-2.68 Å and bite angles are 78.58-79.2 °. The extended polyaromatic tpphz bridging ligand is almost completely planar. In contrast to the previously reported structure for the co-crystallised Λ,Λ and Δ,Δ isomers, in which dinuclear cations are orthogonally stacked over each other (figure 2.9), the cations in the Λ,Δ-[**1**].4Cl structure are arranged in tightly packed offset coplanar stacks. This results in the formation of layers with the chloride counter ions in channels between them (figure 2.12). The stacks have tpphz ligands in close contact with a parallel-displaced π-stacking mode. Further crystallographic data is shown in the appendix.

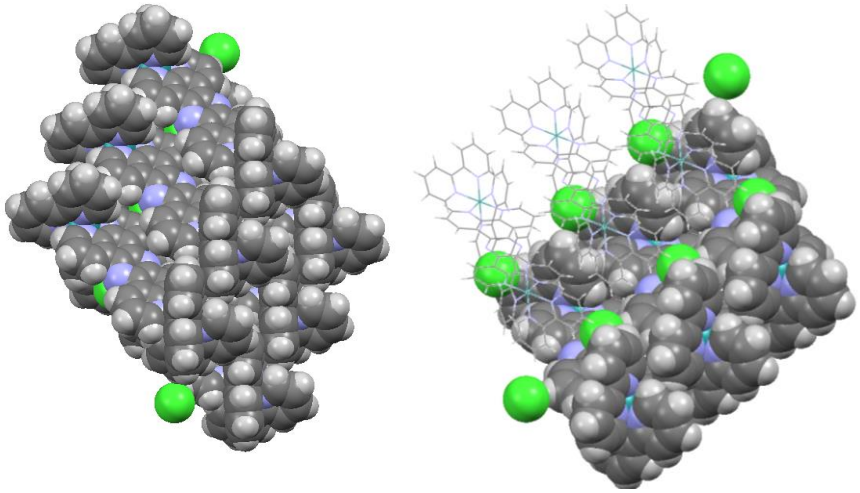


Figure 2.12. Crystallographic structure of Λ,Δ-[1].4Cl showing packing with stacks of Λ,Δ-[1] with chloride ions shown in between.

2.4.2 X-Ray Crystallographic Studies of Δ,Δ-[1]

Crystals of Δ,Δ-[1].4PF_6 suitable for X-ray structure determination were successfully obtained by vapour diffusion of diethyl ether into a nitromethane solution containing Δ,Δ-[1].4PF_6. A summary of the crystallographic data can be found in the appendix. The crystal structure shown in figure 2.13 below is the first reported of this stereoisomer grown from a pre-resolved solution and not co-crystallised with its enantiomer as previously reported.[17]

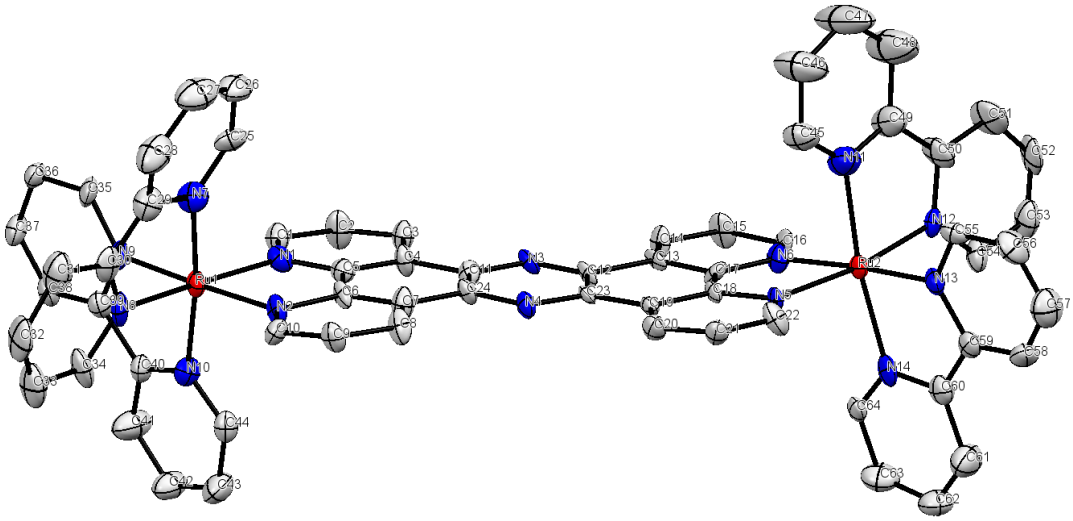


Figure 2.13. ORTEP plot of the X-ray crystallographic structure of Δ,Δ-[1] with thermal ellipsoids indicating 50% probability.

The asymmetric unit contains a full Δ,Δ-[1] complex (figure 2.14). The crystals form a monoclinic system with the space group $P2_1$. The flack parameter is obtained by comparing the crystal to a hypothetic crystal with an applied inversion. The closer the value is to zero, the higher the enantiopurity. The flack parameter here is quite low with a value of 0.086(17) which suggests that the complex is resolved and is the Δ,Δ enantiomer.[19]

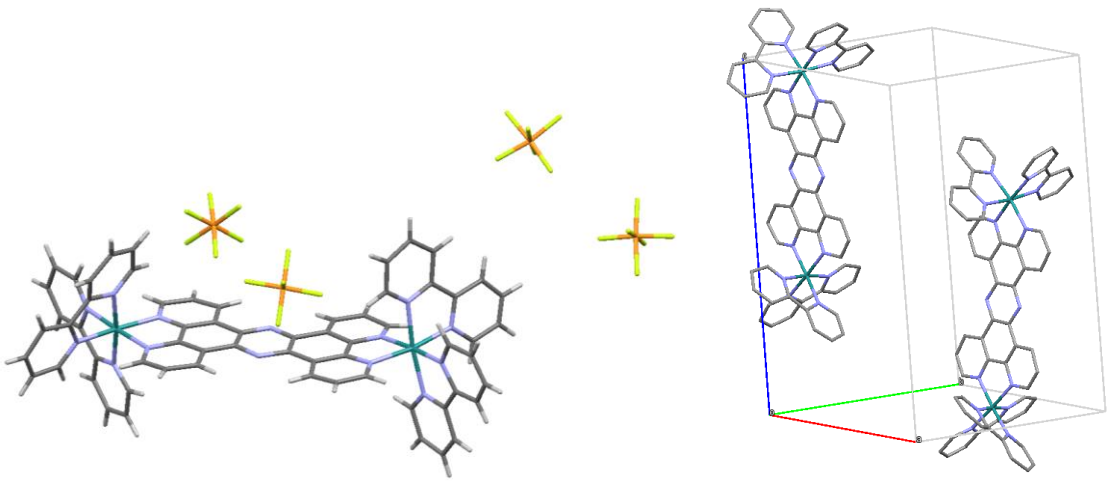


Figure 2.14. Left: The asymmetric unit and right: the full unit cell packing diagram of the monoclinic PF_6 salt of Δ,Δ-[1].4PF_6 with solvent molecules and hydrogen atoms omitted for clarity.

The geometry of the Ru(bipy)$_2$(tpphz) stereocentre is very similar to that found in the previously solved Δ,Δ-[1], [17] with ruthenium-nitrogen bond lengths of 2.038-2.113 Å and bite angles of 77.89-80.12 °. The tpphz bridging ligand is far more planar than that seen in the previously reported structure by Bolger, *et al.* As with the Λ,Δ-[1] structure, the Δ,Δ-[1] complexes are arranged in offset coplanar stacks, however they are far less tightly packed due to the larger PF_6 anions forcing the stacks further apart (figure 2.15).

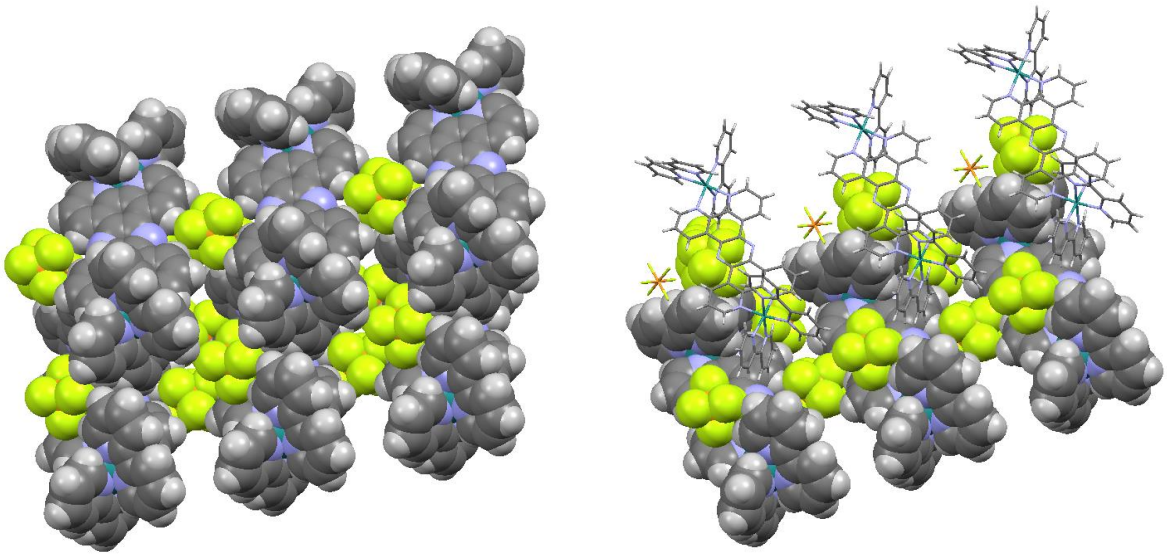


Figure 2.15. Crystallographic structure of Δ,Δ-[1].4PF_6 showing packing with stacks of Δ,Δ-[1] with PF_6 ions sitting in the channels.

2.4.3 X-ray Crystallographic Studies of Λ,Λ-[1]

Crystals of Λ,Λ-[1].4PF_6 suitable for X-ray structure determination were grown by the same method as that of its enantiomer, by vapour diffusion of diethyl ether into a nitromethane solution containing Λ,Λ-[1].4PF_6. A summary of the crystallographic data can be found in the appendix. The crystal structure is shown in figure 2.16 and is the first reported for this stereoisomer grown from a pre-resolved solution.

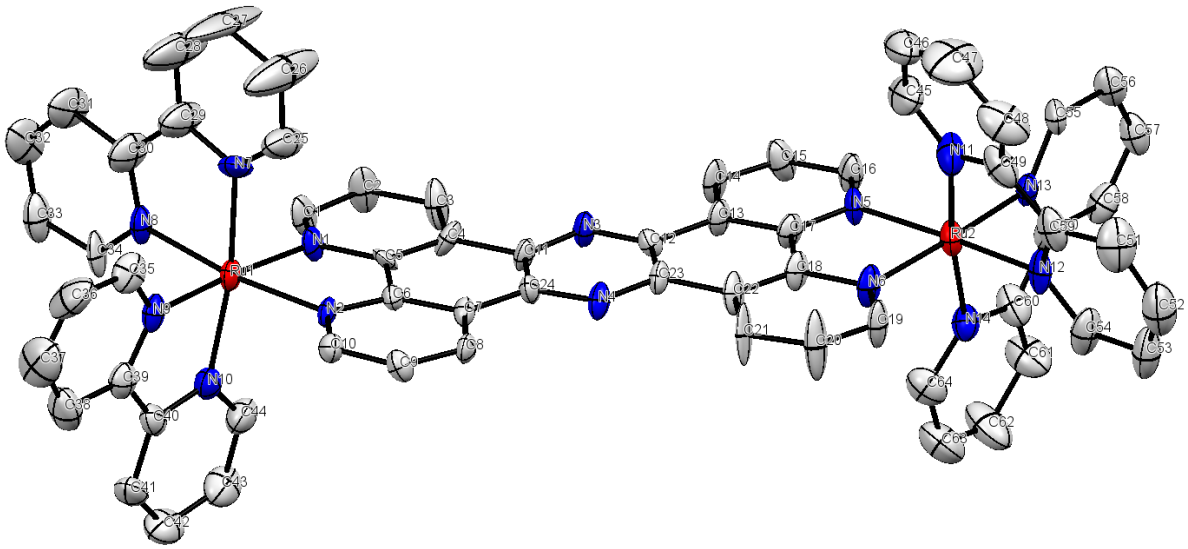


Figure 2.16. ORTEP plot of the X-ray crystallographic structure Λ,Λ-[1] with thermal ellipsoids indicating 50% probability.

The asymmetric unit contains a full Λ,Λ-[1] complex (figure 2.17). The crystals form the monoclinic space group $P2_1$. The low flack parameter of 0.065(9) indicates that the crystal is the correct enantiomer and of high enantiopurity.

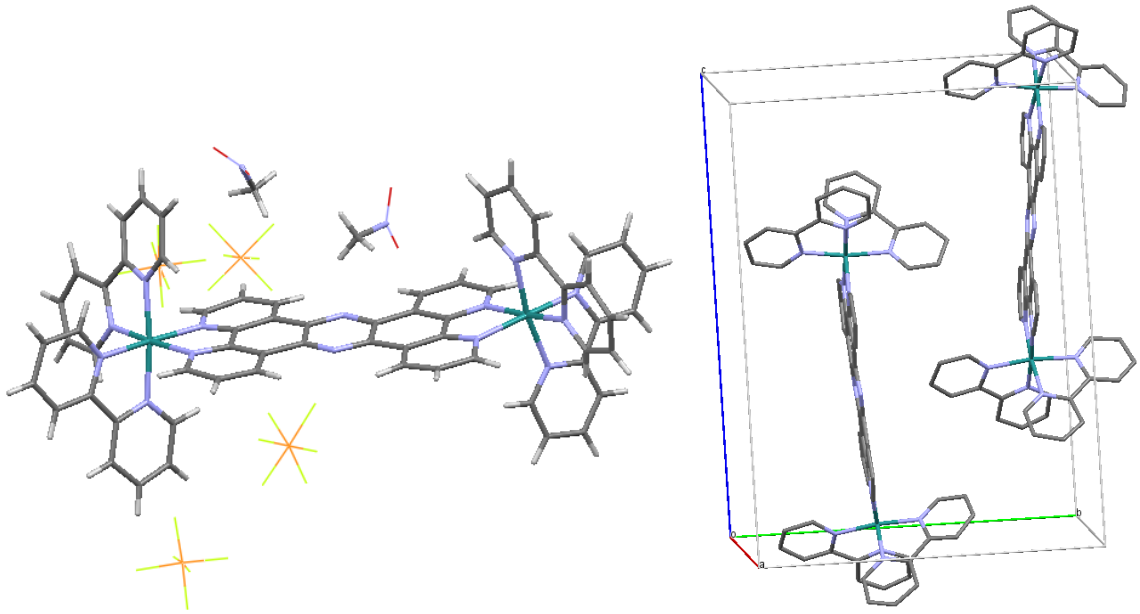


Figure 2.17. Left: The asymmetric unit and right: the full unit cell packing diagram of the monoclinic PF_6 salt of Λ,Λ-[1].4PF_6 with solvent molecules and hydrogen atoms omitted for clarity.

The structure is the inverted form of that seen for the Δ,Δ-[1] species. Thus, the geometry of the Ru(bipy)$_2$(tpphz) stereocentre is very similar to that found in Δ,Δ-[1], with ruthenium-nitrogen bond lengths of 2.035-2.093 Å and bite angles of 78.36-79.91 °. The structures are arranged in offset coplanar stacks (figure 2.18), again as observed with the Δ,Δ-[1] complex.

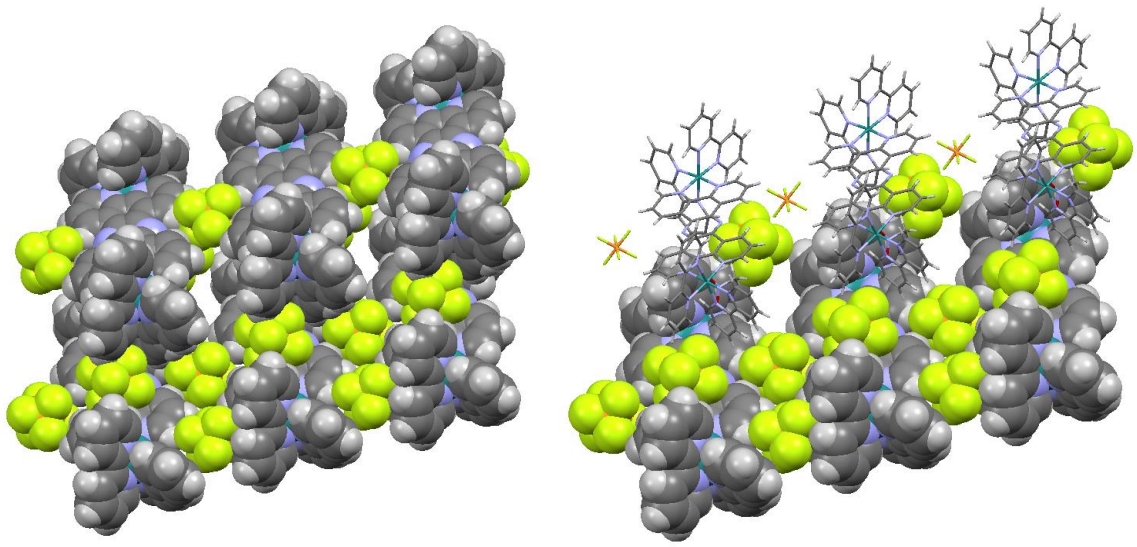


Figure 2.18. Crystallographic structure of Λ,Λ-[1].4PF_6 showing packing with stacks of Λ,Λ-[1] with PF_6 ions sitting in the channels.

2.4.4 X-ray Crystallographic Studies of Δ,Δ-[2]

Crystals of Δ,Δ-[2].4Cl suitable for X-ray structure determination were successfully obtained by vapour diffusion of acetone into a methanolic solution of the chloride salt Δ,Δ-[2].4Cl. A summary of the crystallographic data can be found in the appendix. The crystal structure is shown in figure 2.19 below is the first reported of this metal complex in any form.[17] The crystal shows two different packing sites A and B and therefore ORTEP plots for both sites are shown.

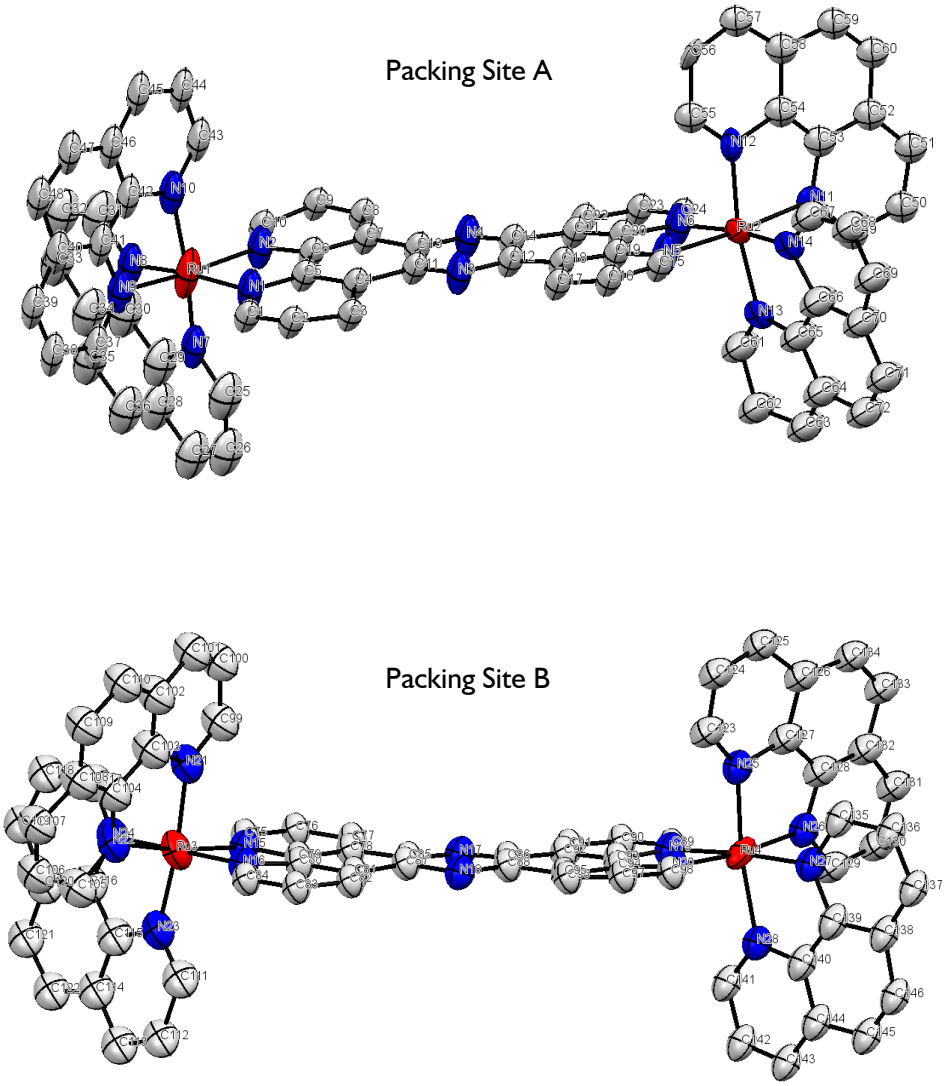


Figure 2.19. ORTEP plot of the X-ray crystallographic structure of Δ,Δ-[2] with thermal ellipsoids indicating 50% probability. There are two different packing sites as illustrated above.

The asymmetric unit is complicated containing Δ,Δ-[2] complexes with two distinct sites A and B (figure 2.20). The crystals form an orthorhombic system with the space group $P2_12_12_1$. The flack parameter is 0.157.

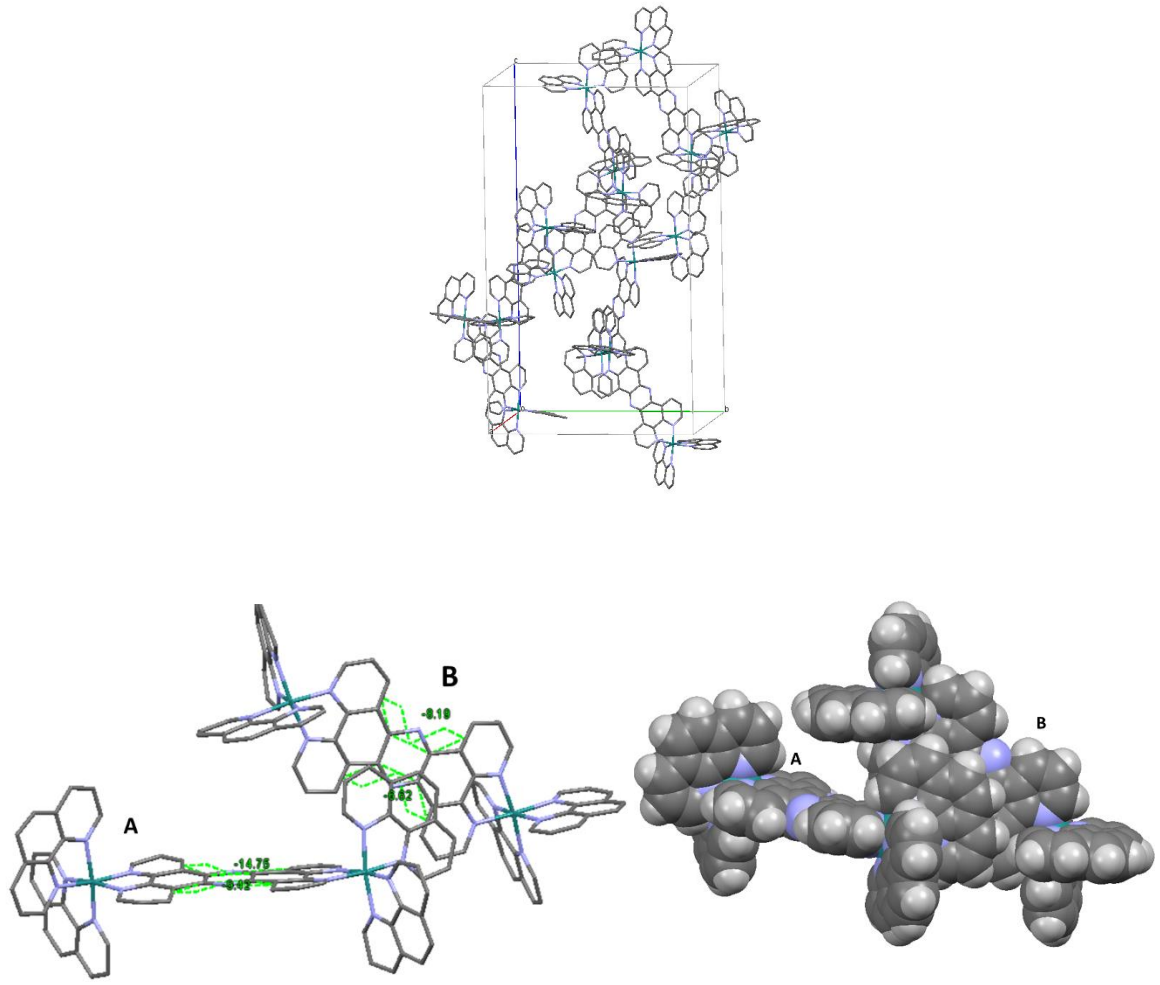


Figure 2.20. Packing diagram of orthorhombic [2] (Top). Different packing sites A and B (Bottom). Solvent molecules and hydrogen atoms are omitted for clarity.

Both A and B sites show Δ,Δ-[2] in distorted octahedral geometries around the ruthenium centres, with Ru-N bond lengths of A, 1.936-2.088 Å, and B, 1.89-2.093 Å. The tpphz ligand displays distortion from planarity in both sites, with site A exhibiting the largest twist of 14.75 °. This large deviation from planarity is likely caused by crystal packing forces and is seen to a lesser extent in the parent mononuclear dppz complex.[20]

The packing of Δ,Δ-[2] has multiple screw axes with sites A and B tightly packed together. The phen ligands of the site A show π stacking in close contact to the tpphz of the site B complexes in figure 2.21.

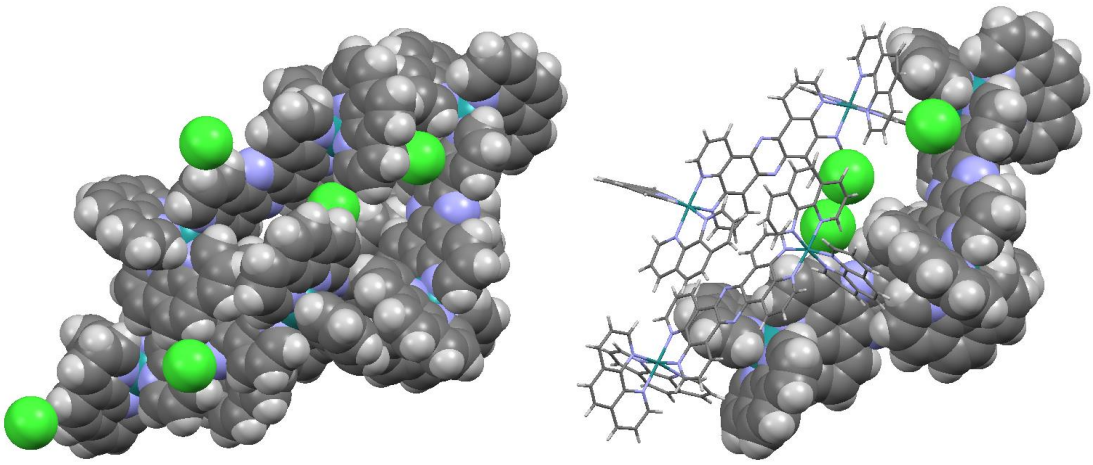


Figure 2.21. Crystallographic structure of Δ,Δ-[2].4Cl showing packing interlocking screw axis Δ,Δ-[2] A and B with chloride ions shown .

2.5 Summary

This chapter reports the successful synthesis and subsequent separation of the dinuclear ruthenium complexes, $[(Ru(bipy)_2)_2tpphz][Cl]_4$ and $[(Ru(phen)_2)_2tpphz][Cl]_4$ into their three corresponding stereoisomers, Δ,Δ, Λ,Λ and Λ,Δ. The separation into the three isomers was carried out using column chromatography employing a SP-sephadex C-25 cation exchange resin and was confirmed by CD studies.

The stereospecific structures determined by X-ray crystallography within this chapter are the first to be reported. The X-ray crystal structures of Λ,Δ-[1] and Δ,Δ-[1] indicate that different isomers can produce changes in crystal packing, although direct comparison is made difficult as different anions were employed. Although both structures display similar ordered packing of the cations with regular channels within the unit cell, they each exhibit different space groups. The crystal structure of, Δ,Δ-[2], containing phen ligands displays more complicated packing with two different packing sites. The crystal samples of Λ,Δ-[2] and Λ,Λ-[2] are currently being resolved which will enable a more in depth comparison.

2.6 References

1 J. Bolger, A. Gourdon, E. Ishow and J.-P. Launay, *Inorg. Chem.*, 1996, **35**, 2937–2944.

2 R. H. Zheng, H. C. Guo, H. J. Jiang, K. H. Xu, B. B. Liu, W. L. Sun and Z. Q. Shen, *Chinese Chem. Lett.*, 2010, **21**, 1270–1272.

3 B. P. Sullivan, D. J. Salmon and T. J. Meyer, *Inorg. Chem.*, 1978, **17**, 3334–3341.

4 B. Kolp, H. Viebrock, A. von Zelewsky and D. Abeln, *Inorg. Chem.*, 2001, **40**, 1196–1198.

5 C. Hiort, P. Lincoln and B. Norden, *J. Am. Chem. Soc.*, 1993, **115**, 3448–3454.

6 B. Bosnich and F. Dwyer, *Aust. J. Chem.*, 1966, **19**, 2229.

7 X. Hua and A. von Zelewsky, *Inorg. Chem.*, 1995, **34**, 5791–5797.

8 M. G. Finn, *Chirality*, 2002, **14**, 534–540.

9 J. K. Barton and J. S. Nowick, *J. Chem. Soc. Chem. Commun.*, 1984, 1650.

10 J. Lacour and V. Hebbe-Viton, *Chem. Soc. Rev.*, 2003, **32**, 373–382.

11 L. M. Sweeting, D. C. Crans and G. M. Whitesides, *J. Org. Chem.*, 1987, **52**, 2273–2276.

12 L. Salassa, E. Borfecchia, T. Ruiu, C. Garino, D. Gianolio, R. Gobetto, P. J. Sadler, M. Cammarata, M. Wulff and C. Lamberti, *Inorg. Chem.*, 2010, **49**, 11240–11248.

13 N. C. Fletcher, P. C. Junk, D. A. Reitsma and F. R. Keene, *J. Chem. Soc. Dalt. Trans.*, 1998, **0**, 133–138.

14 F. Richard Keene, *Chem. Soc. Rev.*, 1998, **27**, 185.

15 F. M. Foley, F. R. Keene and J. G. Collins, *J. Chem. Soc. Dalt. Trans.*, 2001, 2968.

16 N. C. Fletcher and F. R. Keene, *J. Chem. Soc. Dalt. Trans.*, 1999, 683–690.

17 J. Bolger, A. Gourdon, E. Ishow and J. P. Launay, *J. Chem. Soc., Chem. Commun.*, 1995, **0**, 1799–1800.

18 M. Biner, H. B. Buergi, A. Ludi and C. Roehr, *J. Am. Chem. Soc.*, 1992, **114**, 5197–5203.

19 H. D. Flack and G. Bernardinelli, *J. Appl. Crystallogr.*, 2000, **33**, 1143–1148.

20 F. E. Poynton, J. P. Hall, P. M. Keane, C. Schwarz, I. V. Sazanovich, M. Towrie, T. Gunnlaugsson, C. J. Cardin, D. J. Cardin, S. J. Quinn, C. Long and J. M. Kelly, *Chem. Sci.*, 2016, **7**, 3075–3084.

Chapter 3. NMR Studies

3.1 Background

Structural determination of nucleic acids has traditionally been carried out by X-ray crystallography. Proceeding from the first X-ray data based on fibre diffraction of B-DNA from Crick, Watson and Franklin, to the much studied Dickerson dodecamer which was the first canonical B-DNA structure to be derived from a single crystal (figure 3.1).[1,2] In this method, data is collected from X-ray beams diffracted by electrons in a crystal according to Bragg's law.[3] This can be used to generate electron density maps, to which a structure can be fitted and repeatedly refined until the structure fits the diffraction data.[4] Whilst X-ray crystallographic methods are a powerful tool in the structural determination of biomolecules it is not infallible.[5] Nucleic acids, in the absence of proteins, are often highly flexible and growth of suitable crystals is not trivial. As crystallographic studies are based on molecules in the crystalline phase they are intrinsically subjected to crystal packing forces. Dynamic processes in solution, such as ligand binding, need to be fortuitously captured in the crystal for structural assignment.[6]

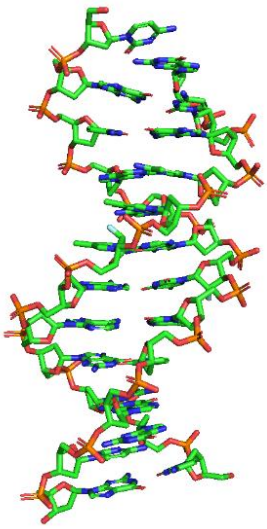


Figure 3.1. The Dickerson dodecamer d(CGCGAATTCGCG) solved by X-ray crystallography to 1.1 Å resolution. [7]

NMR spectroscopy is a technique which probes the chemical environment of nuclei subjected to a radio frequency pulse, with nuclear spins aligned by an external magnetic field. This technique enables the study of molecules in solution

in their natural state, without the influence of crystal packing forces. Whilst the resolution obtained is generally lower than that of the direct structural information given by crystallography, it is possible to probe dynamic systems with multiple conformations. A variety of NMR experiments have been developed to provide structural and dynamic information about biomolecules.[8] Thus, NMR has become a powerful tool in the study of binding events and nucleic acid conformation, with 1256 out of 3262 nucleic acid structures in the nucleic acids database determined through NMR techniques.[9] The technique has been used previously to determine the binding mode of several small molecules to DNA.[10,11]

3.2 B-DNA Duplex Selection

A non-self-complementary DNA octanucleotide sequence d(GCATATCG).(CGATATGC) was selected for this binding study for several reasons. A palindromic self-complementary DNA sequence, whilst initially simple to assign in the free duplex, causes ambiguity when the degeneracy of the strands is broken. This occurs upon asymmetrical binding of a ligand, and for this reason a non-complementary sequence was selected. The short octanucleotide sequence was also chosen to provide simplicity in assignment over a longer sequence. An octanucleotide is about the shortest duplex that can readily form a dimer solution at room temperature with confidence.[12] Looking in more detail, the sequence itself was chosen to include both A and T base pairs as well as C and G base pairs, with the intention of deducing whether [**1**] binds at AT rich or GC rich sequences. AT base pairs are intrinsically weaker than GC pairs due to fewer hydrogen bonding interactions. Therefore GC pairs were selected as the terminal nucleotides in order to provide stability of the duplex and prevent fraying.[13]

3.2.1 Oligonucleotide Preparation

DNA oligonucleotides of the sequences d(GCATATCG) and d(CGATATGC) were purchased from Eurogentec Biotechnology (Southampton) and made up to 0.5 mM in D_2O with d_4-trimethylsilylpropanoic acid (TSP) as a reference. One strand d(GCATATCG) was titrated into the other strand d(CGATATGC) until a 1:1 ratio was achieved which was monitored by NMR. It was noted that a 0.5 mM solution showed multiple broad peaks due to exchange broadening. When the sample was concentrated to a 1 mM solution, the extra peaks from the melted DNA disappeared, leaving just the sharpened peaks of the duplex (figure 3.2). From this, it can be deduced that at lower concentrations the DNA is melted due to the inherent instability of the short sequence. Therefore, in order to stabilise the duplex, low temperatures and concentrations greater than 2 mM were employed. A solution of 25 mM NaCl was also added to increase the ionic strength of the solution, in order to minimise the effect of repulsion from the negatively charged phosphates between the two strands.

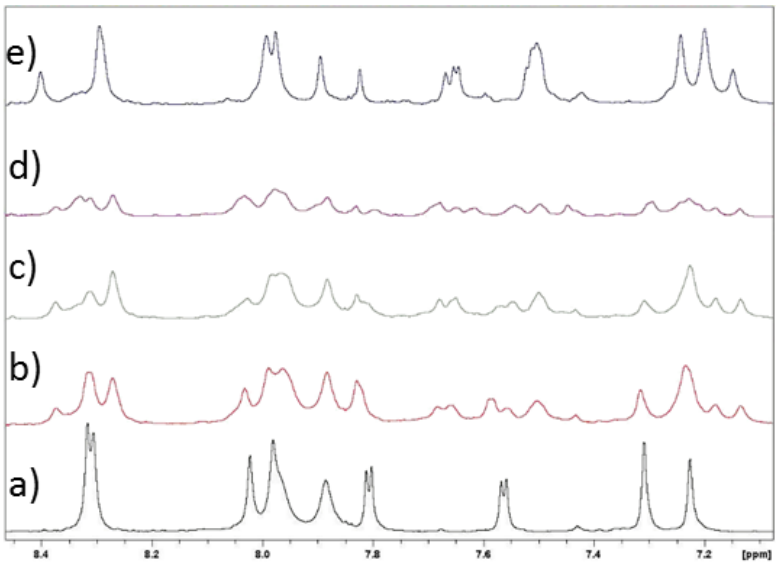


Figure 3.2. The aromatic region of the ^{1}H NMR spectrum showing the addition of d(CGATATGC) into d(GCATATCG) and the effect of concentration on duplex stability. a) 0.5 mM solution of d(GCATATCG) in D_2O at 298 K, b) the addition of 1/3 molar equivalents of d(CGATATGC) in D_2O at 298 K, c) 2/3 molar equivalents of d(CGATATGC), d) 0.5 mM solution containing both strands and e) 1 mM solution containing both strands.

3.2.1.1 Oligonucleotide Non-Exchangeable Proton Assignment

With a suitably stable oligonucleotide solution to hand, chemical shift assignment of the DNA sequence was necessary. As a starting point, average chemical shifts from the Biological Magnetic Resonance Data Bank were used to research where the general chemical shifts of the non-exchangeable protons in a B-DNA sample arise (figure 3.3).[14]

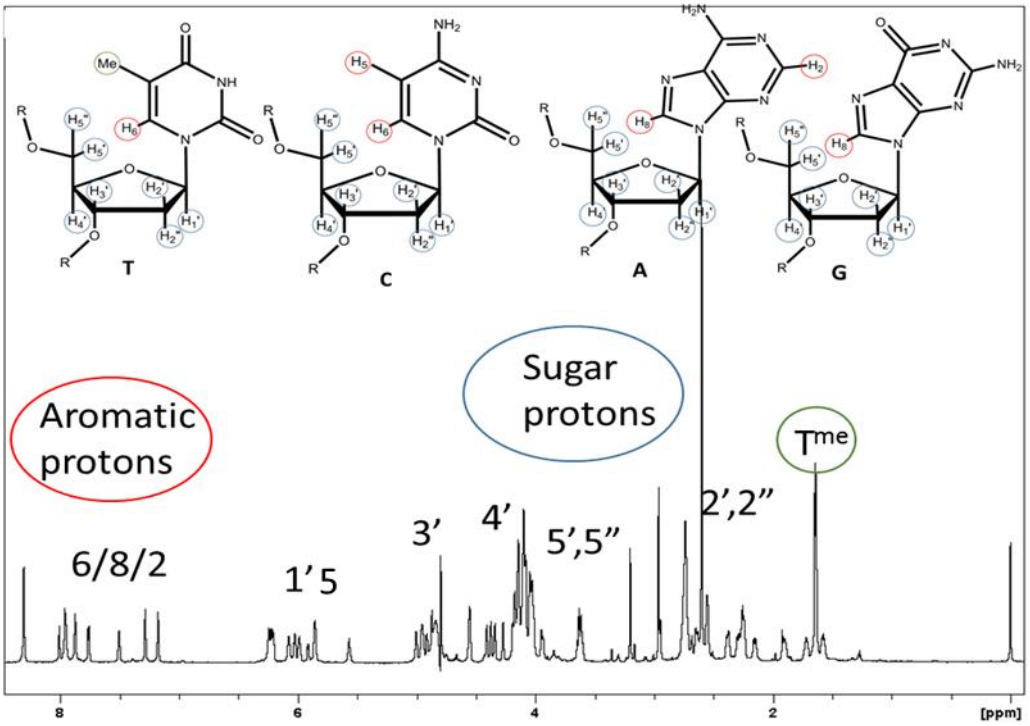


Figure 3.3. ^{1}H NMR spectrum at 298 K of 2 mM solution of DNA duplex d(GCATATCG).(CGATATGC) in D_2O with 25 mM NaCl outlining the general regions of the chemical shifts of non-exchangeable protons in a B-DNA sample.[14]

The exchangeable protons of DNA are very useful for assignment and structure determination. In duplex B-DNA the imino protons involved in the base pairing give rise to sharp signals in the 12-14 ppm region, which indicates stable Watson and Crick base pairing. In order to observe these signals, it is necessary for the sample to be in H_2O rather than D_2O.[15] However, the presence of H_2O obscures the 3', 4' and 5' signals and would make it impossible for the assignment of phosphorus chemical shift to be carried out later in this study. For this reason, the non-exchangeable protons were instead assigned in D_2O.

Full assignment of the non-exchangeable protons of the double helix was carried out using a standard procedure of a combination of COSY, TOCSY and

NOESY 2D NMR experiments (figure 3.4), using NOESY and TOCSY mixing times of 200 ms and 60 ms respectively.[16–18] Experimental details are outlined in chapter 8. COSY experiments are used to probe scalar 'through bond' coupling of protons generally three bonds apart or less. TOCSY experiments, through the application of a spin-lock, allow the magnetization of J coupled protons to diffuse across the whole spin system to give a series of cross peaks between sets of protons. The combination of these two experiments allowed the stepwise assignment of the sugar spin systems and the aromatic spin systems of the T (TMe-H6) and C (H5-H6) nucleotides shown in red (figure 3.4). On the other hand, NOESY spectra exploit dipolar 'through space' coupling giving a cross peak between protons with intensities proportional to the interatomic distances (r) by a factor of r^{-6}. These usually occur with protons which are less than 5 Å apart.[19] The NOEs were used to link up the sugar spin systems to the corresponding base protons, thereby enabling the assignment of each nucleotide.

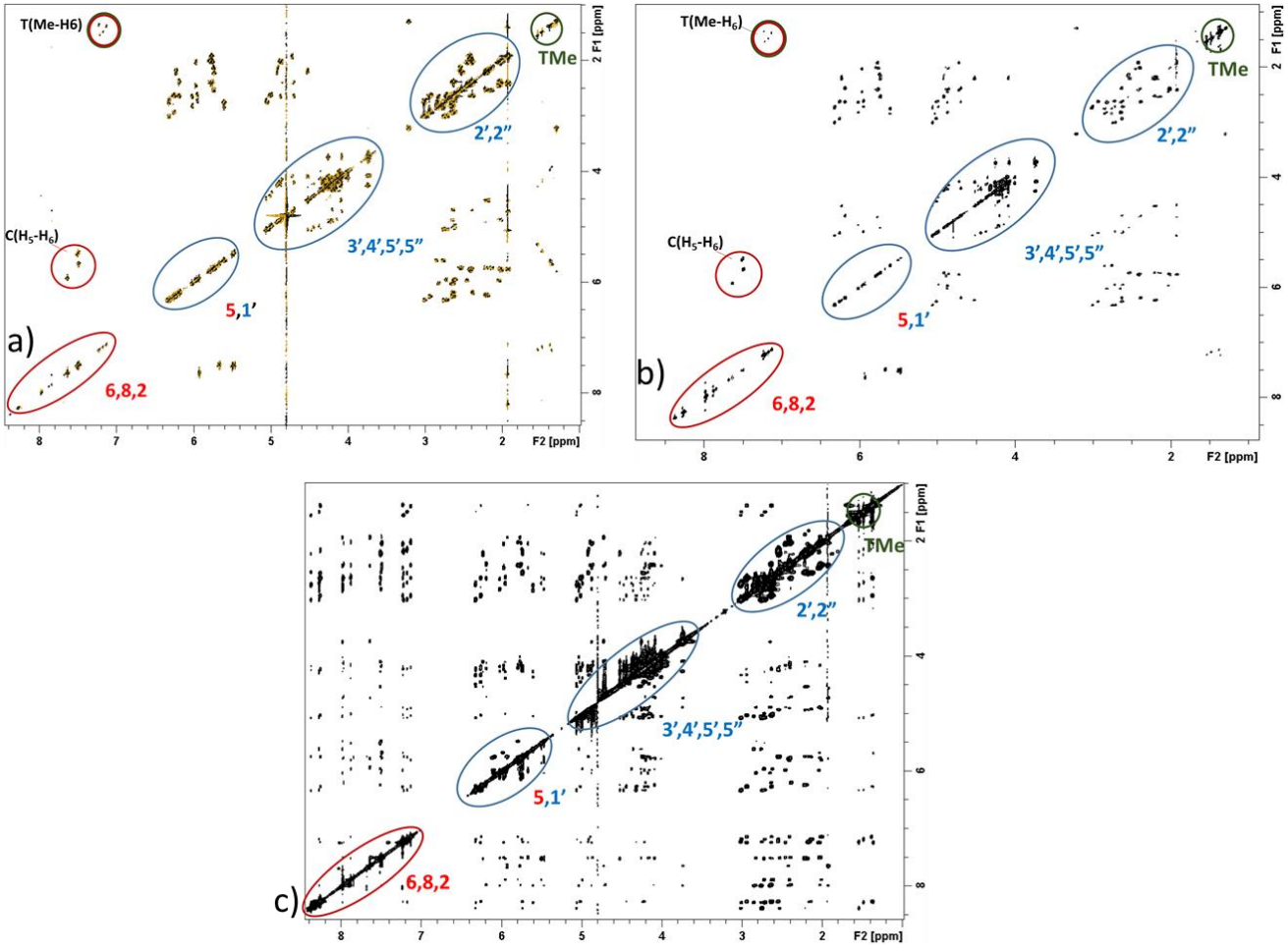


Figure 3.4. The 2D NMR spectra used in the assignment of oligonucleotides. Indicated are the key regions identifiable in each spectrum d($G_1C_2A_3T_4A_5T_6C_7G_8$) and d($C_9G_{10}A_{11}T_{12}A_{13}T_{14}G_{15}C_{16}$). a) COSY, b) TOCSY and c) NOESY

In B-DNA the inter-proton distances are such that base protons (H6/8) are not only close enough to the sugar protons (H1', H2' and H2") on their own nucleotide (i), but also the previous nucleotide (i-1) to give NOEs. Therefore, inter-nucleotide NOEs provide a method of sequential assignment referred to as an "NOE walk".[18] The base protons of the nucleotide at the 5' terminus show only NOEs to the sugar protons of their own nucleotide and this is used as a starting point for the walk. This is correlated to the 1' proton of the sugar which is then correlated to the base proton of the next nucleotide iteratively until the sequential assignment is complete for both strands. Full NOE walks were completed for both H6/8-1' and H6/8-2',2" indicating a stable duplex and allowing, through connection of the previously assigned spin systems, the chemical shift assignments which are seen in figure 3.5.

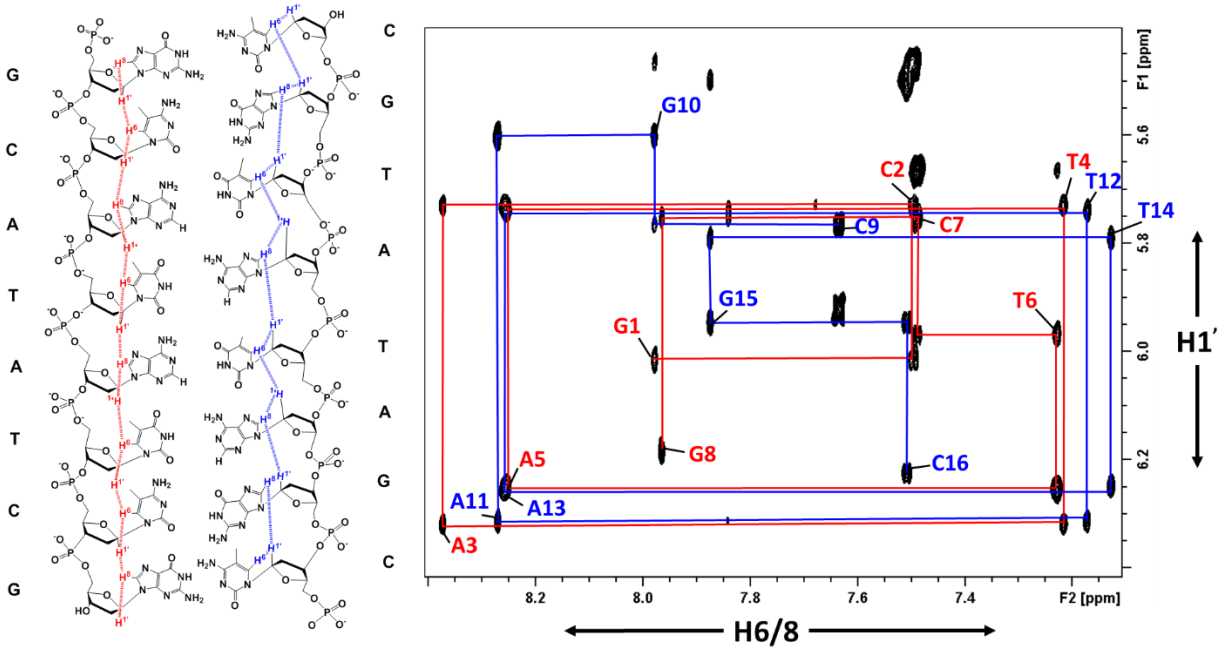


Figure 3.5. Expansion of the H1'-H6/8 region of the NOESY spectrum of a d($G_1C_2A_3T_4A_5T_6C_7G_8$) and d($C_9G_{10}A_{11}T_{12}A_{13}T_{14}G_{15}C_{16}$) duplex. The lines indicate the DNA 1'-H6/8 internucleotide NOE walk.

The chemical shifts of all non-exchangeable assigned protons are shown in table 3.1. Of all the protons, 118 out of 136 were successfully assigned. The 18 unassigned protons were from the 4', 5' and 5" protons of certain residues that fall in a very crowded region of the spectrum.

Table 3.1. The full chemical shift assignment (δ/ppm) in D_2O at 298 K for all non-exchangeable protons of the DNA duplex $d(G_1C_2A_3T_4A_5T_6C_7G_8).(C_9G_{10}A_{11}T_{12}A_{13}T_{14}G_{15}C_{16})$. Na indicates the chemical shift could not be assigned due to spectral overlap.

Residue	H8	H6	H5	H1'	H2'	H2"	H3'	H4'	H5'/5"	Me
δ G1	**7.98**	**-**	**-**	**6.01**	**2.64**	**2.82**	**4.87**	**4.25**	**3.73**	**-**
δ C2	**-**	**7.49**	**5.46**	**5.73**	**2.20**	**2.52**	**4.92**	**Na**	**Na**	**-**
δ A3	**8.37**	**-**	**-**	**6.32**	**2.74**	**3.00**	**5.06**	**4.47**	**4.14,4.22**	**-**
δ T4	**-**	**7.22**	**-**	**5.73**	**2.16**	**2.54**	**4.92**	**Na**	**Na**	**1.53**
δ A5	**8.26**	**-**	**-**	**6.25**	**2.67**	**2.93**	**5.01**	**4.45**	**4.21**	**-**
δ T6	**-**	**7.23**	**-**	**5.97**	**2.03**	**2.45**	**4.87**	**Na**	**Na**	**1.35**
δ C7	**-**	**7.49**	**5.67**	**5.75**	**2.02**	**2.38**	**4.86**	**Na**	**Na**	**-**
δ G8	**7.96**	**-**	**-**	**6.18**	**2.63**	**2.38**	**4.69**	**4.20**	**4.10**	**-**
δ C9	**-**	**7.64**	**5.92**	**5.76**	**1.92**	**2.41**	**4.71**	**4.08**	**3.72**	**-**
δ G10	**7.98**	**-**	**-**	**5.60**	**2.76**	**2.85**	**5.03**	**4.35**	**4.00,4.11**	**-**
δ A11	**8.27**	**-**	**-**	**6.31**	**2.70**	**3.00**	**5.06**	**4.52**	**4.26**	**-**
δ T12	**-**	**7.17**	**-**	**5.74**	**2.12**	**2.54**	**4.89**	**Na**	**Na**	**1.47**
δ A13	**8.26**	**-**	**-**	**6.25**	**2.63**	**2.93**	**5.02**	**4.45**	**4.21**	**-**
δ T14	**-**	**7.13**	**-**	**5.79**	**1.99**	**2.40**	**4.89**	**Na**	**Na**	**1.37**
δ G15	**7.87**	**-**	**-**	**5.94**	**2.64**	**2.75**	**4.98**	**4.38**	**4.16**	**-**
δ C16	**-**	**7.50**	**5.50**	**6.22**	**2.19**	**2.19**	**4.51**	**4.25**	**4.08**	**-**

3.3 Preliminary Studies of [1] with DNA oligonucleotide

With the NMR assignment of the DNA duplex completed, the NMR assignment of [1] was also required prior to addition of the complex to the oligo solution to enable comparison. The assignment of the bipy and tpphz spin systems was completed using COSY and TOCSY spectra and was aided by an inter ligand NOE linking the two spin systems together as shown in figure 3.6.

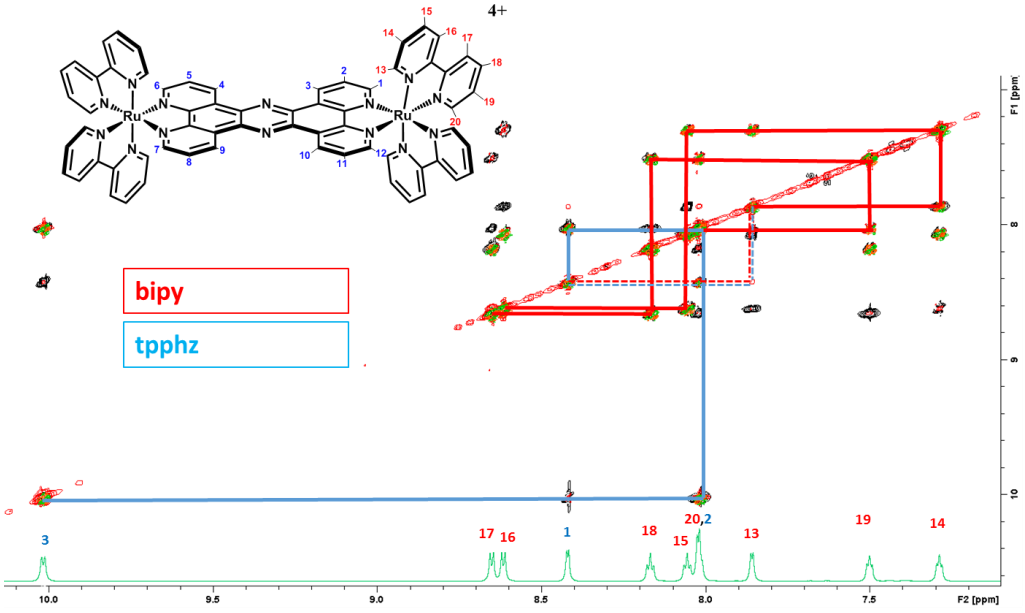


Figure 3.6. The overlaid NMR spectra of [1] in D_2O. Outlined is the chemical shift assignment with the spectra consisting of TOCSY (black), COSY (green) and NOESY (red). The red line indicates COSY couplings in the bipy spin system and the blue line indicates COSY couplings the tpphz spin system. The red and blue dashed lines indicate the inter ligand NOE coupling from bipy H_1 to tpphz H_{13}.

Preliminary binding studies of [1] to the DNA oligo sequence using the resolved complexes Λ,Δ-[1], Λ,Λ-[1] and Δ,Δ-[1] were carried out. For the free metal complex, the sets of pseudo-symmetry related protons (i.e the tpphz proton sets [1, 6, 7, 12], [2, 5, 8, 11] and [3, 4, 9, 10] as well as many of the bipy protons) have identical chemical shifts. However, in the DNA complex each group of protons gives at least 4 different bound signals, for each stereoisomer and bound conformer. Therefore, it was necessary to use the resolved compounds in this study.

3.3.1 Titration of Λ,Δ-[1] into DNA

Titration of Λ,Δ-[1] into the DNA solution was monitored by 1D ^{1}H NMR and is shown in figure 3.7. Upon inspection of the spectra, the appearance of new peaks from the bound complex became apparent. This shows the DNA to be in a slow exchange regime, as the bound and unbound DNA peaks were not averaged, implying that the dissociation rate is slow on the NMR timescale. As the ratio of Λ,Δ-[1] to DNA increased past 50%, the NMR spectra broadened and decreased in intensity as a result of precipitation of the DNA. For this reason, DNA duplex:[1] ratios of 2:1 were used in the rest of this study.

From comparison with the NMR of the free Λ,Δ-[1], under these conditions, even a cursory inspection of the 1D NMR reveals large upfield shifts (>1.5 ppm) for tpphz protons 3,4,9,10 seen in an uncrowded region of the spectrum. This large change in chemical shift has been previously reported to be an indication of intercalation in the parent mononuclear complex, as a ring current shift effect when stacking between the aromatic bases.[20]

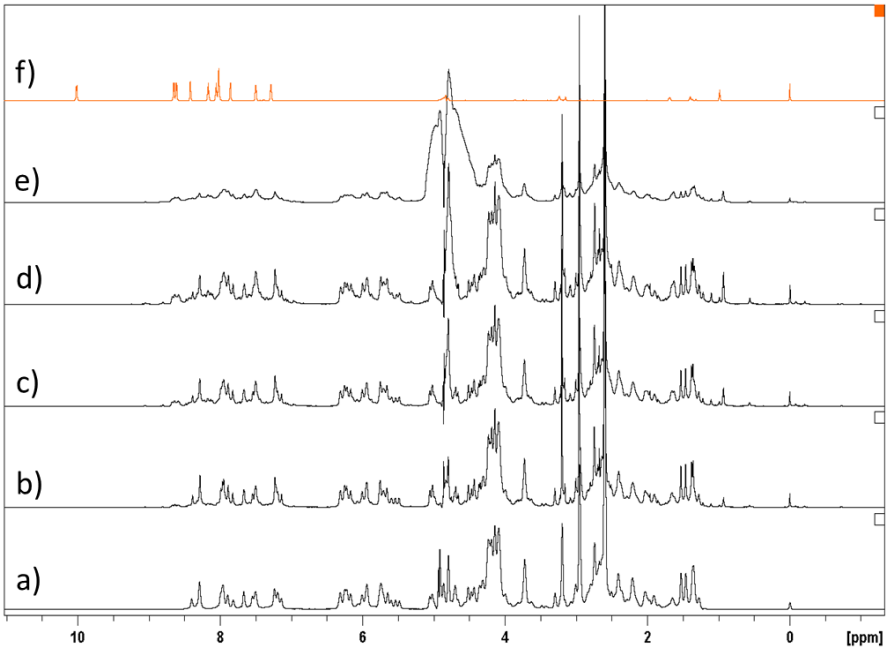


Figure 3.7. ^{1}H NMR showing the addition of Λ,Δ-[1] to DNA in D_2O at 298 K. a) 2 mM d(GCATATCG).(CGATATGC) with b) 0.33 mM Λ,Δ-[1], c) 0.66 mM Λ,Δ-[1], d) 1 mM Λ,Δ-[1] and e) 1.5 mM Λ,Δ-[1]. The orange spectrum is that of the free Λ,Δ-[1].

The 2D ^{1}H NMR experiments reveal a rather crowded aromatic region due to a non-equivalence of the bipyridyl protons upon binding (figure 3.8).

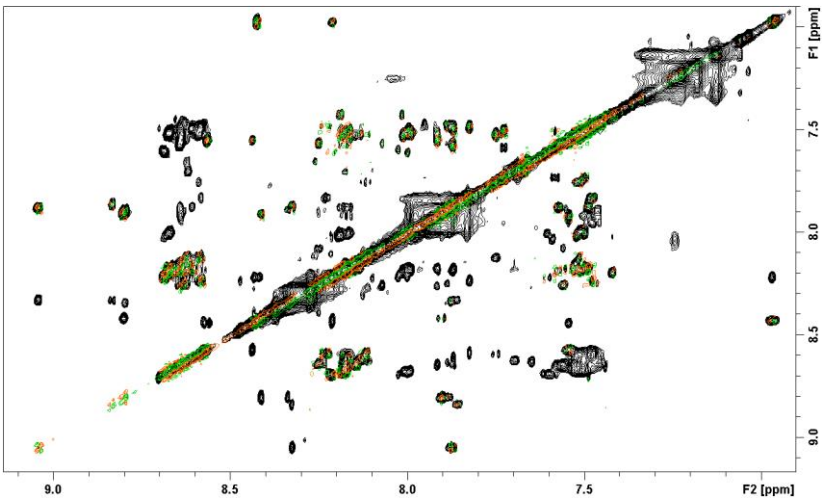


Figure 3.8. 2D ^{1}H NMR COSY (green) and TOCSY (black) spectra of a 2 mM solution of d(GCATATCG).(CGATATGC) in D_2O and 25 mM NaCl with the addition of non-deuterated Λ,Δ-[1] showing a crowded aromatic region.

Assignment of the Λ,Δ-[1]:DNA complex was aided by selective deuteration using d8-bipy, to simplify the spectra by removing the bipy proton resonances leaving only the tpphz protons as NMR active (figure 3.9). Using perdeuterated 2,2-bipyridine-d8, the synthesis and resolution of $[Ru((d8\text{-}bipy)_2)_2(tpphz)]^{4+}$ was achieved by the same procedure as [1]. NMR assignment is shown below. All further studies from this point were then carried out using deuterated resolved complexes.

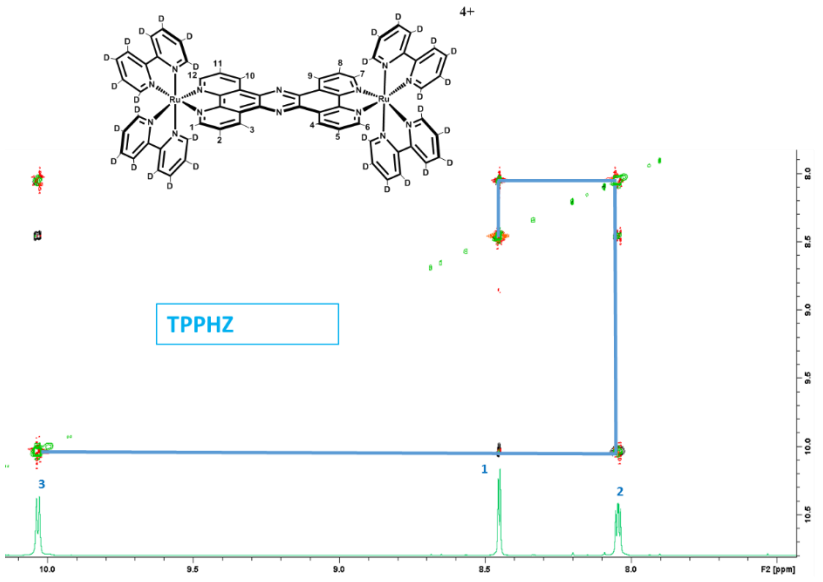


Figure 3.9. The overlaid TOCSY (black) and COSY (green) NMR spectra of deuterated [1] in D_2O. Outlined is the chemical shift assignment with proton sets 1 (1,6,7,12), 2 (2,5,8,11) and 3 (3,4,9,10) symmetrically equivalent.

3.4 Addition of d_{32}-Λ,Δ -[1] to the Oligonucleotide

Upon addition of the deuterated Λ,Δ-[1], the spectra became more assignable due to the removal of the 32 aromatic protons from the bipyridyl moieties. Several intermolecular NOEs are observed between Λ,Δ-[1] and the DNA as well as large changes in chemical shift. The TOCSY and COSY spectra of the aromatic region show two sets of 4 spin system signals for the tpphz protons with integration ratios of 2:1 (figure 3.10). This indicates that the complex binds in two conformations, a major **(Λ,Δ-[1]a)** and a minor form **(Λ,Δ- [1]b)** in slow exchange.

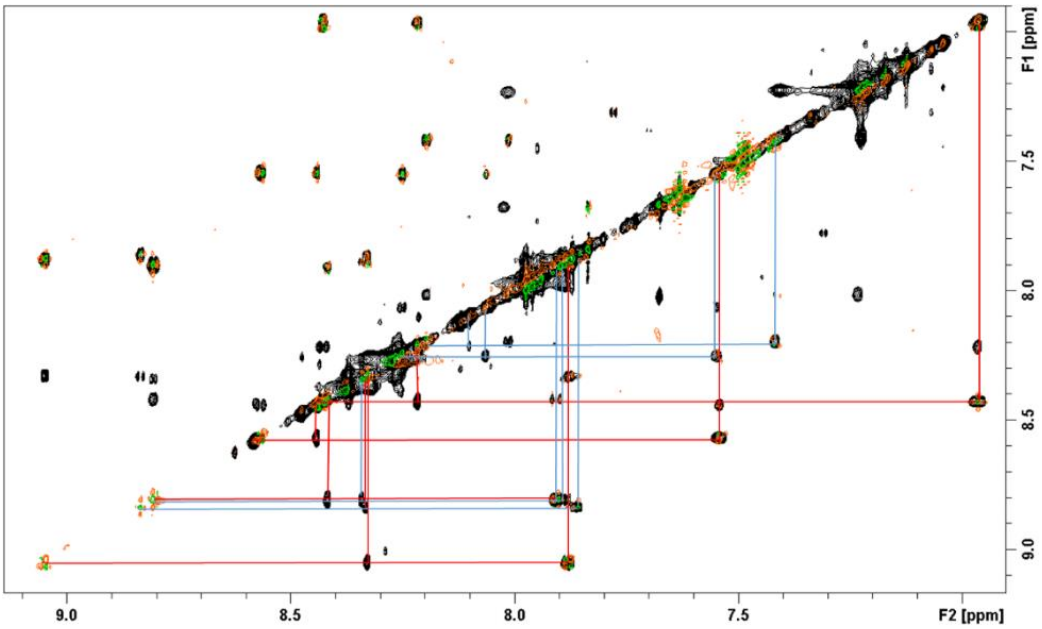


Figure 3.10. Expansion of the aromatic region of the COSY (green) and TOCSY (black) spectra of a d(GCATATCG).(CGATATGC) DNA/Λ,Δ-[1] complex. The lines indicate the discrete sets of spin systems with population ratios 2:1 for the tpphz protons in the major Λ,Δ-[1]a (red) and minor Λ,Δ-[1]b (blue) forms.

Chemical shift assignment of Λ,Δ-[1]a shows the largest upfield shifts for the central protons (3,4,9,10) then (2,5,8,11), with (1,6,7,12) protons least shifted (table 3.2). This suggests a threaded binding orientation, as the central tpphz protons are most strongly affected by the ring current shift of the bases, due to their alignment over the base stack. Comparison to the spectra of non-deuterated Λ,Δ-[1] (figure 3.8) suggests smaller chemical shift changes to the unassigned bipyridyl protons. The protons were initially assigned ambiguously and specific resonances were elucidated by molecular modelling, discussed in chapter 5.

Table 3.2. The changes in chemical shift (δ/ppm) of the major conformation of Δ,Λ-[1] upon binding to DNA duplex d(GCATATCG).(CGATATGC) in D_2O at 298 K.

Δ,Λ-[1]a	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
δ/ppm	8.22	6.96	8.43	9.05	7.88	8.33	8.42	7.90	8.81	8.57	7.54	8.44
Δδ	-0.24	-1.09	-1.61	-0.98	-0.17	-0.13	-0.04	-0.15	-1.23	-1.46	-0.51	-0.02

3.4.1 Chemical Shift Assignment of the DNA:Λ,Δ-[1]a Complex

Chemical shift assignment of the Λ,Δ-[1]:DNA complex was made using the method applied to the free duplex, but this was not trivial as the spectra contained 50% of the free duplex d(GCATATCG).(CGATATGC) and two extra sets of signals from bound conformations. However, there was sufficient spectral dispersion to enable the assignment of the sugar protons of the major Λ,Δ-[1]a conformation (table 3.3).

Table 3.3. The changes in chemical shift (δ/ppm) of the major conformation of DNA duplex d(GCATATCG).(CGATATGC) upon addition of Λ,Δ-[1] in D_2O at 298 K.

Λ,Δ-[1]a	H8	H6	H5	H1'	H2'	H2"	H3'	Me
G1	7.94	-	-	5.97	2.65	2.76	4.82	-
Δδ	-0.04	-	-	-0.05	0.01	-0.06	-0.06	-
C2	-	7.43	5.31	5.73	2.08	2.41	4.91	-
Δδ	-	-0.06	-0.15	0.01	-0.13	-0.12	-0.01	-
A3	8.12	-	-	6.14	2.43	2.83	5.01	-
Δδ	-0.25	-	-	-0.17	-0.31	-0.18	-0.06	-
T4	-	7.20	-	5.62	2.55	2.55	5.11	0.57
Δδ	-	-0.02	-	-0.11	0.39	0.01	0.19	-0.96
A5	8.44	-	-	5.66	2.64	2.76	4.85	-
Δδ	0.18	-	-	-0.59	-0.04	-0.17	-0.17	-
T6	-	7.04	-	5.73	1.98	2.35	4.85	1.12
Δδ	-	-0.18	-	-0.24	-0.05	-0.09	-0.02	-0.23
C7	-	7.48	5.63	5.76	1.99	2.36	4.87	-
Δδ	-	0.00	-0.04	0.00	-0.03	-0.02	0.01	-
G8	7.94			6.14	2.60	2.37	4.67	-
Δδ	-0.02			-0.04	-0.03	-0.02	-0.01	-
C9		7.65	5.87	5.73	2.01	2.43	4.71	-
Δδ		0.01	-0.05	-0.03	0.09	0.02	-0.01	-
G10	7.95	-	-	5.65	2.72	2.83	5.02	-
Δδ	-0.03	-	-	0.05	-0.04	-0.01	-0.01	-
A11	8.10	-	-	6.19	2.49	2.91	5.05	-
Δδ	-0.17	-	-	-0.13	-0.21	-0.10	-0.01	-
T12	-	7.34	-	5.72	2.65	2.40	5.14	0.98
Δδ	-	0.17	-	-0.02	0.53	-0.14	0.25	-0.49
A13	8.48	-	-	5.67	2.75	2.82	4.97	-
Δδ	0.21	-	-	-0.58	0.12	-0.12	-0.06	-
T14	-	7.07	-	5.64	1.96	2.37	4.88	1.23
Δδ	-	-0.05	-	-0.14	-0.03	-0.03	-0.01	-0.15
G15	7.90	-	-	6.00	2.65	2.75	5.00	-
Δδ	0.03	-	-	0.06	0.01	0.00	0.02	-
C16	-	7.47	5.48	6.17	2.19	2.19	4.51	-
Δδ	-	-0.03	-0.02	-0.05	0.00	0.00	0.00	-

Attempts to assign the chemical shifts for the Λ,Δ-[1]b conformation were unsuccessful due to the significant spectral overlap of many of the resonances. However, the chemical shift changes for Λ,Δ-[1]a can be visualised more clearly in the graphs below in figure 3.11.

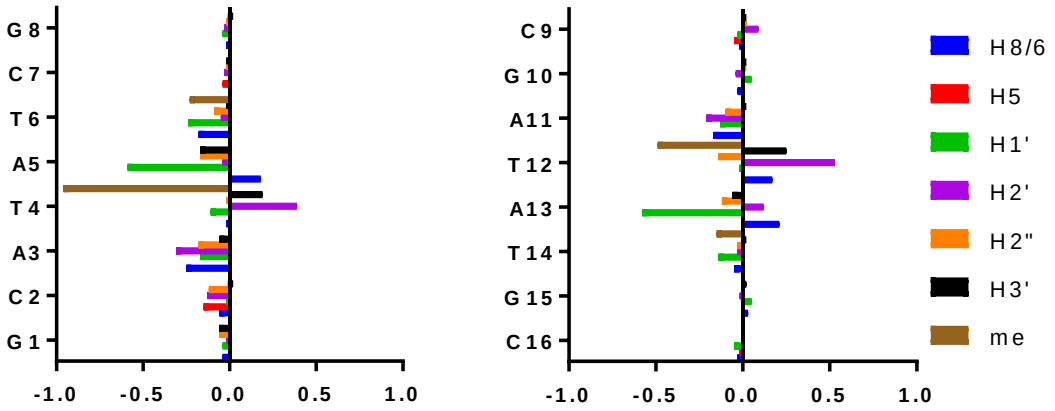


Figure 3.11. The chemical shift perturbations (δ/ppm) of the DNA duplex d($G_1C_2A_3T_4A_5T_6C_7G_8$).($C_9G_{10}A_{11}T_{12}A_{13}T_{14}G_{15}C_{16}$) for the major bound form seen upon addition of half an equivalent of Λ,Δ-[1] in 298 K.

The largest changes in chemical shift are observed for the central adenine and thymine residues, which indicates that binding occurs at the central TA base step. This preference is likely due to the flexibility of AT rich tracts observed in previous studies with similar compounds.[23] The large chemical shift changes to the thymidine methyl (TMe) protons (largest Δδ, T_4Me -0.96 ppm) located in the major groove and to the sugar protons (largest Δδ, A_5H1' -0.59 ppm) located in the minor groove is a good indication of threading intercalation or a conformational change in the DNA upon binding.

An expansion of the spectra (figure 3.12) revealed that the sequential assignment of the nucleotide strands was incomplete as the H1'-H6/8 NOE walk broke down at the central (T_4,A_{13}) | (T_{12},A_5) binding site. This result is expected for an intercalating compound, as intercalation of Λ,Δ-[1] at the (T_4,A_{13}) | (T_{12},A_5) step would cause helical unwinding, leading to an increase of distance between these protons to over 5Å.[21] Interestingly, the sequential assignment could be completed by walking through the intermolecular NOEs from Λ,Δ-[1]a (figure 3.12). The NOE close contacts between (T_4H1'-A_5H8) and (T_{12}H1'-A_{13}H8) disappeared upon binding and were replaced with new intermolecular NOEs from the tpphz protons

which completed the sequential assignment. This is a very strong indication of intercalation, as this shows that the distance between the bases has increased upon insertion of the tpphz ligand. A schematic representation is shown on the left hand side of figure 3.12.

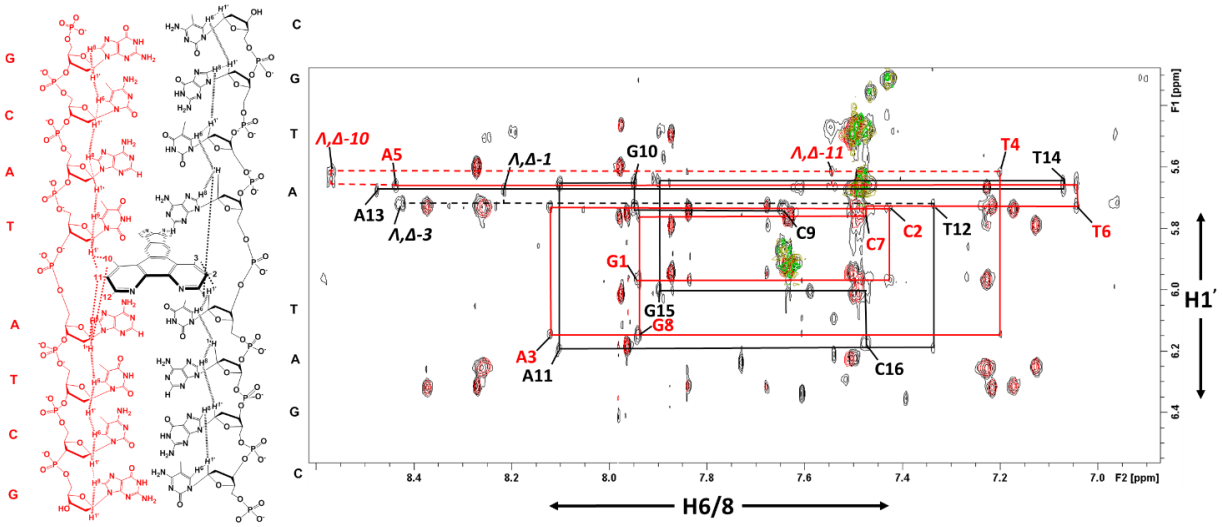


Figure 3.12. Expansion of the H1'-H6/8 region of the NOESY and COSY spectrum of a d($G_1C_2A_3T_4A_5T_6C_7G_8$).($C_9G_{10}A_{11}T_{12}A_{13}T_{14}G_{15}C_{16}$) DNA/Λ,Δ-[1] complex in D_2O at 298 K. The solid lines show the DNA 1'-H6/8 internucleotide NOE walk and dashed lines show NOE connectivities between Λ,Δ and the DNA 1' protons. The overlaid NMR spectra show the NOESY spectra of 2 mM DNA (red), 2 mM DNA + 1 mM Λ,Δ-[1] (black) and the COSY spectrum 2 mM DNA + 1 mM Λ,Δ-[1] (green).

As well as the intermolecular NOEs to the 1' protons located in the minor groove there are several NOEs to the TMe groups which protrude into the major groove (figure 3.13). In total, there are 23 assignable intermolecular NOEs between the oligonucleotide and Λ,Δ-[1] as shown in (table 3.4). These extend into both the major and minor grooves and from either end of the tpphz ligands. On comparison with the non-deuterated sample, few NOEs were seen from the bipy ancillary ligands with only two to the thymine methyl protons in the major groove.

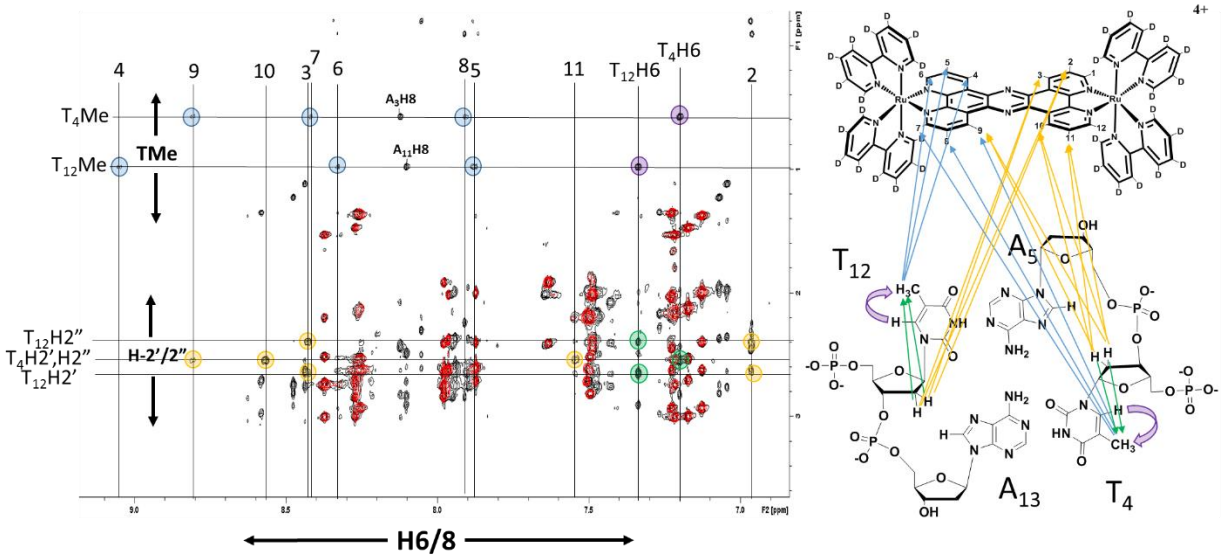


Figure 3.13. Expansion of the 2',2"and TMe-aromatic region of the NOESY spectrum of a 2 mM D_2O solution containing duplex sequence d(GCATATCG).(CGATATGC) in (red), overlayed onto a spectrum containing half an equivalent of Λ,Δ-[1] (black), indicating the intramolecular NOE correlations H6/2',2" (green) and H6/TMe (purple) and the intermolecular NOE correlations Λ,Δ-[1]a/TMe (blue) and Λ,Δ-[1]a/H2',2" (orange) located in both grooves of the DNA. A schematic of the (T_4,A_{13})|(T_{12},A_5) binding site is given for reference.

Table 3.4. Intermolecular NOEs between the DNA duplex d(GCATATCG).(CGTATAGC) and Λ,Δ-[1]a in D_2O in 298 K.

Λ,Δ-[1]a	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12
T4							Me	Me	Me,2',2"	1',2',2"	1',2',2"	2"
A5												
T12		1',2',2"	1',2',2"	Me,2'	Me	Me						
A13	1'		1'									

3.4.2 DNA:Λ,Δ-[1]b Complex

Although full chemical shift assignment of the DNA:Λ,Δ-[1]b complex could not be completed due to significant spectral overlap of the signals, an inspection of the H6/8-TMe region of the spectrum in comparison to DNA:Λ,Δ-[1]a revealed similarities between the two species (figure 3.14).

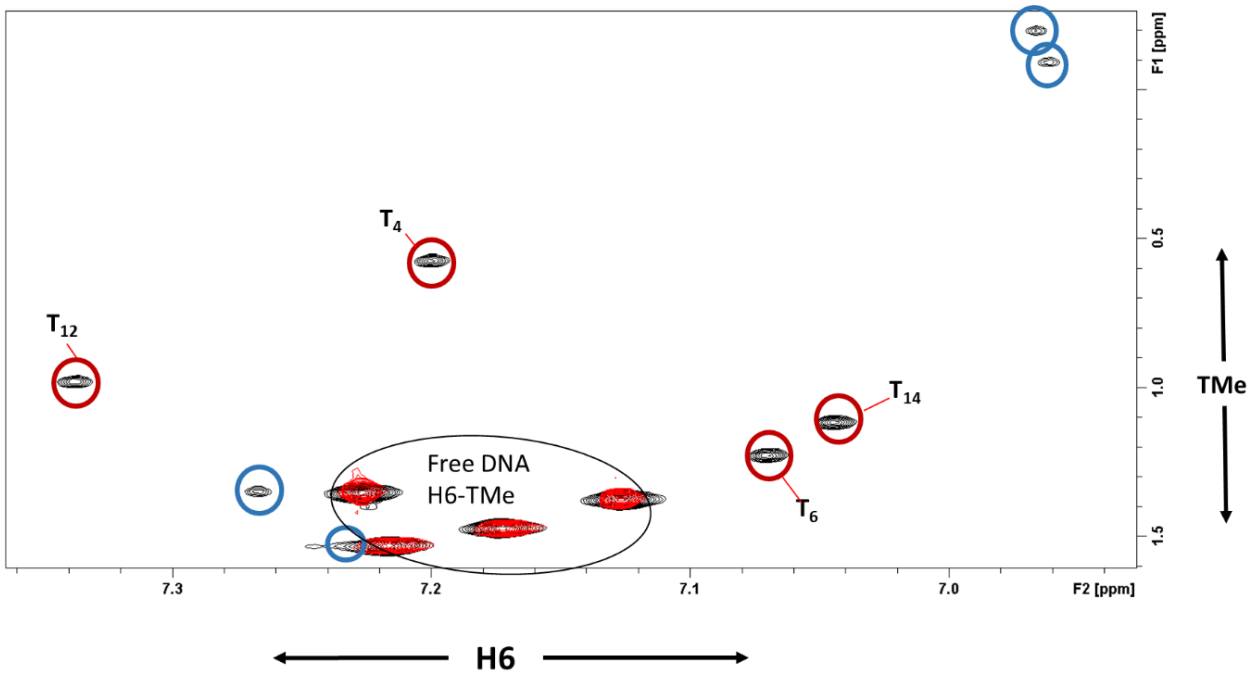


Figure 3.14. The expansion of the H6-TMe region of the TOCSY spectrum of a 2 mM D_2O solution containing duplex sequence d(GCATATCG).(CGATATGC) in (red), overlayed onto a spectrum containing half an equivalent of Λ,Δ-[1] (black). The red circles indicate the Λ,Δ-[1]a H6-TMe cross peaks, the blue circles indicate the Λ,Δ-[1]b H6-TMe cross peaks.

Two thymidine methyl peaks are shifted very far upfield in a similar fashion to that seen for the major conformation. These peaks are assignable as the thymidine methyl peaks due to the J-coupled cross peaks (H6-Me) in the TOCSY spectrum. There is also a similar pattern of NOEs to these thymidine methyl protons (figure 3.15) as well as the associated 2" protons as seen for the Λ,Δ-[1]a complex.

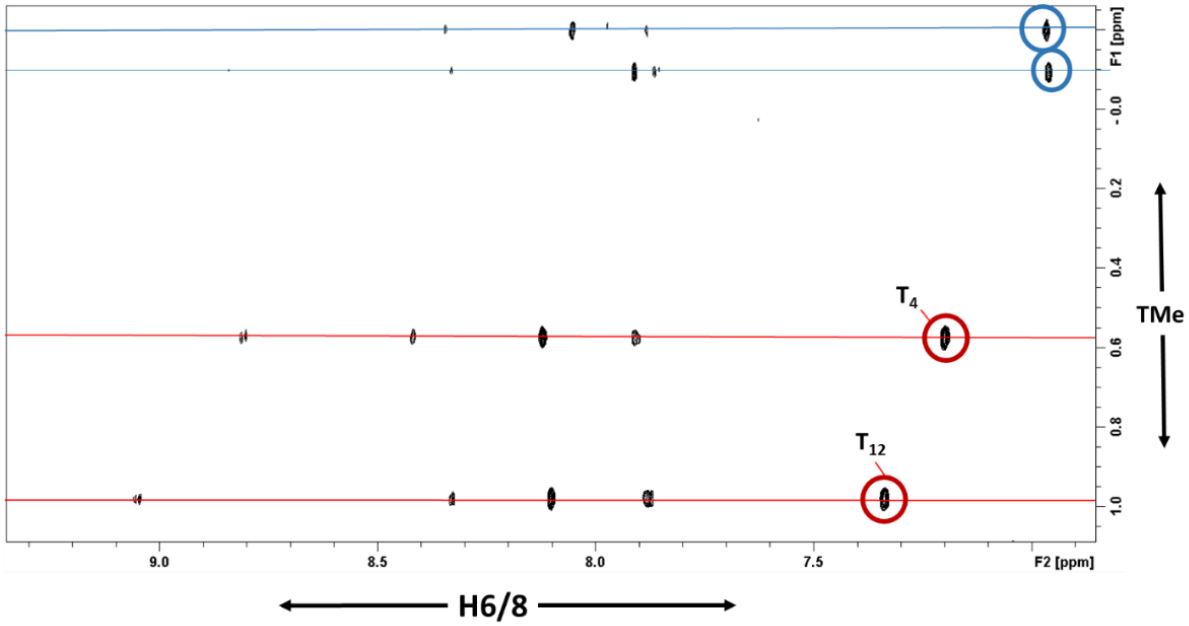


Figure 3.15. The similar array of intermolecular NOEs observed for the Λ,Δ-[1]a (red) and Λ,Δ-[1]b (blue) species. The figure shows an expansion of the TMe-aromatic region of the NOESY spectrum of a 2 mM D_2O solution containing duplex sequence d(GCATATCG).(CGTAGTAGC) and half an equivalent of Λ,Δ-[1]. The red circles indicate the Λ,Δ-[1]a H6-TMe cross peaks and the blue circles indicate the Λ,Δ-[1]b H6-TMe cross peaks.

These shifted peaks and NOEs would suggest an intercalation binding mode at a site adjacent to two thymine residues. In the DNA sequence d(GCATATCG).(CGATATGC) there are three possible binding sites fulfilling this criterion, the central TA|TA step and two sets of AT|AT base steps. In the crystal structures of the related Λ-$[Ru(phen)_2dppz]^{2+}$ complex reported by the Cardin group,[22] which was described in chapter 1, it was shown that although intercalation takes place at the central TA|TA base steps it did not occur when substituted by an AT|AT step. This study outlines the accessibility to intercalation at these base steps and suggests that the most likely binding site for Λ,Δ-[1]b is at the central (T_4,A_{13})|(T_{12},A_5) step, but with a different conformation to the major species. As Λ,Δ-[1] has two different Ru^{II} chiral centres (Λ and Δ) this phenomenon could likely be explained by conformations in which the Λ and Δ centres exchange their positions within the major and minor grooves of the DNA, leading to the two separate sets of resonances. A schematic of the binding modes proposed from this data is shown in figure 3.16.

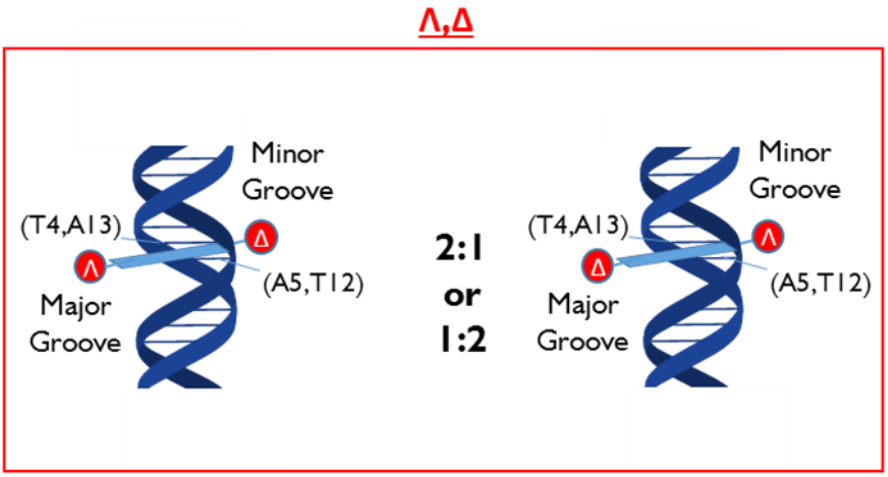


Figure 3.16. Schematic of the proposed binding modes of Λ,Δ-[1].

3.5 Addition of d_{32}-Λ,Λ-[1] to the Oligonucleotide

Upon titration of deuterated Λ,Λ-[1] into the DNA duplex, there was again two distinct sets of resonances for the bound complex with a ratio of 2:1. This indicates the existence of two binding sites or conformations for this complex (figure 3.17); a major form **Λ,Λ-[1]a** and a minor form **Λ,Λ-[1]b**. These two sets of signals are in slow exchange with each other and with the free signals, permitting an almost complete assignment of all signals.

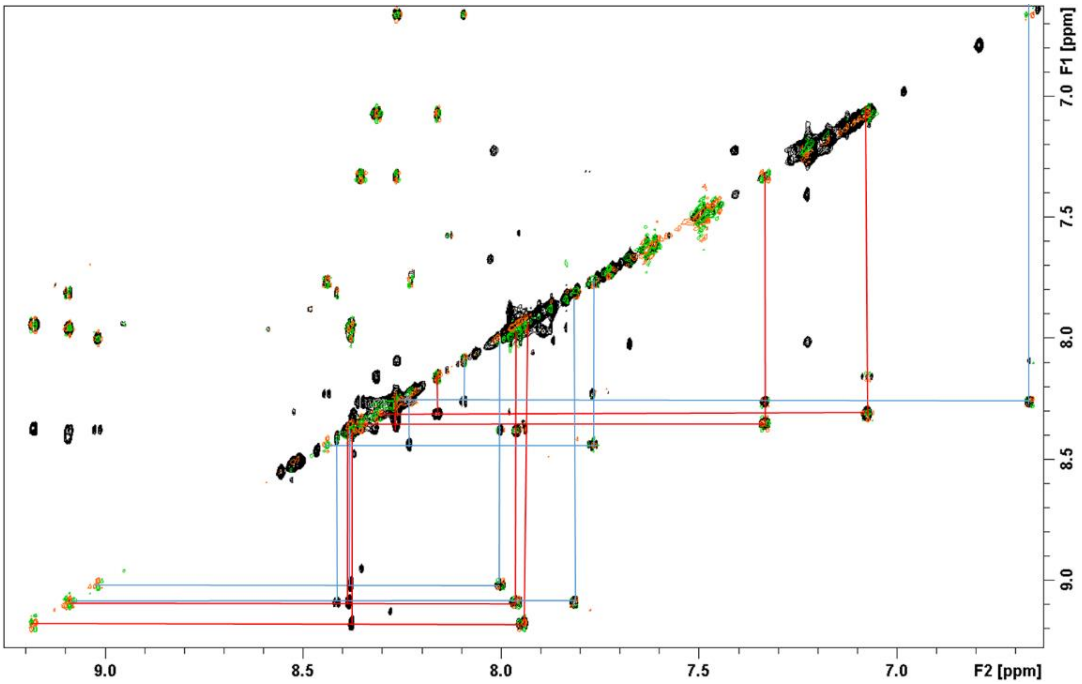


Figure 3.17. Expansion of the aromatic region of the COSY (green) and TOCSY (black) spectra of a d(GCATATCG).(CGATATGC) DNA/Λ,Λ-[1] complex. The lines indicate the spin system assignment of the tpphz in the major Λ,Λ-[1]a (red) and minor Λ,Λ-[1]b (blue) forms.

Λ,Λ-[1] resonances are shifted upfield upon binding to the DNA. The largest upfield shifts to the central protons (3,4,9,10) are also consistent with an intercalation binding mode with these protons aligned between the base stack.[20] Both of these bound forms are in a slow exchange with well dispersed peaks allowing for chemical shift assignment (table 3.5).

Table 3.5. The changes in chemical shift (δ/ppm) of Λ,Λ-[1] upon binding to d(GCATATCG).(CGATATGC) in D_2O at 298 K.

Λ,Λ-1a	1	2	3	4	5	6	7	8	9	10	11	12
δ/ppm	8.16	7.07	8.31	9.18	7.95	8.38	8.38	7.96	9.09	8.35	7.33	8.27
Δδ	-0.30	-0.98	-1.72	-0.85	-0.10	-0.08	-0.07	-0.09	-0.94	-1.68	-0.72	-0.19

Λ,Λ-1b	1	2	3	4	5	6	7	8	9	10	11	12
δ/ppm	8.38	8.00	9.02	8.44	7.77	8.23	8.42	7.81	9.09	8.26	6.67	8.10
Δδ	-0.08	-0.05	-1.01	-1.59	-0.28	-0.22	-0.04	-0.24	-0.94	-1.77	-1.38	-0.36

The predominant bound form, Λ,Λ-[1]a, showed large changes in chemical shift at the central (T_4,A_{13}) | (T_{12},A_5) base step as a result of the ring current from the aromatic complex (figure 3.18). As discussed previously, the preference for (T_4,A_{13}) | (T_{12},A_5) is due to the lower stability and higher flexibility of the AT base pairing; an observation consistent with similar compounds in previous studies.[23] Notably, very large chemical shift changes (largest Δδ, T_4CH_3' -1.61 ppm) are seen for the thymidine methyl protons located in the major groove of the DNA. Sizeable shifts are also observed for the sugar protons located in the minor groove (largest Δδ, A_{13}H1' -0.61 ppm), which is also indicative of an intercalation binding mode.

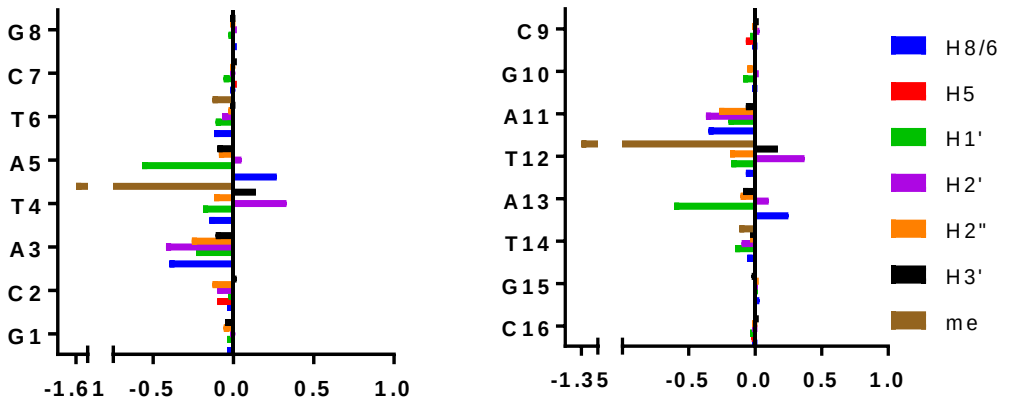


Figure 3.18. The chemical shift perturbations (δ/ppm) of the DNA duplex d($G_1C_2A_3T_4A_5T_6C_7G_8$).($C_9G_{10}A_{11}T_{12}A_{13}T_{14}G_{15}C_{16}$) upon binding of Λ,Λ-[1] in the major conformation of Λ,Λ-[1]a in D_2O at 298 K.

Interestingly, the minor bound species (Λ,Λ-[1]b) displays large changes in chemical shift at the (C2,G15)|(A3,T14) binding site with almost zero perturbation to chemical shift at the other end of the duplex (figure 3.19). The largest change in chemical shift (largest Δδ, T_{14}Me -2.17 ppm) is assigned to the thymine methyl proton in the major groove along with other significant perturbations for chemical shifts of protons located in the minor groove (largest Δδ, G_{15}H1' -0.61 ppm). This would suggest that, unlike the Λ,Δ-[1]b, where it seems two forms of binding occur at the same site, the Λ,Λ-[1]b binding site is at a completely different section of the sequence. The changes in chemical shift for the Λ,Λ-[1]:DNA complexes are shown on the following pages (tables 3.6 and 3.7).

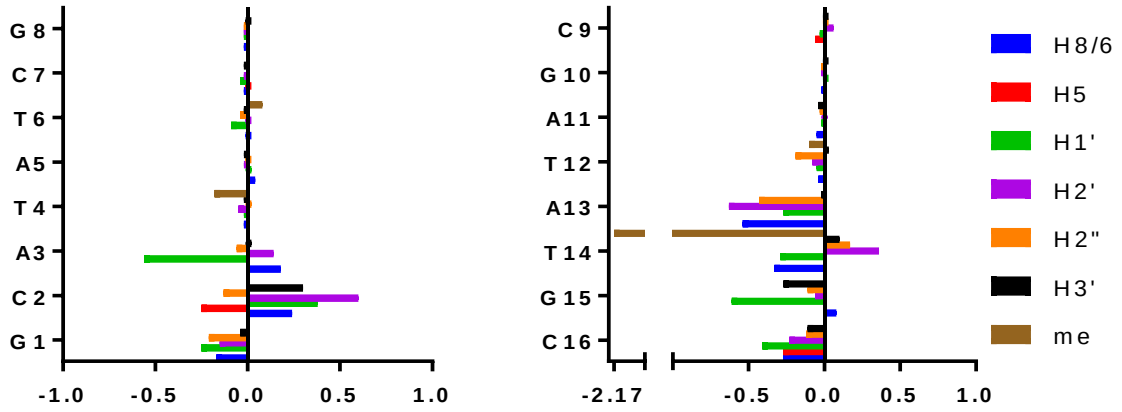


Figure 3.19. The chemical shift perturbations (δ/ppm) of the DNA duplex d($G_1C_2A_3T_4A_5T_6C_7G_8$).($C_9G_{10}A_{11}T_{12}A_{13}T_{14}G_{15}C_{16}$) upon binding of Λ,Λ-[1] in the minor conformation of Λ,Λ-[1]b.

Table 3.6. The changes in chemical shift (δ/ppm) of the major conformation of DNA duplex d(GCATATCG).(CGATATGC) upon addition of Λ,Λ-[1]a in D_2O at 298 K.

Λ,Λ-[1]a	H8	H6	H5	H1'	H2'	H2''	H3'	Me
G1	7.94	-	-	5.97	2.64	2.76	4.82	-
Δδ	-0.04	-	-	-0.04	0.01	-0.06	-0.05	-
C2	-	7.46	5.36	5.69	2.10	2.39	4.93	-
Δδ	-	-0.04	-0.10	-0.03	-0.10	-0.13	0.02	-
A3	7.98	-	-	6.09	2.33	2.74	4.95	-
Δδ	0.40	-	-	-0.23	-0.42	-0.26	-0.11	-
T4	-	7.07	-	5.53	2.49	2.42	5.06	-0.08
Δδ	-	-0.15	-	-0.19	0.33	-0.12	0.14	-1.61
A5	8.52	-	-	5.68	2.72	2.84	4.91	-
Δδ	0.27	-	-	-0.57	0.05	-0.09	-0.10	-
T6	-	7.11	-	5.86	1.96	2.42	4.84	1.23
Δδ	-	-0.12	-	-0.11	-0.07	-0.03	-0.02	-0.13
C7	-	7.50	5.66	5.70	2.03	2.39	4.84	-
Δδ	-	0.01	-0.01	-0.06	0.01	0.01	-0.01	-
G8	7.95	-	-	6.15	2.62	2.36	4.67	-
Δδ	-0.01	-	-	-0.03	-0.01	-0.03	-0.02	-
C9	-	7.61	5.85	5.73	1.95	2.39	4.70	-
Δδ	-	-0.02	-0.07	-0.04	0.03	-0.02	-0.02	-
G10	7.96	-	-	5.51	2.75	2.79	5.03	-
Δδ	-0.02	-	-	-0.09	-0.01	-0.06	0.00	-
A11	7.93	-	-	6.11	2.33	2.74	4.99	-
Δδ	-0.35	-	-	-0.20	-0.37	-0.27	-0.07	-
T12	-	7.10	-	5.56	2.48	2.34	5.06	0.12
Δδ	-	-0.07	-	-0.18	0.37	-0.19	0.17	-1.35
A13	8.51	-	-	5.64	2.73	2.83	4.93	-
Δδ	0.25	-	-	-0.61	0.10	-0.11	-0.09	-
T14	-	7.07	-	5.64	1.89	2.36	4.86	1.26
Δδ	-	-0.06	-	-0.15	-0.10	-0.04	-0.04	-0.12
G15	7.90	-	-	5.97	2.66	2.75	5.00	-
Δδ	0.03	-	-	0.02	0.02	-0.01	0.01	-
C16	-	7.49	5.51	6.18	2.17	2.17	4.50	-
Δδ	-	-0.02	0.01	-0.05	-0.02	-0.02	-0.01	-

Table 3.7. The changes in chemical shift (δ/ppm) of the minor conformation of DNA duplex d(GCATATCG).(CGATATGC) upon addition of Λ,Λ-[1]b in D_2O at 298 K.

Λ,Λ-[1]b	H8	H6	H5	H1'	H2'	H2"	H3'	Me
G1	7.81	-	-	5.76	2.48	2.61	4.83	-
Δδ	-0.17	-	-	-0.25	-0.15	-0.21	-0.04	-
C2	-	7.73	5.21	6.11	2.80	2.40	5.22	-
Δδ	-	0.24	-0.25	0.38	0.60	-0.13	0.30	-
A3	8.56	-	-	5.76	2.88	2.95	5.08	-
Δδ	0.18	-	-	-0.56	0.14	-0.06	0.02	-
T4	-	-	-	5.74	2.11	2.56	4.90	1.35
Δδ	-	-	-	0.01	-0.05	0.02	-0.02	-0.18
A5	8.30	-	-	6.27	2.68	2.92	5.02	-
Δδ	0.04	-	-	0.02	0.01	-0.01	0.01	-
T6	-	7.22	-	5.88	2.02	2.41	4.84	1.43
Δδ	-	-0.01	-	-0.09	-0.01	-0.04	-0.02	0.08
C7	-	7.47	5.66	5.72	1.99	2.38	4.83	-
Δδ	-	-0.02	-0.01	-0.04	-0.02	0.00	-0.02	-
G8	7.95	-	-	6.17	2.61	2.36	4.68	-
Δδ	-0.02	-	-	-0.02	-0.02	-0.02	-0.01	-
C9	-	7.64	5.86	5.73	1.98	2.40	4.70	-
Δδ	-	0.00	-0.06	-0.03	0.06	-0.01	-0.02	-
G10	7.96	-	-	5.59	2.72	2.83	5.02	-
Δδ	-0.02	-	-	-0.01	-0.03	-0.02	-0.01	-
A11	8.22	-	-	6.29	2.72	2.97	5.02	-
Δδ	-0.05	-	-	-0.02	0.02	-0.03	-0.04	-
T12	-	7.14	-	5.69	2.03	2.35	4.93	1.37
Δδ	-	-0.04	-	-0.05	-0.08	-0.19	0.03	-0.10
A13	7.72	-	-	5.98	1.99	2.51	5.00	-
Δδ	-0.54	-	-	-0.27	-0.63	-0.43	-0.02	-
T14	-	6.79	-	5.50	2.35	2.57	4.99	-0.79
Δδ	-	-0.33	-	-0.29	0.36	0.17	0.10	-2.17
G15	7.96	-	-	5.34	2.57	2.64	4.72	-
Δδ	0.08	-	-	-0.61	-0.07	-0.11	-0.27	-
C16	-	7.23	5.22	5.82	1.96	2.07	4.40	-
Δδ	-	-0.27	-0.27	-0.41	-0.23	-0.12	-0.11	-

3.5.1 Λ,Λ-[1]a NOEs

To gain further insight into the DNA binding of Λ,Λ-[1]a, NOESY spectra were inspected. For the Λ,Λ-[1]a-DNA complex, 30 new intermolecular NOEs are assignable from the DNA to the tpphz protons, with 23 out of the 30 also present for Λ,Δ-[1]a (table 3.8). These NOEs localize at the (T_4,A_5)|(T_{12},A_{13}) binding site and extend into both the major and minor grooves and from either end of the tpphz ligands (figure 3.20).

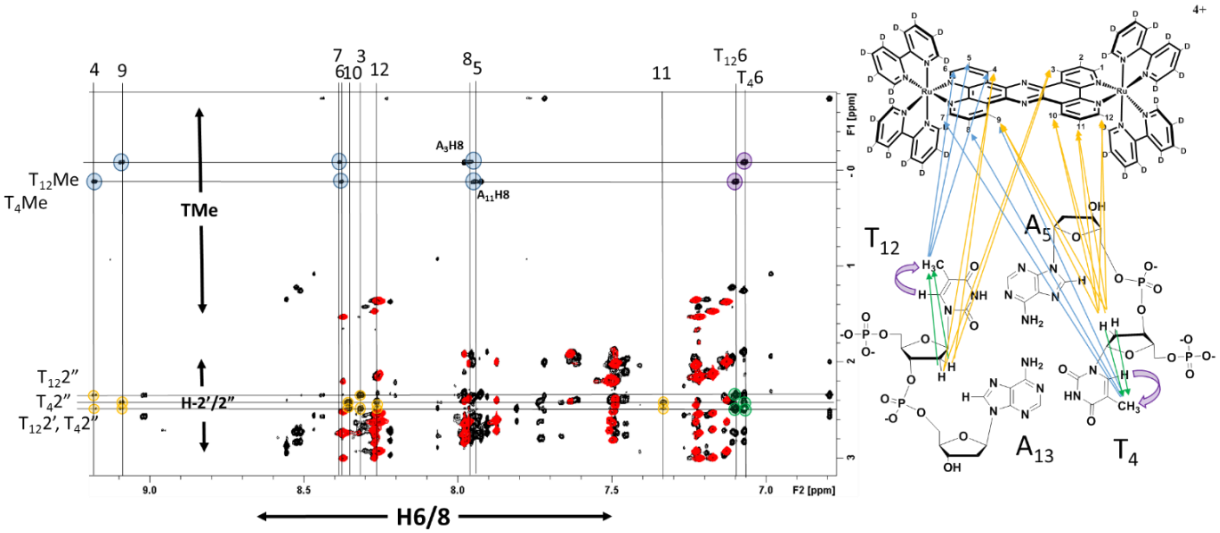


Figure 3.20. The expansion of the 2',2"and TMe-aromatic region of the NOESY spectrum of duplex sequence d(GCATATCG).(CGATATGC) in (red), overlayed onto a spectrum containing half an equivalent of Λ,Λ-[1] (black). The intramolecular NOE correlations H6/2',2" (green) and H6/TMe (purple) and the intermolecular NOE correlations Λ,Λ-[1]a/TMe (blue) and Λ,Λ-[1]a/H2',2" (orange) located in both grooves of the DNA. A schematic of the (T_4,A_{13})|(T_{12},A_5) binding site is given for reference.

Table 3.8. The intermolecular NOEs between the DNA duplex d(GCATATCG).(CGTATAGC) and Λ,Λ-[1]a in D_2O at 298 K

Λ,Λ-[1]a	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12
T4							Me	Me	Me,2',2"	1',2',2"	1',2',2"	1',2',2"
A5										1'	1'	1'
T12	1'	1'	1',2',2"	Me,2',2"	Me	Me						
A13	1'	1'	1'									

Binding of the metal complex again causes a break in the internucleotide 5'-3' NOE walk between the base pairs (T_4,A_{13})|(T_{12},A_5). This indicates that the binding of the complex causes the separation of T_4|A_5 and T_{12}|A_{13} at this site, as

was seen for the Λ,Λ-[1]a:DNA complex. It is also possible to use internucleotide NOEs from d-Λ,Λ-[1]a to complete the sequential assignment (figure 3.21). This is a very strong indication of intercalation with Λ,Λ-[1] residing in between the base pairs causing a lengthening of the DNA at this site.

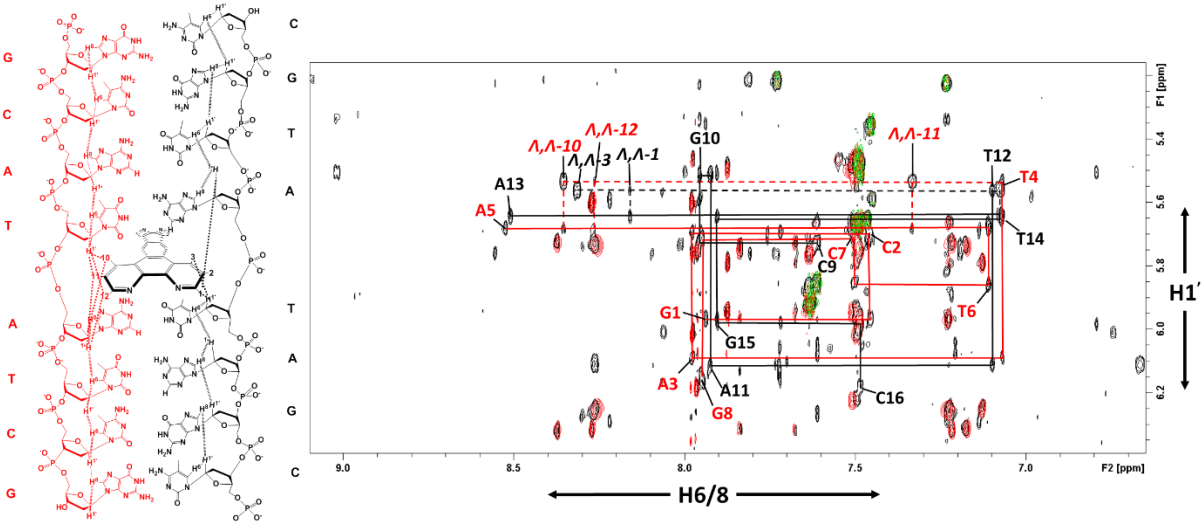


Figure 3.21. Expansion of the H1'-H6/8 region of the NOESY and COSY spectrum of a d($G_1C_2A_3T_4A_5T_6C_7G_8$).($C_9G_{10}A_{11}T_{12}A_{13}T_{14}G_{15}C_{16}$) DNA/Λ,Λ complex. The solid lines show the DNA 1'-H6/8 internucleotide NOE walk. Dashed lines show NOE connectivities between Λ,Λ-[1]a and the DNA 1' protons. The overlaid NMR spectra show the NOESY spectra of 2 mM DNA (red) 2 mM DNA + 1 mM Λ,Λ-[1] (black) and the COSY spectrum 2 mM DNA + 1 mM d-Λ,Λ-1 (green).

3.5.2 Λ,Λ-[1]b NOEs

The minor bound species, Λ,Λ-[1]b, exhibits 24 assignable intermolecular NOEs (figure 3.22). Unlike the other binding mode, the NOEs reside at one end of the DNA duplex with connectivities to C_2,A_3,T_{14} and G_{15} with a similar pattern to that seen for the Λ,Λ-[1]a species, extending into both grooves of the DNA. Contrasting to the NOE walk of the major species, a break occurs at C_2|A_3 and T_{14}|G_{15} with the sequential assignment being completed through close contacts to the Λ,Λ-[1]b protons. This indicates that the ligand is stacking between the base pairs and increasing the internucleotide separation at the (C_2,G_{15})|(T_{14},A_3) step.

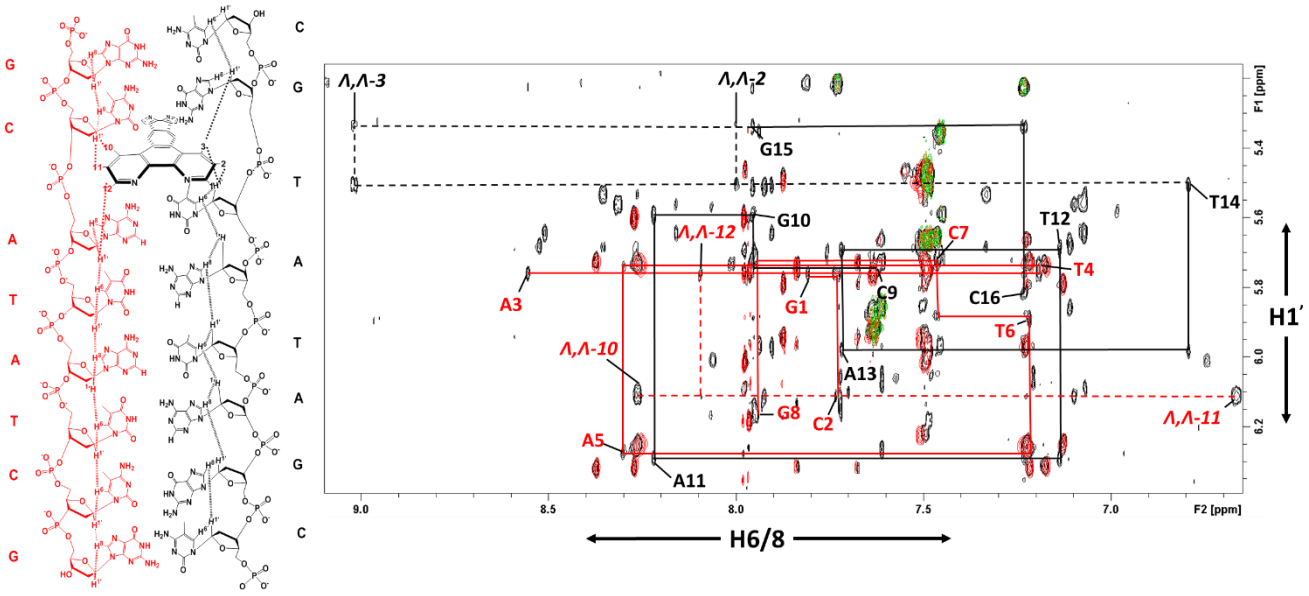


Figure 3.22. Expansion of the H1'-H6/8 region of the NOESY and COSY spectrum of a d($G_1C_2A_3T_4A_5T_6C_7G_8$).($C_9G_{10}A_{11}T_{12}A_{13}T_{14}G_{15}C_{16}$) DNA/Λ,Λ complex. The solid lines show the DNA 1'-H6/8 internucleotide NOE walk. Dashed lines show NOE connectivities between Λ,Λ-[1]b and the DNA 1' protons. The overlaid NMR spectra show the NOESY spectra of 2 mM DNA (red) 2 mM DNA + 1 mM Λ,Λ-[1] (black) and the COSY spectrum 2 mM DNA + 1 mM Λ,Λ-[1] (green).

The NOE network seen for the DNA:Λ,Λ-[1]b conformation is shown in figure 3.23 with the complete table of observable intermolecular NOEs shown in table 3.9.

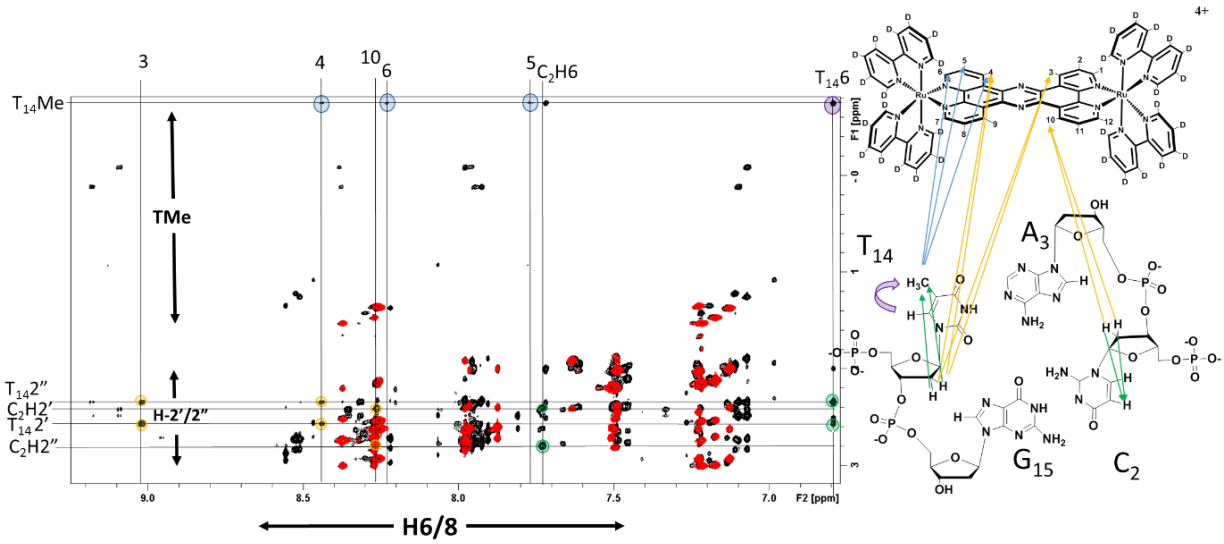


Figure 3.39. The expansion of the 2',2"and TMe-aromatic region of the NOESY spectrum of duplex sequence d(GCATATCG).(CGATATGC) in (red), overlayed onto a spectrum containing half an equivalent of Λ,Λ-[1] (black). The intramolecular NOE correlations H6/2',2" (green) and H6/TMe (purple) and the intermolecular NOE correlations Λ,Λ-[1]b/TMe (blue) and Λ,Λ-[1]b/H2',2" (yellow) located in both grooves of the DNA. A schematic of the (T_4,A_{13}) | (T_{12},A_5) binding site is given for reference

Table 3.9. The intermolecular NOEs between the DNA duplex d(GCATATCG).(CGTATAGC) and Λ,Λ-[1]b

Λ,Λ-[1]b	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12
C2							H5	H5	H5,1'	1',2',2"	1',2',2"	1'
A3										1'	1'	1'
T14			1',2',2"	Me,2',2"	Me	Me						
G15		1'	1'									

3.5.3 Proposed Binding Modes of Λ,Λ-[1]

The proposed binding mode of Λ,Λ-[1] to oligonucleotide sequence d(GCATATCG).(GCATATGC) is outlined in figure 3.24. It shows two threaded conformations populated in a 2:1 ratio, in slow exchange. This is in contrast to the binding mode of [1] to genomic CT-DNA which interacts entirely through groove binding.[24] Lincoln, *et al*. have shown that shortening of the bridging ligands in dinuclear ruthenium polypyridyl complexes slows threading to CT-DNA.[25,26] With even shorter bridges, threading is inhibited entirely and threading is only possible in the presence of more flexible poly d(AT) DNA.[27] The tpphz bridge in [1] is shorter than the dppzip ligand discussed in the study by Lincoln, *et al*. and therefore it is reasonable to predict that threading would only be possible in very flexible or unstable sequences, such as this short oligonucleotide.

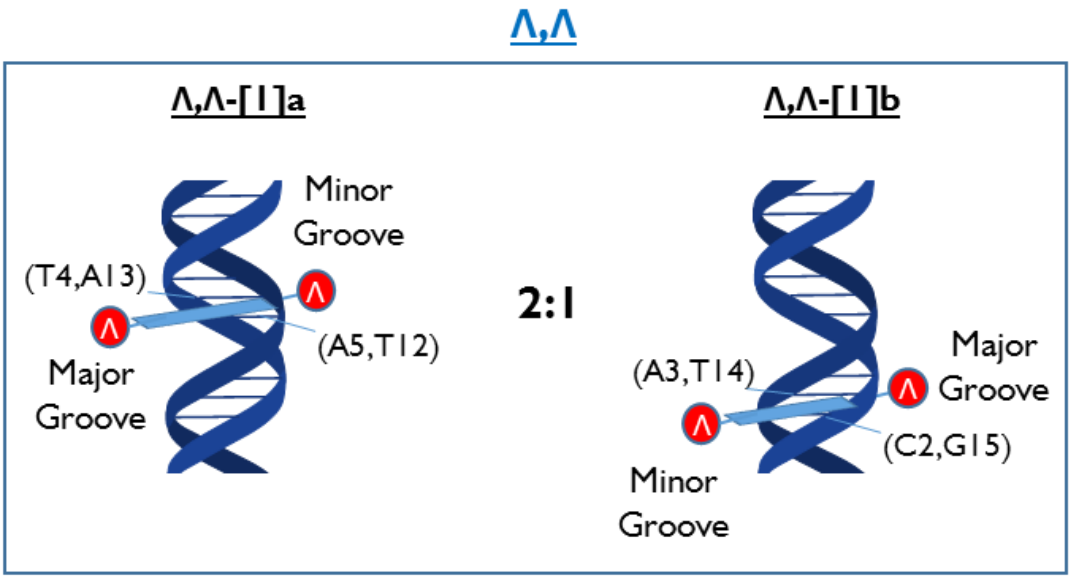


Figure 3.24. A schematic of the proposed binding modes of Λ,Λ-[1]

A comparison of the intercalation sites occupied by Λ,Λ-[1] can be carried out by looking at dinucleotide step parameters which have been calculated and compared to experimental X-ray data for all possible step combinations by Packer, *et al.* (figure 3.25).[28] Looking at the TA/TA and CA/TG intercalation sites, it has been shown that the CA/TG step ($\boldsymbol{\phi}_{slide}$ 1.35 kJ mol^{-1} Å^{-2}) is the most flexible base step and the TA/TA step is the fourth most flexible ($\boldsymbol{\phi}_{slide}$ 4.34 kJ mol^{-1} Å^{-2}). TA/TA and CA/TG have the largest and second largest roll at 8.4 ° and 7.4 ° respectively. CA/TG has the smallest roll and the third smallest twist of all sequences, 32.2° and 34.6° respectively. This information coupled with the observed intercalated species Λ,Λ-[1]b and Λ,Λ-[1]a suggests a large roll, small twist and high flexibility dinucleotide step is needed for intercalation to occur. Interestingly, an AT/AT base step is classified as rigid ($\boldsymbol{\phi}_{slide}$ 16.69 kJ mol^{-1} Å^{-2}) with the most negative roll (-8.1 °) and largest twist (39.7 °) of any dinucleotide step. This reinforces the suggested binding site of the Λ,Λ-[1]b being at the central TA/TA site and not AT/AT.

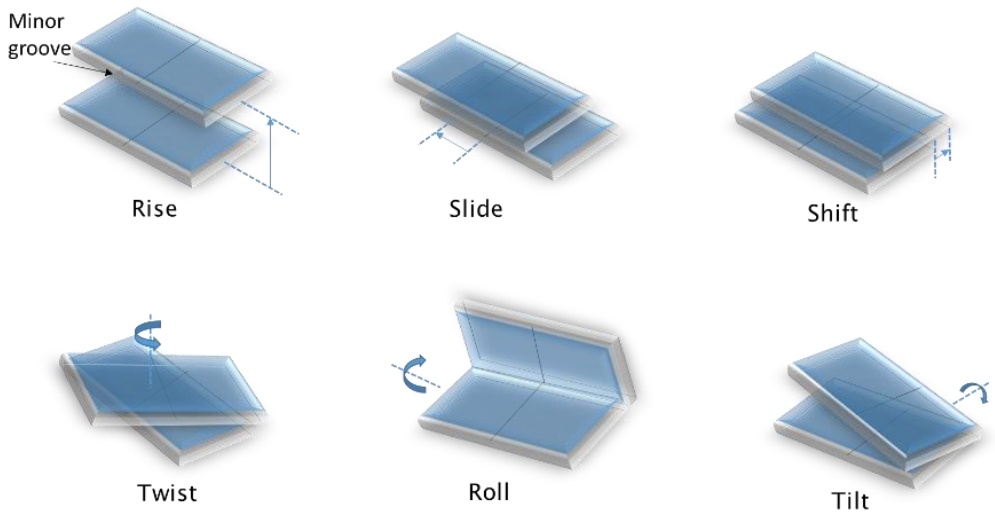


Figure 3.25. A schematic display of the six independent dinucleotide step parameters as viewed from the minor groove adapted from reference 130. [29]

3.6 Addition of d_{32}- Δ,Δ -[1] to the Oligonucleotide

Upon addition of Δ,Δ-1 to the DNA duplex, two sets of broadened signals in slow to intermediate chemical exchange on the NMR timescale are observed. The two sets of DNA resonances are equally populated and show exchange doublets in the NOESY spectrum characterised by off-diagonal peaks correlating to the two species.[30] Attempts to push the exchange into a slow exchange regime by lowering the temperature to 283 K were unsuccessful, with the chemical exchange still occurring in the NOESY spectrum at all temperatures. The clearest set of exchange doublets can be seen for the thymidine methyl (TMe) residues, as they are upfield in an uncrowded region of the spectrum (figure 3.26). Whilst the spectrum is generally unassignable due to spectral overcrowding from the multiple exchange peaks, the TMe residues can be assigned from the TMe-H6 J-coupled spin systems in the TOCSY and COSY spectrum. Intermolecular NOEs between TMe protons and Δ,Δ-[1] exist in one of the two conformations similar to those seen for Λ,Δ-[1]a (figure 3.26). There are also large chemical shift differences between the two bound conformations, with one set of TMe resonances shifted up field by at least 0.6 ppm. The TMe protons are located in the major groove of the DNA duplex, and one of the two conformations has largely shifted TMe residues and

NOE close contacts to the Δ,Δ-[1] ligand indicating a threaded state. The other conformation has comparatively small chemical shifts to the TMe protons which could suggest binding of Δ,Δ-[1] in the minor groove of the DNA. It has been previously reported that binuclear ruthenium complexes can associate with DNA in its minor groove and this becomes the starting point for the intercalative threading at AT residues.[31] This dynamic exchange behaviour suggests that the threading process is occurring with Δ,Δ-[1] translating from the minor groove through to the threaded state. This will be discussed further in the next chapter. A comparison of the chemical shifts of the Thymidine methyl protons from the intercalated conformation (Δ,Δ-[1]a), to that of the major conformation of the Λ,Δ-[1]a, reveals many similarities that suggest a similar intercalated binding mode.

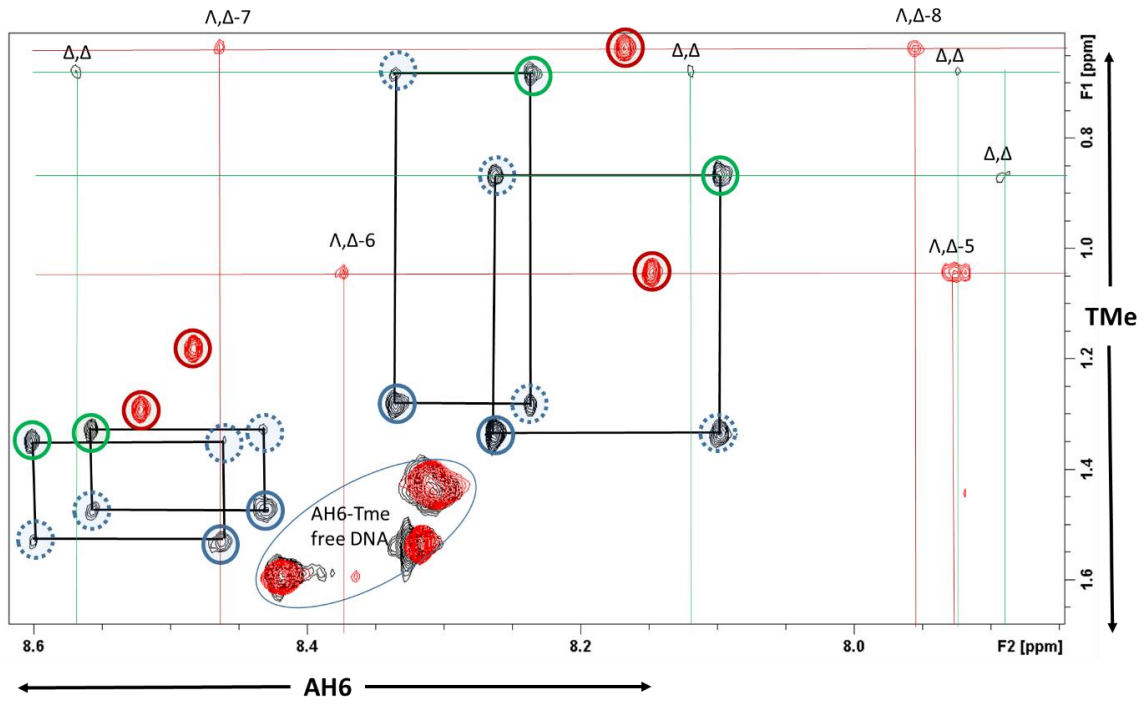


Figure 3.26. An expansion of the NOESY NMR spectra showing the AH6-TMe region of a solution containing 2 mM d(GCATATCG).(CGATATGC) and 1 mM Δ,Δ-[1] in D_2O (black), the DNA:Λ,Δ-[1] spectrum is overlaid for comparison (red). The circles outline AH6-TMe internucleotide couplings with the green circles show Δ,Δ-[1]a (intercalated) conformation, the blue circles show the Δ,Δ-[1]b form (minor groove) and the red circles show the Λ,Δ-[1]a for comparison. The dashed blue circles are the exchange peaks seen between the two conformations, green and red lines indicate the Δ,Δ-[1]a and Δ,Δ-[1]b intermolecular NOE networks respectively.

3.6.1 Proposed Binding Modes of Δ,Δ-[1]

The proposed binding mode of Δ,Δ-[1] to oligonucleotide sequence d(GCATATCG).(GCATATGC) is a dynamic equilibrium between species in the minor groove in a slow-intermediate exchange regime with a threaded complex. This is outlined in figure 3.27.

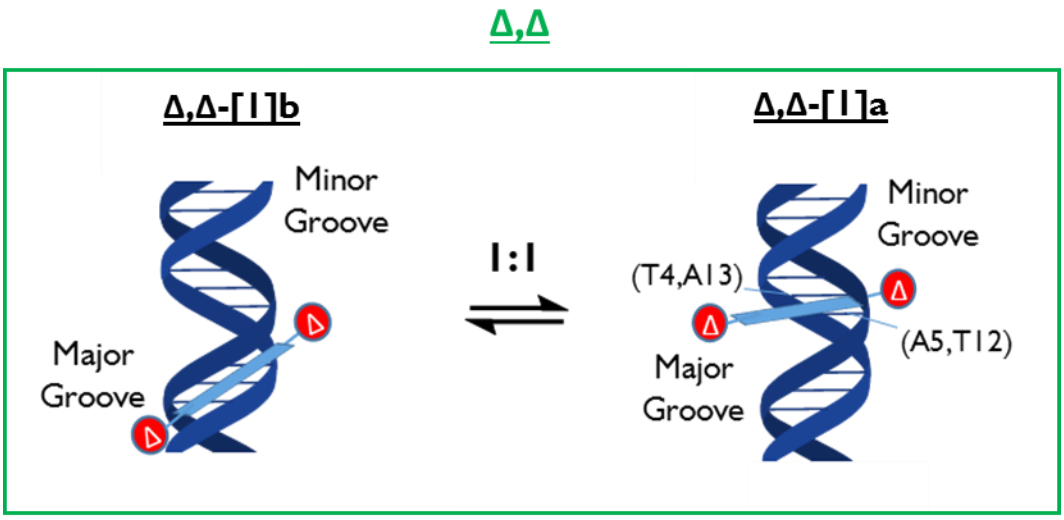


Figure 3.27. A schematic of the proposed binding mode of Δ,Δ-[1].

In the model reported by Lincoln, *et al.* for a related dinuclear ruthenium complex, threading is proposed to occur by initial association in the minor groove of the DNA at long AT tracts. This association in the minor groove is more favourable due to the higher electronegative potentials of AT base pairs attracting the cationic units.[32] The threading itself then requires a large deformation of the DNA in order for the bulky ruthenium centre to thread through the DNA. The two binding modes in exchange, suggest this process is happening for the Δ,Δ-[1] complex during the mixing time of the NOESY experiment. The ability for this to occur is probably due to the high flexibility of this short DNA sequence. It can be predicted that the Δ,Δ-[1] isomer has a lower threading activation energy to that of the other stereoisomers, as has been demonstrated in the related complexes by Lincoln, *et al.* This lower activation energy can be interpreted by the exchange behaviour of [1]. The 1:1 populations seen in both conformations also suggests that Δ,Δ-[1]a and Δ,Δ-[1]b are equally stable, in contrast to Λ,Λ-[1] and Λ,Δ-[1], which display a strong preference for an intercalated structure. Moreover, the exchange

rate for reversible intercalation of Δ,Δ-[1] is faster than the other stereoisomers implying that this difference is due to an energetically less favourable intercalated structure. In other words, combined, the two Δ units provide greater steric hindrance and thus a less favourable fit into the DNA grooves when intercalated.

3.7 ^{31}P NMR Studies

Phosphorous NMR experiments have proven a useful tool in the conformational assignment of DNA, as the chemical shift of the phosphate backbone is very sensitive to changes in backbone torsional angles. As a result, the chemical shifts of backbone phosphates can probe conformational changes in DNA structure.[33] The chemical shift of a ^{31}P nucleus is most affected by stereoelectronic effects of rotation about the phosphodiester bonds α (P-O5') and ζ (PO-3') angles in (figure 3.28).[34] The chemical shifts of the nuclei are shielded in the most energetically favourable α-ζ gauche-gauche (± 60 °) conformation resonating several ppm upfield from non-gauche conformations.[35]

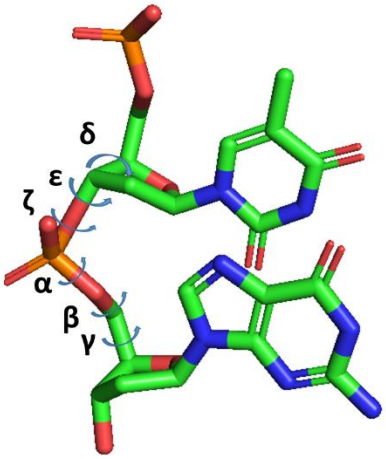


Figure 3.28. The backbone torsional angles of a polynucleotide.

More recently this ^{31}P sensitivity to backbone torsional angles has been applied to the detection of intercalation of small molecules into B-DNA.[36,37] Intercalation results in elongation of the DNA upon insertion of the molecule into the base stack. Rotation of the phosphodiester bonds accommodates this separation, which results in a downfield shift at the intercalation site.[11,38]

^{31}P-NMR experiments were carried out on the free d(GCATATCG).(CGATATGC) duplex and solutions containing 2:1 DNA:[1], with the spectra for the three stereoisomers of [1] given in figure 3.29. Upon addition of the Λ,Λ-[1] and Λ,Δ-[1], a downfield change in chemical shift is apparent (1.86 ppm). Observation of the Λ,Λ-[1]:DNA complex reveals four new peaks downfield with a ratio of 2:1. The Δ,Δ:DNA complex is again less resolvable but from a cursory inspection downfield peaks are observed which indicates that at least one of the conformations has the Δ,Δ-[1] intercalated into the DNA.

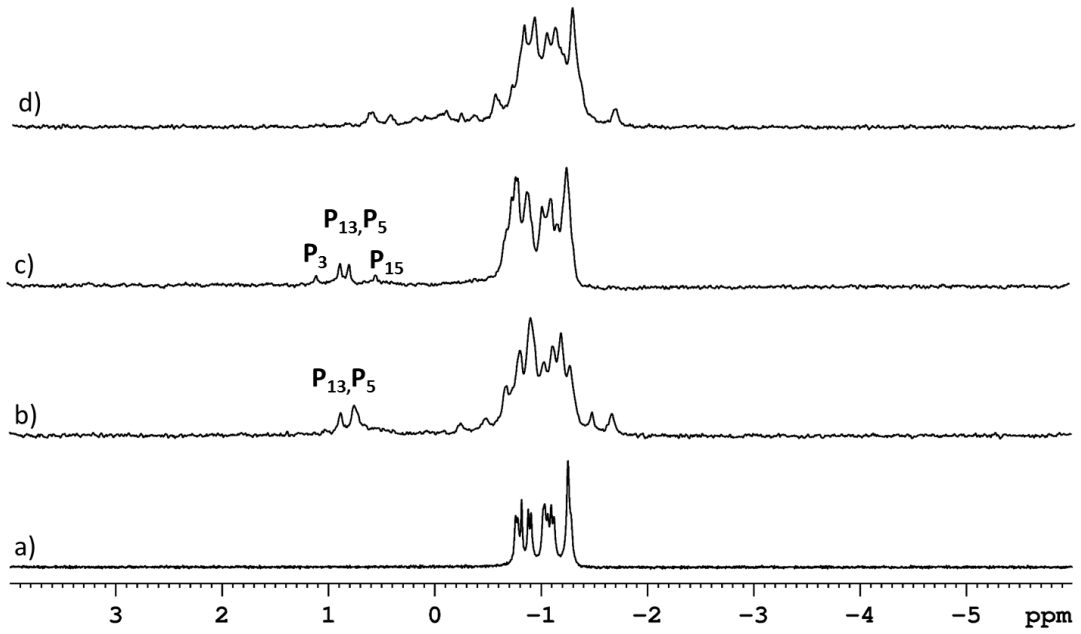


Figure 3.29. Spectra showing the proton decoupled 1D ^{31}P spectrum of **a)** 2 mM oligonucleotide sequence d(GCATATCG).(CGATATGC) in D_2O containing b**)** Λ,Δ-[1] **c)** Λ,Λ-[1] **d)** Δ,Δ-[1], assignments of b) and c) are deduced from their respective ^{31}P-^{1}H COSY spectra.

In order to further inspect the ^{31}P spectra of the [1]:DNA complexes, assignment was carried out using 2D ^{31}P-^{1}H heteronuclear correlation experiments. ^{31}P-^{1}H COSY experiments were first used by Pardi, et *al*. in 1983 for the sequential assignments of backbone phosphates in short oligonucleotide sequences.[39] In these spectra, three bond couplings between the phosphate and H3', 5' and 5" sugar protons give cross peaks in the 2D spectrum. The assignment was aided by comparison of the previously assigned homonuclear ^{1}H-spectra. In the case of these experiments some of the 5',5" couplings could not be resolved due to significant overlap in the proton dimension. However, it is possible to assign the

P-H3'(i-1) couplings, with the assignment of the free DNA oligonucleotide shown in figure 3.30.

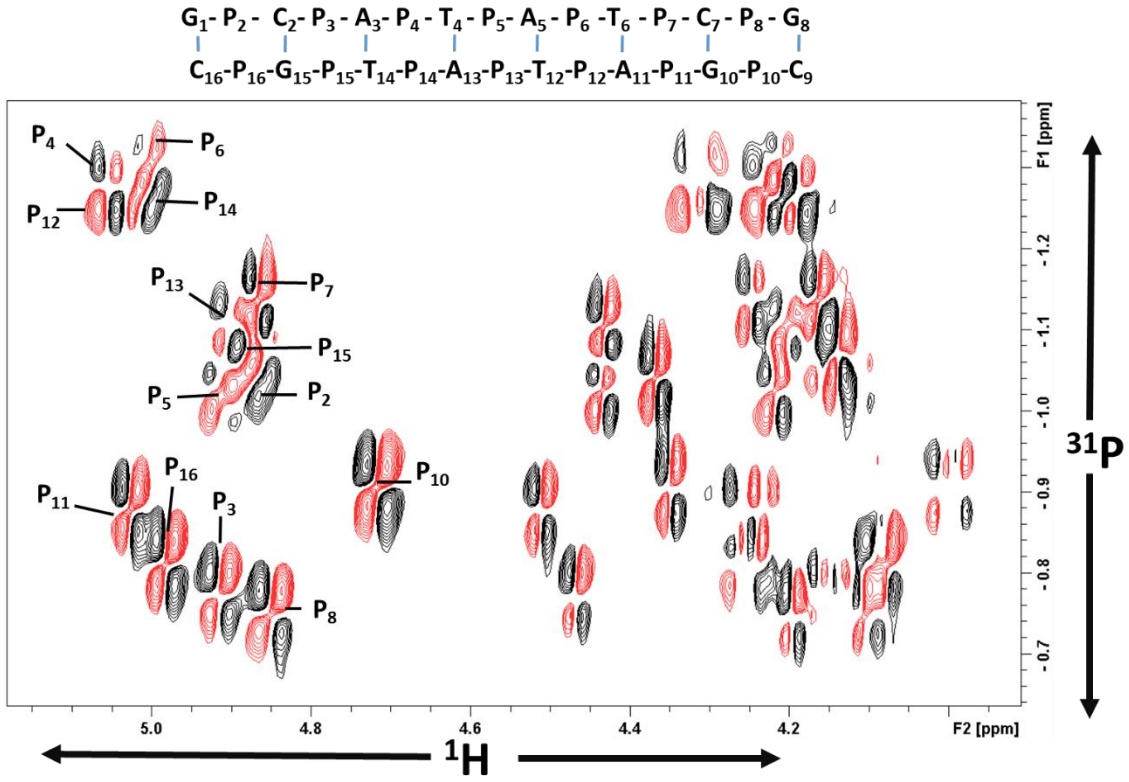


Figure 3.30. The ^{31}P-^{1}H COSY spectrum of 2 mM oligonucleotide sequence (GCATATCG) (CGATATGC) in D_2O showing the P-H3'(i-1) couplings, 5' and 5" couplings have been unassigned due to spectral overcrowding in this region.

Despite the fact that many of the phosphate residues could not be assigned due to spectral crowding, ^{31}P-^{1}H COSY experiments on the Λ,Λ-[1]:DNA complex revealed a very interesting observation. Although the major peaks of Λ,Λ-[1]a were overlapping in the 2D COSY, the downfield shifted P-H3'(i-1) couplings were dispersed enough to be assigned in the 1D spectrum. The peaks indicate the presence of two ^{31}P nuclei coupled to the H3' protons of T4 (5.06 ppm) and T12 (5.06 ppm) which can be assigned as P_5 and P_{13} respectively (figure 3.31).

The minor peaks of Λ,Λ-[1]b are coupled to the 3' protons of C_2 (5.22 ppm) and T_{14} (4.99 ppm) which correlate to P_3 and P_{15}. This is in agreement with the suggested intercalation sites identified from the homonuclear experiments.

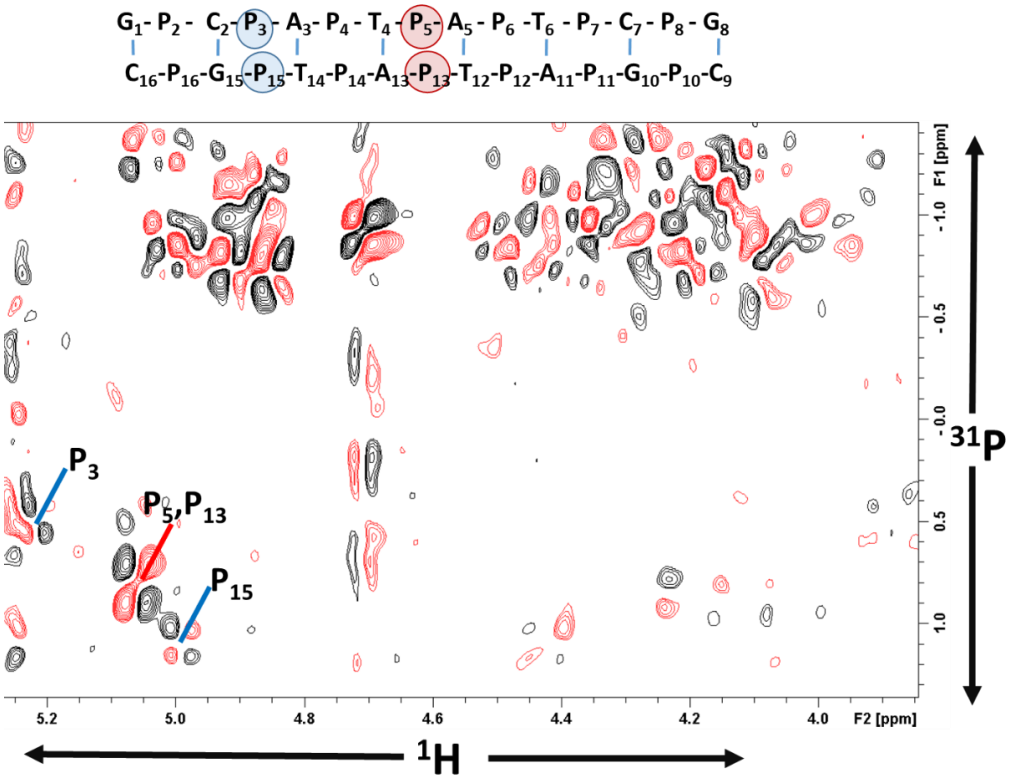


Figure 3.31. The ^{31}P-^{1}H COSY spectrum of 2 mM oligonucleotide sequence (GCATATCG) (CGATATGC) in D_2O containing 1 mM Λ,Λ-[1]. The spectrum indicates the largely shifted P-H3'(i+1) couplings with the blue and red lines denoting the major (Λ,Λ-[1]a) and minor (Λ,Λ-[1]b) forms respectively.

^{31}P-^{1}H COSY experiments on the Λ,Δ-[1]:DNA complex also show downfield shifted P-H3'(i-1) cross peaks which were dispersed enough to be assigned. The peaks were coupled to the H3' protons of T4 (5.11 ppm) and T12 (5.14 ppm), which correlates with P_5 and P_{13} in figure 3.32. The peaks from the minor species Λ,Δ -[1]b cannot be assigned due to the lack of clear chemical shift assignment in the ^{1}H homonuclear spectra. But the large chemical shift difference (1.86 ppm in table 3.10) in the phosphorus spectrum is almost identical to that seen for the major species, which suggests a similar conformational change at the phosphate residues and thus a similar intercalative binding mode.

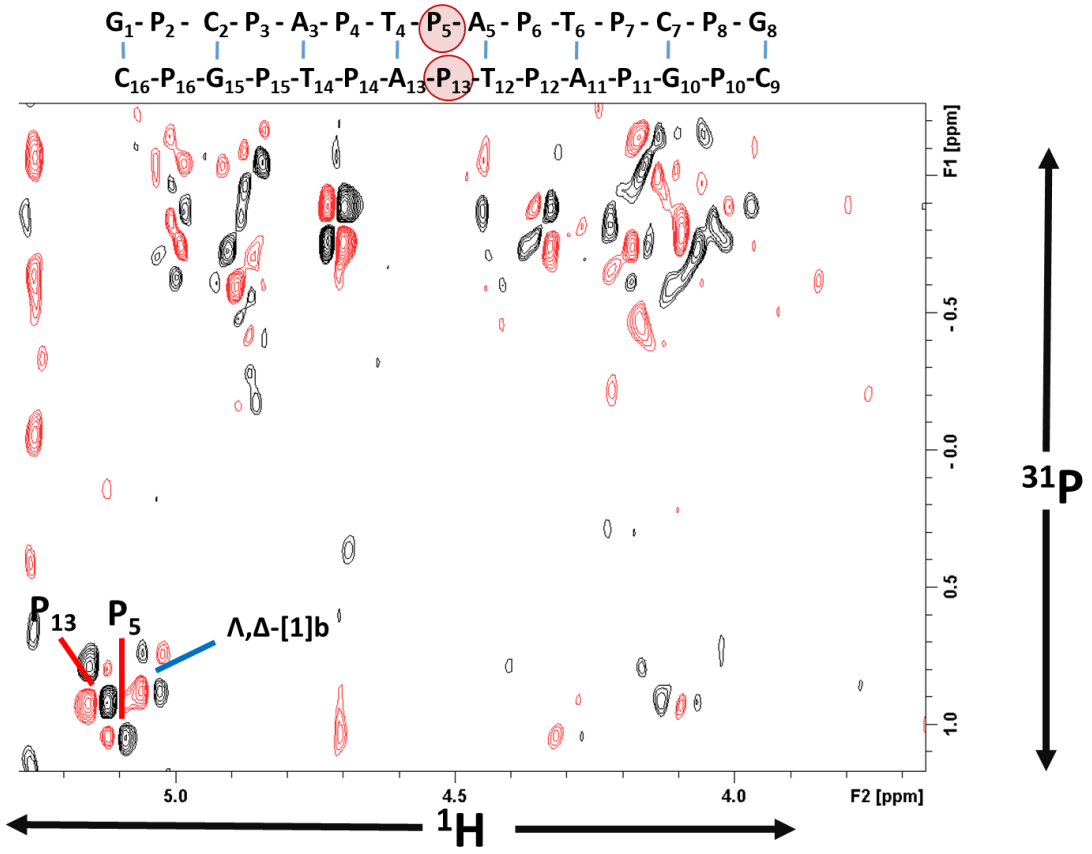


Figure 3.32. The ^{31}P-^{1}H COSY spectrum of 2 mM oligonucleotide sequence d(GCATATCG).(CGATATGC) in D_2O containing 1 mM Λ,Λ-[1]. The spectrum indicates the largely shifted P-H3'(i+1) couplings with the blue and red lines denoting the major (Λ,Δ-[1]a) and minor (Λ,Δ-[1]b) forms respectively.

Table 3.10. The chemical shift δ of the phosphorus nuclei of a 2 mM solution in D_2O d(GCATATCG).(CGATATGC) upon addition of the 3 stereoisomers of [1].

	P2	P3	P4	P5	P6	P7	P8	P10	P11	P12	P13	P14	P15	P16
Free oligo δ	-1.03	-0.78	-1.25	-1.01	-1.28	-1.13	-0.76	-0.9	-0.88	-1.25	-1.09	-1.25	-0.81	-0.88
Λ,Λ-[1]b		0.54											1.095	
Δδ		1.32											1.905	
Λ,Λ-[1]a				0.83							0.83			
Δδ				1.84							1.92			
Λ,Δ-[1]a				0.83							0.83			
Δδ				1.84							1.92			

3.8 Initial NMR Studies of $[Ru_2(phen)_2(tpphz)]$ [2]

Initial binding studies were carried out with racemic non-deuterated [2]. Further studies were not carried out due to time constraints and the lack of available (d8-1,10-phenanthroline). But an initial inspection of the NOESY spectrum in figure 3.33 shows a comparable chemical shift profile to that of [1], with largely shifted thymine methyls and tracks of what appear to be intermolecular NOEs indicative of threading. Interestingly, upon initial inspection there does not appear to be any exchange doublets occurring as seen in the Δ,Δ-[1], which would suggest that a more bulky phen ancillary ligand slows the rate of threading. Further NMR experiments using resolved deuterated complexes are needed in order to investigate this phenomenon further.

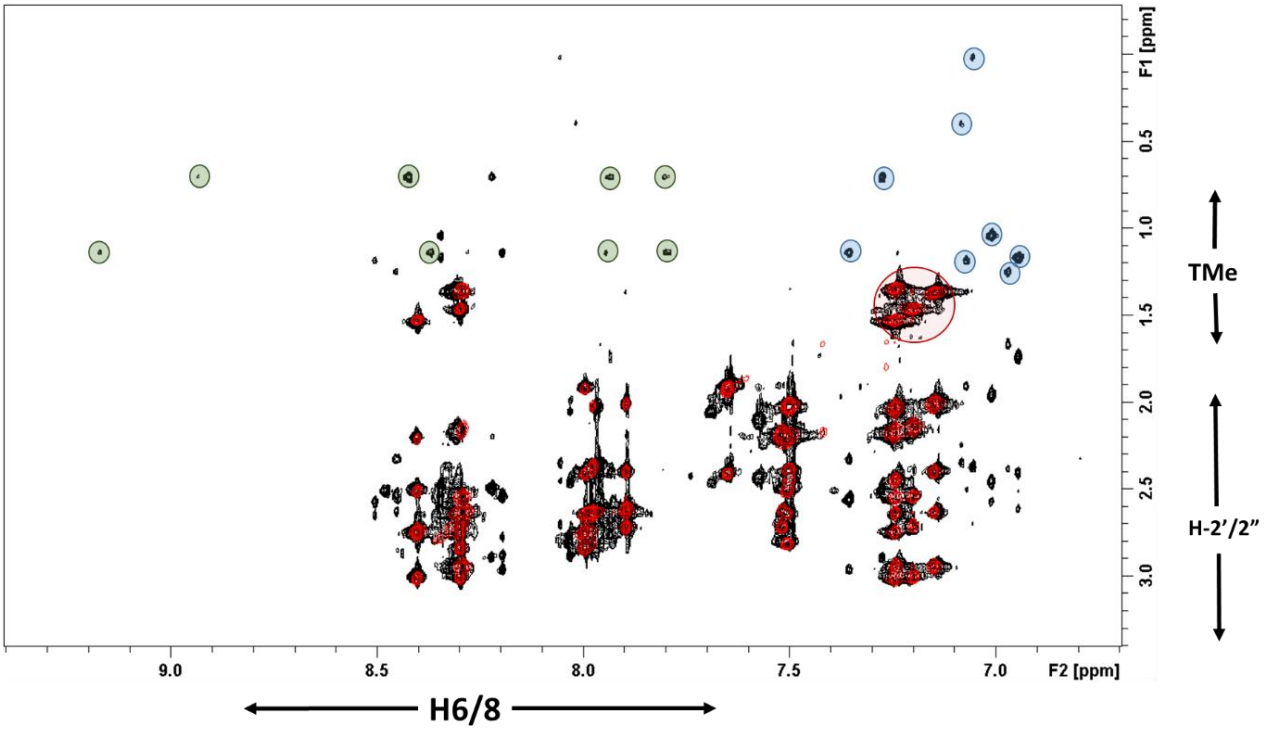


Figure 3.33. The expansion of the 2',2"and TMe-aromatic region of the NOESY spectrum of a 2 mM D_2O solution containing the duplex sequence d(GCATATCG).(CGATATGC) in (red), overlayed onto a spectrum containing half an equivalent of racemic non-deuterated [2] (black). The red circle shows the H6-TMe peaks of the free DNA, the blue circles indicate the H6-TMe peaks for the DNA:[2] complex green circles indicate possible [2]-TMe intermolecular NOEs.

3.9 Discussion

The data from the NMR studies suggest that all three stereoisomers of [1] have different binding preferences when added to the oligonucleotide sequence d(GCATATCG).(CGATATGC). Λ,Λ-[1] shows two bound conformations, Λ,Λ-[1]a, and Λ,Λ-[1]b, which are stable in slow exchange on the NMR timescale. They are threaded at two sites with the major form at the (T_4A_5)/($T_{12}A_{13}$) base step and the minor form at the (C_2A_3)/($T_{14}G_{15}$) base step, with no detectable binding elsewhere. The Λ,Λ-[1] also exhibits two bound conformations in slow exchange with the Λ,Δ-[1]a major species intercalated at the (T_4A_5)/($T_{12}A_{13}$) base step. The Λ,Δ-[1]b conformation is proposed to bind at the same site, but with the Λ and Δ ends interchanged. The Δ,Δ-[1] : DNA complex is shown to be in two conformations. These are suggested to exhibit an equally populated threaded form (Δ,Δ-[1]a) and minor groove-bound form (Δ,Δ-[1]b) in a slow-intermediate exchange regime on the NMR timescale. The proposed binding modes of the complexes are shown in figure 3.34.

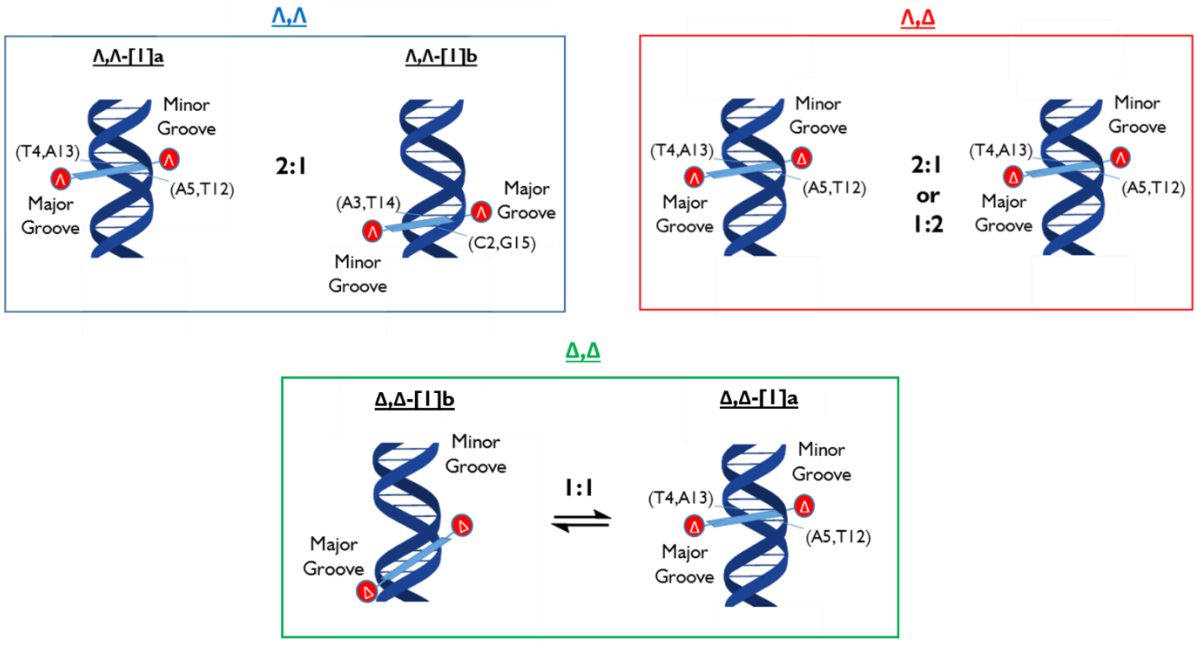


Figure 3.34. Proposed binding modes of the three stereoisomers of [1].

Although [1] interacts with CT-DNA entirely through groove binding,[24] these studies indicate that [1] intercalates into an octanucleotide fragment

through a threading interaction. As discussed previously, it has been shown that for dinuclear ruthenium complexes linked by a bridging ligand, the length and rigidity of the bridge play an important role in the selectivity and kinetics of threading. Shorter and more rigid bridges cannot thread into calf thymus DNA, but threading through more flexible sequences such as poly d(AT) DNA can still occur.[23,27] It has been suggested that as the bridging moieties become shorter it becomes progressively less favourable to have a Δ ruthenium centre in the major groove.[40] The tpphz bridging moiety in [1] is far shorter and more rigid than previously studied threading complexes, where the curvature of the molecule matches the pitch of the DNA.[31] This means that in the groove bound state of [1], unfavourable steric clashes from a Λ configured metal centre, to the right handed DNA duplex would be increased; this is consistent with the findings reported in this chapter, as it appears Δ,Δ-[1] displays a better binding fit in the groove than the other two stereoisomers. The Δ,Δ complex also shows faster kinetics than the Λ,Λ and Λ,Δ forms, with exchange occurring on the NMR timescale which suggests that the Λ,Λ and Λ,Δ are kinetically locked into the threaded mode, as well as thermodynamically less favourable in the groove.

Slow kinetics is typical of threading intercalation as the bulky ruthenium centres need to pass through the duplex in order to dissociate. This is reflected in the Λ,Λ and Λ,Δ isomers. However, the proposed initial groove bound state is not observed in this study. In genomic DNA, threading has been shown to occur on a very slow timescale, sometimes in the order of weeks at room temperature.[41] It is normally preceded by an initial groove bound state as reported in the NMR study by Lincoln, *et al.* discussed in chapter 1.[42] The DNA oligonucleotide sequence used in this study is of low stability and is mostly single stranded at lower concentrations. It is proposed that the nucleotide sequence can thus form a lowly populated single stranded state which can anneal around [1] and reach thermodynamic equilibrium before any NMR experiments are carried out. Therefore, the systems are in their most thermodynamically stable forms.

3.10 References

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

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

3 W. H. Bragg and W. L. Bragg, *Proc. R. Soc. A Math. Phys. Eng. Sci.*, 1913, **88**, 428–438.

4 S. Arnott and D. W. Hukins, *J. Mol. Biol.*, 1973, **81**, 93–105.

5 G. Wagner, S. G. Hyberts and T. F. Havel, *Annu. Rev. Biophys. Biomol. Struct.*, 1992, **21**, 167–198.

6 L. Krause, R. Herbst-Irmer, G. M. Sheldrick and D. Stalke, *J. Appl. Crystallogr.*, 2015, **48**, 3–10.

7 V. Tereshko, G. Minasov and M. Egli, *J.Am.Chem.Soc.*, 1999, **121**, 470–471.

8 O. Zerbe, *BioNMR in Drug Research*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, FRG, 2002.

9 B. Coimbatore Narayanan, J. Westbrook, S. Ghosh, A. I. Petrov, B. Sweeney, C. L. Zirbel, N. B. Leontis and H. M. Berman, *Nucleic Acids Res.*, 2014, **42**, D114–D122.

10 A. Gelasco and S. J. Lippard, *Biochemistry*, 1998, **37**, 9230–9239.

11 C. M. Dupureur and J. K. Barton, *Inorg. Chem.*, 1997, **36**, 33–43.

12 R. Owczarzy, P. M. Vallone, F. J. Gallo, T. M. Paner, M. J. Lane and A. S. Benight, *Biopolymers*, 1997, **44**, 217–39.

13 K. J. Breslauer, R. Frank, H. Blöcker and L. A. Marky, *Proc. Natl. Acad. Sci. U. S. A.*, 1986, **83**, 3746–50.

14 E. L. Ulrich, H. Akutsu, J. F. Doreleijers, Y. Harano, Y. E. Ioannidis, J. Lin, M. Livny, S. Mading, D. Maziuk, Z. Miller, E. Nakatani, C. F. Schulte, D. E. Tolmie, R. Kent Wenger, H. Yao and J. L. Markley, *Nucleic Acids Res.*, 2007, **36**, D402–D408.

15 E. L. Ulrich, E. M. M. John, G. R. Gough, M. J. Brunden, P. T. Gilham, W. M. Westler and J. L. Markley, *Biochemistry*, 1983, **22**, 4362–4365.

16 D. J. Patel, L. Shapiros and D. Hare, *Journal*, 1986, **261**, 1223–1229.

17 F. J. M. van de Ven and C. W. Hilbers, *Nucleic Acids Res.*, 1988, **16**, 5713–5726.

18 R. M. Scheek, R. Boelens, N. Russo, J. H. Van Boom and R. Kaptein, *Biochemistry*, 1984, **23**, 1371–1376.

19 J. C. Lindon, G. E. Tranter and J. L. Holmes, Academic Press, 1st edition., 2000.

20 C. M. Dupureur and J. K. Barton, *J. Am. Chem. Soc.*, 1994, **116**, 10286–10287.

21 S. S. David and J. K. Barton, *J. Am. Chem. Soc.*, 1993, **115**, 2984–2985.

22 H. Niyazi, J. P. Hall, K. O'Sullivan, G. Winter, T. Sorensen, J. M. Kelly and C. J. Cardin, *Nat. Chem.*, 2012, **4**, 621–628.

23 P. Nordell, F. Westerlund, L. M. Wilhelmsson, B. Nordén and P. Lincoln, *Angew. Chemie*, 2007, **119**, 2253–2256.

24 D. A. Lutterman, A. Chouai, Y. Liu, Y. Sun, C. D. Stewart, K. R. Dunbar and C. Turro, *J. Am. Chem. Soc.*, 2008, **130**, 1163–1170.

25 L. M. Wilhelmsson, F. Westerlund, P. Lincoln and B. Nordén, *J. Am. Chem. Soc.*, 2002, **124**, 12092–12093.

26 F. Westerlund, P. Nordell, J. Blechinger, T. M. Santos, B. Nordén and P. Lincoln, *J. Phys. Chem. B*, 2008, **112**, 6688–6694.

27 J. Andersson, M. Li and P. Lincoln, *Chem. A Eur. J.*, 2010, **16**, 11037–11046.

28 M. J. Packer, M. P. Dauncey and C. A. Hunter, *J. Mol. Biol.*, 2000, **295**, 71–83.

29 M. A. El Hassan and C. R. Calladine, *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.*, 1997, **355**, 43–100.

30 W. Lee and N. R. Krishna, *J. Magn. Reson.*, 1992, **98**, 36–48.

31 L. Wu, A. Reymer, C. Persson, K. Kazimierczuk, T. Brown, P. Lincoln, B. Nordén and M. Billeter, *Chem. A Eur. J.*, 2013, **19**, 5401–5410.

32 T. A. Fairley, R. R. Tidwell, I. Donkor, N. A. Naiman, K. A. Ohemeng, R. J. Lombardy, J. A. Bentley and M. Cory, *J. Med. Chem.*, 1993, **36**, 1746–1753.

33 D. G. Gorenstein, *Methods in Enzymology*, Academic Press, 1992, vol. 211, pp. 254–286.

34 D. G. Gorenstein, B. A. Luxon, E. M. Goldfield, K. Lai and D. Vegeais, *Biochemistry*, 1982, **21**, 580–589.

35 D. G. Gorenstein, S. A. Schroeder, J. M. Fu, J. T. Metz, V. Roongta and C. R. Jones, *Biochemistry*, 1988, **27**, 7223–7237.

36 R. L. Jones and W. D. Wilson, *J. Am. Chem. Soc.*, 1980, **102**, 7776–7778.

37 C. M. Dupureur and J. K. Barton, *J. Am. Chem. Soc.*, 1994, **116**, 10286–10287.

38 E. M. Goldfield, B. A. Luxon, V. Bowie and D. G. Gorenstein, *Biochemistry*, 1983, **22**, 3336–3344.

39 A. Pardi, R. Walker, H. Rapoport, G. Wider and K. Wüthrich, *J. Am. Chem. Soc.*, 1983, **105**,

1652–1653.

40 J. Andersson and P. Lincoln, *J. Phys. Chem. B*, 2011, **115**, 14768–14775.

41 A. A. Almaqwashi, T. Paramanathan, P. Lincoln, I. Rouzina, F. Westerlund and M. C. Williams, *Nucleic Acids Res.*, 2014, **42**, 11634–11641.

42 L. Wu, A. Reymer, C. Persson, K. Kazimierczuk, T. Brown, P. Lincoln, B. Nordén and M. Billeter, *Chem. A Eur. J.*, 2013, **19**, 5401–5410.

Chapter 4. EXSY NMR Studies of Δ,Δ-[1]

4.1 Chemical Exchange

The chemical shift of a nucleus is sensitive to the surrounding environment. Thus, changes to this environment, either through a chemical reaction or a conformational change, result in an alteration of chemical shift. Exchange in NMR is a dynamic process, whereby the NMR probe is exposed to multiple environments or conformations in a manner that is time dependent. One simple model is an interconversion of a two state system (figure 4.1). In this model, the proton exists in one of two conformations, either A or B. Each of these conformations has an associated resonance frequency, ν, and the difference in frequency between the two states is Δν. The exchange regime can be governed by the temperature, as such, an increase in temperature elevates the rate of interconversion and thus the exchange rate constant (k_r). The exchange regime can be described by comparison of the values for k_r and Δν in hertz.[1] The effect of increasing temperature (T) for a two state system with equal populations is shown in figure 4.1.

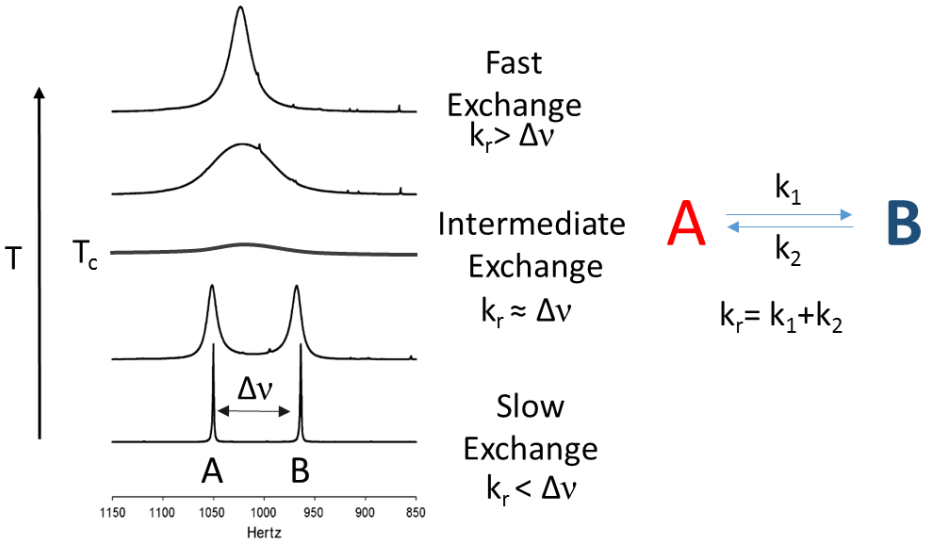


Figure 4.1. The effect of increasing temperature on chemical exchange in a two state system. Adapted from reference 2.[2]

When undergoing slow exchange there is no significant exchange on the NMR timescale and this is reflected in the observation of both distinct sets of resonances. As the temperature is increased, the line widths of each resonance broaden in an effect known as exchange broadening. This phenomenon is observed as depopulation of each state during acquisition results in a loss of transverse magnetization and a reduction in the transverse relaxation time (T_2). The line width ($\nu_{1/2}$) is defined as the peak width at half height and is inversely proportional to T_2 as shown in (equation 4.1).

$$\boldsymbol{\nu_{1/2} = \frac{1}{\pi T_2}}$$

Equation 4.1. Definition of line width.

As the temperature increases, $\nu_{1/2}$ also increases due to faster exchange resulting in line broadening. This moves the system into an intermediate exchange regime as the trough between the two peaks reduces. Broadening occurs until eventually the two peaks coalesce into a single peak at the temperature of coalescence (T_c). At T_c, the exchange rate constant (k_c) is given by equation 4.2.

$$\boldsymbol{k_c = \frac{\pi \Delta \nu}{\sqrt{2}}}$$

Equation 4.2. Definition of the exchange rate constant at the temperature of coalescence.

From this point, the interconversion of the two states becomes rapid on the NMR timescale. This leads to an averaging of signals giving a resultant peak occuring at the population weighted average of the individual chemical shifts. As the temperature increases further, the signal continues to sharpen until a limit of an entirely fast exchange regime with complete averaging of states is reached.

Multiple NMR experiments have been designed to study exchange phenoma occuring over the different exchange regime timescales.[3–5] One method of determining the exchange rate of a system that is in chemical exchange, is the calculation of k_c at the coalesence temperature. However, the drawback of this method is that the exact temperature of T_c can be hard to determine, and if the chemical shifts are temperature dependent then $\Delta\nu$ can also be hard to predict.[6]

This research applies the use of 2D EXchange SpectroscopY (EXSY) for the dynamic study of the Δ,Δ-[1] DNA complex. Here, the slow to intermediate conformational change seen in the NMR is used to determine activation energy parameters.[7,8]

4.2 EXSY NMR

Since 1979, EXSY spectroscopy has been used to quantify dynamic processes in the slow to intermediate time window of 10-5000 ms.[8] The technique can be applied to systems that yield sharp separate peaks in the NOESY spectrum ($k_r << \Delta\nu$), but also yield exchange cross peaks, as seen for the Δ,Δ-[1] DNA complex in figure 4.2.

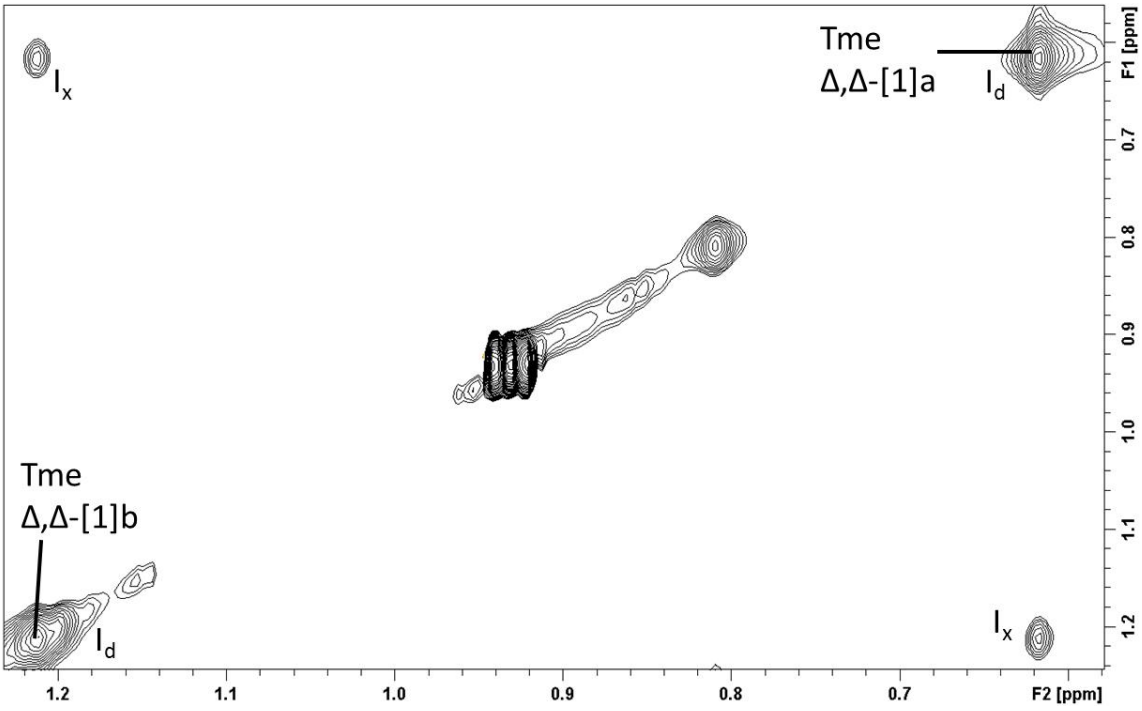


Figure 4.2. Expansion of the TMe region of the NOESY spectrum of 2:1 Δ,Δ-[1]: DNA complex. The cross peaks (I_x) shown in the spectrum appear as a result of chemical exchange between two bound conformations Δ,Δ-[1]a and Δ,Δ-[1]b.

The cross peaks are seen in the NOESY spectrum as a result of cross relaxation due to dipolar coupling between nuclear spins that are close in space, but also between exchanging molecules. EXSY experiments look to probe the latter.

A NOESY NMR experiment takes the general form shown in figure 4.3.

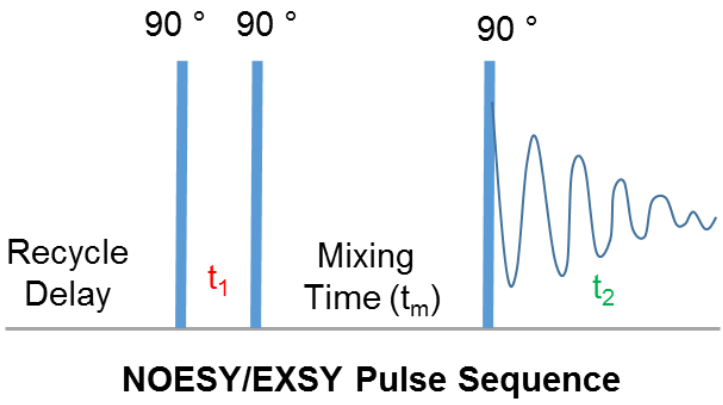


Figure 4.3. Illustration of the general pulse sequence applied during a NOESY experiment.

A molecule undergoing slow-intermediate chemical exchange over the NOESY mixing time (t_m) (200 ms for this study) will display two distinct sets of resonances with cross peaks between them. A proton that has the chemical shift of conformation (A) prior to the mixing time, experiences a conformational change throughout the mixing time and takes on the chemical shift of conformation (B). This means a peak labelled with chemical shift frequency, A, during t_1 has the chemical shift, B, during t_2 acquisition which results in a cross peak between the two chemical shifts. With longer mixing times the cross peak intensity increases, as more conformational change occurs during this time. This effect can be exploited to determine rates of reactions for equilibrium processes using an experiment called 2D EXchange SpectroscopY, EXSY.[8] An EXSY uses exactly the same pulse sequence as a NOESY. By varying the t_m, the build-up rate of the exchange cross peak can be measured as a function of time. Thus, this can be used to derive k_r for the system, which, if measured at different temperatures, can be applied to estimate kinetic parameters.[9]

4.2.1 Studies of Δ,Δ-[1]

The two conformations of the DNA/Δ,Δ-[1] are equally populated with integration of the two sets of resonances showing a 1:1 ratio. For systems of spins A and B which are equally populated with equal longitudinal relaxation times ($T_{1A} = T_{1B} = T_1$) it has been shown that the diagonal peak intensity (I_d) in an EXSY experiment is described by equation 4.3.[10]

$$I_d = \frac{1}{4}\left[e^{\left(-\frac{t_m}{T_1}\right)}\right]\left[1 + e^{\left(-k_r\frac{t_m}{T_1}\right)}\right]$$

Equation 4.3. The diagonal peak intensity, where t_m is the mixing time and k_r is the exchange rate constant and T_1 is the longitudinal relaxation time

The cross peak intensity (I_x) can be described by the following equation.

$$I_x = \frac{1}{4}\left[e^{\left(-\frac{t_m}{T_1}\right)}\right]\left[1 - e^{\left(-k_r\frac{t_m}{T_1}\right)}\right]$$

Equation 4.4. The cross peak intensity.

For the DNA/Δ,Δ-[1] it was possible to measure I_d and I_x for the exchanging thymidine methyl protons as there was sufficient spectral dispersion. EXSY experiments were carried out at various mixing times (figure 4.4) in order to measure the build up of cross peak intensity with increasing mixing time (t_m = 50 ms ,100 ms, 200 ms, 500 ms and 1s).

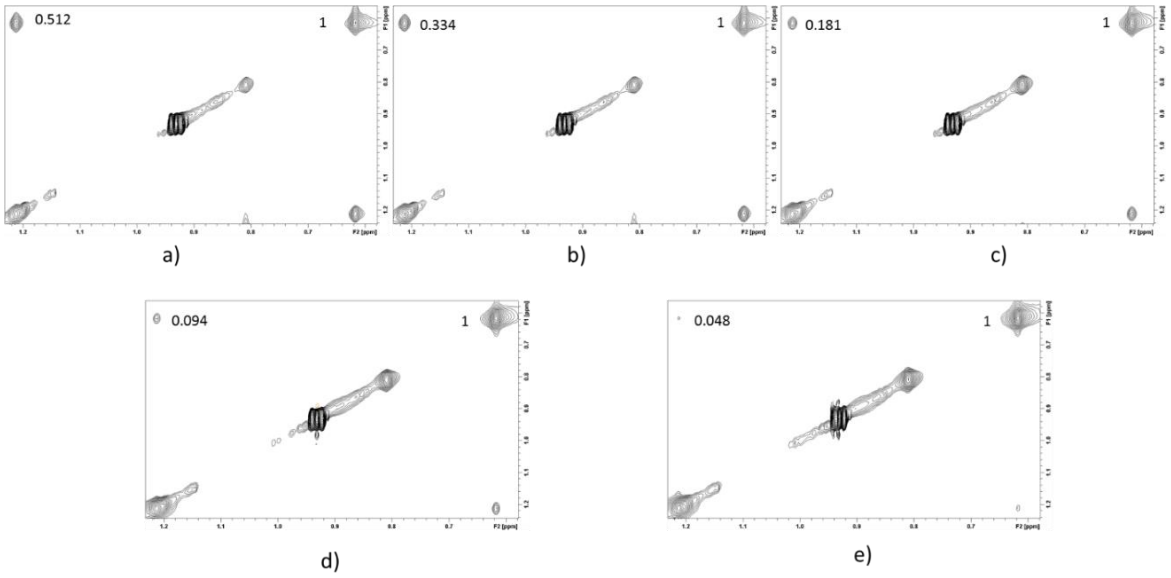


Figure 4.4. EXSY NMR experiment from the uncrowded thymidine methyl region of the NOESY spectrum of the Δ,Δ-[1]/DNA complex at 298 K showing the relative cross peak and diagonal peak intensities using mixing times of a) 1 s, b) 500 ms, c) 200 ms, d)100 ms and e) 50 ms.

As t_m is increased, the I_x increases relative to I_d as more exchange has occurred over the longer mixing time. The ratio of the two can be measured, I_x/I_d, which is described by equation 4.5.

$$I_x/I_d = \left[\frac{1 - e^{(-k_r t_m)}}{1 + e^{(-k_r t_m)}}\right]$$

Equation 4.5. The ratio of cross peak intensity to diagonal peak intensity

In this form of the equation, the T_1 relaxation terms cancel out and thus it is possible to determine k_r by fitting the data to the equation. This was carried out using sigma plot and is seen is figure 4.5.

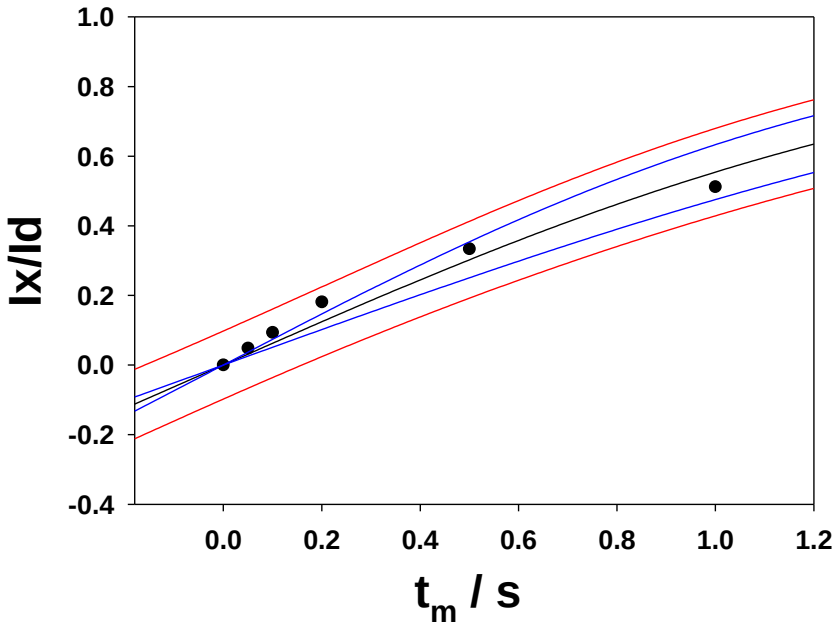


Figure 4.5. Δ,Δ-[1]/DNA EXSY build up data at 298 K giving (I_x/I_d) at different mixing times (t_m) given by fitting to equation 4.5. This estimates k_r to be 1.25 s^{-1}

The evaluation of the cross peak data at 298 K gave the value of k_r to be 1.25 ± 0.089 s^{-1}. In order to provide a more accurate estimate of thermodynamic parameters, multiple values of k_r are needed at different temperatures to observe how the exchange rate constant varies with temperature. Therefore, the EXSY experiments used to produce k_r at 298 K were repeated at temperatures of 288 K, 293 K, 303 K and 308 K. The analysis of the data is displayed in figure 4.6.

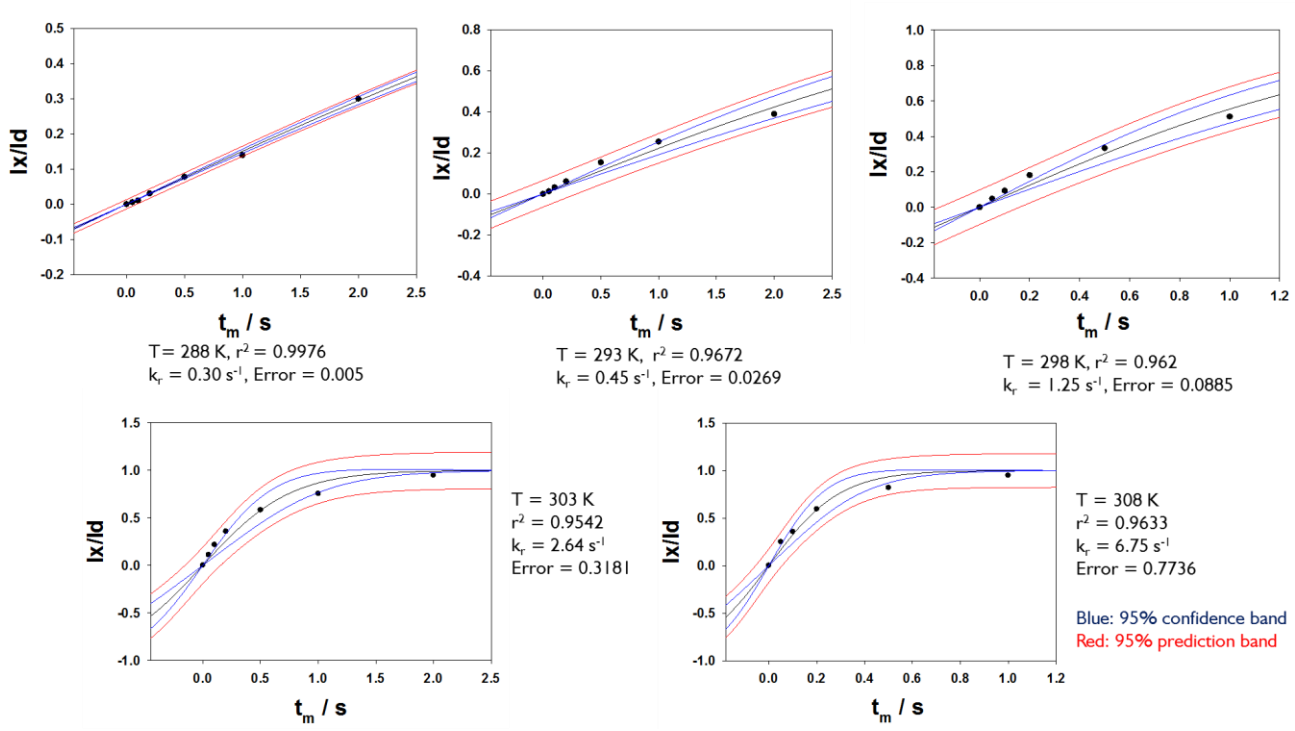


Figure 4.6. Δ,Δ-[1]/DNA EXSY build up data giving (I_x/I_d) at different mixing times (t_m) (50, 100, 200, 500, 1000, 2000) ms and temperatures a) 288 K, b) 293 K, c) 298 K, d) 303 K and e) 308 K fitted to equation 4.5. 95% confidence bands are shown in blue and 95% prediction bands are shown in red.

The temperature dependent rate constants as derived by the EXSY experiments are listed below in table 4.1.

Table 4.1. Temperature dependent rate constants from EXSY.

T / K	k_r / s^{-1}
288	0.3 ± 0.005
293	0.45 ± 0.026
298	1.25 ± 0.089
303	2.64 ± 0.318
308	6.75 ± 0.774

4.2.2 Arrhenius Analysis

The Arrhenius equation describes the temperature dependence of reaction rates and is given by equation 4.6.

$$\mathbf{k_r = A.\, e^{-\frac{Ea}{RT}}}$$

Equation 4.6. The Arrhenius equation where E_a is the activation energy, R is the gas constant (8.3145 J K^{-1} mol^{-1}, T is the temperature and A is the Arrhenius frequency factor).

Application of the Arrhenius equation can be used to determine the activation energy, E_a, for the transition between Δ,Δ-[1]a and Δ,Δ-[1]b by plotting $\ln(k_r)$ against 1/T to generate a linear Arrhenius plot (equation 4.7).

$$\mathbf{\ln(k_r) = \ln(A) - \frac{E_a}{RT}}$$

Equation 4.7. The linear form of the Arrhenius equation

In this form of the equation, the E_a is obtained from the slope of the plot ($-E_a/R$) and the y intercept is the pre exponential factor A. The experimentally derived rate constants at each temperature are plotted in this form and fitted to the linear equation. The plot is shown in figure 4.7.

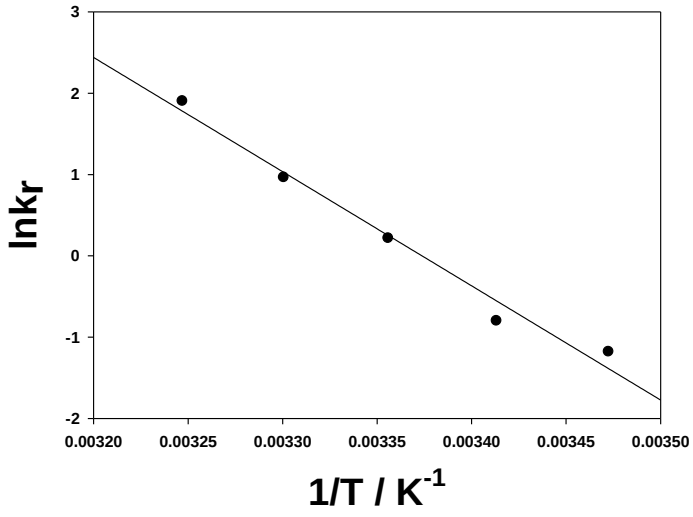


Figure 4.7. Arrhenius plot used to derive the activation energy of the conformational change seen in the Δ,Δ-[1]/DNA complex.

Using this Arrhenius analysis, the activation energy (E_a) for the transition was determined to be 116 kJ mol^{-1}.

4.2.3 Eyring Analysis

The Eyring equation (4.8) can also be employed to describe the temperature dependence of reaction rates in terms of the Gibbs free energy of activation.

$$\mathbf{k_r = \kappa \frac{k_b T}{h} e^{-\frac{\Delta G^{\ddagger}}{RT}}}$$

Equation 4.8. The Eyring equation, where κ is the transmission coefficient, k_b is the Boltzmann constant, h is Planck's constant, R is the gas constant and $G^{\ddagger}$ is the gibbs free energy of activation.

The standard equation for calculating the change in Gibbs free energy is shown below. Thus, if the enthalpy change and entropy change for a given reaction are known, then the Gibbs free energy change can be determined.

$$\mathbf{\Delta G^{\ddagger} = \Delta H^{\ddagger} - T\Delta S^{\ddagger}}$$

Equation 4.9. The Gibb's free energy equation where $H^{\ddagger}$ is the activation enthalpy and $S^{\ddagger}$ is the activation entropy.

Using a transmission coefficient (κ) of 1 assumes the flux through the reaction coordinate is complete with no re-crossing of the transition state occurring.[11] These two equations can be combined to provide estimates for the enthalpy of activation ($\Delta H^{\ddagger}$) and activation entropy ($\Delta S^{\ddagger}$). The linear form of the Eyring equation (4.10) enables the determination of enthalpy and entropy terms by plotting $\ln(k_r/T)$ against 1/T (figure 4.8).

$$\mathbf{\ln\left(\frac{k_r}{T}\right) = \ln\left(\frac{k_b}{h}\right) - \frac{\Delta H^{\ddagger}}{RT} + \frac{\Delta S^{\ddagger}}{R}}$$

Equation 4.10. Linear form of the Eyring equation described in terms of enthalpy and entropy of activation.

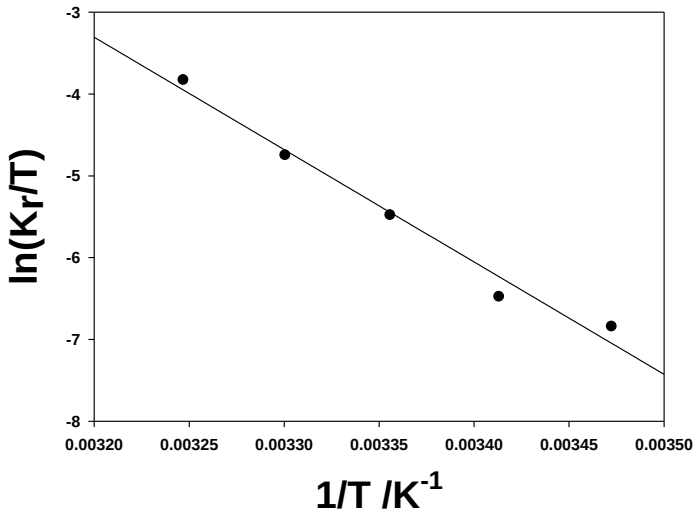


Figure 4.8. Eyring plot produced from exchange rate constants for the Δ,Δ-[1]:DNA complex at varied temperatures, giving E_a (116.7 kJ mol^{-1}), $\Delta H^{\ddagger}$ (114.1 kJ mol^{-1}) and $\Delta S^{\ddagger}$ (140.3 J mol^{-1} K^{-1}).

Employment of the Eyring equation using this figure led to enthalpy ($\Delta H^{\ddagger}$) and entropy ($\Delta S^{\ddagger}$) estimates of 114.1 kJ mol^{-1} and 140.3 J mol^{-1} K^{-1} respectively. This suggests that the energetic barrier for the transition is mainly enthalpic, but without the kinetic data from the other stereoisomers to compare it to, little else can be inferred.

4.3 Discussion

Δ,Δ-[1] binds to DNA in two conformations, intercalated (Δ,Δ-[1]a) and groove bound (Δ,Δ-[1]b) these are equally populated as shown experimentally in their equal intensity NMR signals. These conformations are shown to be in slow dynamic exchange, observed by well resolved separate NOESY spectra showing exchange peaks. It is therefore observable that these two species are thermodynamically similar in energy, as well as being kinetically accessible. The simplified reaction coordinate can thus be predicted to be as shown in figure 4.9.

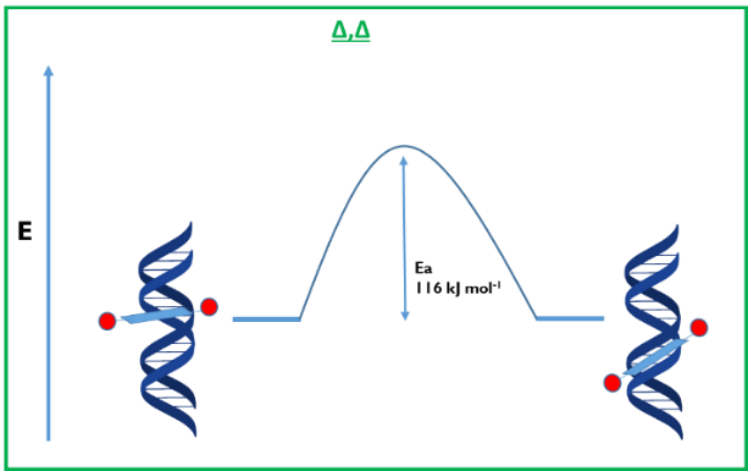


Figure 4.9. The simplified reaction coordinate energy level diagram for threading of Δ,Δ-[1] into d(GCATATCG).(CGATATGC).

The system is clearly not a simple two-site exchange, with only two defined conformations significantly populated, which is evident from the systematic errors seen in the fits of figures 4.5 and 4.6. It is predicted that there are several intermediate conformations that the system adopts whilst the bulky ruthenium centre attempts to worm its way through the partially dissociated duplex. The data obtained is not complete enough to allow a more detailed analysis of this system and it is therefore described as such.

Using an Arrhenius analysis of the experimentally derived rate constants, the activation energy (E_a) for the transition was found to be 116 kJ mol^{-1}. Employment of the Eyring equation led to the enthalpy ($\Delta H^{\ddagger}$) and entropy ($\Delta S^{\ddagger}$) estimates of 114.1 kJ mol^{-1} and 140.3 $Jmol^{-1}$ K^{-1} respectively, values that are comparable to those obtained in previous studies. Specifically, in work on their more extended threading complex $[Ru_2(bipy)_2bidppz]^{4+}$, which incorporates the same $Ru^{II}(bipy)_2$ moieties, the Lincoln group reported that the Λ,Λ diastereomer threads through d(AT) DNA with first order kinetics and an activation energy of 132 $kJmol^{-1}$.[12] The lower activation energy for threading of Δ,Δ-[1] is reflected in the faster exchange kinetics seen in the NMR. This is consistent with previous studies showing that the Δ,Δ stereoisomers of threading complexes display faster kinetics than those of the Λ,Λ forms.[13] One explanation for this effect is that the Δ,Δ forms a more stable transition state than that of the other stereoisomers, thus lowering the activation energy of the reaction. The hypothesis, proposed by Lincoln, *et al.* is illustrated in figure 4.10. The transition state is proposed to contain at least one flipped out base,[14] stabilised by π-stacking interactions to the aromatic ancillary ligands, which then leads to threading.

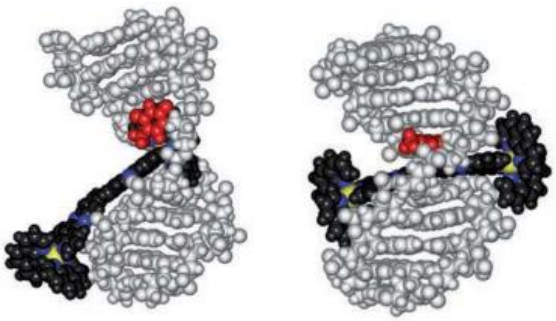


Figure 4.10. A model of the proposed transition state structure (left) adopted by $[Ru_2(phen)_2bidppz]^{4+}$ in the threading to poly d(AT) DNA and (right) the final threaded structure. Adapted from reference 12.[12]

Λ,Λ-[1] and Λ,Δ-[1] are rather different to Δ,Δ-[1] as there is no observable groove bound form. Therefore in this case, a groove bound conformation must be much less stable than the intercalated species. This thermodynamic preference for the intercalated form must be from a combination of unfavourable interactions from Δ,Δ in the threaded state and also an unfavourable steric interaction from having a single Λ centre in the groove bound state. The slow rate of dethreading

for these species is predicted to be due to a combination of unfavourable thermodynamics and a higher activation energy to threading/dethreading. A simplified energetic profile for the dethreading is shown below.

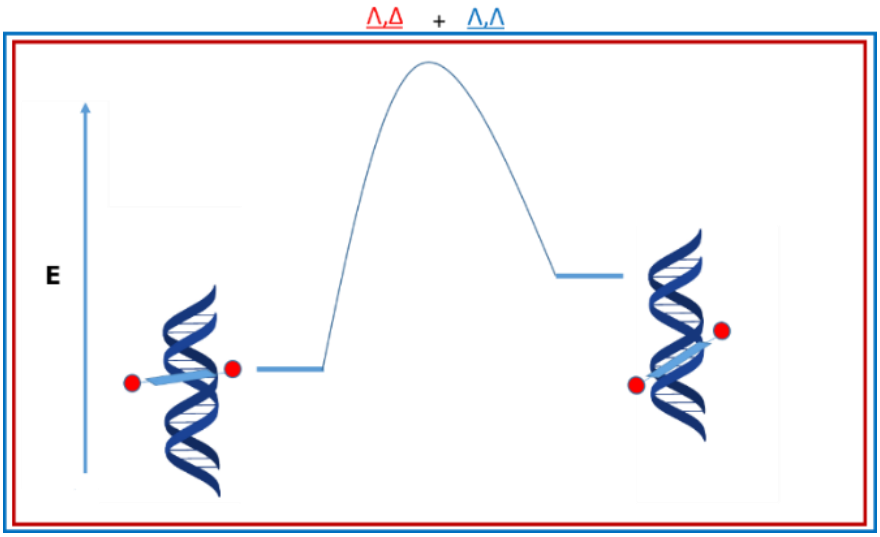


Figure 4.11. The simplified reaction coordinate energy level diagram for threading of Λ,Δ-[1] and Λ,Λ-[1] into d(GCATATCG).(CGATATGC)

4.4 References

1 I. R. Kleckner and M. P. Foster, *Biochim. Biophys. Acta - Proteins Proteomics*, 2011, **1814**, 942–968.

2 A. D. Bain, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2003, **43**, 63–103.

3 A. G. Palmer, C. D. Kroenke and J. Patrick Loria, *Methods in Enzymology*, Academic Press, 2001, **339**, 204–238.

4 A. Schmidt and S. Vega, *J. Chem. Phys.*, 1987, **87**, 6895–6907.

5 A. Allerhand, H. S. Gutowsky, J. Jonas and R. A. Meinzer, *J. Am. Chem. Soc.*, 1966, **88**, 3185–3194.

6 H. S. Gutowsky and H. N. Cheng, *J. Chem. Phys.*, 1975, **63**, 2439–2441.

7 J.-L. Yan, X.-A. Mao and L.-F. Shen, *Chem. Phys. Lett.*, 1997, **272**, 278–283.

8 J. Jeener, B. H. Meier, P. Bachmann and R. R. Ernst, *J. Chem. Phys.*, 1979, **71**, 4546–4553.

9 K. D. Zimmer, R. Shoemaker and R. R. Ruminski, *Inorganica Chim. Acta*, 2006, **359**, 1478–1484.

10 G. Fischer and E. Kleinpeter, *Magn. Reson. Chem.*, 1991, **29**, 204–206.

11 E. M. Mortensen and H. Eyring, *J. Phys. Chem.*, 1960, **64**, 846–849.

12 P. Nordell, F. Westerlund, L. M. Wilhelmsson, B. Nordén and P. Lincoln, *Angew. Chemie*, 2007, **119**, 2253–2256.

13 J. R. Johansson, Y. Wang, M. P. Eng, N. Kann, P. Lincoln and J. Andersson, *Chem. A Eur. J.*, 2013, **19**, 6246–6256.

14 T. Paramanathan, F. Westerlund, M. J. McCauley, I. Rouzina, P. Lincoln and M. C. Williams, *J. Am. Chem. Soc.*, 2008, **130**, 3752–3753.

Chapter 5. Simulated Annealing

5.1. Molecular Mechanics

The determination of a biomolecular structure through NMR techniques is generally done by augmenting traditional computational molecular modelling functions with experimentally derived restraints. Computational *ab initio* molecular modelling studies involve the application of a model that describes intra and intermolecular interactions of a chosen system. One of the most common models chosen for this utilises molecular mechanics. A Newtonian approach employs classical mechanics to calculate the forces and masses that define the potential energy of a system. The total energy, $\mathbf{E_{tot}}$, of a system is described by the sum of the contributing energy functions. The energy terms can be broken down into force potentials which together describe the overall force field, a general form of which is given below in equation 5.1.

$$\mathbf{E_{tot} = E_{bond} + E_{angle} + E_{dihedral} + E_{improper} + E_{electrostatic} + E_{van\ der\ Waals}}$$

Equation 5.1. The potential energy of a system

Modelling programs are then employed to solve individual contributing energy equations which describe these force potentials to calculate the $\mathbf{E_{tot}}$ of the system. The energy functions describe molecules as functions of their atomic coordinates modelling the atoms as point charges with set mass and radii linked by harmonic springs. The bonded energy terms, $\mathbf{E_{bond}}$, $\mathbf{E_{angle}}$ $\mathbf{E_{dihedral}}$ and $\mathbf{E_{improper}}$, are modelled as harmonic potentials with energies calculated by the differences between bond lengths and angles from the ideal equilibrium values.[1] Therefore, if the bond length, or other conformational term, is increased or decreased with respect to the expected value, the energy of the system increases. These bonding energy terms are described by the following general equations (5.2-5.5).

$$\mathbf{E_{bond}} = \sum \boldsymbol{k_l}(\boldsymbol{l} - \boldsymbol{l_0})^2$$

Equation 5.2. Potential energy term for the contribution of bond length to the force field, where k_l = force constant, l_0= expected bond length and l = bond length.

$$\mathbf{E_{angle}} = \sum \boldsymbol{k_\theta}(\boldsymbol{\theta} - \boldsymbol{\theta_0})^2$$

Equation 5.3. Potential energy term for the contribution of bond angle to the force field, where k_θ = force constant, θ_0= expected bond angle and θ = bond angle.

$$\mathbf{E_{dihedral}} = \sum \boldsymbol{k_\phi}[\mathbf{1} + \boldsymbol{cos}\,(\boldsymbol{n\phi} - \boldsymbol{\phi_0})]$$

Equation 5.4. Potential energy term for the contribution of the dihedral angle to the force field, where k_ϕ = force constant for each Fourier term, $\boldsymbol{\phi}_0$ = expected dihedral, $\boldsymbol{\phi}$ = dihedral angle and n= multiplicity of potential minimum conformation.

$$\mathbf{E_{improper}} = \sum \boldsymbol{k_\omega}(\boldsymbol{\omega} - \boldsymbol{\omega_0})^2$$

Equation 5.5. Potential energy term for the contribution of the improper angle to the force field, where k_ω = force constant, $\boldsymbol{\omega}_0$ = expected improper and $\boldsymbol{\omega}$ = improper.

The non-bonded potential energy terms $\mathbf{E_{van\ der\ Waals}}$ and $\mathbf{E_{elecrostatic}}$ are described by the equation below, where summation occurs over all atoms in the system.

$$\mathbf{E_{nonbond}} = \sum \boldsymbol{\varepsilon_{ij}}\left[\left(\frac{r_{ij}^{min}}{r_{ij}}\right)^{12} - 2\left(\frac{r_{ij}^{min}}{r_{ij}}\right)^{6}\right] + \sum\left(\frac{q_i q_j}{4\pi\varepsilon_r\varepsilon_o r_{ij}}\right)$$

Equation 5.6. Potential energy term for the non-bonding contributions to the field, where $\boldsymbol{\varepsilon_{ij}}$ = potential well depth, r_{ij} = distance between atoms, r_{ij}^{min} =distance at potential minima, ε_r = dielectric constant, ε_o= vacuum permittivity and q = charge

Energy terms from the van der Waals (vdW) components are described as Lennard Jones potentials which are attractive forces at r^{-6} and repulsive at r^{-12}. The electrostatic energy function is defined by Coulomb's law. Often during molecular modelling calculations, the $\mathbf{E_{elecrostatic}}$ is left out of the force field equations as defining the dielectric constant of a system can be very computationally intensive with the needed addition of solvent molecules to the calculation. Attractive vdW terms can also be exclude to further simplify the forcefield with the application of a purely repulsive vdW component.

5.2. Molecular Modelling

Molecular modelling programs apply force field calculations to determine the energy of chosen systems such as biomolecules. This can be applied to studying the interaction of an oligonucleotide sequence with the various diastereomers of [1]. Computer programs are written to attempt to optimise the structure of the system in order to produce the lowest energy conformation. There are two main methods of doing this which can be used in conjuction, energy minimisation and molecular dynamics.

5.2.1 Energy Minimisation

Energy minimisation is the process of optimization of the structure to minimise $\mathbf{E_{tot}}$. Various changes are applied to the atomic coordinates within the system and checked to see if these movements result in any improvement to the potential energy. If a decrease in potential energy is observed, the newly generated coordinates are accepted and another step can then be undertaken. Once a set amount of steps is completed or a predefined value of the energy minimum is reached, then the minimisation can be considered complete. Several algorithms have been written for this gradient to a potential minimum and applications of these can give a rapid convergence to a local minimum.[2] Whilst energy minimisation is useful in the determination of local energy minima, quite often these are not the global minimum. If there is an energetic penalty in order to transition from one state to another, calculations using only this method will not be able to overcome this barrier (figure 5.1).

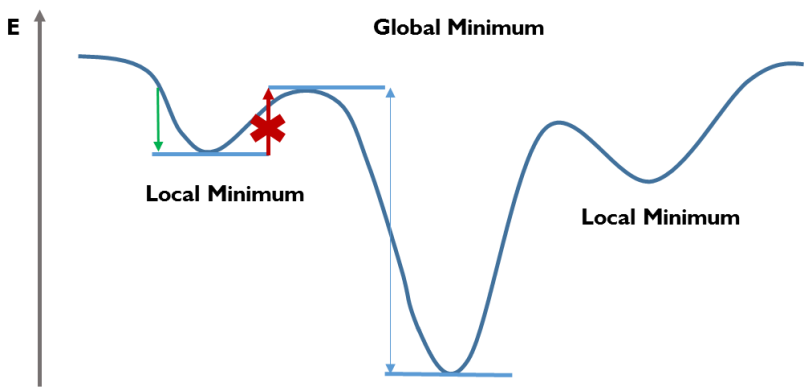


Figure 5.1. 2D illustration of energy minima. Energy minimisation enables convergence to a local minimum but this is not necessarily the global minimum.

5.2.2 Molecular Dynamics

If the structural landscape contains many local maxima that cannot be overcome by a simple energy minimisation protocol, which is the case for molecular modelling simulations of biomolecules, another method of structural optimization is needed. Molecular dynamics calculations are employed to model the behaviour of a system with regards to time. This adds a kinetic aspect to the calculations with changes in the position of an atom with respect to time giving the atoms velocity. Complementary to this, the forces on an atom can be computed though the potential energy function. The averaged velocities of the atoms within the system are affected by the temperature in a Boltzmann distribution and therefore it is possible to "heat" a system by increasing the velocities of the atoms. These kinetic components can be incorporated into the calculations within molecular dynamics simulations, thus eliminating the problem of escaping local minima and facilitating exploration of the entire energy landscape, including higher energy species. A common method of molecular modelling is simulated annealing, whereby a high temperature molecular dynamics phase allows molecules to escape local minima, with a slow cooling phase allowing the system to move towards a globally minimised structure (figure 5.2). This is followed by final energy minimisation algorithms to optimize the final structure.

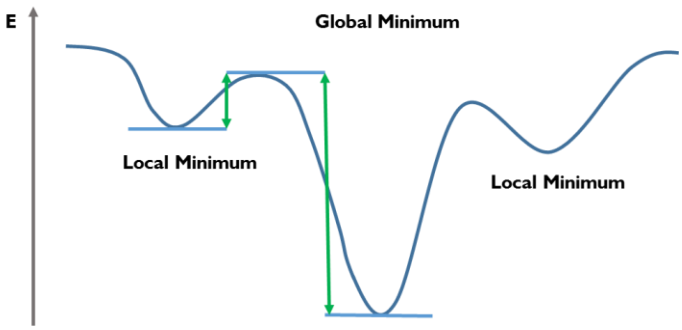


Figure 5.2. 2D illustration of the energy landscape. Molecular dynamics enables exploration of the entire surface.

Complete molecular dynamics can be very computationally demanding as simulations require high levels of computing power with many calculations being performed at very short time periods. Therefore, calculations are often simplified by reducing the amount of energy functions in the force field calculations. Non-bonding potential energy terms, such as the electrostatic and van der Waals terms

are often omitted, with solvent molecules also removed from calculations. These final models are therefore not an entirely accurate representation of the structure, but can provide useful insight when compared with experimentally derived data.

It is also possible to introduce experimentally derived restraints into the force field equations as pseudo energy terms in a binding study. These restraints can be weighted and used in tandem with the empirical energy terms in order to generate representative structures. This is the method employed in NMR based models, which are generated with terms based on observed phenomena, such as distance restraints from hydrogen bonds, and NOEs or dihedral angle restraints from J-couplings.[3]

5.2.3 Calibration of NOE Restraints

Before experimental data such as NOEs can be input as restraints into structure calculation programs, they first need to be interpreted as internuclear distances. There are multiple ways of doing this; one is to directly calculate the value of each individual NOE by using an internal reference distance. The interproton distances within DNA that are not dependant on conformation are (H2'-H2", 1.8 Å), (CH5-CH6, 2.5 Å) and (TH6-TMe, 2.7 Å). The unknown distance ($r_{i,j}$) relative to the reference distance (r_{ref}) can then be calculated using equation 5.7:

$$\frac{I_{ref}}{I_{i,j}} = \left(\frac{r_{i,j}}{r_{ref}}\right)^6$$

Equation 5.7. Where I_{ref} and $I_{i,j}$ are the NOE intensities for the reference and the unknown correlations respectively.

This equation assumes that the only difference between I_{ref} and $I_{i,j}$ is the distance dependence of r^{-6}, the dipolar cross relaxation rates. This mostly holds true for short mixing times and isotropic tumbling conditions. However errors in distance measurement can arise due to factors such as spin diffusion or overlapping peaks causing an error in the measurement of I_{ref} or I_{ij}. Therefore, rather than measuring each individual interproton distance and using these values as restraints, sets of NOEs are often approximated into groups classified as strong, medium or weak depending on their relative intensities. In this way,

these approximations allow for uncertainties in NOE measurement whilst giving reasonable distance restraints.[4]

5.3. Simulated Annealing using CNS

The Crystallography and NMR System (CNS) software suite (a development of XPLOR) has been routinely used for its function as a simulated annealing and energy minimisation package.[1,5,6] These computational methods combined with experimentally derived restraints have been used to effectively model the DNA binding of small molecules.[7,8] In these studies, the simulated annealing and energy minimisation functionality of CNS has been applied to the binding event of [1] to the oligonucleotide sequence d(GCATATCG).(CGATATGC).

5.3.1 Generation of Structure Files

The CNS program requires the application of suitable structure files which include topology, parameter, and coordinate files. These are necessary to define the structures of both the DNA and the ligands before the simulated annealing can be undertaken. Molecular topology files (MTF) define all of the atoms in the molecule by both mass and charge and then go on to define which atoms are bonded together with bond angles, improper and dihedral angles. Parameter files (PAR) then give values to these defined geometric functions, enabling the program to generate the structures in space when combined with a coordinate file in the protein data bank (PDB) format.

5.3.1.1 Structure Files for the DNA sequence d(GCATATCG).(CGATATGC)

There are utilities within the CNS suite for the generation of structure files for a B-DNA sequence, but first a suitable PDB file with the coordinates of the sequence is required. The coordinates for the B-DNA were generated using the DNA sequence-to-structure web tool, which produced a coordinate PDB file of the nucleotide sequence d(GCATATCG).(CGATATGC).[9] The web programme utilises conformational parameters taken from a database of fibre-diffraction studies. This file was then edited to be readable by the CNS software using the 'fix_dna_rna'

utility. Following this, the MTF file for the DNA was generated using the 'generate_easy.inp' function which also performs an energy minimisation to give a suitable structure and coordinate file for use in the simulation (figure 5.3). Throughout the study the standard parameter file 'dna-rna-allatom.param' was included.

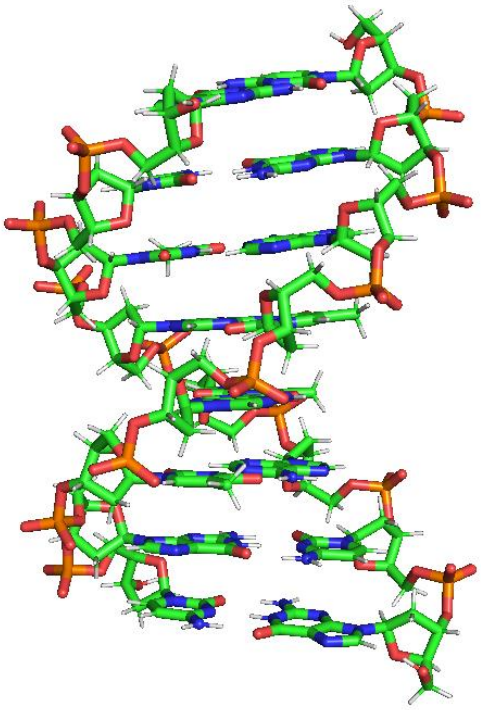


Figure 5.3. The energy minimised structure of the B-DNA helix d(GCATATCG).(CGATATGC), coordinate file generated from conformational parameters taken from fibre-diffraction studies.[9]

5.3.1.2 Structure Files for [1]

Whilst there are utilities for the generation of structure files for common biomolecules within the CNS software suite, generating the files for hetero compounds is less trivial. These files need to be converted from a PDB file to a CNS readable form using external utilities such as XPLO2D, a command line program written for this function by N. Kleywegt.[10] The coordinates from the crystal structures of [1] discussed in chapter 2 were used as a starting point for the generation of structure files. The generated topology and parameter files were stripped of their hydrogen atoms by the program. Therefore, these were re-added manually along with some other corrections needed to correct the syntax of the

files. The final structures were then energy minimised to produce the final coordinates (figure 5.4).

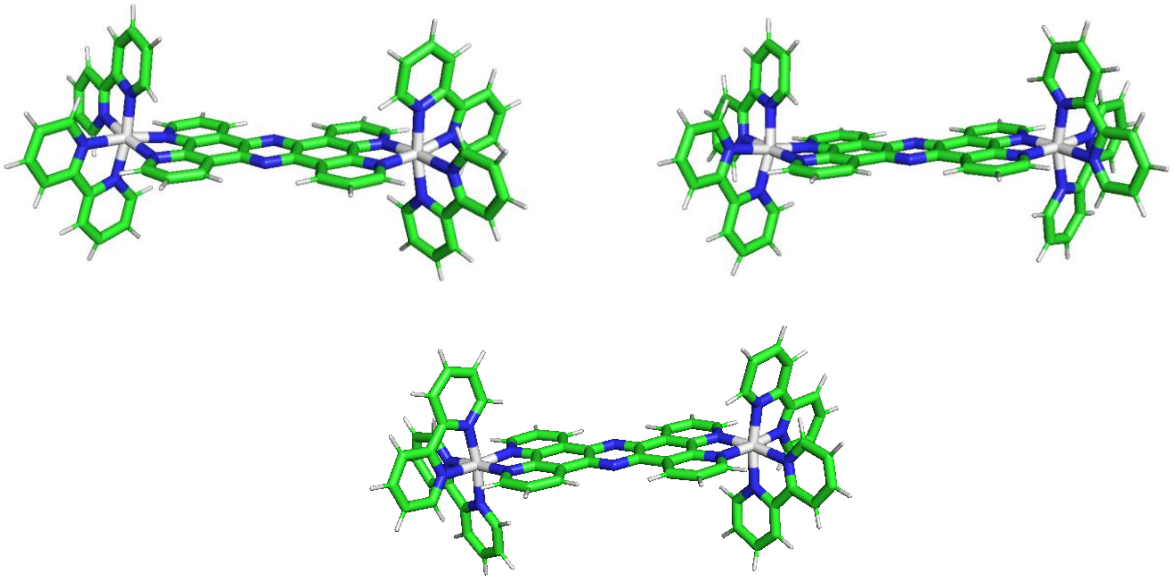


Figure 5.4. The energy minimised structures of Λ,Λ-[1], Λ,Δ-[1] and Δ,Δ-[1] generated from x-ray crystallographic structures.

5.4 Simulated Annealing of Λ,Δ-[1]

Once the appropriate structure files for the complexes had been generated, restrained molecular dynamics could be performed. The two PDB files (Λ,Δ-[1] and the DNA) were combined into one coordinate file to provide the starting point for the simulated annealing run (figure 5.5). The file was set up so that the two molecules were far enough apart so as to avoid a high energy steric clash during the dynamics calculation.

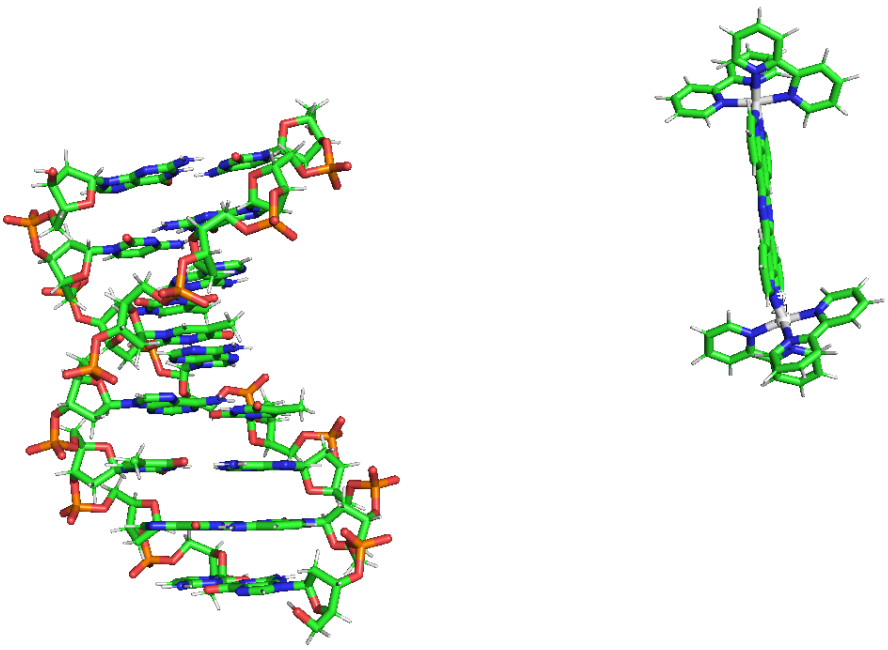


Figure 5.5. The starting coordinate file used to perform the restrained molecular dynamics simulations on the binding of Λ,Δ-[1] to d(GCATATCG).(CGATATGC)

The structure of the bound complex was calculated through the introduction of 24 intermolecular experimentally derived NOEs from the NMR data. These NOEs were grouped based on their observed intensities and classified as either strong (1.8 - 3 Å), medium (1.8 - 4 Å), or weak (1.8 - 5 Å).[11] Ambiguous NOEs were applied for the thymidine methyl protons with an additional 0.5 Å added to the upper bound to allow for this. Initially, each of the sets of Λ,Δ-[1] protons (1,6,7,12) (2,5,8,11) and (2,11,5,8) were treated as ambiguous NOEs (table 5.1). Test MD simulation runs suggested that further restraints were needed to keep the DNA as a double helix with the structure splitting apart upon calculation. Therefore, planarity restraints for the bases and interbase hydrogen bonding restraints were added to prevent the duplex DNA splitting into single strands during the simulation. Upon further simulations, large scale helical unwinding was seen, producing high energy structures and largely irregular DNA conformations. Thus, further dihedral restraints were utilised to maintain the C2' endo sugar pucker conformations. Backbone torsional angles were measured from the starting ideal B-DNA structure and restrained to be within ± 5° of this conformation (α = -46.8°, β = -146.1°, γ = 36.4°, δ = 156.4°, ε = 155°, ζ = -97.8°). The range of the backbone dihedral restraints were increased to ± 30° for the residues of the binding site, as it was predicted from the phosphorus NMR chemical shift data that these would

deviate more from their starting conformation. A full set of restraint files used in this study is listed in the appendix.

Table 5.1. Ambiguous NOE restraints used for simulated annealing of the Λ,Δ-[1]:DNA complex - (s), (m) and (w) indicate strong weak or medium NOE groupings.

Λ,Δ-[1]a	H1 or 6 or 7 or12	H2 or 5 or 8 or11	H3 or 4 or 9 or10
T_4	Me(m), 2"(w)	Me(s) , 1'(m), 2'(m), 2"(m)	Me(s), 2'(m), 2"(m) 1'(S), 2'(s), 2"(s)
A_5	1'(w)	1'(w)	
T_{12}	Me(m)	1'(m), 2'(m), 2"(m), Me(s)	1'(m), 2'(s), 2"(s), Me(m), 2'(m)
A_{13}	1'(w)		1'(w)

The simulated annealing itself was performed using the model_anneal.inp task file within CNS. Complex calculation was started with well-separated ligand and duplex structures and an energy minimisation was performed. This was followed by a high temperature annealing stage with heating up to 1000 K in 1000 steps. During the 15 picosecond high temperature dynamics phase, the van der Waals energy term was reduced to 0.1, thus facilitating threading of Λ,Δ-[1]. This was followed by a 15 picosecond cooling phase over 1000 steps in which the van der Waals energy term was scaled back up to one. This allows the ligand to sit in place and the reintroduction of the vdW repulsions pushes the adjacent bases apart to create space for [1]. This was followed by a final energy minimisation.

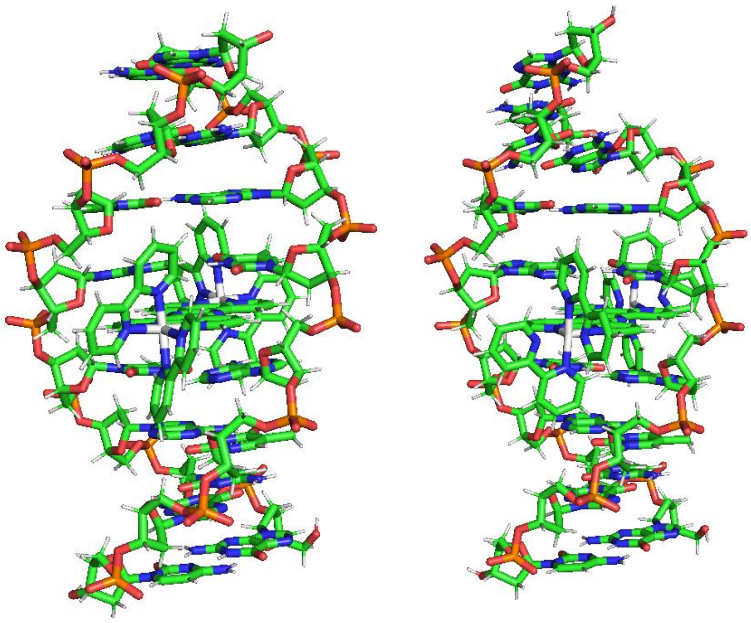


Figure 5.6. Resultant structures of molecular dynamics calculations annealed using ambiguous NOE restraints showing the two possible orientations of Λ,Δ-[1].

The restrained MD elucidated two energy minimised structures that satisfied all the NOE restraints and gave no close van der Waals contacts. One structure shows the Λ ruthenium centre in the minor groove (figure 5.6: left) and the other with the Λ in the major groove (figure 5.6: right). Although the ambiguous NOEs in table 5.1 were used in these initial calculations, the NMR data shows the NOEs are not entirely ambiguous, as the protons in each symmetry related spin system i.e. (1,2,3), (4,5,6), (7,8,9) and (10,11,12) are coupled to the same DNA proton within their groups. This means that it is valid to assign non-ambiguous NOEs by inspection of the close contacts observed in the simulated models, which allows for refinement to the specific NOEs given in table 5.2 and 5.3.

Table 5.2. Specific NOE restraints used for the simulated annealing of the Λ,Δ-[1]:DNA complex – (s), (m) and (w) indicate strong, medium or weak NOE groupings. Λ centre in the minor groove.

Λ,Δ-[1]	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12
T_4							Me(m)	Me(s)	Me(s), 1'(s), 1'(m),	2'(m), 2'(s), 2'(m),	2"(m) 2"(s) 2"(m)	2"(w)
A_5										1'(w)	1'(w)	
T_{12}		1'(m), 2'(m), 2"(m)	1'(m), 2'(s), 2"(s)	Me(m), 2'(m)	Me(s)	Me(m)						
A_{13}	1'(w)		1'(w)									

Table 5.3. Specific NOE restraints used for the simulated annealing of the Λ,Δ-[1]:DNA complex – (s), (m) and (w) indicate strong, medium or weak NOE groupings. Λ centre in the major groove.

Λ,Δ-[1]	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12
T_4	Me(m)	Me(s)	Me(s), 2'(m), 2"(m)	1'(s), 2'(s), 2"(s)	1'(m), 2'(m), 2"(m)	2"(w)						
A_5				1'(w)	1'(w)							
T_{12}								1'(m), 2'(m), 2"(m)	1'(m), 2'(s), 2"(s)	Me(m), 2'(m)	Me(s)	
A_{13}							1'(w)		1'(w)			

The simulations were then repeated using the specific NOEs. The final lowest energy minimised structural models are given in figure 5.7.

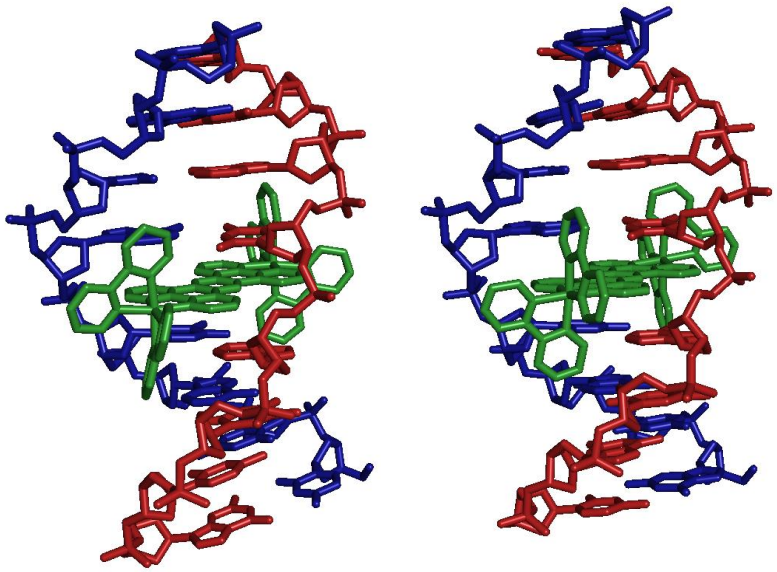


Figure 5.7. Resultant structures of restrained molecular dynamics using specific NOEs calculations showing the two possible orientations of Λ,Δ-[1].

The existence of these two structures is concordant with the NMR spectra that show two bound conformations. In both of the structures the ligand is intercalated at the central $(T_4,A_{13})|(T_{12},A_5)$ base step. The two structures differ in that they have the Λ and Δ Ru^{II} centres interchanged within the grooves of the DNA. Both of the structures have Λ,Δ-[1] embedded tightly into the minor groove maximising electrostatic and van der Waals contacts in this region, with the aromatic tpphz ligand over the base stack maximising π-π stacking interactions. The other ruthenium centre projects into the major groove (figure 5.8). The structure also shows a widening of the base pairs parallel to the long axis of the DNA as the tpphz ligand is added to the stack. The total energy for the structure with the Λ Ru^{II} centre in the minor groove was 681 kcal mol^{-1}, and the structure with the Δ Ru^{II} centre in the major groove was 683 kcal mol^{-1}. When compared to the total energy of the starting structure, 659.63 kcal mol^{-1}, this gives an energy increase upon complexation of 22 and 24 kcal mol^{-1} respectively. Whilst these numbers have no experimental value they do show that both structures are plausible and conform well to the input parameters.

The structure itself looks very plausible as it is comparable to that of the crystal structure of the mononuclear parent dppz compound.[12] The mononuclear compound binds deeply into the minor groove in a head on fashion with the end of

the dppz ligand protruding into the major groove in a similar location to that of the other ruthenium centre in the meso complex (figure 5.8).

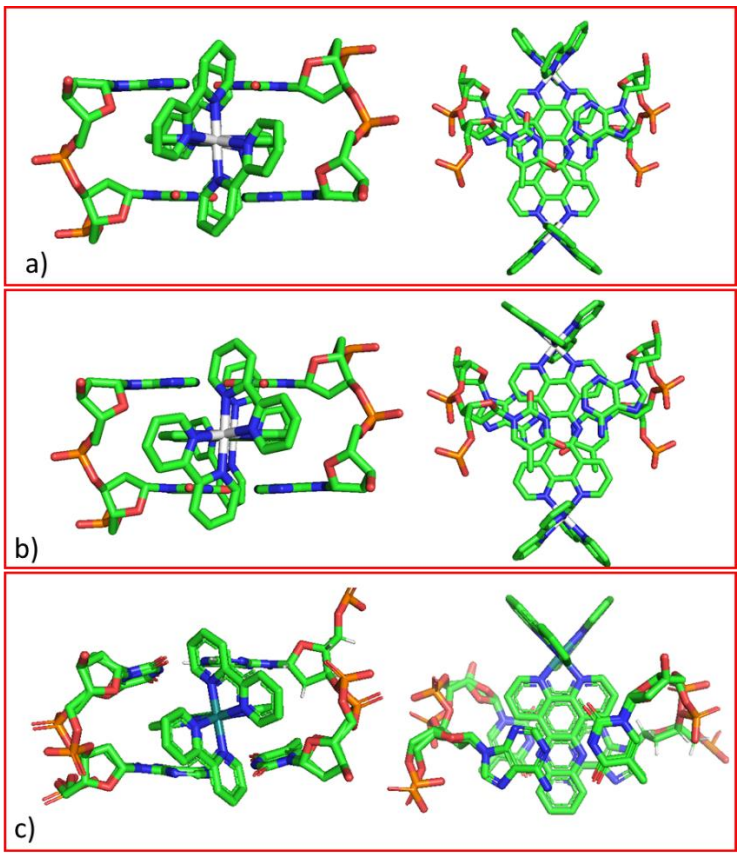


Figure 5.8. Views from minor groove and down the helical axis of DNA duplexes showing the Λ,Δ-[1] intercalation site (T_4,A_{13})|(T_{12},A_5) with (a) Λ-Ru(bipy)$_2$ in the minor groove and (b) with Δ-Ru(bipy)$_2$ in minor groove and (c) The Δ-[Ru(bpy)$_2$dppz]$^{2+}$ between AT base pairs from x-ray crystallographic studies by Barton et al.[12]

The most favourable form of the two species can be predicted from inspection of the major and minor grooves of the two conformations showing the Van der Waals surfaces (figure 5.9). The bipyridyl ligands of the Δ centre exhibit a twist perpendicular to the DNA with protons facing towards the DNA strand causing unfavourable steric clashes. However, the Λ centre has favourable vdW contacts to the DNA strand along the entire surface of the bipyridyl unit. As the complex is bound deeply in the minor groove it can be predicted that the most stable conformation (Λ,Δ-[1]a) is bound with the Λ centre in the minor groove and the Δ extended into the major groove. Consequently, the less stable and hence less populated (Λ,Δ-[1]b) is bound with the Δ centre in the minor groove and the Λ centre extended into the major groove as shown in figure 5.9.

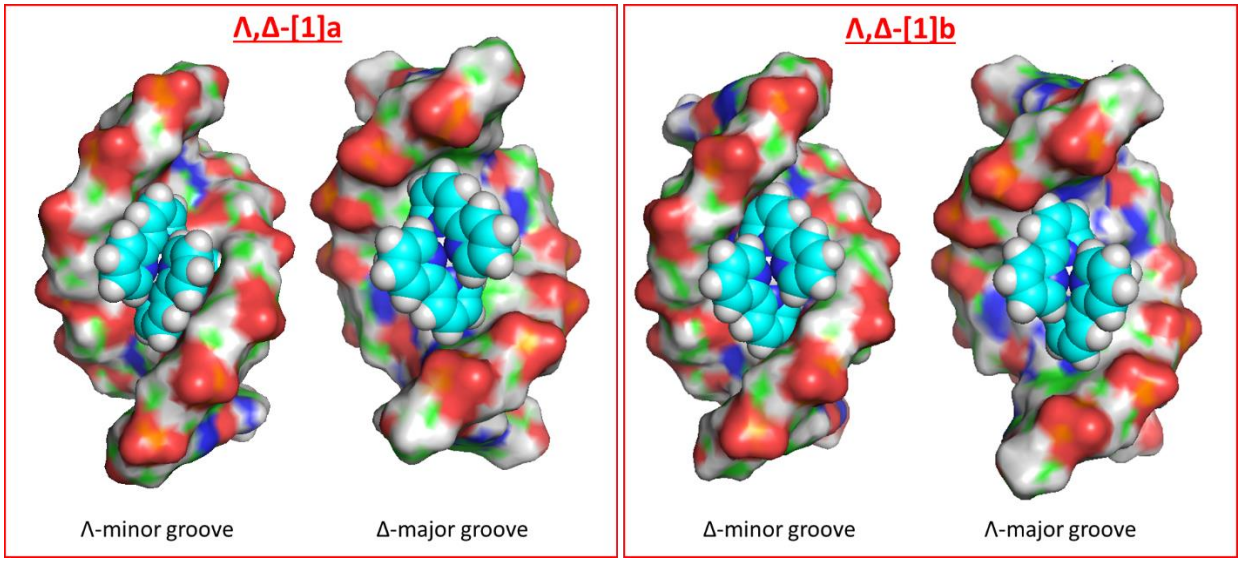


Figure 5.9. Views of the major and minor groove of the Λ,Δ-[1]a and Λ,Δ-[1]b complex conformations.

This prediction of the Λ,Δ-[1]a with the Λ centre in the minor groove and the Λ,Δ-[1]b with the Λ in the major groove is in agreement with the NMR data that shows the Λ,Δ-[1]b conformation has the TMe protons shifted very far upfield in comparison to the Λ,Δ-[1]a conformation. Inspection of the intercalation site with a view down the helical axis (figure 5.10) shows both conformations have TMe aligned directly over the tpphz aromatic system producing a shielding effect as a result of the induced magnetic field of the circulating π electrons. In the Λ,Δ-[1]b conformation, the bipyridyl units in the major groove also give a shielding effect with the TMe protons aligned over the aromatic system. The two aromatic ring current shift effects give rise to upfield resonances seen in the NMR data (T_4 and T_{12}, $\Delta\delta \sim -1.6$ ppm). However, in the Λ,Δ-[1]a system the TMe are adjacent to the circulating π electrons resulting in a deshielding effect. This is opposed to the shielding achieved through the alignment with the central tpphz system which is reflected in the more moderate upfield perturbation to TMe chemical shifts seen for this conformation (T_4, $\Delta\delta$ -0.96 ppm, T_{12}, $\Delta\delta$ -0.49 ppm).

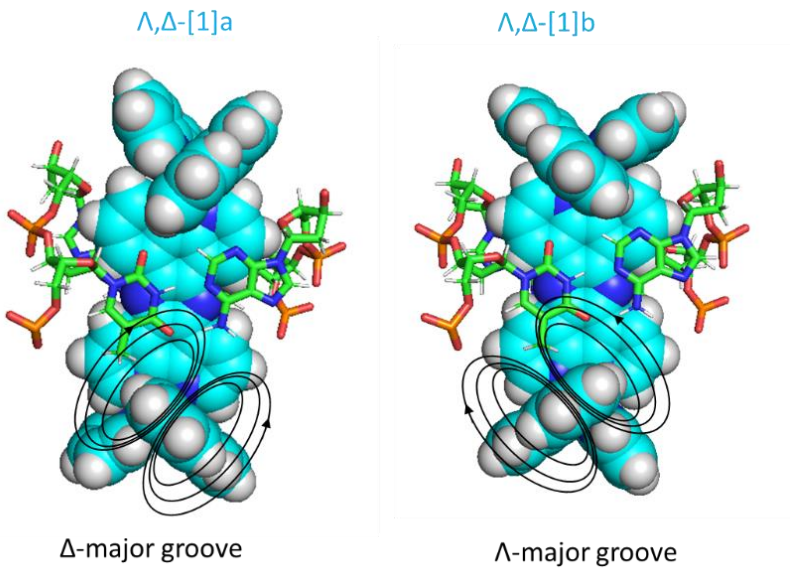


Figure 5.10. Views down the helical axis of the Λ,Δ-[1]a and Λ,Δ-[1]b intercalation sites showing the deshielding (left) and shielding (right) ring current effect of the ancillary bipyridyl ligands in the major groove. Indicated in black are the induced magnetic field lines from the circulating π systems.

5.5 Simulated Annealing of Λ,Λ-[1]

The simulated annealing calculations for the Λ,Λ-[1] DNA complex were performed in the same way as described previously starting from the coordinate file containing both d(GCATATCG).(CGATATGC) and Λ,Λ-[1] (figure 5.11).

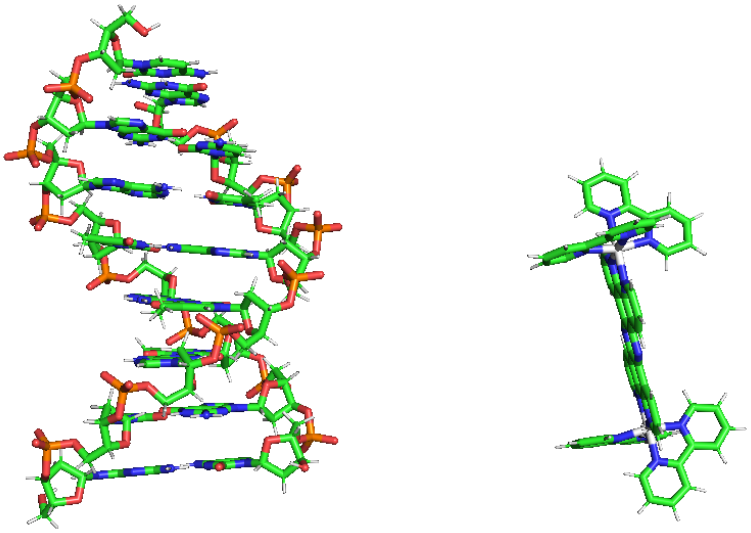


Figure 5.11. The starting coordinate file used to perform the restrained molecular dynamics simulations on the binding Λ,Λ-[1] to d(GCATATCG).(CGATATGC)

Restraint files were created from the experimentally derived NOEs from Λ,Λ-[1]a and Λ,Λ-[1]b. The files contained 30 experimentally derived NOE

restraints for the Λ,Λ-[1]a conformation and 23 for the Λ,Λ-[1]b form. Initially these were provided as ambiguous restraints and after subsequent MD runs were refined as the specific NOEs give in. These NOEs were again classified as either strong (1.8 – 3 Å), medium (1.8 – 4 Å), or weak (1.8 – 5 Å) and given in tables 5.4 and 5.5. The same set of DNA restraints were used to maintain a duplex structure during the MD runs.

Table 5.4. The specific intermolecular NOEs between the DNA duplex d(GCATATCG).(CGTATAGC) and Λ,Λ-[1]a used as restraints for restrained molecular dynamics calculations. (s), (m) and (w) indicate strong, medium or weak NOE groupings.

Λ,Λ-[1]a	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12
T4							Me(m)	Me(m)	Me(m) 2'(w) 2"(w)	1'(s) 2'(s) 2"(s)	1'(m) 2'(m) 2"(m)	1'(m) 2'(m) 2"(m)
A5										1'(m)	1'(m)	1'(m)
T12	1'(m)	1'(w)	1'(s) 2'(s) 2"(s)	Me(m) 2'(w) 2"(w)	Me(m)	Me(m)						
A13	1'(m)	1'(m)	1'(m)									

Table 5.5. The intermolecular NOEs between the DNA duplex d(GCATATCG).(CGTATAGC) and Λ,Λ-[1]b used as constraints for restrained molecular dynamics calculations. (s), (m) and (w) indicate strong, medium or weak NOE groupings.

Λ,Λ-[1]b	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12
C2							5(w)	5(w)	5(m) 1'(w)	1'(s) 2'(s) 2"(s)	1'(s) 2'(s) 2"(s)	1'(m)
A3										1'(w)	1'(w)	1'(m)
T14			1'(s) 2'(s) 2" (s)	Me(m) 2'(m) 2"(m)	Me(m)	Me(m)						
G15		1'(m)	1'(w)									

The resultant lowest energy structures of these MD simulations are given in figure 5.12. As predicted from the NMR studies, the addition of NOE restraints for the Λ,Λ-[1]a conformation generated the structure threaded into the central (T_4,A_{13}) | (T_{12},A_5) binding site. The Λ,Λ-[1]b conformation yielded a structure bound at the (C_2,G_{15}) | (A_3,T_{14}) site. The ambiguous NOEs used in the initial simulation were then converted into the non-ambiguous NOEs by inspection of the bound structures and the MD repeated to give the final energy minimised conformations seen in (figure 5.12).

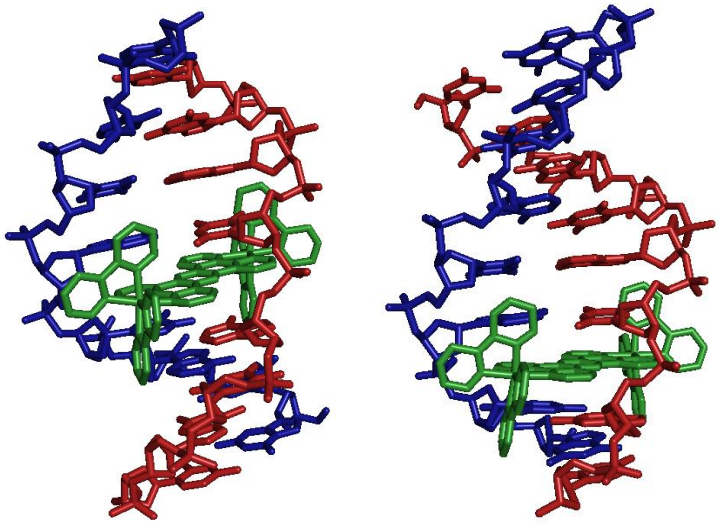


Figure 5.12. Resultant structures of restrained molecular dynamics calculations showing (left) Λ,Λ-[1]a and (right) Λ,Λ-[1]b.

A closer look at the binding site shows that both of the conformations of Λ,Λ-[1] are embedded deeply into the minor groove of the DNA with the other ruthenium centre projecting into the major groove. This maximises electrostatic interactions between the cationic ruthenium and anionic phosphate backbone. The aromatic tpphz ligand is again inserted into the base stack, maximising π-π stacking interactions and extending the long axis of the DNA at the intercalation site (figure 5.13). A slight angled intercalation is observed, similar to that of the binding orientation seen for Λ-$[Ru(phen)_2dppz]^{2+}$.[13]

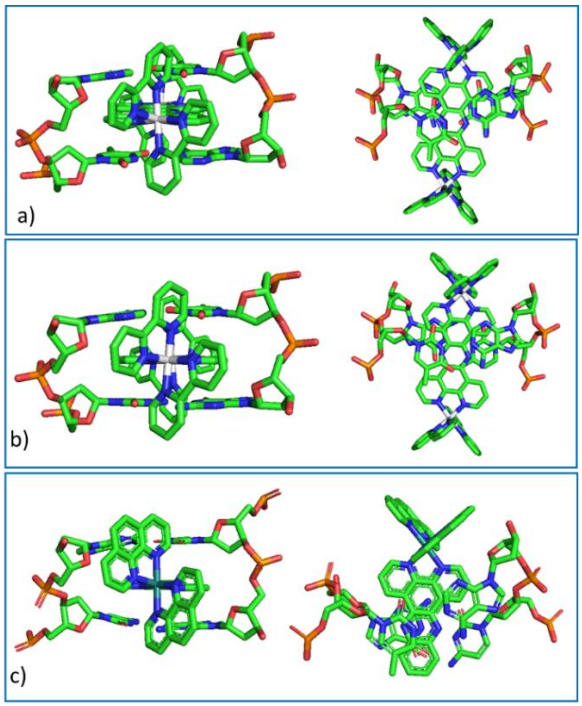


Figure 5.13. Views from the minor groove and down the helical axis of DNA duplexes showing (top) Λ,Δ-[1] intercalation site (T_4,A_{13})|(T_{12},A_5) from restrained molecular dynamics simulations and (bottom) showing the canted form of Λ-$[Ru(phen)_2dppz]^{2+}$.[13]

The surface filling views from the major and minor groove of both Λ,Λ-[1]a and Λ,Λ-[1]b show the deep binding of Λ,Λ-[1] in the minor groove at both binding sites (figure 5.14). By inspection it can be seen that the bipyridyl ligands run with a twist parallel to those of the grooves giving rise to favourable contacts.

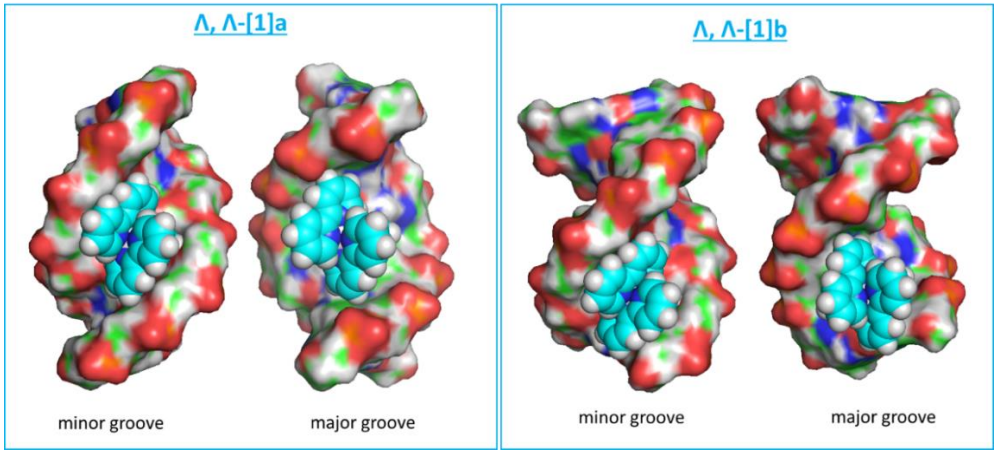


Figure 5.14. Views from the major and minor grooves of simulated annealing studies of the Λ,Λ-[1]a and Λ,Λ-[1]b conformations.

Inspection of the two intercalation sites of the Λ,Λ-[1]a and Λ,Λ-[1]b with a view down the helical axis gives more information as to the chemical shift of the TMe groups in major groove (figure 5.15). Both aromatic systems of the tpphz and the bipyridyl units are cooperatively shielding the TMe protons from the external magnetic field. This is due to the alignment of the TMe directly over the circulating π electrons from both moieties and in both conformations. This gives rise to the large chemical shift perturbations seen in both Λ,Λ-[1]a (T_4, Δδ -1.61 ppm, T_{12}, Δδ -1.35 ppm) and Λ,Λ-[1]b (T_{14}, Δδ -2.17 ppm).

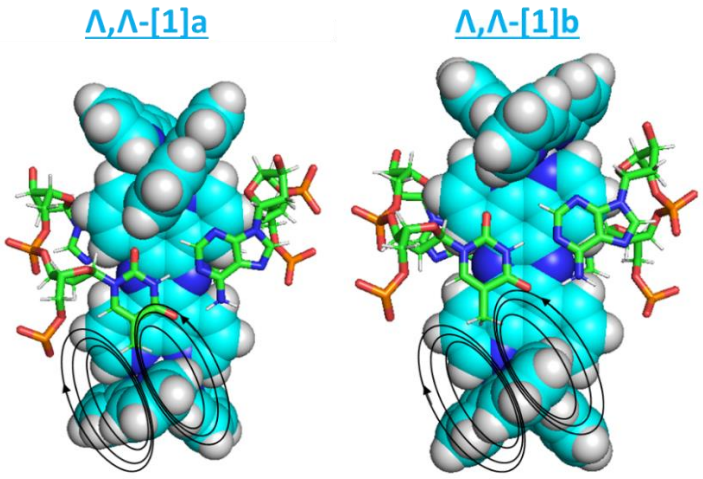


Figure 5.15. Views down the helical axis of the Λ,Λ-[1]a and Λ,Λ-[1]b intercalation sites showing the shielding ring current effect of the ancillary bipyridyl ligands in the major groove. Indicated in black are the induced magnetic field lines from the circulating π systems.

5.6 Comparison of Δ,Δ-[1] to Λ,Δ-[1]a

Although an assignment of the Δ,Δ-[1]:DNA complex could not be made due to the multiple exchange peaks, NMR experiments suggest that at least one of the exchanging species must be an intercalated site. Chemical shifts of the TMe protons in the Δ,Δ-[1]a species show similarities to that of the Λ,Δ-[1]a and thus these NOEs were used as a guide. With this in mind a simulated annealing was attempted using the experimentally derived NOEs from Λ,Δ -[1]a to visualise the binding of Δ,Δ-[1]. The simulated annealing was carried out using exactly the same set of parameter topology and restraint files as the Λ,Δ-[1] complex. The Δ,Δ-[1] file was used instead of the Λ,Δ-[1]a. However, the resultant structures were not intercalated and gave unusual DNA conformations. Therefore, in order to visualise how Δ,Δ-[1] might interact with DNA in the Δ,Δ-[1]a conformer, a simple pair fitting was done through PyMOL software, replacing Λ,Δ-[1] with Δ,Δ-[1] from the restrained MD refined model (figure 5.16).

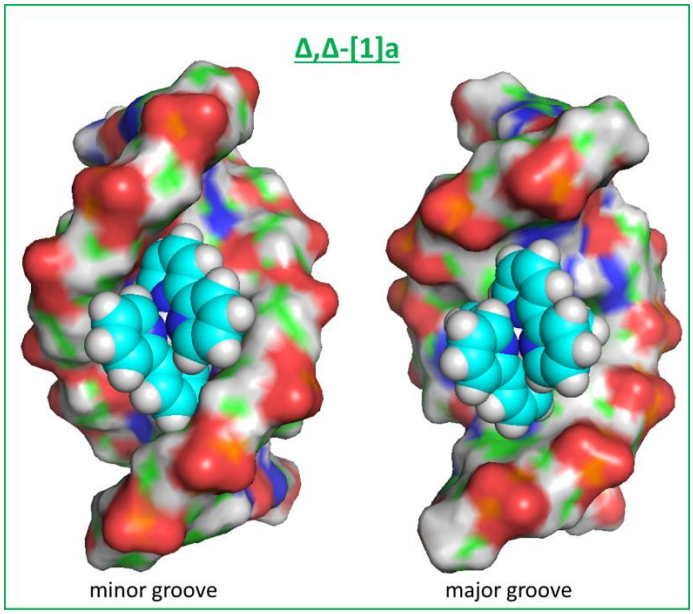


Figure 5.16. Model of the binding of Δ,Δ-[1] to d(GCATATCG).(CGATATGC). Views from the minor and major groove manually docked using pair fitting from the Λ,Δ-[1]a structure.

The resultant model shows unfavourable close contacts within the vdw radii between the bipyridyl protons and the DNA. To accommodate the Δ,Δ-[1]a conformation, the DNA would require a larger deformation than that of the Λ,Δ -[1]a complex. The twist of the bipyridyl ligands in both chiral centres run

perpendicular to that of the grooves, whereby the protons point in towards the proximal DNA strand (figure 5.17). In Λ,Δ-[1]a, a Λ chiral centre is present and can compensate for the unfavourable steric interactions caused by the Δ centre. This means it can still form very a stable complex. Conversely, Δ,Δ-[1] with two Δ centres make a less stable threaded species which exhibits a similar energy to the groove bound form.

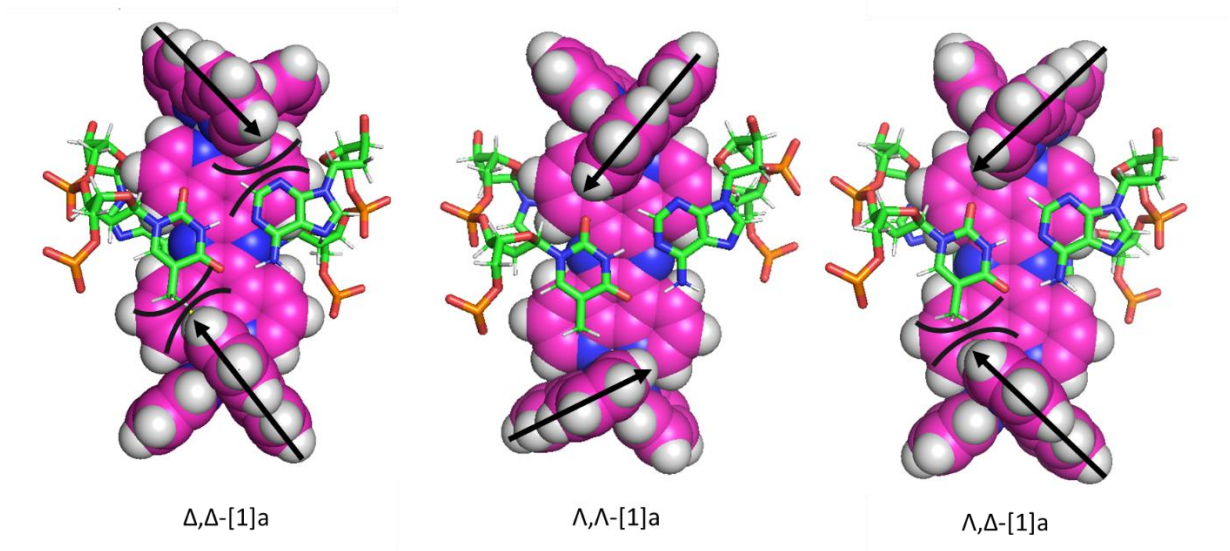


Figure 5.17. Views down the helical axis of the 3 stereoisomers of [1] bound to d(GCATATCG).(CGATATGC). The arrows indicate the orientations of the bipyridyl ancillary ligands and the curves indicate unfavourable steric clashes.

5.7 Discussion

Restrained molecular dynamics have been employed to model both the Λ,Δ-[1] and Λ,Λ-[1] using experimentally derived NOE restraints. The models observed for each conformation adhere well to the experimental restraints and show plausible structures in which the metal centres are deeply inserted into the minor groove of the DNA, similar to that seen from monointercalating parent complexes.[14,15] This deep insertion into the minor groove has also been proposed as the binding mode for other bis-intercalating threading complexes.[16] Care must be taken when looking at the larger structure, as many restraints put on these models assume a conformation close to that of B-DNA outside of the intercalation site, which may not be the case in solution. That said, the ^{31}P and ^{1}H chemical shifts of residues outside of the intercalation site are broadly similar to the free DNA structure suggesting this assumption is correct. The models show that the alignment of the bipyridyl units relative to the duplex in each chiral centre play a large role in the binding event, with Λ chiral centres being favourable. This accounts for the observed binding of Λ,Λ-[1] at the (C_2A_3)/($T_{14}G_{15}$) base step as well as the most accessible (T_4A_5)/($T_{12}A_{13}$) step. It is predicted that two Δ units, as in Δ,Δ-[1]a, create unfavourable steric interactions, lowering the difference in stability of the groove bound and intercalated forms of Δ,Δ-[1], which allows for equal population of both conformations.

5.8 References

1 A. T. Brünger, *X-PLOR Version 3.1: A system for X-ray crystallography and NMR*, Yale University Press., 1992.

2 R. L. Fox and E. L. Stanton, *AIAA J.*, 1968, **6**, 1036–1042.

3 H. J. Dyson and P. E. Wright, *Methods Enzymol.*, 2005, **394**, 299–321.

4 D. Neuhaus and M. Williamson, *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, Wiley-VCH, 2nd edition, 2000.

5 A. T. Brünger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson and G. L. Warren, *Acta Crystallogr., Sect. D Biol. Crystallogr.*, 1998, **54**, 905–921.

6 A. T. Brunger, *Nat. Protoc.*, 2007, **2**, 2728–2733.

7 T. Wilson, P. J. Costa, V. Félix, M. P. Williamson and J. A. Thomas, *J. Med. Chem.*, 2013, **56**, 8674–83.

8 P. Waywell, J. A. Thomas and M. P. Williamson, *Org. Biomol. Chem.*, 2010, **8**, 648–54.

9 S. Arnott, P. J. Campbell-smith and R. Chandrasekaran, *In Handbook of Biochemistry and Molecular Biology*, CRC Press, Cleveland, 3rd edition., 1976.

10 G. J. Kleywegt, *Newsl. Protein Crystallograpy*, 1995, **31**, 45–50.

11 G. M. Clore and A. M. Gronenborn, *FEBS Lett.*, 1984, **175**, 117–123.

12 H. Song, J. T. Kaiser and J. K. Barton, *Nat. Chem.*, 2012, **4**, 615–20.

13 J. P. Hall, D. Cook, S. R. Morte, P. McIntyre, K. Buchner, H. Beer, D. J. Cardin, J. A. Brazier, G. Winter, J. M. Kelly and C. J. Cardin, *J. Am. Chem. Soc.*, 2013, **135**, 12652–12659.

14 H. Niyazi, J. P. Hall, K. O'Sullivan, G. Winter, T. Sorensen, J. M. Kelly and C. J. Cardin, *Nat. Chem.*, 2012, **4**, 621–628.

15 J. P. Hall, D. Cook, S. R. Morte, P. McIntyre, K. Buchner, H. Beer, D. J. Cardin, J. A. Brazier, G. Winter, J. M. Kelly and C. J. Cardin, *J. Am. Chem. Soc.*, 2013, **135**, 12652–9.

16 B. Önfelt, P. Lincoln and B. Nordén, *J. Am. Chem. Soc.*, 2001, **123**, 3630–3637.

Chapter 6. Induced Circular Dichroism

6.1 Threading Intercalation Studies

NMR data has revealed that all three stereoisomers of [1], bind in some form, to the short oligonucleotide sequence d(GCATATCG).(CGATATGC), through a threading mode. As discussed in chapter 1 it has been shown by Turro, *et al*. that threading of [1] into both genomic CT and herring sperm DNA is not possible, even with high temperature annealing.[1] However, Lincoln, *et al*. have shown that shortening of the bridge in dinuclear species inhibits threading into less flexible DNA such as CT.[2] The shorter bridged systems only thread into more flexible poly d(AT) DNA with annealing.[3,4] Therefore, a study aimed at delineating whether [1] can also thread into flexible extended DNA sequences was carried out. Although these systems are too complex to be studied by NMR spectroscopy, due to the significant spectral overlap associated with long DNA sequences, it is possible to get some indication of binding orientation by other spectroscopic methods. Lincoln, *et al*. have used two main spectroscopic methods employing polarised light spectroscopy to study this threading phenomenon: linear dichroism (LD) and circular dichroism (CD).[5–7]

6.2 Induced Circular Dichroism

As discussed earlier, CD is a spectroscopic technique that probes the differential absorbance of left (A_L) and right (A_R) circularly polarised light, by chiral molecules in solution and is defined by equation 6.1.

$$CD = A_L - A_R$$

Equation 6.1. Definition of circular dichroism.

Although isolated DNA bases are not chiral, they show a small CD signal between 180-300 nm when a chiral sugar moiety is introduced to the structure as a nucleotide. The strength of this CD signal increases dramatically due to the formation of higher order chiral structures such as the right handed helix of B-DNA. The orientation of these structures and thus DNA conformation can be

predicted by observation of the CD signature.[8] This effect can also be observed for small molecules upon binding to DNA, as the molecule is aligned in some way with the chiral DNA structure; even achiral molecules may display a CD signal. This effect, known as induced circular dichroism (ICD), can be used to monitor the binding of small molecules to DNA. It has been shown that the ICD for an intercalating molecule is sensitive to the orientation of the bound species. A negative ICD is observed for an electronic transition parallel to the base pair long axis, whilst a positive ICD is seen for a transition perpendicular to long axis of the DNA (figure 6.1).[9,10]

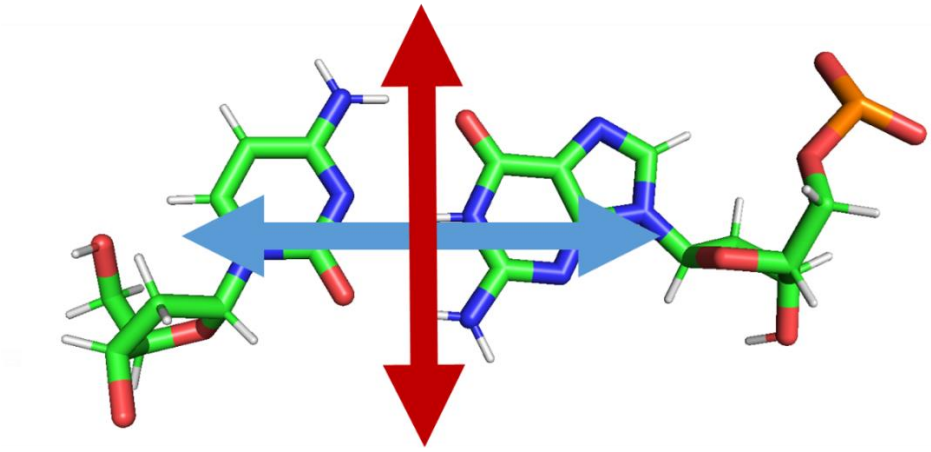


Figure 6.1. A schematic illustration of the orientation of electronic transition moment relative to the long axis of a base pair for an intercalating molecule. The blue arrow represents a parallel transition with an associated negative ICD and the red arrow a perpendicular transition with a positive ICD.

With this in mind, it is possible to look at the ICD of [1] whilst binding to the oligonucleotide sequence d(GCATATCG).(CGATATGC). The MLCT transition of the Λ,Δ-[1] discussed in chapter 1, absorbs at around 450 nm upon binding to the oligonucleotide sequence. This is useful as the DNA has no absorbance at this wavelength and thus zero CD signal. The symmetry of Λ,Δ-[1] dictates that the chirality of both ruthenium centres are equal and opposite and therefore cancel out and show no CD signal from the free [1]. Upon addition of 500 μM Λ,Δ-[1] to the 2 mM solution of d(GCATATCG).(CGATATGC) in 150 mM NaCl, 5 mM tris buffer solution at pH 7.4, a structured positive ICD is seen with a maximum (λ_{max}) 465 nm. There are also two shoulder peaks at (λ_{max}) 408 nm and (λ_{max}) 370 nm (figure 6.2).

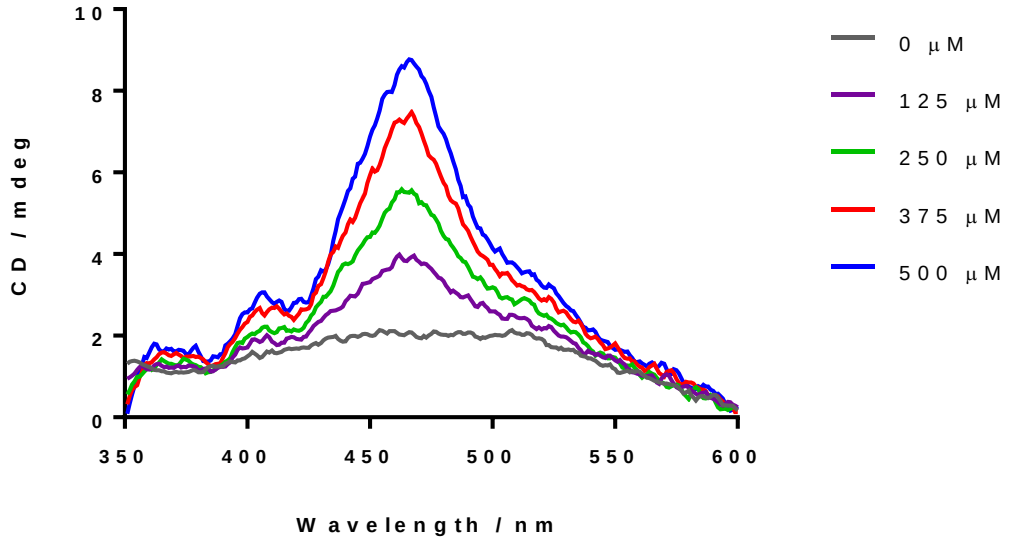


Figure 6.2. The addition of 500 μm Λ,Δ-[1] in increments of 125 μM to a solution of 2 mM d(GCATATCG).(CGATATGC) in 5 mM tris buffer at pH 7.4 with 150 mM NaCl monitored by CD.

The above data is consistent with the NMR based models reported in this thesis, with the MLCT transition being perpendicular to the base pair long axis in a threaded mode and giving a positive CD. This structured CD signal is strikingly similar to that seen by Lincoln, *et al.* for their threaded species.[11] Lincoln, *et al.* have also demonstrated that threading intercalation is associated with a shift to a positive more structured ICD.[12] The CD spectrum of Λ,Δ-[1] additions to d(GCATATCG).(CGATATGC) in buffer shows a very similar ICD trace (figure 6.3). They have also shown that threading intercalation can be monitored by the increase in the ICD upon annealing of the threaded species, resulting in a strong positive ICD in the visible region. With this in mind the binding of Λ,Δ-[1] to CT and poly d(AT) DNA was tested in order to see if threading occurs in either of these species (figure 6.3).

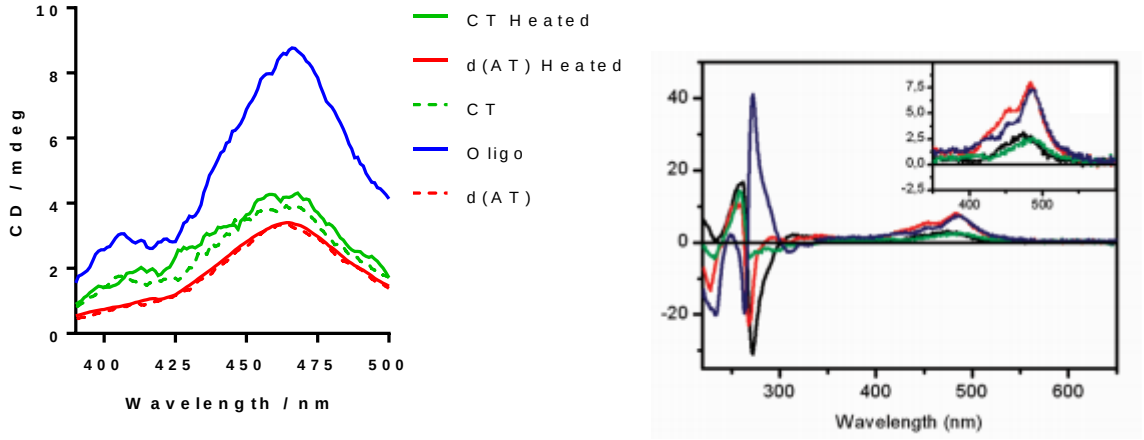


Figure 6.3. Right: the ICD of 0.5 mM Λ,Δ-[1] in 5 mM tris buffer at pH 7.4 with 150 mM containing (red) 2 mM per base pair poly d(AT) DNA before (solid) and after (dashed) heating at 60° for 48 hours, (green) 2 mM calf thymus DNA before (solid) and after (dashed) heating at 60° for 48 hours, (blue) 2 mM d(GCATATCG).(CGATATCG). Left: for comparison the ICD seen for $[\{Ru(phen)_2\}_2bidppz]^{4+}$ undergoing threading where the green and black lines are the non-threaded dinuclear ruthenium complexes and the blue and red are threaded species from reference blank.[11]

Inspection of the initial CD spectrum after addition of Λ,Δ-[1] to AT and CT DNA reveals a small positive ICD upon complexation to both AT and CT DNA. This is similar to that seen for the initial groove bound $[\{(Ru(phen)_2\}_2(bidppz)_2]^{2+}$ complexes.[11] To see if it was possible to generate a threaded species for these DNA sequences, an annealing was attempted by heating the samples containing the DNA to 60 °C and leaving them for 48 hours. After slow cooling to room temperature, the CD spectra were observed again to see if any reorientation of the complex had occurred. There was no identifiable change in the ICD for either CT DNA or poly d(AT). This is an indication that no threading occurred.

6.3 Discussion

Whilst circular dichroism can probe the binding orientations of small molecules to DNA, it is far from the most accurate method of doing so. A more accurate method of this would be through the employment of linear dichroism as discussed in chapter 1.[13] Results using CD to predict orientation should be treated with caution as there are multiple binding orientations that give a positive or negative signal. Thus CD can be ambiguous with regards to conformation. Although LD could not be obtained for these systems without the facility, the CD profile of the complexes do look strikingly similar to those of related compounds. The effect of heating the sample to 60 °C for 48 hours seems to have no noticeable effect on the CD spectrum of Λ,Δ-[1] with both CT and poly d(AT) DNA. This indicates that in order to bind through a threading mode, [1] must need a very unstable and flexible DNA sequence. It has been shown that for ruthenium polypyridyl complexes, reducing the length and flexibility of the bridging ligand inhibits threading through CT DNA and slows the threading through polyd(AT). The tpphz bridge in [1] is the shortest bridge that has been shown to bind to DNA by a threading mode. Compared to previous threading complexes, such as the bidppz threading unit, it is also more rigid. It is therefore possible that this is the shortest bridged ruthenium complex that can feasibly bind with this mode.

6.4 References

1 D. A. Lutterman, A. Chouai, Y. Liu, Y. Sun, C. D. Stewart, K. R. Dunbar and C. Turro, *J. Am. Chem. Soc.*, 2008, **130**, 1163–1170.

2 J. R. Johansson, Y. Wang, M. P. Eng, N. Kann, P. Lincoln and J. Andersson, *Chem. A Eur. J.*, 2013, **19**, 6246–6256.

3 P. Nordell, F. Westerlund, L. M. Wilhelmsson, B. Nordén and P. Lincoln, *Angew. Chemie*, 2007, **119**, 2253–2256.

4 J. Andersson, M. Li and P. Lincoln, *Chem. - A Eur. J.*, 2010, **16**, 11037–11046.

5 Rodger A. and B. Norden, *Circular Dichroism and Linear Dichroism*, Oxford University Press, 1997.

6 F. Westerlund, P. Nordell, J. Blechinger, T. M. Santos, B. Nordén and P. Lincoln, *J. Phys. Chem. B*, 2008, **112**, 6688–6694.

7 F. Westerlund, M. P. Eng, M. U. Winters and P. Lincoln, *J. Phys. Chem. B*, 2007, **111**, 310–317.

8 J. Kypr, I. Kejnovska, D. Renciuk and M. Vorlickova, *Nucleic Acids Res.*, 2009, **37**, 1713–1725.

9 O. Buchardt, T. Koch, M. Wirth, B. Nordén and P. E. Nielsen, *J. Am. Chem. Soc.*, 1988, **110**, 932–939.

10 H.-C. Becker and B. Nordén, *J. Am. Chem. Soc.*, 1997, **119**, 5798–5803.

11 J. Andersson and P. Lincoln, *J. Phys. Chem. B*, 2011, **115**, 14768–14775.

12 L. M. Wilhelmsson, E. K. Esborner, F. Westerlund, B. Norden and P. Lincoln, *J. Phys. Chem. B*, 2003, **107**, 11784–11793.

13 M. Eriksson and B. Nordén, *Methods Enzymol.*, 2001, **340**, 68–98.

Chapter 7. Conclusions and Future Work

7.1 Conclusions

The results reported in this thesis demonstrate that the three stereoisomers of the DNA binding compound $[\{Ru(bipy)_2\}_2(tpphz)]^{4+}$ all bind differently to the octanucleotide sequence d(GCATATCG).(CGATATGC). Despite the many studies on such systems, this is the first time threading interactions of ruthenium polypyridyl complexes have been observed through NMR experiments in solution. Further inspection of the structures illustrate that chirality plays an important role in the binding of such complexes to B-DNA. It seems that ancillary ligand orientation is a key factor in the threading process, with the Δ ruthenium centre displaying unfavorable steric interactions when intercalated.

The three stereoisomers display different binding selectivity and kinetics: Λ,Λ-[1] and Δ,Λ-[1] exhibit slow exchange kinetics as they bind to the octanucleotide through threaded intercalation. The Δ,Δ-[1]/DNA complex exists in two bound states in intermediate exchange on the NMR timescale, which suggests a dynamic exchange between minor groove binding and intercalation. A comparison between these results and previous studies[6,7,27] suggests that although intercalation into stable duplex DNA is very slow, intercalation is very much faster when the DNA is less stable and thus more flexible (as with poly d(AT)) or is destabilised (here, by being very short). It is interesting that even though the complexes are shown to bind through threading to the short oligonucleotide studied, they do not thread in to poly d(AT) DNA even with heating. The systems studied here have very short rigid bridges and thus need extremely flexible or denatured DNA in order to bind by a threaded mode. It is therefore possible that these compounds could be used as leads in the development of probes that label DNA where local destabilisation of the duplex has occurred, such as transcription bubbles.[1] It is also possible that threading may occur by insertion into destabilised mismatch sites as seen for the parent mononuclear $[Ru(bipy)_2dppz]^{2+}$ complex.[2,3]

7.2 Future Work

Future work with regards to these systems could include the growth of crystals containing the threaded complexes bound to DNA suitable for X-ray crystallographic studies. Structures from crystals would complement and validate the NMR derived models. To further this research, it would also be beneficial to investigate the effect of changing the structure of the dinuclear threading complex on the binding mode to the oligonucleotide sequence. This could be done by altering the ancillary ligands to make them bulkier. Initial studies with the unresolved non-deuterated phenanthroline system [2] suggest that a bigger ancillary ligand may slow the rate of threading due to the lack of exchange doublets in the NOESY NMR spectrum. Equally it might be useful to inhibit threading by increasing the stability of the DNA, using a longer sequence or a sequence without T/A or CA/TG base steps as suitable intercalation sites. This could trap the proposed initial groove bound form of the complex and allow structural investigation.

7.3 References

1 M. D. Bick, C. S. Lee and C. A. Thomas, *J. Mol. Biol.*, 1972, **71**, 1–9.

2 H. Song, J. T. Kaiser and J. K. Barton, *Nat. Chem.*, 2012, **4**, 615–20.

3 P. Nordell, F. Westerlund, A. Reymer, A. H. El-Sagheer, T. Brown, B. Nordén and P. Lincoln, *J. Am. Chem. Soc.*, 2008, **130**, 14651–14658.

Chapter 8. Methods

8.1. General

8.1.1 Chemicals

Chemicals purchased from commercial vendors were used as supplied unless otherwise specified. All heavy metal compounds were assumed and treated as toxic.

8.1.2 Mass Spectrometry

Electrospray ionisation mass spectra (ESI) were recorded using a Micromass LCT ES-TOF machine. The spectra were obtained by the University of Sheffield Mass Spectrometry Service located within the Chemistry department.

8.1.3 Nuclear Magnetic Resonance (NMR) Spectroscopy

^{1}H-NMR experiments were carried out on a Bruker Avance 800 MHz or 400 MHz spectrometer working in fourier transform mode. ^{31}P and ^{31}P-^{1}H experiments were carried out on a Bruker Avance 500 MHz spectrometer. The abbreviations associated with the ^{1}H-NMR spectra are as follows; s-singlet, d-doublet, dd-doublet of doublet, t-triplet, q-quartet and m-multiplet.

8.1.4 X-Ray Crystallography

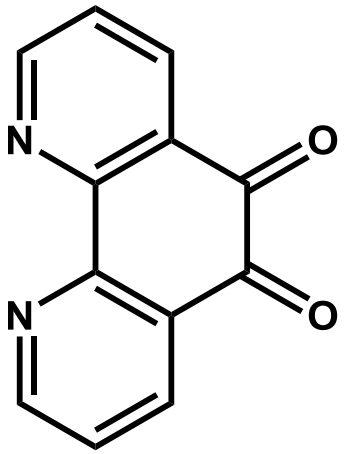
X-ray quality crystals of the chloride salts of ruthenium complexes were grown using slow vapour diffusion of acetone anti-solvent into a concentrated solution of pure complex in methanol, which evolved ruby red crystals suitable for structure determination. X-ray quality crystals of the hexafluorophosphate salts of ruthenium complexes were grown by a similar method using nitromethane solvent and diethyl ether anti-solvent.

Crystal structures were then determined within the Sheffield Chemistry Department by Dr. Craig Robertson. The intensity data was obtained on either a Bruker Kappa Apex-II CCD or Bruker Kappa Apex-II diffractometer operating

with a MoKα sealed-tube X-ray source, or a Bruker D8 Venture diffractometer equipped with a Photon 100 CMOS detector using a CuKα microfocus X-ray source at 100 K. The reflections were corrected for absorption via empirical methods (SADABS) based upon symmetry equivalent reflections combined with measurements at varied azimuthal angles.[1,2] The crystal structure was solved and refined against F^2 values using ShelXT for solution and ShelXL for refinement through the Olex2 program.[170,171,5] Non-hydrogen atoms were refined anisotropically and hydrogen atoms were placed in calculated geometries and refined utilizing a riding model and isotropic displacement parameters. Crystallographic data is recorded in the appendix.

8.2. Synthesis

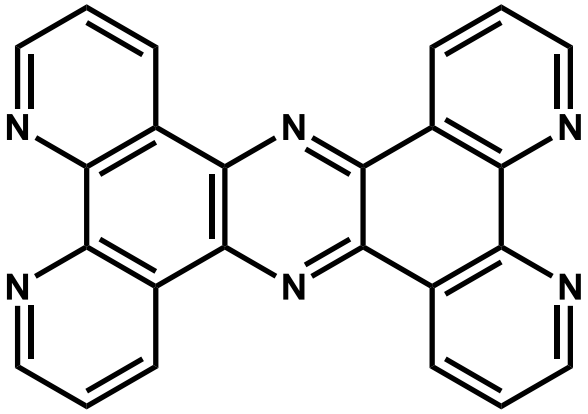
8.2.1 1,10-Phenanthroline-5,6-Dione (dpq)



1,10-phenanthroline (18.21 g, 100 mmol) was dissolved in 125 mL of 60 % sulphuric acid. Addition of potassium bromate (66.8 g, 400 mmol) was done slowly over 10 minutes with vigorous stirring. The remaining suspension was left to cool slowly to room temperature. The reaction mixture was then cooled in an ice bath followed by the addition of ice (100 g) prior to careful neutralization to pH 5-6 using a saturated sodium hydroxide (400 mL, 20 M) solution. The yellow precipitate was filtered and washed with water (1000 mL) and diethyl ether (100 mL) to yield dpq as a yellow powder.[6]

Yield: 15.87 g, 76.21 %; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.63 (dd, J = 4.7, 7.9 Hz, 2H), 8.51 (dd, J = 1.8, 7.9 Hz, 2H), 9.12 (dd, J = 1.8, 4.7 Hz, 2H). ESI-MS: m/z (%): 211 (100) $[MH]^+$.

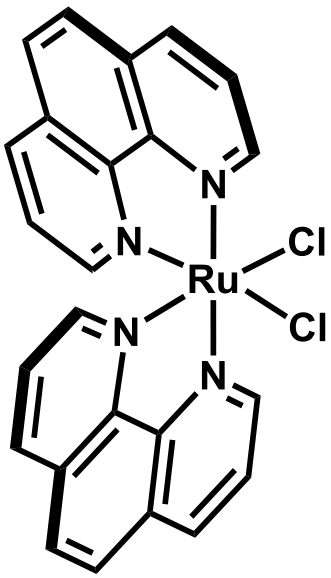
8.2.2 Tetrapyrido[3,2-a:2',3'-c:3",2"-h:2"',3"'-j]phenazine (tpphz)



1,10-phenanthroline-5,6-dione (2.90 g, 13.6 mmol), sodium hydrosulfite (0.30 g, 1.72 mmol) and ammonium acetate (15.0 g), were heated under inert argon to 180 °C. This temperature was maintained for 2 hours with occasional stirring. Once cool, following the addition of water (20 mL), the precipitate was collected by filtration. This solid was then washed with methanol, water and acetone and triturated in refluxing ethanol (100 mL) followed by a hot filtration. The product was then washed with ethanol, and dried *in vacuo* giving tpphz as a pale yellow powder.[7]

Yield: 1.53 g, 28.7 %; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.88 (dd, J = 4.4, 8.1 Hz 4H), 9.42 (dd J = 1.8, 4.4 Hz, 4H), 9.7(dd, J = 1.8, 8.1 Hz, 4H). ESI-MS: m/z (%): 385 (100) $[MH]^+$.

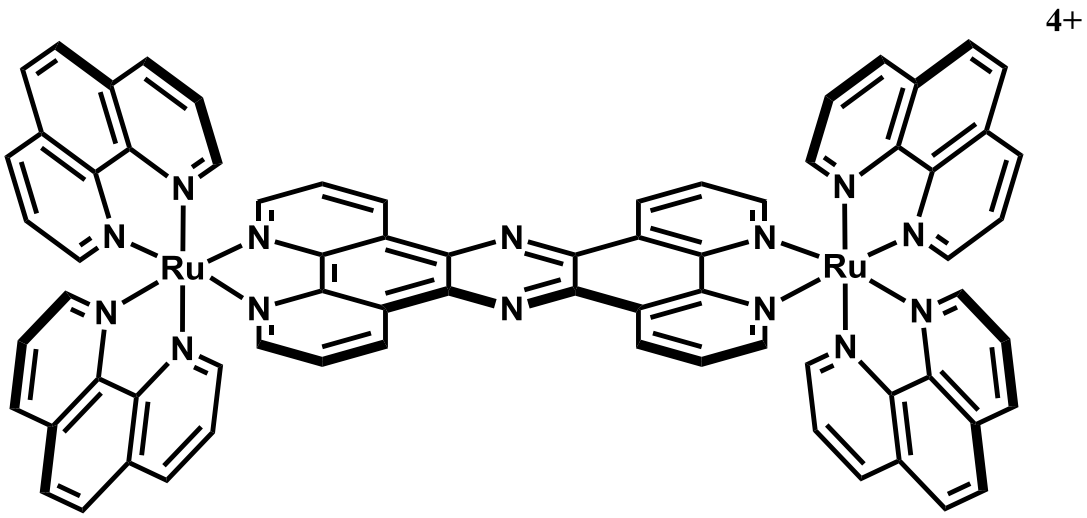
8.2.3 $[Ru(phen)_2Cl_2].2H_2O$



$RuCl_3.3H_2O$ (3 g, 11.53 mmol), LiCl (3.1 g, 73.8 mmol) and 1,10-phenanthroline (4.1 g, 22.7 mmol) were heated under reflux for a period of 8 hours in dimethylformamide (DMF) (40 mL). The solution was then left to cool to room temperature. After addition of acetone (200 mL), the solution was kept at 4 °C for 14 hours. The precipitate was then collected by filtration and washed with water followed by diethyl ether. The solid was dried under vacuum to yield $[Ru(phen)_2Cl_2].2H_2O$ as a black powder.[8]

Yield: 3.98 g, 3.98 g, 64.5 %; ^1H NMR (400 MHz, CD_3CN) δ 8.65 (d, J = 9.6 Hz, 4H), 8.27 (s, 4H), 8.05 (d J = 4.8 Hz, 4H), 7.64 (dd, J = 4.8, 9.6 Hz, 4H). ESI-MS m/z: 497 (100) $[M^+-Cl]$.

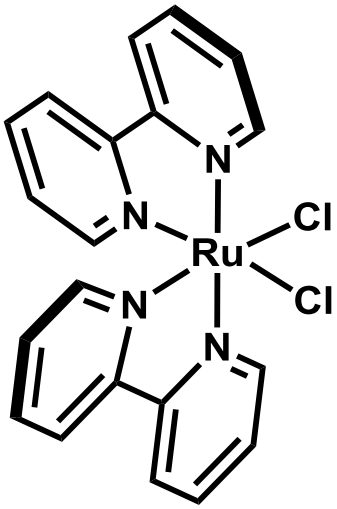
8.2.4 Racemic $[(Ru(phen)_2)_2tpphz][PF_6]_4$



$Ru(phen)_2Cl_2.2H_2O$ (1.0 g, 1.9 mmol) and tpphz (0.3 g, 0.78 mmol) were suspended in a 1 : 1 solution of EtOH:H_2O (50 mL). The suspension was refluxed for 12 hours under inert conditions. After cooling to room temperature, the suspension was then kept at 4 °C overnight. The resulting brown coloured solution was filtered and concentrated under vacuum. Addition of ammonium hexafluorophosphate (NH_4PF_6) produced an orange precipitate of $[(Ru(phen)_2)_2(tpphz)][PF_6]_4$ as the hexafluorophosphate salt. This solid was collected by filtration and washed with water (50 mL) before being dried under vacuum.[9]

Yield: 0.4 g, 62 %; ^{1}H NMR (400 MHz, CD_3CN) δ 9.95 (d, J = 8.3 Hz, 8H), 8.66 (d, J = 8.3 Hz, 8H), 8.31 (s, 8H), 8.28 – 8.23 (m, 2H), 8.07 (d, J = 4.9 Hz, 4H), 7.90 (dd, J = 7.8, 5.4 Hz, 4H), 7.71 – 7.65 (m, 8H). ESI-MS; *m/z* (%): 1737 (20) $[M^+-PF_6]$, 799 (100) $[M^{2+}-2PF_6]$, 484 (100) $[M^{3+}-3PF_6]$, 327 (100) $[M^{4+}-4PF_6]$.

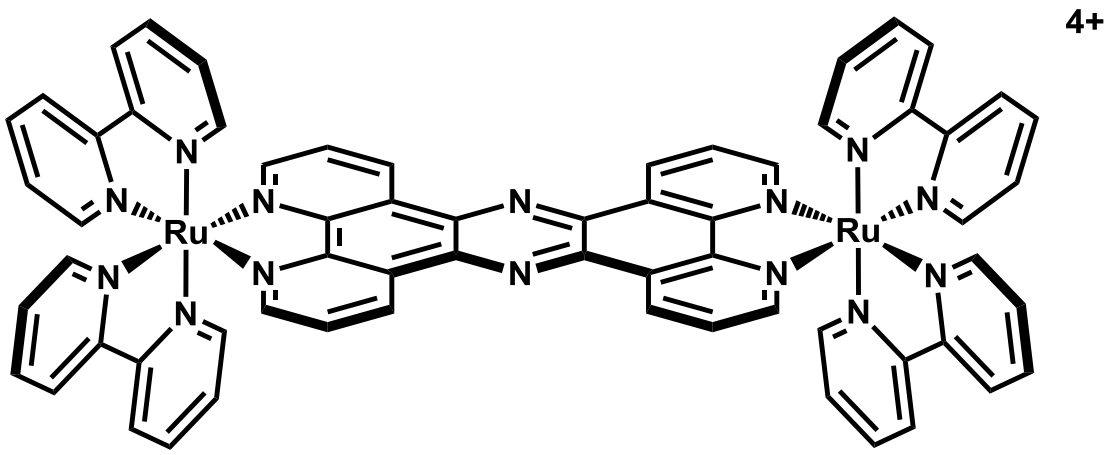
8.2.5 $[Ru(bipy)_2Cl_2].2H_2O$



$RuCl_3.3H_2O$ (3 g, 11.53 mmol), LiCl (3 g, 70.96 mmol) and 2,2-bipyridine (3.6 g, 22.8 mmol) were heated to reflux in DMF (30 mL) for 8 hours. After cooling to room temperature, acetone (200 mL) was added and the mixture was stored at 4°C overnight. The resulting suspension was then filtered and the precipitate was washed with three portions of diethyl ether (25 mL each) and dried under vacuum to give $[Ru(bipy)_2Cl_2].2H_2O$ as a black insoluble powder.[8]

Yield: 4 g, 71.44 %; MS m/z (%): 449 (100) $[M\text{-}Cl]^+$.

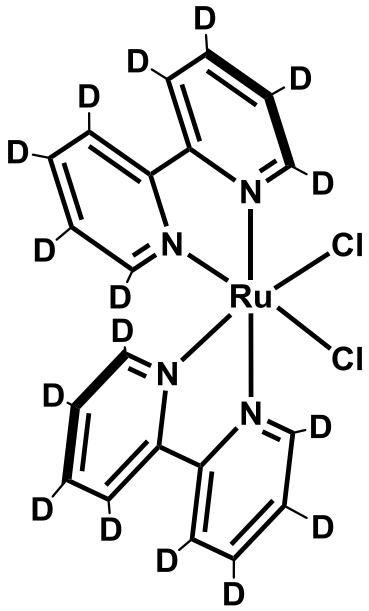
8.2.6 $[(Ru(bipy)_2)_2tpphz][Cl]_4$



$Ru(bipy)_2Cl_2.2H_2O$ (3 g, 2.22 mmol) and tpphz (0.38 g, 1 mmol) were suspended in a 1:1 solution of EtOH:H_2O (40 mL). The suspension was refluxed for 12 hours under inert gas and left to cool to room temperature followed by refrigeration overnight. The resulting brown solution was filtered and then concentrated *in vacuo*. Excess ammonium hexafluorophosphate (NH_4PF_6) was added to the solution and the precipitate was collected, washed with water (50 mL) and then dried. The PF_6 salt was then converted to the chloride salt via metathesis on a dowex exchange resin to produce $[(Ru(bipy)_2)_2tpphz][Cl]_4$ as an orange powder.[9]

Yield: 0.8 g, 59 %; ^{1}H NMR (800 MHz, D_2O) δ 10.03 (d, *J* = 7.8 Hz, 4H), 8.66 (d, *J* = 8.1 Hz, 4H), 8.63 (d, *J* = 8.1 Hz, 4H), 8.44 (d, *J* = 4.8 Hz, 4H), 8.18 (t, *J* = 7.8 Hz, 4H), 8.14 – 7.95 (m, 12H), 7.88 (d, *J* = 5.2 Hz, 4H), 7.51 (t, *J* = 6.3 Hz 4H), 7.30 (t, *J* = 6.3 Hz, 4H). MS; m/z (%): 303 (100) [M^{4+}-4Cl].

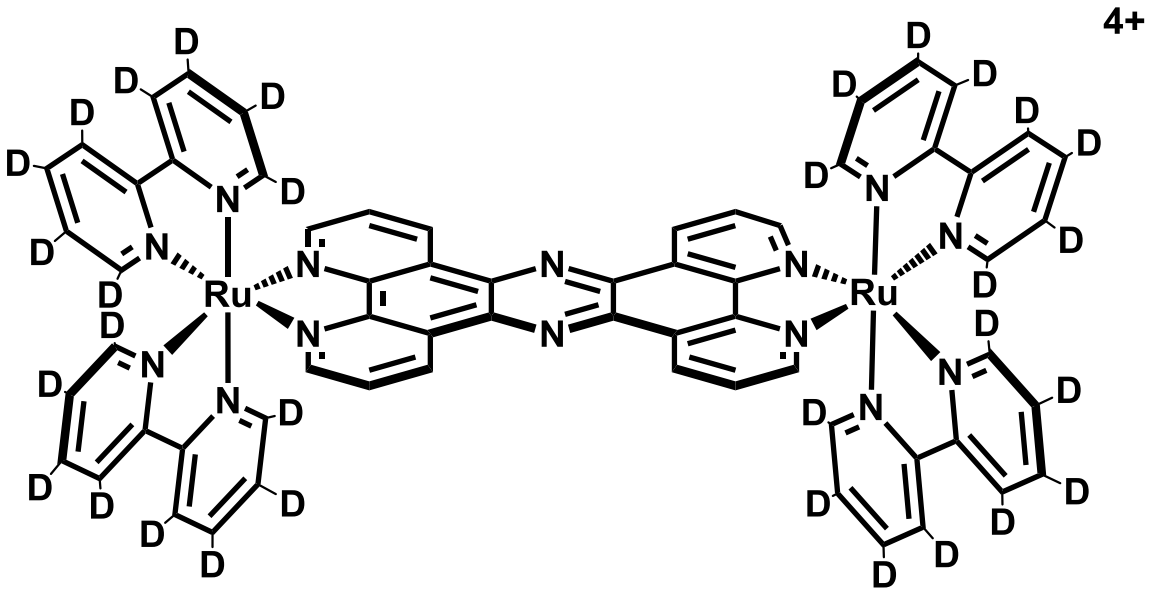
8.2.7 Deuterated-[Ru(d$_8$-bipy)$_2$Cl$_2$].2H$_2$O



$RuCl_3.3H_2O$ (0.8 g, 3.07 mmol), LiCl (0.9 g, 21.29 mmol) and 2,2'-bipyridine-d8 (1 g, 6.09 mmol) were heated to reflux in DMF (15 mL) for 8 hours. After cooling to room temperature acetone (200 mL) was added and the mixture stored at 4°C overnight. The resulting suspension was then filtered and the precipitate washed with three portions of diethyl ether (15 mL each) and dried under vacuum to give $[Ru(bipy)_2Cl_2].2H_2O$ as a black powder.

Yield: 0.968 g, 59.26 %; ESI-MS *m/z* (%): 523 (100) $[M+Na]^+$, 465 (30) $[M-Cl]^+$.

8.2.8 Deuterated-[(Ru(d_8-bipy)$_2$)$_2$tpphz][Cl]$_4$

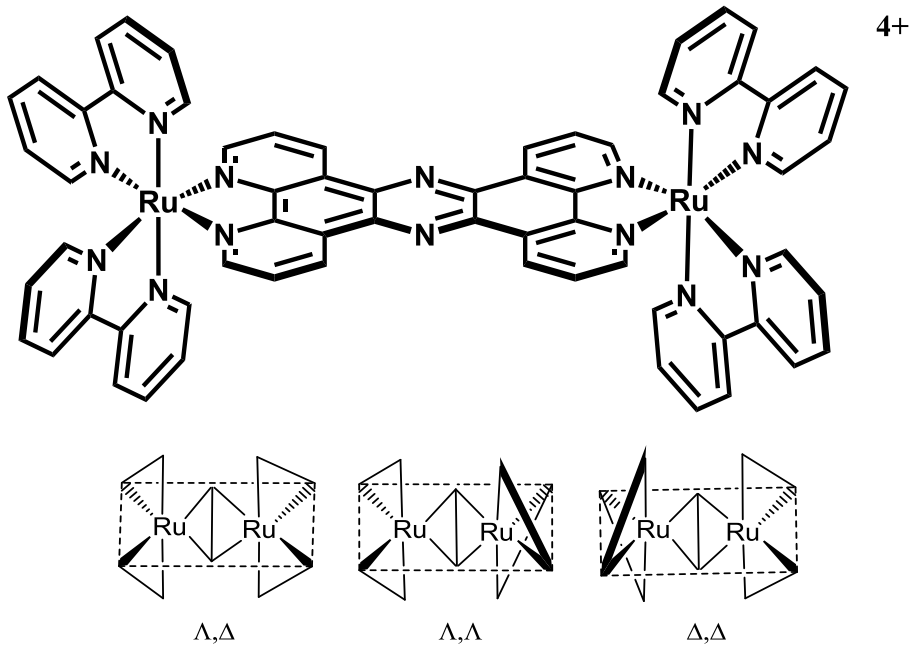


Ru(d_8-bipy)$_2$Cl$_2$.2H$_2$O (0.49 g, 0.9 mmol) and tpphz (0.1536 g, 0.4 mmol) were suspended in a 1:1 solution of EtOH:H$_2$O (40 mL). The reaction mixture was refluxed for 12 hours under argon. The suspension was left to cool to room temperature and stored at 4 °C overnight. The cold brown solution was filtered and concentrated under reduced pressure. After addition of excess NH_4PF_6, the resulting precipitate was collected, washed with water (50 mL) and then dried under vacuum. The PF_6 salt was then dissolved in acetone (20 mL) and precipitated with excess tetrabutylammonium chloride to give [(Ru(d_8-bipy)$_2$)$_2$tpphz][Cl]$_4$ as an orange powder.

Yield: 0.34 g, 61.36 %; ^{1}H NMR (800 MHz, D$_2$O) δ 10.03 (d, J = 8.1 Hz, 4H), 8.45 (d, J = 5.4 Hz, 4H), 8.05 (dd, J = 8.1, 5.4 Hz, 4H). MS; m/z (%): 310.8 (100) [M-4Cl]$^{4+}$.

8.2.9 Separation of Stereoisomers

8.2.9.1 Column Chromatography of $[(Ru(bipy)_2)_2tpphz][Cl]_4$

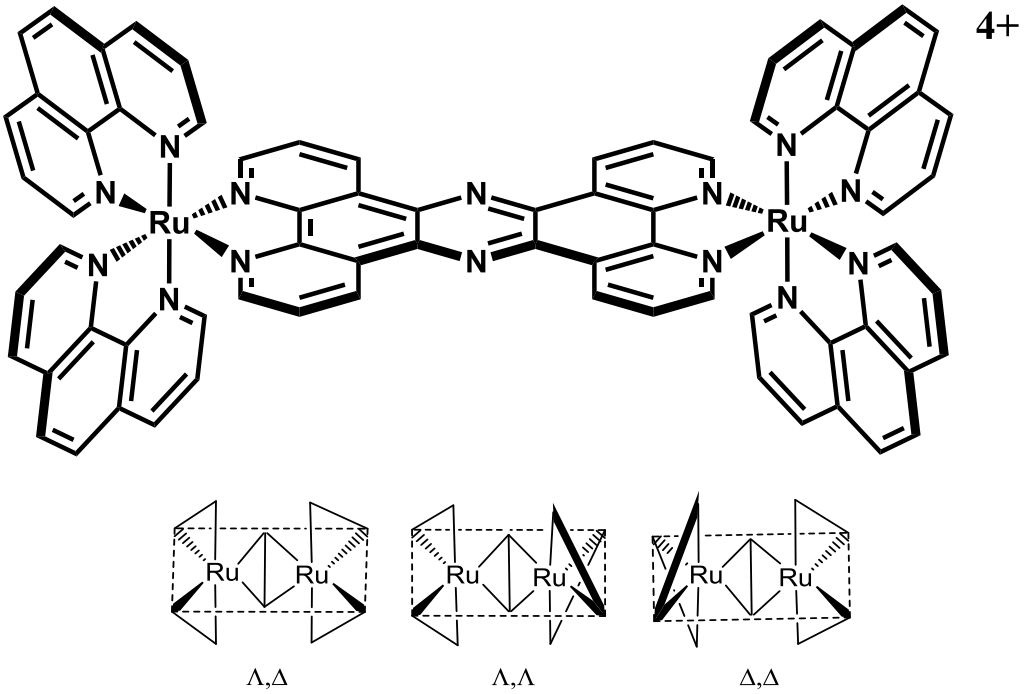


$[(Ru(bipy)_2)_2tpphz][Cl]_4$ was dissolved in water (50 mL) and loaded onto a column system containing SP sephadex C-25 with the flow controlled and circulated around the column by a peristaltic pump at a flow rate of 0.75 ml min^{-1}. The addition of an eluent solution of sodium octanoate (0.15 mM) at pH 10.5 caused a band to progress down the column and the system was set up to recycle. After several passes of the column there was clear separation into two red bands. The first (Λ,Δ) and second (Δ,Δ Λ,Λ) fractions were collected and extracted with the addition of NH_4PF_6 into dichloromethane (DCM) and dried using magnesium sulphate ($MgSO_4$). The organic layer was filtered and the solvent removed under reduced pressure to give the PF_6 salts. The salts were then converted to the chloride salts by dissolution in acetone followed by precipitation using tetrabutylammonium chloride.

The racemic band was then reloaded onto the column using the same method described above and eluted using a solution of sodium dibenzoyl-l-tartrate (0.05 M) at pH 10.5. One passage of the column was needed to see complete separation with the Δ,Δ eluted first and Λ,Λ eluted second. Fractions were collected and extracted, with the addition of NH_4PF_6, into DCM and dried using

$MgSO_4$). The organic layer was filtered and the solvent removed under reduced pressure to give the PF_6 salts. The salts were then converted to the chloride salts by dissolution in acetone followed by precipitation using tetrabutylammonium chloride.

8.2.9.2 Column Chromatography of $[(Ru(phen)_2)_2tpphz][Cl]_4$

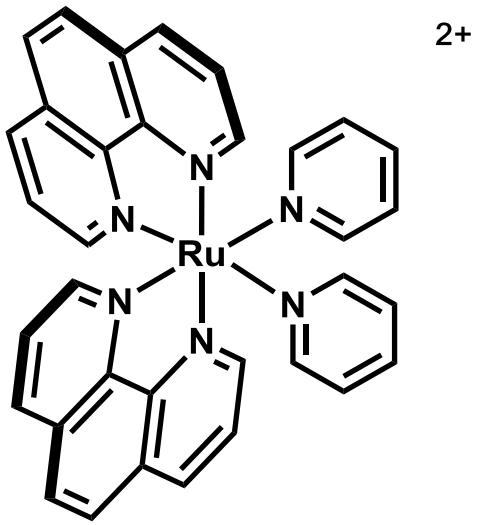


$[(Ru(phen)_2)_2tpphz][Cl]_4$ was dissolved in water (50 mL) and loaded onto a column system containing SP sephadex C-25 with the flow controlled and circulated around the column by a peristaltic pump at a flow rate of 0.75 ml min^{-1}. The addition of an eluent solution of sodium dibenzoyl-L-tartrate (0.05 mM) at pH 10.5 caused a band to progress down the column and the system was set up to recycle. After several passes of the column there was clear separation into three red bands. These fractions were collected and extracted, with the addition of NH_4PF_6, into dichloromethane (DCM) and dried using magnesium sulphate ($MgSO_4$). The organic layer was filtered and the solvent removed under reduced pressure to give the PF_6 salts. The salts were then converted to the chloride salts by dissolution in acetone followed by precipitation using tetrabutylammonium chloride

8.2.10 Attempted Resolution by Crystallisation

As discussed in chapter 2, the experimental for the attempted resolution is detailed below.

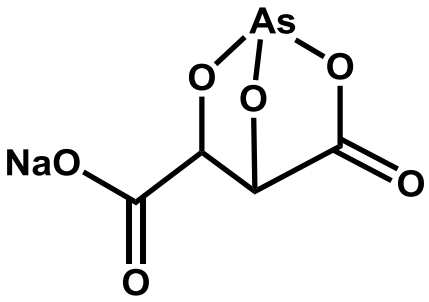
8.2.10.1 Synthesis of $[Ru(phen)_2(py)_2][Cl]_2$[5]



$[Ru(phen)_2Cl_2]$ (0.5 g) was refluxed for 3 hours in a solution of pyridine (5 mL) and water (10 mL). The filtered solution was removed under reduced pressure, the brown residue dissolved in methanol (10 mL) and precipitated out with addition of diethyl ether. The solid was dissolved in a solution of water:methanol (50:1) and precipitated using ammonium hexafluorophosphate. Crystal growth took place as the methanol was evaporated slowly. The crystals were then collected, dried under vacuum and dissolved in acetone (90 mL). Tetrabutylammonium chloride was added in excess, precipitating the product as a chloride salt. The solid was filtered and washed with acetone to give Cis-$[(phen)_2Ru(py)_2][Cl_2]$ as an orange powder.

Yield: 0.3 g, 49 %; ^{1}H NMR (400 MHz, CD3CN) δ 9.40 (dd, J = 5.2, 1.2 Hz, 2H), 8.77 (dd, J = 8.3, 1.2 Hz, 2H), 8.51 – 8.37 (m, 6H), 8.27 – 8.06 (m, 6H), 8.03 (dd, J = 5.2, 1.2 Hz, 2H), 7.81 (t, J = 8.0 Hz, 2H), 7.53 (dd, J = 8.2, 5.3 Hz, 2H), 7.27 (t, J = 8.0 Hz, 4H). MS m/z (%): 655 (10) $[M^+-Cl]$, 310 (100) $[M^{2+}-2Cl]$.

8.2.10.2 Sodium Arsenyl-D(-)-Tartrate[10]



NaOH (2.62 g, 0.067 mol) and D-(-)-tartaric acid (10 g, 0.067 mol) were dissolved in water (75 mL), and heated under reflux. As_2O_3 (6.55 g, 0.033 mol) was added to the solution and refluxed for 45 minutes. As the reaction progressed the solution turned clear. A hot filtration was carried out, and ethanol (150 mL) was added to the filtrate. The mixture was cooled to 4 °C for 14 hours. The white crystals were collected by filtration, washed with cold ethanol, and left to dry.

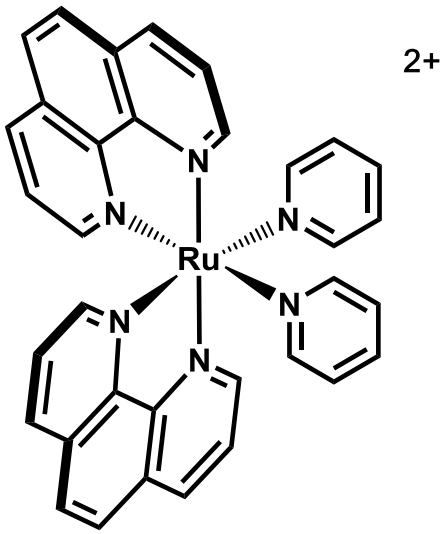
Yield: 10 g, 61.9%; ^{1}H NMR (400 MHz, D_2O) δ 4.65 (s, 1H) 4.45 (s, 1H). MS m/z (%): 221 (10) [M^-- Na^+].

8.2.10.3 Sodium Arsenyl-L(-)-Tartrate[10]

The above synthesis can be employed using L-(-)-tartaric acid as the starting material to yield the sodium arsenyl-L(-)-tartrate. The crystals were collected by filtration, washed with cold ethanol, and left to dry.

Yield: 25g, 77%; ^{1}H NMR (D_2O) δ 4.65 (s, 1H) 4.45 (s, 1H). MS m/z (%): 221 (10) [M^-- Na^+].

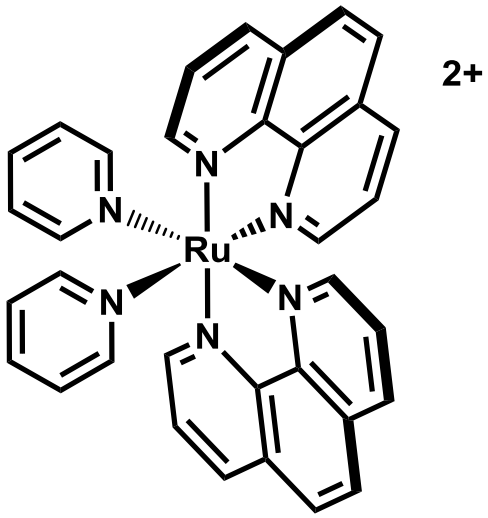
8.2.10.4 Λ-[(phen)$_2$Ru(py)$_2$][Cl]$_2$[11]



Cis-[Ru(phen)$_2$(py)$_2$]Cl$_2$ (0.5 g) was added to water (20 mL) and the solution heated to boiling. Sodium arsenyl-(+)-tartrate (1.0 g) in hot water (5 mL) was added to the heating solution. Crystallisation was induced through scratching using a glass rod and proceeded overnight with slow cooling to 4 °C. The orange crystals were collected and washed with ice cold water. For purification the crystals were dissolved in hot DMSO (5 mL) and filtered. Boiling water (15 mL) was added to the filtrate followed by a hot aqueous solution of sodium arsenyl-(+)-tartrate (0.2 g, 2 mL). Crystallisation was again initiated by scratching with a glass rod and allowed to proceed overnight with slow cooling to 4°C. The orange crystals were collected and washed with ice cold water. The crystals were then dissolved in hot 20% acetic acid solution (20 mL) and precipitated with excess NH_4PF_6. The precipitate was collected and dried and then recrystallized from acetone/water. The resulting precipitate was filtered and dried under vacuum. The PF_6 salt was then converted to the Cl_2 salt by addition of tetrabutylammonium chloride in minimal acetone solution to yield the Λ-Cis-[(phen)$_2$Ru(py)$_2$].2[Cl] as an orange powder.

Yield: 0.13 g, 26%; ^{1}H NMR (400 MHz, CD3CN) δ 9.40 (dd, J = 5.2, 1.2 Hz, 2H), 8.77 (dd, J = 8.3, 1.2 Hz, 2H), 8.51 – 8.37 (m, 6H), 8.27 – 8.06 (m, 6H), 8.03 (dd, J = 5.2, 1.2 Hz, 2H), 7.81 (t, J = 8.0 Hz, 2H), 7.53 (dd, J = 8.2, 5.3 Hz, 2H), 7.27 (t, J = 8.0 Hz, 4H). ESI-MS *m/z* (%): 655 (15) [M$^+$-Cl], 310 (100)[M^{2+}-2Cl].

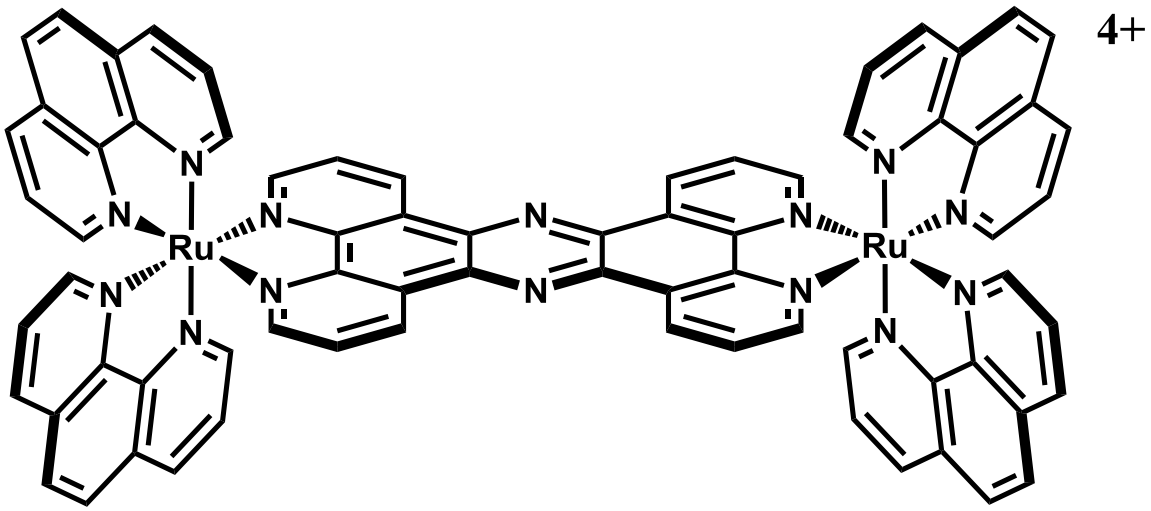
8.2.10.5 Resolution of Δ-$[phen)_2Ru(py)_2][Cl]_2$[11]



A solution of cis-$[Ru(phen)_2(py)_2]Cl_2$ (0.5 g) in water (20 mL) was heated to boiling. To this solution sodium arsenyl-D(-)-tartrate (1.0 g) in hot water (5 mL) was added. Crystallisation was induced through scratching using a glass rod and proceeded overnight with slow cooling to 4°C. The orange crystals were collected and washed with ice cold water. For purification the crystals were dissolved in hot DMSO (5 mL) and filtered. Boiling water (15 mL) was added to the filtrate prior to addition of a hot aqueous solution of sodium arsenyl-(-)-tartrate (0.2 g, 2 mL). Crystallisation was again induced by scratching with a glass rod and proceeded overnight with slow cooling to 4°C. The orange crystals were collected and washed with ice cold water. The crystals were then dissolved in hot 20% acetic acid solution (20 mL) and precipitated with excess of NH_4PF_6 which were collected and dried. The precipitate was then recrystallised from acetone/water and filtered and dried under vacuum. The PF_6 salt was dissolved in minimal acetone solution and converted to the Cl_2 salt by addition of tetrabutylammonium chloride to give Δ-Cis-$[(phen)_2Ru(py)_2].2[Cl]$ as an orange powder.

Yield: 0.1 g, 20%. ^{1}H NMR (400 MHz, CD3CN) δ 9.40 (dd, J = 5.2, 1.2 Hz, 2H), 8.77 (dd, J = 8.3, 1.2 Hz, 2H), 8.51 – 8.37 (m, 6H), 8.27 – 8.06 (m, 6H), 8.03 (dd, J = 5.2, 1.2 Hz, 2H), 7.81 (t, J = 8.0 Hz, 2H), 7.53 (dd, J = 8.2, 5.3 Hz, 2H), 7.27 (t, J = 8.0 Hz, 4H). MS m/z (%): 655 (8) $[M^+-Cl]$, 310 (100)$[M^{2+}-2Cl]$.

8.2.10.6 Δ,Δ – $[(phen)_2Ru(tpphz)Ru(phen)_2][(PF_6)_4]$



Δ-$[Ru(phen)_2(py)_2]Cl_2$ (0.0677 g) and tpphz (0.0277 g) were added to a 1:1 solution of EtOH:H_2O (10mL). The suspension was refluxed for 4 days in reduced light conditions under argon. After cooling to room temperature, the suspension was kept at 4 °C overnight. The resulting brown solution was filtered and concentrated under reduced pressure. Addition of NH_4PF_6 produced an orange precipitate which was collected and washed with water. However, CD measurements revealed that the product was not fully resolved.

8.3. Techniques

8.3.1 DNA Sample Preparation

Calf thymus DNA was purchased from Sigma-Aldrich as a freeze dried sodium salt. The DNA was prepared by the method described by JB Chaires, *et al.* with dissolution in 5 mM tris buffer solution (pH 7.4) followed by sonication for 30 minutes in an ice bath using a soniprep 150 ultrasonic disintegrator, giving DNA of average length 100-200 bases. [12,13] The quality of the DNA was then assessed by UV spectroscopy, as DNA absorbs at 260 nm, and single stranded RNA and protein contaminants absorb at 280 nm, an A_{260}/A_{280} being greater than 1.8 indicating a protein free sample. The concentrations per base pair of the final solutions was then determined by UV/Vis spectroscopy with $\varepsilon_{260} = 13200\ M^{-1}\ cm^{-1}$ for CT DNA and $\varepsilon_{258} = 13200\ M^{-1}\ cm^{-1}$ for poly d(AT) DNA.

8.3.2 Circular Dichroism (CD)

Circular dichroism measurements were recorded using 15 μM solutions of [1] or [2] in H_2O on a Jasco J-810 spectropolarimeter at 25 °C, scanning at 100 nm/min from 500-200 nm. The final CD spectra were the average of three scans. CD experiments using the oligonucleotide sequence d(GCATATCG).(CGATATGC) were carried out in solutions containing 2 mM DNA with 150 mM NaCl in 5 mM tris buffer at pH 7.4 and 0.5 mM Λ,Δ-[1] was added to these solutions. Threading experiments using calf thymus and poly d(AT) DNA were carried out in solutions containing 2 mM DNA per base pair with 25 mM NaCl in 5 mM tris buffer at pH 7.4 and 0.5 mM Λ,Δ-[1] was added to the solutions. After initial CD spectra were performed, the samples were heated to 60 °C in a water bath and left for 24 hours. After cooling, subsequent CD spectra were recorded.

Threading experiments were carried out in the sample conditions employed by Lincoln, *et al.* which consisted of solutions containing 60 μM per base pair poly d(AT) DNA, with 150 mM NaCl, 5 mM tris buffer at pH 7.4, and 11.25 μM metal complex.[14]

8.3.3 NMR

All studies using oligonucleotide were carried by dissolution of d(GCATATCG) and d(CGATATGC) in 250 µL of D_2O. Additions of one strand into the other were monitored by NMR until a stable duplex was formed at a 1:1 ratio. The concentration of the samples and the chemical shifts were measured against the 5 mM TSP which was used as a reference in the solution. 2 mM solutions of DNA were required for stability with 25 mM NaCl in D_2O. Concentrated stock solutions (20 mM) of ruthenium complexes were added to the DNA in D_2O up to a half ratio (1 mM).

^{1}H-NMR experiments were carried out on a Bruker Avance 800 MHz spectrometer. 1D experiments (zgpr) were carried out with 16 scans recorded over 25556 points with a spectral width of 12.4946 ppm and a relaxation delay of 4 s. Phase sensitive double quantum filtered correlation spectroscopy (COSY) (cosydfphpr), total correlated spectroscopy (TOCSY)(dipsi2phpr) and Nuclear Overhauser effect spectroscopy (NOESY)(NOESYphpr) experiments were carried out with 16 scans collecting 4096 points in f2 and 800 points in f1, a spectral width of 12.0141 ppm in both dimensions and a relaxation delay of 1.5 s. NOESY mixing times were typically 100 ms and TOCSY mixing times were 60 ms. Water suppression was carried out using a pre-saturation.

^{31}P 1D NMR and ^{1}H-^{31}P correlated spectroscopy ^{31}P-^{1}H COSY experiments,[15,16] were carried out on a BrukerAvance 500 MHz spectrometer acquired with a triple resonance broadband probe. ^{31}P 1D NMR experiments (zgig) were recorded with proton decoupling. 512 scans were acquired over 2048 points with a spectral width of 9.9906 ppm and a relaxation delay of 0.5 s. ^{31}P-^{1}H COSY experiments were carried out with 400 scans collecting 2048 points in F2 and 200 in F1 with a spectral width of 9.9974 ppm in F2 and 4.0000 ppm in F1 using a relaxation delay of 1 s.

2D spectra were typically processed by applying a shifted sine bell window function and baseline corrected by a polynomial in both dimensions.

8.4 References

1 Bruker, *SADABS*, Bruker Axs Inc., Madison, wisconsin, USA, 2016.

2 L. Krause, R. Herbst-Irmer, G. M. Sheldrick and D. Stalke, *J. Appl. Crystallogr.*, 2015, **48**, 3–10.

3 G. M. Sheldrick, *Acta Crystallogr. Sect. A Found. Adv.*, 2015, **71**, 3–8.

4 G. M. Sheldrick, *Acta Crystallogr. Sect. C Struct. Chem.*, 2015, **71**, 3–8.

5 O. V. Dolomanov, L. J. Bourhis, R. J. Gildea, J. A. K. Howard and H. Puschmann, *J. Appl. Crystallogr.*, 2009, **42**, 339–341.

6 R. H. Zheng, H. C. Guo, H. J. Jiang, K. H. Xu, B. B. Liu, W. L. Sun and Z. Q. Shen, *Chinese Chem. Lett.*, 2010, **21**, 1270–1272.

7 J. Bolger, A. Gourdon, E. Ishow and J. P. Launay, *J. Chem. Soc., Chem. Commun.*, 1995, **0**, 1799–1800.

8 B. P. Sullivan, D. J. Salmon and T. J. Meyer, *Inorg. Chem.*, 1978, **17**, 3334–3341.

9 J. Bolger, A. Gourdon, E. Ishow and J.-P. Launay, *Inorg. Chem.*, 1996, **35**, 2937–2944.

10 M.-Y. Tong, T. Payagala, S. Perera, F. M. MacDonnell and D. W. Armstrong, *J. Chromatogr. A*, 2010, **1217**, 1139–1148.

11 B. Bosnich and F. Dwyer, *Aust. J. Chem.*, 1966, **19**, 2229.

12 J. B. Chaires, N. Dattagupta and D. M. Crothers, *Biochemistry*, 1982, **21**, 39333940.

13 M. Hogan, N. Dattagupta and D. M. Crothers, *Proc. Natl. Acad. Sci.*, 1978, **75**, 195–199.

14 J. Andersson and P. Lincoln, *J. Phys. Chem. B*, 2011, **115**, 14768–14775.

15 V. Sklenár , H. Miyashiro, G. Zon, H. Todd Miles and A. Bax, *FEBS Lett.*, 1986, **208**, 94–98.

16 P. Waywell, J. A. Thomas and M. P. Williamson, *Org. Biomol. Chem.*, 2010, **8**, 648–654.

Chapter 9. Appendix

9.1 X-ray Crystallographic Data

The general crystal data, bond lengths and angles obtained from X-ray crystal structures are reported in this section.

9.1.1 Λ,Δ-[1]

Table 9.1. A summary of crystallographic data and structure refinement of Λ,Δ-[1]

Identification code	iaj646k_0m
Empirical formula	$C_{64}H_{44}Cl_4N_{14}Ru_2$
Formula weight	1353.07
Temperature / K	100
Crystal system	Triclinic
Space group	P-1
a / Å	8.3032(16)
b / Å	11.378(2)
c / Å	18.250(4)
α / °	95.542(7)
β / °	92.729(7)
γ / °	95.095(7)
Volume / Å^3	1706.6(6)
Z	1
ρ_{calc} / gcm^{-3}	1.317
μ / mm^{-1}	0.646
F(000)	682.0
Crystal size / mm^3	0.4 × 0.212 × 0.188
Radiation	MoKα (λ = 0.71073)
2θ Range for Data Collection / °	3.612 to 55.112
Index Ranges	-10 ≤ h ≤ 10, -14 ≤ k ≤ 14, -23 ≤ l ≤ 23
Reflections Collected	46946
Independent Reflections	7843 [R_{int} = 0.0683, R_{sigma} = 0.0575]
Data / Restraints / Parameters	7843/381/389
Goodness-of-fit on F^2	1.022
Final R indexes [I>=2σ (I)]	R_1 = 0.0492, wR_2 = 0.1183
Final R indexes [all data]	R_1 = 0.0707, wR_2 = 0.1296
Largest diff. peak/hole / e Å^{-3}	1.84/-0.68

Table 9.2. Bond lengths (Å) for Λ,Δ-[1]

Atom 1	Atom 2	Length / Å	Atom 1	Atom 2	Length / Å
Ru1	N1	2.061(3)	C6	C7	1.393(5)
Ru1	N2	2.068(3)	C7	C8	1.403(6)
Ru1	N4	2.045(3)	C7	C11	1.467(5)
Ru1	N5	2.062(3)	C8	C9	1.376(6)
Ru1	N6	2.062(3)	C9	C10	1.398(6)
Ru1	N7	2.056(3)	C11	C12^{1}	1.402(5)
N1	C1	1.334(5)	C12	C4^{1}	1.459(5)
N1	C5	1.370(4)	C12	C11^{1}	1.402(5)
N2	C6	1.364(4)	C13	C14	1.372(7)
N2	C10	1.338(5)	C14	C15	1.368(9)
N3	C11	1.343(5)	C15	C16	1.385(8)
N3	C12	1.335(4)	C16	C17	1.386(6)
N4	C13	1.355(5)	C17	C18	1.480(5)
N4	C17	1.350(6)	C18	C19	1.387(5)
N5	C18	1.354(5)	C19	C20	1.384(6)
N5	C22	1.347(4)	C20	C21	1.383(5)
N6	C23	1.341(5)	C21	C22	1.380(5)
N6	C27	1.370(5)	C23	C24	1.401(6)
N7	C28	1.356(5)	C24	C25	1.365(6)
N7	C32	1.348(4)	C25	C26	1.387(6)
C1	C2	1.398(5)	C26	C27	1.381(5)
C2	C3	1.386(5)	C27	C28	1.482(5)
C3	C4	1.404(5)	C28	C29	1.393(5)
C4	C5	1.394(5)	C29	C30	1.386(6)
C4	C12^{1}	1.459(5)	C30	C31	1.381(6)
C5	C6	1.440(5)	C31	C32	1.390(5)

Table 9.3. Bond Angles (°) for Λ,Δ-[1] (^{1}-X, 1-Y, 1-Z)

Atom 1	Atom 2	Atom 3	Angle / °	Atom 1	Atom 2	Atom 3	Angle / °
N1	Ru1	N2	79.50(11)	N2	C6	C5	116.3(3)
N1	Ru1	N5	96.55(11)	N2	C6	C7	122.9(3)
N1	Ru1	N6	97.22(12)	C7	C6	C5	120.8(3)
N4	Ru1	N1	86.42(12)	C6	C7	C8	118.3(3)
N4	Ru1	N2	97.41(14)	C6	C7	C11	118.4(3)
N4	Ru1	N5	78.96(12)	C8	C7	C11	123.3(4)
N4	Ru1	N6	176.14(12)	C9	C8	C7	118.7(4)
N4	Ru1	N7	97.88(12)	C8	C9	C10	119.9(4)
N5	Ru1	N2	174.85(12)	N2	C10	C9	122.3(4)
N5	Ru1	N6	99.28(11)	N3	C11	C7	117.7(3)
N6	Ru1	N2	84.57(13)	N3	C11	C12^{1}	121.9(3)
N7	Ru1	N1	174.47(12)	C12^{1}	C11	C7	120.5(3)
N7	Ru1	N2	96.44(11)	N3	C12	C4^{1}	117.9(3)
N7	Ru1	N5	87.72(11)	N3	C12	C11^{1}	121.8(3)
N7	Ru1	N6	78.56(12)	C11^{1}	C12	C4^{1}	120.3(3)
C1	N1	Ru1	128.0(2)	N4	C13	C14	122.7(5)
C1	N1	C5	117.9(3)	C15	C14	C13	118.5(5)
C5	N1	Ru1	114.1(2)	C14	C15	C16	120.3(5)
C6	N2	Ru1	113.8(2)	C17	C16	C15	118.4(6)
C10	N2	Ru1	128.2(3)	N4	C17	C16	121.8(4)
C10	N2	C6	118.0(3)	N4	C17	C18	114.6(3)
C12	N3	C11	116.4(3)	C16	C17	C18	123.6(4)
C13	N4	Ru1	125.7(3)	N5	C18	C17	114.8(3)
C17	N4	Ru1	116.1(2)	N5	C18	C19	121.6(3)
C17	N4	C13	118.1(4)	C19	C18	C17	123.5(4)
C18	N5	Ru1	115.3(2)	C20	C19	C18	119.1(4)
C22	N5	Ru1	126.1(2)	C21	C20	C19	119.3(4)
C22	N5	C18	118.6(3)	C22	C21	C20	118.8(3)
C23	N6	Ru1	125.4(3)	N5	C22	C21	122.4(3)
C23	N6	C27	118.4(3)	N6	C23	C24	121.6(4)
C27	N6	Ru1	115.5(2)	C25	C24	C23	119.5(4)
C28	N7	Ru1	116.2(2)	C24	C25	C26	119.5(4)
C32	N7	Ru1	125.3(2)	C27	C26	C25	119.0(4)
C32	N7	C28	118.4(3)	N6	C27	C26	121.9(3)
N1	C1	C2	122.4(3)	N6	C27	C28	113.9(3)
C3	C2	C1	120.2(4)	C26	C27	C28	124.1(3)
C2	C3	C4	118.0(3)	N7	C28	C27	114.7(3)
C3	C4	C12^{1}	123.0(3)	N7	C28	C29	121.8(3)
C5	C4	C3	118.7(3)	C29	C28	C27	123.4(3)
C5	C4	C12^{1}	118.3(3)	C30	C29	C28	119.3(4)
N1	C5	C4	122.7(3)	C31	C30	C29	118.7(3)
N1	C5	C6	115.8(3)	C30	C31	C32	119.6(4)
C4	C5	C6	121.5(3)	N7	C32	C31	122.1(4)

9.1.2 Δ,Δ-[1]

Table 9.4. A summary of crystallographic data and structure refinement of Δ,Δ-[1]

Identification code	IAJ672k_0m
Empirical formula	$C_{32}H_{22}F_{12}N_7P_2Ru$
Formula weight	895.57
Temperature / K	100
Crystal system	Monoclinic
Space group	$P2_1$
a / Å	12.5327(14)
b / Å	16.3645(19)
c / Å	21.809(3)
α / °	90
β / °	97.129(5)
γ / °	90
Volume / Å^3	4438.3(9)
Z	4
ρ_{calc} / gcm^{-3}	1.340
μ / mm^{-1}	0.504
F(000)	1780.0
Crystal size / mm^3	0.51 × 0.17 × 0.104
Radiation	MoKα (λ = 0.71073)
2θ Range for Data Collection / °	3.12 to 55.158
Index Ranges	$-16 \leq h \leq 14$, $-21 \leq k \leq 21$, $-24 \leq l \leq 28$
Reflections Collected	77631
Independent Reflections	20302 [R_{int} = 0.1023, R_{sigma} = 0.1197]
Data / Restraints / Parameters	20302/1151/973
Goodness-of-fit on F^2	0.984
Final R indexes [I>=2σ (I)]	R_1 = 0.0758, wR_2 = 0.1894
Final R indexes [all data]	R_1 = 0.1158, wR_2 = 0.2114
Largest diff. peak/hole / e Å^{-3}	1.16/-0.88
Flack Parameter	0.086(17)

Table 9.5. Bond lengths (Å) for Δ,Δ-[1]

Atom 1	Atom 2	Length / Å
Ru1	N1	2.072(9)
Ru1	N2	2.065(9)
Ru1	N7	2.041(10)
Ru1	N8	2.065(9)
Ru1	N9	2.041(9)
Ru1	N10	2.061(10)
Ru2	N5	2.080(7)
Ru2	N6	2.075(9)
Ru2	N11	2.114(12)
Ru2	N12	2.037(9)
Ru2	N13	2.061(8)
Ru2	N14	2.108(10)
N1	C1	1.347(13)
N1	C5	1.348(14)
N2	C6	1.402(13)
N2	C10	1.317(14)
N3	C11	1.332(13)
N3	C12	1.341(13)
N4	C23	1.367(12)
N4	C24	1.353(13)
N5	C18	1.362(12)
N5	C22	1.272(13)
N6	C16	1.346(14)
N6	C17	1.331(13)
N7	C25	1.330(16)
N7	C29	1.460(16)
N8	C30	1.324(17)
N8	C34	1.347(17)
N9	C35	1.362(15)
N9	C39	1.338(15)
N10	C40	1.335(14)
N10	C44	1.307(15)
N11	C45	1.244(17)
N11	C49	1.370(18)
N12	C50	1.360(15)
N12	C54	1.382(14)
N13	C55	1.384(15)
N13	C59	1.400(15)
N14	C60	1.333(15)
N14	C64	1.239(15)
C1	C2	1.411(15)
C20	C21	1.378(14)
C21	C22	1.409(15)
C25	C26	1.408(17)
C26	C27	1.40(2)
C27	C28	1.43(2)
C28	C29	1.417(19)
C29	C30	1.399(18)
C30	C31	1.397(18)
C31	C32	1.45(2)
C32	C33	1.32(2)
C33	C34	1.380(16)
C35	C36	1.368(15)
C36	C37	1.459(16)
C37	C38	1.334(16)
C38	C39	1.416(15)
C39	C40	1.462(16)
C40	C41	1.391(18)
C41	C42	1.384(19)
C42	C43	1.365(19)
C43	C44	1.448(18)
C45	C46	1.40(2)
C46	C47	1.37(3)
C47	C48	1.40(3)
C48	C49	1.49(2)
C49	C50	1.40(2)
C50	C51	1.445(18)
C51	C52	1.34(2)
C52	C53	1.25(2)
C53	C54	1.391(17)
C55	C56	1.402(16)
C56	C57	1.34(2)
C57	C58	1.375(19)
C58	C59	1.390(17)
C59	C60	1.400(17)
C60	C61	1.454(18)
C61	C62	1.329(19)
C62	C63	1.37(2)
C63	C64	1.408(18)
P1	F1	1.623(11)
P1	F2	1.612(9)
P1	F3	1.593(10)

C2	C3	1.393(16)
C3	C4	1.389(15)
C4	C5	1.380(14)
C4	C11	1.449(15)
C5	C6	1.452(16)
C6	C7	1.375(15)
C7	C8	1.361(16)
C7	C24	1.432(15)
C8	C9	1.412(15)
C9	C10	1.374(15)
C11	C24	1.444(15)
C12	C13	1.500(12)
C12	C23	1.425(14)
C13	C14	1.338(13)
C13	C17	1.362(14)
C14	C15	1.352(13)
C15	C16	1.420(15)
C17	C18	1.471(14)
C18	C19	1.419(12)
C19	C20	1.367(14)
C19	C23	1.441(13)

P1	F4	1.627(10)
P1	F5	1.641(11)
P1	F6	1.640(9)
P2	F7	1.591(11)
P2	F8	1.622(10)
P2	F9	1.580(10)
P2	F10	1.618(9)
P2	F11	1.592(12)
P2	F12	1.593(8)
P3	F13	1.536(10)
P3	F14	1.634(11)
P3	F15	1.607(10)
P3	F16	1.591(8)
P3	F17	1.572(8)
P3	F18	1.574(9)
P4	F19	1.573(13)
P4	F20	1.573(15)
P4	F21	1.563(15)
P4	F22	1.626(14)
P4	F23	1.589(15)
P4	F24	1.586(15)

Table 9.6. Bond Angles (°) for Δ,Δ-[1]

Atom 1	Atom 2	Atom 3	Angle / °
N2	Ru1	N1	79.3(3)
N7	Ru1	N1	99.7(4)
N7	Ru1	N2	82.5(4)
N7	Ru1	N8	77.9(5)
N7	Ru1	N9	97.3(4)
N7	Ru1	N10	173.6(4)
N8	Ru1	N1	173.5(3)
N8	Ru1	N2	94.3(3)
N9	Ru1	N1	97.9(3)
N9	Ru1	N2	177.1(4)
N9	Ru1	N8	88.4(3)
N9	Ru1	N10	78.4(4)
N10	Ru1	N1	85.7(4)
N10	Ru1	N2	102.0(4)
N10	Ru1	N8	97.1(4)
N5	Ru2	N11	92.7(4)
N5	Ru2	N14	92.3(4)
N6	Ru2	N5	79.2(3)
N6	Ru2	N11	88.9(5)
N6	Ru2	N14	98.1(4)
N12	Ru2	N5	171.8(4)
N12	Ru2	N6	96.7(3)
N12	Ru2	N11	80.1(4)
N12	Ru2	N13	87.4(3)
N12	Ru2	N14	95.3(4)
N13	Ru2	N5	97.0(3)
N13	Ru2	N6	175.6(4)
N13	Ru2	N11	93.5(5)
N13	Ru2	N14	79.8(4)
N14	Ru2	N11	172.1(4)
C1	N1	Ru1	126.1(7)
C1	N1	C5	118.4(9)
C5	N1	Ru1	115.3(7)
C6	N2	Ru1	114.0(7)
C10	N2	Ru1	130.1(7)
C10	N2	C6	115.5(9)
C11	N3	C12	116.7(9)
C24	N4	C23	116.3(9)
C18	N5	Ru2	113.0(6)
C22	N5	Ru2	128.5(7)
C22	N5	C18	118.4(8)

Atom 1	Atom 2	Atom 3	Angle / °
N7	C25	C26	122.7(12)
C27	C26	C25	120.0(14)
C26	C27	C28	119.2(13)
C29	C28	C27	119.5(14)
C28	C29	N7	118.6(12)
C30	C29	N7	113.6(11)
C30	C29	C28	127.4(13)
N8	C30	C29	116.5(11)
N8	C30	C31	124.6(13)
C31	C30	C29	118.8(13)
C30	C31	C32	112.9(14)
C33	C32	C31	123.7(13)
C32	C33	C34	116.9(16)
N8	C34	C33	123.7(15)
N9	C35	C36	124.8(12)
C35	C36	C37	114.8(10)
C38	C37	C36	120.9(10)
C37	C38	C39	119.6(11)
N9	C39	C38	121.0(11)
N9	C39	C40	116.4(9)
C38	C39	C40	122.1(11)
N10	C40	C39	113.5(10)
N10	C40	C41	120.4(11)
C41	C40	C39	126.0(11)
C42	C41	C40	120.4(13)
C43	C42	C41	119.4(12)
C42	C43	C44	116.9(12)
N10	C44	C43	122.3(11)
N11	C45	C46	122.7(14)
C47	C46	C45	121.4(17)
C46	C47	C48	115.6(18)
C47	C48	C49	121.2(17)
N11	C49	C48	114.7(14)
N11	C49	C50	120.3(13)
C50	C49	C48	123.7(14)
N12	C50	C49	115.1(11)
N12	C50	C51	118.1(13)
C49	C50	C51	126.7(13)
C52	C51	C50	119.8(14)
C53	C52	C51	123.1(13)
C52	C53	C54	119.3(13)

C16	N6	Ru2	127.0(7)
C17	N6	Ru2	115.6(7)
C17	N6	C16	117.2(9)
C25	N7	Ru1	126.4(8)
C25	N7	C29	119.3(10)
C29	N7	Ru1	114.1(8)
C30	N8	Ru1	117.3(9)
C30	N8	C34	118.0(11)
C34	N8	Ru1	124.7(9)
C35	N9	Ru1	126.3(8)
C39	N9	Ru1	115.1(7)
C39	N9	C35	118.6(10)
C40	N10	Ru1	116.2(8)
C44	N10	Ru1	123.3(8)
C44	N10	C40	120.5(10)
C45	N11	Ru2	127.2(10)
C45	N11	C49	123.4(13)
C49	N11	Ru2	109.4(10)
C50	N12	Ru2	115.0(8)
C50	N12	C54	116.9(10)
C54	N12	Ru2	128.1(8)
C55	N13	Ru2	125.5(8)
C55	N13	C59	121.7(9)
C59	N13	Ru2	112.6(8)
C60	N14	Ru2	111.4(8)
C64	N14	Ru2	125.2(8)
C64	N14	C60	123.4(11)
N1	C1	C2	120.9(10)
C3	C2	C1	119.7(10)
C4	C3	C2	118.0(10)
C3	C4	C11	121.4(10)
C5	C4	C3	118.9(10)
C5	C4	C11	119.6(10)
N1	C5	C4	123.5(10)
N1	C5	C6	116.1(9)
C4	C5	C6	120.4(10)
N2	C6	C5	114.9(9)
C7	C6	N2	123.0(10)
C7	C6	C5	122.0(9)
C6	C7	C24	117.9(10)
C8	C7	C6	119.9(10)
C8	C7	C24	122.2(10)
C7	C8	C9	117.5(11)
C10	C9	C8	119.4(10)
N2	C10	C9	124.3(10)
N3	C11	C4	119.5(10)
N12	C54	C53	122.8(12)
N13	C55	C56	119.1(12)
C57	C56	C55	121.4(13)
C56	C57	C58	117.6(12)
C57	C58	C59	125.4(12)
C58	C59	N13	114.7(10)
C58	C59	C60	129.6(11)
C60	C59	N13	115.5(10)
N14	C60	C59	120.0(11)
N14	C60	C61	116.0(11)
C59	C60	C61	123.9(11)
C62	C61	C60	121.7(13)
C61	C62	C63	117.6(13)
C62	C63	C64	118.7(12)
N14	C64	C63	122.4(12)
F1	P1	F4	91.7(6)
F1	P1	F5	92.0(6)
F1	P1	F6	87.8(5)
F2	P1	F1	176.9(6)
F2	P1	F4	90.5(5)
F2	P1	F5	90.0(6)
F2	P1	F6	89.8(5)
F3	P1	F1	89.0(5)
F3	P1	F2	89.0(5)
F3	P1	F4	88.1(5)
F3	P1	F5	178.6(6)
F3	P1	F6	89.0(5)
F4	P1	F5	92.9(6)
F4	P1	F6	177.1(6)
F6	P1	F5	90.0(5)
F7	P2	F8	177.4(5)
F7	P2	F10	92.5(5)
F7	P2	F11	89.9(7)
F7	P2	F12	86.8(5)
F9	P2	F7	92.2(6)
F9	P2	F8	88.6(4)
F9	P2	F10	92.7(5)
F9	P2	F11	177.8(7)
F9	P2	F12	88.7(5)
F10	P2	F8	89.9(5)
F11	P2	F8	89.3(6)
F11	P2	F10	87.7(6)
F11	P2	F12	90.9(5)
F12	P2	F8	90.7(5)
F12	P2	F10	178.5(5)
F13	P3	F14	175.0(6)

N3	C11	C24	122.1(9)
C24	C11	C4	118.3(9)
N3	C12	C13	118.6(9)
N3	C12	C23	122.6(8)
C23	C12	C13	118.8(8)
C14	C13	C12	122.1(9)
C14	C13	C17	118.4(9)
C17	C13	C12	119.1(9)
C13	C14	C15	120.7(10)
C14	C15	C16	118.4(9)
N6	C16	C15	120.1(9)
N6	C17	C13	123.5(9)
N6	C17	C18	114.9(9)
C13	C17	C18	120.7(8)
N5	C18	C17	116.9(8)
N5	C18	C19	121.5(9)
C19	C18	C17	121.6(8)
C18	C19	C23	117.2(9)
C20	C19	C18	116.8(9)
C20	C19	C23	126.0(9)
C19	C20	C21	122.4(9)
C20	C21	C22	115.0(10)
N5	C22	C21	125.8(10)
N4	C23	C12	121.0(8)
N4	C23	C19	116.8(9)
C12	C23	C19	122.2(8)
N4	C24	C7	117.1(10)
N4	C24	C11	121.1(9)
C7	C24	C11	121.5(9)

F13	P3	F15	91.6(5)
F13	P3	F16	90.2(6)
F13	P3	F17	92.9(5)
F13	P3	F18	94.7(6)
F15	P3	F14	86.7(6)
F16	P3	F14	85.0(6)
F16	P3	F15	88.5(6)
F17	P3	F14	88.6(5)
F17	P3	F15	175.1(6)
F17	P3	F16	89.6(5)
F17	P3	F18	93.3(5)
F18	P3	F14	90.0(7)
F18	P3	F15	88.2(6)
F18	P3	F16	174.2(7)
F19	P4	F20	175.6(13)
F19	P4	F22	87.9(9)
F19	P4	F23	89.4(10)
F19	P4	F24	92.5(12)
F20	P4	F22	89.3(13)
F20	P4	F23	93.9(12)
F20	P4	F24	90.9(14)
F21	P4	F19	97.6(10)
F21	P4	F20	79.3(11)
F21	P4	F22	95.2(10)
F21	P4	F23	172.1(11)
F21	P4	F24	95.2(11)
F23	P4	F22	88.8(10)
F24	P4	F22	169.5(12)
F24	P4	F23	80.7(11)

9.1.3 Δ,Δ-[2]

Table 9.7. A summary of crystallographic data and structure refinement of Δ,Δ-[2]

Identification code	iaj665v_3rdCrx_0m
Empirical formula	$C_{72}H_{44}Cl_4N_{14}Ru_2$
Formula weight	1449.15
Temperature / K	99.99
Crystal system	Orthorhombic
Space group	$P2_12_12_1$
a / Å	19.7100(9)
b / Å	22.1364(10)
c / Å	37.4364(17)
α / °	90
β / °	90
γ / °	90
Volume / Å^3	16333.8(13)
Z	8
ρ_{calc} / gcm^{-3}	1.179
μ / mm^{-1}	4.549
F(000)	5840.0
Crystal size / mm^3	0.46 × 0.113 × 0.099
Radiation	CuKα (λ = 1.54178)
2θ Range for Data Collection / °	4.638 to 80.034
Index Ranges	$-16 \leq h \leq 16$, $-18 \leq k \leq 18$, $-31 \leq l \leq 30$
Reflections Collected	58623
Independent Reflections	9515 [R_{int} = 0.0547, R_{sigma} = 0.0436]
Data / Restraints / Parameters	9515/1848/607
Goodness-of-fit on F^2	1.083
Final R indexes [I>=2σ (I)]	R_1 = 0.0766, wR_2 = 0.1952
Final R indexes [all data]	R_1 = 0.0817, wR_2 = 0.1996
Largest diff. peak/hole / e Å^{-3}	0.61/-0.53
Flack Parameter	0.157(5)

Table 9.8. Bond lengths (Å) for Δ,Δ-[2]

Atom 1	Atom 2	Length / Å	Atom 1	Atom 2	Length / Å
Ru1	N1	2.01(2)	Ru3	N15	1.936(19)
Ru1	N2	2.087(18)	Ru3	N16	1.963(19)
Ru1	N7	1.87(3)	Ru3	N21	2.01(2)
Ru1	N9	2.092(11)	Ru3	N22	2.06(2)
Ru1	N8	2.057(13)	Ru3	N23	1.93(2)
Ru1	N10	2.060(15)	Ru3	N24	2.03(2)
Ru2	N5	2.062(19)	Ru4	N19	2.028(19)
Ru2	N6	2.02(2)	Ru4	N20	2.034(19)
Ru2	N12	2.01(2)	Ru4	N25	1.96(2)
Ru2	N13	2.003(19)	Ru4	N26	2.046(11)
Ru2	N14	1.89(2)	Ru4	N28	2.062(12)
Ru2	N11	2.069(11)	Ru4	N27	2.088(11)
N1	C1	1.35(3)	N15	C75	1.40(3)
N1	C5	1.35(3)	N15	C79	1.34(3)
N2	C6	1.46(3)	N16	C80	1.39(3)
N2	C10	1.26(3)	N16	C84	1.41(3)
N3	C11	1.31(3)	N17	C85	1.33(3)
N3	C12	1.35(3)	N17	C86	1.33(3)
N4	C13	1.31(3)	N18	C87	1.45(3)
N4	C14	1.30(3)	N18	C88	1.32(3)
N5	C15	1.34(3)	N19	C89	1.37(3)
N5	C19	1.35(3)	N19	C93	1.32(3)
N6	C20	1.41(3)	N20	C94	1.35(3)
N6	C24	1.35(3)	N20	C98	1.37(3)
N7	C25	1.38(3)	N21	C99	1.29(3)
N7	C29	1.50(4)	N21	C103	1.45(2)
N9	C41	1.39	N22	C104	1.31(2)
N9	C37	1.39	N22	C105	1.35(3)
C41	C40	1.39	N23	C111	1.27(3)
C41	C42	1.302(17)	N23	C115	1.41(2)
C40	C39	1.39	N24	C117	1.36(3)
C40	C48	1.48(3)	N24	C116	1.32(2)
C39	C38	1.39	N25	C123	1.29(3)
C38	C37	1.39	N25	C127	1.49(3)
N12	C54	1.45(3)	N26	C129	1.39
N12	C55	1.30(3)	N26	C128	1.39
N13	C61	1.36(3)	C129	C130	1.39
N13	C65	1.37(2)	C130	C131	1.39
N14	C66	1.45(2)	C131	C132	1.39
N14	C67	1.35(3)	C132	C128	1.39
C1	C2	1.33(3)	C132	C133	1.43(3)

C2	C3	1.41(3)
C3	C4	1.47(3)
C4	C5	1.42(3)
C4	C11	1.51(3)
C5	C6	1.33(3)
C6	C7	1.44(3)
C7	C8	1.41(3)
C7	C13	1.44(3)
C8	C9	1.34(3)
C9	C10	1.50(3)
C11	C13	1.37(3)
C12	C14	1.39(3)
C12	C18	1.48(3)
C14	C21	1.55(3)
C15	C16	1.32(3)
C16	C17	1.36(3)
C17	C18	1.43(3)
C18	C19	1.37(3)
C19	C20	1.38(3)
C20	C21	1.39(3)
C21	C22	1.48(3)
C22	C23	1.36(3)
C23	C24	1.35(3)
C25	C26	1.53(4)
C26	C27	1.41(4)
C27	C28	1.39(4)
C28	C29	1.23(4)
C28	C36	1.42(4)
C29	C30	1.42(3)
C30	C34	1.39
C30	N8	1.39
C34	C33	1.39
C34	C35	1.32(3)
C33	C32	1.39
C32	C31	1.39
C31	N8	1.39
C35	C36	1.43(4)
C43	C44	1.39
C43	N10	1.39
C44	C45	1.39
C45	C46	1.39
C46	C42	1.39
C46	C47	1.36(3)
C42	N10	1.39

C128	C127	1.33(3)
N28	C140	1.39
N28	C141	1.39
C140	C144	1.39
C140	C139	1.332(16)
C144	C143	1.39
C144	C145	1.39(3)
C143	C142	1.39
C142	C141	1.39
C75	C76	1.40(3)
C76	C77	1.42(3)
C77	C78	1.40(3)
C78	C79	1.35(3)
C78	C85	1.47(3)
C79	C80	1.40(3)
C80	C81	1.47(3)
C81	C82	1.39(3)
C81	C87	1.40(3)
C82	C83	1.33(3)
C83	C84	1.35(3)
C85	C87	1.45(3)
C86	C88	1.40(3)
C86	C92	1.46(3)
C88	C95	1.50(3)
C89	C90	1.44(3)
C90	C91	1.34(3)
C91	C92	1.44(3)
C92	C93	1.41(3)
C93	C94	1.38(3)
C94	C95	1.43(3)
C95	C96	1.40(3)
C96	C97	1.37(3)
C97	C98	1.33(3)
C99	C100	1.43(3)
C100	C101	1.34(3)
C101	C102	1.40(3)
C104	C103	1.39
C104	C108	1.39
C103	C102	1.39
C102	C110	1.39
C110	C109	1.39
C109	C108	1.39
C108	C107	1.44(3)
C105	C106	1.46(3)

C47	C48	1.30(3)
C49	C50	1.39
C49	N11	1.39
C50	C51	1.39
C51	C52	1.39
C52	C53	1.39
C52	C60	1.46(3)
C53	N11	1.39
C53	C54	1.32(3)
C54	C58	1.46(3)
C55	C56	1.44(3)
C56	C57	1.35(3)
C57	C58	1.39(3)
C58	C59	1.39(3)
C59	C60	1.29(3)
C61	C62	1.47(3)
C62	C63	1.54(4)
C63	C64	1.35(3)
C65	C66	1.39
C65	C64	1.39
C66	C70	1.39
C70	C71	1.39
C70	C69	1.49(3)
C71	C72	1.39
C72	C64	1.39
C67	C68	1.40(4)
C68	C69	1.28(3)

C106	C107	1.31(3)
C111	C112	1.46(3)
C112	C113	1.45(3)
C113	C114	1.41(3)
C117	C118	1.50(4)
C118	C119	1.32(3)
C119	C120	1.50(3)
C120	C116	1.39
C120	C121	1.39
C116	C115	1.39
C115	C114	1.39
C114	C122	1.39
C122	C121	1.39
C123	C124	1.36(3)
C124	C125	1.39(3)
C125	C126	1.40(3)
C126	C127	1.47(3)
C126	C134	1.50(3)
C133	C134	1.32(3)
C135	C136	1.39
C135	N27	1.39
C136	C137	1.39
C137	C138	1.39
C138	C139	1.39
C138	C146	1.44(3)
C139	N27	1.39
C145	C146	1.26(3)

Table 9.9. Bond Angles (°) for Δ,Δ-[2]

Atom 1	Atom 2	Atom 3	Angle / °
N1	Ru1	N2	81.2(8)
N1	Ru1	N9	93.3(7)
N1	Ru1	N8	171.0(9)
N1	Ru1	N10	91.9(8)
N2	Ru1	N9	171.3(9)
N7	Ru1	N1	93.1(9)
N7	Ru1	N2	88.3(9)
N7	Ru1	N9	98.7(8)
N7	Ru1	N8	79.1(9)
N7	Ru1	N10	174.1(8)
N8	Ru1	N2	94.1(7)
N8	Ru1	N9	92.3(6)
N8	Ru1	N10	96.2(7)
N10	Ru1	N2	95.5(8)
N10	Ru1	N9	77.9(7)
N5	Ru2	N11	175.1(8)
N6	Ru2	N5	80.8(8)
N6	Ru2	N11	100.2(7)
N12	Ru2	N5	95.7(9)
N12	Ru2	N6	90.7(8)
N12	Ru2	N11	79.5(7)
N13	Ru2	N5	92.1(8)
N13	Ru2	N6	96.1(9)
N13	Ru2	N12	170.4(8)
N13	Ru2	N11	92.6(7)
N14	Ru2	N5	93.4(8)
N14	Ru2	N6	172.5(8)
N14	Ru2	N12	94.6(8)
N14	Ru2	N13	79.3(9)
N14	Ru2	N11	86.0(6)
C1	N1	Ru1	130.9(16)
C5	N1	Ru1	114.4(16)
C5	N1	C1	115(2)
C6	N2	Ru1	107.9(15)
C10	N2	Ru1	129.2(18)
C10	N2	C6	123(2)
C11	N3	C12	112(2)
C14	N4	C13	114(2)
C15	N5	Ru2	129.5(18)
C15	N5	C19	117(2)
C19	N5	Ru2	113.0(16)

Atom 1	Atom 2	Atom 3	Angle / °
N15	Ru3	N16	77.4(8)
N15	Ru3	N21	90.6(8)
N15	Ru3	N22	169.8(8)
N15	Ru3	N24	97.5(8)
N16	Ru3	N21	85.7(8)
N16	Ru3	N22	95.6(8)
N16	Ru3	N24	174.8(8)
N21	Ru3	N22	81.4(8)
N21	Ru3	N24	95.5(9)
N23	Ru3	N15	92.2(8)
N23	Ru3	N16	100.5(9)
N23	Ru3	N21	173.6(9)
N23	Ru3	N22	96.5(9)
N23	Ru3	N24	78.4(9)
N24	Ru3	N22	89.5(8)
N19	Ru4	N20	77.7(8)
N19	Ru4	N26	97.8(7)
N19	Ru4	N28	98.7(7)
N19	Ru4	N27	174.4(7)
N20	Ru4	N26	172.5(7)
N20	Ru4	N28	91.5(7)
N20	Ru4	N27	98.2(7)
N25	Ru4	N19	90.5(8)
N25	Ru4	N20	91.6(8)
N25	Ru4	N26	82.4(7)
N25	Ru4	N28	170.7(7)
N25	Ru4	N27	93.4(7)
N26	Ru4	N28	95.2(6)
N26	Ru4	N27	86.7(6)
N28	Ru4	N27	77.5(6)
C75	N15	Ru3	128.6(16)
C79	N15	Ru3	120.6(16)
C79	N15	C75	110.8(19)
C80	N16	Ru3	117.2(15)
C80	N16	C84	110.0(19)
C84	N16	Ru3	132.6(16)
C86	N17	C85	115(2)
C88	N18	C87	112.3(19)
C89	N19	Ru4	124.6(16)
C93	N19	Ru4	116.8(16)
C93	N19	C89	118(2)

C20	N6	Ru2	111.6(16)
C24	N6	Ru2	133.1(16)
C24	N6	C20	115(2)
C25	N7	Ru1	133(2)
C25	N7	C29	106(3)
C29	N7	Ru1	121.5(19)
C41	N9	Ru1	115.3(10)
C41	N9	C37	120
C37	N9	Ru1	124.6(10)
N9	C41	C40	120
C42	C41	N9	112.2(16)
C42	C41	C40	127.8(16)
C41	C40	C48	115.1(17)
C39	C40	C41	120
C39	C40	C48	124.8(17)
C38	C39	C40	120
C39	C38	C37	120
C38	C37	N9	120
C54	N12	Ru2	113.8(16)
C55	N12	Ru2	128.8(17)
C55	N12	C54	117(2)
C61	N13	Ru2	125.2(18)
C61	N13	C65	121(2)
C65	N13	Ru2	113.2(16)
C66	N14	Ru2	121.3(16)
C67	N14	Ru2	137.4(19)
C67	N14	C66	101(2)
C2	C1	N1	125(2)
C1	C2	C3	123(2)
C2	C3	C4	112(2)
C3	C4	C11	119(2)
C5	C4	C3	118(2)
C5	C4	C11	122(2)
N1	C5	C4	125(2)
C6	C5	N1	119(2)
C6	C5	C4	117(2)
C5	C6	N2	118(2)
C5	C6	C7	126(2)
C7	C6	N2	116(2)
C6	C7	C13	116(2)
C8	C7	C6	116(2)
C8	C7	C13	129(2)
C9	C8	C7	131(2)
C8	C9	C10	109(2)
C94	N20	Ru4	113.5(16)
C94	N20	C98	114(2)
C98	N20	Ru4	132.1(16)
C99	N21	Ru3	135.0(19)
C99	N21	C103	116(2)
C103	N21	Ru3	109.2(14)
C104	N22	Ru3	114.7(15)
C104	N22	C105	120(2)
C105	N22	Ru3	124.6(19)
C111	N23	Ru3	131(2)
C111	N23	C115	113(2)
C115	N23	Ru3	116.3(17)
C117	N24	Ru3	125.1(18)
C116	N24	Ru3	117.2(17)
C116	N24	C117	118(2)
C123	N25	Ru4	136.0(19)
C123	N25	C127	114(2)
C127	N25	Ru4	109.5(15)
C129	N26	Ru4	127.2(9)
C129	N26	C128	120
C128	N26	Ru4	112.7(9)
C130	C129	N26	120
C129	C130	C131	120
C132	C131	C130	120
C131	C132	C133	120.4(16)
C128	C132	C131	120
C128	C132	C133	119.6(16)
C132	C128	N26	120
C127	C128	N26	114.5(17)
C127	C128	C132	125.0(17)
C140	N28	Ru4	114.2(9)
C140	N28	C141	120
C141	N28	Ru4	125.8(9)
C144	C140	N28	120
C139	C140	N28	117.4(13)
C139	C140	C144	122.3(13)
C140	C144	C145	114.9(16)
C143	C144	C140	120
C143	C144	C145	124.5(16)
C142	C143	C144	120
C143	C142	C141	120
C142	C141	N28	120
C76	C75	N15	126(2)
C75	C76	C77	119(2)

N2	C10	C9	126(2)
N3	C11	C4	117(2)
N3	C11	C13	126(2)
C13	C11	C4	116(2)
N3	C12	C14	121(2)
N3	C12	C18	115(2)
C14	C12	C18	123(2)
N4	C13	C7	116(2)
N4	C13	C11	121(2)
C11	C13	C7	123(2)
N4	C14	C12	125(2)
N4	C14	C21	120(2)
C12	C14	C21	114(2)
C16	C15	N5	122(2)
C15	C16	C17	124(2)
C16	C17	C18	113(2)
C17	C18	C12	121(2)
C19	C18	C12	118(2)
C19	C18	C17	120(2)
N5	C19	C18	122(2)
N5	C19	C20	117(2)
C18	C19	C20	121(2)
C19	C20	N6	118(2)
C19	C20	C21	124(2)
C21	C20	N6	119(2)
C20	C21	C14	118(2)
C20	C21	C22	122(2)
C22	C21	C14	120(2)
C23	C22	C21	117(2)
C24	C23	C22	118(2)
C23	C24	N6	130(2)
N7	C25	C26	125(3)
C27	C26	C25	119(3)
C28	C27	C26	115(3)
C27	C28	C36	113(4)
C29	C28	C27	122(4)
C29	C28	C36	124(3)
C28	C29	N7	133(3)
C28	C29	C30	118(3)
C30	C29	N7	108(3)
C34	C30	C29	124(2)
C34	C30	N8	120
N8	C30	C29	116(2)
C30	C34	C33	120
C35	C34	C30	114(2)
C35	C34	C33	126(2)
C78	C77	C76	114(2)
C77	C78	C85	114(2)
C79	C78	C77	122(2)
C79	C78	C85	124(2)
N15	C79	C78	127(2)
N15	C79	C80	112(2)
C78	C79	C80	119(2)
N16	C80	C79	113(2)
N16	C80	C81	124(2)
C79	C80	C81	123(2)
C82	C81	C80	118(2)
C82	C81	C87	128(2)
C87	C81	C80	113(2)
C83	C82	C81	117(2)
C82	C83	C84	123(2)
C83	C84	N16	126(2)
N17	C85	C78	124(2)
N17	C85	C87	123(2)
C87	C85	C78	113(2)
N17	C86	C88	124(2)
N17	C86	C92	116(2)
C88	C86	C92	120(2)
N18	C87	C85	119(2)
C81	C87	N18	113(2)
C81	C87	C85	126(2)
N18	C88	C86	125(2)
N18	C88	C95	115(2)
C86	C88	C95	119(2)
N19	C89	C90	125(2)
C91	C90	C89	115(2)
C90	C91	C92	122(2)
C91	C92	C86	124(2)
C93	C92	C86	118(2)
C93	C92	C91	118(2)
N19	C93	C92	121(2)
N19	C93	C94	114(2)
C94	C93	C92	124(2)
N20	C94	C93	118(2)
N20	C94	C95	123(2)
C93	C94	C95	120(2)
C94	C95	C88	118(2)
C96	C95	C88	122(2)
C96	C95	C94	119(2)
C97	C96	C95	118(2)
C98	C97	C96	121(2)
C97	C98	N20	126(2)

C32	C33	C34	120
C33	C32	C31	120
N8	C31	C32	120
C30	N8	Ru1	115.1(12)
C31	N8	Ru1	124.9(12)
C31	N8	C30	120
C34	C35	C36	125(3)
C28	C36	C35	114(3)
C44	C43	N10	120
C43	C44	C45	120
C46	C45	C44	120
C42	C46	C45	120
C47	C46	C45	121.8(19)
C47	C46	C42	118.2(19)
C41	C42	C46	115.9(15)
C41	C42	N10	124.1(15)
C46	C42	N10	120
C43	N10	Ru1	129.5(9)
C42	N10	Ru1	110.5(9)
C42	N10	C43	120
C48	C47	C46	128(3)
C47	C48	C40	114(3)
C50	C49	N11	120
C51	C50	C49	120
C50	C51	C52	120
C51	C52	C53	120
C51	C52	C60	127.6(15)
C53	C52	C60	112.4(15)
N11	C53	C52	120
C54	C53	C52	121.5(16)
C54	C53	N11	118.5(16)
C49	N11	Ru2	127.2(8)
C53	N11	Ru2	112.5(9)
C53	N11	C49	120
N12	C54	C58	119(2)
C53	C54	N12	115(2)
C53	C54	C58	125(2)
N12	C55	C56	123(2)
C57	C56	C55	122(3)
C56	C57	C58	118(3)
C57	C58	C54	120(2)
C57	C58	C59	128(3)
C59	C58	C54	112(2)
C60	C59	C58	122(3)
C59	C60	C52	126(2)
N13	C61	C62	124(3)
C61	C62	C63	111(3)
C64	C63	C62	121(2)

N21	C99	C100	126(3)
C101	C100	C99	116(3)
C100	C101	C102	122(3)
N22	C104	C103	115.2(15)
N22	C104	C108	124.7(15)
C103	C104	C108	120
C104	C103	N21	118.9(15)
C104	C103	C102	120
C102	C103	N21	121.1(15)
C103	C102	C101	117.5(17)
C103	C102	C110	120
C110	C102	C101	122.3(17)
C109	C110	C102	120
C110	C109	C108	120
C104	C108	C107	116.8(16)
C109	C108	C104	120
C109	C108	C107	123.0(16)
N22	C105	C106	116(3)
C107	C106	C105	125(3)
C106	C107	C108	117(2)
N23	C111	C112	127(3)
C113	C112	C111	119(3)
C114	C113	C112	112(2)
N24	C117	C118	124(3)
C119	C118	C117	112(3)
C118	C119	C120	125(3)
C116	C120	C119	112.9(17)
C116	C120	C121	120
C121	C120	C119	127.1(17)
N24	C116	C120	126.4(17)
N24	C116	C115	113.6(17)
C120	C116	C115	120
C116	C115	N23	114.4(17)
C114	C115	N23	125.5(17)
C114	C115	C116	120
C115	C114	C113	122.2(18)
C115	C114	C122	120
C122	C114	C113	117.8(18)
C114	C122	C121	120
C122	C121	C120	120
N25	C123	C124	130(3)
C123	C124	C125	117(3)
C124	C125	C126	123(3)
C125	C126	C127	115(2)
C125	C126	C134	133(2)
C127	C126	C134	111(2)
C128	C127	N25	119(2)
C128	C127	C126	120(2)

N13	C65	C66	119.9(16)
N13	C65	C64	119.9(16)
C66	C65	C64	120
C65	C66	N14	106.2(16)
C70	C66	N14	133.6(16)
C70	C66	C65	120
C66	C70	C71	120
C66	C70	C69	109.2(17)
C71	C70	C69	130.7(17)
C72	C71	C70	120
C71	C72	C64	120
C63	C64	C65	122.5(19)
C63	C64	C72	117.5(19)
C72	C64	C65	120
N14	C67	C68	137(3)
C69	C68	C67	114(3)
C68	C69	C70	125(3)

C126	C127	N25	120(2)
C134	C133	C132	116(3)
C133	C134	C126	128(3)
C136	C135	N27	120
C135	C136	C137	120
C136	C137	C138	120
C137	C138	C146	129.3(15)
C139	C138	C137	120
C139	C138	C146	110.2(15)
C140	C139	C138	123.8(13)
C140	C139	N27	116.2(13)
N27	C139	C138	120
C135	N27	Ru4	125.5(9)
C139	N27	Ru4	113.8(9)
C139	N27	C135	120
C146	C145	C144	122(3)
C145	C146	C138	127(3)

9.1.4 Λ,Λ-[1]

Table 9.10. A summary of crystallographic data and structure refinement of Λ,Λ-[1]

Identification code	IAJ674k_2_0m
Empirical formula	$C_{66}H_{50}F_{24}N_{16}O_4P_4Ru_2$
Formula weight	1913.24
Temperature / K	100
Crystal system	Monoclinic
Space group	$P2_1$
a / Å	12.4658(7)
b / Å	16.2947(9)
c / Å	21.7440(12)
α / °	90
β / °	97.257(3)
γ / °	90
Volume / Å^3	4381.4(4)
Z	2
ρ_{calc} / gcm^{-3}	1.450
μ / mm^{-1}	0.520
F(000)	1908.0
Crystal size / mm^3	0.5 × 0.2 × 0.18
Radiation	MoKα (λ = 0.71073)
2θ Range for Data Collection / °	3.132 to 55.274
Index Ranges	-16 ≤ h ≤ 16, -21 ≤ k ≤ 21, -28 ≤ l ≤ 28
Reflections Collected	132518
Independent Reflections	20322 [R_{int} = 0.0879, R_{sigma} = 0.0750]
Data / Restraints / Parameters	20322/1190/1105
Goodness-of-fit on F^2	1.030
Final R indexes [I>=2σ (I)]	R_1 = 0.0710, wR_2 = 0.1735
Final R indexes [all data]	R_1 = 0.1063, wR_2 = 0.1925
Largest diff. peak/hole / e Å^{-3}	2.38/-0.80
Flack Parameter	0.065(9)

Table 9.11. Bond lengths (Å) for Λ,Λ-[1]

Atom 1	Atom 2	Length / Å	Atom 1	Atom 2	Length / Å
Ru1	N1	2.071(8)	C29	C30	1.477(19)
Ru1	N2	2.064(6)	C30	C31	1.415(16)
Ru1	N7	2.083(9)	C31	C32	1.38(2)
Ru1	N8	2.060(7)	C32	C33	1.40(2)
Ru1	N9	2.053(8)	C33	C34	1.349(15)
Ru1	N10	2.074(9)	C35	C36	1.362(17)
Ru2	N5	2.092(7)	C36	C37	1.29(2)
Ru2	N6	2.069(8)	C37	C38	1.43(2)
Ru2	N11	2.055(10)	C38	C39	1.378(16)
Ru2	N12	2.025(8)	C39	C40	1.433(16)
Ru2	N13	2.035(8)	C40	C41	1.389(15)
Ru2	N14	2.063(9)	C41	C42	1.362(18)
N1	C1	1.327(12)	C42	C43	1.426(19)
N1	C5	1.364(10)	C43	C44	1.369(18)
N2	C6	1.370(10)	C45	C46	1.37(2)
N2	C10	1.325(11)	C46	C47	1.31(2)
N3	C11	1.320(11)	C47	C48	1.41(2)
N3	C12	1.309(12)	C48	C49	1.385(18)
N4	C23	1.326(12)	C49	C50	1.441(18)
N4	C24	1.356(10)	C50	C51	1.435(17)
N5	C16	1.290(12)	C51	C52	1.40(2)
N5	C17	1.344(11)	C52	C53	1.33(2)
N6	C18	1.365(11)	C53	C54	1.401(14)
N6	C19	1.337(12)	C55	C56	1.425(15)
N7	C25	1.280(16)	C56	C57	1.408(17)
N7	C29	1.350(15)	C57	C58	1.351(17)
N8	C30	1.352(15)	C58	C59	1.427(14)
N8	C34	1.332(13)	C59	C60	1.457(18)
N9	C35	1.337(15)	C60	C61	1.392(17)
N9	C39	1.345(15)	C61	C62	1.415(19)
N10	C40	1.339(14)	C62	C63	1.37(2)
N10	C44	1.343(15)	C63	C64	1.40(2)
N11	C45	1.311(16)	P1	F1	1.604(13)
N11	C49	1.360(16)	P1	F2	1.560(12)
N12	C50	1.379(16)	P1	F3	1.593(8)
N12	C54	1.344(17)	P1	F4	1.575(8)
N13	C55	1.389(14)	P1	F5	1.571(10)
N13	C59	1.318(15)	P1	F6	1.591(7)
N14	C60	1.370(14)	P2	F7	1.559(13)
N14	C64	1.318(15)	P2	F8	1.612(11)
C1	C2	1.372(14)	P2	F9	1.614(11)

C2	C3	1.381(13)
C3	C4	1.376(13)
C4	C5	1.369(12)
C4	C11	1.480(11)
C5	C6	1.452(12)
C6	C7	1.400(11)
C7	C8	1.391(12)
C7	C24	1.433(11)
C8	C9	1.377(11)
C9	C10	1.408(12)
C11	C24	1.422(13)
C12	C13	1.486(13)
C12	C23	1.414(13)
C13	C14	1.398(14)
C13	C17	1.384(12)
C14	C15	1.353(14)
C15	C16	1.395(13)
C17	C18	1.453(13)
C18	C22	1.391(14)
C19	C20	1.400(15)
C20	C21	1.353(13)
C21	C22	1.412(15)
C22	C23	1.454(12)
C25	C26	1.39(2)
C26	C27	1.40(2)
C27	C28	1.37(2)
C28	C29	1.34(2)

P2	F10	1.586(12)
P2	F11	1.595(10)
P2	F12	1.603(11)
P3	F13	1.584(7)
P3	F14	1.583(9)
P3	F15	1.530(9)
P3	F16	1.592(11)
P3	F17	1.580(8)
P3	F18	1.561(9)
P4B	F19B	1.68(3)
P4B	F20B	1.69(2)
P4B	F21B	1.67(3)
P4B	F22B	1.66(3)
P4B	F23B	1.67(3)
P4B	F24B	1.66(3)
O1S	N1S	1.211(15)
O2S	N1S	1.177(14)
N1S	C1S	1.41(2)
O3S	N2S	1.146(18)
O4S	N2S	1.153(17)
N2S	C2S	1.46(2)
F19A	P4A	1.66(3)
F20A	P4A	1.51(3)
F21A	P4A	1.63(3)
F22A	P4A	1.59(2)
F23A	P4A	1.46(3)
F24A	P4A	1.50(3)

Table 9.12. Bond Angles (°) for Λ,Λ-[1]

Atom 1	Atom 2	Atom 3	Angle / °
N1	Ru1	N7	89.4(4)
N1	Ru1	N10	98.7(4)
N2	Ru1	N1	79.6(3)
N2	Ru1	N7	93.9(4)
N2	Ru1	N10	91.7(3)
N8	Ru1	N1	96.4(3)
N8	Ru1	N2	172.7(4)
N8	Ru1	N7	80.0(4)
N8	Ru1	N10	94.8(3)
N9	Ru1	N1	175.9(4)
N9	Ru1	N2	97.2(3)
N9	Ru1	N7	93.2(4)
N9	Ru1	N8	87.1(3)
N9	Ru1	N10	78.9(4)
N10	Ru1	N7	170.8(4)
N6	Ru2	N5	79.7(3)
N11	Ru2	N5	98.6(4)
N11	Ru2	N6	83.4(4)
N11	Ru2	N14	173.2(3)
N12	Ru2	N5	175.1(3)
N12	Ru2	N6	95.5(3)
N12	Ru2	N11	79.8(4)
N12	Ru2	N13	88.4(3)
N12	Ru2	N14	94.3(4)
N13	Ru2	N5	96.4(3)
N13	Ru2	N6	176.1(3)
N13	Ru2	N11	98.0(4)
N13	Ru2	N14	78.3(4)
N14	Ru2	N5	87.6(4)
N14	Ru2	N6	100.6(3)
C1	N1	Ru1	128.5(6)
C1	N1	C5	117.0(8)
C5	N1	Ru1	114.4(6)
C6	N2	Ru1	114.2(5)
C10	N2	Ru1	127.5(5)
C10	N2	C6	118.3(6)
C12	N3	C11	116.4(8)
C23	N4	C24	117.4(8)
C16	N5	Ru2	127.2(6)
C16	N5	C17	119.5(7)
C17	N5	Ru2	112.9(6)
C18	N6	Ru2	114.0(6)
C19	N6	Ru2	128.5(6)
C19	N6	C18	117.3(8)
C25	N7	Ru1	126.7(8)

Atom 1	Atom 2	Atom 3	Angle / °
C31	C32	C33	120.7(11)
C34	C33	C32	118.0(11)
N8	C34	C33	124.2(12)
N9	C35	C36	122.7(14)
C37	C36	C35	117.3(15)
C36	C37	C38	123.4(13)
C39	C38	C37	116.7(13)
N9	C39	C38	119.1(11)
N9	C39	C40	117.2(9)
C38	C39	C40	123.7(11)
N10	C40	C39	115.4(9)
N10	C40	C41	119.2(10)
C41	C40	C39	125.4(10)
C42	C41	C40	122.3(11)
C41	C42	C43	117.7(11)
C44	C43	C42	117.4(12)
N10	C44	C43	123.4(12)
N11	C45	C46	125.1(15)
C47	C46	C45	115.5(15)
C46	C47	C48	125.5(16)
C49	C48	C47	113.2(15)
N11	C49	C48	123.6(13)
N11	C49	C50	116.5(10)
C48	C49	C50	120.0(13)
N12	C50	C49	114.3(10)
N12	C50	C51	120.5(12)
C51	C50	C49	125.1(12)
C52	C51	C50	116.3(13)
C53	C52	C51	124.2(13)
C52	C53	C54	116.4(16)
N12	C54	C53	124.7(15)
N13	C55	C56	121.9(11)
C57	C56	C55	116.7(11)
C58	C57	C56	120.4(10)
C57	C58	C59	120.3(11)
N13	C59	C58	121.3(12)
N13	C59	C60	115.8(8)
C58	C59	C60	122.9(11)
N14	C60	C59	114.0(10)
N14	C60	C61	121.7(12)
C61	C60	C59	124.3(11)
C60	C61	C62	118.4(12)
C63	C62	C61	118.6(13)
C62	C63	C64	119.1(13)
N14	C64	C63	122.3(12)

C25	N7	C29	119.1(11)
C29	N7	Ru1	114.2(8)
C30	N8	Ru1	113.7(7)
C34	N8	Ru1	128.5(8)
C34	N8	C30	117.7(9)
C35	N9	Ru1	125.2(9)
C35	N9	C39	120.8(9)
C39	N9	Ru1	113.7(7)
C40	N10	Ru1	114.3(7)
C40	N10	C44	120.0(10)
C44	N10	Ru1	125.6(8)
C45	N11	Ru2	129.2(9)
C45	N11	C49	117.0(12)
C49	N11	Ru2	113.6(8)
C50	N12	Ru2	115.3(8)
C54	N12	Ru2	126.9(9)
C54	N12	C50	117.8(10)
C55	N13	Ru2	123.9(7)
C59	N13	Ru2	116.9(7)
C59	N13	C55	119.1(9)
C60	N14	Ru2	114.9(7)
C64	N14	Ru2	125.8(8)
C64	N14	C60	118.7(10)
N1	C1	C2	121.8(8)
C1	C2	C3	120.1(9)
C4	C3	C2	118.3(9)
C3	C4	C11	122.1(8)
C5	C4	C3	118.4(8)
C5	C4	C11	119.5(8)
N1	C5	C4	123.3(8)
N1	C5	C6	115.7(7)
C4	C5	C6	120.9(7)
N2	C6	C5	116.1(7)
N2	C6	C7	122.4(8)
C7	C6	C5	121.3(7)
C6	C7	C24	117.9(7)
C8	C7	C6	117.2(7)
C8	C7	C24	124.9(7)
C9	C8	C7	121.0(7)
C8	C9	C10	117.8(8)
N2	C10	C9	122.9(8)
N3	C11	C4	119.1(8)
N3	C11	C24	122.4(7)
C24	C11	C4	118.1(7)
N3	C12	C13	118.0(8)
N3	C12	C23	123.3(8)
C23	C12	C13	118.7(8)
C14	C13	C12	122.2(8)

F2	P1	F1	176.0(7)
F2	P1	F3	91.2(7)
F2	P1	F4	86.3(5)
F2	P1	F5	93.0(8)
F2	P1	F6	89.4(5)
F3	P1	F1	92.0(7)
F4	P1	F1	91.3(6)
F4	P1	F3	90.0(5)
F4	P1	F6	89.4(4)
F5	P1	F1	89.5(8)
F5	P1	F3	89.7(6)
F5	P1	F4	179.2(7)
F5	P1	F6	91.0(5)
F6	P1	F1	87.4(5)
F6	P1	F3	179.1(8)
F7	P2	F8	179.1(7)
F7	P2	F9	88.7(7)
F7	P2	F10	92.0(7)
F7	P2	F11	92.4(6)
F7	P2	F12	88.9(7)
F8	P2	F9	91.4(6)
F10	P2	F8	88.9(6)
F10	P2	F9	89.9(7)
F10	P2	F11	91.0(6)
F10	P2	F12	177.8(6)
F11	P2	F8	87.5(5)
F11	P2	F9	178.5(7)
F11	P2	F12	91.0(5)
F12	P2	F8	90.2(6)
F12	P2	F9	88.1(6)
F13	P3	F16	88.3(4)
F14	P3	F13	175.5(7)
F14	P3	F16	87.8(6)
F15	P3	F13	92.9(6)
F15	P3	F14	89.3(7)
F15	P3	F16	88.0(8)
F15	P3	F17	175.7(9)
F15	P3	F18	95.3(8)
F17	P3	F13	89.6(4)
F17	P3	F14	88.0(6)
F17	P3	F16	88.7(7)
F18	P3	F13	91.7(4)
F18	P3	F14	92.1(5)
F18	P3	F16	176.8(7)
F18	P3	F17	88.1(5)
F19B	P4B	F20B	173(3)
F21B	P4B	F19B	72(3)
F21B	P4B	F20B	114(4)

C17	C13	C12	119.6(9)
C17	C13	C14	118.2(9)
C15	C14	C13	117.3(8)
C14	C15	C16	121.5(9)
N5	C16	C15	121.0(8)
N5	C17	C13	122.4(9)
N5	C17	C18	117.6(8)
C13	C17	C18	120.0(9)
N6	C18	C17	115.5(8)
N6	C18	C22	122.4(9)
C22	C18	C17	122.0(8)
N6	C19	C20	122.2(9)
C21	C20	C19	121.3(10)
C20	C21	C22	117.0(10)
C18	C22	C21	119.4(8)
C18	C22	C23	118.2(9)
C21	C22	C23	122.2(9)
N4	C23	C12	120.5(8)
N4	C23	C22	118.1(8)
C12	C23	C22	121.3(8)
N4	C24	C7	118.0(8)
N4	C24	C11	119.9(8)
C11	C24	C7	122.1(7)
N7	C25	C26	125.7(13)
C25	C26	C27	116.7(15)
C28	C27	C26	115.2(16)
C29	C28	C27	125.1(14)
N7	C29	C30	115.0(12)
C28	C29	N7	118.0(13)
C28	C29	C30	126.7(12)
N8	C30	C29	117.0(10)
N8	C30	C31	123.0(12)
C31	C30	C29	119.9(12)
C32	C31	C30	116.1(13)

F22B	P4B	F19B	99(3)
F22B	P4B	F20B	81(2)
F22B	P4B	F21B	99(4)
F22B	P4B	F23B	92(3)
F22B	P4B	F24B	176(4)
F23B	P4B	F19B	95(3)
F23B	P4B	F20B	78.3(19)
F23B	P4B	F21B	164(4)
F24B	P4B	F19B	77(3)
F24B	P4B	F20B	102(3)
F24B	P4B	F21B	81(3)
F24B	P4B	F23B	87(3)
O1S	N1S	C1S	111.9(16)
O2S	N1S	O1S	132(2)
O2S	N1S	C1S	116.2(17)
O3S	N2S	O4S	109(3)
O3S	N2S	C2S	128(3)
O4S	N2S	C2S	118(2)
F20A	P4A	F19A	173(2)
F20A	P4A	F21A	89.4(17)
F20A	P4A	F22A	96.0(16)
F21A	P4A	F19A	84.2(17)
F22A	P4A	F19A	82.5(14)
F22A	P4A	F21A	93.0(13)
F23A	P4A	F19A	90.3(18)
F23A	P4A	F20A	96(2)
F23A	P4A	F21A	171.1(19)
F23A	P4A	F22A	93.2(14)
F23A	P4A	F24A	82.0(16)
F24A	P4A	F19A	94.6(16)
F24A	P4A	F20A	87.4(15)
F24A	P4A	F21A	91.4(17)
F24A	P4A	F22A	174.4(15)

9.2 CNS Scripts

The following scripts were used in the restrained molecular dynamics studies discussed in section 5. These files can be found in the MBB NMR server: /home1/simon/cns_work

Other standard files such as dna-rna-allatom.param can be found online at cns-online.org/v1.3/.

The model_anneal.inp input file was the same used for all simulations. The full input file can be found online at cns-online.org/v1.3/ and the edited block parameter definition with the restraint functions applied are shown below:

```
{- begin block parameter definition -} define(

{====================== molecular structure ========================}

{* parameter file(s) *}
{===>} par.1="dna-rna-allatom.param";
{===>} par.2="mbipyh.par";
{===>} par.3="";
{===>} par.4="";
{===>} par.5="";

{* structure file(s) *}
{===>} struct.1="s12nop.mtf";
{===>} struct.2="mbipyh_min.mtf";
{===>} struct.3="";
{===>} struct.4="";
{===>} struct.5="";

{* input coordinate file(s) *}
{===>} pdb.in.file.1="s12nop.pdb";

{===>} pdb.in.file.2="mbipyh_min.pdb";
{===>} pdb.in.file.3="";

{========================= atom selection ==========================}

{* input "backbone" selection criteria for average structure generation *}
{* for protein      (name n or name ca or name c)
   for nucleic acid (name O5' or name C5' or name C4' or name C3'
                     or name O3' or name P) *}
{===>} pdb.atom.select=(name O5' or name C5' or name C4' or name C3'
                     or name O3' or name P);

{====================== refinement parameters ======================}

{* type of molecular dynamics for hot phase *}
{+ choice: "torsion" "cartesian" +}
{===>} md.type.hot="torsion";

{* type of molecular dynamics for cool phase *}
```

```
{+ choice: "torsion" "cartesian" +}
{===>} md.type.cool="torsion";

{* refine using different initial velocities or coordinates
   (enter base name in "input coordinate files" field) *}
{+ choice: "veloc" "coord" +}
{===>} md.type.initial="veloc";

{* seed for random number generator *}
{* change to get different initial velocities *}
{===>} md.seed=5111;

{* select whether the number of structures will be either trial or
   accepted structures and whether to print only the trial, accepted,
   both sets of structures. *}
{+ list: The printing format is as follows:
         trial = pdb.out.name + _#.pdb , accepted = pdb.out.name + a_#.pdb
+}

{* are the number of structures to be trials or accepted? *}
{+ choice: "trial" "accept" +}
{===>} flg.trial.struc="trial";
{* number of trial or accepted structures *}
{===>} pdb.end.count=15;

{* print accepted structures *}
{+ choice: true false +}
{===>} flg.print.accept=false;
{* print trial structures *}
{+ choice: true false +}
{===>} flg.print.trial=true;

{* calculate an average structure for either the trial or
   accepted structure.  If calculate accepted average is false then
   an average for the trial structures will be calculated. *}

{* calculate an average structure? *}
{+ choice: true false +}
{===>} flg.calc.ave.struct=false;
{* calculate an average structure for the accepted structures? *}
{+ choice: true false +}
{===>} flg.calc.ave.accpt=false;
{* minimize average coordinates? *}
{+ choice: true false +}
{===>} flg.min.ave.coor=true;

{=================== torsion dynamics parameters ====================}

{* maximum unbranched chain length *}
{* increase for long stretches of polyalanine or for nucleic acids *}
{===>} md.torsion.maxlength=50;

{* maximum number of distinct bodies *}
{===>} md.torsion.maxtree=4;

{* maximum number of bonds to an atom *}
{===>} md.torsion.maxbond=6;

{========== parameters for high temperature annealing stage ==========}

{* temperature (proteins: 50000, dna/rna: 20000) *}
```

```
{===>} md.hot.temp=10000;
{* number of steps (proteins: 1000, dna/rna: 4000) *}
{===>} md.hot.step=1000;
{* scale factor to reduce van der Waals (repel) energy term *}
{===>} md.hot.vdw=0.1;
{* scale factor for NOE energy term *}
{===>} md.hot.noe=150;
{* scale factor for dihedral angle energy term (proteins: 100, dna/rna: 5)
*}
{===>} md.hot.cdih=5;
{* molecular dynamics timestep *}
{===>} md.hot.ss=0.015;

{======== parameters for the first slow-cool annealing stage =========}

{* temperature (cartesian: 1000, torsion: [proteins: 50000, dna/rna:
20000]) *}
{===>} md.cool.temp=1000;
{* number of steps *}
{===>} md.cool.step=1000;
{* scale factor for final van der Waals (repel) energy term
   (cartesian: 4.0, torsion: 1.0) *}
{===>} md.cool.vdw=1.0;
{* scale factor for NOE energy term *}
{===>} md.cool.noe=150;
{* scale factor for dihedral angle energy term *}
{===>} md.cool.cdih=200;
{* molecular dynamics timestep (cartesian: 0.005, torsion: 0.015) *}
{===>} md.cool.ss=0.015;
{* slow-cool annealing temperature step (cartesian: 25, torsion: 250) *}
{===>} md.cool.tmpstp=250;

{========= parameters for a second slow-cool annealing stage ==========}

{* cartesian slow-cooling annealing stage to be used only with torsion
   slow-cool annealing stage *}
{* this stage is only necessary when the macromolecule is a protein
   greater than 160 residues or in some cases for nucleic acids *}

{* use cartesian cooling stage? *}
{+ choice: true false +}
{===>} md.cart.flag=false;
{* temperature *}
{===>} md.cart.temp=2000;
{* number of steps *}
{===>} md.cart.step=3000;
{* scale factor for initial van der Waals (repel) energy term *}
{===>} md.cart.vdw.init=1.0;
{* scale factor for final van der Waals (repel) energy term *}
{===>} md.cart.vdw.finl=4.0;
{* scale factor for NOE energy term *}
{===>} md.cart.noe=150;
{* scale factor for dihedral angle energy term *}
{===>} md.cart.cdih=200;
{* molecular dynamics timestep *}
{===>} md.cart.ss=0.005;
{* slow-cool annealing temperature step *}
{===>} md.cart.tmpstp=25;

{============== parameters for final minimization stage ==============}
```

```
{* scale factor for NOE energy term *}
{===>} md.pow.noe=75;
{* scale factor for dihedral angle energy term *}
{===>} md.pow.cdih=400;
{* number of minimization steps *}
{===>} md.pow.step=400;
{* number of cycles of minimization *}
{===>} md.pow.cycl=45;

{============================== noe data ==============================}

{- Important - if you do not have a particular data set then
   set the file name to null ("") -}

{* NOE distance restraints files. *}

{* restraint set 1 file *}
{===>} nmr.noe.file.1="mbipymin_noe1.tbl";
{* restraint set 2 file *}
{===>} nmr.noe.file.2="";
{* restraint set 3 file *}
{===>} nmr.noe.file.3="";
{* restraint set 4 file *}
{===>} nmr.noe.file.4="";
{* restraint set 5 file *}
{===>} nmr.noe.file.5="";
{* NOE averaging modes *}

{* restraint set 1 *}
{+ choice: "sum" "cent" "R-6" "R-3" "symm" +}
{===>} nmr.noe.ave.mode.1="sum";
{* restraint set 2 *}
{+ choice: "sum" "cent" "R-6" "R-3" "symm" +}
{===>} nmr.noe.ave.mode.2="sum";
{* restraint set 3 *}
{+ choice: "sum" "cent" "R-6" "R-3" "symm" +}
{===>} nmr.noe.ave.mode.3="R-6";
{* restraint set 4 *}
{+ choice: "sum" "cent" "R-6" "R-3" "symm" +}
{===>} nmr.noe.ave.mode.4="sum";
{* restraint set 5 *}
{+ choice: "sum" "cent" "R-6" "R-3" "symm" +}
{===>} nmr.noe.ave.mode.5="sum";

{======================== hydrogen bond data ========================}

{* hydrogen-bond distance restraints file. *}
{===>} nmr.noe.hbnd.file="hbonds_b-dna_default";

{* enter hydrogen-bond distance averaging mode *}
{+ choice: "sum" "cent" "R-6" "R-3" "symm" +}
{===>} nmr.noe.ave.mode.hbnd="sum";

{======================== other restraint data ========================}

{* dihedral angle restraints file *}
{* Note: the restraint file MUST NOT contain restraints
         dihedral or end *}
{===>} nmr.cdih.file="dihedral530.tbl";

{* DNA-RNA base planarity restraints file *}
```

```
{* Note: include weights as $pscale in the restraint file *}
{===>} nmr.plan.file="plane_B-dna_default.tbl";
{* input planarity scale factor - this will be written into $pscale *}
{===>} nmr.plan.scale=150;

{* NCS-restraints file *}
{* example is in inputs/xtal_data/eg1_ncs_restrain.dat *}
{===>} nmr.ncs.file="";

{======================== input/output files ==========================}

{* base name for input coordinate files *}
{===>} pdb.in.name="mbipy_complex_start.pdb";

{* base name for output coordinate files *}
{===>} pdb.out.name="complex_new";

flg.cv.flag=false;
flg.cv.noe=false;
flg.cv.coup=false;
flg.cv.cdih=false;
flg.dgsa.flag=false;
nmr.cv.numpart=10;

 ) {- end block parameter definition -}
```

9.2.1 NOE Restraints for Λ,Δ-[1]a

The NOE restraints used in the final MD simulations for Λ,Δ-[1]a are given below.

```
!mbipymin_noe1
!LD1a
ASSI ((resid 12 and name H71) or (resid 12 and name H72) or (resid 12 and
name H73)) (resid 23 and name H71) 3.0 1.2 1.5
ASSI ((resid 12 and name H71) or (resid 12 and name H72) or (resid 12 and
name H73)) (resid 23 and name H72) 2.0 0.2 1.5
ASSI ((resid 12 and name H71) or (resid 12 and name H72) or (resid 12 and
name H73)) (resid 23 and name H73) 3.0 1.2 1.5
ASSI (resid 12 and name H2') (resid 23 and name H73) 3.0 1.2 1.0

ASSI ((resid 4 and name H71) or (resid 12 and name H72) or (resid 12 and
name H73)) (resid 23 and name H65) 3.0 1.2 1.5
ASSI ((resid 4 and name H71) or (resid 12 and name H72) or (resid 12 and
name H73)) (resid 23 and name H66) 2.0 0.2 1.5
ASSI ((resid 4 and name H71) or (resid 12 and name H72) or (resid 12 and
name H73)) (resid 23 and name H67) 3.0 1.2 1.5
ASSI (resid 4 and name H2') (resid 23 and name H67) 3.0 1.2 1.0
ASSI (resid 4 and name H2'') (resid 23 and name H67) 3.0 1.2 1.0

ASSI (resid 4 and name H2'') (resid 23 and name H97) 4.0 2.2 1.0
ASSI (resid 4 and name H1') (resid 23 and name H98) 3.0 1.2 1.0
ASSI (resid 4 and name H2') (resid 23 and name H98) 3.0 1.2 1.0
ASSI (resid 4 and name H2'') (resid 23 and name H98) 3.0 1.2 1.0
ASSI (resid 5 and name H1') (resid 23 and name H98) 4.0 2.2 1.0
ASSI (resid 4 and name H1') (resid 23 and name H99) 2.0 0.2 1.0
ASSI (resid 4 and name H2') (resid 23 and name H99) 2.0 0.2 1.0
ASSI (resid 4 and name H2'') (resid 23 and name H99) 2.0 0.2 1.0
ASSI (resid 5 and name H1') (resid 23 and name H99)   4.0 2.2 1.0
```

```
ASSI (resid 12 and name H1') (resid 23 and name H104) 3.0 1.2 1.0
ASSI (resid 12 and name H2') (resid 23 and name H104  3.0 1.2 1.0
ASSI (resid 12 and name H2'') (resid 23 and name H104) 3.0 1.2 1.0
ASSI (resid 12 and name H1') (resid 23 and name H105) 2.0 0.2 1.0
ASSI (resid 12 and name H2') (resid 23 and name H105) 2.0 0.2 1.0
ASSI (resid 12 and name H2'') (resid 23 and name H105) 3.0 1.2 1.0
ASSI (resid 13 and name H1') (resid 23 and name H105) 4.0 0.2 1.0
```

9.2.2 NOE Restraints for Λ,Δ-[1]b

```
!mbipy_noe1
!LD1b
ASSI ((resid 12 and name H71) or (resid 12 and name H72) or (resid 12 and
name H73)) (resid 23 and name H97) 3.0 1.2 1.5
ASSI ((resid 12 and name H71) or (resid 12 and name H72) or (resid 12 and
name H73)) (resid 23 and name H98) 2.0 0.2 1.5
ASSI ((resid 12 and name H71) or (resid 12 and name H72) or (resid 12 and
name H73)) (resid 23 and name H99) 3.0 1.2 1.5
ASSI (resid 12 and name H2') (resid 23 and name H99) 3.0 1.2 1.0

ASSI ((resid 4 and name H71) or (resid 12 and name H72) or (resid 12 and
name H73)) (resid 23 and name H103) 3.0 1.2 1.5
ASSI ((resid 4 and name H71) or (resid 12 and name H72) or (resid 12 and
name H73)) (resid 23 and name H104) 2.0 0.2 1.5
ASSI ((resid 4 and name H71) or (resid 12 and name H72) or (resid 12 and
name H73)) (resid 23 and name H105) 3.0 1.2 1.5
ASSI (resid 4 and name H2') (resid 23 and name H105) 3.0 1.2 1.0
ASSI (resid 4 and name H2'') (resid 23 and name H105) 3.0 1.2 1.0

ASSI (resid 4 and name H2'') (resid 23 and name H71) 4.0 2.2 1.0
ASSI (resid 4 and name H1') (resid 23 and name H72) 3.0 1.2 1.0
ASSI (resid 4 and name H2') (resid 23 and name H72) 3.0 1.2 1.0
ASSI (resid 4 and name H2'') (resid 23 and name H72) 3.0 1.2 1.0
ASSI (resid 5 and name H1') (resid 23 and name H72) 4.0 2.2 1.0
ASSI (resid 4 and name H1') (resid 23 and name H73) 2.0 0.2 1.0
ASSI (resid 4 and name H2') (resid 23 and name H73) 2.0 0.2 1.0
ASSI (resid 4 and name H2'') (resid 23 and name H73) 2.0 0.2 1.0
ASSI (resid 5 and name H1') (resid 23 and name H73)   2.0 0.2 1.0

ASSI (resid 12 and name H1') (resid 23 and name H66) 3.0 1.2 1.0
ASSI (resid 12 and name H2') (resid 23 and name H66) 3.0 1.2 1.0
ASSI (resid 12 and name H2'') (resid 23 and name H66) 3.0 1.2 1.0
ASSI (resid 12 and name H1') (resid 23 and name H67) 2.0 0.2 1.0
ASSI (resid 12 and name H2') (resid 23 and name H67) 2.0 0.2 1.0
ASSI (resid 12 and name H2'') (resid 23 and name H67) 3.0 1.2 1.0
ASSI (resid 13 and name H1') (resid 23 and name H67) 4.0 0.2 1.0
```

9.2.3 NOE Restraints for Λ,Λ-[1]a

```
!lbipymaj_noe1
!LL1a
ASSI ((resid 12 and name H71) or (resid 12 and name H72) or (resid 12 and
name H73)) (resid 23 and name H97) 3.0 1.2 1.0
ASSI ((resid 12 and name H71) or (resid 12 and name H72) or (resid 12 and
name H73)) (resid 23 and name H98)  3.0 1.2 1.0
```

```
ASSI ((resid 12 and name H71) or (resid 12 and name H72) or (resid 12 and
name H73)) (resid 23 and name H99) 3.0 1.2 1.0
ASSI (resid 12 and name H2') ((resid 23 and name H99) 4.0 2.2 1.0
ASSI (resid 12 and name H2'') ((resid 23 and name H99) 4.0 2.2 1.0

ASSI ((resid 4 and name H71) or (resid 12 and name H72) or (resid 12 and
name H73)) (resid 23 and name H103) 3.0 1.2 1.0
ASSI ((resid 4 and name H71) or (resid 12 and name H72) or (resid 12 and
name H73)) (resid 23 and name H104) 3.0 1.2 1.0
ASSI ((resid 4 and name H71) or (resid 12 and name H72) or (resid 12 and
name H73)) (resid 23 and name H105) 3.0 1.2 1.0
ASSI (resid 4 and name H2') (resid 23 and name H105) 4.0 2.2 1.0
ASSI (resid 4 and name H2'') (resid 23 and name H105) 4.0 2.2 1.0

ASSI (resid 4 and name H1') (resid 23 and name H71) 3.0 1.2 1.0
ASSI (resid 4 and name H2') (resid 23 and name H71) 3.0 1.2 1.0
ASSI (resid 4 and name H2'') (resid 23 and name H71) 3.0 1.2 1.0
ASSI (resid 5 and name H1') (resid 23 and name H71) 3.0 1.2 1.0
ASSI (resid 4 and name H1') (resid 23 and name H72) 3.0 1.2 1.0
ASSI (resid 4 and name H2') (resid 23 and name H72) 3.0 1.2 1.0
ASSI (resid 4 and name H2'') (resid 23 and name H72) 3.0 1.2 1.0
ASSI (resid 5 and name H1') (resid 23 and name H72) 3.0 1.2 1.0

ASSI (resid 4 and name H1') (resid 23 and name H73) 2.0 0.2  1.0
ASSI (resid 4 and name H2') (resid 23 and name H73) 2.0 0.2  1.0
ASSI (resid 4 and name H2'') (resid 23 and name H73) 2.0 0.2  1.0
ASSI (resid 5 and name H1') (resid 23 and name H73) 3.0 1.2 1.0

ASSI (resid 12 and name H1') (resid 23 and name H65) 3.0 1.2 1.0
ASSI (resid 13 and name H1') (resid 23 and name H65) 3.0 1.2 1.0
ASSI (resid 12 and name H1') (resid 23 and name H66) 4.0 2.2 1.0
ASSI (resid 13 and name H1') (resid 23 and name H66) 3.0 1.2 1.0

ASSI (resid 12 and name H1') (resid 23 and name H67) 2.0 0.2  1.0
ASSI (resid 12 and name H2') (resid 23 and name H67) 2.0 0.2 1.0
ASSI (resid 12 and name H2'') (resid 23 and name H67) 2.0 0.2  1.0
ASSI (resid 13 and name H1') (resid 23 and name H67) 3.0 1.2 1.0
```

9.2.4 NOE Restraints for Λ,Λ-[1]b

```
!lbipy_noe1
!LL1b
ASSI (resid 2 and name H5) (resid 23 and name H65) 4.0 2.2 1.0
ASSI (resid 2 and name H71) (resid 23 and name H66) 4.0 2.2 1.0
ASSI (resid 2 and name H5) (resid 23 and name H67) 3.0 1.2 1.0
ASSI (resid 2 and name H1') (resid 23 and name H67) 4.0 2.2 1.0

ASSI (resid 15 and name H1') (resid 23 and name H104) 3.0 1.2 1.0
ASSI (resid 15 and name H1') (resid 23 and name H105) 3.0 1.2 1.0
ASSI (resid 14 and name H1') (resid 23 and name H105) 2.0 0.2 1.0
ASSI (resid 14 and name H2') (resid 23 and name H105) 2.0 0.2 1.0
ASSI (resid 14 and name H2'') (resid 23 and name H105) 2.0 0.2 1.0

ASSI ((resid 14 and name H71) or (resid 12 and name H72) or (resid 12 and
name H73)) (resid 23 and name H71) 3.0 1.2 1.5
ASSI ((resid 14 and name H71) or (resid 12 and name H72) or (resid 12 and
name H73)) (resid 23 and name H72) 3.0 1.2 1.5
ASSI ((resid 14 and name H71) or (resid 12 and name H72) or (resid 12 and
name H73)) (resid 23 and name H73) 3.0 1.2 1.0
```

```
ASSI (resid 14 and name H2') (resid 23 and name H73) 3.0 1.2 1.0
ASSI (resid 14 and name H2'') (resid 23 and name H73) 3.0 1.2 1.0

ASSI (resid 2 and name H1') (resid 23 and name H97) 2.0 0.2 1.0
ASSI (resid 3 and name H1') (resid 23 and name H97) 3.0 1.2 1.0
ASSI (resid 2 and name H1') (resid 23 and name H98) 2.0 0.2 1.0
ASSI (resid 2 and name H2') (resid 23 and name H98) 2.0 0.2 1.0
ASSI (resid 2 and name H2'') (resid 23 and name H98)2.0 0.2 1.0

ASSI (resid 3 and name H1') (resid 23 and name H98)) 4.0 2.2 1.0
ASSI (resid 2 and name H1') (resid 23 and name H99) 2.0 0.2 1.0
ASSI (resid 2 and name H2') (resid 23 and name H99) 2.0 0.2 1.0
ASSI (resid 2 and name H2'') (resid 23 and name H99) 2.0 0.2 1.0
ASSI (resid 3 and name H1') ((resid 23 and name H99) 4.0 2.2 1.0
```

9.2.5 H-bond Restraints

The following H-bond restraints were used in all MD simulations.

```
! H_bonds_b-dna-default
!1GC
assign (resid 1 and name O6) (resid 16 and name N4) 2.91 0.01 0.01
assign (resid 1 and name N1) (resid 16 and name N3) 2.95 0.01 0.01
assign (resid 1 and name N2) (resid 16 and name O2) 2.86 0.01 0.01

!2CG
assign (resid 2 and name N4) (resid 15 and name O6) 2.91 0.01 0.01
assign (resid 2 and name N3) (resid 15 and name N1) 2.95 0.01 0.01
assign (resid 2 and name O2) (resid 15 and name N2) 2.86 0.01 0.01

!3AT
assign (resid 3 and name N6) (resid 14 and name O4) 2.95 0.01 0.01
assign (resid 3 and name N1) (resid 14 and name N3) 2.82 0.01 0.01

!4TA
assign (resid 4 and name O4) (resid 13 and name N6) 2.95 0.01 0.01
assign (resid 4 and name N3) (resid 13 and name N1) 2.82 0.01 0.01

!5AT
assign (resid 5 and name N6) (resid 12 and name O4) 2.95 0.01 0.01
assign (resid 5 and name N1) (resid 12 and name N3) 2.82 0.01 0.01

!6TA
assign (resid 6 and name O4) (resid 11 and name N6) 2.95 0.01 0.01
assign (resid 6 and name N3) (resid 11 and name N1) 2.82 0.01 0.01

assign (resid 7 and name N4) (resid 10 and name O6) 2.91 0.01 0.01
assign (resid 7 and name N3) (resid 10 and name N1) 2.95 0.01 0.01
assign (resid 7 and name O2) (resid 10 and name N2) 2.86 0.01 0.01

!8GC
assign (resid 8 and name O6) (resid 9 and name N4) 2.91 0.01 0.01
assign (resid 8 and name N1) (resid 9 and name N3) 2.95 0.01 0.01
assign (resid 8 and name N2) (resid 9 and name O2) 2.86 0.01 0.01
```

9.2.6 Planarity Restraints

The following planarity restraints were used in all MD simulations.

```
!planarity restaints
      group
            selection=((not( hydro or name P or name O1P or name O2P or
name O5' or name H5T or name H1' or name H2' or name H3' or name H4' or
name H5' or name H2""" or name H5""" or name O3' {or name C1'} or name C2'
or name C3' or name O2' or name C4' or name C5' or name O4' or name H3T)
and resid 1 ) or  (not( hydro or name P or name O1P or name O2P or name O5'
or name H5T or name H1' or name H2' or name H3' or name H4' or name H5' or
name H2""" or name H5""" or name O3' {or name C1'} or name C2' or name C3'
or name O2' or name C4' or name C5' or name O4' or name H3T) and resid 16 )
      weight=20.0
     end

        group
                selection=( (not( hydro or name P or name O1P or name O2P
or name O5' or name H5T or name H1' or name H2' or name H3' or name H4' or
name H5' or name H2""" or name H5""" or name O3' {or name C1'} or name C2'
or name C3' or name O2' or name C4' or name C5' or name O4' or name H3T)
and resid 2 ) or  (not( hydro or name P or name O1P or name O2P or name O5'
or name H5T or name H1' or name H2' or name H3' or name H4' or name H5' or
name H2""" or name H5""" or name O3' {or name C1'} or name C2' or name C3'
or name O2' or name C4' or name C5' or name O4' or name H3T) and resid 15 )
        weight=20.0
     end

        group
                selection=((not( hydro or name P or name O1P or name O2P or
name O5' or name H5T or name H1' or name H2' or name H3' or name H4' or
name H5' or name H2""" or name H5""" or name O3' {or name C1'} or name C2'
or name C3' or name O2' or name C4' or name C5' or name O4' or name H3T)
and resid 3 ) or  (not( hydro or name P or name O1P or name O2P or name O5'
or name H5T or name H1' or name H2' or name H3' or name H4' or name H5' or
name H2""" or name H5""" or name O3' {or name C1'} or name C2' or name C3'
or name O2' or name C4' or name C5' or name O4' or name H3T) and resid 14 )
        weight=20.0
     end

      group
            selection=((not( hydro or name P or name O1P or name O2P or
name O5' or name H5T or name H1' or name H2' or name H3' or name H4' or
name H5' or name H2""" or name H5""" or name O3' {or name C1'} or name C2'
or name C3' or name O2' or name C4' or name C5' or name O4' or name H3T)
and resid 4 ) or  (not( hydro or name P or name O1P or name O2P or name O5'
or name H5T or name H1' or name H2' or name H3' or name H4' or name H5' or
name H2""" or name H5""" or name O3' {or name C1'} or name C2' or name C3'
or name O2' or name C4' or name C5' or name O4' or name H3T) and resid 13 )
      weight=20.0
     end

        group
                selection=((not( hydro or name P or name O1P or name O2P or
name O5' or name H5T or name H1' or name H2' or name H3' or name H4' or
name H5' or name H2""" or name H5""" or name O3' {or name C1'} or name C2'
or name C3' or name O2' or name C4' or name C5' or name O4' or name H3T)
and resid 5 ) or  (not( hydro or name P or name O1P or name O2P or name O5'
or name H5T or name H1' or name H2' or name H3' or name H4' or name H5' or
```

```
name H2""" or name H5""" or name O3' {or name C1'} or name C2' or name C3'
or name O2' or name C4' or name C5' or name O4' or name H3T) and resid 12 )
        weight=20.0
     end
      group
                selection=((not( hydro or name P or name O1P or name O2P or
name O5' or name H5T or name H1' or name H2' or name H3' or name H4' or
name H5' or name H2""" or name H5""" or name O3' {or name C1'} or name C2'
or name C3' or name O2' or name C4' or name C5' or name O4' or name H3T)
and resid 6 ) or  (not( hydro or name P or name O1P or name O2P or name O5'
or name H5T or name H1' or name H2' or name H3' or name H4' or name H5' or
name H2""" or name H5""" or name O3' {or name C1'} or name C2' or name C3'
or name O2' or name C4' or name C5' or name O4' or name H3T) and resid 11 )
        weight=20.0
     end

      group
            selection=((not( hydro or name P or name O1P or name O2P or
name O5' or name H5T or name H1' or name H2' or name H3' or name H4' or
name H5' or name H2""" or name H5""" or name O3' {or name C1'} or name C2'
or name C3' or name O2'or name C4' or name C5' or name O4' or name H3T)and
resid 7 ) or  (not( hydro or name P or name O1P or name O2P or name O5' or
name H5T or name H1' or name H2' or name H3' or name H4' or name H5' or
name H2""" or name H5""" or name O3' {or name C1'} or name C2' or name C3'
or name O2' or name C4' or name C5' or name O4' or name H3T) and resid 10 )
      weight=20.0
     end

      group
                selection=((not( hydro or name P or name O1P or name O2P or
name O5' or name H5T or name H1' or name H2' or name H3' or name H4' or
name H5' or name H2""" or name H5""" or
name O3' {or name C1'} or name C2' or name C3' or name O2'
or name C4' or name C5' or name O4' or name H3T)and resid 8 )or  (not(
hydro or name P or name O1P or name O2P or name O5' or name H5T or name H1'
or name H2' or name H3' or name H4' or name H5'or name H2""" or name H5"""
or name O3' {or name C1'} or name C2' or name C3' or name O2'or name C4' or
name C5' or name O4' or name H3T)and resid 9 )
        weight=20.0
     end
```

9.2.7 Dihedral Restraints

Dihedral restraints used in MD simulation of Λ,Δ-[1]a, Λ,Δ-[1]b and Λ,Λ-[1]a are listed below.

The restraints for Λ,Λ-[1]b are identical except for the range for dihedral angle restraints is changed to 30 for the residues 2,3,14 and 15 instead of 4,512 and 13.

```
!dihedral530
!dihedral restraints for sugar pucker
!1
assign (resid 1 and name C1') (resid 1 and name C2')
(resid 1 and name C3') (resid 1 and name C4') 20.0 -34.9 5.0 2
assign (resid 1 and name C5') (resid 1 and name C4')
```

```
(resid 1 and name C3') (resid 1 and name C2') 20.0 -86.4 5.0 2
assign (resid 1 and name C1') (resid 1 and name O4')
(resid 1 and name C4') (resid 1 and name C5') 20.0 106.4 5.0 2

!2
assign (resid 2 and name C1') (resid 2 and name C2')
(resid 2 and name C3') (resid 2 and name C4')  20.0  -34.9 5.0 2
assign (resid 2 and name C5') (resid 2 and name C4')
(resid 2 and name C3') (resid 2 and name C2') 20.0  -86.4 5.0 2
assign (resid 2 and name C1') (resid 2 and name O4')
(resid 2 and name C4') (resid 2 and name C5') 20.0 106.4 5.0 2

!3
assign (resid 3 and name C1') (resid 3 and name C2')
(resid 3 and name C3') (resid 3 and name C4') 20.0 -34.9 5.0 2
assign (resid 3 and name C5') (resid 3 and name C4')
(resid 3 and name C3') (resid 3 and name C2') 20.0 -86.4 5.0 2
assign (resid 3 and name C1') (resid 3 and name O4')
(resid 3 and name C4') (resid 3 and name C5') 20.0 106.4 5.0 2

!4
assign (resid 4 and name C1') (resid 4 and name C2')
(resid 4 and name C3') (resid 4 and name C4') 20.0 -34.9 5.0 2
assign (resid 4 and name C5') (resid 4 and name C4')
(resid 4 and name C3') (resid 4 and name C2') 20.0 -86.4 5.0 2
assign (resid 4 and name C1') (resid 4 and name O4')
(resid 4 and name C4') (resid 4 and name C5') 20.0 106.4 5.0 2

!5
assign (resid 5 and name C1') (resid 5 and name C2')
(resid 5 and name C3') (resid 5 and name C4')  20.0  -34.9 5.0 2
assign (resid 5 and name C5') (resid 5 and name C4')
(resid 5 and name C3') (resid 5 and name C2')  20.0  -86.4 5.0 2
assign (resid 5 and name C1') (resid 5 and name O4')
(resid 5 and name C4') (resid 5 and name C5')  20.0 106.4 5.0 2

!6
assign (resid 6 and name C1') (resid 6 and name C2')
(resid 6 and name C3') (resid 6 and name C4')  20.0  -34.9 5.0 2
assign (resid 6 and name C5') (resid 6 and name C4')
(resid 6 and name C3') (resid 6 and name C2')  20.0  -86.4 5.0 2
assign (resid 6 and name C1') (resid 6 and name O4')
(resid 6 and name C4') (resid 6 and name C5')  20.0 106.4 5.0 2

!7
assign (resid 7 and name C1') (resid 7 and name C2')
(resid 7 and name C3') (resid 7 and name C4') 20.0 -34.9 5.0 2
assign (resid 7 and name C5') (resid  7 and name C4')
(resid 7 and name C3') (resid 7 and name C2') 20.0 -86.4 5.0 2
assign (resid 7 and name C1') (resid 7 and name O4')
(resid 7 and name C4') (resid 7 and name C5') 20.0 106.4 5.0 2

!8
assign (resid 8 and name C1') (resid 8 and name C2')
(resid 8 and name C3') (resid 8 and name C4') 20.0 -34.9 5.0 2
assign (resid 8 and name C5') (resid 8 and name C4')
(resid 8 and name C3') (resid 8 and name C2') 20.0 -86.4 5.0 2

!9
assign (resid 9 and name C1') (resid 9 and name C2')
(resid 9 and name C3') (resid 9 and name C4') 20.0 -34.9 5.0 2
```

```
assign (resid 9 and name C5') (resid 9 and name C4')
(resid 9 and name C3') (resid 9 and name C2') 20.0 -86.4 5.0 2
assign (resid   9   and name C1') (resid 9 and name O4')
(resid 9 and name C4') (resid 9 and name C5') 20.0 106.4 5.0 2

!10
assign (resid 10  and name C1') (resid 10 and name C2')
(resid 10 and name C3') (resid 10 and name C4') 20.0 -34.9 5.0 2
assign (resid 10 and name C5') (resid 10 and name C4')
(resid 10 and name C3') (resid 10 and name C2') 20.0 -86.4 5.0 2
assign (resid   10  and name C1') (resid   10  and name O4')
(resid 10 and name C4') (resid 10 and name C5') 20.0 106.4 5.0 2

!11
assign (resid 11 and name C1') (resid 11 and name C2')
(resid 11 and name C3') (resid 11 and name C4') 20.0 -34.9 5.0 2
assign (resid 11  and name C5') (resid   11  and name C4')
(resid 11 and name C3') (resid 11 and name C2') 20.0 -86.4 5.0 2
assign (resid 11 and name C1') (resid 11 and name O4')
(resid  11 and name C4') (resid 11 and name C5') 20.0 106.4 5.0 2

!12
assign (resid 12 and name C1') (resid 12 and name C2')
(resid 12 and name C3') (resid 12 and name C4') 20.0 -34.9 5.0  2
assign (resid 12 and name C5') (resid 12 and name C4')
(resid 12 and name C3') (resid 12 and name C2') 20.0 -86.4 5.0 2
assign (resid 12 and name C1') (resid   12  and name O4')
(resid 12 and name C4') (resid 12 and name C5') 20.0 106.4 5.0 2

!13
assign (resid 13 and name C1') (resid 13 and name C2')
(resid 13 and name C3') (resid 13 and name C4') 20.0 -34.9 5.0 2
assign (resid 13 and name C5') (resid 13 and name C4')
(resid 13 and name C3') (resid 13 and name C2') 20.0 -86.4 5.0 2
assign (resid 13 and name C1') (resid 13 and name O4')
(resid 13 and name C4') (resid 13 and name C5') 20.0 106.4 5.0 2

!14
assign (resid 14 and name C1') (resid 14 and name C2')
(resid 14 and name C3') (resid 14 and name C4') 20.0  -34.9 5.0 2
assign (resid 14 and name C5') (resid 14 and name C4')
(resid 14  and name C3') (resid 14 and name C2') 20.0 -86.4 5.0 2
assign (resid 14 and name C1') (resid 14 and name O4')
(resid 14 and name C4') (resid 14 and name C5')  20.0 106.4 5.0 2

!15
assign (resid   15  and name C1') (resid   15  and name C2')
(resid 15 and name C3') (resid 15 and name C4') 20.0 -34.9 5.0  2
assign (resid   15  and name C5') (resid   15  and name C4')
(resid 15 and name C3') (resid 15 and name C2') 20.0 -86.4 5.0  2
assign (resid 15 and name C1') (resid 15 and name O4')
(resid 15 and name C4') (resid 15 and name C5') 20.0 106.4 5.0 2

!16
assign (resid 16 and name C1') (resid 16 and name C2')
(resid 16 and name C3') (resid 16 and name C4') 20.0 -34.9 5.0 2
assign (resid 16 and name C5') (resid 16 and name C4')
(resid 16 and name C3') (resid 16 and name C2') 20.0 -86.4 5.0 2
assign (resid 16 and name C1') (resid 16 and name O4')
(resid 16 and name C4') (resid 16 and name C5') 20.0 106.4 5.0 2
```

```
!dihedral restraints for back bone
!1
assign (resid 1 and name O5') (resid 1 and name C5')
(resid 1 and name C4') (resid 1 and name C3')  5.0 36.4 5.0 2
assign (resid 1 and name C5') (resid 1 and name C4')
(resid 1 and name C3') (resid 1 and name O3')  5.0 156.4 5.0 2
assign (resid 1 and name C4') (resid 1 and name C3')
(resid 1 and name O3') (resid 2 and name P  )  5.0 155.0 5.0 2
assign (resid 1 and name C3') (resid 1 and name O3')
(resid 2 and name P) (resid 2 and name O5')  5.0 -95.2 5.0 2
assign (resid 1 and name O4') (resid 1 and name C1')
(resid 1 and name N1 ) (resid 1 and name C2 )  5.0 -97.8 5.0 2

!2
assign (resid 14 and name O3') (resid 15 and name P  )
(resid 15 and name O5') (resid 15 and name C5')  5.0 -46.8 5.0 2
assign (resid 15 and name P  ) (resid 15 and name O5')
(resid 15 and name C5') (resid 15 and name C4')  5.0 -146.1 5.0 2
assign (resid 15 and name O5') (resid 15 and name C5')
(resid 15 and name C4') (resid 15 and name C3')  5.0 36.4 5.0 2
assign (resid 15 and name C5') (resid 15 and name C4')
(resid 15 and name C3') (resid 15 and name O3')  5.0 156.4 5.0 2
assign (resid 15 and name C4') (resid 15 and name C3')
(resid 15 and name O3') (resid 16 and name P  )  5.0 155.0 5.0 2
assign (resid 15 and name C3') (resid 15 and name O3')
(resid 16 and name P  ) (resid 16 and name O5')  5.0 -95.2 5.0 2
assign (resid 15 and name O4') (resid 15 and name C1')
 (resid 15 and name N9 ) (resid 15 and name C2 )  5.0 -97.8 5.0 2

!3
assign (resid 14 and name O3') (resid 15 and name P  )
 (resid 15 and name O5') (resid 15 and name C5')  5.0 -46.8 5.0 2
assign (resid 15 and name P  ) (resid 15 and name O5')
(resid 15 and name C5') (resid 15 and name C4')  5.0 -146.1 5.0 2
assign (resid 15 and name O5') (resid 15 and name C5')
(resid 15 and name C4') (resid 15 and name C3')  5.0 36.4 5.0 2
assign (resid 15 and name C5') (resid 15 and name C4')
(resid 15 and name C3') (resid 15 and name O3')  5.0 156.4 5.0 2
assign (resid 15 and name C4') (resid 15 and name C3')
(resid 15 and name O3') (resid 16 and name P  )  5.0 155.0 5.0 2
assign (resid 15 and name C3') (resid 15 and name O3')
(resid 16 and name P  ) (resid 16 and name O5')  5.0 -95.2 5.0 2
assign (resid 15 and name O4') (resid 15 and name C1')
 (resid 15 and name N9 ) (resid 15 and name C2 )  5.0 -97.8 5.0 2

!4
assign (resid 4 and name O3') (resid   5   and name P  )(resid 5 and name
O5') (resid   5   and name C5')  5.0 -46.8 30.0 2
assign (resid 5 and name P  ) (resid   5   and name O5')(resid 5 and name
C5') (resid   5   and name C4')  5.0 -146.1 30.0 2
assign (resid 5 and name O5') (resid   5   and name C5')(resid 5 and name
C4') (resid   5   and name C3')  5.0 36.4 30.0 2
assign (resid 5 and name C5') (resid   5   and name C4')(resid 5 and name
C3') (resid   5   and name O3')  5.0 156.4 30.0 2
assign (resid 5 and name C4') (resid   5   and name C3')(resid 5 and name
O3') (resid   6   and name P  )  5.0 155.0 30.0 2
assign (resid 5 and name C3') (resid   5   and name O3')(resid 6 and name P
) (resid   6   and name O5')  5.0 -95.2 30.0 2
assign (resid 5 and name O4') (resid   5   and name C1' (resid 5 and name
N9 ) (resid 5 and name C2 )  5.0 -97.8 30.0 2
```

```
!5
assign (resid 4 and name O3') (resid   5   and name P  )(resid 5 and name
O5') (resid   5   and name C5')  5.0 -46.8 30.0 2
assign (resid 5 and name P  ) (resid   5   and name O5')(resid 5 and name
C5') (resid   5   and name C4')  5.0 -146.1 30.0 2
assign (resid 5 and name O5') (resid   5   and name C5')(resid 5 and name
C4') (resid   5   and name C3')  5.0 36.4 30.0 2
assign (resid 5 and name C5') (resid   5   and name C4')(resid 5 and name
C3') (resid   5   and name O3')  5.0 156.4 30.0 2
assign (resid 5 and name C4') (resid   5   and name C3')(resid 5 and name
O3') (resid   6   and name P  )  5.0 155.0 30.0 2
assign (resid 5 and name C3') (resid   5   and name O3')(resid 6 and name P
) (resid   6   and name O5')  5.0 -95.2 30.0 2
assign (resid 5 and name O4') (resid   5   and name C1' (resid 5 and name
N9 ) (resid 5 and name C2 )  5.0 -97.8 30.0 2

!6
assign (resid 14 and name O3') (resid 15 and name P  )
 (resid 15 and name O5') (resid 15 and name C5')  5.0 -46.8 5.0 2
assign (resid 15 and name P  ) (resid 15 and name O5')
(resid 15 and name C5') (resid 15 and name C4')  5.0 -146.1 5.0 2
assign (resid 15 and name O5') (resid 15 and name C5')
(resid 15 and name C4') (resid 15 and name C3')  5.0 36.4 5.0 2
assign (resid 15 and name C5') (resid 15 and name C4')
(resid 15 and name C3') (resid 15 and name O3')  5.0 156.4 5.0 2
assign (resid 15 and name C4') (resid 15 and name C3')
(resid 15 and name O3') (resid 16 and name P  )  5.0 155.0 5.0 2
assign (resid 15 and name C3') (resid 15 and name O3')
(resid 16 and name P  ) (resid 16 and name O5')  5.0 -95.2 5.0 2
assign (resid 15 and name O4') (resid 15 and name C1')
 (resid 15 and name N9 ) (resid 15 and name C2 )  5.0 -97.8 5.0 2

!7
assign (resid 14 and name O3') (resid 15 and name P  )
 (resid 15 and name O5') (resid 15 and name C5')  5.0 -46.8 5.0 2
assign (resid 15 and name P  ) (resid 15 and name O5')
(resid 15 and name C5') (resid 15 and name C4')  5.0 -146.1 5.0 2
assign (resid 15 and name O5') (resid 15 and name C5')
(resid 15 and name C4') (resid 15 and name C3')  5.0 36.4 5.0 2
assign (resid 15 and name C5') (resid 15 and name C4')
(resid 15 and name C3') (resid 15 and name O3')  5.0 156.4 5.0 2
assign (resid 15 and name C4') (resid 15 and name C3')
(resid 15 and name O3') (resid 16 and name P  )  5.0 155.0 5.0 2
assign (resid 15 and name C3') (resid 15 and name O3')
(resid 16 and name P  ) (resid 16 and name O5')  5.0 -95.2 5.0 2
assign (resid 15 and name O4') (resid 15 and name C1')
 (resid 15 and name N9 ) (resid 15 and name C2 )  5.0 -97.8 5.0 2


!8
assign (resid 7 and name O3') (resid 8 and name P  )
(resid 8 and name O5') (resid 8 and name C5')  5.0 -46.8 5.0 2
assign (resid 8   and name P  ) (resid 8 and name O5')
(resid 8 and name C5') (resid 8 and name C4')  5.0 -146.1 5.0 2
assign (resid 8 and name O5') (resid 8 and name C5')
(resid 8 and name C4') (resid 8 and name C3')  5.0 36.4 5.0 2
assign (resid 8 and name C5') (resid 8 and name C4')
(resid 8 and name C3') (resid 8 and name O3')  5.0 156.4 5.0 2
assign (resid 8 and name O4') (resid 8 and name C1')
(resid 8 and name N9 ) (resid 8 and name C2 )  5.0 -97.8 5.0 2
```

```
! 2nd strand segment B
!9
assign (resid 9 and name O5') (resid 9   and name C5')
(resid 9 and name C4') (resid 9   and name C3')  5.0 36.4 5.0 2
assign (resid 9 and name C5') (resid 9   and name C4')
(resid 9 and name C3') (resid 9   and name O3')  5.0 156.4 5.0 2
assign (resid 9 and name C4') (resid 9   and name C3')
(resid 9 and name O3') (resid 10  and name P  )  5.0 155.0 5.0 2
assign (resid 9 and name C3') (resid 9 and name O3')
(resid 10 and name P ) (resid 10 and name O5')  5.0 -95.2 5.0 2
assign (resid 9 and name O4') (resid 9 and name C1')
(resid 9 and name N1 ) (resid 9 and name C2 )  5.0 -97.8 5.0 2

!10
assign (resid 14 and name O3') (resid 15 and name P  )
 (resid 15 and name O5') (resid 15 and name C5')  5.0 -46.8 5.0 2
assign (resid 15 and name P  ) (resid 15 and name O5')
(resid 15 and name C5') (resid 15 and name C4')  5.0 -146.1 5.0 2
assign (resid 15 and name O5') (resid 15 and name C5')
(resid 15 and name C4') (resid 15 and name C3')  5.0 36.4 5.0 2
assign (resid 15 and name C5') (resid 15 and name C4')
(resid 15 and name C3') (resid 15 and name O3')  5.0 156.4 5.0 2
assign (resid 15 and name C4') (resid 15 and name C3')
(resid 15 and name O3') (resid 16 and name P  )  5.0 155.0 5.0 2
assign (resid 15 and name C3') (resid 15 and name O3')
(resid 16 and name P  ) (resid 16 and name O5')  5.0 -95.2 5.0 2
assign (resid 15 and name O4') (resid 15 and name C1')
 (resid 15 and name N9 ) (resid 15 and name C2 )  5.0 -97.8 5.0 2

!11
assign (resid 14 and name O3') (resid 15 and name P  )
 (resid 15 and name O5') (resid 15 and name C5')  5.0 -46.8 5.0 2
assign (resid 15 and name P  ) (resid 15 and name O5')
(resid 15 and name C5') (resid 15 and name C4')  5.0 -146.1 5.0 2
assign (resid 15 and name O5') (resid 15 and name C5')
(resid 15 and name C4') (resid 15 and name C3')  5.0 36.4 5.0 2
assign (resid 15 and name C5') (resid 15 and name C4')
(resid 15 and name C3') (resid 15 and name O3')  5.0 156.4 5.0 2
assign (resid 15 and name C4') (resid 15 and name C3')
(resid 15 and name O3') (resid 16 and name P  )  5.0 155.0 5.0 2
assign (resid 15 and name C3') (resid 15 and name O3')
(resid 16 and name P  ) (resid 16 and name O5')  5.0 -95.2 5.0 2
assign (resid 15 and name O4') (resid 15 and name C1')
 (resid 15 and name N9 ) (resid 15 and name C2 )  5.0 -97.8 5.0 2

!12
assign (resid 4 and name O3') (resid   5   and name P  )(resid 5 and name
O5') (resid   5   and name C5')  5.0 -46.8 30.0 2
assign (resid 5 and name P  ) (resid   5   and name O5')(resid 5 and name
C5') (resid   5   and name C4')  5.0 -146.1 30.0 2
assign (resid 5 and name O5') (resid   5   and name C5')(resid 5 and name
C4') (resid   5   and name C3')  5.0 36.4 30.0 2
assign (resid 5 and name C5') (resid   5   and name C4')(resid 5 and name
C3') (resid   5   and name O3')  5.0 156.4 30.0 2
assign (resid 5 and name C4') (resid   5   and name C3')(resid 5 and name
O3') (resid   6   and name P  )  5.0 155.0 30.0 2
assign (resid 5 and name C3') (resid   5   and name O3')(resid 6 and name P
) (resid   6   and name O5')  5.0 -95.2 30.0 2
assign (resid 5 and name O4') (resid   5   and name C1' (resid 5 and name
N9 ) (resid 5 and name C2 )  5.0 -97.8 30.0 2
```

```
!13
assign (resid 4 and name O3') (resid   5   and name P  )(resid 5 and name
O5') (resid   5   and name C5')  5.0 -46.8 30.0 2
assign (resid 5 and name P  ) (resid   5   and name O5')(resid 5 and name
C5') (resid   5   and name C4')  5.0 -146.1 30.0 2
assign (resid 5 and name O5') (resid   5   and name C5')(resid 5 and name
C4') (resid   5   and name C3')  5.0 36.4 30.0 2
assign (resid 5 and name C5') (resid   5   and name C4')(resid 5 and name
C3') (resid   5   and name O3')  5.0 156.4 30.0 2
assign (resid 5 and name C4') (resid   5   and name C3')(resid 5 and name
O3') (resid   6   and name P  )  5.0 155.0 30.0 2
assign (resid 5 and name C3') (resid   5   and name O3')(resid 6 and name P
) (resid   6   and name O5')  5.0 -95.2 30.0 2
assign (resid 5 and name O4') (resid   5   and name C1' (resid 5 and name
N9 ) (resid 5 and name C2 )  5.0 -97.8 30.0 2

!14
assign (resid 14 and name O3') (resid 15 and name P  )
 (resid 15 and name O5') (resid 15 and name C5')  5.0 -46.8 5.0 2
assign (resid 15 and name P  ) (resid 15 and name O5')
(resid 15 and name C5') (resid 15 and name C4')  5.0 -146.1 5.0 2
assign (resid 15 and name O5') (resid 15 and name C5')
(resid 15 and name C4') (resid 15 and name C3')  5.0 36.4 5.0 2
assign (resid 15 and name C5') (resid 15 and name C4')
(resid 15 and name C3') (resid 15 and name O3')  5.0 156.4 5.0 2
assign (resid 15 and name C4') (resid 15 and name C3')
(resid 15 and name O3') (resid 16 and name P  )  5.0 155.0 5.0 2
assign (resid 15 and name C3') (resid 15 and name O3')
(resid 16 and name P  ) (resid 16 and name O5')  5.0 -95.2 5.0 2
assign (resid 15 and name O4') (resid 15 and name C1')
 (resid 15 and name N9 ) (resid 15 and name C2 )  5.0 -97.8 5.0 2

!15
assign (resid 14 and name O3') (resid 15 and name P  )
 (resid 15 and name O5') (resid 15 and name C5')  5.0 -46.8 5.0 2
assign (resid 15 and name P  ) (resid 15 and name O5')
(resid 15 and name C5') (resid 15 and name C4')  5.0 -146.1 5.0 2
assign (resid 15 and name O5') (resid 15 and name C5')
(resid 15 and name C4') (resid 15 and name C3')  5.0 36.4 5.0 2
assign (resid 15 and name C5') (resid 15 and name C4')
(resid 15 and name C3') (resid 15 and name O3')  5.0 156.4 5.0 2
assign (resid 15 and name C4') (resid 15 and name C3')
(resid 15 and name O3') (resid 16 and name P  )  5.0 155.0 5.0 2
assign (resid 15 and name C3') (resid 15 and name O3')
(resid 16 and name P  ) (resid 16 and name O5')  5.0 -95.2 5.0 2
assign (resid 15 and name O4') (resid 15 and name C1')
 (resid 15 and name N9 ) (resid 15 and name C2 )  5.0 -97.8 5.0 2

!16
assign (resid 15 and name O3') (resid 16 and name P  )
(resid 16 and name O5') (resid 16 and name C5')  5.0 -46.8 5.0 2
assign (resid 16 and name P  ) (resid 16 and name O5')
(resid 16 and name C5') (resid 16 and name C4')  5.0 -146.1 5.0 2
assign (resid 16 and name O5') (resid 16 and name C5')(resid 16 and name
C4') (resid 16 and name C3')  5.0 36.4 5.0 2
assign (resid 16 and name C5') (resid 16 and name C4')(resid 16 and name
C3') (resid 16 and name O3')  5.0 156.4 5.0 2
assign (resid 16 and name O4') (resid 16 and name C1') (resid 16 and name
N1 ) (resid 16 and name C2 )  5.0 -97.8 5.0 2
```