Structural Insights into the DNA Binding Properties of Ruthenium Polypyridyl Complexes

A thesis submitted for the degree of Doctor of Philosophy in the School of Chemistry, Food and Pharmacy

Kane Thomas McQuaid

Supervised by Prof. Christine Cardin, Dr John Brazier, and Dr Dave Allan

November 2019

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Kane Thomas McQuaid

Acknowledgements

First and foremost I would like to thank my amazing supervisors; Prof. Christine Cardin, Dr John Brazier, and Dr Dave Allan. I could not have asked for a more supportive network of advisors. Without their continuous support, wisdom, and patience throughout my studies I would have not been able to achieve a fraction of what I have. Thanks to them I have been able to grow into a more independent and critical researcher, travel around the globe to conduct my research, and finish this journey with a smile still on my face. A special thanks goes to Christine for putting up with me and for continually nudging me in the right direction; but also to my pseudo-supervisor, emeritus Prof. David Cardin, for his limitless intellect, insights, and sarcasm. In addition I would like to truly thank my assessors, Prof. Jim Thomas and Prof. Kenneth Shankland, for both sacrificing their time to read through my thesis, and for providing feedback that has ultimately developed the current work and the future directions of these projects.

Secondly I cannot thank Dr James Hall enough for his efforts throughout my studies. He has acted as an inspiration, a mentor, and as a friend. Without his crystallographic expertise and remarkable fortitude in the face of a petulant student, I quite literally would not have got this far. He has also taught me a good lesson about watching what you wish for. Banter is banter until you rip River and Simon from me!

I would also like to thank everyone that I had the pleasure to collaborate with over my time as a PhD student. I would like to thank Dr Edith Glazer and Dr Erin Wachter (University of Kentucky) for welcoming me into their wonderful group/state and allowing me loose in the labs. I would especially like to thank the wonderful group of esteemed researchers at the FIBER institute at Konan University, Japan. I owe a huge debt of gratitude to Dr Shuntaro Takahashi, Dr Hisae Tateishi-Karimata, Dr Endoh Tamaki, and Prof. Naoki Sugimoto who hosted me during my three month fellowship there and taught me so much that has allowed me to expand the scope of my research. I would also like to thank Prof. Ramon Vilar and Timothy Kench for their continued collaboration and their expertise in the use of high throuput FRET techniques. I hope and fully expect that these wonderful collaborations will continue far into the future.

In addition I would like to take this opportunity to thank the final year project students, NUIST students, and German Erasmus+ technicians that I have had the pleasure to work with. Specific thanks go to my co-authors Holly Abell and Lena Baumgärtner who, through both raw talent and hard work, achieved many successes and who separately contributed to this body of work. Big shout out to

Abstract

Ruthenium polypyridyl complexes are a class of ruthenium centred coordination complexes that are identified by their coordinate sphere of polypyridine ligands, such as phen or TAP (phen = 1,10-phenanthroline, TAP = 1,4,5,8-tetraazaphenanthrene), and by their wealth of photophysical properties. Heteroleptic variations on this, such as $[Ru(phen)_2(dppz)]^{2+}$, have shown prospective utility in a range of remedial therapeutics and diagnostics based predominantly on their ability to bind to and intercalate into DNA and its many morphologies. In this thesis, a deeper understanding of these binding modes is strived for so as to better develop a next generation of metallic therapeutic agents.

A small range of derivatives, based on the parent complex $[Ru(TAP)_2(dppz)]^{2+}$, are structurally characterised in absence and presence of the DNA decamer d(TCGGCGCCGA) to better understand how substitution affects intercalation. X-ray crystallographic data of the systems shows that incorporation of simple electron withdrawing substituents onto the distal ring of dppz (such as -C≡N or -NO$_2$) can direct base pairing at adjacent steps, causing previously flipped out nucleobases to reform a complete binding cavity.

 Next the complex rac-$[Ru(TAP)_2(11\text{-}CN\text{-}dppz)]^{2+}$ is shown to stabilise and bind, with topological preference, to the G-quadruplex forming sequence d(TAGGGTTA), binding adjacent to the G-stack. Crystallography elucidates a number of enantiospecific interactions that direct the folded topology, and is used to explain the motif-specific luminescence response of a light switch analogue complex. Further structural studies led to a second G-quadruplex structure containing the parent complex and a truncated sequence. Unlike the first, this structure contains no guanine interaction but contains multiple novel T/A binding modes such as semi-intercalation, mismatch binding, and major groove binding, all of which are compared to the potential binding pockets in the loop regions of telomeric DNA.

Lastly, a number of polypyridyl complexes are investigated in relation to their G-quadruplex binding efficacy. Of particular note is the complex Λ-$[Ru(phen)_2(Aqphen)]^{2+}$ which is demonstrated to greatly inhibit polymerisation of a G-quadruplex sequence in $vitro$ and then shown using an immunofluorescence assay, to bind strongly to G-quadruplexes in $vivo$, displacing the G-quadruplex specific BG4 antibody.

Contents

Abbreviations Used

A – adenine

AA – acetic acid

abs – absorption

ACN – acetonitrile

Aqphen – 10,11-[1,4-naphtalenedione]dipyridophenazine

ATR – attenuated total reflectance

bp – base pairs

bpy – bipyridine

BSA – bovine serum albumin

C – cytosine

CD – circular dichroism

CSD – Cambridge structural database

CT-DNA – calf thymus DNA

Da – Daltons (g/mol)

DFT – density functional theory

DIP – 4,7-diphenyl-1,10-phenanthroline

DMA – dimethylacetamide

DMSO – dimethyl sulfoxide

DNA – deoxyribonucleic acid

dNTPs – deoxyribonucleotide triphosphates

dppz – dipyridophenazine

ds – double-stranded DNA

EDTA – ethylenediaminetetraacetic acid

emi – emission

ESI – electrospray ionisation

EWG – electron withdrawing group

FAM – fluorescein amidite

FRET – Förster resonance energy transfer

G – guanine

HPLC – high performance liquid chromatography

HRMS – high resolution mass spectrometry

HTS – human telomeric sequence

IC – internal conversion

IR – infrared

ISC – intersystem crossing

KF – Klenow fragment

L – ligand

MAD – multi-wavelength anomalous dispersion

MBP – maltose-binding protein

MD – molecular dynamics

MLCT – metal-to-ligand charge transfer

MM – molecular mechanics

MPD – (+/-)-2-methyl-2,4-pentanediol

MR – molecular replacement

MS – mass spectrometry

Mw – molecular weight

NDB – nucleic acid database

NMR – nuclear magnetic resonance MM

NOE(SY) – nuclear Overhauser effect (spectroscopy)

PAGE – polyacrylamide gel electrophoresis

PBS – phosphate-buffered saline

PCR – polymerase chain reaction

PDB – protein database

phen – 1,10-phenanthroline

PFA - paraformaldehyde

PMT – photomultiplier tube

RMSD – root-mean-square deviation

RNA – ribonucleic acid

ROS – reactive oxygen species

SAD – single-wavelength anomalous dispersion

SAR – structure-activity relationship

SG – space group

SRCD – synchrotron radiation circular dichroism

ss – single-stranded DNA

T – thymine

TAMRA - tetramethylrhodamine

TAP – 1,4,5,8-tetraazaphenanthrene

TBE – Tris-borate-EDTA

TEA – triethylamine

tel – telomeric

T_m – DNA melting temperature

TMS – tetramethylsilane

TOF – time of flight

TRAP – telomerase repeated amplification protocol

UFF – universal force field

UV – ultraviolet

wt – wild type

1 Introduction

1.1 Introduction

The ability to selectively alter and manipulate DNA, the very scaffold of our existence and wellbeing, has dominated the cutting edge of biological and medicinal research since its discovery. The complexity and vast length of the polynucleotide allows it to store all essential information required to sustain functionality and development in living organisms. Amino acid sequences are synthesised via the transcription and translation of these genetic codes which, through additional conformational changes such that primary and secondary structure is achieved, become proteins and enzymes which govern the manner in which the organism functions. As a result however, damage or inconsistencies in these chromosomal DNA blueprints via genetic variation, disease, or environmental damage, can upset the normal functionality of the cell often leading to chronic cellular disease.
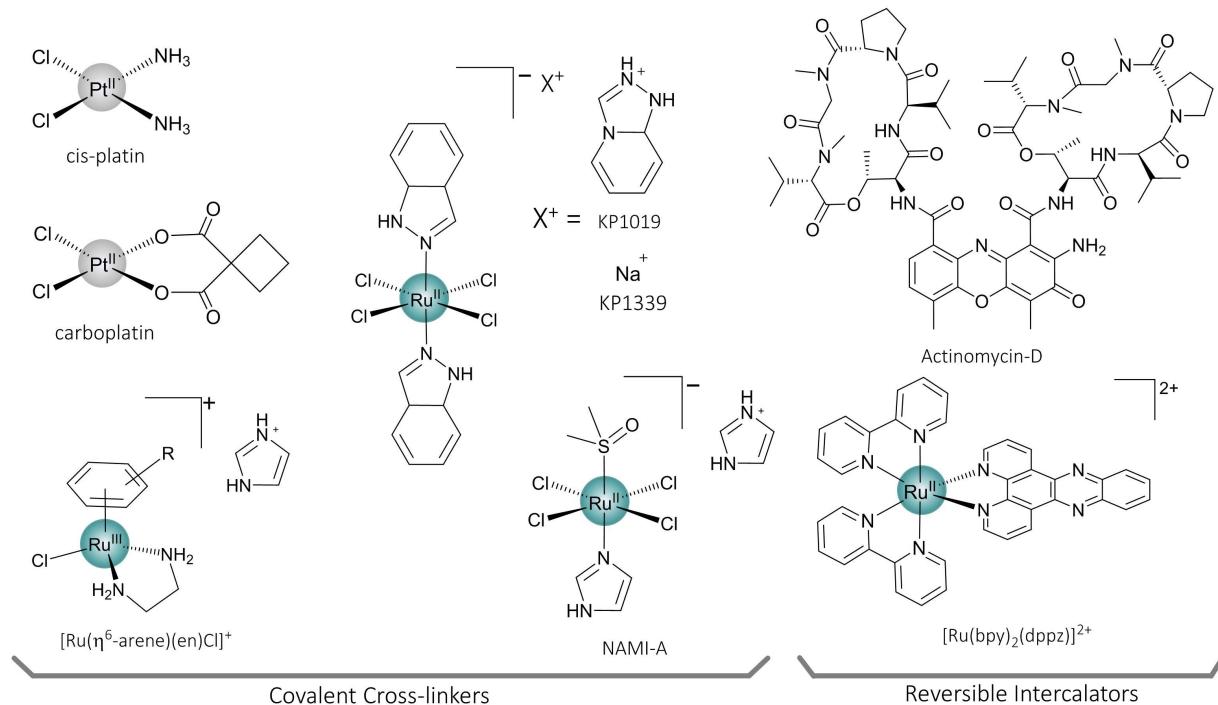
Figure 1.1 – *Well documented DNA cross-linkers and reversible intercalators. All except [Ru(bpy)$_2$(dppz)]$^{2+}$ are either common chemotherapeutics or have been clinically assessed.*

It is evident then, that disruption of this sequencing by the binding of small molecules could dictate the ability to successfully translate the DNA into transcribable RNA, and thus, the ability to modulate cell replication or assemble the applicable protein. This manner of DNA inhibition is

intensively researched and is the most prevalent mode of action for a spectrum of anti-tumour drugs such as cisplatin and Actinomycin D.[1]

The medical and socio-economic importance of such replication inhibitors is evidently important; as the prevalence of tumour related diseases continue to increase, the pressure to provide therapies which selectively halt (or reverse) tumour growth; or to visualise the structural behaviour of DNA in the cell, is also increasing.

Famously discovered by chance, cisplatin remains one of the most widely used and medically effective DNA binders to date. The structurally simple square planar platinum complex promotes irrepairable *pseudo*-alkylation *in vivo*, causing cross-linking within the DNA, inducing apoptosis and ultimately cell death.[2] The mode of action of cisplatin is fairly well understood, and is fundamentally a result of the lability of the chloride ligands; as such, once administered, a chloride ion is displaced in favour of an aqua ligand, revising the ionic state of the complex. Consequently, the charged cationic species can then bind to the electron rich guanines of the DNA as a result of the nucleobases strong electron donating properties, where further loss of a chloride can cause intra/interstrand crosslinking and induce strong kinks in the DNA structure. These alkylations/crosslinks cause severe impairment during the damaged DNAs repair routines, provoking the observed apoptosis.[3] This mode of action has been thoroughly studied, and adaptations to the classic cisplatin model have yielded pharmaceuticals with increased activity or, as with oxaliplatin, less severe side effects of treatment. In addition, alternative metal centres of the metallodrugs have been probed for their efficacy, with an increasing number of ruthenium centred drugs having been clinically assessed in recent years, such as NAMI-A and KP1339 (Figure 1).[4] Covalent binders are of course very effective inducers of cell apoptosis, however, owing to the simplicity of the pro-drug, they act indiscriminately, often killing healthy and unhealthy cells alike. As such, increasing attention has been directed towards the development of reversible DNA binders that are less toxic to somatic cells and that assert their desired effect with more specificity.

Interestingly, DNA binders are rarely analogous to each other and often found to be wildly diverse in both structure and composition. In comparison to the morphologically elementary cisplatin, the comparably active actinomycins are far more sophisticated in architecture and are wholly organic in nature. Dactinomycin (Actinomycin D) is the most significant member of the actinomycins and the first antibiotic to have proven anti-tumour benefits.[1] The precise mode of action of dactinomycin is reasonably well understood as of late and many hypotheses have been published indicating a multitude of cytotoxic modes; however, it is widely acknowledged that the most prominent of these modes is the molecule's capacity to intercalate duplex DNA and quench the operation of

topoisomerases; similar to the consequences of cisplatin binding. Intercalation of reversible binders has garnered much research interest partly due to their often low inherent toxicities, and partly due to their structural diversity and DNA binding specificity. It is reasonable to assume however, that obvious divergences in structure would lead to diverse binding mechanisms as a direct result of steric inhibitions and electrical functionality. Appropriately then, a sound knowledge of DNA morphology is crucial to better understanding the processes that such drugs undergo when binding to DNA; so as to construct more logical pathways for drug discovery and a more informed grasp of such molecule's effects on DNAs structure and chemical behaviour.

1.2 DNA Structure

Figure 1.2 – *Molecular schematics showing the deoxyribose sugar backbone with 3'-5' phosphodiester linkage; and the Watson-Crick complementary base pairing of guanine and cytosine, and adenine and thymine.*

Watson and Crick first proposed the double helix model for DNA in 1953 which has since been seen as the standard model for double stranded DNA in its native form (B-DNA).[5] Deoxyribonucleic acid consists of two polynucleotide strands in an anti-parallel formation; these strands are formed of a deoxy-ribose sugar backbone to which each sugar is covalently linked to a nitrogenous base and bonded through a 3'-5' phosphodiester linkage, forming the DNA monomer substituent known as a nucleotide. The nitrogenous bases (nucleobases) found on the nucleotide can be one of four possible nitrogen containing heteroatomic rings; Adenine (A), Guanine (G), Cytosine (C) or Thymine (T); where

adenine and guanine are double-ringed purine derivatives and cytosine and thymine are single-ringed pyrimidine derivatives. Independent strands of DNA are held together by hydrogen bonds between complementary nucleobases, perpendicular to the strand axis, Chargaff's rules apply such that A and T pair, and conversely, G and C pair; albeit, mismatches are not prohibited, and are observed.[6] This preference for base pairing is by virtue of an additional hydrogen bond between G and C when paired and as such an observed 30 % increase in interaction strength.[7] The thermodynamic stability of DNA is predominantly determined however by the favourable stacking interactions between adjacent nucleobases; accounting for well over half of the stability of polymeric DNA, and defining the sequence dependency on stability.[8]

Deviations from the canonical Watson-Crick base pairing model in duplex DNA are possible. Spurred by structural data of a co-crystallised H-bonded A·T base pair, Hoogsteen proposed that rotation around the *N*-glycosidic bond (χ) of the adenine base provides an additional face of H-bonding capability that could pair with pyrimidine bases.[9] Now an observed and understood pairing motif, the coined Hoogsteen and reverse-Hoogsteen base pairs respectively, both contain 180° flipped purines about χ (i.e *anti* to *syn* conformer), with the latter also exhibiting a base that has rotated 180° about the helical axis (i.e asymmetric pairing). More recently it has been shown that A·T and G·C Watson-Crick bps will transiently form Hoogsteen bps with sequence specific lifetimes.[10] Since this transition modifies the morphology and chemical presentation of the DNA it can affect DNA recognition processes, repair mechanisms, and may be an interesting target site for DNA damage induction.[11,12,13]

Founded on X-Ray data of DNA fibres collected by R. Franklin, Watson and Crick correctly postulated that the hydrophobic nucleobases would additionally stack with neighbouring base pairs so as to reduce hydrophobic interaction.[14] Subsequently, a skew and a twist to the native backbone structure were seen to further reduce solvent interaction, introducing an additional hydrophobic attraction between adjacent base pairs, and ultimately the right-handed Watson-Crick B-DNA model that is so well known.[15]

1.2.1 B-DNA structural parameters

To gain an accurate insight into the steric availability of the native DNA, it would be fitting to examine and define the internal constraints and torsions within the helix itself. B-DNA adopts a right-handed double helical structure as a result of, *inter alia*, the hydrophobicity of the chiral base pairs; and, as a result of the glycosidic sugar linkages subsisting on the same face of the hydrogen bonded base pairs, two grooves of differing capacity on the helical axis are observed.[15] Known appropriately as the major and minor grooves due to the difference in width and depth, these helical

channels can accommodate ions, crucial proteins (such as transcription factors) or more relevantly to this project, DNA binding agents.[16] The minor groove runs approximately 7.5 Å deep and 5.7 Å wide whereas in comparison the major groove may be as large as 12 Å in depth and 8.8 Å wide in a B-DNA system. Of course, these grooves are also offset by the twists of the backbone, and as such are subjected to 35.9° average twist per Watson-Crick base pair, equating to 10.5 base pairs per helical rotation and a total helical pitch (full turn length) of 3.57 nm in canonical duplex B-DNA[17]
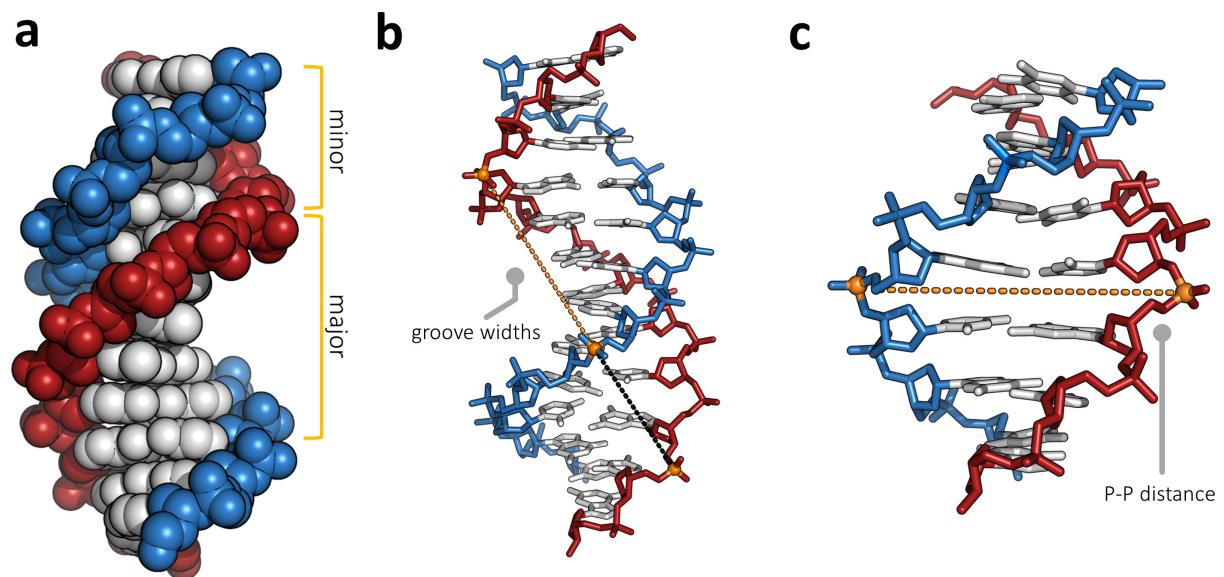
Figure 1.3 – Graphical representations of the secondary structure of DNA; (a) van der Waals (vdW) space-filling model highlighting the widths and solvent accessibility of the two asymmetric grooves of DNA; (b) groove widths are usually defined as the lowest absolute interstrand P-P distance minus the vdW radius of the phosphate groups (5.8 Å); (c) P-P vector described as the P-P distance within a base pair. Groove width and P-P separation measurements shown as orange dotted lines.

In the circumstance that differing DNA structures are to be compared, i.e. distinct base sequences or in the presence of external ligation, it is essential to have defined parameters to correlate structures. In the case of DNA it is possible to assign a set of torsion angles to conveniently portray the obliquities of the nucleotides in question. When compared with the torsions of the native structure it is an effective method in presenting distortions in a sequence.[18] The measurements required to gather these torsional angle sets for the phosphate backbone are displayed in figure 1.4a.
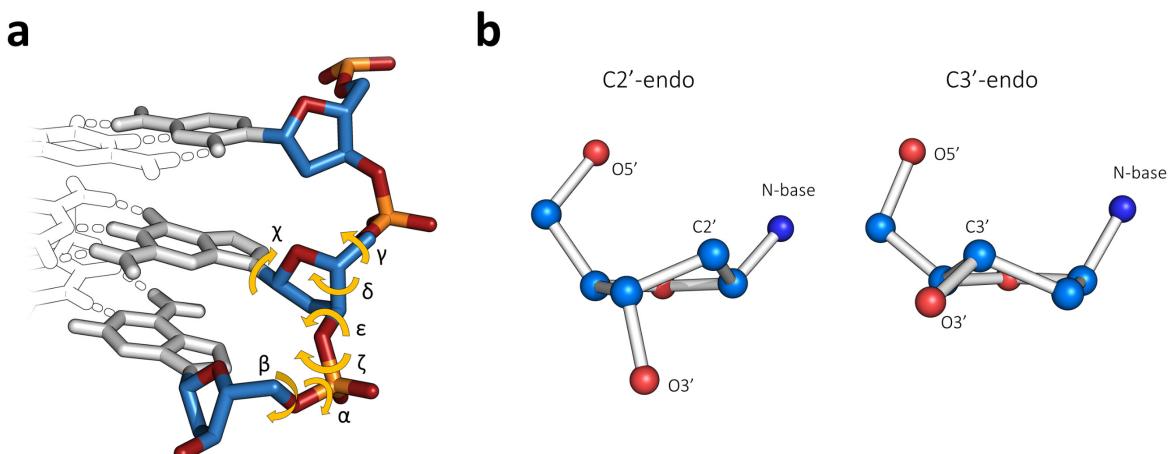
Figure 1.4 – *Molecular schematics showing; (a) defined torsional angles of the phosphodiester backbone and (b) the two most common sugar pucker conformations of deoxyribose in natural DNA.*

An important factor that affects the global structural morphology of DNA is sugar pucker. It is named as such as it describes the manner in which the five membered ring of the sugar deforms out of planarity. The pucker type characterizes the orientation of the atom bent out of plane as either *endo* or *exo* dependent on the direction the ring has warped in relation to the plane of the C5' atom. Figure 1.4b highlights the two most common deoxyribose puckering types; C2'-endo and C3'-endo. The phase angle of pseudorotation (*P*), defined by the endocyclic torsional angles of the deoxyribose (and described in figure 1.5a) can be used to assign the pucker type and this type can be used to designate the global DNA conformation.[19] Interestingly, B-DNA in its native form holds a perfectly homologous set of C2'-endo sugar puckers; however, deformations in sugar pucker provoked by extraneous factors such as the presence of bound complexes can cause base-to-base discrepancies in pucker, that can lead to hybrid global conformations. All possible sugar puckers are expressed by their pseudorotational angle in figure 1.5b; C2'-endo (140°≤*P*≤185°) and C3'-endo (-10°≤*P*≤40°) pseudorotational angles have been highlighted.[20]

$$\tan P = \frac{(\nu_2 + \nu_4) - (\nu_1 + \nu_3)}{2\nu_0(\sin 36 + \sin 72)}$$
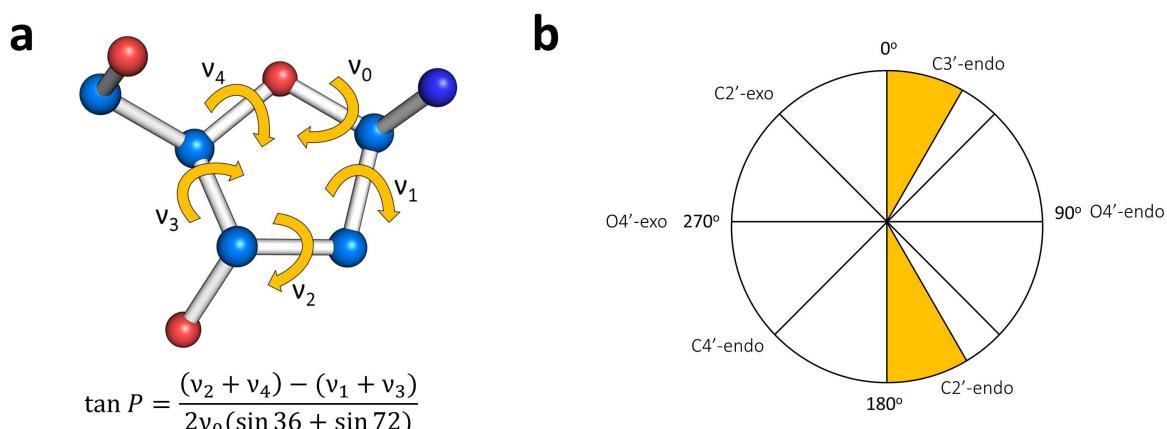
Figure 1.5 – *Molecular schematics showing; (a) defined internal torsional angles of the deoxyribose with the equation for the phase angle of pseudorotation P, and (b) the pseudorotational wheel for assignment of pucker denotation from P.*
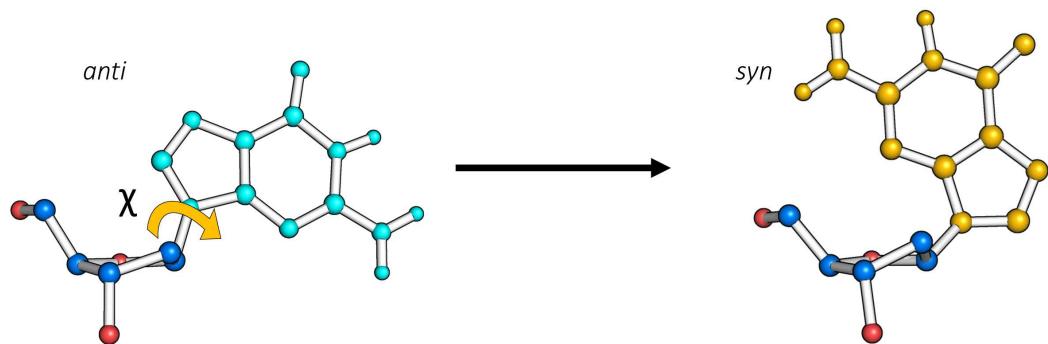
Figure 1.6 – *Molecular schematics illustrating how rotation around the N-glycosidic bond (χ) yields the less common syn-conformer of guanosine.*

1.2.2 Tertiary DNA structure

Although DNA often adopts the B-DNA conformation mentioned earlier; variations in the nucleotide sequencing, or differing salt/ionic concentrations in the surrounding solution, can have adverse effects on the fundamental tertiary structure of the polynucleotide.[21,22] Many structural variants of DNA, both natural and synthetic, have been found and examined; however, to avoid any confusion, only the biologically active A, B and Z conformations will be discussed.

A-DNA is similar to B-DNA; it exhibits the same right-handedness and accommodates major and minor grooves. A-DNA is often observed in conditions of low humidity or high salt concentrations in an attempt to protect the DNA.[23] The spatial arrangement of A-DNA however, differs primarily from B-DNA in that it has a larger median magnitude of base pairs per turn of the backbone (11 bp/turn) and a base tilt of ~20°; as such, A-DNA displays a smaller twist angle and thus a decrease of 0.9 Å in rise per base pair is also observed. Due to the collaborative effect of a tighter coil and an increased inclination angle, the sugars on the backbone pucker in a C3'-endo fashion and a large channel is observed running down the double helix; as can be seen in figure 1.7. This is similarly found in duplex RNA as a result of the axial 2'-OH groups.
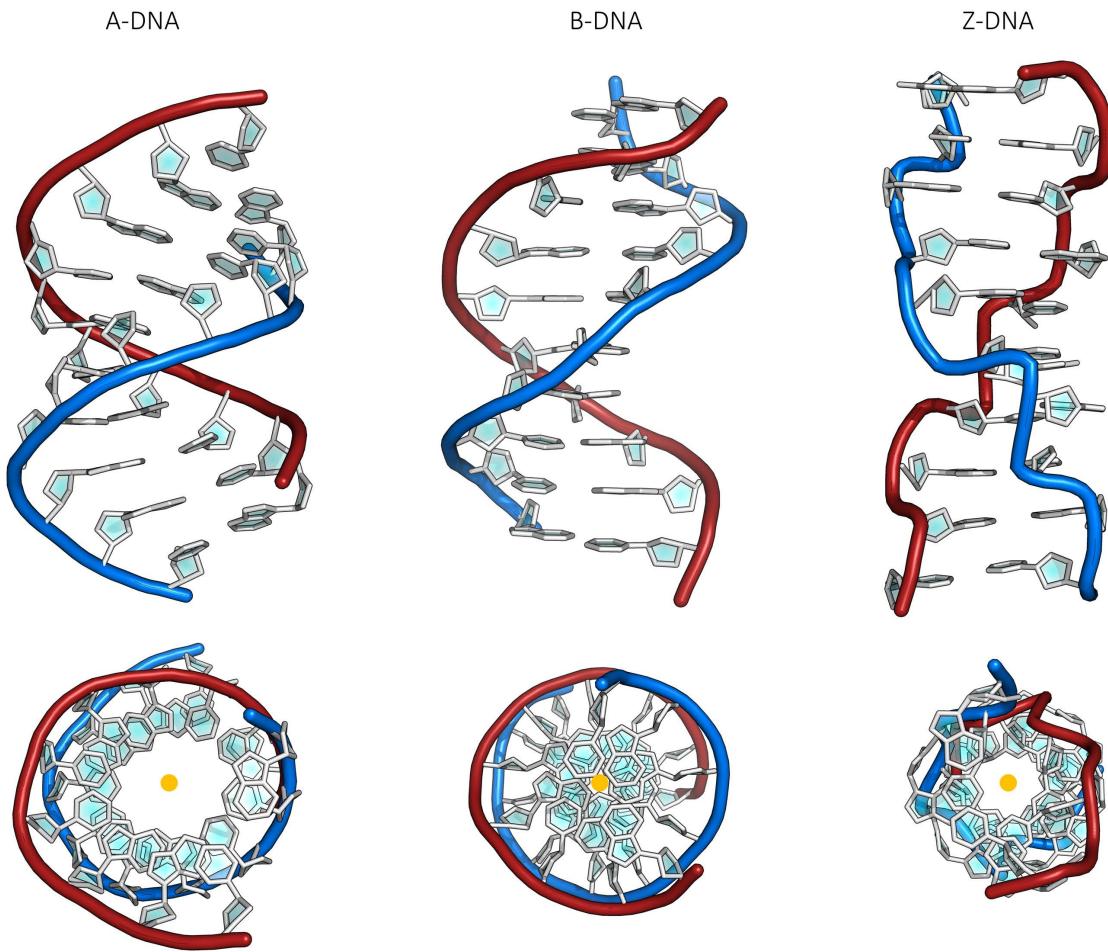
Figure 1.7 – *Crystallographic models showing the structural forms of natural duplex DNA, perpendicular to (top) and down (bottom) the helical axis; this axis is highlighted by an orange circle. Of special note is the difference in groove widths between A and B-DNA, and the large helical void observed in the comparatively lesser wound A-DNA. Phosphate backbones are coloured in blue and red with the nucleosides in white. Structural coordinates taken from 5MVT, 1BNA, and 4OCB PDB entries.*

The final biologically active duplex DNA conformation is that of Z-DNA. Unlike that of both A and B-DNA, this conformation is constituted of a left-handed double helix that is often generated transiently from B-DNA *in vitro* in certain electrostatic environments such as high ionic strength or by specific cations. Z-DNA formation is especially promoted in sequences containing repeating purine-pyrimidine residues i.e. poly-d(GC)$_2$. In this instance the guanine undergoes a base inversion due to the reversal in handedness and moves into a *syn*-conformation whereas a full sugar-base inversion is identified with the cytosine, thus keeping its *anti*-conformation.[24] As a consequence of this *anti-syn* alternation, the backbone adopts a local 'zig-zag' formation; hence the name Z-DNA. Furthermore, this consistent local structure exhibited in Z-DNA tends to form minor and major grooves of similar volumes.

Z-DNA has been shown to be an active component in transcriptional processes *in vivo*; providing torsional strain relief, in the form of negative supercoiling, downstream to gene promoter regions.[25] Although considered a normal regulatory process in healthy cells, higher expositions of this morphology have been found to increase genetic instability, often promoting mutagenesis that manifests as gene deletions and translocations. Such over-expression, and subsequent genetic modification downstream to promotor regions, has been linked with cancers such as lymphoma and leukaemia.[26] In addition, due to the base expulsion often observed at the B-to-Z-form junction, it has also been speculated that these motifs have additional susceptibility to chemical damage or enzymatic transformation.[27]

Table 1.1 – *Canonical helical parameters for different duplex DNA forms; derived from crystal structure analyses.[28]*

Helix Parameter	A-DNA	B-DNA	Z-DNA
Helix Sense	Right	Right	Left
Residues per turn	11	10.5	12
Axial Rise, Å	2.55	3.4	3.7
Helical Pitch, Å	28	36	45
Base pair tilt, °	20	-6	7
Rotation per residue, °	33	34.3	-30
Helical diameter, Å	23	20	18
Glycosidic bond configuration			
dA, dT, dC	*anti*	*anti*	*anti*
dG	*anti*	*anti*	*syn*
Sugar Pucker			
dA, dT, dC	C3'-endo	C2'-endo	C2'-endo
dG	C3'-endo	C2'-endo	C3'-endo
Intrastrand phosphate-phosphate distance, Å			
dA, dT, dC	5.9	7	7

1.3 Higher Order DNA

1.3.1 G-quadruplex structure and topology

Although initially dismissed as novel but of tentative biological relevance, the G-quadruplex has been a focal point in molecular biology and therapeutical.[29] Formed in nucleic acid sequences that are rich in guanine, such as seen at the ends of chromosomes and in gene promoter regions, G-quadruplexes are quadruple helical structures that express variable and dynamic morphology dependent on sequence and environment. Initially they were hypothesised to explain the helical aggregation of GMP, and later the self-association of short G-rich oligonucleotides.[30,31]
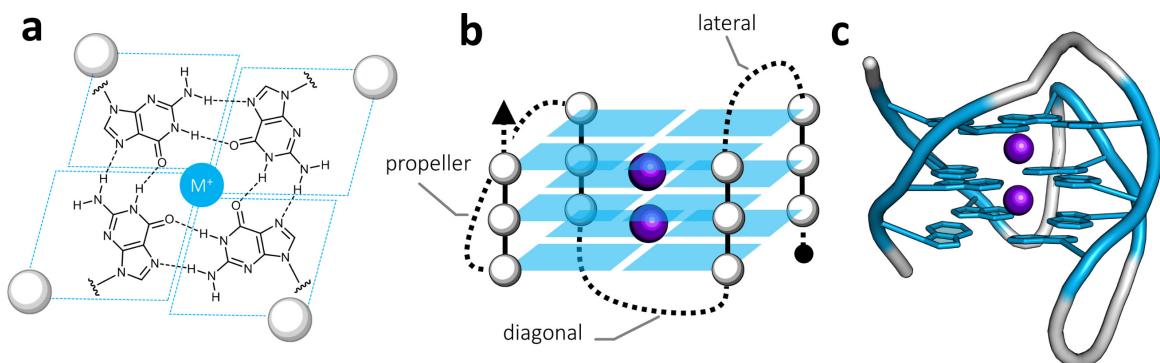
Figure 1.8 *– (a) skeletal representation of the Hoogsteen G-tetrad; (b) graphical representation of the G-quadruplex highlighting the stacking of tetrads and the three common loop varieties; and (c) crystallographic model of a G-quadruplex (PDB: 2JPZ). Guanines are coloured blue, K+ ions are coloured purple, and loop regions are white. All ribose sugars, and the loop nucleosides have been removed for clarity.*

Due to the possession of two perpendicular faces of H-bonding functionality, both with opposing polarities, guanine is able to base pair with 2 separate guanine residues. As such, four guanines associated cyclically through Hoogsteen H-bonding on both faces can form square planar constructions known as guanine tetrads (G-tetrads) (see figure 1.8a). A parallel stack of two or more of these tetrads, formed from the contiguous runs of guanines, constitutes the core of a G-quadruplex (figure 1.8b). This core is often stabilised by monovalent cations situated down the helical axis that, dependent on ionic radius and charge, either interpolate between neighbouring G-tetrads, or coordinate in the plane of the G-tetrad. These cations, most commonly K^+ or Na^+, are often coordinated, in square antiprismatic geometry (interpolated) or square planar geometry (in-plane), with the O6 atoms in each respective guanine residue. Interestingly, this cationic interaction has been demonstrated using *ab initio* methods to be more energetically supportive to quadruplex formation than H-bonding or π-stacking interactions.[32] Molecularity, in respect to the stoichiometry of nucleic acid strands, also plays a pivotal role in the stability and morphology of the folded quadruplex.
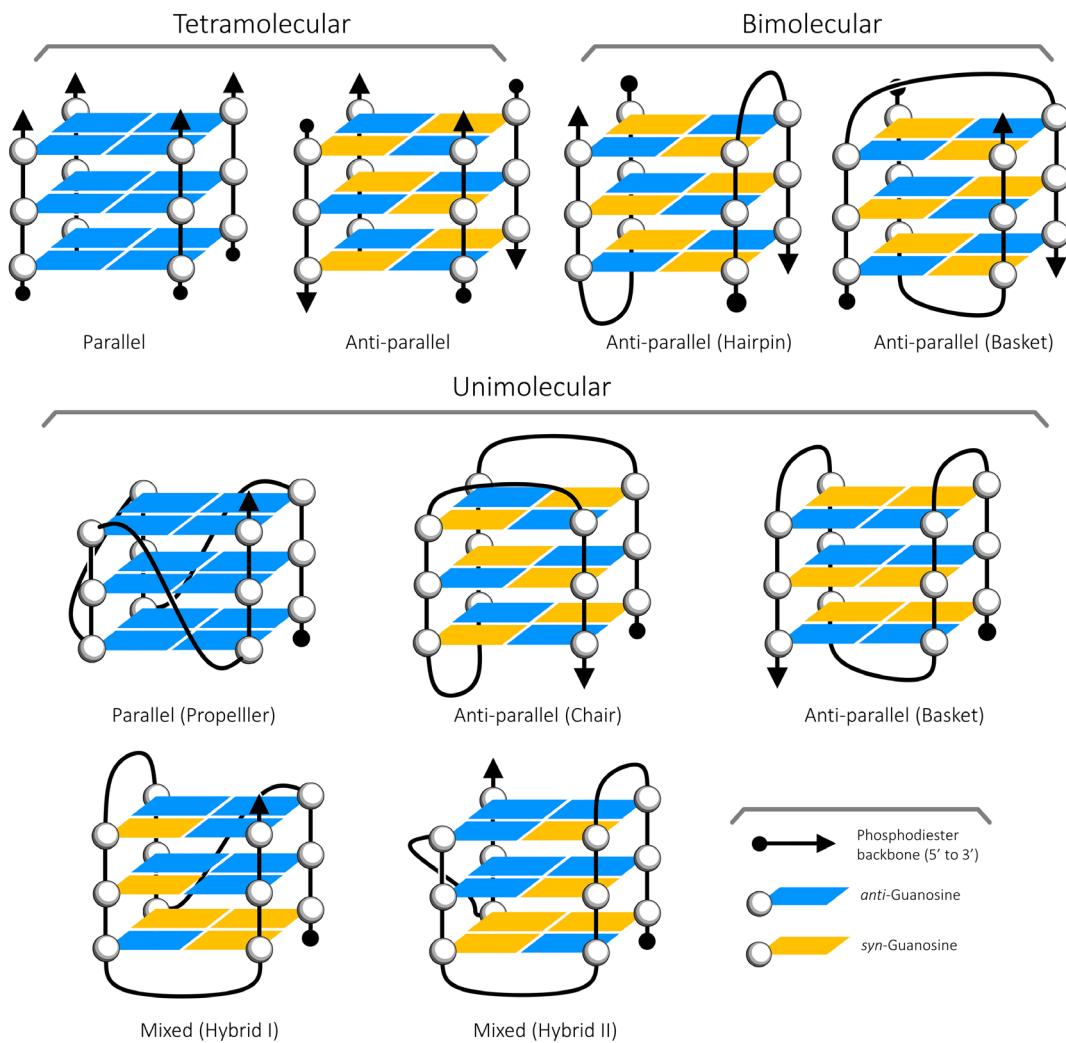
Figure 1.9 – *Graphics highlighting strand stoichiometries and the common structural topologies of a range of commonly observed G-quadruplexes.*

Biologically relevant structures contain one, two or four strands, known as unimolecular, bimolecular, and tetramolecular, respectively; and can interact intramolecularly or intermolecularly to form G-quadruplexes. Unlike duplex DNA, the strand 5'-to-3' directionalities (or polarity) are not generally constrained (to antiparallel), and collectively their orientation is useful in the topological classification of the quadruplex. If the polarities of all strands are orientated in the same direction then the quadruplex is said to be parallel in nature. In contrast, if the strands are orientated so that their polarities are opposite to their neighbours, the quadruplex is termed to be anti-parallel in nature. Naturally, deviations from this architecture are observed that contain both parallel and anti-parallel characteristics, and these are termed mixed or hybrid topologies. In quadruplex systems with less than four separate chains, short loop regions, of predominantly thymine and adenine content, link associating G-tracts. The length, composition, and nucleotide order of these loop regions necessitates energetic preference for topology alongside salt character/concentration. Loop regions

can generally adopt one of three loop characteristics; propeller, connecting adjacent G-tracts and preserving parallel polarity; lateral, connecting adjacent G-tracts but reversing polarity; and diagonal, connecting opposite G-tracts and also reversing polarity (see figure 1.8b/c).

In parallel systems, irrespective of strand stoichiometry, all guanosines adopt an *anti* *N*-glycosidic form; however, in anti-parallel and mixed systems *syn*-guanosine is prevalent, much more so than is found in canonical forms. In addition, characteristic *N*-glycosidic conformational patterns can be observed as dependent on the polarity of the strand, such that 5'-*syn-anti*-3' is the preferred base step conformation found in crystallographic data. These observations have been accounted for using MD and free energetic analyses, where it was identified that in terms of base step energetic stabilities, 5'-*syn-anti* > *anti-anti* > *anti-syn* > *syn-syn*-3'. As such, the authors concluded that anti-parallel systems will, when possible, adopt this 5'-*syn-anti*-3' repeating pattern down the G-tracts; and postulated that the large polymorphism observed in telomeric quadruplex sequences is a result of having an odd number of guanines in the G-tract, and thus a higher ratio of unfavourable 5'-*anti-syn*-3' steps.[33] Common examples of G-quadruplex topologies are illustrated in figure 1.9. Note: Due to the extensive polymorphism exhibited by G-quadruplexes, attempting to generalise tertiary structure using helical parametrics would be inadequate.

Tetramolecular species are the most elementary of topologies by structure; they are often all parallel in nature, with all guanosines in the *anti* conformation. Despite suffering from low kinetic association (fourth order in monomer), tetramolecular G-quadruplexes are studied *in vitro* due to the motif's simplicity, stiffness, and analogy to the intramolecular parallel telomeric quadruplexes.[34,35] As such they make ideal target systems for better understanding drug-DNA interactions, but lack greater biological relevance than the genomic unimolecular systems formed from single stranded DNA.

1.3.2 Biological importance

Sequence analysis of the human genome has predicted over 380,000 G-quadruplex forming regions; predominantly clustered in influential regions such as telomeres, replication origins, untranslated regions, and gene exons and introns.[36] As such, G-quadruplex formation has been implicated in numerous cellular processes; transcription, translation, recombination, and in epigenetic stability. *In vivo* evidence for the motif has been elucidated by utilising structure-specific cellular antibodies that visualise G-quadruplex locations; and by proxy through observing the regulatory effects of quadruplex binding small molecules.[37,38] As a result of its importance in the regulation of replication mechanisms, the motif has become an intensively studied target in oncological research.
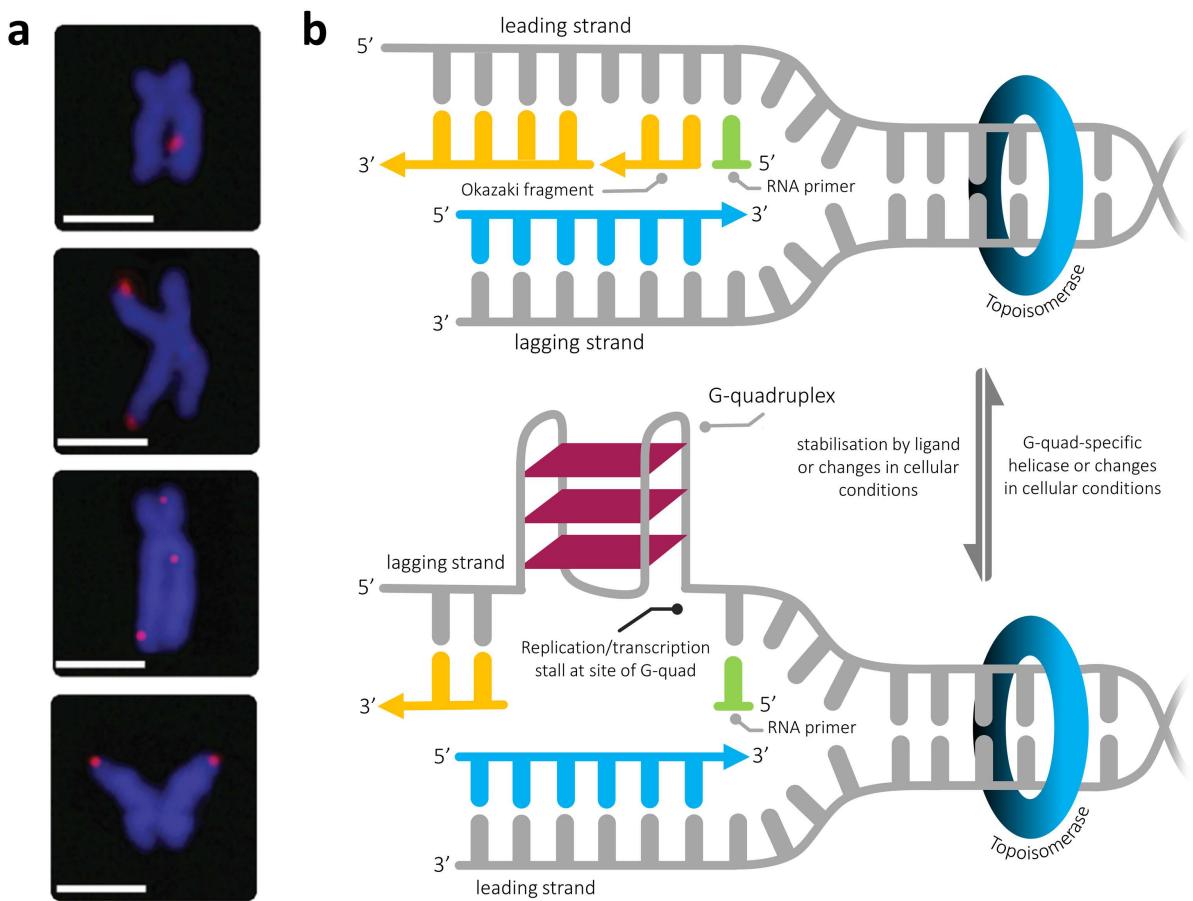
Figure 1.10 – (a) Fluorescence micrographs showing localisation of G-quadruplex structures using the immunofluorescence of BG4 (red) on metaphase chromosomes isolated from Hela cervical cancer cells. Discrete foci can be observed in both telomeric and non-telomeric regions. Chromosomes are counterstained with DAPI (blue). Scale bars represent 2.5 µm. Nuclei are counterstained with DAPI (blue). Scale bar represents 20 µm. (b) Schematics depicting the effects of G-quadruplex formation on the replication of DNA on the lagging strand. Figure 1.10a reprinted/adapted with permission from Springer Nature: Biffi, G et al. Nat. Chem. **5,** 182–186 (2013). Copyright (2013).

The telomeric regions of human chromosomes terminate with a long 3' single stranded overhang (2-20 kbases) consisting of a tandem repeating pattern of 5'-TTAGGG-3' nucleotides capable of folding into G-quadruplexes. The repeating sequence is highly polymorphic and sensitive to local conditions, such as the ionic type/strength, molecular crowding, superhelical stress and pH. The telomeres in somatic cell chromosomes act as the buffer regions of the genome that prevent recombination and gene truncation at the expense of a systematic loss of terminal non-coding bases during replication. This loss is due, in part, to the limitations of 3'-5' replication on the lagging strand of the replication fork, where Okazaki fragments cannot be synthesised without RNA primers downstream to polymerase. As such, over time, the telomeres wither and the essential genetic code is no longer protected; where further replication causes damage and mutations which lead to the natural death of the cell. As a principal determinant of cell immortality, the telomerase complex, which is over-expressed in over 80 % of hard tumours, maintains telomere homeostasis via reverse transcription, protecting cancerous cells from this natural regulatory senescence. Folded

G-quadruplexes cannot be elongated by telomerase and extensive investigations have supported the notion that Inhibition of telomerase activity by virtue of targeting and stabilising G-quadruplex formation in the telomeres can be an effective antitumour strategy.[39,40]

G-quadruplex structure formation within promoter regions of proto-oncogenes is associated with transcription repression/regulation, a result of the motif's morphology and unfavourable unwinding kinetics which effectively halt the action of polymerases. Additional stabilisation of the motif upon binding of small molecules, in proto-oncogenes such as the gene encoding tumour angiogenesis growth factor VEGF, has been shown to inhibit protein expression effectively, and ultimately lead to cell apoptosis.[41] In another study, cervical carcinoma cells (HeLa-S3) were exposed to the porphyrin TMPyP$_4$; the compound was found to inhibit the basal transcriptional activity of the *c-myc* gene promoter via ligand-mediated G-quadruplex stabilisation, validating the promotor regions function as a transcriptional repressor element and as a prospective site for gene control.[42]

Current research continues to assume that G-quadruplexes are very much a promising target for drug intervention in anti-cancer therapy, but development of small molecule clinical candidates is often decelerated due to lack of binding selectivity, or bioavailability. As a result of their morphologically dynamic but discrete folding patterns, G-quadruplexes hold an advantage over canonical DNA structures in that the design of molecular binders should, in theory, be more easily developed to provide topologically selective interactions. In practice however, complete G-quadruplex topological specificity in binding has not been reached, despite huge research efforts.[43] The first clinically evaluated therapeutic agent designed specifically for G-quadruplex targeting is quarfloxin (CX-3543).[44] Despite exhibiting impressive *in vitro* cytotoxicity in a number of carcinomic cell lines, and a marked selectivity towards nucleolin/G-quadruplex structures in the nucleolus, the prospective drug failed in phase II clinical trials against carcinoid/neuroendocrine tumours.

1.4 DNA Binding and Intercalation

As a result of the widely variable local structure, electrostatic behaviour, and polymorphism of DNA, ligands of broadly disparate disposition possess the ability to bind to, and often alter, the architecture of the DNA secondary structure.[45] When compounds that are electrostatically attracted to DNA come in close contact to the polynucleotide, dependent on the manner of the attraction, the molecules may bind to particular sites or grooves; three such small molecule binding modes are defined below.[46]

1.4.1 DNA binding modes

Groove Binding – This mode is usually preferred by small molecules with low functionality, where the molecule's attraction to the DNA is primarily through van der Waals contacts, electrostatic interactions and occasionally hydrogen bonding. Ligands that are susceptible to groove binding are more often than not, situated in the minor groove of the DNA (figure 1.11a). Common groove binders include the classes of arginines and the lexitropsins.[47,48]

Covalent Binding – In this mode, the molecule in question is required to establish covalent bonds between the ligand and the DNA (typically the N7 atoms of guanine residues) and as such, usually; but not always, causes irreversible distortions to the double helix. Such interactions often promote bending to the duplex at the site of binding, often causing kinking in the backbone (figure 1.11b). Prevalent examples of such binders include; the previously mentioned platinum complexes, such as cisplatin and oxaliplatin, and the nitrogen mustards (melphalan).[3,49,50]

Intercalation and Insertion – This mode of binding is defined as the insertion of a ligand(s), interspatially, between the planar nucleosidic base pairs of the DNA (figure 1.11c). Since the DNA is often coiled quite tightly, the insertion is usually preceded by a partial unwinding of the double helix, in which the magnitude of this uncoiling can be represented as an angle known as the degree of unwinding. This structural change tends to induce a lengthening effect on the strand and promotes twisting of the base pairs adjacent to the site which can, in turn, provoke functionality revisions that inhibit or alter biological operations.[51] Such intercalators include the anthracyclines and acridines.[52,53]
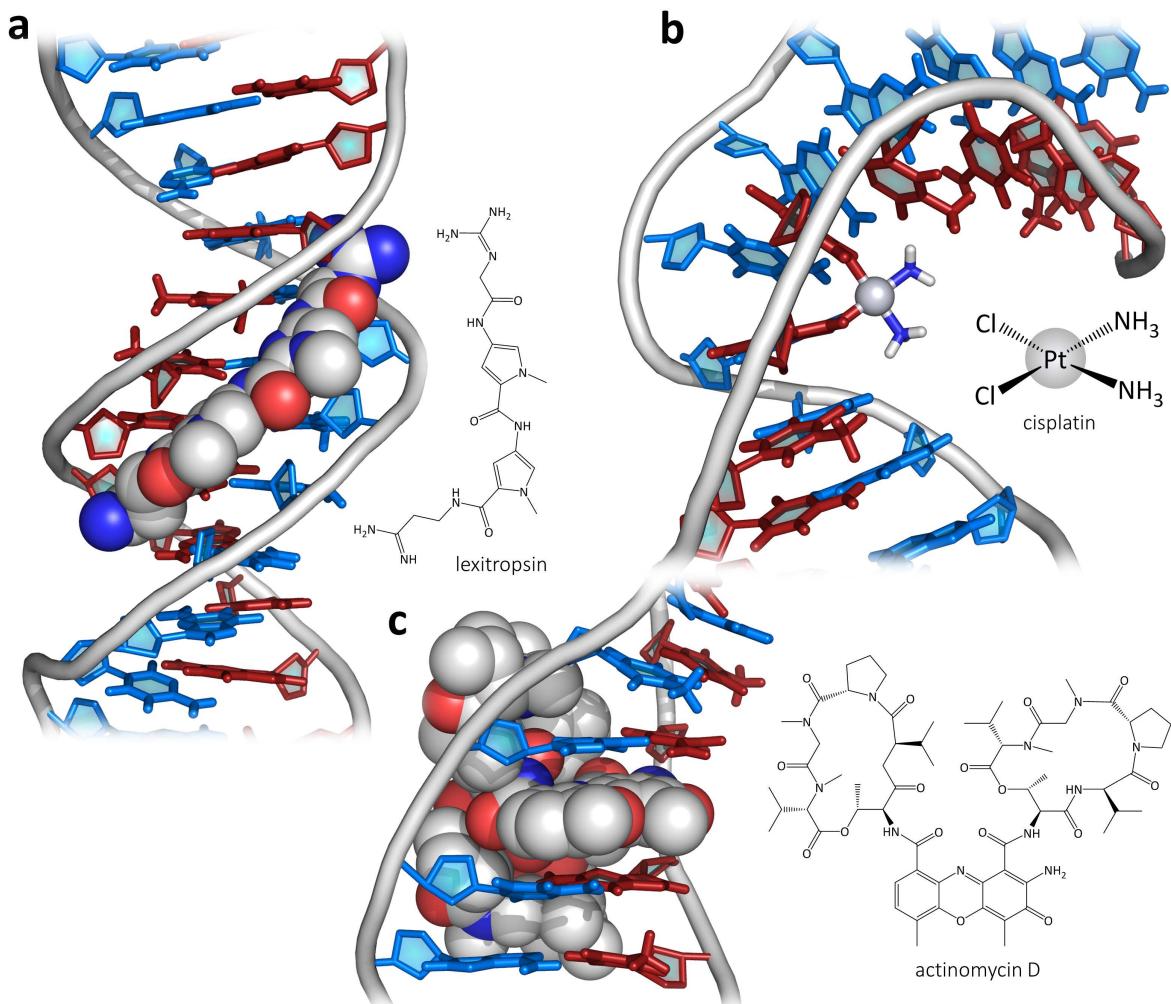
Figure 1.11 – *Crystal structures highlighting three common modes of DNA binding; (a) non-reversible covalent binding of cisplatin, (b) reversible intercalation of actinomycin D, (c) groove binding of lexitropsin.*

1.4.2 Intercalation

Originally hypothesised as the binding mode for acridines to explain the diminished length-specific mass and markedly reduced sedimentation coefficient of CT-DNA upon binding; intercalation has since been shown to be the major binding mode of a range of targeted molecules.[54] Intercalation, or the ability to 'insert' between adjacent base pairs, has been shown to induce frameshift mutations due to the intercalator poorly imitating a nucleotide and leading to polymerases skipping or adding additional bases in replication. As such many intercalators, which are often large polyaromatic cations, are known to be carcinogenic.

The ability to intercalate effectively is dependent on a slew of considerations; both the sterics and the electrostatics of the ligand and DNA play a large role in successful intercalation; where local base step geometry often dictates the binding interaction, and deviations from the standard model of insertion can occur. There are many submodes of intercalation that have been observed as a result of this interplay; the relevant intercalatory forms are described (figure 1.12).
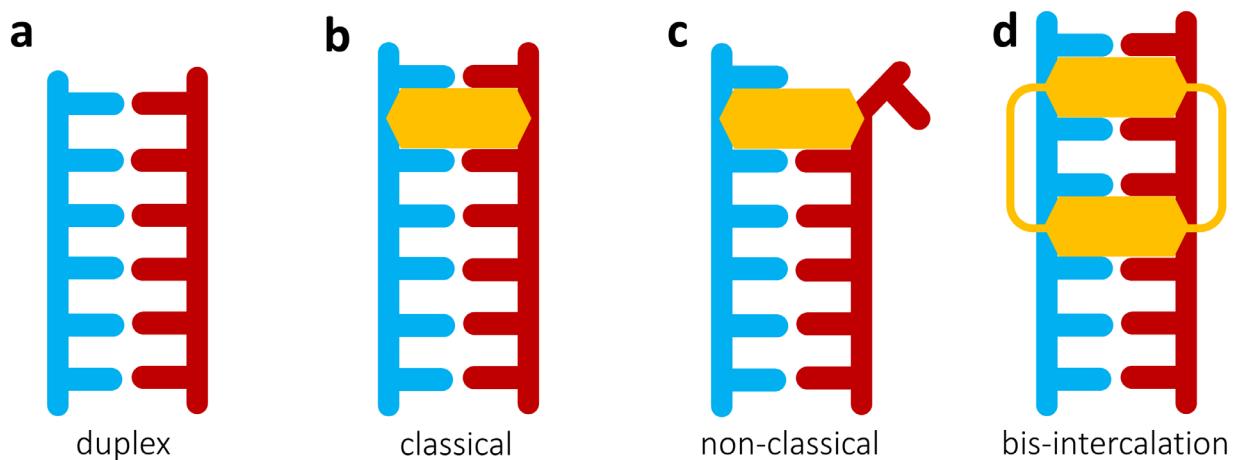
Figure 1.12 – *Simple schematical diagram highlighting the main differences between the 3 most common modes of intercalation; (a) DNA hexamer base stack, (b) 'classical' mono-intercalation, (c) 'non-classical' mono-intercalation, and (d) bis-intercalation. DNA chains are shown in blue or red, and orange blocks signify intercalating moities.*

Classical mono-intercalation, as touched upon previously, involves the insertion of a ligand into the DNA central base stack, subsequently displacing two base pairs and inducing an unwinding of the duplex. In this form, the intercalating agent may have ancillary groups that protrude into the major or minor grooves where further ligand-DNA interactions may materialize. The dislocation of the base pairs may be followed by twisting, but the base pairs must never flip. This is in contrast with the non-classical mode of intercalation, where base flipping may transpire, and, in extreme cases such as intercalation at a terminal step, base steps are permitted to flip outside of the π-stacking arrangement, reducing the sum of nucleobases per intercalation. Non-classical intercalation is especially prevalent in systems where all four nucleobases do not derive from the same duplex.

Intercalation of two separate sites from the same ligand is a viable mode of binding for some molecules and occurs when the compound in question possesses two planar intercalatory ancillary groups joined by a molecular linker such as an alkyl chain. Such systems are known as bis-intercalators and are found fairly commonly in the literature as extended organic molecules but far less so as binuclear metallic complexes.

1.5 Transition Metal Polypyridyl Complexes

Organic based intercalators constitute, by far, the largest group of DNA binders. They are relatively easy to design and modifications to said systems can be reasonably trivial; however, they often lack functionality past molecular recognition. For this reason, over the last three decades, a subsidiary field of transition metal intercalation research has been thriving. Transition metal complexes, along with offering a unique, modular approach to binder construction; are more often than not, excellent sources of rich photochemical and electrochemical properties.[55] As such, metal based intercalators hold the ability to serve many biologically significant functions, for instance, as site specific fluorescent DNA markers, electrochemical probes[56] and photoactivated DNA strand cleavers.[57]
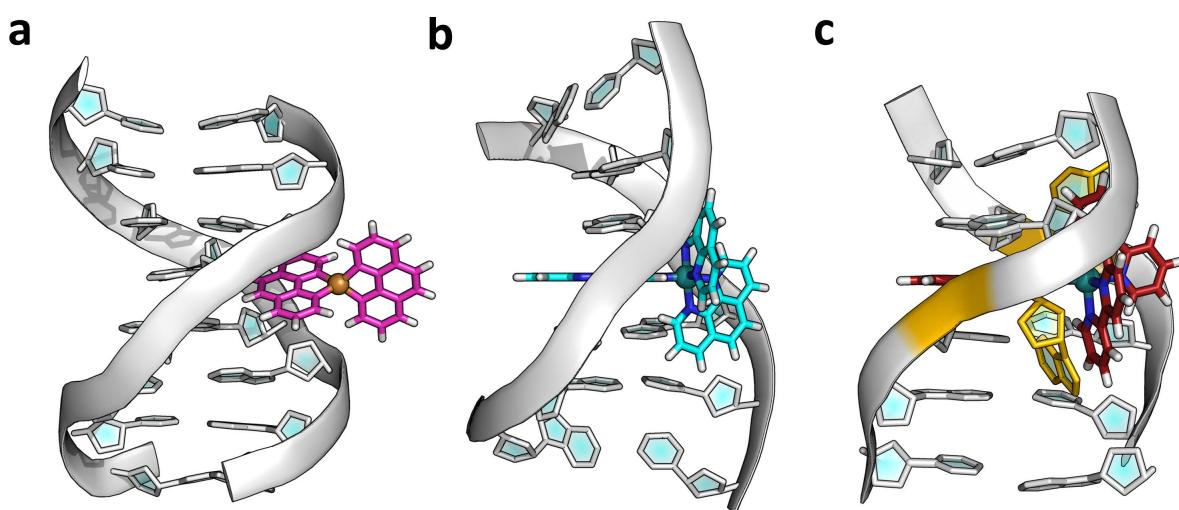
Figure 1.13 *– Models representing three non-covalent modes of metalloorganic DNA binders; (a) groove binding, (b) intercalation, and (c) metallo-insertion. Orange colouring denotes a flipped-out residue.*

Often comprised of metal centres of stiff d^6 octahedral or d^8 square planar geometry and frequently containing planar aromatic ligands, these complexes are attractive intercalators since they are often effectively substitutionally and kinetically inert in biological environments.[58] This rigidity in structure bodes well for the complexes, since large vibrational variations in the structure could reduce the selectivity of the binding. Additionally, modularity in DNA interaction properties can be achieved through the permutation of ancillary ligands and/or functional groups, effecting an option of a dynamic/adjustable chemical environment. Such metal centred complexes often also benefit from substantial photoelectric potential and as such, frequently exhibit intense metal-to-ligand charge transfer (MLCT) that can facilitate luminescence and/or DNA damage responses. Finally, just as a DNA duplex itself exhibits chirality, so can asymmetrical octahedral metallic complexes; such compounds are significant, since this chirality has been shown to induce further binding specificity.[59]

1.5.1 Ruthenium polypyridyl intercalators

Although many transition metal centres have been probed for their biological and photo-activities in the presence of DNA, a large portion of the research effort has been concentrated on the group 8 metals and more specifically ruthenium-based compounds. Ruthenium centred complexes were initially developed as alternatives for platinum-based medicines as a result of their functionalised DNA binding properties, antineoplastic activity, and inherently low cytotoxicity.

Ruthenium(II) polypyridyl complexes are not new, they have been the focal point of much investigation since the late 1950s when the room-temperature luminescent property of $[Ru(bpy)_3]^{2+}$ in aprotic solvents was reported.[60] Although these compounds eventually became the subject of detailed research, it was not for another two decades that their photophysical properties were investigated in the presence of DNA.
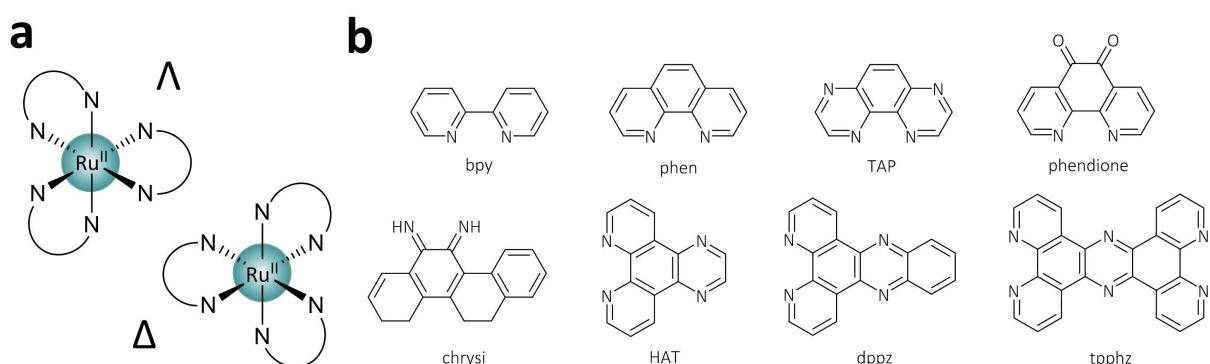
Figure 1.14 – *(a) Enantiomeric definition of fully coordinated ruthenium(II) complexes with bidentate ligands; (b) molecular structures of a small range of common bidentate ancillary ligands found on investigated ruthenium polypyridyl complexes.*

The homoleptic compound $[Ru(phen)_3]^{2+}$ was first isolated as enantiomerically pure isomers before being proven as an effective DNA binder.[61] It was ascertained; using spectroscopic, hydrodynamic, and binding parameter studies, that the complex bound to DNA by a mode of intercalation, kinking the duplex backbone as first postulated. Hypochromicity data showing a 17 % decrease in MLCT intensity, coupled with a large increase in luminescence confirmed the tethering of the molecule to the poly d(GC) duplex. Furthermore, the team noted that for any particular Ru concentration, a greater luminescence is exhibited for the Δ enantiomer when compared to the Λ; thus, when combined with the experimental excited state lifetimes of the stereoisomers, it was suggested that the Δ bound to the duplex with more affinity than the Λ.[61] Further experimentation elucidated that the Δ-$[Ru(phen)_3]^{2+}$, due to the right handed propeller twist, better conformed to the duplex and allowed the ancillary ligands to fit into the major groove (figure 1.15); moreover, this binding also unwound the double helix by 22° supporting an intercalative binding mode.[61] Later studies utilizing equilibrium binding, viscosity measurements, and competition dialysis, contradicted

this characterisation of binding mode.[62,63] By application of polyelectrolytic theory the experiments concluded that both enantiomers essentially electrostatically bound to dsDNA through the minor groove, and exhibited positive exotherms implying modes driven by entropics. 2D-NOESY and simulated docking studies following this both separately indicated similar 'non-classical' modes, with the latter suggesting an additional pseudo-intercalated mode through the major groove.[64,65]

a

Δ

b

Λ
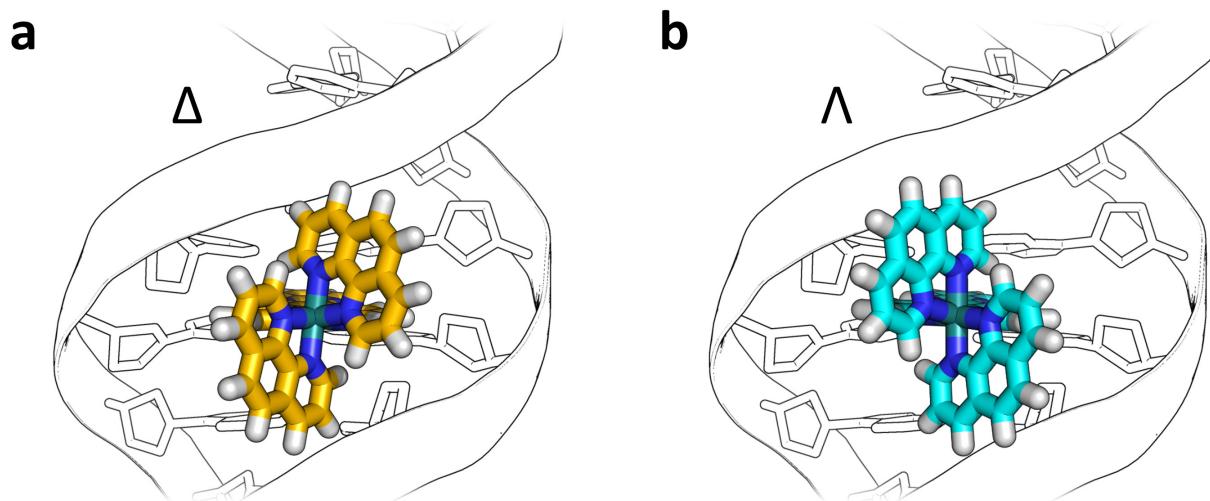
Figure 1.15 – *Graphical model highlighting the steric fit differences between the two enantiomers of a ruthenium polypyridyl complex symmetrically intercalated into a DNA duplex.*

Extensive analysis of homoleptic *tris*-polypyridyl compounds as ligands for ruthenium based intercalators continued, and $[Ru(TAP)_3]^{2+}$ was soon demonstrated as a prime candidate. The investigation inquiring into the TAP ligands role in DNA interaction found that the TAP ligand could bind efficiently and even acted as a sensitizer for photoactivity induced cleavage of a poly d(GC) duplex sequence.[66] Perhaps, equally exciting was the fact that, unlike the $[Ru(bpy)_3]^{2+}$ and $[Ru(phen)_3]^{2+}$, inherent luminescence was quenched on binding with DNA. It was later found that bis-TAP bound complexes with a second polypyridyl ligand (heteroleptic) yielded novel interactions. Subsequent investigation on heteroleptic complexes showed that $[Ru(TAP)_2(bpy)]^{2+}$ could form covalent bonds with guanine residues on the duplex DNA as a result of a photoinduced electron transfer, effectively forming a photo-adduct and a novel mode of covalent attachment.[67]

1.5.1.1 DNA 'light-switch'

Findings such as there above, sparked further interest, and soon, more photophysically sophisticated ruthenium complexes were investigated. $[Ru(bpy)_2(dppz)]^{2+}$ was a first of its kind; coined as a molecular 'light switch' it was discovered that the complex showed little-to-no room temperature photoluminescence in aqueous media, but, when bound to calf thymus DNA, presented an increase of four magnitudes in luminescence.[68] This intriguing discovery led to the compound being hailed as an

excellent alternative to radioactive DNA probes; however, the absolute rationale behind the mode of the quenching was unclear. Further investigation found that the luminescent properties of the complex were left intact whilst suspended in apolar solvents; this also explained the increase in luminosity when bound to DNA, since whilst enveloped by the organic based duplex, water molecules could not interact with the intercalated dppz moiety.[68] It was later found, using phosphorescence spectroscopy to elucidate luminescent lifetimes, that the quenching of luminescence in protic media could be attributed to an increase in the extent of equatorial hydrogen bonding on the phenazine rings.[69]
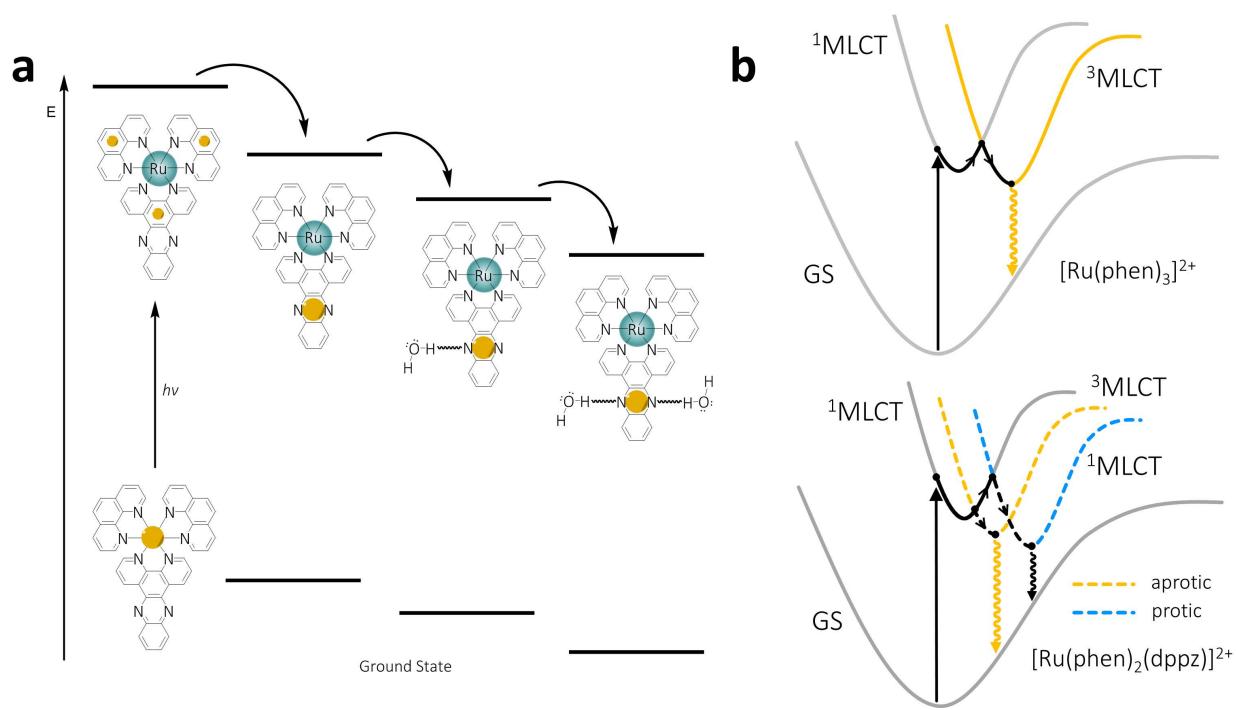
Figure 1.16 – *Graphical representation of the proposed excited state relaxation pathway of [Ru(phen)$_2$(dppz)]$^{2+}$ in the presence of protic media. Blue circles indicate regions of high quantum electron probability (b) Simplified potential energy landscapes of [Ru(phen)$_3$]$^{2+}$ and [Ru(phen)$_2$(dppz)]$^{2+}$. Wavy lines represent relaxation to ground state, where orange denotes a radiative pathway. Figure 1.16a is a reproduction of a figure from reference 69 - Olofsson, J., Önfelt, B. & Lincoln, P., J. Phys. Chem. A 108, 4391–4398 (2004).*

Upon excitation, electron transfer from the central ruthenium to the π-deficient dppz ligand occurs. This ^{1}MLCT state rapidly transposes to a ^{3}MLCT excited state, located primarily on the phenazine moiety of dppz (π*), via intersystem crossing (ISC). Relaxation via phosphorescence is vibrationally quenched in the presence of water due to H-bonding with the phenazine nitrogens but is preserved upon intercalation into DNA. This observation was especially valuable since it further supports the notion of disparate binding modes between the Δ or Λ enantiomers and DNA; it has also been shown that despite similar binding affinities, the stereoisomers exhibit very different quantum yields (almost an order of magnitude higher for the Δ).[70] With such knowledge, it is not difficult to see that, if the Λ complex did not penetrate as deeply into the duplex as the Δ, then more of the complex

would be accessible to the aqueous solvent, thus forming more hydrogen bonding and hence a higher fraction of relaxation through non-radiative pathways. The precise energetic rationale for the 'light-switch' is however not without dispute, with more recent developments implying that the solvent and π-stacking environment around the ancillary ligands can play a part as well as the envelopment of the intercalating ligand.

| MO-phen | MO-dppz(phen) | MO-dppz(phz) |
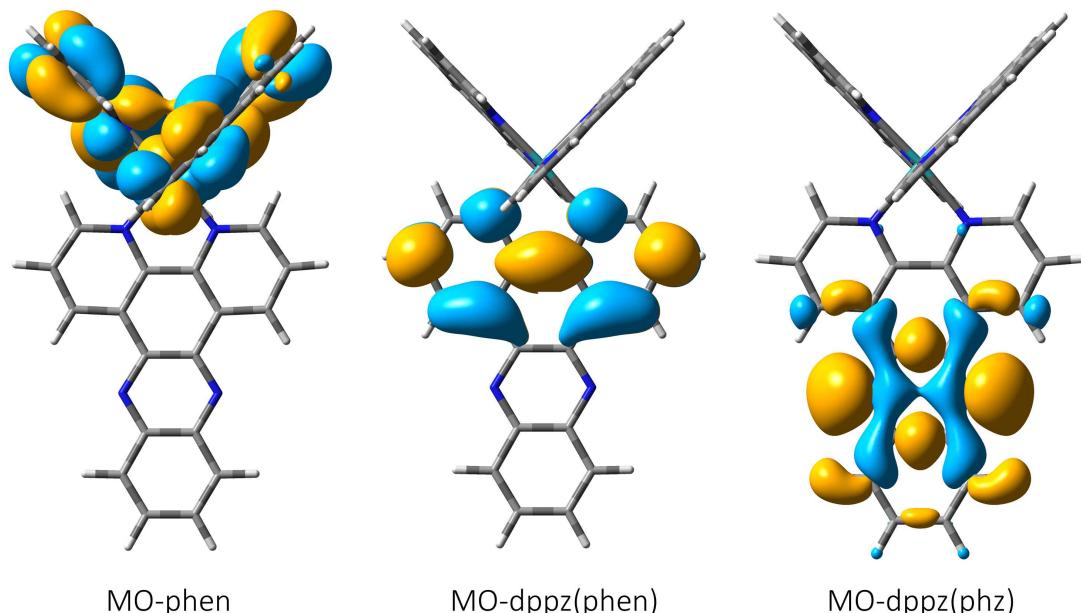
Figure 1.17 – *Electronic molecular orbitals calculated at the DFT level that depict different general excited state localisations around [Ru(phen)$_2$(dppz)]$^{2+}$. phen and dppz(phen) localised excited states result in radiative relaxation pathways whereas dppz(phz) localised states are non-radiative.*

Perhaps one of the more interesting aspects of this field of study, is the effect of seemingly small alterations in ancillary ligand structure on the photophysical behaviour of the complexes and the subsequent behavioural changes when bound to DNA. This was briefly touched upon earlier when it was shown that the luminescent properties of a tris-polypyridyl based ruthenium system could be reversed from quenched in the presence of DNA, to amplified when bound to DNA, by changing the ancillary ligands from TAP to phen.

As a result of the burgeoning interest in the area, [Ru(TAP)$_2$(dppz)]$^{2+}$ was synthesized and consequently investigated for its binding affinities and expected powerful photophysical performance.[71] The complex's lowest excited state was found to be strongly luminescent, even in the vicinity of protic media; furthermore, employing laser flash photolysis data, the TAP ligands were found to be formidably oxidising, and as such can oxidise guanine-containing polynucleotides such as poly d(GC)$_2$ or calf thymus DNA. Interestingly, it was also established that when using guanosine-5'-monophosphate (GMP), in a deuterated aqueous solution as a means of quenching luminescence, the room temperature rate constant was nearly half of the similarly obtained H$_2$O data.

This finding, along with spectroelectrochemical correlation data, suggests that the quenching of the triplet MLCT state occurs through a proton-coupled electron transfer process involving a TAP ligand and a local guanine nucleobase. Such conclusions are in stark contrast to the proven MLCT states of $[Ru(phen)_2(dppz)]^{2+}$, whose excited triplet state exhibit charge transfer to the dppz ligand.[71]

Seemingly small alterations in the electronic states of the coordination sphere can clearly have large implications on the excited state photophysics, and thus potential medicinal application, of the complex. In recent years, many iterative studies have built on this, developing ruthenium-based drug candidates with perspective application as anti-microbials, *in vivo* structural probes, and PDT agents.[58,72,73]

1.5.2 Binding modes

Extensive research into the exact binding modes of ruthenium polypyridyl complexes has been described in the literature; however, with some dispute as to the explicit intercalatory mode and ancillary ligand positions. One of the recurring issues within this field is the dependence of spectroscopic and solution NMR studies in assigning binding locations; and while being strong techniques, there can sometimes be contrasting interpretations, especially in systems where multiple binding geometries may be observed. Since many of these techniques eventually rely on a direct comparison of titration data of known groove and intercalative binders, which are themselves usually not completely sequence specific, it can be very challenging to determine with any real conviction, the precise binding location of the intercalator. This notion of uncertainty is assuredly valid and, historically, a cause of concern. Before the emission anisotropy and flow linear dichroism measurements were presented it was reasonably assumed that $[Ru(bpy)_3]^{2+}$ and $[Ru(phen)_3]^{2+}$ were both intercalators. However, after observing a growth in DNA length after addition of $[Ru(phen)_2(dppz)]^{2+}$ but only observing a helical kink after inclusion of $[Ru(phen)_3]^{2+}$, it was postulated that there may be additional manners in which, dependent on the ancillary ligands, the complexes could intercalate into the DNA duplex.[74] Such variations on the 'classical' intercalation are now better known, but nonetheless, it is a reminder that where possible, spectroscopic measurements should be supported by structural data.

A full understanding of the intricacies of binding are necessary when interpreting photophysical data; of particular contention has been the major/minor groove distinction. Initially it was proposed that $[Ru(L)_2(dppz)]^{2+}$ complexes intercalated into duplex DNA via the major groove since by measurement the minor groove P-P distance is too small to accommodate deep intercalation. Spectroscopic displacement assays and NMR NOE measurements have been in stark contrast, suggesting both major and minor groove binding in solution.[75–79] This contention persisted until the

emergence of crystallographic evidence showing the intercalation of Λ-[Ru(TAP)$_2$(dppz)]$^{2+}$ to the decamer d(TCGGCGCCGA).[80] The structure, and subsequent structures, have consequently shown that such complexes bind to well-matched sequences through the minor groove and can bind with enantiomeric, and sequence specificity.
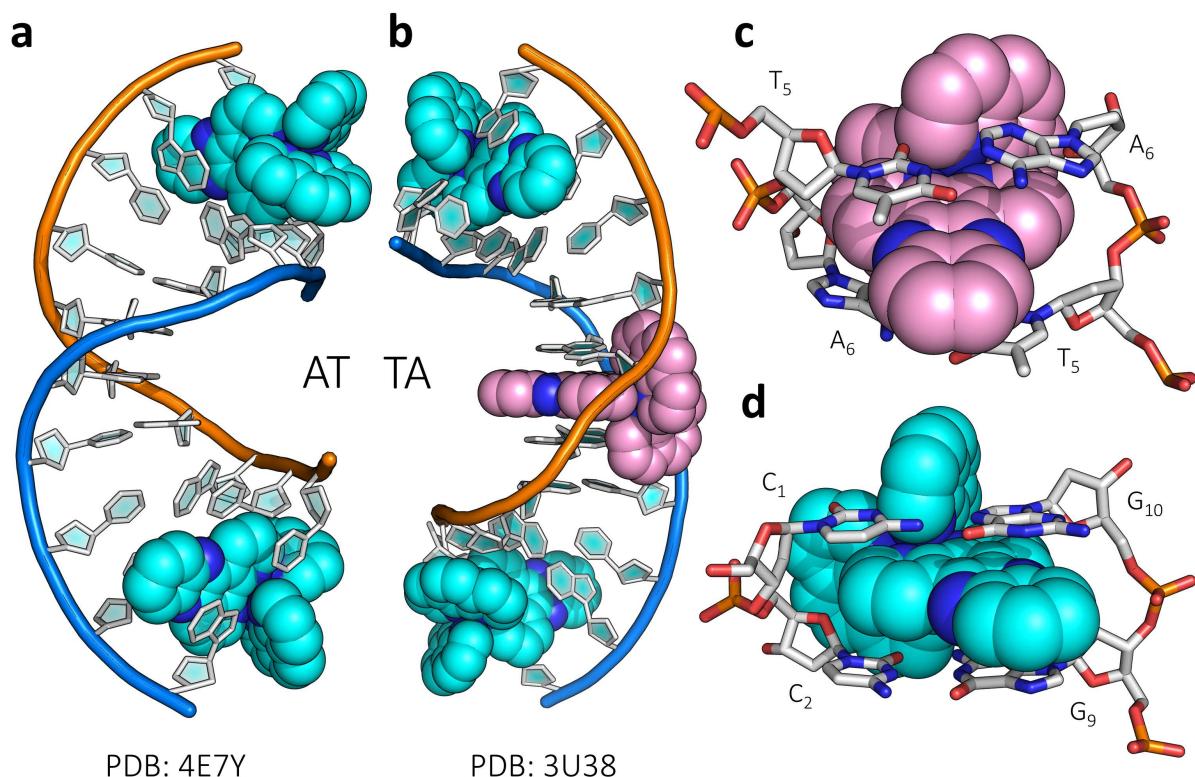
Figure 1.18 – *Crystal structures of Λ-[Ru(phen)$_2$(dppz)]$^{2+}$ bound to (a) d(CCGGATCCGG); and (b) d(CCGGTACCGG), highlighting how intercalation into the central step is only observed the base step is TA/TA View from the major groove perpendicular to the P-P plane showing the different intercalative angles exhibited by the complex to the discrete base steps in 3U38. Complexes bound in a 'canted' mode are shown in cyan; complexes symmetrically bound are shown in light pink.*

Structural investigations of this class of complexes have since shown many differing intercalative modes and novelties. Sequence specificity of binding was observed in structures of Λ-[Ru(phen)$_2$(dppz)]$^{2+}$ bound to the decamer sequences d(CCGGATCCGG) and d(CCGGTACCGG) (figure 1.18).[59] In both cases, shallow angled intercalation is observed into the terminal CC/GG base steps, whilst the ancillary phen ligands semi-intercalate into a GG step of an adjacent duplex. The orientation of the chromophore is directed by a secondary π-stacking interaction between an ancillary ligand and a neighbouring ribose; a clear additional specificity from using octahedral salts and one which will be dependent on ancillary ligand type, optical isomerism, ribose pucker, and DNA/RNA target. Interestingly however, binding at the central step is only observed into the TA/TA step and not into the AT/AT. Intercalation at this step forms a symmetrical binding cavity as a result of an absence of steric interaction between the ancillary ligands and DNA, which in comparison to the angled 'canted'

mode allows the complex to penetrate more deeply, exposing more of the phenazine region of the dppz into the major groove. Such disparities in binding have been used to explain the observation of biexponential decay of two different emission lifetimes for the enantiomer when in the presence of synthetic homogenous polynucleotides.[81] Perpendicular intercalation allows for a deeper penetration of the complex but in doing so exposes the ligand to polar solvent which in turn is expected to reduce the lifetime and quantum yield of the radiative excited state.

Intriguingly, despite the delta isomers consistently yielding stronger binding affinities to canonical DNA than the lambda counterparts, in most crystallographic studies to date the lambda has preferentially crystallised even though most are grown from racemic mixtures. Despite this, a few structures have been used to aid in the description of the delta binding modes. Hall *et al.* used these structures to propose five discrete binding models of the delta isomer and then utilised these structural insights to explain the differences in luminescence lifetimes between the models.[82] Relating the angle of insertion to the degree of chromophore encapsulation and phenazine nitrogen exposure, the authors postulate that the luminescence lifetimes would follow the following orientation dependent pattern: mismatch (complete flipping of base pair) > well-matched 'non-classical' intercalation (non CG pair followed by flipping out of a single base) > canted intercalation > symmetrical intercalation > semi-intercalation.

1.5.3 Therapeutic potential of ruthenium polypyridyl complexes

1.5.3.1 antineoplastic activity

Since the development of *cis*platin and its subsequent derivatives for chemotherapy, the exploration of metal-based DNA binders with prospective diagnostic and remedial utility has grown in importance. Ruthenium centred metallo-pharmaceuticals such as NAMI-A imidazolium and KP1019 indazolium for example have been clinically assessed as a result of their promising antineoplastic activity.[4] However, despite their advantageous redox capabilities and inherently lower toxicity, these complexes are not as cytotoxic as the platinum drugs and suffer from similar non-specificity due to their often similar ligand-exchange mediated modes of action. Kinetically inert Ru(II) complexes with saturated coordination spheres have since been examined as target specific DNA agents due in part to their structural diversity and high DNA binding affinities.

Despite its known DNA affinity and characterised intercalation properties, $[Ru(bpy)_2(dppz)]^{2+}$ has little to no cytotoxic behaviour towards a range of cancerous cell lines.[68] Not too structurally dissimilar complexes however, such as *rac*-$[Ru(bpy)_2(pztp)]^{2+}$ have conversely been shown to express cytotoxicities that rival those of established pharmaceuticals.[83] Interestingly, it was shown that

changing the ancillary co-ligands from bpy to phen led to observed IC$_{50}$ values over a magnitude larger in all tested lines; an observation that certainly highlights the ambiguities of efficacy prediction. In another study the antineoplastic effect of the presence of *rac*-[Ru(bpy/phen)$_2$(ρ-tFMPIP)]$^{2+}$ (figure 1.19) was investigated on MDA-MB-231, a breast carcinoma cell line.[84] The phenanthroline complexes outperformed cisplatin, exhibiting an IC$_{50}$ value less than half of that of the covalent binder, and were subsequently shown to be inherently non-toxic to somatic cells. In this case however, the bipyridine complex is five times less effective than the phenanthroline derivative. Further study implicates the prospective drug in modes of action regarding *c-myc* oncogene suppression via G-quadruplex stabilisation by groove binding.[84]
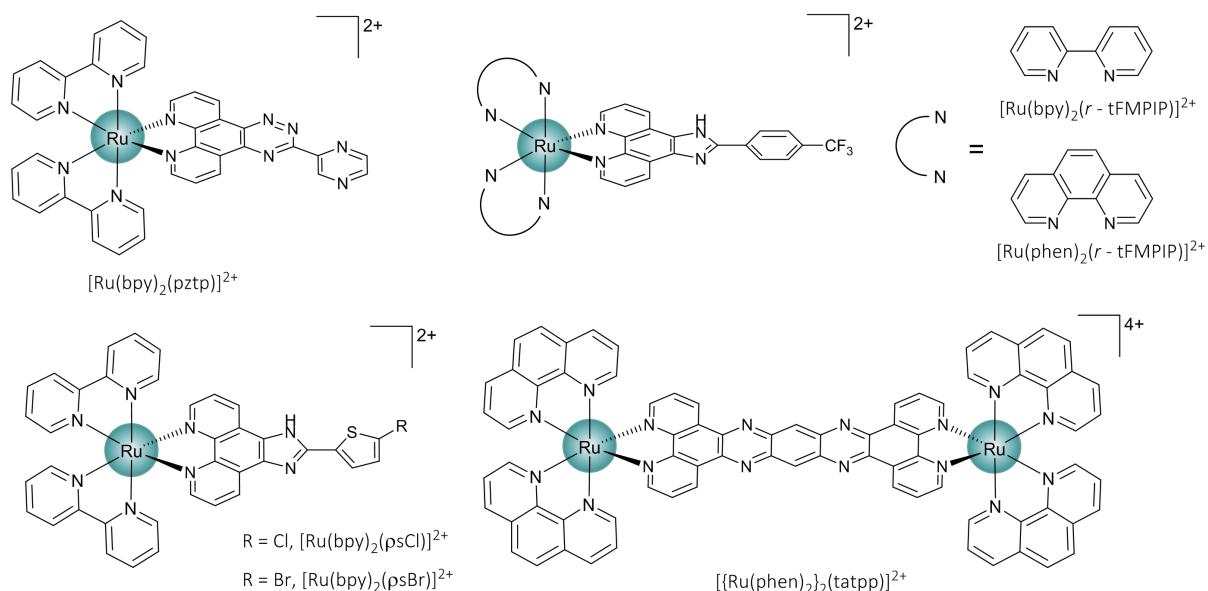
Figure 1.19 – *Structures of mono and dinuclear Ru(II) polypyridyls that exhibit antineoplastic behaviour.*

As echoed before, comparatively few investigations have studied enantiomeric differences to the same extent. One such study examined the complexes Λ/Δ-[Ru(bpy)$_2$(psCl/psBr)]$^{2+}$ (figure 1.19) against a range of metastatic cell lines in comparison to cisplatin and NAMI-A.[85] In all cases the delta isomer was more active than the lambda and exhibited a 10-fold improvement on NAMI-A on HeLa cell proliferation; however, in all lines cisplatin was more effective. This enantiomeric difference correlated well with cellular uptake measurements and nucleus accumulation. DNA religation assays revealed topoisomerase I and II inhibition activity as a major mode for the induced apoptosis; it may be postulated that the tighter binding of delta seen in previous structural investigations could be a reason.

Dinuclear complexes have also been probed for their antineoplastic activity, and in general exhibited promise often surpassing that of mononuclear variants. A well examined example is that of

[{Ru(phen)$_2$}$_2$(tatpp)]$^{4+}$ (figure 1.19). The extended complex has been shown to be a strong DNA binder and potent cytotoxin whose activity is amplified in the anaerobic conditions of malignant cells.[86] Further studies described DNA cleavage mechanisms of the complex involving intermediary radical carbon species, and compelling cancer cell specificity that translated to significant *in vivo* tumour growth suppression in mice xenograft models.[87]

1.5.3.2 biological probes

Fully coordinated luminescent Ru(II) polypyridyl complexes have garnered significant interest as probes for *in cellulo* structure, due in part to their spectroscopically active centres and their tunability as DNA binders. Specifically, MLCT-originated light-switching systems offer advantages such as: low excitation energies (i.e. visible wavelengths), high luminescence quantum yields, long phosphorescence lifetimes, sensitivity to local binding environment, large Stokes' shifts, and high photostabilities. Such properties have been exploited in the development of cellular imaging agents that stain nucleic acid structures; visualising critical biological processes such as transcription and translation, and giving a route to spectrophotometric assessment of cellular uptake and intracellular localisation.[72]

The uptake efficiency, that is, the ability to cross the phospholipid bilayer membrane of the cell, has been shown to be correlated to the hydrophobicity of the ancillary ligands in such complexes. Using confocal microscopy and flow cytometry, complexes such as the highly emissive *rac*-[Ru(DIP)$_2$(dppz)]$^{2+}$ have been shown to passively diffuse and internalise in HeLa cervical carcinoma cells (figure 1.20a).[88] Interestingly, its MLCT emission fingerprint implies full localisation in the cytoplasm with no accumulation from within the nuclear envelope. Indeed this is not always the case and complexes such as *rac*-[Ru(phen)$_2$(11,12-bis-ethoxymethyl-dppz)]$^{2+}$ have been shown to accumulate in and stain the nucleoplasm of permeabilised cells (figure 1.20).[89] In the same study the authors investigated the effect of increasing hydrophobicity on the localisation and concluded that the more hydrophilic derivatives were poorly transported through the membrane and tend to stain the nucleus, whereas the more hydrophobic variants tended towards punctate staining of the cytoplasm. This leads to an interesting dilemma in the field where high hydrophobicity allows for membrane permeation but often leads to localisation outside of the target organelle whereas conversely hydrophilicity prevents poor passive permeation but improved nuclear localisation. These views may be a correlative rule of thumb but is not a discrete rule; as evidenced by the work of Thomas *et al* on tpphz containing ruthenium complexes.[90] Here they have shown that the dinuclear [{Ru(phen)$_2$}$_2$(tpphz)]$^{4+}$ is actively transported into the cells of the MCF-7 breast carcinoma line and rapidly localises in the nucleus despite the hydrophilic nature of the cation. Subsequent studies have

confirmed a similar *in cellulo* fate for the mononuclear structural analogue [Ru(phen)$_2$(tpphz)]$^{2+}$; where this complex exhibits a higher inherent cytotoxicity which is postulated to be a result of the increased nuclear accumulation.[91] Past investigations from the group have proven the complexes' selective enantiomeric affinity and dramatically increased luminescence response for G-quadruplexes (x150 in the presence of anti-parallel topology) over B-DNA helices.[92] Such interactions with higher order nuclear DNA have been demonstrated to be responsible for the observed blue-shifted emission in the *in cellulo* response; exhibiting a response from one chromophore that can be used to differentiate DNA structures *in vivo*. Structurally discriminating luminescent complexes have since been intensively researched and examples such as [Ru(bpy)$_2$(dppz-idzo)]$^{2+}$ have been shown to be compelling *in vitro* probes for DNA tertiary structure. The complex exhibits a very intense 'light-switch' (~5 times greater than [Ru(bpy)$_2$(dppz)]$^{2+}$) in the presence of G-quadruplexes; with marked emission in the presence of hybrid topologies.[93] More recently, a systematic study investigated the effects of enantiospecificity and G-quadruplex multi-factors on the colorimetric probes 'light-switch' response; working towards the overarching aim of designing metal-based probes that report with complete topological specificity.[94]
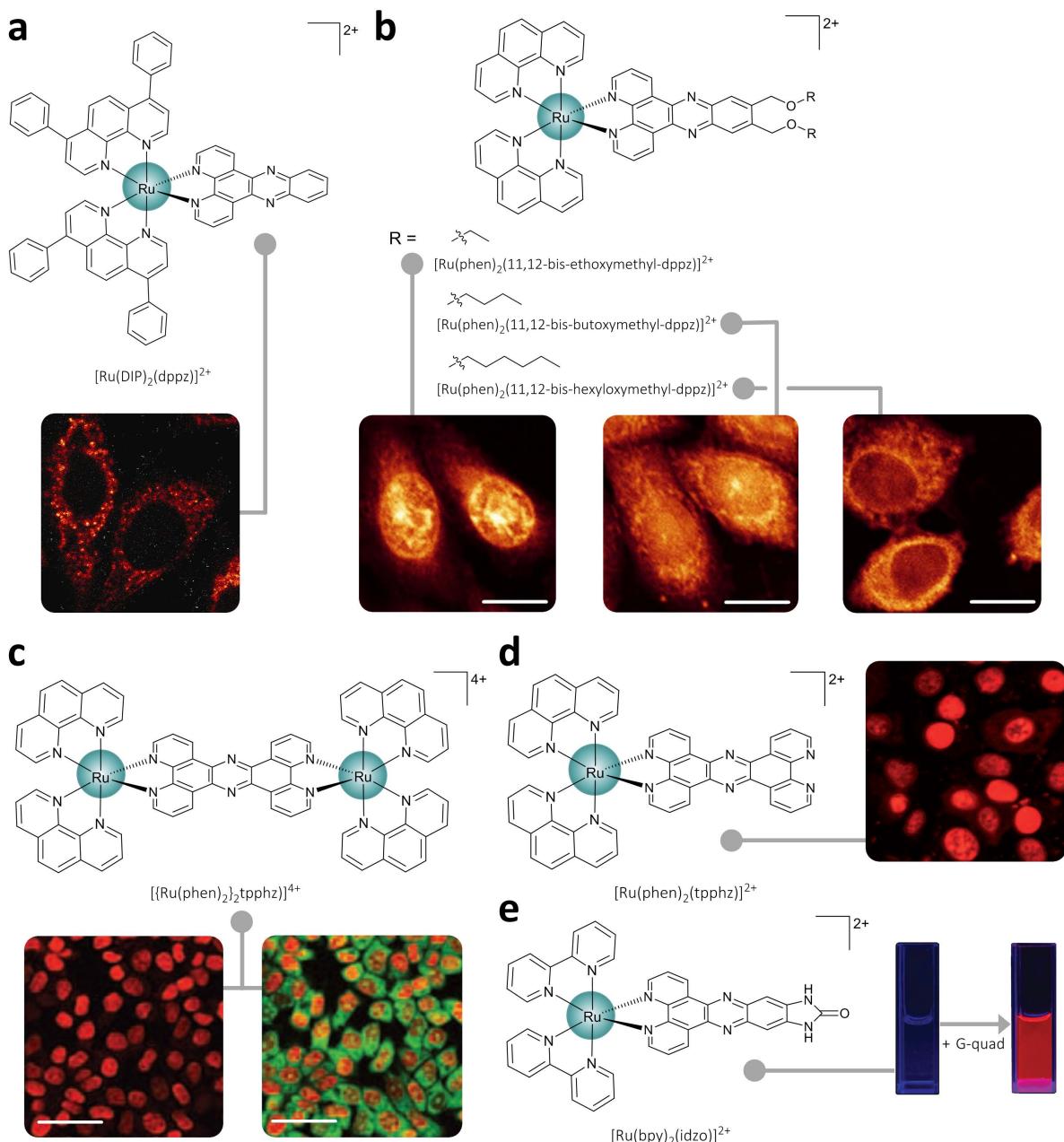
Figure 1.20 *– Structures and in cellulo luminescence response of: (a) [Ru(DIP)₂(dppz)]²⁺ in HeLa cells; (b) bis-oxymethyl derivatives of [Ru(phen)₂(dppz)]²⁺ and the effect of lipophilicity on localisation in fixed CHO-K1 cells; (c) [{Ru(phen)₂}₂(tpphz)]⁴⁺ in MCF-7 breast cancer cells with and without overlay of live cell stain STYO-9 (green); (d) [Ru(phen)₂(tpphz)]²⁺ incubated MCF-7 cells. (e) Structure of [Ru(bpy)₂(dppz-idzo)]²⁺ and the complexes colorimetric 'light-switch' in the presence of wtTel22 G-quadruplex (K⁺). Scale bars represent 10 µm. Adapted by permission of ACS publications, Elsevier publishing, Springer Nature publishing, John Wiley and Sons publishing, and The Royal Society of Chemistry, for (a)-(e) from references 88,89,90,91 and 93 respectively.*

1.6 Binding to G-quadruplexes – Structural Insights

In the past few years, due to the increased understanding of the motif's role in critical cell functions, the therapeutic potential of G-quadruplex targeting drugs has become apparent. Many compounds of disparate architecture have been probed for their efficiency as G-quadruplex binders. As such, libraries of small molecule binders have been explored for their capability to bind to, induce, stabilise, or provide structure specific damage to G-quadruplexes *in vitro*, and to a lesser extent *in vivo*. However, despite the huge interest in this area, a comparatively small fraction of small molecule candidates have been structurally characterised with DNA. A thorough search of the Nucleic Acid Database (NDB) and the Protein Data Bank (PDB) yielded 44 sets of experimentally derived structural coordinates; of which 27 are of X-ray crystallographic origin, and 17 are from solution NMR studies. In the following sections, classes of investigated G-quadruplex binders will be explored with particular focus on how structural studies have aided in the understanding of binding interactions and modes.

1.6.1 Substituted acridines
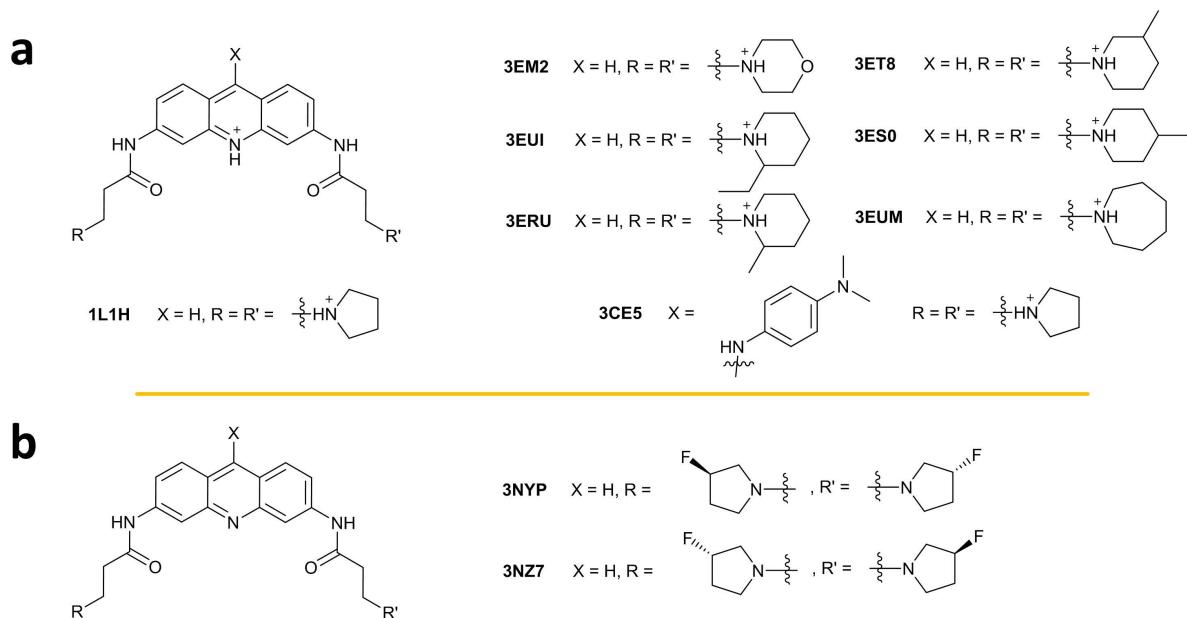
Figure 1.21 – *3,6-bi and 3,6,9-tri-substituted acridines successfully crystallised with G-quadruplex forming oligonucleotides. (a) species containing protonated acridine core; and (b) species containing origin species. PDB codes of structural coordinates are given in bold.*

L. Hurley and S. Neidle *et al* were the first to describe the ability of a small molecule to inhibit telomerase activity selectively. Aided by a structure-based approach to drug development and spurred by previous investigations with triplex interactions, they discovered that a class of compounds, the 2,6-diamidoanthraquinones, were potent G-quadruplex binders that could subsequently disturb the enzymatic action of telomerase (IC$_{50}$ of 23 µm).[95] A systematic analysis of

analogues ensued, where both the π-stacking heterocyclic and the flexible amido chain regions were varied to promote more selective interaction. It was demonstrated that an acridine based aromatic core was more active than the anthraquinone moiety in terms of telomeric G-quadruplex binding. This was associated with the introduction of a nitrogenous heterocycle that, under physiological pH, could be protonated to provide an electron deficient chromophore; it was speculated that this increases stabilisation by acting as a pseudocation, complementing the ion channel running down the helical axis.[96] Further SAR analyses led to a second generation of bi- and tri-substituted amidoacridines.[97] It was determined that bulky non-aromatic substituents on the side chains destabilize the G-quadruplex, but it was concluded, using a combination of solution and *in silico* experiments, that the 3,6,9-trisubstituted aminoalkylamido acridines were the most potent of the three regioisomeric series examined. Specifically, the compound coined BRACO-19 (3CE5 in figure 1.21) emerged as the top candidate, exhibiting high target selectivity/affinity to G-quadruplexes, as well as rapidly relocating to host cell nuclei.[98]

Initially the 3,6-disubstituted acridines were postulated to intercalate/thread the guanine tetrad stack, however these modes were shortly rejected on grounds of poor energetic probability. This was later supported by MD simulations on fully solvated models and fibre diffraction data obtained from unsuccessful crystallisation which convincingly supposed that the chromophore stacks (end-caps) the terminal G-tetrads of the anti-parallel d(GGGGTTTTGGGG).[99] The first G-quadruplex-ligand crystal structure followed shortly, and ultimately verified this understanding. The structure presents the interaction of a 3,6-disubstituted acridine (compound 1L1H) with the dimeric intermolecular quadruplex forming sequence, d(GGGGTTTTGGGG) found in the telomeres of *Oxytricha nova*, a ciliate protozoan (PDB: 1L1H).
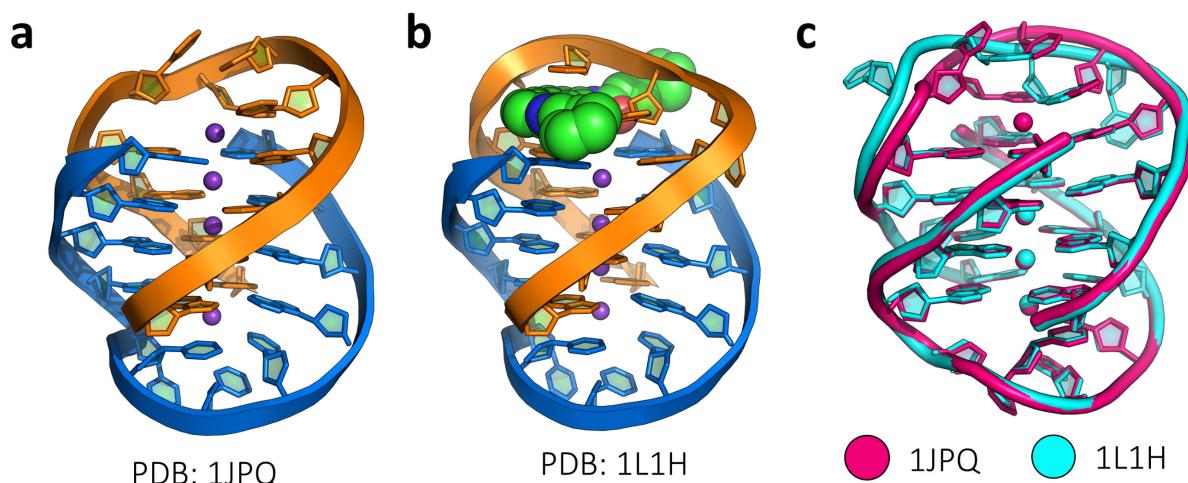
Figure 1.22 – *Crystallographic models of: (a) native bimolecular d(GGGGTTTTGGGG); (b) 3,6-disubstituted acridine bound to d(GGGGTTTTGGGG); and (c) superimposition of both native and bound complexes to highlight the similarity in DNA morphology.*

Crystallographic models of the native DNA show that the G-quadruplex takes on a bimolecular anti-parallel arrangement with diagonal loops; corroborating previous solution state NMR studies of the sequence in both Na^+ and K^+ ionic environments (PDB: 1JPQ).[100] Interestingly, but perhaps not surprisingly, very little change to the DNA global structure is observed upon binding of the acridine derivative. The ligand binds at a 1:1 stoichiometry to the biological unit and binds through one diagonal loop in an end-capping/threading mode; π-stacking predominantly on two *anti*-guanosines located on one side of the terminal tetrad. The pyrrolidinopropioamide chains, thought to interact with the grooves, are splayed out towards the grooves but are not long enough to penetrate them. The protonated ends do however take part in weak H-bonding with an exocyclic N2 of guanine and with local ordered water. Although global DNA structure is conserved upon binding, local divergences in loop geometry are observed. Namely, thymine-3 in the binding loop pocket rotates to stack on top of the threading acridine core; and thymine-4 in the loop is forced to flip out of plane into the mouth of the wide groove to accommodate the ligand.

a PDB: 1L1H

b

c
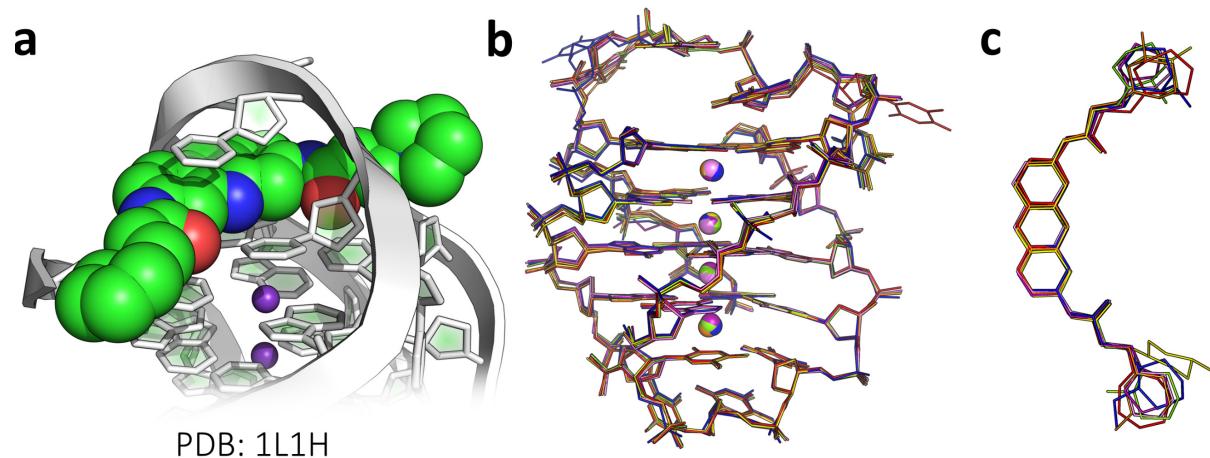
Figure 1.23 – *Crystallographic models of: (a) the general loop threading mechanism of the disubstituted acridines to d(GGGGTTTTGGGG). Superimposition of: (b) all 7 crystallographic models of DNA-drug complexes containing disubstituted acridines; and (c) the compound geometries from the crystallographic models. Note the similarity in DNA architecture, and the rigidity in binding of the π-stacked acridine core.*

Systematic crystallographic studies have been conducted on this class of complexes, where the effect of derivatisation, by means of adding steric bulk to the end of the amide chains, on binding were investigated.[101] The study concluded that the addition of steric constraints to the chains, in the form of progressively larger pyrrolidino rings, does not hinder the ability to bind to the diagonal anti-parallel loops. The seven available structures of disubstituted acridines are superimposed in figure 1.23b, and this highlights how analogous the binding modes for the derivatives are. The authors then postulated that short diagonal loops (≤4 nucleotides) could allow for even more expansive ligands to bind, observing that short propeller type loops impose a far more constricted electrostatic

binding pocket. They used this hypothesis to explain the 7-fold, and additional 3-fold, drops in association constant (K_a) with the parallel stranded human telomeric sequence when the amido chains are substituted with piperidino and azocano rings, respectively. In a similar study, fluorination of the peripheral pyrrolidine moieties on the alkylamido chains was investigated due to the distinctly different ring conformations compared to the parent molecule.[102] Although binding affinities were not calculated, thermal melting analysis suggested that the β-fluorinated analogues exhibited at least half the stabilisation effect on the bimolecular anti-parallel quadruplex in comparison to the parent. Crystallographic analysis of the two derivatives' binding pattern to the same *oxytricha* G-quadruplex yielded almost isomorphous binding morphology to the parent complex; albeit with a change in terminal heterocyclic pucker which produced subtle differences in H-bonding environment. The authors attributed the loss in stability to this new H-bonding network since the principal interactions were now associated only with the uppermost loop; and postulated that the ligand is no longer anchoring the two strands together so adequately.

Qualitative molecular modelling led to the development of BRACO-19, a 3,6,9-trisubstituted acridine whose three amine/amido linked substituents were designed to occupy the grooves of G-quadruplexes whilst the acridine end-caps the G-tetrad stack.[103] BRACO-19 has since been shown to inhibit the activity of telomerase effectively and to disrupt the uncapping of proteins associated with the telomeric overhang region, generating end-to-end chromosomal fusions in prostate cancer lines.[97,104] BRACO-19 exhibited an impressive *in vivo* antineoplastic activity in a xenograft model of uteral carcinoma (96 % inhibition). In the same study, a loss of hTERT protein expression and a heightened prevalence of atypical mitosis which is indicative of telomere disfunction led to crystallographic study of the interactions of the potent candidate with telomeric DNA.[98] BRACO-19 has since become one of the most carefully studied quadruplex ligands, and has subsequently been shown to induce long-term growth arrest and replicative senescence in carcinoma cell lines *in vivo,* as well as being used to demonstrate the positive regulatory role of G-quadruplexes in the transcription of the hepatitis B virus.[105,106]

The complex between BRACO-19 and a bimolecular human telomeric-derived G-quadruplex formed by d(TAGGGTTAGGGT) was evaluated.[107] The quadruplex, formed in physiological concentrations of K⁺, folded into a bimolecular parallel topology with two opposing propeller type loops; reminiscent of the famous unimolecular sequence of duplicate sequence.[35] BRACO-19 can be found bound in an end-capping fashion between the 3' G-tetrad face and a reverse Watson-Crick TATA quartet of a symmetry related biological unit. Despite the complexation occurring with an all-parallel motif, the structure shares commonalities with the previous structures containing

disubstituted acridines; namely the ordering of the cationic ring nitrogen with the ion channel through water-mediated H-bonding, and the flipping in of a loop thymine that interacts with the drug. Distinctly however, the structure shows how the aniline moiety in the 9-position is an almost perfect fit for the hydrophobic pocket at the groove entrance and elucidates a rationale for the 10-fold affinity increase when compared to the disubstituted analogue. Interestingly, it was found that increasing the alkyl length of the side chains in the 3/6-positions dramatically decreased binding by up to 50-fold when up to three -CH_2- units were added. The authors postulated that this was a result of the longer chains spaying out of the G-stack too far to participate in the stabilising H-bond network. The structure especially highlights the three-dimensional complexity of the possible binding sites in telomeric DNA, deviating from preceding uncomplicated end-capping modes.

1.6.2 Metal complexes

As described earlier in the chapter fully coordinated metalloorganic complexes have attracted attention for their therapeutic potential; in most part because binder construction can be more easily tailored to include functionalised π-stacking ligands alongside charged reactive centres. However, despite the large interest in such complexes very little structural information has been published elucidating their interaction with G-quadruplexes; currently only nine structures are present (four X-ray diffraction and five solution NMR).

1.6.2.1 metal salphens

Initially a product of an extensive qualitative modelling investigation, the metal-salphens have since been shown to be strong G-quadruplex binders and potent inhibitors of telomerase.[108–110] Consisting of a heteroaromatic bis-Schiff base derivative tetradentatively coordinated to a square-planar/pyramidal metal centre, the complexes have been systematically optimised to be proficient binders of telomeric G-quadruplexes (figure 1.24a). Central metal ion type, coordination geometry, and substituent effects have all been investigated.

In addition to increasing solubility, the number and nature of the substituents located on the salphen ligand considerably influence the resulting affinity and structural selectivity of the complex. As with the acridines, FRET melting analysis helped to show that pyrrolidinium and piperidinium were the most suitable heterocyclic ends for the ethereal linked alkyl substituents.[109] However, they found that derivatisation around the central phenyl was of larger consequence. Interestingly, despite modelling implications for increased favourable π-π stacking and more than adequate space, posterior phenyl substitution always led to a decrease in FRET melting temperature with the wtTel22 sequence (d(AG$_3$(T$_2$AG$_3$)$_3$)).[111] Substitution in these cases does however often heighten selectivity between

duplex and quadruplex DNA. Investigation into the effect of central metal type revealed that square planar Ni^{II} and Cu^{II} complexes were more stabilising and delivered higher antiproliferative properties and efficacy/propensity for telomerase inhibition over the pseudo-square pyramidal Zn^{II} and V^{IV} complexes; presumably because the planar equatorial coordination could place the metal closer to the ion channel.[112]
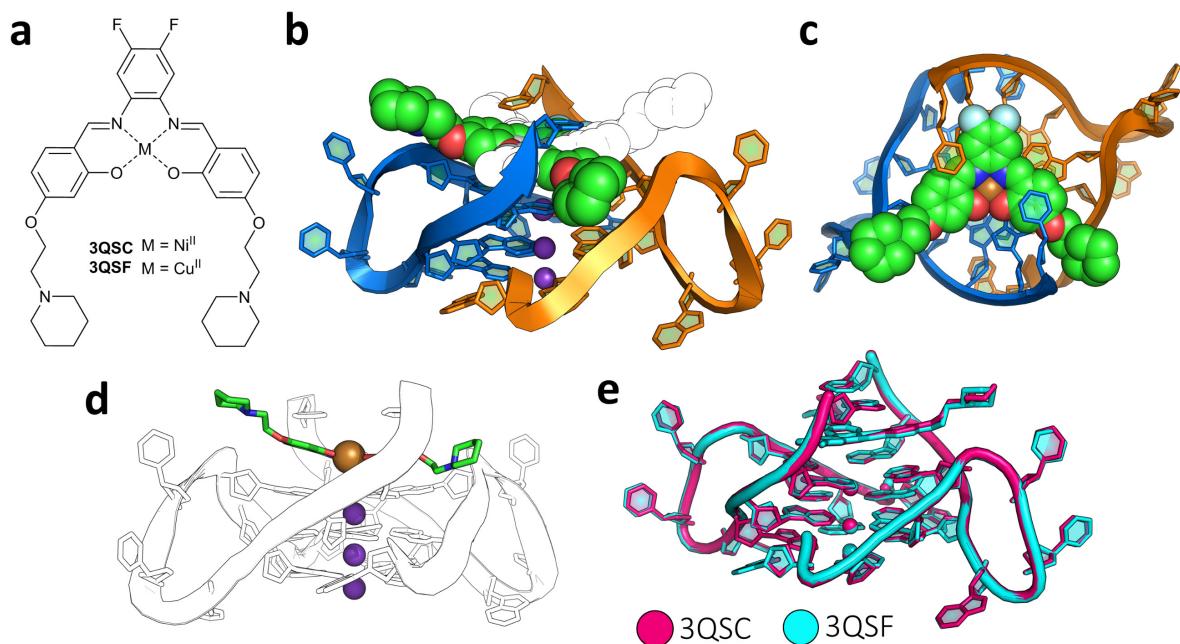
Figure 1.24 – (a) Structures of crystallographically resolved salphens. Crystallographic models of 3QSC bound to d(AGGGT^{Br}UAGGGTT), highlighting (b) the overall complexed structure; (c) the binding site as viewed from the 3' face; and (d) the extension of the ion channel down the helical axis by the central coordinated metal. (e) Superimposition of the two salphen structures.

Crystallographic study yielded two crystal structures of a square planar Ni^{II}/Cu^{II} metal salphen bound to d(AGGGT^{Br}UAGGGTT) (figure 1.24b-e).[111] Both structures contain biological units comprised of a bimolecular all-parallel quadruplex formed by the sequence with a two-fold symmetry axis running down the helical axis. The complexes are seen to bind in an end-capping fashion, as previously postulated from molecular modelling experiments; however, unexpected is the flipping in of a terminal thymine from both strand termini, π-stacking on the faces of the diametric phenyl moieties on the salphen. As designed, the central metal ions are situated in-line with the ionic channel but are not in direct coordination with any O6 guanines (3.8+ Å) since the complexes are held in plane by the salphen. Despite containing different metal centres, the overall structures are effectively isostructural; albeit for a small but significant difference in ligand conformation. The authors noted deviations from planarity of the salphen ligand when comparing the bound Ni^{II} and Cu^{II} complexes, where the Cu^{II} was more severely bend out of plane. This additional bowing could affect π-stacking overlap with the G-tetrad, and indeed a difference in stacking distance of 0.2-0.3 Å is observed and was attributed to

the stark reduction in the binding affinity of the Cu^{II} analogue. In addition, the structure allowed the authors to propose a reason for the decrease in affinity upon substitution of fluorine on the central phenyl. Although postulated to increase favourable π-stacking electronics on the bonded ring, the structure shows that the substituted ring is only partially overlapping the base below and unfavourable repulsive interactions occur between a fluorine and an O6 atom situated on a proximal guanine (substitution of donor groups at this position could exploit this interaction).

1.6.2.2 gold *N*-heterocyclic complexes

Presented as novel alternatives to the archetypical platinum-based chemotherapeutics, gold centred organometallics have recently been established as potent cytotoxins with structural selectivity. Gold(I) mono/dicarbene species (figure 1.25a-b) are especially promising candidates due to their physiological stability, antineoplastic activity, and lower systematic toxicity than previously studied cytotoxic gold complexes.[113] *N*-heterocyclic gold(I) carbenes (NHC) as such have been shown to be potent inhibitors of mitochondrial selenoenzymes, however a multitude of activities are believed to be relevant, including inhibition of proteasome and telomerase activity.[114,115] Interestingly, antitelomerase activity has also been shown to be a distinct mode of action for the antiproliferative effect of Auranofin; a repurposed gold(I) thiolate-based antirheumatic agent that is currently under clinical trial for the treatment of ovarian cancers.[116] Originally a point of contention, gold(I) species were believed to exhibit binding modes similar to that of platinated species, however contrary observations of poor duplex reactivity has led to the divergence from this hypothesis.

The cationic gold(I) bis-carbene, [Au(9-methylcaffein-8-ylidene)$_2$]$^+$ has been evaluated, using an *in vitro* FRET melting assay, to be completely quadruplex specific in its binding and has since been shown to be a selective cytotoxin to cancerous cells.[117,118] The binding mode to the telomeric G-quadruplex forming sequence d(TAGGG(TTAGGG)$_3$) (wtTel23) has been structurally characterised by a combined X-ray crystallographic and ESI-MS study.[119] The crystal structure (PDB: 5CCW) shows how the global topology of the quadruplex is conserved upon binding of the complex in a parallel K$^+$ stabilised conformation. Also supported by solution MS, the structure shows how the complex binds by end-capping on both 5' and 3' tetrad faces in a maximum 3:1 stoichiometry across the biological unit. Interestingly, in each of bound species only one of the caffeinylidene moieties is interacting with the surface of the G-tetrad, leaving the second protruded outside of the tetrad surface, not hindered by the all-parallel loops. Unlike the Ni and Cu containing species, the central metal ion is not aligned or interacting with the central ion channel and instead, in all cases, is stacked between pyrimidines, interacting weakly with the N1 of a guanosine.
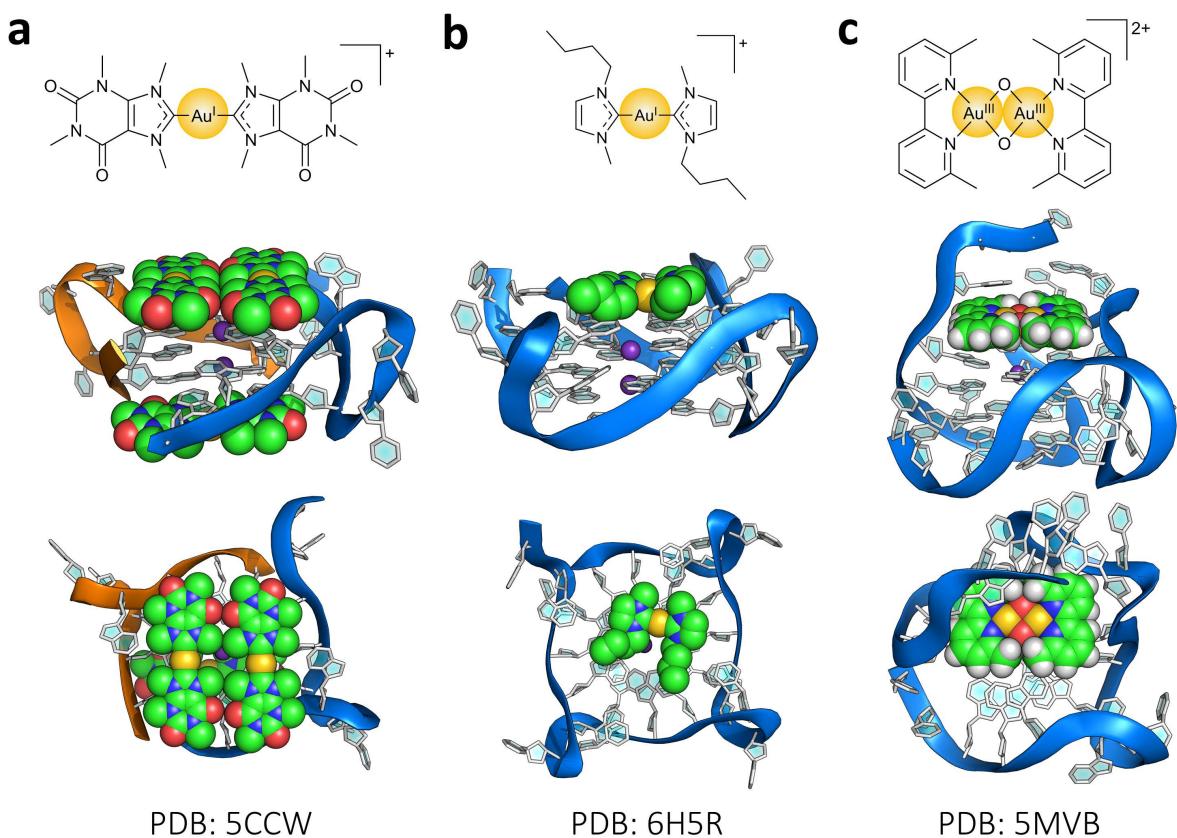
Figure 1.25 – *Chemical structures and crystallographic/NMR models of the gold complexes that have been structurally characterised with G-quadruplex DNA; (a) [Au(9-methylcaffein-8-ylidene)₂]⁺ with d(GGGTTAGGG); (b) [Au(NHC)₂]⁺ with d(TAGGG(TTAGGG)₃T); and (c) [Au₂(6,6'-dimethylbipyridine)₂(μ-O)₂]²⁺ (Auoxo6) with d((TTAGGG)₄TT).*

In a recent analogous combined experiment, the interaction of a simple gold(I) bis-carbene [Au(NHC)₂]⁺ was investigated in the presence of different telomeric G-quadruplexes.[120] Successful crystallisation was achieved with the wtTel24 sequence d(TAGGG(TTAGGG)₃T) (PDB: 6H5R) and similarly produced an all-parallel topology, however in this case the complex is observed at a stoichiometry of 1:1 to biological unit. Interestingly, MS also implied a 1:1 binding stoichiometry, and melting analysis indicated no perceivable increase in ΔT_m; in stark contrast to the stabilisation of the bound caffeinylidene complex. The complex which is symmetrically disordered (four-fold), is π-stacked across two guanine residues on the 3' tetrad; as with 5CCW the metal centre is not aligned with the central ion channel but a weak interaction with N1 is also observed.

Metalloorganic gold(III) complexes, such as the *N*-heterocyclic dioxo bridged binuclear complex Auoxo6 (figure 1.25c), have been shown to exhibit similar *in vivo* cytotoxicity (low μM), as well as inhibition of selenoenzymes, proteasome action and telomerase, to the gold(I) carbene species.[121,122] The complex, which exhibited marked G-quadruplex affinity and selectivity, was a focal point of structural investigation; the interaction of the complex with the wtTel26 sequence d((TTAGGG)₄TT) was probed using solution NMR techniques.[123] The ligand π-stacks asymmetrically on

the 5' tetrad in a pseudo threading/end-capping fashion but, due to its larger footprint, interacts with 3 separate facial guanines. In comparison with the native NMR structure (PDB: 2JPZ) the DNA loop regions in the bound complex have undergone structural rearrangement (RMSD = 3.4 Å) to accommodate the ligand but overall topology is conserved. The compound is shifted in relation to the ion channel axis and no discernible gold-guanine interactions are observed.

Unlike the metal salphens, and other planar cationic species, the central metal ions in the available gold structures show no obvious propensity to stack in line of the ion channel as initially postulated. In the absence of steric limitations, in these cases electrostatics and favourable π overlap influence the binding pocket more than metallic coordination.

1.6.2.3 octahedral ruthenium polypyridyl complexes

Owing to their tractable three-dimensional shape and broad electronic diversity, octahedral metal complexes have been examined as promising G-quadruplex binders. They are especially pertinent when considering topological specificity since secondary DNA interactions with ancillary ligands could, enantiospecifically, mediate preference between morphologies by interaction with strand polarity, *syn/anti* sugars, and loop regions above the binding site. Expansion of the intercalatory ligand scaffold such that larger π-extended ligands are incorporated, tends to generate complexes that exhibit a higher affinity to G-quadruplexes in relation to duplex DNA. Inclusion of such large hydrophobic ligands often negatively affects solubility and subsequently bioavailability, so additional metal centres are often coordinated to offset this.
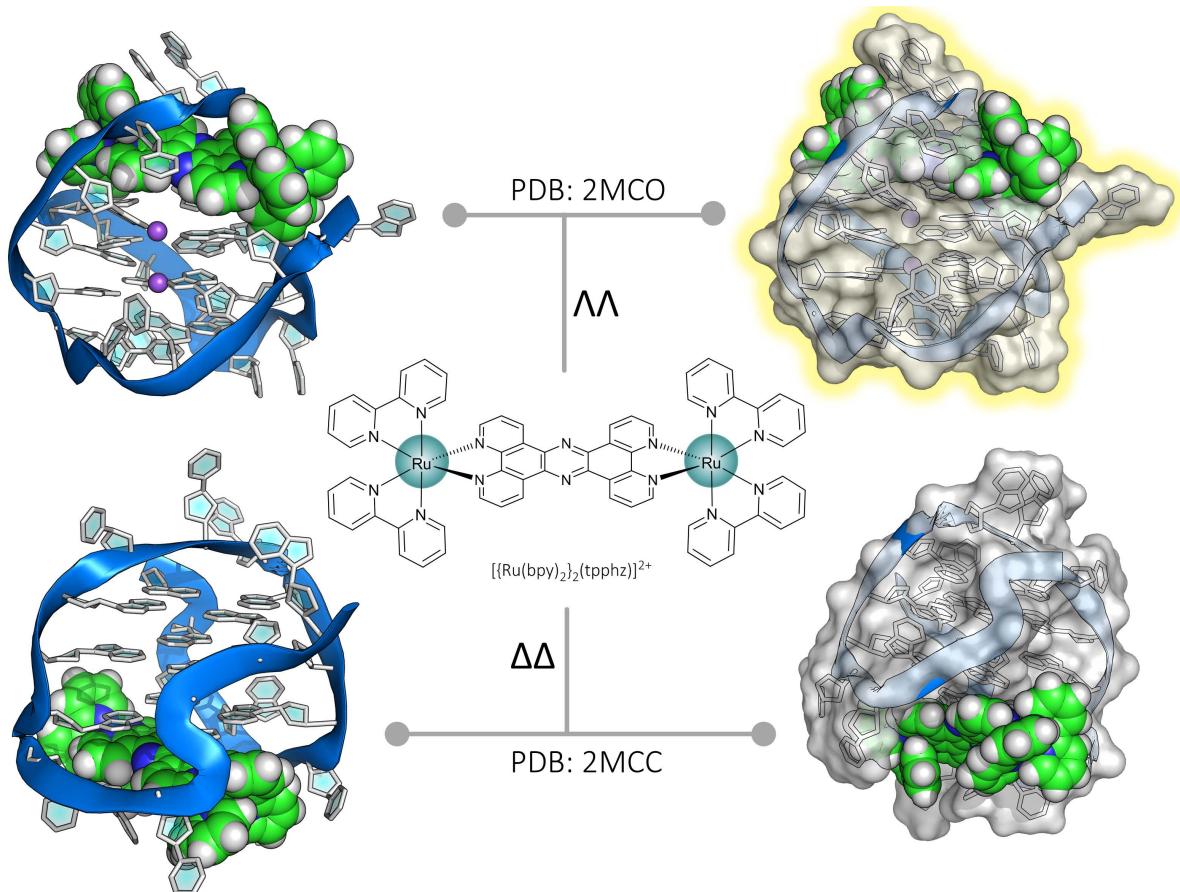
Figure 1.26 – *Chemical structure of the Ru complex that has been structurally characterised with G-quadruplex DNA and NMR models of the interaction of the two enantiomers with the wtTel22 sequence, d(AG₃(T₂AG₃)₃). In both cases the complex endcaps the tetrad stack, however the ΛΛ complex is threading through a diagonal loop and contains additional π-stacking with ancillary ligands. The increased luminescence response of the ΛΛ-Ru in relation to the ΔΔ-Ru isomer has been attributed to the additional encapsulation of the chromophore by the diagonal loop.*

Species such as $[Ru(phen)_2(tpphz)]^{2+}$ and the dinuclear derivative $[\{Ru(phen)_2\}_2(tpphz)]^{4+}$ were reported earlier for their explicit quadruplex luminescence responses and *in cellulo* probe potentials; the latter of which has also been structurally evaluated by Thomas *et al*. Using a combined NMR-MM methodology, the group investigated the isomerically pure binding of $\Delta\Delta/\Lambda\Lambda$-$[\{Ru(bpy/phen)_2\}_2(tpphz)]^{4+}$ to the anti-parallel basket forming telomeric sequence d(AG₃(TTAG₃)₃) (wtTel22).[124] In previous work they noted high affinities for the system, and determined an intense blue-shifted luminescence response of the racemate that had been attributed almost entirely to binding to antiparallel topologies with longer diagonal loops (≥ 3 nucleotides).[125] They later observed that the ΛΛ isomer of the phen analogue is responsible for the bulk of the response (6-fold higher than ΔΔ at saturation), highlighting firstly the enantiomeric disparities in the interaction, and secondly the implication of a highly hydrophobic binding mode for ΛΛ. Unfortunately, unfavourable relaxation rates hampered the NMR studies of the phen analogue, but primary NOE signals supported disparate binding of the enantiomers to different tetrad faces. NOE derived

structures and unconstrained MD simulations were successfully executed on the bpy analogue however. Little perturbation of the native DNA conformation is noted upon binding of ΔΔ and the complex stacks on the opposite 'loopless' tetrad in an end-capping fashion. Conversely, the ΛΛ complex is found to be threaded through the diagonal loop of the basket topology, as originally postulated. The guanosines connectively adjacent to the diagonal loop have been somewhat perturbed by the threading, buckling the distal G-tetrad and creating a tight binding cavity around the chromophore. Ancillary ligand interactions with neighbouring riboses are observed and it is clear to see that assuming the same binding modes, these secondary interactions would be enriched by the larger π-surface of the phen analogue. Generation of the Van der Waals surfaces for both structures highlights the greater envelopment of the ΛΛ chromophore compared to the solvent accessible binding of the ΔΔ; this offers a structural rationale for the observed divergence in luminescence.

1.6.2.4 platinum tripods

The general chemical structure of the platinum tripods consists of a central tertiary amine in possession of three long pendant arms comprising of three aromatic rings and capped with platinum centred units. The three-fold symmetric complex exhibits promising photodynamic therapeutic potential both *in vitro* and *in vivo*, and has been shown to damage DNA rapidly, specifically G-quadruplexes, after light irradiation.[126,127] It has been demonstrated to bind, with fair specificity, to the hybrid-1 telomeric G-quadruplex (Tel26) and this ligand-mediated stability effectively inhibits the activity of telomerase *in vitro* (IC_{50} = 1.22 µM) shown by a TRAP-LIG amplification assay.

Structural investigation of the complexation of the Pt-tripod with the same sequence has been presented, utilising solution NMR studies to elucidate the molecular structures induced by varying DNA-ligand stoichiometries.[128] Binding of the Pt-tripod to the hybrid-1 quadruplex seems to occur favourably on the 5'-end of the tetrad stack where the complex stacks off centre from the helical axis equidistantly above a G·G pair of the top tetrad. The pendant platinated arms align roughly to the π-surface of two edges of the tetrad and protrude through the grooves of the quadruplex, with one of the arms partially enveloped by an A·A·T triad; this partial threading could potentially be a leading factor towards the facial selectivity. A second structure was also solved where at a higher stoichiometry of ligand, a second Pt-tripod unit stacks in a similar fashion but on the 3' tetrad face, inducing a dimeric quadruplex stack in the crystal structure. Interestingly, the authors concluded that despite the observation that at a 1:1 stoichiometry the complex is found bound at the 5' face, binding at the 3' face seemed favoured because the T·A·T triad that formed around the complex was better defined in the structural map and included a more diverse H-bonding environment. It is worth noting however that at a 4:2 stoichiometry the quadruplexes are stacked together by both of their

3'-faces/ligands, which would be likely to form a much tighter binding cavity for the complexes; whereas the 5' tetrad ends, and their three nucleotide overhangs, are fully solvated and devoid of base pairing, thus more disordered around the complex.
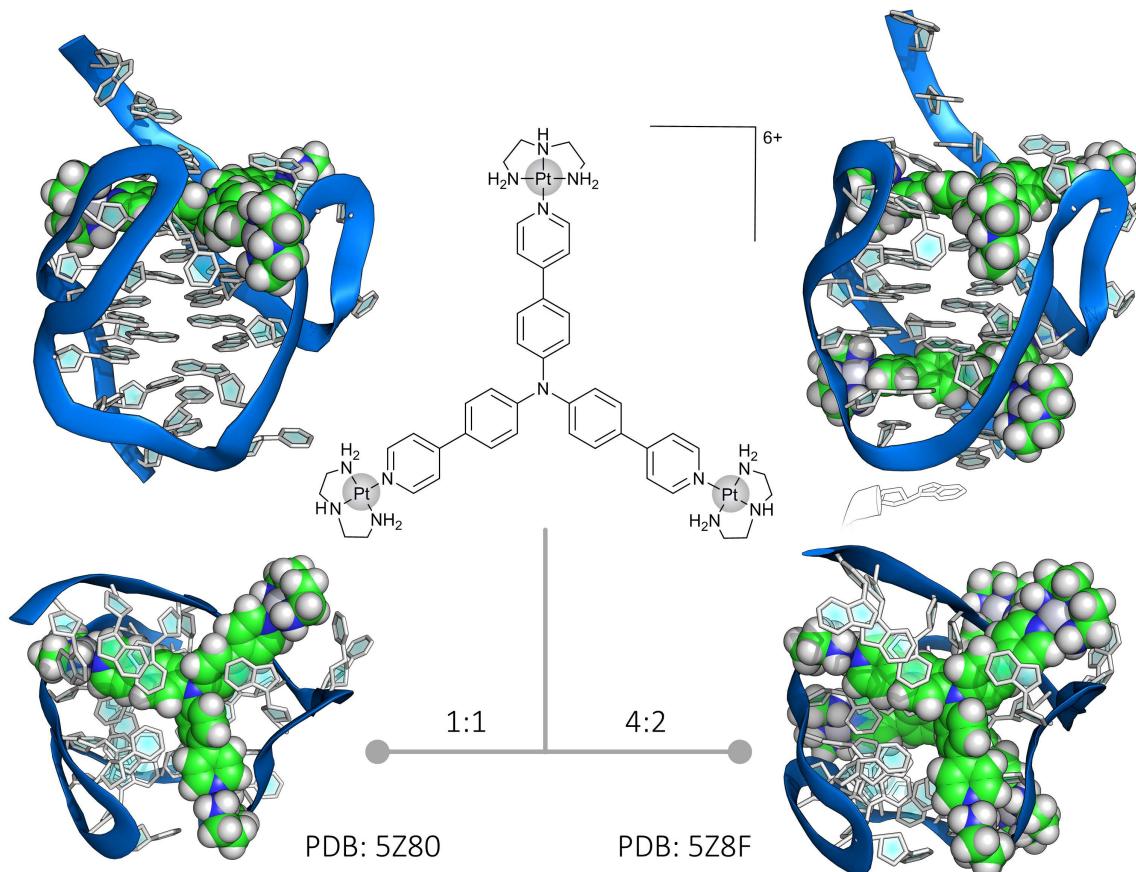
Figure 1.27 – *The non-planar Pt-tripod complex has been investigated in the presence of the Tel26 human telomeric sequence, d(A₃(G₃T₂A)₃G₃A₂), using NMR and subsequent NOE restrained MD. Two discrete structures were obtained from different stoichiometries highlighting two distinct binding pockets but an overarching preference for binding to the 5' tetrad. Note the shifted location of the tripod in relation to the terminal tetrad; presumably to increase π-stacking and allow the platinated arms to fit into the grooves without dislocating the DNA backbone.*

The complex itself is an interesting departure from the norm for quadruplex binding agents; containing no extended planar π-surfaces, and containing active metal centres that are designed to protrude away from the central quadruplex stack.

1.6.3 Structural survey summary

In the grand scheme of the large field of G-quadruplex targeting ligands, very little structural information is present in the literature. Furthermore, only a small percentage of these are pertaining to the interactions of metal centred species, where most of the available data sets describe the interaction of mostly planar organics and their derivatives. Despite this, where present, the structural results have imparted useful insights, explaining conformational specificities in binding and providing

useful leads to develop more targeted complexes. A summary of all publicly available G-quadruplex-ligand structures are given in tables 1.4 (XRD) and A1.1 (solution NMR) and these attempt to outline some of the key features of the ligands and their subsequent binding interactions.

One of the leading aims of current research efforts in this field is to develop ligands that not only display motif selectivity but also topological selectivity in binding. Of course there are lots of conformational features that alter between folded unimolecular topologies, even if the sequence is constant. Loop regions in relevant sequences are often short, and between conformations can switch from providing an optimal binding cavity, to looping over the tetrad and blocking bulky ligand threading. Summarising the available structures however highlights just how flexible these loops are, in most cases they at least interact with the ligands electrostatically, but more often than not they directly π-stack, even flipping bases against the natural torsion to do so. Such accommodation of the ligand by the DNA surely makes it difficult to design planar complexes that exhibit topological specificity. The effect of adding long and often charged side chains is observed frequently in the structural investigations where enhanced G-quadruplex specific interactions with groove regions and phosphates are often seen. A trend away from the planar organic framework is noted as well, with more recent structures containing more metal centres. In some cases these charged metals are designed to interact directly with the central ionic channel, and they often do so. This spatial understanding should provide a predictable starting model for designing derivatives or similar complexes with topology specific interactions in mind (i.e. *syn/anti* placement, phosphate geometry, and H-bonding with adjacent loop bases). In other cases, such as with the dinuclear ruthenium polypyridyls, the complexes exhibit very different binding modes, with one enantiomer discriminating topology and exhibiting enantiospecific interactions with the guanine stack. This not only reiterates the importance of working with enantiopure material, but more loosely highlights the advantage of G-quadruplex binders with rigid three-dimensional profile.

In summary, XRD and NMR solution studies have proven to be an indispensable technique for understanding how G-quadruplex ligands bind to their target and have often aided in the development of more specific binding agents by elucidating structure specific binding features.

Table 1.4 – *Summary of all publicly available G-quadruplex-ligand XRD structures*

PDB	Date	Experiment	Sequence	Salt	Topology	Ligand	Ligand Class	Ligand Types	Binding Interactions	Res(Å)	R_{work}/R_{free}
1L1H	2003	X-ray Diffraction	GGGGTTTTGGGG	20 mM K+	Anti-parallel (hairpin)	BSI-6039 (3,6-disubstituted acridine)	Organic Heterocycle w/ saturated chain substituents (cationic)	P/nP/Hc/HA/HD/3+	Tetrad π-stacking and H-bond with thymines in loops.	1.75	0.139/0.219
1O0K	2003	X-ray Diffraction	TGGGGT	? Na+	Parallel	Daunomycin (Anthracycline)	Organic heterocycle (Non-planar)	P/nP/Hc/HA/HD	Tetrad π-stacking	1.17	0.157/0.197
2HRI	2007	X-ray Diffraction	TAGGGTAGGG	70 mM K+	Parallel	TMPyP4 (Tetrapyridyl porphyrin)	Organic Macrocycle (Planar/Cationic)	P/nP/Hc/HA/HD/4+	Tetrad π-stacking	2.09	0.198/0.248
3CCO	2008	X-ray Diffraction	TAGGG(TTAGGG)3	100 mM K+, 50 mM Na+	Parallel (Propeller)	tetra-substituted naphthalene diimide	Organic Heterocycle w/ saturated chain substituents	P/nP/Hc/HA/HD	Tetrad π-stacking and H-bond with TTA loops	2.2	0.254/0.303
3CDM	2008	X-ray Diffraction	TAGGGTAGGGT	350 mM Na+	Parallel	tetra-substituted naphthalene diimide	Organic Heterocycle w/ saturated chain substituents	P/nP/Hc/HA/HD	Tetrad π-stacking and H-bond with TTA loops	2.1	0.234/0.295
3CE5	2008	X-ray Diffraction	TAGGGTAGGGT	20 mM K+, 20 mM Na+	Parallel	BRACO-19 (3,6,9-trisubstituted acridine)	Organic Heterocycle w/ saturated chain substituents (charged)	P/nP/Hc/HA/HD/3+	Tetrad π-stacking/T1 π-stacking	2.5	0.183/0.213
3EM2	2008	X-ray Diffraction	GGGGTTTTGGGG	40 mM K+	Anti-parallel	BSI-6038 (3,6-disubstituted acridine)	Organic Heterocycle w/ saturated chain substituents (charged)	P/nP/Hc/HA/HD/3+	Tetrad π-stacking	2.3	0.203/0.293
3EUI	2008	X-ray Diffraction	GGGGTTTTGGGG	40 mM K+	Anti-parallel	BSI-6042 (3,6-disubstituted acridine)	Organic Heterocycle w/ saturated chain substituents (charged)	P/nP/Hc/HA/HD/3+	Tetrad π-stacking	2.2	0.234/0.285
3ERU	2008	X-ray Diffraction	GGGGTTTTGGGG	40 mM K+	Anti-parallel	BSI-6045 (3,6-disubstituted acridine)	Organic Heterocycle w/ saturated chain substituents (charged)	P/nP/Hc/HA/HD/3+	Tetrad π-stacking	2	0.199/0.249
3ES0	2008	X-ray Diffraction	GGGGTTTTGGGG	40 mM K+	Anti-parallel	BSI-6048 (3,6-disubstituted acridine)	Organic Heterocycle w/ saturated chain substituents (charged)	P/nP/Hc/HA/HD/3+	Tetrad π-stacking	2.2	0.195/0.245
3ET8	2008	X-ray Diffraction	GGGGTTTTGGGG	40 mM K+	Anti-parallel	BSI-6054 (3,6-disubstituted acridine)	Organic Heterocycle w/ saturated chain substituents (charged)	P/nP/Hc/HA/HD/3+	Tetrad π-stacking	2.45	0.207/0.277
3EUM	2008	X-ray Diffraction	GGGGTTTTGGGG	40 mM K+	Anti-parallel	BSI-6066 (3,6-disubstituted acridine)	Organic Heterocycle w/ saturated chain substituents (charged)	P/nP/Hc/HA/HD/3+	Tetrad π-stacking	1.78	0.222/0.270
3NYP	2010	X-ray Diffraction	GGGGTTTTGGGG	20 mM K+	Anti-parallel	(R, R)-β-fluorinated 3,6-disubstituted acridine)	Organic Heterocycle w/ saturated chain substituents (charged)	P/nP/C/Hc/HA/HD	Tetrad π-stacking and Loop threading	1.18	0.166/0.188
3NZ7	2010	X-ray Diffraction	GGGGTTTTGGGG	10 mM K+	Anti-parallel	(S, S)-β-fluorinated 3,6-substituted acridine)	Organic Heterocycle w/ saturated chain substituents (charged)	P/nP/C/Hc/HA/HD	Tetrad π-stacking and Loop threading	1.1	0.138/0.159
3QSC	2012	X-ray Diffraction	AGGGTUbrAGGGTT	20 mM K+	Parallel	Copper(II)(salphen)	Metal-Organic (Square-planar)	M/P/nP/Hc/HA/2+	Tetrad π-stacking	2.4	0.213/0.234
3QSF	2012	X-ray Diffraction	AGGGTAGGGTT	16.7 mM K+	Parallel	Nickel(II)(salphen)	Metal-Organic (Square-planar)	M/P/nP/Hc/HA/2+	Tetrad π-stacking	2.4	0.237/0.320
3R6R	2012	X-ray Diffraction	TAGGG(T2AG3)3	125 mM K+	Parallel	Berberine (benzyltetrahydroisoquinoline alkaloid)	Organic Heterocycle (Planar)/Quat	M/P/nP/Hc/HA/1+	Tetrad π-stacking	2.4	0.215/0.252
3TSE	2012	X-ray Diffraction	AGGG(TTAGGG)3	182.5 mM K+	Parallel	BMSG-SH-4(/2) (Naphthalene diimide)	Organic Heterocycle w/ saturated chain substituents (charged)	P/nP/Hc/HA/HD	Tetrad π-stacking and electrostatic groove interaction	2.1	0.245/0.282
3TVB	2012	X-ray Diffraction	GGGG	30 mM Na+	Parallel	Daunomycin (Anthracycline)	Organic heterocycle (Non-planar)	P/nP/Hc/HA/HD	Tetrad π-stacking and H-bond with O1P	1.08	0.159/0.202
3UYH	2012	X-ray Diffraction	AGGG(TTAGGG)3	75 mM Na+, 50 mM K+	Parallel	MM41 (tetra-substituted Naphthalene dimide)	Organic Heterocycle w/ saturated chain substituents (charged)	P/nP/Hc/HA/HD	Tetrad π-stacking	1.95	0.231/0.275
4DA3	2012	X-ray Diffraction	GGG(TTAGGG)3	100 mM Na+	Parallel	MM41 (tetra-substituted Naphthalene dimide)	Organic Heterocycle w/ saturated chain substituents (charged)	P/nP/Hc/HA/HD	Tetrad π-stacking	2.4	0.242/0.282
4DAQ	2012	X-ray Diffraction	AGGG(TTAGGG)3	65 mM K+	Parallel	BMSG-SH-3 (Naphthalene dimide)	Organic Heterocycle w/ saturated chain substituents (charged)	P/nP/Hc/HA/HD	Tetrad π-stacking	2.3	0.258/0.288
4FXM	2012	X-ray Diffraction	AGGG(TTAGGG)3	50 mM K+	Parallel	N-Methyl mesoporphyrin IX (NMM)	Organic Heterocycle w/ saturated chain substituents (charged)	P/nP/Hc/HA/HD	Tetrad π-stacking	1.65	0.222/0.262
5CCW	2016	X-ray Diffraction	GGGTAGGG	50 mM K+	Parallel	Au(I)(methylcaffeinylidene) (Gold(I) bis-carbene)	Metalloheterocyclic carbene (Square-planar)	M/P/nP/Hc/HA/1+	Tetrad π-stacking	1.89	0.225/0.260
5HIX	2016	X-ray Diffraction	GGGGTTTTGGGG	90 mM K+	Anti-parallel (hairpin)	oligo-quinolinecarbozamide foldamer	Chiral organic foldamer	P/nP/Hc/HA/HD	Ammonium-phosphate electrostatic interactions	2.48	0.227/0.288
5LS8	2018	X-ray Diffraction	TAGGGTTA	50 mM K+	Anti-parallel	[Ru(TAP)2(11-CN-dppz)]2+ (Ruthenium polypyridyl)	Metal-Organic (Octahedral/charged)	M/P/nP/Hc/HA/2+	Tetrad π-stacking/GpA intercalation and 5' sugar stacking	1.76	0.214/0.233
6HSR	2018	X-ray Diffraction	TAGGG(TTAGGG)T	50 mM K+	Parallel	Au(I)bis-(1-butyl-3-methyl-imadazole-2-ylidene) (Gold(I) bis-carbene)	Metal N-heterocyclic carbene	M/P/nP/Hc/HA/1+	Tetrad π-stacking	2	0.228/0.253

1.7 Summary and Project Aims

Ruthenium centred polypyridyls hold a great deal of prospective diagnostic and therapeutic utility, as is evident in the literature; from use as luminescent handles of DNA structure and cellular function, to candidates for photosensitised therapies. Understanding how this class of compounds interact with DNA on an atomic scale yields informative insights that can help explain solution observations, provide leads for increasing specificity, and aid in the design of the next generation of metal centred DNA binders. Structural analysis of such complexes in the presence of duplex DNA has been successful in recent years and several structures are available where commonly observed binding traits are now being extrapolated to predict and/or explain the behaviour of octahedral metal polypyridyls in solution. Despite this, an imbalance still lies between their catalogued activity in the solution state and the available structural knowledge pertaining their exact binding modes and intricacies of interaction. This is especially true in relation to binding to higher order DNA motifs, where no structural data is present in the literature regarding the interaction of mononuclear complexes.

Broadly speaking, the overall aim of this thesis is to aid in the development of our understanding of the binding of ruthenium polypyridyls to canonical and non-canonical DNA. It aims to provide structural evidence of binding patterns, and to investigate how binding to DNA affects photophysical responses and biological processes. Specifically, the aims of this work are as follows:

- **Structurally define the binding modes of substituted polypyridyl complexes to duplex DNA.** Previous structures from the group were the first to elucidate the binding of dppz containing complexes to duplex DNA; however, little information exists regarding to derivatised chromophores and their effect on the intercalation cavity. Investigation into the effect of π-directing substitutions has implicated stronger affinities and thermal stabilities for interactions with π-deficient intercalatory ligands. Structural insights may clarify binding geometries and provide logic for the perceived effects of substitution pattern, symmetry, and type, on thermodynamic observations.

- **Use crystallography to characterise the binding modes of octahedral ruthenium complexes to quadruplex DNA**. End-capping, intercalation, and groove binding have all been postulated to be the principal binding mode of Ru-dppz complexes to quadruplex DNA. Here we aim to define some of the possible binding modes and help deconvolute the important interactions necessary for topological specificity of binding. In addition, substantial enantiomeric differences in affinity and luminescence yield have been noted in the literature; with often disparate binding modes postulated for each isomer.

Understanding the manner in which these complexes bind, and the potential interactions that guide preferential binding, will aid in the design of more specialised and specific quadruplex binding agents.

- **Examine the effects of binding of ruthenium complexes to G-quadruplexes on their ability to be replicated by polymerases and to study their *in cellulo* binding efficacy**. G-quadruplex formation has been shown in some cases to hinder the progression of polymerases, reverse transcriptases, and telomerases by a multitude of modes. Small molecule interactions with these motifs can stabilise them and lead to an inhibition of enzyme activity. Thus the aim is to evaluate Ru-dppz complexes in this context and derive enantiomeric and substitutional patterns, and deduce the effects of topology and irradiation on these processes. In addition, it would be desirable to investigate the ability for this class of compounds to bind to G-quadruplexes *in vivo*.

1.8 References

1   Sobell, H. M. Actinomycin and DNA transcription. *Proc. Natl. Acad. Sci.* **82**, 5328–5331 (1985).

2   Siddik, Z. H. Cisplatin: mode of cytotoxic action and molecular basis of resistance. *Oncogene.* **22**, 7265 (2002).

3   Florea, A.-M. & Büsselberg, D. Cisplatin as an anti-tumor drug: cellular mechanisms of activity, drug resistance and induced side effects. *Cancers (Basel).* **3**, 1351–1371 (2011).

4   Alessio, E. & Messori, L. NAMI-A and KP1019/1339, Two iconic ruthenium anticancer drug candidates face-to-face: A case story in medicinal inorganic chemistry. *Molecules* **24**, (2019).

5   Watson, J. D. & Crick, F. H. C. Molecular Structure of Nucleic Acids: A structure for deoxyribose nucleic acid. *Nature* **171**, 737–738 (1953).

6   Elson, D. & Chargaff, E. On the desoxyribonucleic acid content of sea urchin gametes. *Experientia* **8**, 143–145 (1952).

7   Chalikian, T. V, Völker, J., Plum, G. E. & Breslauer, K. J. A more unified picture for the thermodynamics of nucleic acid duplex melting: A characterization by calorimetric and volumetric techniques. *Proc. Natl. Acad. Sci.* **96**, 7853–7858 (1999).

8   Yakovchuk, P., Protozanova, E. & Frank-Kamenetskii, M. D. Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic Acids Res.* **34**, 564–574 (2006).

9   Hoogsteen, K. The crystal and molecular structure of a hydrogen-bonded complex between 1-methylthymine and 9-methyladenine. *Acta Crystallogr.* **16**, 907–916 (1963).

10   Nikolova, E. N. *et al.* Transient Hoogsteen base pairs in canonical duplex DNA. *Nature* **470**, 498-502 (2011).

11   Rice, P. A., Yang, S., Mizuuchi, K. & Nash, H. A. Crystal structure of an IHF-DNA complex: A protein-induced DNA U-turn. *Cell* **87**, 1295–1306 (1996).

12   Aishima, J. *et al.* A Hoogsteen base pair embedded in undistorted B-DNA. *Nucleic Acids Res.* **30**, 5244–5252 (2002).

13   Zerbe, B. *et al.* Computational mapping reveals dramatic effect of Hoogsteen breathing on duplex DNA reactivity with formaldehyde. *Nucleic Acids Res.* **40**, 7644–7652 (2012).

14   Franklin, R. E. & Gosling, R. G. Molecular configuration in sodium thymonucleate. *Nature* **171**, 740–741 (1953).

15   Chargaff, E., Lipshitz, R. & Green, C. Composition of the desoxypentose nucleic acids of four genera of sea-urchin. *J. Biol. Chem.* **195**, 155–160 (1952).

16   Suh, D. & Chaires, J. B. Criteria for the mode of binding of DNA binding agents. *Bioorg. Med. Chem.* **3**, 723–728 (1995).

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

18   Keepers, J. W., Kollman, P. A., Weiner, P. K. & James, T. L. Molecular mechanical studies of DNA

flexibility: Coupled backbone torsion angles and base-pair openings. *Proc. Natl. Acad. Sci.* **79**, 5537–5541 (1982).

19   Altona, C. & Sundaralingam, M. Conformational analysis of the sugar ring in nucleosides and nucleotides. New description using the concept of pseudorotation. *J. Am. Chem. Soc.* **94**, 8205-8212 (1972).

20   Neidle, S. *DNA Structure and Recognition*. (Oxford University Press Oxford, 1994).

21   Steffen, N. R., Murphy, S. D., Tolleri, L., Hatfield, G. W. & Lathrop, R. H. DNA sequence and structure: direct and indirect recognition in protein-DNA binding. *Bioinforma.* **18**, S22–S30 (2002).

22   Dmitrieva, N. I., Cui, K., Kitchaev, D. A., Zhao, K. & Burg, M. B. DNA double-strand breaks induced by high NaCl occur predominantly in gene deserts. *Proc. Natl. Acad. Sci.* **108**, 20796-20801 (2011).

23   Whelan, D. R. *et al.* Detection of an en masse and reversible B- to A-DNA conformational transition in prokaryotes in response to desiccation. *J. R. Soc. Interface* **11**, 1-6 (2014).

24   Thamann, T. J., Lord, R. C., Wang, A. H. J. & Rich, A. The high salt form of poly(dG-dC)·poly(dG-dC) is left-handed Z-DNA: Raman spectra of crystals and solutions. *Nucleic Acids Res.* **9**, 5443-5458 (1981).

25   Wittig, B., Dorbic, T. & Rich, A. Transcription is associated with Z-DNA formation in metabolically active permeabilized mammalian cell nuclei. *Proc. Natl. Acad. Sci.* **88**, 2259-2263 (1991).

26   Wang, G., Christensen, L. A. & Vasquez, K. M. Z-DNA-forming sequences generate large-scale deletions in mammalian cells. *Proc. Natl. Acad. Sci.* **103**, 2677 LP – 2682 (2006).

27   Crawford, J. L. *et al.* The tetramer d(CpGpCpG) crystallizes as a left-handed double helix. *Proc. Natl. Acad. Sci.* **77**, 4016 LP – 4020 (1980).

28   Sinden, R. R., Pearson, C. E., Potaman, V. N. & Ussery, D. W. DNA: Structure and function. in *Genes and Genomes* **5**, 1–141 (1998).

29   Henderson, E., Hardin, C. C., Walk, S. K., Tinoco Jr., I. & Blackburn, E. H. Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine-guanine base pairs. *Cell* **51**, 899–908 (1987).

30   Gellert, M., Lipsett, M. N. & Davies, D. R. Helix formation by guanylic acid. *Proc. Natl. Acad. Sci.* **48**, 2013–2018 (1962).

31   Sen, D. & Gilbert, W. Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature* **334**, 364–366 (1988).

32   Gu, J., Leszczynski, J. & Bansal, M. A new insight into the structure and stability of Hoogsteen hydrogen-bonded G-tetrad: an ab initio SCF study. *Chem. Phys. Lett.* **311**, 209–214 (1999).

33   Cang, X., Šponer, J. & Cheatham 3rd, T. E. Explaining the varied glycosidic conformational, G-tract length and sequence preferences for anti-parallel G-quadruplexes. *Nucleic Acids Res.* **39**, 4499–4512 (2011).

34 Špačková, N., Berger, I. & Šponer, J. Nanosecond molecular dynamics simulations of parallel and antiparallel guanine quadruplex DNA molecules. *J. Am. Chem. Soc.* **121**, 5519–5534 (1999).

35 Parkinson, G. N., Lee, M. P. H. & Neidle, S. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* **417**, 876–880 (2002).

36 Balasubramanian, S. & Huppert, J. L. G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res.* **35**, 406–413 (2006).

37 Biffi, G., Tannahill, D., McCafferty, J. & Balasubramanian, S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* **5**, 182–186 (2013).

38 Rodriguez, R. *et al.* Small-molecule–induced DNA damage identifies alternative DNA structures in human genes. *Nat. Chem. Biol.* **8**, 301 (2012).

39 Zahler, A. M., Williamson, J. R., Cech, T. R. & Prescott, D. M. Inhibition of telomerase by G-quartet DNA structures. *Nature* **350**, 720 (1991).

40 Sun, Z.Y., Wang, X.N., Cheng, S.Q., Su, X.X. & Ou, T.M. Developing novel G-quadruplex ligands: from interaction with nucleic acids to interfering with nucleic acid-protein interaction. *Molecules* **24**, 396 (2019).

41 Rizzo, A. *et al.* Evidence for G-quadruplex in the promoter of vegfr-2 and its targeting to inhibit tumor angiogenesis. *Nucleic Acids Res.* **42**, 2945–2957 (2013).

42 Siddiqui-Jain, A., Grand, C. L., Bearss, D. J. & Hurley, L. H. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci.* **99**, 11593–11598 (2002).

43 Asamitsu, S., Bando, T. & Sugiyama, H. Ligand Design to Acquire Specificity to Intended G-Quadruplex Structures. *Chem. Eur. J.* **25**, 417–430 (2019).

44 Drygin, D. *et al.* Anticancer activity of CX-3543: A direct inhibitor of rRNA biogenesis. *Cancer Res.* **69**, 7653–7661 (2009).

45 Kumar, S., Pandya, P., Pandav, K., Gupta, S. & Chopra, A. Structural studies on ligand–DNA systems: A robust approach in drug design. *J. Biosci.* **37**, 553-561 (2012).

46 Chen, A. Y., Yu, C., Gatto, B. & LIu, L. F. DNA minor groove-binding ligands: a different class of mammalian DNA topoisomerase I inhibitors. *Proc. Natl. Acad. Sci.* **90**, 8131–8135 (1993).

47 West, S. M., Rohs, R., Mann, R. S. & Honig, B. Electrostatic interactions between arginines and the minor groove in the nucleosome. *J. Biomol. Struct. Dyn.* **27**, 861–866 (2010).

48 Lown, J. W. DNA recognition by lexitropsins, minor groove binding agents. *J. Mol. Recognit.* **7**, 79–88 (1994).

49 Graham, M. A. *et al.* Clinical pharmacokinetics of Oxaliplatin: A critical review. *Clin. Cancer Res.* **6**, 1205–1218 (2000).

50 Polavarapu, A., Stillabower, J. A., Stubblefield, S. G. W., Taylor, W. M. & Baik, M.H. The mechanism of guanine alkylation by nitrogen mustards: A computational study. *J. Org. Chem.* **77**, 5914–5921 (2012).

51 Williams, L. D., Egli, M. & Rich, A. DNA Intercalation: Helix Unwinding and Neighbor-Exclusion. *Structure and Function.* **1**, 107-125 (1992).

52 Chow, K.-C., Macdonald, T. L. & Ross, W. E. DNA binding by epipodophyllotoxins and N-acyl anthracyclines: implications for mechanism of topoisomerase II inhibition. *Mol. Pharmacol.* **34**, 467–473 (1988).

53 Crenshaw, J. M., Graves, D. E. & Denny, W. A. Interactions of acridine antitumor agents with DNA: binding energies and groove preferences. *Biochemistry* **34**, 13682–13687 (1995).

54 Lerman, L. S. Structural considerations in the interaction of DNA and acridines. *J. Mol. Biol.* **3**, 18-30 (1961).

55 Liu, H.K. & Sadler, P. J. Metal complexes as DNA intercalators. *Acc. Chem. Res.* **44**, 349–359 (2011).

56 Drummond, T. G., Hill, M. G. & Barton, J. K. Electrochemical DNA sensors. *Nat. Biotechnol.* **21**, 1192 (2003).

57 Maheswari, P. U. & Palaniandavar, M. DNA binding and cleavage properties of certain tetrammine ruthenium(II) complexes of modified 1,10-phenanthrolines–effect of hydrogen-bonding on DNA-binding affinity. *J. Inorg. Biochem.* **98**, 219–230 (2004).

58 Li, F., Collins, J. G. & Keene, F. R. Ruthenium complexes as antimicrobial agents. *Chem. Soc. Rev.* **44**, 2529–2542 (2015).

59 Niyazi, H. *et al.* Crystal structures of Λ-[Ru(phen)$_2$dppz]$^{2+}$ with oligonucleotides containing TA/TA and AT/AT steps show two intercalation modes. *Nat. Chem.* **4**, 621–628 (2012).

60 Paris, J. P. & Brandt, W. W. Charge transfer luminescence of a ruthenium(II) chelate. *J. Am. Chem. Soc.* **81**, 5001–5002 (1959).

61 Barton, J. K., Danishefsky, A. & Goldberg, J. Tris (phenanthroline) ruthenium(II): stereoselectivity in binding to DNA. *J. Am. Chem. Soc.* **106**, 2172–2176 (1984).

62 Satyanarayana, S., Dabrowiak, J. C. & Chaires, J. B. Neither delta- nor lambda-tris(phenanthroline)ruthenium(II) binds to DNA by classical intercalation. *Biochemistry* **31**, 9319–9324 (1992).

63 Satyanarayana, S., Dabrowiak, J. C. & Chaires, J. B. Tris(phenanthroline)ruthenium(II) enantiomer interactions with DNA: Mode and specificity of binding. *Biochemistry* **32**, 2573-2584 (1993).

64 Eriksson, M., Leijon, M., Hiort, C., Norden, B. & Graeslund, A. Binding of delta- and lambda-[Ru(phen)$_3$]$^{2+}$ to [d(CGCGATCGCG)]$_2$ Studied by NMR. *Biochemistry* **33**, 5031–5040 (1994).

65 Coggan, D. Z. M., Haworth, I. S., Bates, P. J., Robinson, A. & Rodger, A. DNA binding of ruthenium tris(1,10-phenanthroline): Evidence for the dependence of binding mode on metal complex concentration. *Inorg. Chem.* **38**, 4486–4497 (1999).

66 Kelly, J. M. *et al.* Ruthenium polypyridyl complexes; their interaction with DNA and their role as sensitisers for its photocleavage. *J. Chem. Soc. Chem. Commun.* 1821–1823 (1987).

67   Jacquet, L., Davies, R. J. H., Kirsch-De Mesmaeker, A. & Kelly, J. M. Photoaddition of [Ru(tap)$_2$(bpy)]$^{2+}$ to DNA: A new Mode of covalent attachment of metal complexes to duplex DNA. *J. Am. Chem. Soc.* **119**, 11763–11768 (1997).

68   Friedman, A. E., Chambron, J. C., Sauvage, J. P., Turro, N. J. & Barton, J. K. A molecular light switch for DNA: [Ru(bpy)$_2$(dppz)]$^{2+}$. *J. Am. Chem. Soc.* **112**, 4960–4962 (1990).

69   Olofsson, J., Önfelt, B. & Lincoln, P. Three-state light switch of [Ru(phen)$_2$dppz]$^{2+}$: Distinct excited-state species with two, one, or no hydrogen bonds from solvent. *J. Phys. Chem. A* **108**, 4391–4398 (2004).

70   Hiort, C., Lincoln, P. & Norden, B. DNA binding of delta- and lambda-[Ru(phen)$_2$DPPZ]$^{2+}$. *J. Am. Chem. Soc.* **115**, 3448–3454 (1993).

71   Ortmans, I., Elias, B., Kelly, J. M., Moucheron, C. & Kirsch-DeMesmaeker, A. [Ru(TAP)$_2$(dppz)]$^{2+}$: a DNA intercalating complex, which luminesces strongly in water and undergoes photo-induced proton-coupled electron transfer with guanosine-5'-monophosphate. *Dalt. Trans.* 668–676 (2004).

72   Poynton, F. E. *et al.* The development of ruthenium(ii) polypyridyl complexes and conjugates for in vitro cellular and in vivo applications. *Chem. Soc. Rev.* **46**, 7706–7756 (2017).

73   Monro, S. *et al.* Transition metal complexes and photodynamic therapy from a tumor-centered approach: Challenges, opportunities, and highlights from the development of TLD1433. *Chem. Rev.* **119**, 797–828 (2019).

74   Lincoln, P. & Nordén, B. DNA binding geometries of ruthenium(II) complexes with 1,10-phenanthroline and 2,2'-bipyridine ligands studied with linear dichroism spectroscopy. Borderline cases of intercalation. *J. Phys. Chem. B* **102**, 9583–9594 (1998).

75   Holmlin, R. E., Stemp, E. D. A. & Barton, J. K. [Ru(phen)$_2$dppz]$^{2+}$ luminescence: Dependence on DNA sequences and groove-binding agents. *Inorg. Chem.* **37**, 29–34 (1998).

76   Dupureur, C. M. & Barton, J. K. Use of selective deuteration and 1H-NMR in demonstrating major groove binding of delta-[Ru(phen)$_2$dppz]$^{2+}$ to d(GTCGAC)$_2$. *J. Am. Chem. Soc.* **116**, 10286–10287 (1994).

77   Dupureur, C. M. & Barton, J. K. Structural studies of Λ- and Δ-[Ru(phen)$_2$dppz]$^{2+}$ bound to d(GTCGAC)$_2$: Characterization of enantioselective intercalation. *Inorg. Chem.* **36**, 33–43 (1997).

78   Tuite, E., Lincoln, P. & Nordén, B. Photophysical evidence That Δ- and Λ-[Ru(phen)$_2$(dppz)]$^{2+}$ intercalate DNA from the minor groove. *J. Am. Chem. Soc.* **119**, 239–240 (1997).

79   Greguric, A., Greguric, I. D., Hambley, T. W., Aldrich-Wright, J. R. & Collins, J. G. Minor groove intercalation of Δ-[Ru(Me$_2$phen)$_2$dppz]$^{2+}$ to the hexanucleotide d(GTCGAC)$_2$. *J. Chem. Soc. Dalton.* 849 (2002).

80   Hall, J. . *et al.* Structure determination of an intercalating ruthenium dipyridophenazine complex which kinks DNA by semiintercalation of a tetraazaphenanthrene ligand. *Proc. Natl. Acad. Sci. USA* **108**, 17610–17614 (2011).

81   McKinley, A. W., Andersson, J., Lincoln, P. & Tuite, E. M. DNA sequence and ancillary ligand

modulate the biexponential emission decay of intercalated [Ru(L)$_2$dppz]$^{2+}$ enantiomers. *Chemistry* **18**, 15142–50 (2012).

82   Hall, J. P. *et al.* Delta chirality ruthenium 'light-switch' complexes can bind in the minor groove of DNA with five different binding modes. *Nucleic Acids Res.* **44**, 9472-9482 (2016).

83   Liu, J. *et al.* Synthesis, characterization and antitumor activity of a series of polypyridyl complexes. *Met. Based. Drugs* **7**, 343–348 (2000).

84   Zhang, Z. *et al.* Ruthenium(II) complexes as apoptosis inducers by stabilizing c-myc G-quadruplex DNA. *Eur. J. Med. Chem.* **80**, 316–324 (2014).

85   Wang, Y.-C. *et al.* Dual topoisomerase I and II poisoning by chiral Ru(II) complexes containing 2-thiophenylimidazo[4,5-f][1,10]phenanthroline derivatives. *J. Inorg. Biochem.* **130**, 15–27 (2014).

86   Janaratne, T. K., Yadav, A., Ongeri, F. & MacDonnell, F. M. Preferential DNA cleavage under anaerobic conditions by a DNA-binding ruthenium dimer. *Inorg. Chem.* **46**, 3420–3422 (2007).

87   Yadav, A. *et al.* Regression of lung cancer by hypoxia-sensitizing ruthenium polypyridyl complexes. *Mol. Cancer. Ther.* **12**, 643–653 (2013).

88   Puckett, C. A. & Barton, J. K. Methods to explore cellular uptake of ruthenium complexes. *J. Am. Chem. Soc.* **129**, 46–47 (2007).

89   Svensson, F. R., Matson, M., Li, M. & Lincoln, P. Lipophilic ruthenium complexes with tuned cell membrane affinity and photoactivated uptake. *Biophys. Chem.* **149**, 102–106 (2010).

90   Gill, M. R. *et al.* A ruthenium(II) polypyridyl complex for direct imaging of DNA structure in living cells. *Nat. Chem.* **1**, 662-667 (2009).

91   Gill, M. R., Derrat, H., Smythe, C. G. W., Battaglia, G. & Thomas, J. A. Ruthenium(II) metallo-intercalators: DNA imaging and cytotoxicity. *ChemBioChem* **12**, 877–880 (2011).

92   Rajput, C., Rutkaite, R., Swanson, L., Haq, I. & Thomas, J. A. Dinuclear monointercalating RuII complexes that display high affinity binding to duplex and quadruplex DNA. *Chem. Eur. J.* **12**, 4611–4619 (2006).

93   Yao, J.-L., Gao, X., Sun, W., Shi, S. & Yao, T.-M. [Ru(bpy)$_2$dppz-idzo]$^{2+}$: a colorimetric molecular 'light switch' and powerful stabilizer for G-quadruplex DNA. *Dalt. Trans.* **42**, 5661–5672 (2013).

94   Hu, X. *et al.* Regulation of multi-factors (tail/loop/link/ions) for G-quadruplex enantioselectivity of Δ- and Λ- [Ru(bpy)$_2$(dppz-idzo)]$^{2+}$. *Dalt. Trans.* **47**, 5422–5430 (2018).

95   Sun, D. *et al.* Inhibition of human telomerase by a G-quadruplex interactive compound. *J. Med. Chem.* **40**, 2113–2116 (1997).

96   Read, M. A. *et al.* Molecular modeling studies on G-quadruplex complexes of telomerase inhibitors: Structure–activity relationships. *J. Med. Chem.* **42**, 4538–4546 (1999).

97   Harrison, R. J. *et al.* Trisubstituted acridine derivatives as potent and selective telomerase inhibitors. *J. Med. Chem.* **46**, 4463–4476 (2003).

98   Burger, A. M. *et al.* The G-quadruplex-interactive molecule BRACO-19 inhibits tumor growth,

consistent with telomere targeting and interference with telomerase function. *Cancer Res.* **65**, 1489–1496 (2005).

99 Read, M. A. & Neidle, S. Structural characterization of a guanine–quadruplex ligand complex. *Biochemistry* **39**, 13422–13432 (2000).

100 Schultze, P., Smith, F. W. & Feigon, J. Refined solution structure of the dimeric quadruplex formed from the Oxytricha telomeric oligonucleotide d(GGGGTTTTGGGG). *Structure* **2**, 221-233 (1994).

101 Campbell, N. H. *et al.* Selectivity in ligand recognition of G-quadruplex loops. *Biochemistry*. **48**, 1675–1680 (2009).

102 Campbell, N. H., Smith, D. L., Reszka, A. P., Neidle, S. & O'Hagan, D. Fluorine in medicinal chemistry: β-fluorination of peripheral pyrrolidines attached to acridine ligands affects their interactions with G-quadruplex DNA. *Org. Biomol. Chem.* **9**, 1328–1331 (2011).

103 Read, M. *et al.* Structure-based design of selective and potent G quadruplex-mediated telomerase inhibitors. *Proc. Natl. Acad. Sci.* **98**, 4844 LP – 4849 (2001).

104 Incles, C. M. *et al.* A G-quadruplex telomere targeting agent produces p16-associated senescence and chromosomal fusions in human prostate cancer cells. *Mol. Cancer Ther.* **3**, 1201-1206 (2004).

105 Gowan, S. M. *et al.* A G-quadruplex-interactive potent small-molecule inhibitor of telomerase exhibiting in vitro and in vivo antitumor activity. *Mol. Pharmacol.* **61**, 1154-1162 (2002).

106 Biswas, B., Kandpal, M. & Vivekanandan, P. A G-quadruplex motif in an envelope gene promoter regulates transcription and virion secretion in HBV genotype B. *Nucleic Acids Res.* **45**, 11268–11280 (2017).

107 Campbell, N. H., Parkinson, G. N., Reszka, A. P. & Neidle, S. Structural basis of DNA quadruplex recognition by an acridine drug. *J. Am. Chem. Soc.* **130**, 6722–6724 (2008).

108 Reed, J. E., Arnal, A. A., Neidle, S. & Vilar, R. Stabilization of G-quadruplex DNA and inhibition of telomerase activity by square-planar nickel(II) complexes. *J. Am. Chem. Soc.* **128**, 5992–5993 (2006).

109 Arnal, A. A, Benet-Buchholz, J., Neidle, S. & Vilar, R. Effects of metal coordination geometry on stabilization of human telomeric quadruplex DNA by square-planar and square-pyramidal metal complexes. *Inorg. Chem.* **47**, 11910–11919 (2008).

110 Lecarme, L. *et al.* Efficient inhibition of telomerase by nickel–salophen complexes. *ChemMedChem* **11**, 1133–1136 (2016).

111 Campbell, N. H. *et al.* Molecular basis of structure–activity relationships between salphen metal complexes and human telomeric DNA quadruplexes. *J. Med. Chem.* **55**, 209–222 (2012).

112 Abd Karim, N. H. *et al.* Salphen metal complexes as tunable G-quadruplex binders and optical probes. *RSC Adv.* **4**, 3355–3363 (2014).

113 Bertrand, B. & Casini, A. A golden future in medicinal inorganic chemistry: the promise of anticancer gold organometallic compounds. *Dalt. Trans.* **43**, 4209–4219 (2014).

114   Rubbiani, R., Salassa, L., de Almeida, A., Casini, A. & Ott, I. Cytotoxic gold(I) N-heterocyclic carbene complexes with phosphane ligands as potent enzyme inhibitors. *ChemMedChem* **9**, 1205–1210 (2014).

115   Messori, A. C. and L. Molecular mechanisms and proposed targets for selected anticancer gold compounds. *Curr. Top. Med. Chem.* **11**, 2647–2660 (2011).

116   Kim, Nam-Hoon; Park, Ho Jung; Oh, Mi-Kyung; Kim, I.S. Antiproliferative effect of gold(I) compound auranofin through inhibition of STAT3 and telomerase activity in MDA-MB 231 human breast cancer cells. *BMB Rep.* **46**, (2013).

117   Stefan, L. *et al.* Assessing the differential affinity of small molecules for noncanonical DNA structures. *ChemBioChem* **13**, 1905–1912 (2012).

118   Bertrand, B. *et al.* Caffeine-based gold(I) N-heterocyclic carbenes as possible anticancer agents: Synthesis and biological properties. *Inorg. Chem.* **53**, 2296–2303 (2014).

119   Bazzicalupi, C. *et al.* Determinants for tight and selective binding of a medicinal dicarbene gold(I) complex to a telomeric DNA G-quadruplex: A joint ESI MS and XRD investigation. *Angew. Chem. Int. Ed.* **55**, 4256–4259 (2016).

120   Guarra, F. *et al.* Interaction of a gold(I) dicarbene anticancer drug with human telomeric DNA G-quadruplex: solution and computationally aided X-ray diffraction analysis. *Dalt. Trans.* **47**, 16132–16138 (2018).

121   Micale, N. *et al.* Selected cytotoxic gold compounds cause significant inhibition of 20S proteasome catalytic activities. *J. Inorg. Biochem.* **141**, 79–82 (2014).

122   Gratteri, P. *et al.* Interactions of selected gold(III) complexes with DNA G quadruplexes. *Dalt. Trans.* **44**, 3633–3639 (2015).

123   Wirmer-bartoschek, J. *et al.* Solution NMR structure of a ligand/hybrid-2-G-quadruplex complex reveals rearrangements that affect ligand binding. *Angew. Chem. Int. Ed.* **56**, 7102-7106 (2017).

124   Wilson, T., Costa, P. J., Félix, V., Williamson, M. P. & Thomas, J. A. Structural studies on dinuclear ruthenium(II) complexes that bind diastereoselectively to an antiparallel folded human telomere sequence. *J. Med. Chem.* **56**, 8674–8683 (2013).

125   Wilson, T., Williamson, M. P. & Thomas, J. A. Differentiating quadruplexes: binding preferences of a luminescent dinuclear ruthenium(II) complex with four-stranded DNA structures. *Org. Biomol. Chem.* **8**, 2617–2621 (2010).

126   Zheng, X.-H. *et al.* Trigeminal star-like platinum complexes induce cancer cell senescence through quadruplex-mediated telomere dysfunction. *Chem. Commun.* **52**, 14101–14104 (2016).

127   Zhong, Y.-F. *et al.* A Platinum(II)-based photosensitive tripod as an effective photodynamic anticancer agent through DNA damage. *Chem. Eur. J.* **23**, 16442–16446 (2017).

128   Liu, W. *et al.* Solution structures of multiple G-quadruplex complexes induced by a platinum(II)-based tripod reveal dynamic binding. *Nat. Commun.* **9**, 3496 (2018).

2 Structural Studies of Substituted Ruthenium Polypyridyl Complexes

Contribution statement

Complex synthesis, enantiomeric purification, characterisation, crystallisation, data collection, structure solution, and data analysis were performed by Kane McQuaid.

Parts of the work presented in this chapter were featured in the following publication:

McQuaid, K., Hall, J.P., Brazier, J.A., Cardin, D.J., and Cardin, C.J. X-ray crystal structures show DNA stacking advantage of terminal nitrile substitution in Ru-dppz complexes. *Chem. Eur. J.* **24**, 15859, (2018)

The structures listed below were a direct output from the work presented here and were entered into the Protein Data Bank with the following identifications:

5NBE: Λ-[Ru(TAP)$_2$(dppz-11-CN)$^{2+}$ bound to d(TCGGCGCCGA)

6R6D: Λ-[Ru(TAP)$_2$(11,12-CN-dppz)$^{2+}$ bound to d(TCGGCGCCGA)

6RSO: Λ-[Ru(phen)$_2$(10-NO$_2$-dppz)$^{2+}$ bound to the oligonucleotide sequence d(TCGGCGCCGA)

6RSP: Λ-[Ru(phen)$_2$(11-NO$_2$-dppz)$^{2+}$ bound to the oligonucleotide sequence d(TCGGCGCCGA)

6GLD: Intercalation of Λ-[Ru(TAP)$_2$(11-Br-dppz)]$^{2+}$ bound to d(TCGGCGCCGA)

6G8S: Λ-[Ru(TAP)$_2$(11,12-CN-dppz)]$^{2+}$ bound to d(CCGGACCCGG/CCGGGTCCGG)

2.1 Introduction

Ru(II) complexes based on the [Ru(L)$_2$(dppz)]$^{2+}$ framework have shown significant promise as probes for nucleic acid structure and as site-specific damage agents.[1,2] Simple modification by way of distal ring substitution on the dppz has led to complexes that exhibit higher DNA binding affinity, significantly better structural specificity in luminescence response, and more effective $^1O_2$ generation for prospective use as photosensitisers in photodynamic therapies.[3,4] The structural consequences of such substitution upon intercalation are however relatively unexplored; previous investigations of methyl and chloro substituted complexes allude to the general conclusion that very little structural change is expected upon binding of simple derivatives.[5,6] The former study showed the effect of

alteration of the substitution pattern of methyl groups around the distal ring of dppz in the parent [Ru(TAP)$_2$(dppz)]$^{2+}$, and structurally characterised their interaction with d(TCGGCGCCGA). The authors noted that despite the near magnitude increase in binding constant, and the incorporated steric mass of the methyl groups, no discernible change in binding mode was observed when comparing the structure of Λ-[Ru(TAP)$_2$(11,12-Me-dppz)]$^{2+}$ to the parent complex bound to the DNA decamer. Interestingly however they did note that asymmetric substitution on the dppz always led to a single orientation of the complex in the intercalation cavity; that is the asymmetric substituent(s) were always ordered in one direction as opposed to an equal distribution of disorder across both possibilities. The logic for this preference is not inherently clear but was hypothesised to be electronic in nature since the addition of the mildly e$^-$ donating groups at the 11 position still generated a single orientation despite the absence of potential steric clashes with the DNA. In a separate study two derivatives were analysed for their binding efficacy and ability to cleave dsDNA.[7] The complexes studied, *rac*-[Ru(phen)$_2$(10-OH-dppz)]$^{2+}$ and *rac*-[Ru(phen)$_2$(10-NO$_2$-dppz)]$^{2+}$, contain either a resonance donor or resonance acceptor moiety respectively, significantly altering the electrostatic potential of the intercalating ligand. The authors found that the nitro species exhibited a five-fold increase in binding affinity when compared to the hydroxyl species; in addition, the nitro complex exhibited significantly more potent photocleavage properties, even doing so in the dark. The presence of a single binding mode is favourable for many reasons; one of these being the ability to interpret solution state measurements with more insight. As an example, using TRIR along with crystallographic data and the subsequent knowledge of a single binding mode, one particular investigation identified the precise guanosine that was reversibly reduced in a photoinduced electron transfer process.[8] A later study built on this; using time resolved IR and ps/ns transient absorption spectroscopy it was found that substitution of the guanine in the intercalation pocket (G$_9$) with inosine leads to a system far more susceptible to Ru(II) sensitised photooxidation, despite the lower sensitivity of the inosine base to oxidation.[9] In this case, the crystallographic coordinates showed that [Ru(TAP)$_2$(dppz)]$^{2+}$ could penetrate the intercalation site deeper when inosine was the adjacent base on the 3' side. This is due to the absence of 2-NH$_2$ on the Watson-Crick edge of guanine that protrudes into the minor groove, limiting the depth of intercalation. Such structural analysis allowed an informed assignment of binding mode in solution.[10]

In this chapter several complexes with substituted dppz ligands (figure 2.1) have been synthesised and subsequently crystallised in both the absence and presence of duplex forming oligonucleotides. A particular focus has been drawn to substituents with electron withdrawing properties, and with altering the position/number of substituents on the intercalating chromophore.

This study aims to investigate further the effect of substitution on the spatial intricacies of intercalation, but also to begin to make more general expectations for similar systems.
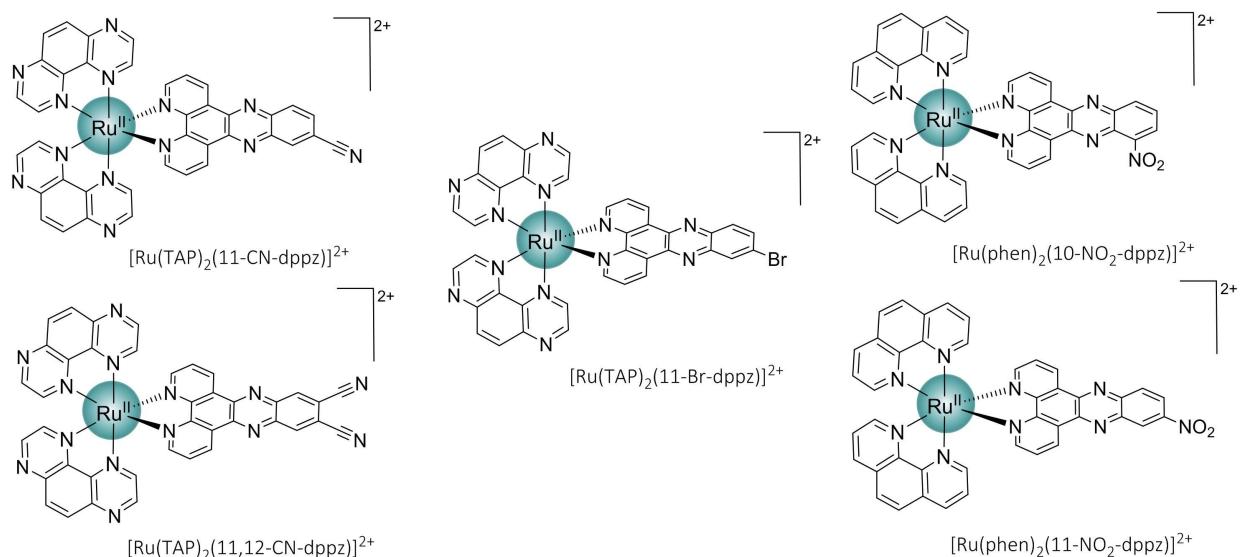
Figure 2.1 – Complexes described within this chapter and successfully crystallised with dsDNA.

2.2 Methodology

Oligonucleotides were purchased as double HPLC-purified solids from Eurogentec Ltd, and were used without further purification. Unless otherwise stated, all other materials and chemicals were sourced from Sigma Aldrich or Honeywell research chemicals.

2.2.1 Synthesis, characterisation, and enantiomeric separation

Synthesis of *rac*-[Ru(TAP)₂(11-CN-dppz)]·Cl₂, *rac*-[Ru(TAP)₂(11,12-CN-dppz)]·Cl₂, *rac*-[Ru(phen)₂(10-NO₂-dppz)]·Cl₂, and *rac*-[Ru(TAP)₂(11-NO₂-dppz)]·Cl₂ were performed using modifications on previously published methodology and are described in detail in chapter 7. Enantiomeric resolution of the nitrile derivatives is described in section 7.2.4 with their elution conditions, and subsequent circular dichroism spectra of the optically pure enantiomers displayed in figure A3.1.

2.2.2 Small molecule X-ray crystallography

Following purification *rac*-[Ru(TAP)₂(11-CN-dppz)]·Cl₂, Δ-[Ru(TAP)₂(11,12-CN-dppz)]·Cl₂, *rac*-[Ru(phen)₂(10-NO₂-dppz)]·Cl₂, and *rac*-[Ru(phen)₂(11-NO₂-dppz)]·Cl₂ were all metathesised to their hexafluorophosphate salt form by dropwise addition of a supersaturated aqueous solution of KPF₆ to the complexes dissolved in water.

2.2.2.1 nitrile derivatives

Crystals suitable for X-ray diffraction experiments were obtained by the vapour diffusion of diethyl ether into saturated solutions of *rac*-[Ru(TAP)$_2$(11-CN-dppz)]·(PF$_6$)$_2$ or Δ-[Ru(TAP)$_2$(11,12-CN-dppz)]·(PF$_6$)$_2$ in acetonitrile. Dark red rods were obtained after roughly 2 days at 291 K.

Data collections were performed on an Oxford Diffraction Gemini-S Ultra diffractometer fitted with a 135 mm Atlas CCD detector using graphite-monochromated Mo *K*α radiation (λ = 0.7107 Å) for [Ru(TAP)$_2$(11-CN-dppz)]·(PF$_6$)$_2$; and a Rigaku Synergy-S diffractometer with a HyPix-6000 pixel detector using Cu *K*α (λ = 1.5418 Å) for crystals of [Ru(TAP)$_2$(11,12-CN-dppz)]·(PF$_6$)$_2$; both crystals were cooled to 150 K using a stream of nitrogen. Screening, strategy determination, data collection, and data reduction/processing was achieved using the CrysAlisPro program package. Using the Olex2 software package, the structures were solved using either Superflip for [Ru(TAP)$_2$(11-CN-dppz)]·(PF$_6$)$_2$ or SHELXT for [Ru(TAP)$_2$(11,12-CN-dppz)]·(PF$_6$)$_2$ and refined using a weighted full-matrix least-squares minimisation against F^2 using the SHELXL package.[11–13] Hydrogen positions were first calculated then refined using a riding model. Subsequently, non-hydrogen atoms were refined anisotropically. Data collection and final refinement statistics are shown in table A2.1. Selected bond lengths and angles are given in table A2.2, and unit cell contents are shown in figures A2.3-5. Experimental data and refined structures were uploaded to the IUCr checkCIF server and no warnings above level C were reported.

2.2.2.2 nitro derivatives

Crystals suitable for X-ray diffraction experiments were obtained by the vapour diffusion of diethyl ether into saturated solutions of *rac*-[Ru(phen)$_2$(10-NO$_2$-dppz)]·(PF$_6$)$_2$ or *rac*-[Ru(phen)$_2$(11-NO$_2$-dppz)]·(PF$_6$)$_2$ in acetonitrile. Red rods grew after roughly two days at 277 K.

Data collection for both structures was performed on a Rigaku Synergy-S diffractometer with the diffraction detected with a HyPix-6000 pixel detector. Cu *K*α radiation (λ = 1.5418 Å) was used on crystal fragments cooled to 150 K using a stream of nitrogen gas. Data collection, reduction, and processing was performed using the CrysAlisPro program package. SHELXT was utilised in the Olex2 software suite to solve the structures and they were refined against F^2 using a full-matrix least-square minimisation in SHEXL. All non-hydrogen atoms were refined anisotropically, and hydrogen atoms were first placed in calculated positions and refined using a riding model. Data collection and final refinement statistics are shown in table A2.1. Selected bond lengths and angles are given in

table A2.3, and unit cell contents are shown in figures A2.6-8. Experimental data and refined structures were uploaded to the IUCr checkCIF server and no warnings above level C were reported.

2.2.2.3 aqueous crystallisations

Crystals of suitable quality for both enantiomers of $[Ru(TAP)_2(11\text{-}CN\text{-}dppz)]\cdot Cl_2$ were grown from Natrix HT screening conditions in crystallisation plates. $\Lambda\text{-}[Ru(TAP)_2(11\text{-}CN\text{-}dppz)]\cdot Cl_2$ was grown via vapour diffusion of water from a 4 μL drop containing 500 μM $\Lambda\text{-}[Ru(TAP)_2(11\text{-}CN\text{-}dppz)]\cdot Cl_2$, 62.5 μM $d(AG_3(T_2AG_3)_3)$, 150 mM KCl, 25 mM TRIS hydrochloride pH 7.5, 25 mM magnesium chloride, and 5 % v/v MPD. $\Delta\text{-}[Ru(TAP)_2(11\text{-}CN\text{-}dppz)]\cdot Cl_2$ was similarly grown by vapour diffusion from a sitting drop but did so in the presence of a different oligonucleotide. The drop contained 250 μM $\Delta\text{-}[Ru(TAP)_2(11\text{-}CN\text{-}dppz)]\cdot Cl_2$, 250 μM $d(TAG_3TTA)$, 150 mM KCl, 25 mM TRIS hydrochloride pH 7.5, 25 mM magnesium chloride, and 5 % v/v MPD.

Crystals were initially believed to be macromolecular (lighter colour and observed crystal birefringence) and were thus initially treated as such. Data collection was performed at Diamond Light Source Ltd. at beamline I03 on single, flash-cooled crystals using radiation of 0.557 Å wavelength. After screening, unit cell determination, and initial map fitting, the detector resolution limit was extended out to the maximum for collection (0.68 Å). All data were processed with the xia2 pipeline; indexing and scaling using XDS and XSCALE for the delta enantiomer dataset, and DIALS and Aimless for the lambda isomer.[14–18] The data diffracted isotropically to 1.42 and 1.49 Å respectively; lower than expected for a small molecule dataset. As a result, macromolecular techniques were used in data manipulations (note that direct methods or charge-flipping were not attempted). Large anomalous signal was observed for both sets (1.903 and 1.811 respectively); as such initial phases were calculated using the anomalous diffraction of ruthenium ($|\Delta f| \approx 7.5$ e$^-$) using HySS and Phaser-EP in the Phenix suite.[19,20] In both cases, two 100 % occupancy ruthenium atoms were found in the asymmetric units; subsequent model building was performed in WinCoot, using Phenix.refine to refine against the original data.[21,22] Final R_{cryst}/R_{free} of 0.170/0.210 and 0.164/0.187 respectively were observed for the data. Data collection and final refinement statistics are displayed in table A2.4. As a result of the high e.s.d's associated with the lower resolution structures, bond length tables have been omitted, but derived polyatomic parameters will be discussed below.

2.2.3 Macromolecular X-ray crystallography

2.2.3.1 crystallisation of Λ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ and Λ-[Ru(TAP)$_2$(11,12-CN-dppz)]$^{2+}$ with d(TCGGCGCCGA)

2.2.3.1.1 crystallisation conditions

Crystals containing the oligonucleotide d(TCGGCGCCGA)$_2$ and ligand Λ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ were grown via the vapour diffusion method from sitting drops at 291 K. Crystallisation was observed in a number of conditions from the Nucleic Acid Mini-Screen from Hampton Research, where the best diffracting example came from a 8 µL drop containing, as final concentrations; 250 µM d(TCGGCGCCGA), 750 µM *rac*-[Ru(TAP)$_2$(11-CN-dppz)]·Cl$_2$, 7.5 % v/v MPD, 30 mM pH 7 sodium cacodylate, 9 mM spermine tetrahydrochloride, 60 mM KCl and 15 mM BaCl$_2$, all equilibrated against 500 µL of 35 % v/v MPD. Orange/red rods grew following roughly 3 weeks of incubation at 291 K.

Similarly, crystals containing the oligonucleotide d(TCGGCGCCGA)$_2$ and ligand Λ-[Ru(TAP)$_2$(11,12-CN-dppz)]$^{2+}$ were grown via the vapour diffusion method from sitting drops at 291 K. Crystallisation was observed from an 8 µL drop containing, as final concentrations; 250 µM d(TCGGCGCCGA), 500 µM *rac*-[Ru(TAP)$_2$(11,12-CN-dppz)]·Cl$_2$, 7.5 % v/v MPD, 30 mM pH 7 sodium cacodylate, 9 mM spermine tetrahydrochloride, 60 mM KCl and 15 mM BaCl$_2$, all equilibrated against 500 µL of 35 % v/v MPD. Orange/red rhombohedra grew following roughly 6 weeks of incubation at 291 K.

2.2.3.1.2 data collection, refinement, and analysis

Diffraction data for the Λ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$/d(TCGGCGCCGA) system were collected from a single, liquid nitrogen flash-cooled crystal fragment at 100 K on beamline I02 at Diamond Light Source Ltd. Data were processed with xia2, using XDS and XSCALE to integrate and merge peaks from all collected images; yielding 5452 unique reflections at a maximum resolution of 1.51 Å. The structure was solved using single-wavelength anomalous dispersion (SAD), using the anomalous diffraction of barium, with the SHELXC/D/E package.[14,15,23] Chemical refinement restraints for the ligand [Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ were calculated using the eLBOW/REEL programs in the PHENIX software package using a molecular model derived from a mixture of small molecule X-ray crystallographic and molecular modelling data.[22,24] The small molecule restraint model was built in Avogadro and its geometry optimised within Gaussian09 using the B3LYP/6-31G(d',p') functional basis set and the LAN2LZ pseudopotential whilst constraining the ruthenium geometry and Ru-N distance to the small molecule crystallographic data.[25] The model was built by hand, using Wincoot, and refined against the original data using Refmac 5.0 in the CCP4i suite.[21,26,27] 5 % of reflections were reserved for

the R_{free} set. The final model has an R_{cryst}/R_{free} of 0.165/0.190 and has been deposited in the Protein Data Bank with ID 5NBE. Full conformational analysis of the asymmetric unit is displayed in table A2.5.

Data for Λ-[Ru(TAP)$_2$(11,12-CN-dppz)]·Cl$_2$ with d(TCGGCGCCGA) were collected from a single, flash-cooled crystal at 100 K on beamline I03 at Diamond Light Source, Ltd at a wavelength of 0.9763 Å. Data were pipeline processed with xia2, using DIALS and Aimless to integrate and merge peaks from the collected images; 6316 unique observations were yielded at a maximum resolution of 1.84 Å, and a mid-slope of anomalous probability of 1.348 was observed.[14,17,18] As such, a joint SAD-MR approach, using a partial model from 5NBE and the anomalous scattering of barium, was employed to phase the data using Phaser-MR/EP in the Phenix software suite.[20,22] Structural restraints were calculated from a mixed X-ray-DFT optimised structure using the eLBOW/REEL programs, as above.[24] The subsequent model was fitted by hand, using Wincoot, and refined against the original data using Phenix.refine, yielding a final model R_{cryst}/R_{free} of 0.179/0.187, with 5.1 % of the reflections being reserved for the R_{free} set.[21,28] The structure has been deposited in the Protein Data Bank with the ID 6R6D. Full conformational analysis of the asymmetric unit is displayed in table A2.6.

2.2.3.2 crystallisation of Λ-[Ru(phen)$_2$(10-NO$_2$-dppz)]$^{2+}$ and Λ-[Ru(phen)$_2$(11-NO2-dppz)]$^{2+}$ with d(TCGGCGCCGA)

2.2.3.2.1 crystallisation conditions

Diffraction quality crystals of the complex of oligonucleotide d(TCGGCGCCGA)$_2$ with ligand Λ-[Ru(phen)$_2$(10-NO$_2$-dppz)]$^{2+}$ were grown via the vapour diffusion method from sitting drops at 291 K. Crystallisation was observed in a number of conditions from the Nucleic Acid Mini-Screen from Hampton Research, where the best diffracting example came from a 8 µL drop containing, as final concentrations; 250 µM d(TCGGCGCCGA), 625 µM rac-[Ru(phen)$_2$(10-NO$_2$-dppz)]·Cl$_2$, 7.5 % v/v MPD, 30 mM pH 7 sodium cacodylate, 9 mM spermine tetrahydrochloride, 60 mM KCl and 15 mM BaCl$_2$, all equilibrated against 500 µL of 35 % v/v MPD. Orange/red rods grew following roughly 4 weeks of incubation at 291 K.

Similarly, crystals containing the oligonucleotide d(TCGGCGCCGA)$_2$ and ligand Λ-[Ru(phen)$_2$(11-NO$_2$-dppz)]$^{2+}$ were grown via the vapour diffusion method from sitting drops at 291 K. Crystallisation was observed from an 8 µL drop containing; 250 µM d(TCGGCGCCGA), 500 µM rac-[Ru(phen)$_2$(11-NO$_2$-dppz)]·Cl$_2$, 7.5 % v/v MPD, 30 mM pH 7 sodium cacodylate, 9 mM spermine tetrahydrochloride, 60 mM KCl and 15 mM BaCl$_2$, all equilibrated against 500 µL of 35 % v/v MPD. Orange/red rhombohedra grew following roughly 7 weeks of incubation at 291 K

Diffraction images were collected using X-rays of 0.9763 Å wavelength, for both crystal systems, on beamline I03 at Diamond Light Source Ltd. from single crystals flash cooled in liquid nitrogen. Data were processed using xia2 indexing with peaks found from all images, utilising XDS and XSCALE for the system containing [Ru(phen)$_2$(10-NO$_2$-dppz)]$^{2+}$; and using DIALS and Aimless to index and scale the frames for the system containing [Ru(phen)$_2$(11-NO$_2$-dppz)]$^{2+}$.[14,15,17,18] The experiments yielded 5004 and 5650 unique reflections, at an overall resolution of 1.97 and 1.91 Å for the 10-NO$_2$ and 11-NO$_2$ systems respectively. Both structures were solved by SAD methodology in the Phaser-EP package using the anomalous diffraction of barium.[20] Chemical refinement restraints for both Λ-[Ru(phen)$_2$(10-NO$_2$-dppz)]$^{2+}$ and Λ-[Ru(phen)$_2$(11-NO$_2$-dppz)]$^{2+}$ were calculated using the eLBOW/REEL programs in the PHENIX software package using a molecular model derived from a mixture of small molecule X-ray crystallographic and molecular modelling data (as described before).[24,28] The crystallographic models were subsequently built from the substructure sites in Wincoot and refined against the original data using Phenix.refine, retaining 4.67 % of the reflections for the R_{free} set. The final models have R_{cryst}/R_{free} of 0.200/0.241 and 0.190/0.207 respectively, and were published on the PDB under codes 6RSO and 6RSP.

2.2.3.3 crystallisation of Λ-[Ru(TAP)$_2$(11-Br-dppz)]$^{2+}$ with d(TCGGCGCCGA), and crystallisation of Λ-[Ru(TAP)$_2$(11,12-CN-dppz)]$^{2+}$ with d(CCGGACCCGG/CCGGGTCCGG)

2.2.3.3.1 crystallisation conditions

Crystals containing d(TCGGCGCCGA) and Λ-[Ru(TAP)$_2$(11-Br-dppz)]$^{2+}$ were grown by vapour diffusion from sitting drops at 291 K. Data were collected from a crystal grown from a 8 µL drop containing, as final concentrations; 250 µM d(TCGGCGCCGA), 625 µM *rac*-[Ru(TAP)$_2$(11-Br-dppz)]·Cl$_2$, 7.5 % v/v MPD, 30 mM pH 7 sodium cacodylate, 9 mM spermine tetrahydrochloride, 60 mM KCl, and 15 mM BaCl$_2$; all equilibrated against 500 µL of 35 % v/v MPD. Orange/red rods grew following roughly 3 weeks of incubation at 291 K.

Crystals containing d(CCGGACCCGG/CCGGGTCCGG) and Λ-[Ru(TAP)$_2$(11,12-CN-dppz)]$^{2+}$ were grown from sitting drops by vapour diffusion at 291 K. The best diffracting crystal grew from a 8 µL drop containing 125 µM of a pre annealed mixture of d(CCGGACCCGG) and d(CCGGGTCCGG), 125 µM *rac*-[Ru(TAP)$_2$(11,12-CN-dppz)]$^{2+}$, 7.5 % v/v MPD, 30 mM pH 7 sodium cacodylate, 9 mM spermine tetrahydrochloride, 60 mM KCl, and 15 mM BaCl$_2$; all equilibrated against 500 µL of 35 % v/v MPD. Large orange/red rhombohedra grew following 4 weeks of incubation at 291 K. The subsequent data were collected on a crystal fragment broken from the larger crystal.

2.2.3.3.2 data collection, refinement, and analysis

Diffraction data for d(TCGGCGCCGA) and Λ-[Ru(TAP)$_2$(11-Br-dppz)]$^{2+}$ containing crystals were collected from a single, nitrogen flash-cooled crystal fragment at 100 K on beamline I03 at Diamond Light Source Ltd using 0.9763 Å X-rays. Data were processed using DIALS and Aimless to integrate and merge peaks from all collected images, yielding 16,081 unique reflections.[14,17,18] Phases were solved using single-wavelength anomalous dispersion (SAD), using the anomalous signal of barium using the SHELXC/D/E package.[23] The model was built by hand using WinCoot, and refined against the original data using Phaser.refine in the Phenix software package.[20,21,28] Ligand restraints were calculated using eLBOW from the Phenix package.[24] The final model has a R_{cryst}/R_{free} of 0.15/0.16, with 5 % of the reflections reserved for the R_{free}. The model and structure factors are deposited in the Protein Data Bank with accession ID 6GLD. Full conformational analysis of the asymmetric unit is displayed in table A2.9.

Diffraction data for crystals containing Λ-[Ru(TAP)$_2$(11,12-CN-dppz)]$^{2+}$ and d(CCGGACCCGG/CCGGGTCCGG) were collected from a single crystal fragment at 100 K on beamline I03 at Diamond Light Source Ltd. Data were processed with xia2, using DIALS and Aimless to integrate and merge peaks from all collected images, yielding 9018 unique reflections.[14,17,18] The structure was solved using MR-SAD within Phaser in the Phenix software package, using the anomalous signal of barium and a starting model from 5IP8.[20,28] Ligand restraints were calculated using eLBOW in the Phenix package.[24] 5 % of reflections were reserved for the R_{free} set. The final model has an R_{cryst}/R_{free} of 0.19/0.20 and has been deposited in the Protein Data Bank with accession ID 6G8S. Full conformational analysis of the asymmetric unit is displayed in table A2.10.

2.2.3.4 data collection parameters and refinement statistics

Table 2.1 – *Crystallisation parameters and refinement statistics.*

Crystallisation Parameters

Sequence	TCGGCGCCGA	TCGGCGCCGA	CCGGACCCGG/CCGGTCCCGG
Complex	[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$	[Ru(TAP)$_2$(11,12-CN-dppz)]$^{2+}$	[Ru(TAP)$_2$(11,12-CN-dppz)]$^{2+}$
Crystal Morphology	Orthorhombic Rods	Orthorhombic	Orthorhombic
Growth Temp (K)	291	291	291
Crystal Size (µm)	30x30x120	50x30x100	60x30x100
Growth Time	3 weeks	6 weeks	4 weeks

Data Collection

Beamline	I02	I03	I04
X-Ray Wavelength (Å)	0.9795	0.9763	0.9795
Beamsize (µm)	100x20	20x20	32x20
Exposure Time (s)	0.04	0.05	0.05
Nº Images/oscillation (º)	900/0.10	3600/0.10	3600/0.10
Space Group	$P\,4_32_12$	$P\,4_32_12$	$P\,4_3$
Cell Dimensions a, b, c (Å); α, β, γ (º)	47.88, 47.88, 29.14; 90, 90, 90	46.70, 46.70, 34.31; 90, 90, 90	47.25, 47.25, 33.98; 90, 90, 90

Data Processing

Resolution (Å)	33.86 - 1.51 (1.55 - 1.51)	34.31 - 1.84 (1.87 - 1.84)	27.59 - 1.66 (1.69 - 1.66)
R_{merge}	0.078 (1.318)	0.097 (2.665)	0.038 (0.858)
R_{meas}	0.099 (1.477)	0.100 (2.721)	0.053 (0.942)
R_{pim}	0.039 (0.588)	0.021 (0.547)	0.015 (0.261)
Nº Observations	33,788 (2340)	84,821 (4412)	116,655 (5984)
Nº Unique Observations	5452 (394)	6316 (180)	9018 (460)
I/σI	11.4 (1.4)	17.7 (1.40)	23.85 (2.39)
$CC_{1/2}$	0.998 (0.593)	0.999 (0.748)	0.999 (0.695)
Completeness (%)	95.74 (96.80)	100 (98.90)	100 (100)
Multiplicity	6.2 (5.9)	23.5 (24.5)	12.9 (13)
Mid-slope	1.178	1.348	1.234

* Outer Shell Statistics Shown in Parentheses

Refinement

Phase Solution Method	SAD	SAD-MR	SAD-MR
Resolution	24.89 (1.51)	33.03 (1.84)	23.62 (1.66)
Nº Reflections	5156	6316	17,361
R_{work}/R_{free}	0.165/0.190	0.179/0.187	0.187/0.203
Nº Atoms			
DNA	202	202	448
Metal Complex	53	75	55
Water	65	18	68
Average B Factors (Å^2)			
DNA	23.367	54.94	54.11
Metal Complex	19.5	48.60	43.08
Water	34.72	47.86	53.3
rmsd			
Bond Lengths (Å)	0.016	0.011	0.005
Bond Angles (º)	2.869	1.149	0.939
PDB ID	**5NBE**	**6R6D**	**6G8S**

Table 2.2 *– Crystallisation parameters and refinement statistics.*

Crystallisation Parameters			
Sequence	TCGGCGCCGA	TCGGCGCCGA	TCGGCGCCGA
Complex	$[Ru(phen)_2(10\text{-}NO_2\text{-}dppz)]^{2+}$	$[Ru(phen)_2(11\text{-}NO_2\text{-}dppz)]^{2+}$	$[Ru(TAP)_2(11\text{-}Br\text{-}dppz)]^{2+}$
Crystal Morphology	Orthorhombic	Orthorhombic	Orthorhombic
Growth Temperature (K)	291	291	291
Crystal Size (µm)	90x60x160	80x40x140	100x100x60
Growth Time	4 weeks	7 weeks	3 weeks
Data Collection			
Beamline	I03	I03	I03
X-Ray Wavelength (Å)	0.9763	0.9763	0.9763
Beamsize (µm)	80x20	80x20	80x20
Exposure Time (s)	0.05	0.05	0.05
Nº Images/Oscillation (º)	3600/0.10	3600/0.10	3600/0.10
Space Group	$P\,4_3$	$P\,4_3$	$P\,4_32_12$
Cell Dimensions a, b, c (Å); α, β, γ (º)	46.73, 46.73, 31.99; 90, 90, 90	47.09, 47.09, 32.45; 90, 90, 90	42.16, 42.16, 39.17; 90, 90, 90
Data Processing			
Resolution (Å)	46.73 - 1.97 (2.00 - 1.97)*	33.31 - 1.91 (1.94 - 1.91)	39.09 - 1.08 (2.88 - 1.06)
R_{merge}	0.049 (2.637)	0.057 (1.830)	0.151 (0.847)
R_{meas}	0.051 (2.761)	0.059 (1.905)	0.155 (0.897)
R_{pim}	0.014 (0.804)	0.016 (0.526)	0.032 (0.282)
Nº Observations	64,486 (2838)	72,703 (3398)	336,166 (5414)
Nº Unique Observations	9539 (952)‡	5650 (261)	16081 (596)
I/σI	26.9 (1.0)	19.0 (1.5)	12.1 (1.4)
$CC_{1/2}$	1.000 (0.334)	1.000 (0.586)	0.996 (0.406)
Completeness (%)	100 (99.60)	100 (100)	97.4 (74.5)
Multiplicity	12.9 (11.5)	12.9 (13.0)	20.9 (9.1)
Mid-slope	1.494	1.21	0.896
* Outer Shell Statistics Shown in Parentheses			
Refinement			
Phase Solution Method	SAD	SAD	SAD
Resolution	26.40 (1.97)	33.31 (1.91)	29.80 (1.06)
Nº Reflections	9536 (949)	10800	29,454
R_{work}/R_{free}	0.200/0.241	0.188/0.211	0.149/0.158
Nº Atoms			
DNA	404	404	324
Metal Complex	158	158	77
Water	28	40	88
Average B Factors (Å²)			
DNA	73.97	61.71	23.39
Metal Complex	74.02	57.62	14.49
Water	72.55	64.29	32.90
rmsd			
Bond Lengths (Å)	0.010	0.009	0.010
Bond Angles (o)	1.166	1.127	1.407
PDB ID	**6RSO**	**6RSP**	**6GLD**

2.3 Results

2.3.1 Small molecule crystallisations

The crystal structures derived from the two nitrile derivatives (shown in figure 2.2) are very similar in overall structure; addition of an extra nitrile group has little effect on the geometry of the ruthenium centre, where both exhibit almost identical distorted octahedral geometries with Ru-N distances in the range of 2.04-2.07 Å. Incorporation of an additional substituent does however lengthen the average nitrile C≡N bond distance (along with lowering the C-C length) on the nitriles from 1.12 to 1.14 Å. Selected bond lengths, angles and torsions are summarised in table A2.2.

a

Λ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$

b

Δ-[Ru(TAP)$_2$(11,12-CN-dppz)]$^{2+}$
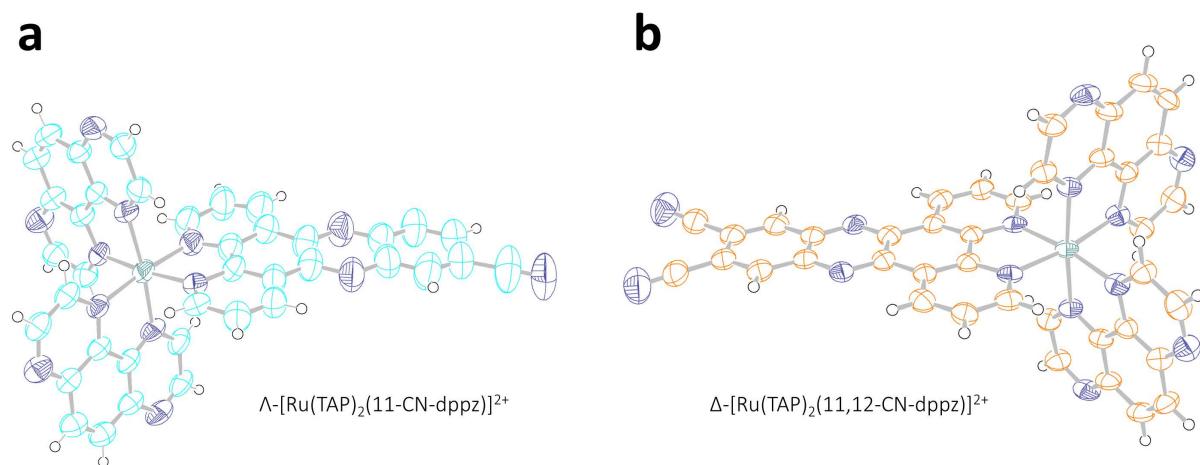
Figure 2.2 *– ORTEP representations of the crystal structures of (a) Λ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$; and (b) Δ-Ru(TAP)$_2$(11,12-CN-dppz)]$^{2+}$. Counter ions and solvent are omitted for clarity. Thermal ellipsoids are represented at a 50 % probability. Carbons are shown as cyan or orange, nitrogen as royal blue, and ruthenium as teal.*

Crystal structures were also obtained for the nitro derivatives (figure 2.3). Distal ring position of the nitro group has very little effect on the global geometry of the complexes; table A8.2 summarises the bond angles and lengths. The crystal structures show no deviation in angle or bond lengths from the distorted octahedral ruthenium environment observed in the nitrile derivative structures (Ru-N distances in the range of 2.06-2.08 Å). Of note however is the out of plane torsions exhibited from substitution at the 10 position. In relation to the rest of the dppz ligand, the 10-NO$_2$ moiety is skewed out of plane by 41°; this is presumably a result of unfavourable electrostatic interactions with the adjacent pyrazine nitrogen. This is not observed with the 11-NO$_2$ however, where the group is planar with respect to the rest of the dppz group. DFT level geometry optimisations also closely mirror the crystal structures (figures A2.10-14), with calculations presenting the same level of out-of-plane propeller twist caused by bay area lone pair repulsion. As such, substitution at the 10- position leads to frontier molecular orbitals that are higher in energy than the 11 substituted counterpart ($|\Delta HOMO|$ = 0.052 eV) (figure A2.13-14). Such twist could however impart

additional specificity not seen in the 11-NO$_2$-dppz system by way of polar contacts that are comparatively closer to the adjacent base pairs, and by way of the twisted moiety creating a tighter binding pocket.
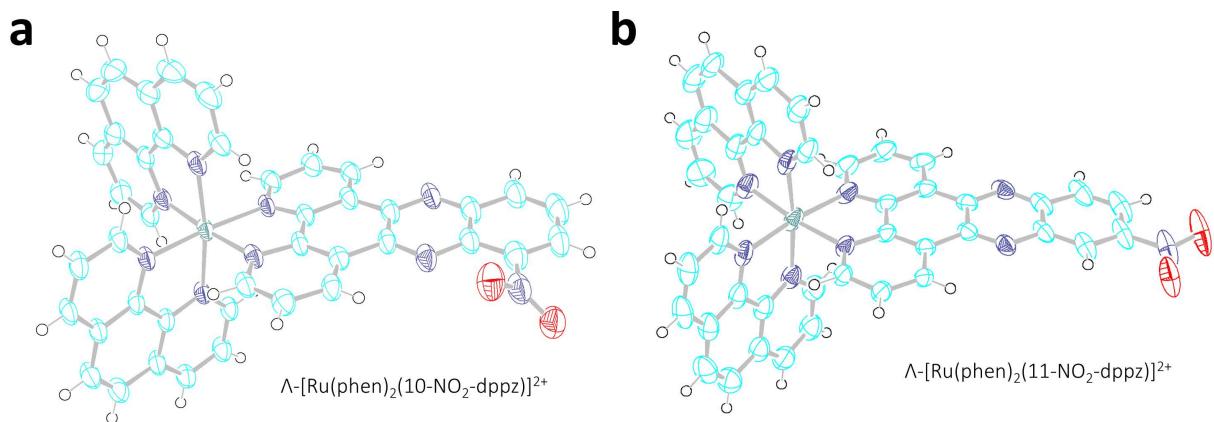
Figure 2.3 – *ORTEP representations of the crystal structures of (a) Λ-[Ru(phen)$_2$(10-NO$_2$-dppz)]$^{2+}$; and (b) Λ -[Ru(phen)$_2$(11-NO$_2$-dppz)]$^{2+}$. Counter ions and solvent are omitted for clarity. Thermal ellipsoids are represented at a 50 % probability. Carbons are shown as cyan or orange, nitrogen as royal blue, and ruthenium as teal.*

The structures above were of course intentionally crystallised from organic media, however, as part of a wider study attempting to co-crystallise ligands with DNA, and due to the high stoichiometries of ligand used in these studies, unintentional crystallisation of the metal complex alone occurred as can occasionally happen. These systems usually crystallise analogously to the organic systems, albeit often in different space groups. However, the two enantiomeric structures presented herein exhibit a stark deviation from this, arranging in such a fashion that is only analogous to a double helix. Interestingly both crystal systems were grown from conditions containing quadruplex DNA and have occurred with a handful of different sequences but never with duplex DNA; the best diffracting examples have been reported (figure 2.4).
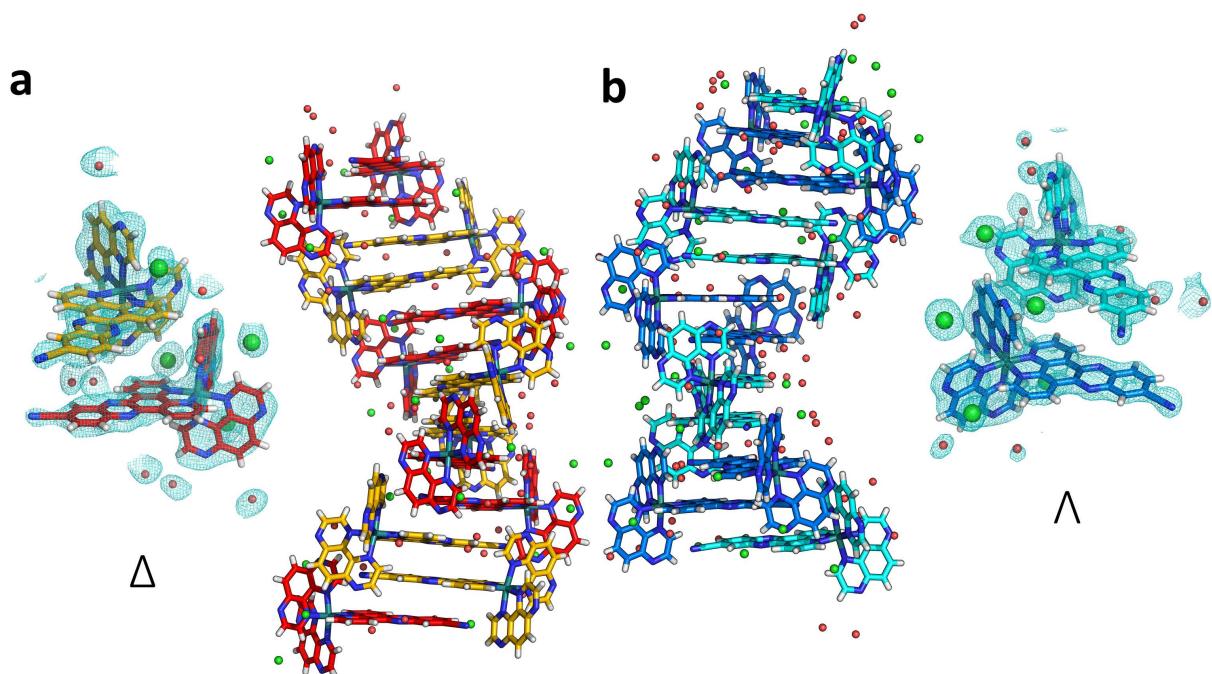
Figure 2.4 – _Crystal structures of (a) Δ-[Ru(TAP)$_2$(11-CN-dppz]·Cl$_2$, and (b) Λ-[Ru(TAP)$_2$(11-CN-dppz)]·Cl$_2$ highlighting the DNA-like helical structures that the complexes form in crystal state; both possessing opposite helical twist directions. Inset shows the asymmetric units with electron density (shown at 1 σ, 0.5 e Å^{-3}) consisting of two complexes._

The structures, although solely of small molecule content, did not diffract to the expected higher resolutions and as such were solved using macromolecular techniques, namely phased by the anomalous diffraction of ruthenium. Both structures are for all intents and purposes isostructural except for opposite axial screw directions (Δ = (-), Λ = (+) coiling); both containing analogous π-stacking environments and local water structure around the complexes. The asymmetric units contain two complexes, charge balanced by ordered chlorine atoms. Expansion of this highlights an interlocked π-stack with the asymmetric dppz groups ordered with complete directional preference towards two ancillary phen ligands; presumably a result of favourable π-sandwich interaction. More intriguingly however is the manner in which the separate enantiomers form extended helical structures in the c direction of the cell, reminiscent of the enantiomeric forms of DNA (figure 2.5). Due to the asymmetry of the stacking interaction this helical morphology is accompanied by the presence of major and minor grooves whose width closely imitates that of canonical B-DNA (24 and 13 Å respectively). These grooves, especially the minor, similarly accommodate a rich shell of ordered crystalline water. Loosely treated as pseudo-DNA, the helix exhibits a helical pitch of 53 Å over 16 complexes to complete a turn; with an average twist angle of 21.2°. Overall, the parameters most closely resemble a partially unwound B-DNA duplex, and to an extent replicates the base step parameters observed at steps adjacent to classical intercalation sites.
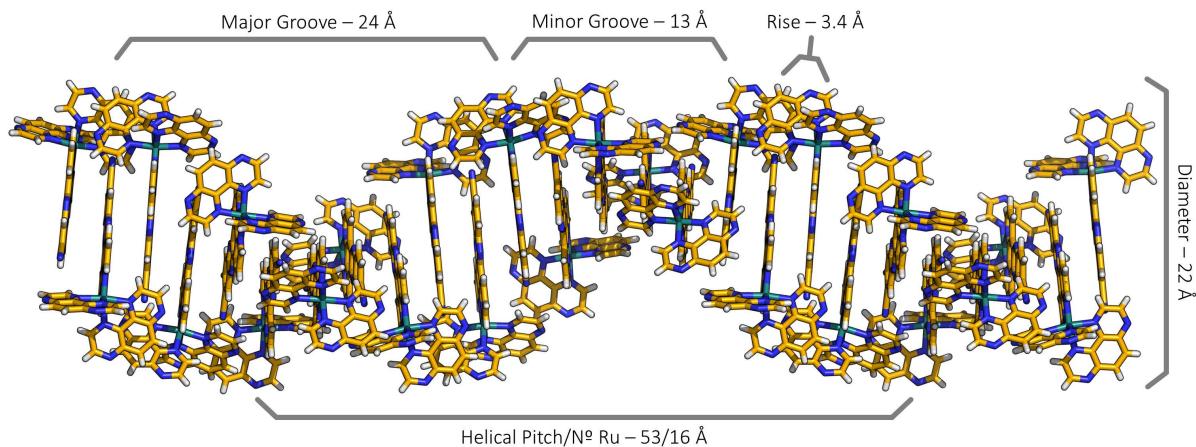
Figure 2.5 – *Crystal structure of Δ-[Ru(TAP)₂(11-CN-dppz)]·Cl₂ elongated in the c axial direction; highlighting its resemblance to natural DNA. Measured helical parameters are reported.*

Extension of the packing in all crystallographic axes highlights an additional feature that is also present in many macromolecular crystal structures (figure 2.6). Viewing down the c axis of the supercell, and concurrently the helical axis of the arrangement, large solvent channels are present (18 Å in width). The large voids, which accommodate a coordination sphere of ordered water and chlorine counterions, is also likely a root of the overall lower resolution of diffraction.
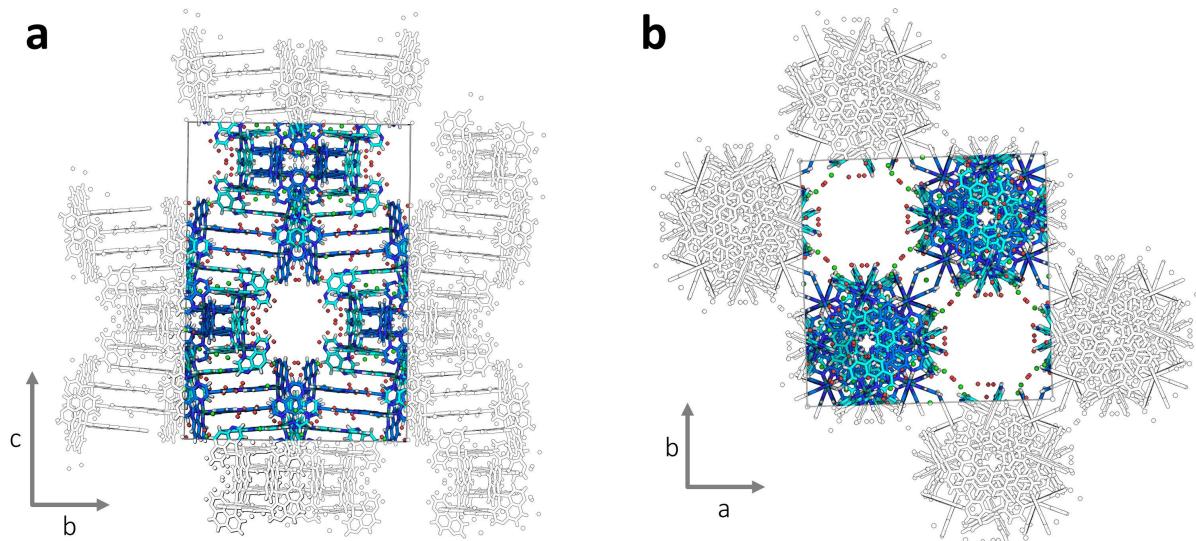
Figure 2.6 – *Crystal structure of Λ-[Ru(TAP)₂(11-CN-dppz)]·2Cl⁻ grown from H₂O; shown down the (a) the a axis, and (b) the c axis. Note the helical nature of assembly, and the large solvent channels formed down the c axis.*

Although of little tangible context, the structures highlight an interesting juxtaposition with the biological target molecules. In a more abstract mindset, the structure with delta absolute configuration, which closely resembles the biologically relevant right-handed duplex, provides some logic to the observed enantiomeric preference in affinity for delta isomers to B-DNA since the complex

emulates the curvature of natural DNA. As previously mentioned, the reported structures were grown from solutions containing quadruplex DNA intended to be crystallised with the ligands but evidently only the complexes crystallised. When repeated in the absence of DNA, crystallisation is not observed; could the DNA be templating the crystallisation? or is it more likely that the DNA aids in the precipitation via increased saturation of the growth medium? A preference for self-aggregation in the presence of the nucleic acid may imply an unsatisfactory binding interaction. However, crystallisation was only observed in conditions with high stoichiometric ratios (≥3:1 Ru:DNA) potentially in excess of the binding saturation point of the DNA. As such the noted crystallisations may have occurred alongside less 'crystallisable' ligand-DNA complexes.

2.3.2 Macromolecular crystallisations

In the presence of Group II ions, the oligonucleotide d(TCGGCGCCGA) inherently forms a stacked-X Holliday junction in the crystal. Initially the sequence was crystallised to elucidate the binding modes of ruthenium complexes to the motif, however, previous studies have shown that in the presence of $[Ru(TAP)_2(dppz)]^{2+}$ the sequence forms a well-matched B-DNA.[29] A total of six discrete ligand-DNA systems are described herein, detailing the duplex binding modes of five derivatised ruthenium species with distal ring substitutions to decamer sequences (primarily d(TCGGCGCCGA) except one with d(CCGGACCCGG/CCGGGTCCGG)). Despite their inherent differences the systems are, for the most part, almost isostructural with each other in regards to the global characteristics but often deviate in local binding environments (see figure A2.9 for a superimposition of the global structures). Further information such as the complete nucleic acid conformational analyses can be found in the appendix (tables A2.5-10).

a   Λ-[Ru(TAP)₂(11-CN-dppz)]²⁺

b   Λ-[Ru(TAP)₂(11,12-CN-dppz)]²⁺

c   Λ-[Ru(phen)₂(10-NO₂-dppz)]²⁺

d   Λ-[Ru(phen)₂(11-NO₂-dppz)]²⁺
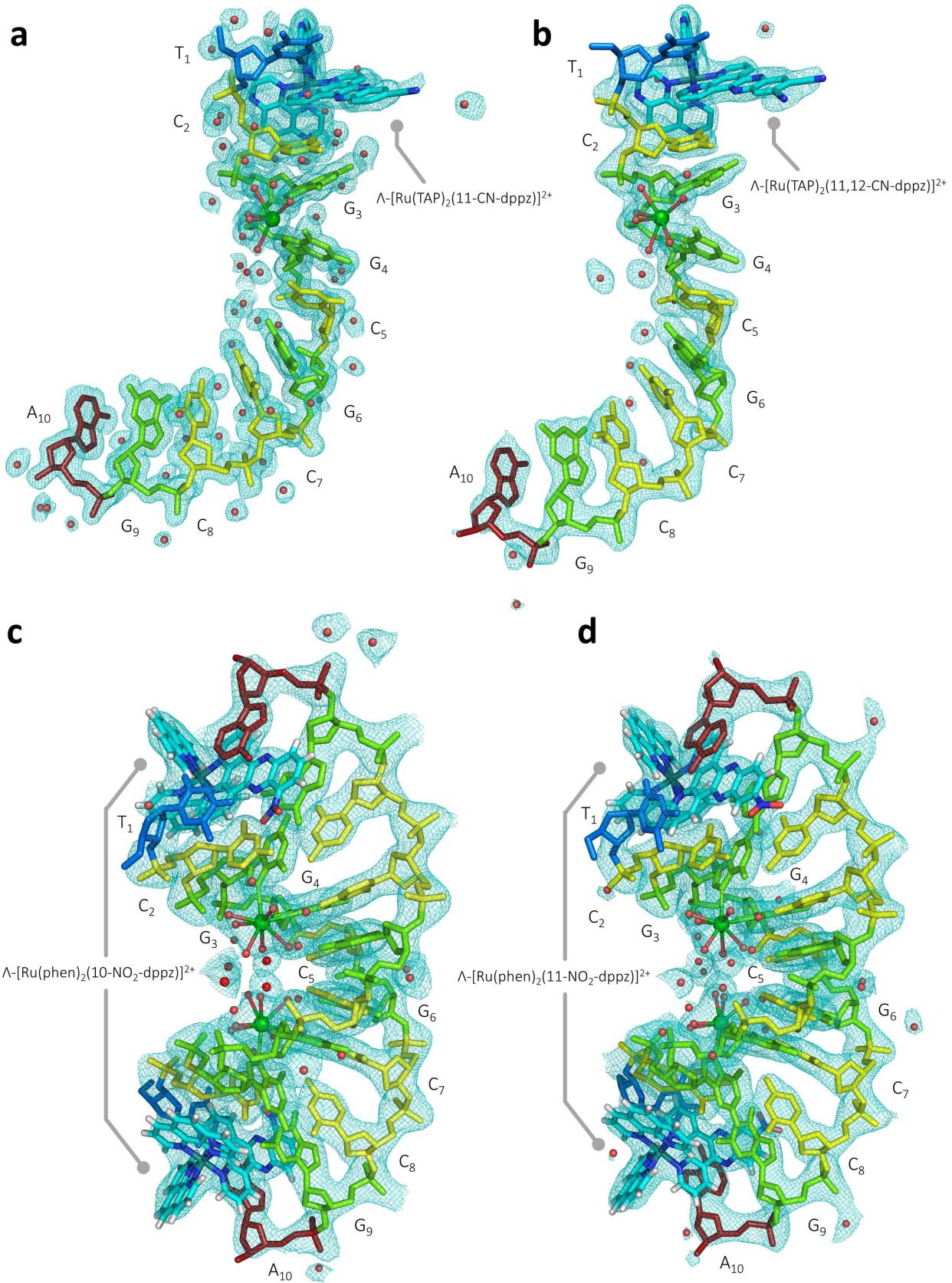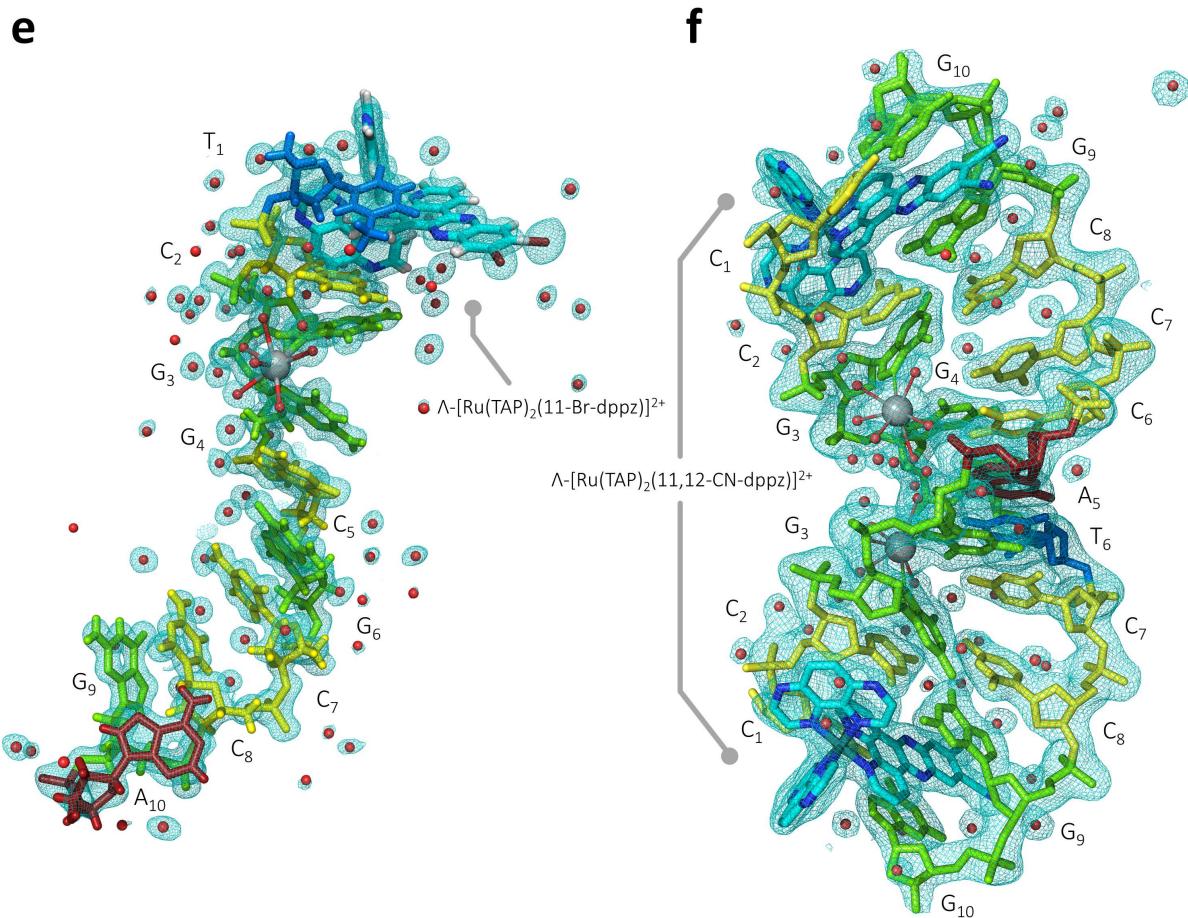
Figure 2.7 – *Asymmetric units containing the crystallographic models and 2F$_o$-F$_c$ electron density maps of (a) Λ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$, (b) Λ-[Ru(TAP)$_2$(11,12-CN-dppz)]$^{2+}$, (c) Λ-[Ru(phen)$_2$(10-NO$_2$-dppz)]$^{2+}$, (d) Λ-[Ru(phen)$_2$(11-NO$_2$-dppz)]$^{2+}$, and (e) Λ-[Ru(TAP)$_2$(11-Br-dppz)]$^{2+}$bound to d(TCGGCGCCGA); whereas (f) shows the interaction of Λ-[Ru(TAP)$_2$(11,12-CN-dppz)]$^{2+}$, with d(CCGGACCCGG/CCGGGTCCGG). Nucleosides are coloured based on the nucleic acid database standards such that: Thymine is blue, cytosine is yellow, guanine is green, and adenine is red. Electron density is contoured at the 1σ level (0.2883 and 0.2388 e/Å^3 respectively).*

2.3.2.1 summary of nitrile derivatives bound to d(TCGGCGCCGA)

Overall, the two nitrile derivatives have complexed with DNA in an isostructural manner. The asymmetric units (shown in figure 2.7a-b) contain a single strand of the decamer oligonucleotide d(TCGGCGCCGA), one copy of either Λ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ or Λ-[Ru(TAP)$_2$(11,12-CN-dppz)]$^{2+}$, one coordinated barium cation, and 65 or 18 ordered water molecules respectively. Stoichiometrically, the complexes interact at a ratio of 1:1 to single strand; generation of the crystallographic symmetry mates presents the full biological unit as a decamer duplex with two intercalating ruthenium complexes. Additionally, the biological units are electrostatically associated through a secondary semi-intercalative binding mode, yielding two discrete non-covalent binding modes per ruthenium cation.

2.3.2.2 summary of nitro derivatives bound to d(TCGGCGCCGA)

Both crystal systems containing the nitro derivatives were solved in the $P4_3$ space group despite the use of the same palindromic sequence used with the nitrile derivatives. When solved in the higher symmetry $P4_32_12$ the electron density maps included chemically implausible artefacts and the R factors did not behave accordingly during refinements; i.e. did not drop despite good agreement between observed F_o-F_c and $2F_o$-F_c maps. The asymmetric units (shown in figures 2.7c-d) contain two strands of the decamer oligonucleotide d(TCGGCGCCGA), two copies of either Λ-[Ru(phen)$_2$(10-NO$_2$-dppz)]$^{2+}$ or Λ-[Ru(phen)$_2$(11-NO$_2$-dppz)]$^{2+}$, two coordinated barium cations bound to the G$_3$ and bifurcated to the G$_4$ of each strand, and 28 or 40 ordered water molecules respectively. Despite this absence in two-fold crystallographic symmetry, the two discrete intercalation sites are almost identical and are markedly similar to the nitrile structures.

2.3.2.3 summary of [Ru(TAP)$_2$(11-Br-dppz)]$^{2+}$ bound to d(TCGGCGCCGA)

Crystals obtained from the bromine derivative diffracted to better than 1.1 Å resolution. The resulting structure has an asymmetric unit containing one strand of the oligonucleotide d(TCGGCGCCGA), one copy of Λ-[Ru(TAP)$_2$(11-Br-dppz)]$^{2+}$, one coordinated barium cation bound to the G$_3$ residue and bifurcated across the G$_4$ residue, and 88 ordered waters across two shells of hydration (figure 2.7e). The structure differs from the other reported structures here in that the terminal A$_{10}$ residue flips out and forms a reverse Watson-Crick pair with a neighbouring duplex as is observed in previously published structures.

2.3.2.4 summary of [Ru(TAP)$_2$(11,12-CN-dppz)]$^{2+}$ bound to d(CCGGACCCGG/CCGGGTCCGG)

The symmetrical dinitrile derivative was successfully crystallised with an oligonucleotide containing the asymmetric central pair step A$_5$C$_6$:G$_5$T$_6$. The structures, which due to its inherent asymmetry was solved in $P4_3$, has an asymmetric unit containing the duplex of d(CCGGACCCGG)/d(CCGGGTCCGG), two copies of Λ-[Ru(TAP)$_2$(11,12-CN-dppz)]$^{2+}$ bound at the terminal base steps, two fully hydrated barium cations bound to G$_3$ and bifurcated to G$_4$, and 68 ordered waters (figure 2.7f). The model contains evidence of backbone conformational disorder and as such was modelled as a 50/50 mixture of B$_I$ and B$_{II}$ conformations around the penultimate phosphates.
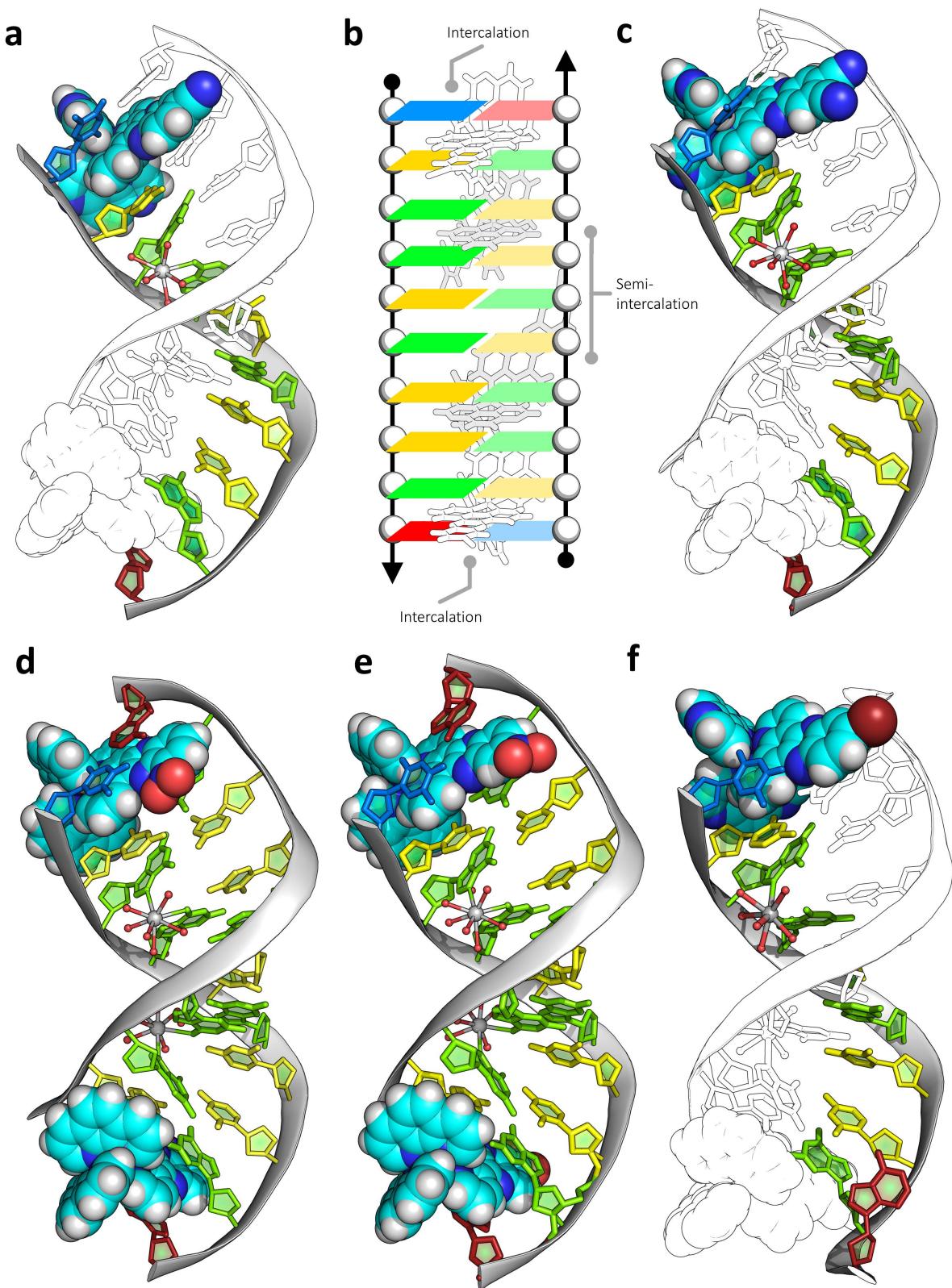
Figure 2.8 – Crystallographic models of the interactions between (a) Λ-[Ru(TAP)₂(11-CN-dppz)]²⁺, (c) Λ-[Ru(TAP)₂(11,12-CN-dppz)]²⁺, (d) Λ-[Ru(phen)₂(10-NO₂-dppz)]²⁺, (e) Λ-[Ru(phen)₂(11-NO₂-dppz)]²⁺, and (f) Λ-[Ru(TAP)₂(11-Br-dppz)]²⁺ with d(TCGGCGCCGA); with (b) showing the general structure and complex binding modes seen in all structures. Asymmetric units are coloured whilst crystallographically related sister strands are outlined in black and white. Nucleosides are coloured based on the nucleic acid database standards such that: Thymine is blue, cytosine is yellow, guanine is green, and adenine is red. Complexes are shown with carbon in cyan, nitrogen in dark blue, oxygen in white, bromine in crimson, and hydrogen in white.

2.3.2.5 intercalation sites and binding cavities

Primarily, the lambda isomers of the complexes are observed to intercalate, from the minor groove, into the terminal 5'-T_1C_2:G_9A_{10}-3' steps of the decamer. Two complexes bind per duplex, intercalating through the minor groove, and are related symmetrically by a two-fold axis that runs orthogonal to the helical axis. Complete binding cavities are conserved in the site of intercalation such that only the two strands in the biological unit form the primary binding environment. As such this mode is termed 'classical' intercalation, an observation that is not seen in the structure containing the parent molecule which incorporates a third neighbouring strand to complete a terminal reverse Watson-Crick base pair. In accordance with previous intercalatory modes, the interaction occurs asymmetrically in relation to the nucleobases, overlapping primarily with the purine side of the cavity. As such, the derivatised dppz moieties are offset by roughly 70° in relation to the H-bonding plane of the $C_2 \cdot G_9$ pair, and by approximately 10° with respect to the terminal $T_1 \cdot A_{10}$ pair. Locally, the intercalative mode causes the T·A and C·G pairs to buckle away from the chromophore (10 and 7° respectively for 11-CN-dppz) in addition to moderate unwinding of local base pairs (average B-DNA is 36° whereas these structures exhibit anywhere from 10-20° of twist at the terminal pair). The intercalatory environment is defined by two major characteristics; (a) a preference for stacking predominantly in between the purine G_9/A_{10} base step over the T_1/C_2 pyrimidine side, presumably to increase favourable π-overlap (figure 2.9a); and by (b) a secondary π-interaction between an ancillary TAP ligand and the 3' sugar ring of C_2, a stacking interaction which could potentially be regulated by ribose pucker or substitution (i.e. 2'-OH in RNA). It is this contact between the TAP ligand and the 3'-deoxyribose that determines the orientation of the dppz ligands in relation to the base pairing plane. Although the orientation is controlled by this π-stacking interaction, the depth of the intercalatory mode is limited by the 2-NH_2 moiety on the G_9 base, resulting in a shallow 'canted' intercalation (figure 2.9b). The depth of intercalation and subsequent envelopment of the π-surfaces of such complexes has been linked to the magnitude of response in their luminescence upon binding in solution. In all these cases the canted mode allows one side of the dppz and one ancillary ligand to be accessible to solvent whilst the rest of molecule is sufficiently shielded.
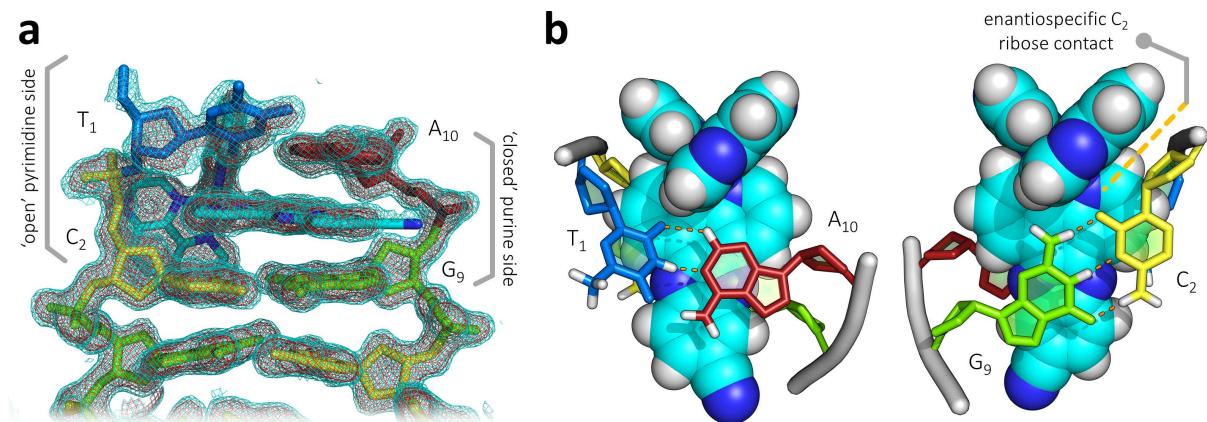
Figure 2.9 – *Crystal structures of (a) Λ-[Ru(TAP)₂(dppz)]²⁺ (PDB: 3QRN) and (b) Λ-[Ru(TAP)₂(11-CN-dppz)]²⁺ (PDB: 5NBE) bound to the duplex forming sequence d(TCGGCGCCGA), highlighting the differences in intercalation cavity completeness. Symmetry related strand is shown in black outline highlighting how the terminal adenine (A₁₀) flips out of the duplex base stack, whereas the A₁₀ in (b) is anchored, forming a full closed cavity and gauche γ dihedral angle.*

The overall intercalation binding cavity is composed of two sides; one of purine composition, and one of pyrimidine composition. Accommodation of the intercalated ligand requires the backbone to distort and conform to the added rise incorporated into the stack (around 7 Å for most whereas canonical B-DNA has a rise of 3.32 Å), thus the torsional angles around the cavity contain information of the effects of intercalation on DNA morphology. In the present examples, the asymmetry of the binding mode also materialises in an asymmetric binding cavity; that is that the complexes tend to elicit an opening of one side of the base pair step. This opening can be described by the γ dihedral angle of the terminal base steps, where in the case of the monosubstituted nitrile complex, the pyrimidine side is opened considerably at 189° (*trans* conformation) at the T_1/C_2 step and is closed on the G_9/A_{10} side with a γ of 63°. The disubstituted complex however is closed on both sides of the cavity, with gauche dihedrals of γ on both (70.4° and 100.1° for the purine and pyrimidine sides respectively in the same structure). Despite also containing complete intercalation cavities, in both structures containing the nitro substituents this asymmetry is reversed compared to the nitrile systems, and is similar to the parent system. The structure containing the mildly electron withdrawing bromo group does not contain a complete cavity but has in fact induced a closing of both sides of the cavity despite the flipped out A_{10}, yielding gauche dihedrals (55.6 and 61.2° for the purine and pyrimidine side respectively).

In addition, two sites of pseudo intercalation are observed between one ancillary TAP/phen ligand of each complex with $G_3G_4:C_7C_8$ base pair steps on neighbouring crystallographic symmetry mates. This interaction non-covalently cross-links orthogonal duplexes in the crystal lattice. Stabilised by coordinated Ba²⁺ ions, these semi-intercalation sites kink the DNA backbone by 46-53° in the structures (shown by the local base step roll angles) and likely accounts for the enantioselectivity of

the overall crystal packing. In previous studies the level of hydration in the crystal has been shown to alter the magnitude of this kink and the subsequent distance between barium cations.[30] All of these structures have Ba-Ba distances close to 8.0 Å, suggesting an intermediate degree of hydration which can also be seen in the shortening of the c axial unit cell dimension (c = 29.14 Å for 11-CN-dppz). The bariums, in all cases, are coordinated in the step to the G_3 and G_4 nucleobases of the base pair step and are fully hydrated by ordered water. The semi-intercalated phen/TAP in all cases is stacked preferentially on the G_4, lying almost parallel to the residue (figure 2.10).
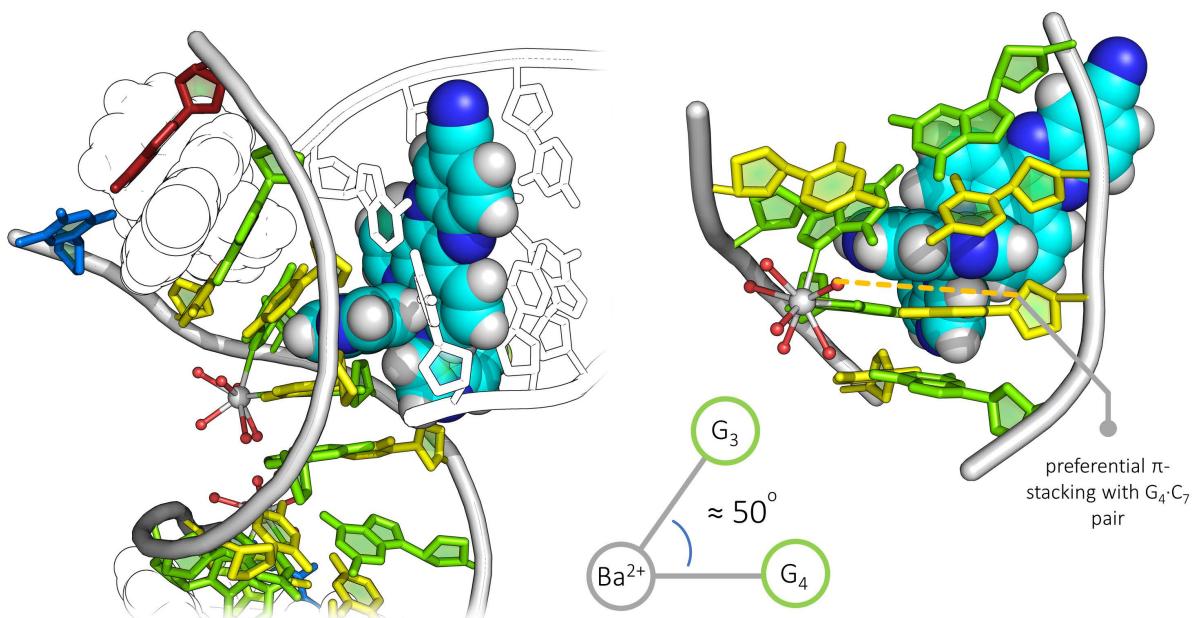
Figure 2.10 – *The semi-intercalation site observed in all structures presented. One ancillary ligand (phen/TAP) is seen to be pseudo-intercalated between G_3G_4:C_6C_7 preferentially stacking on the G_4·C_7 pair, kinking the duplex by roughly 50º. This kink is stabilised by a bound/bifurcated hydrated barium cation. White outlines represent symmetrically equivalent units.*

The directional preference of the asymmetric 11-CN-dppz is observed to be split between two orientations in the electron density map; one directed towards the major groove, and the other directed towards the phosphate backbone. This partial disorder materialises as a 0.64 occupancy directed towards the phosphate backbone on the purine side of the cavity, and a 0.33 occupancy towards the major groove; such disorder patterns are observed in other structures with terminal substitution on the distal ring.[5,6] Mirroring these previously noted patterns, structures containing bromo and nitro substitutions at the same position are also divided across the two rotationally related positions despite the added steric bulk and/or rotational freedom. Conversely, substitution of nitro in the 10 position is 100 % ordered towards the major groove, as was observed with the previously reported methyl substitution.[5] Chemical crystallography revealed the torsional differences between the two nitro moieties in the structures of *rac*-[Ru(phen)$_2$(10-NO$_2$-dppz)]$^{2+}$ and *rac*-[Ru(phen)$_2$(11-NO$_2$-dppz)]$^{2+}$ where when in the 11- position the nitro group is coplanar with the

dppz group whereas when in the 10- position the group is rotated about the C-N bond by roughly 45°. Unfortunately, such precision in atom positions was not possible in the macromolecular structures due to poor electron density maps regarding the oxygen positions. This may be a result, in part, to the directional ambiguity noted in the 11-NO$_2$-dppz structure, but also potentially a larger and more disordered range of motion of the nitro group in the comparatively hydrated 10-NO$_2$-dppz structure.

2.4 Discussion

2.4.1 Complete intercalation cavities

As seen in previous crystallisations with the sequence, the complexes crystallise with complete enantiomeric specificity, from the racemate mixtures. Previously unobserved however, is the deviation in terminal intercalation cavity completeness. In all previous structures involving the sequence d(TCGGCGCCGA) with ruthenium species the terminal base step is incomplete, this is such that the terminal A$_{10}$ base flips out of the base stack and forms a reverse Watson-Crick base pair with a T$_1$ base on a crystallographic symmetry mate (figure 2.11). This additional point of contact between asymmetric units could be a contributing factor to the higher resolution datasets collected with such systems, however, this could be considered a crystallographic artefact and as such make extrapolating binding detail to the dilute solution state more spurious. As such, forming a complete cavity allows for a more representative view of the binding environment outside of the condensed state.
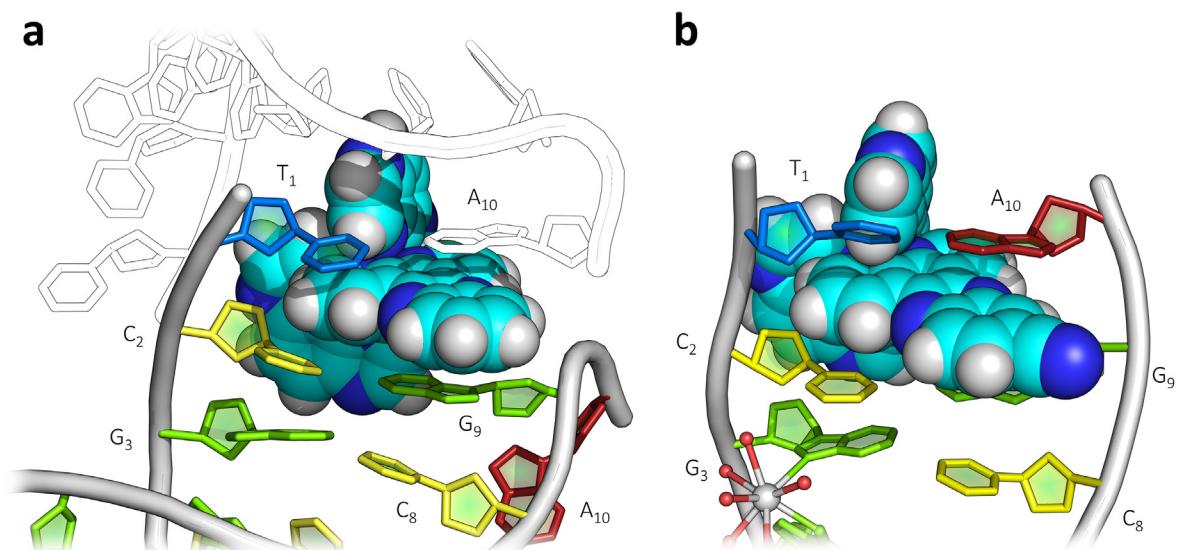
Figure 2.11 *– Crystal structures of (a) Λ-[Ru(TAP)$_2$(dppz)]$^{2+}$ (PDB: 3QRN) and (b) Λ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ (PDB: 5NBE) bound to the duplex forming sequence d(TCGGCGCCGA), highlighting the differences in intercalation cavity completeness. Symmetry related strand is shown in black outline highlighting how the terminal adenine (A$_{10}$) flips out of the duplex base stack, whereas the A$_{10}$ in (b) is anchored, forming a full closed cavity and gauche γ dihedral angle.*

This closure of the intercalation cavities with the reported species suggests that more favourable π-stacking interactions are occurring between the fragile terminal pair and the intercalated ligand. Interestingly, this has only been observed with complexes that contain strong electron withdrawing substituents on the distal ring of the dppz moiety; doing so with complexes incorporating nitrile and nitro groups but not the bromide (or with the previously reported parent/methyl/chloride). Nucleobases are inherently electron rich and, due to the electrostatic repulsions present with direct overlap of populated π orbitals, sandwich and parallel-displaced stacking formations favour intercalation of less negative π quadrupoles. As such, the incorporation of strong electron withdrawing groups could partially polarise the dppz, relocalising π-electron density away from the major interacting π-surface, providing a more attractive interaction. Calculating the RESP partial charges on the dppz atoms (not including substituents) from the DFT level calculations shows that the average partial charge increases as more electron withdrawing groups are added to the parent ligand ($\bar{x}$ = +0.035 and +0.0375 for dppz and 11-CN-dppz complexes respectively). This potentially rationalises the structural features observed herein, but also yields the implication that by altering the π-quadrupole of the intercalating ligand, a certain degree of control of the binding efficacy could also be imparted. Indeed, previous studies have found that the inclusion of a nitro group can increase the intrinsic binding constant to dsDNA by over 4 times when compared to the electron donating hydroxyl group in the 10- position.[7] DFT studies presented here support this claim and suggest substitution of strong electron withdrawing groups in all distal ring locations will further enhance binding. DFT derived molecular orbitals for the four arrangements in [Ru(phen)$_2$(10/11/10,11-X-dppz)]$^{2+}$ (where X = H, CN, or NO$_2$) are shown in figures 2.12 and figure A2.10-14, and display the effect of substitution on frontier molecular orbital localisation. In particular, the LUMO of the substituted species are all lower in energy than the parent and unlike the parent are localised on the intercalating ligand, suggesting a more favourable HOMO$_{nucleobase}$-LUMO$_{dppz}$ overlap. In addition to the relevant LUMO energies of the intercalatory ligand, the planarity of ligand and the localisation of the frontier orbitals all suggest that the electron deficient derivatives would bind with higher affinity when compared to the parent; where localisation of the frontier orbitals are localised predominantly on the ancillary ligands. Methylation of the complex in similar positions has shown to increase the binding strength also, where this has been linked to entropically favoured hydrophobicity of the subsequent binding pocket. X-ray crystallographic studies support this; however the same magnitude of cavity completeness and stabilisation as seen here was not observed.[5] Observations such as these suggest that by the direct modification of the π-framework, binding agents could be tailored to target weaker bound base steps or base mismatches/mutations by the fine tuning of not just the area/shape of π-surface but by the electronic potential of the surface. In addition, such derivatisation, aside from

providing additional polar contacts, steric specificities etc, has been shown to often drastically increase cellular uptake and heighten the potency of photosensitisers.[31,32]

[Ru(phen)$_2$(dppz)]$^{2+}$ [Ru(phen)$_2$(11-CN-dppz)]$^{2+}$ [Ru(phen)$_2$(11,12-CN-dppz)]$^{2+}$
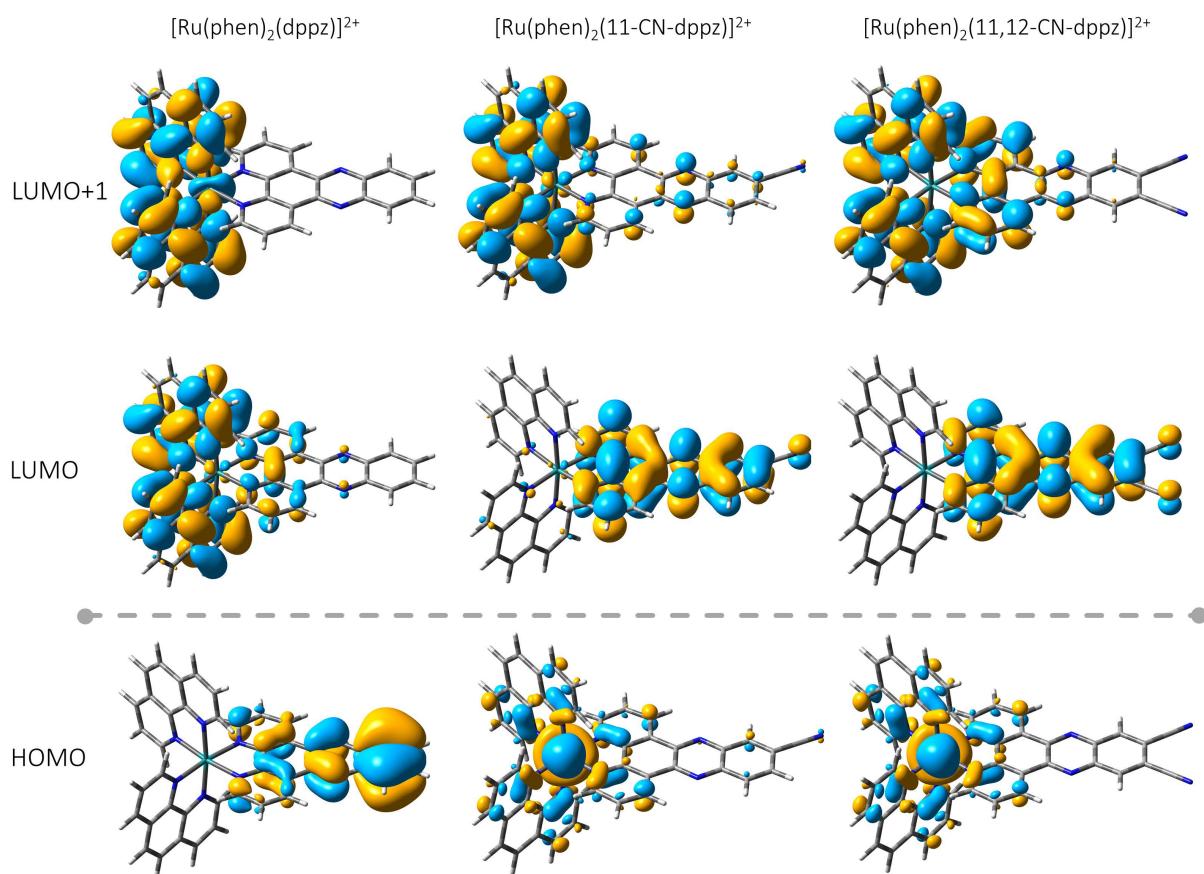
Figure 2.12 – *Selected frontier molecular orbitals calculated at the B3LYP/LAN2LZ DFT level for [Ru(phen)$_2$(dppz)]$^{2+}$ and the nitrile derivatives investigated here. A larger range of FMOs and FMOs for the above and for the nitro species' can be found in figure A2.10-14.*

2.4.2 Asymmetrical DNA torsions

Intercalation of the complexes tends to be accompanied by sizable distortions in the natural torsions of the DNA base steps adjacent to the binding site. The binding cavities in these cases tend to form asymmetrically, effectively favouring the opening of one side of the cavity to promote better overlap with the ligand at the expense of skewing the opposing side of the pair step. This opening or closing of the step can best be described using the γ dihedral angle; a summary is presented in figure 2.13. In the case of all previously studied interactions, this asymmetry is skewed towards closing the T$_1$/C$_2$ step and opening the G$_9$/A$_{10}$ side, all yielding T$_1$/C$_2$ γ dihedrals in the gauche range (57-60°). However, in the case for the two nitrile containing structures, the asymmetry swaps sides or both sides of the cavity close. This conformational switch is not observed with the nitro containing complexes however, where original side preference is conserved despite the complete cavities. For comparison, the intercalation of Actinomycin D into the GC:GC base pair steps of d(ATGCTGCAT)

imparts local antiperiplanar γ dihedrals around 180° across all sites; this is comparable to the backbone opening observed in the intercalation of C·C pairs into CC:CC base pair steps in i-motif structures and is a clear departure from the common gauche γ angles in canonical B-DNA.[33,34] Previous structural investigation of *rac*-[Ru(phen)$_2$(dppz)]$^{2+}$ bound to the hexamer d(ATGCAT) also presented such asymmetric cavities for both enantiomers and this could be related directly to their luminescence response due to the subsequent solvent accessibility to the enantiomers.[35] Interestingly, inspection of the 11-Br-dppz-DNA complex yields the first observation of both the purine and pyrimidine side exhibiting gauche γ dihedrals in an incomplete cavity, with a favourable substitutional interaction between the C3'-OH and the bromine atom potentially explaining this additional closure. The complexes presented here are not light-switch complexes (a result of either containing TAP or electron withdrawing sinks such as nitrile/nitro groups) but if they were substituted for their isoelectronic hydrocarbon equivalents then it could be expected that the complexes forming complete cavities would also have the larger luminescence response.

Complex	PDB Accession Nº		γ (T$_1$/C$_2$)	γ (G$_9$/A$_{10}$)
Λ-[Ru(TAP)$_2$(dppz)]$^{2+}$	3QRN		61.7	182.5
Λ-[Ru(TAP)$_2$(10-Me-dppz)]$^{2+}$	4MJ9		58.4	187.2
Λ-[Ru(TAP)$_2$(11-Me-dppz)]$^{2+}$	4X18		59.7	189.5
Λ-[Ru(TAP)$_2$(10,12-Me-dppz)]$^{2+}$	4X1A		59.9	190.7
Λ-[Ru(TAP)$_2$(11,12-Me-dppz)]$^{2+}$	4E8S		59.0	186.5
Λ-[Ru(TAP)$_2$(11-Cl-dppz)]$^{2+}$	4III		57.7	194.0
Λ-[Ru(TAP)$_2$(11-Br-dppz)]$^{2+}$	6GLD	*	55.6	61.2
Λ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$	5NBE	*	188.9	63.2
Λ-[Ru(TAP)$_2$(11,12-CN-dppz)]$^{2+}$	6R6D	*	70.4	100.4
Λ-[Ru(phen)$_2$(10-NO$_2$-dppz)]$^{2+}$	6RSO	*	67.2	179.3
Λ-[Ru(phen)$_2$(11-NO$_2$-dppz)]$^{2+}$	6RSP	*	53.0	168.4
* presented herein				

Figure 2.13 – *(a) asymmetrical cavity of 5NBE highlighting the γ torsions in the DNA backbone; (b) table summarising the γ torsion angles described in this chapter alongside the previously reported substituted complexes as sourced from the PDB.*

Shallow intercalation (determined by cytosine ribose contact and NH$_2$ on guanine) allows for the semi-intercalative binding mode to occur in these systems, in which these non-covalent cross-links act as the 'molecular glue' between asymmetric units. Anything that greatly affects this depth of intercalation would lead to poorer semi-intercalation. As such, modifications to the sequence such that the G$_9$ were replaced would almost certainly destroy the crystal packing in this system. Interestingly, previous crystallisation attempts with d(TAGGCGCCTA) were unsuccessful, and so were attempts using bpy as an ancillary ligand. This selectivity could be attributed to the binding occurring through the minor groove, since in B-DNA conformation, the adenine and guanines are isostructural when viewed from the major groove.

2.4.3 dppz orientation

The major directional orientation of asymmetric substitutions also exhibit interesting patterns. Considering previous structures and the presented works, in all cases except 11-Me substituents located at the 11- position are disordered unequally with fractional occupancy across the two mirrored positions, and this is true for all the electron withdrawing and halide substitutions. In all these cases, two thirds of the occupancy are directed towards the phosphate backbone and the other third towards the major groove. This does not hold true for ligands substituted at the 10- position however, where complete directional selectivity is observed towards the major groove side. Previously only noted with methylated species, the 10-nitro species exhibits the same 100 % directional preference for the major groove, implying that the directionality is forced mainly by steric inhibition with the backbone and not necessarily an electronic preference for either direction. Indeed if the complex is rotated about its principal axis (formerly the C_2 axis of parent dppz) there is a clear steric clash between the substituent and the backbone (figure 2.14); electrostatically this would also be disfavoured due to lone pair repulsion of the nitro group and the close contact phosphate bonded oxygens. Of course this close clash would only be possible in canted asymmetric intercalation sites; most of the structurally observed binding modes of lambda complexes have done so in a canted fashion except for the symmetrical binding mode to TA:TA base pair steps.[36] Unfortunately, crystallisation trials involving these sequences were unyielding (in fact no asymmetrical complexes could be successfully crystallised with sequences containing TA in the central step).
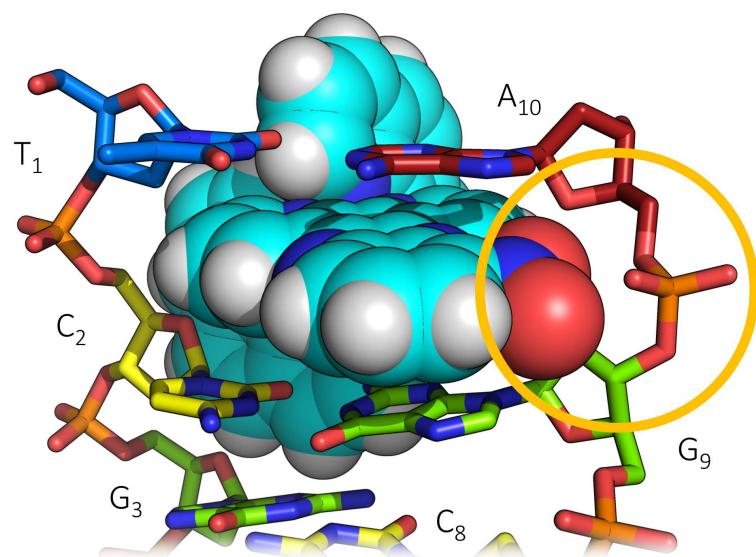

Figure 2.14 *– Fictional model created by rotating Λ-[Ru(phen)$_2$(10-NO$_2$-dppz)]$^{2+}$ by 180° about its former C$_2$ two-fold axis (PDB: 6RSO). An orange ring highlights where the subsequent disfavourable steric/electrostatic clashes would occur potentially explaining the complete directional specificity observed in the structure.*

Substitutions such as the nitrile or nitro groups can be used as effective IR handles to help understand binding environments and electronics in solution, and as such knowing the absolute orientation of a bound ligand can aid in the deconvolution of spectroscopic measurements and/or be used to help direct specific interactions.[37,38] These substitutional trends suggest that it is possible to distinguish steric, hydrophobic, and electronic effects on the orientational selectivity, and implies that asymmetric methyl ordering may be a result of the hydrophobicity of the group favouring the purine side of the intercalation cavity over projection into the major groove.

2.4.4 Polarity alignment and stabilisation of *syn*-guanosine

As discussed before, at the site of intercalation the complexes intercalate at an angle, asymmetrically to the P-P axis and preferentially stack on the side of the step with the most purine character. As a result of this, the terminal purines are almost always well-ordered whereas their paired pyrimidine is often less so (especially the case for the systems with complete cavities presented herein). Interestingly, the purines not only orient to maximise π-stacking with the intercalating ligand but in addition the differing polarities of the nucleobase are aligned consistent to the intercalating ligand. Figure 2.15 summarises the terminal intercalation sites discussed, comparing the difference between the $T_1·A_{10}$ and $C_1·G_{10}$ terminal base pairs, and the addition of the electron withdrawing nitrile substituents. In all the cases containing the T·A terminal pairs the adenine stacks on the dppz ligand with the $6-NH_2$ substituent directed towards the major groove, with the depth of the intercalation determined by the hydrogen in the 2- position (figure 2.15a). In contrast, the C·G terminal base pair observed in the structure between $Λ-[Ru(TAP)_2(11,12-CN-dppz)]^{2+}$ and d(CCGGACCCGG/CCGGGTCCGG) has the guanine stacked directly atop the dppz with the 6-CO directed perpendicular to the dppz axis, above a pyrazine N atom. The amino substituent, now in the 2- position, is again directed towards the major groove. In this case the *syn*-guanosine conformation is stabilised despite the lack of additional hydrogen bonds (stabilisation of *syn*-guanosine is often imparted by bonding to both Watson-Crick and Hoogsteen faces). The same terminal conformation is also observed in the structure containing the parent complex and is presumably a result of the clashes that would result between the nucleobase and TAP ligand with rotation about χ to yield the *anti* conformation. Except in the case of the scarce Z-DNA conformation and in the predicted transient Hoogsteen breathing mechanisms, *syn*-guanosines are relatively rare in duplex DNA; however in G-quadruplex DNA the stabilisation of such conformations is common and key to the observed topological diversity.[39,40]
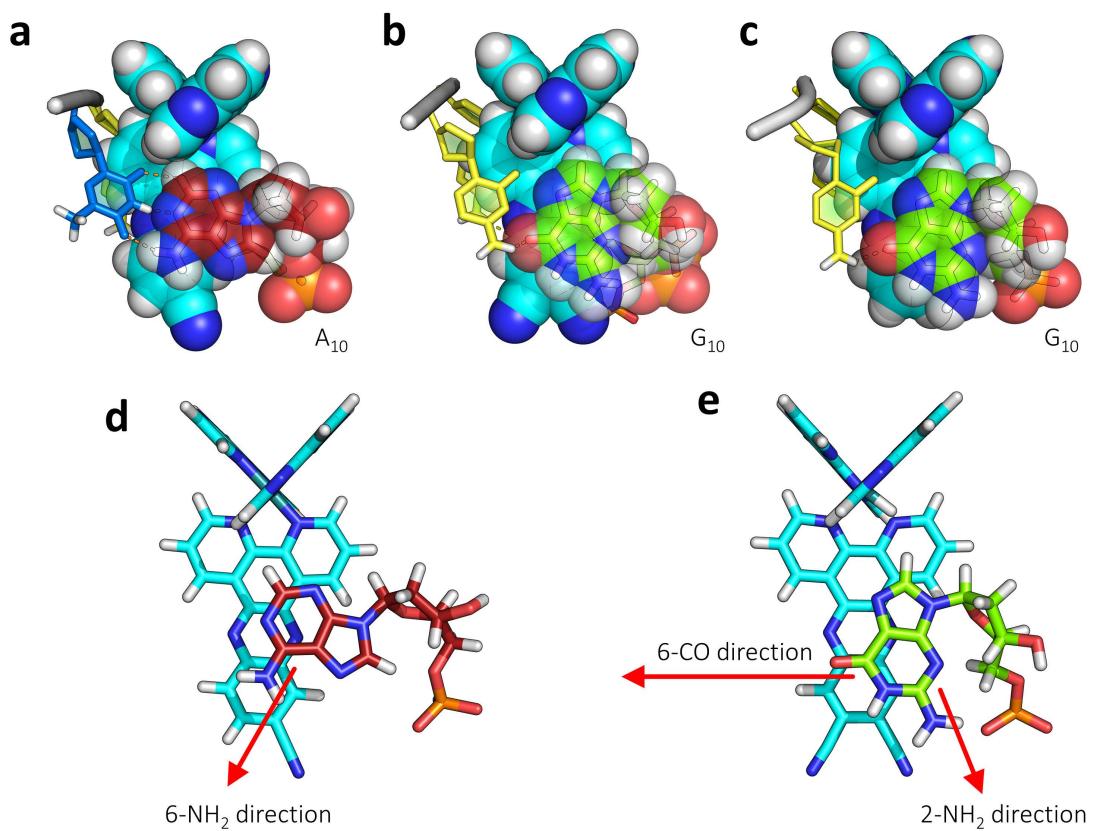
Figure 2.15 – *Selected adenosine and guanosine stacking observed in the presented structures highlighting the connection to base polarity. Adenosine adopts a Watson-Crick anti conformation whereas guanosine adopts a Hoogsteen syn-conformation. This difference can be related not only to the steric inhibitions but by the polarity of the bases. (a) The standard terminal A_{10} stacking on the 11-CN-dppz, seen for the first time as part of a complete terminal cavity; (b) syn-G_{10} stacking onto 11,12-CN-dppz, seen similarly in the parent structure (c); (d and e) highlight the polarity directions of the terminal nucleobases in (a) and (b).*

2.5 Summary

A group of [Ru(L)$_2$(dppz)]$^{2+}$ derivatives containing nitrile, nitro, and halide substitutions on the distal ring of dppz were successfully crystallised in the presence and absence of d(TCGGCGCCGA). It was found that the addition of the strongly withdrawing substituents (nitrile/nitro) enforced a closing of the intercalation cavity when bound to the decamer; an observation not observed with the underivatised parent complex and complexes containing weakly directing groups (halide/methyl). This is the first time such a stabilising effect has been observed in the crystal structures of DNA duplexes. DFT calculations show that seemingly subtle changes in the binder design yield large changes in the electronic structure, and this effect on the electrostatic potential and frontier molecular orbitals of the dppz may be responsible for the observed purine partiality. It is proposed that by utilising such substituent effects to create more favourable π-stacking environments, next generation complexes could be thoughtfully designed to bind selectively to electron rich or deficient base steps, such as mismatches, by altering the electronic quadrupole moment of the interacting ligand.

Observations noted from the collective analysis of the structures presented here and previously:

- Enantiospecific interactions with ribose sugar causes a canted binding mode with the complex stacked predominantly on the purine nucleobases.

- The depth of the interaction is regulated by steric clashes between an ancillary ligand and the terminal purine.

- Strongly electron withdrawing substituents promote complete intercalation cavities in the crystal structures containing T·A terminal base pairs.

- Asymmetry of the intercalating ligand imparts orientational specificity dependent on substitution position (10 position always directed towards major groove); and hydrophobicity (methyl groups always directionally specific, whereas nitro, nitrile, and halides are disordered disproportionally across two positions).

- Substitution with nitrile or bromo causes the asymmetry in DNA torsions of the base step to shift, switching to a gauche (or 'closed') dihedral on the G_9/A_{10} purine side.

An additional structure was presented containing Λ-[Ru(TAP)$_2$(11,12-CN-dppz)]$^{2+}$ bound to an asymmetric decamer d(CCGGACCCGG/CCGGGTCCGG). In this case the complex intercalated into a terminal C_1C_2:G_9G_{10} step instead of a T_1C_2:G_9A_{10} step. The structure is isostructural to that of the system containing the parent species despite containing the additional nitrile moieties. In contrast to the systems containing T·A/C·G terminal pairs, the intercalation cavity is always complete with this terminal base pair (based on a range of published and unpublished structural coordinates), but the preferred canted binding mode of the complex stabilises *syn*-guanosine as a result of the potential steric clash with the *anti* conformer that would prevent such a binding mode.

The works as a whole highlight the importance of electron directing groups not just for their spectroscopic handle properties and potential direct electrostatic interactions, but also their indirect effect on intercalation by adjustment of the dppz electronics. The structures presented are the first to exhibit structurally the thermodynamically stabilising effects of distal ring substitution. Future works will investigate the structural consequences of strongly donating moieties and will expand the oligonucleotides studied to explore further the base step preferences of derivatised dppz systems.

2.6 References

1   Boynton, A. N., Marcélis, L. & Barton, J. K. [Ru(Me$_4$phen)$_2$dppz]$^{2+}$, a light switch for DNA mismatches. *J. Am. Chem. Soc.* **138**, 5020–5023 (2016).

2   Ghizdavu, L. *et al.* Oxidizing Ru(II) complexes as irreversible and specific photo-cross-linking agents of oligonucleotide duplexes. *Inorg. Chem.* **48**, 10988–10994 (2009).

3   Wachter, E., Moyá, D., Parkin, S. & Glazer, E. C. Ruthenium complex 'light switches' that are selective for different G-quadruplex structures. *Chem. Eur. J.* **22**, 550–559 (2016).

4   Hess, J. *et al.* Evaluation of the medicinal potential of two ruthenium(II) polypyridine complexes as one- and two-photon photodynamic therapy photosensitizers. *Chem. Eur. J.* **23**, 9888–9896 (2017).

5   Hall, J. P., Beer, H., Buchner, K., Cardin, D. J. & Cardin, C. J. The structural effect of methyl substitution on the binding of polypyridyl Ru–dppz complexes to DNA. *Organometallics* **34**, 2481–2486 (2015).

6   Hall, J. P., Beer, H., Buchner, K., Cardin, D. J. & Cardin, C. J. Preferred orientation in an angled intercalation site of a chloro-substituted Λ-[Ru(TAP)$_2$(dppz)]$^{2+}$ complex bound to d(TCGGCGCCGA)$_2$. *Philos. Trans. R. Soc. A* **371**, 525 (2013).

7   Liu, X.-W. *et al.* Synthesis, characterization, DNA-binding and photocleavage of complexes [Ru(phen)$_2$(6-OH-dppz)]$^{2+}$ and [Ru(phen)$_2$(6-NO$_2$-dppz)]$^{2+}$. *J. Inorg. Biochem.* **99**, 2372–2380 (2005).

8   Hall, J. P. *et al.* Monitoring one-electron photo-oxidation of guanine in DNA crystals using ultrafast infrared spectroscopy. *Nat. Chem.* **7**, 961–967 (2015).

9   Keane, P. M. *et al.* Inosine can increase DNA's susceptibility to photo-oxidation by a RuII complex due to structural change in the minor groove. *Chem. Eur. J.* **23**, 10344–10351 (2017).

10   Cardin, C. J., Kelly, J. M. & Quinn, S. J. Photochemically active DNA-intercalating ruthenium and related complexes – insights by combining crystallography and transient spectroscopy. *Chem. Sci.* **8**, 4705–4723 (2017).

11   Dolomanov, O. V, Bourhis, L. J., Gildea, R. J., Howard, J. A. K. & Puschmann, H. OLEX2: a complete structure solution, refinement and analysis program. *J. Appl. Crystallogr.* **42**, 339-341 (2009).

12   Palatinus, L. & Chapuis, G. SUPERFLIP– a computer program for the solution of crystal structures by charge flipping in arbitrary dimensions. *J. Appl. Crystallogr.* **40**, 786–790 (2007).

13   Sheldrick, G. Crystal structure refinement with SHELXL. *Acta Crystallogr. Sect. C* **71**, 3–8 (2015).

14   Winter, G. xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43**, 186–190 (2010).

15   Kabsch, W. XDS. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 125–132 (2010).

16   Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. Sect. D* **66**, 133–144 (2010).

17   Winter, G. *et al.* DIALS: implementation and evaluation of a new integration package. *Acta Crystallogr. Sect. D, Struct. Biol.* **74**, 85–97 (2018).

18   Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. Sect. D* **69**, 1204–1214 (2013).

19   Grosse-Kunstleve, R. W. & Adams, P. D. Substructure search procedures for macromolecular structures. *Acta Crystallogr. Sect. D* **59**, 1966–1973 (2003).

20   McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).

21   Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D* **60**, 2126–2132 (2004).

22   Adams, P. D. *et al.* The Phenix software for automated determination of macromolecular structures. *Methods* **55**, 94–106 (2011).

23   Sheldrick, G. M. Experimental phasing with SHELXC/D/E: Combining chain tracing with density modification. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 479–485 (2010).

24   Moriarty, N. W., Grosse-Kunstleve, R. W. & Adams, P. D. electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr. D. Biol. Crystallogr.* **65**, 1074–1080 (2009).

25   Hanwell, M. D. *et al.* Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminform.* **4**, 17 (2012).

26   Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 355–367 (2011).

27   Potterton, L. *et al.* CCP4i2: the new graphical user interface to the CCP4 program suite. *Acta Crystallogr. Sect. D* **74**, 68–84 (2018).

28   Adams, P. D. *et al.* PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 213–221 (2010).

29   Hall, J. P. *et al.* Structure determination of an intercalating ruthenium dipyridophenazine complex which kinks DNA by semiintercalation of a tetraazaphenanthrene ligand. *Proc. Natl. Acad. Sci.* **108**, 17610-17614 (2011).

30   Hall, J. P. *et al.* Controlled dehydration of a ruthenium complex–DNA crystal induces reversible DNA kinking. *J. Am. Chem. Soc.* **136**, 17505–17512 (2014).

31   Mari, C. *et al.* DNA intercalating RuII polypyridyl complexes as effective photosensitizers in photodynamic therapy. *Chem. Eur. J.* **20**, 14421–14436 (2014).

32   Mari, C., Rubbiani, R. & Gasser, G. Biological evaluation of nitrile containing Ru(II) polypyridyl

complexes as potential photodynamic therapy agents. *Inorganica Chim. Acta* **454**, 21–26 (2017).

33   Hou, M.-H., Robinson, H., Gao, Y.-G. & Wang, A. H.-J. Crystal structure of actinomycin D bound to the CTG triplet repeat sequences linked to neurological diseases. *Nucleic Acids Res.* **30**, 4910–4917 (2002).

34   Weil, J. *et al.* Stabilization of the i-motif by intramolecular adenine-adenine-thymine base triple in the structure of d(ACCCT). *Acta Crystallogr. Sect. D Biol. Crystallogr.* **55**, 422–429 (1999).

35   Hall, J. P. *et al.* X-ray crystal structure of rac-[Ru(phen)$_2$dppz]$^{2+}$ with d(ATGCAT)$_2$ shows enantiomer orientations and water ordering. *J. Am. Chem. Soc.* **135**, 12652–12659 (2013).

36   Niyazi, H. *et al.* Crystal structures of Λ-[Ru(phen)$_2$dppz]$^{2+}$ with oligonucleotides containing TA/TA and AT/AT steps show two intercalation modes. *Nat. Chem.* **4**, 621–628 (2012).

37   Hall, J. P. *et al.* Monitoring one-electron photo-oxidation of guanine in DNA crystals using ultrafast infrared spectroscopy. *Nat. Chem.* **7**, 961–967 (2015).

38   Poynton, F. E. *et al.* Direct observation by time-resolved infrared spectroscopy of the bright and the dark excited states of the [Ru(phen)$_2$(dppz)]$^{2+}$ light-switch compound in solution and when bound to DNA. *Chem. Sci.* **7**, 3075–3084 (2016).

39   Gessner, R. V, Frederick, C. A., Quigley, G. J., Rich, A. & Wang, A. H. The molecular structure of the left-handed Z-DNA double helix at 1.0-A atomic resolution. Geometry, conformation, and ionic interactions of d(CGCGCG). *J. Biol. Chem.* **264**, 7921–7935 (1989).

40   Nikolova, E. N. *et al.* Transient Hoogsteen base pairs in canonical duplex DNA. *Nature* **470**, 498-502 (2011).

3 Enantiospecific Recognition of G-quadruplexes by Ruthenium Polypyridyl Complexes

Contribution statement

Complex synthesis, purification and characterisation, enantiomeric purification, HT-fluorescence, SR-circular dichroism, crystallisation of 6HWG, structure solution of 6HWG, data analysis, and molecular modelling, were performed by Kane McQuaid. The crystal of which data were collected for 5LS8 was grown by Holly Abell. Structure solution for 5LS8 was performed by Dr James Hall.

The work presented in this chapter made the basis of the following publication:

McQuaid, K., Abell, H., Gurung, S., Allan, D.R., Winter, G., Sorensen, T., Cardin, D.J., Brazier, J.A., Cardin, C.J., and Hall, J.P. Structural studies reveal enantiospecific recognition of a DNA G-quadruplex by a ruthenium polypyridyl complex. *Angew. Chem. Int. Ed.* **58**, 9881, (2019).

The structures listed below were a direct output from the work presented here and were submitted to the Protein Data Bank with the following identification:

6HWG - Λ-[Ru(phen)$_2$(11-CN-dppz)]$^{2+}$ bound to the B-DNA duplex d(TCGGCGCCGA)

5LS8 - Light-activated ruthenium complex bound to a DNA quadruplex

3.1 Introduction

Owing to their now proven importance in the living cell, G-quadruplexes have become one of the most heavily researched biological motifs. With putative roles in cell division, telomere functionality, and gene expression, the folded assembly has become a target both diagnostically and therapeutically. G-quadruplexes found in the telomeric regions of chromosomal DNA are responsible for the protection of coding genes during replication in healthy cells, but in a number of cancers over-expression of telomerase protects the natural attrition of the telomeres and can lead to the immortalisation of cells. Formed of tandem repeats of the sequence TAGGGT, the telomeric G-quadruplexes have been shown to exhibit wide topological variability dependent on local cellular

conditions. As such, the design of small molecules that can bind to the structures with a degree of morphological selectivity is of key importance. Octahedral ruthenium polypyridyl complexes have shown great promise as G-quadruplex binders and luminescent reporters of structure, but very little is known about the binding modes of the these heavily researched complexes to G-quadruplexes. Currently only 9 sets of structural coordinates are available showing how metal based complexes bind to G-quadruplexes; where two of them are based on the structurally similar dinuclear ruthenium complexes.[1]

This chapter aims to provide structural insights into the manner in which these mononuclear ruthenium species bind to G-quadruplex DNA, and to use these new insights to help better understand observations in the solution state and to provide structural rationale for the development of more specific ruthenium binders.

3.2 Methodology

Oligonucleotides were purchased from Eurogentec as HPLC-purified solids and were used without further purification. Unless otherwise stated, all other starting materials and chemicals were sourced from Sigma Aldrich or Honeywell research chemicals..

3.2.1 Synthesis, characterisation, and enantiomeric separation

Synthesis of the complexes; rac-[Ru(phen)$_2$(dppz)]$^{2+}$, rac-[Ru(phen)$_2$(11-CN-dppz)]$^{2+}$, rac-[Ru(phen)$_2$(11,12-CN-dppz)]$^{2+}$, rac-[Ru(TAP)$_2$(dppz)]$^{2+}$, and rac-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$, were carried out using modifications on previously published methodology and is described in detail in chapter 7. Enantiomeric purification methodology is described in section 7.2.4, with their eluting conditions and the subsequent circular dichroism signals of the optically pure enantiomers shown in Figure A3.1.

3.2.2 Luminescence spectroscopy

3.2.2.1 high-throughput fluorescence screening

High-throughput screening was undertaken as part of a residential trip to the laboratory of Prof. Edith Glazer at the University of Kentucky, USA. DNA sequences were resuspended in buffer, incubated at room temperature for 15 min, vortexed, and annealed prior to measurement. Following heating to the specified annealing temperature, the DNA was cooled slowly to room temperature then stored at 4 °C overnight. Bovine serum albumin (BSA), deoxyadenosine (dA), deoxyguanosine (dG), deoxythymidine (dT) and deoxycytosine (dC) were resuspended and sonicated as reported previously. All biomolecules were stored long term at -20 °C. Luminescence emission was evaluated at

a final concentration of 5 µM complex in the presence of 25 µM biomolecule, where the biomolecule concentration is measured in [bp] yielding a Ru:bp ratio of 1:5. Emission spectra were collected in triplicate from 25 µL wells in Greiner black/UV clear bottom 384-well plates using a λ_{ex} = 440 nm and λ_{em} range of 550-750 nm on a Molecular Devices Spectramax M5 microplate reader. The protein, nucleosides and DNA sequences used as part of the HTS are shown in Table A3.1 along with their buffering/annealing conditions and sequence identities. Raw data were analysed using Graphpad Prism 6 and OriginPro 9.1.

3.2.2.2 focused fluorescence spectroscopy

Further condensed screening of some of the complexes in the presence of a small range of G-quadruplex forming sequences was conducted. Measurements were collected on a Cary Eclipse fluorescence spectrophotometer fitted with a Peltier thermostatted temperature controller. DNA was handled in an analogous fashion to above, storing the DNA at -20 °C when not in use. Complexes and DNA were pre-annealed together (10 mins at 95 °C) prior to measurement (except polynucleotides). Emission spectra were collected on samples containing 25 µM DNA (by [bp]) and 5 µM complex, in 1 cm path length quartz cells using a λ_{ex} = 440 nm and a λ_{em} wavelength range of 550-850 nm.

3.2.3 Synchrotron radiation circular dichroism (SRCD)

SRCD spectra and subsequent CD melting analyses were collected on beamline B23 at Diamond Light Source Ltd. d(TAGGGTTA) concentration was kept constant at 800 µM, giving a final concentration of 200 µM of quadruplex unit. The concentration of complex for a 1:1 molar ratio was therefore 200 µM and for a 4:1 molar ratio was 800 µM. All samples contained 20 mM pH 7 buffered K-cacodylate, and 30 mM KF. All samples were mixed, heated to 95 °C for 5 minutes, and then allowed to cool slowly to room temperature before measurement. Spectra were measured in bespoke quartz plates with 100 µm cell path lengths; whereas the melting analyses were performed in quartz 'H' cells of path length 100 µm. All spectra were acquired using a 1 second integration time per nm, with a 1 nm slit, between 180-350 nm and then were cut according to an appropriate PMT voltage. This results in a cut-off of 193 nm being applied to the spectra. Final plots were both background and offset corrected using the CDApps suite.[2]

3.2.4 X-ray Crystallography

3.2.4.1 crystallisation of Λ-[Ru(phen)₂(11-CN-dppz)]²⁺ with d(TCGGCGCCGA)

3.2.4.1.1 crystallisation conditions

Crystals grown from the oligonucleotide d(TCGGCGCCGA) and ligand, [Ru(phen)₂(11-CN-dppz)]²⁺, were grown from sitting drops via the vapour diffusion method. Crystals suitable for diffraction experiments were grown at 18 °C from a 8 μL drop containing; 250 μM d(TCGGCGCCGA), 750 μM *rac*-[Ru(phen)₂(11-CN-dppz)]·Cl₂, 7.5 % w/v MPD, 30 mM pH 7 sodium cacodylate, 9 mM spermine tetrahydrochloride, 60 mM KCl, and 15 mM BaCl₂; all equilibrated against 1 mL of 35 % v/v MPD. Orange/red rods grew following roughly three weeks at incubation at 291 K.

3.2.4.1.2 data collection, refinement, and analysis

Data were collected on beamline I02 at Diamond Light Source Ltd. using radiation of wavelength 0.9763 Å on a flash-cooled crystal at 100 K. A 180° wedge of data was collected with an oscillation of 0.10° per image, generating a total of 1800 images. The data were processed with xia2, using DIALS and Aimless to integrate, scale, and merge the data; yielding 3816 unique reflections to a maximum outershell resolution of 1.74 Å. Phase determination was achieved using molecular replacement techniques with Phaser-MR, using the predicted isostructural co-ordinates from 5NBE after removing the waters from the model. Chemical refinement restraints for the ligand Λ-[Ru(phen)₂(11-CN-dppz)]²⁺ were calculated from modelling data and crystallographic constraints as has been mentioned in previous chapters, using Gaussian09, Avogadro and the eLBOW/REEL programs in the PHENIX software package. Editing to the top-scoring MR model was done by hand in Wincoot and the subsequent models were refined against the original data using Refmac5, yielding a final R_{work} of 0.20 and R_{free} of 0.23; where 5.2 % of the original reflections were reserved for the generation of the R_{free} flag set. The final model and data were deposited in the protein data bank with PDB ID 6HWG. Table 3.2 summarises the data collection and refinement statistics.

3.2.4.2 crystallisation of Λ-[Ru(TAP)₂(11-CN-dppz)]²⁺ with d(TAGGGTTA)

3.2.4.2.1 crystallisation conditions

Crystals containing the oligonucleotide d(TAGGGTTA) and the ruthenium complex [Ru(TAP)₂(11-CN-dppz)]²⁺ were grown from sitting drops via vapour diffusion of water at 18 °C. Crystals suitable for X-ray diffraction experiments were obtained from various differing conditions, however not all conditions gave rise to well diffracting samples. The solution forming the sitting drops was constituted from two components; 1μL of a pre-annealed mixture of the single stranded oligonucleotide at 250 μM with the complex *rac*-[Ru(TAP)₂(11-CN-dppz)]·Cl₂ at 500 μM in a 20 mM KCl buffer; and 1 μL of a solution containing 25 mM MgCl₂·6H₂O, 20 mM KCl and 10 % w/v PEG

monomethyl ether 550, pH 7.5 buffered using 50 mM TRIS hydrochloride. The sitting drop was equilibrated against 100 μL of the same PEG containing screening condition forming dark orange/red hexagonal prisms within 2 weeks of preparation. No crystals could be obtained using the unsubstituted parent compound, [Ru(TAP)$_2$(dppz)]·Cl$_2$, even as enantiopure material.

3.2.4.2.2 data collection, refinement, and analysis

The data were collected at Diamond Light Source Ltd., on beamline I02 using radiation with a wavelength of 1.65312 Å from a flash cooled crystal at 100K. 720° of data were collected with an oscillation of 0.1° per frame, generating 7200 images. The resulting data were processed using DIALS[3] and Aimless[4] through the xia2[5] pipeline and gave an anomalous signal with a mid-slope of anomalous normal probability[6] of 1.502, finding 9958 unique reflections to a resolution of 1.78 Å. The structure was solved using the anomalous scattering of ruthenium by single wavelength anomalous dispersion with the SHELX/C/D/E[7] pipeline in CCP4.[8] The crystallographic model was built using WinCoot[9] and refined using Refmac5[10] to give a final R_{work} of 0.21 and an R_{free} of 0.23 reserving 5 % of the total reflections for the R_{free} set. Figures were rendered using PyMOL. The structure is deposited in the Protein Data Bank with PDB ID: 5LS8. Table 3.1 highlights the main data collection and refinement statistics.

3.2.4.3 data collection and refinement statistics

Table 3.1 - *Crystallisation, data collection, and processing parameters/refinement results of the crystal structures*

Crystallisation Parameters	[Ru(phen)$_2$(11-CN-dppz)]$^{2+}$ with d(TCGGCGCCGA)	[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ with d(TAGGGTTA)
Crystal Morphology	Orthorhombic	Hexagonal Prismatic
Growth Temperature (K)	291	291
Crystal Size (μm)	60x30x100	60x40x70
Growth Time	5 weeks	4 weeks
Data Collection		
Beamline	I03	I02
X-Ray Wavelength (Å)	0.9763	1.6531
Transmission (%)	10	2.99
Beamsize (μm)	20x20	20x20
Exposure Time (s)	0.02	0.04
№ Images/Oscillation (°)	1800/0.10	7200/0.10
Space Group	$P\,4_3$	$P\,6_5$
Cell Dimensions a, b, c (Å); α, β, γ (°)	47.94, 47.94, 33.63; 90, 90, 90	36.63, 36.63, 136.1; 90, 90,120
Data Processing		
Resolution (Å)	33.90 - 1.74 (1.80 - 1.74)*	68.05 - 1.78 (1.81 - 1.78)
R_{merge}	0.038 (0.358)	0.070 (1.211)
R_{meas}	0.053 (0.506)	0.071 (1.235)
R_{pim}	0.037 (0.358)	0.011 (0.240)
№ Observations	15,941 (1605)	372,526 (13,166)
№ Unique Observations	7989 (805)	9958 (506)
$I/\sigma I$	11.05 (1.76)	27.7 (2.8)
$CC_{1/2}$	0.999 (0.358)	1.000 (0.916)
Completeness (%)	99.94 (100)	99.61 (100.00)
Multiplicity	2.0 (2.0)	37.4 (26.0)
Mid-slope of anom normal probability	1.012	1.502

* Outer shell statistics shown in parentheses

Refinement		
Phase Solution Method	SAD-MR	SAD
Resolution	33.9 (1.74)	31.7 (1.78)
№ Reflections	15338	9857
R_{work}/R_{free}	0.1939/0.2055	0.2124/0.2287
№ Atoms		
DNA	404	699
Metal Complex	108	320
Water	83	24
Average B Factors (Å^2)		
DNA	39.05	51.79
Metal Complex	31.91	40.23
Water	37.86	45.65
rmsd		
Bond Lengths (Å)	0.011	0.010
Bond Angles (°)	1.4	2.4
PDB ID	**6HWG**	**5LS8**

3.3 Results

3.3.1 High-throughput screening trials

To aid in the discovery and evaluation of new structure specific 'light-switch' complexes, a HTS approach was employed. The developed assay allows for the assessment of a prospective probes' luminescence response in the presence of 32 different biomolecules, including: BSA, free nucleotides, natural ds-DNA; non-canonical triplexes and G-quadruplexes; and DNA sequences containing mismatched bases, abasic sites and bulges. The design of the assay allows for the effects on the 'light switch' response of a variety of binding environments to be evaluated in tandem. All complexes were evaluated under DNA saturating conditions (1:5 [Ru]:[DNA bp]) to attempt to reduce the possibility of multiple binding modes occurring; and to allow for the comparison of complexes independent of their respective binding affinities. The photophysical 'light switches' were screened in 96-well plates (figure 3.1); where full emission spectra, collected in triplicate, were used to follow the spectral luminescence response.
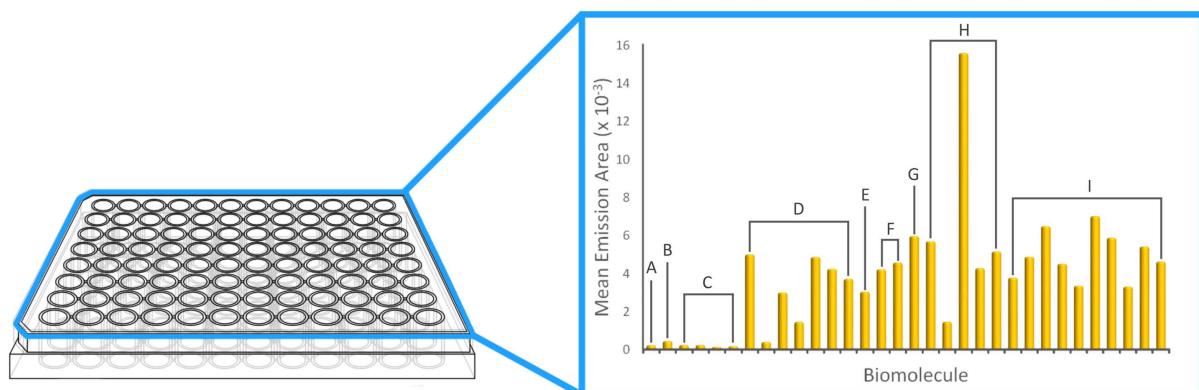
Figure 3.1 – *Experimental scheme of the 96 well screening assay, allowing for the simultaneous evaluation the 'light switch' response in the presence of 32 different biomolecules in triplicate. Inset dataset is an example of the raw luminescence area output and highlights the groups of biomolecules in order as they appear in table A3.1. (A) buffer, (B) BSA, (C) nucleosides, (D) natural and ds-DNA, (E) cruciform DNA, (F) hairpin DNA, (G) triplex DNA, (H) G-quadruplexes and i-motifs, and (I) damaged DNA.*

A range of complexes was investigated but only a handful will be discussed here. Of all the complexes assessed, most exhibited a moderate enhancement of emission in the presence of at least a handful of the motifs or biomolecules. Initially, *rac*-[Ru(phen)$_3$]$^{2+}$ was examined as a control measure. Interestingly however, little motif differentiation was observed in the luminescence response for the racemic mixture; with the mean emission areas deviating only within standard deviation for most biomolecules when compared to buffered complex alone.
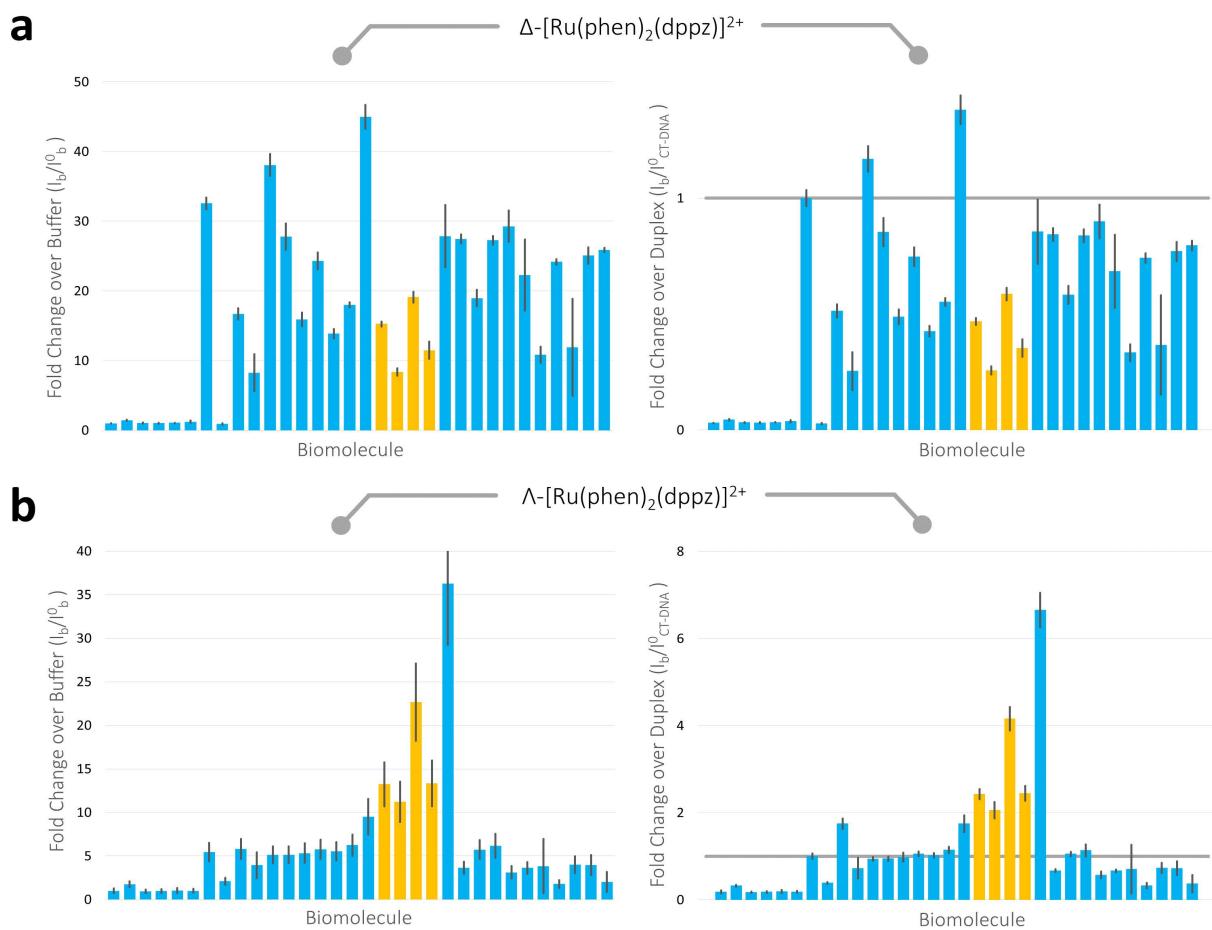
**Figure 3.2** – *High-throughput screening results reporting the luminescence fold changes of (a) Δ-[Ru(phen)₂(dppz)]²⁺, and (b) Λ-[Ru(phen)₂(dppz)]²⁺, when compared to the luminescence response of the complexes in the presence of buffer alone or with natural CT-DNA. G-quadruplex forming sequences are highlighted in orange. Fold change equal to one against duplex DNA ($I_b/I^0_{CT-DNA} = 1$) is shown by a grey line to highlight specificity against natural DNA. Error bars depict the standard deviations of the triplicate results. Note the generally opposite responses to higher-order DNA between the enantiomers.*

Across all 'light-switch' complexes assessed, the free nucleosides and hydrophobic albumin proteins exhibited little difference in the luminescence responses. This highlights the importance of a tight binding cavity, such as is available in nucleic acid base steps, in the mediation of the photophysical 'light-switch' effect. However, when incubated with the different secondary structure forming DNA sequences, motif specific responses become more evident. Initially, the two optical isomers of [Ru(phen)₂(dppz)]²⁺ were evaluated. Whilst neither enantiomer exhibited marked selectivity to any one sequence, the assays highlight a stark preference in response of the delta enantiomer to duplex DNA over G-quadruplexes, and vice versa for lambda and higher order motifs over duplexes (figure 3.2a). In both cases the isomers exhibited a relatively higher luminescence with the unimolecular telomeric sequence folded in the presence of potassium in comparison to other quadruplex forming sequences. This topological preference in response manifests as a 4.2 fold increase in luminescence of the lambda isomer in comparison to CT DNA. Interestingly, the lambda

enantiomer also displayed a 6.6 fold increase in response in the presence of the telomeric i-motif sequence over CT DNA, despite the short diagonal loop systems found in the sequence.

The same assay was conducted on the asymmetrically derivatised $[Ru(phen)_2(11\text{-}CN\text{-}dppz)]^{2+}$ and compared to the parent species (figure 3.3). The cyano complex had not been successfully enantiomerically purified at time of conducting the screens at the University of Kentucky so racemic mixtures were used initially to probe the chromophores specificity.
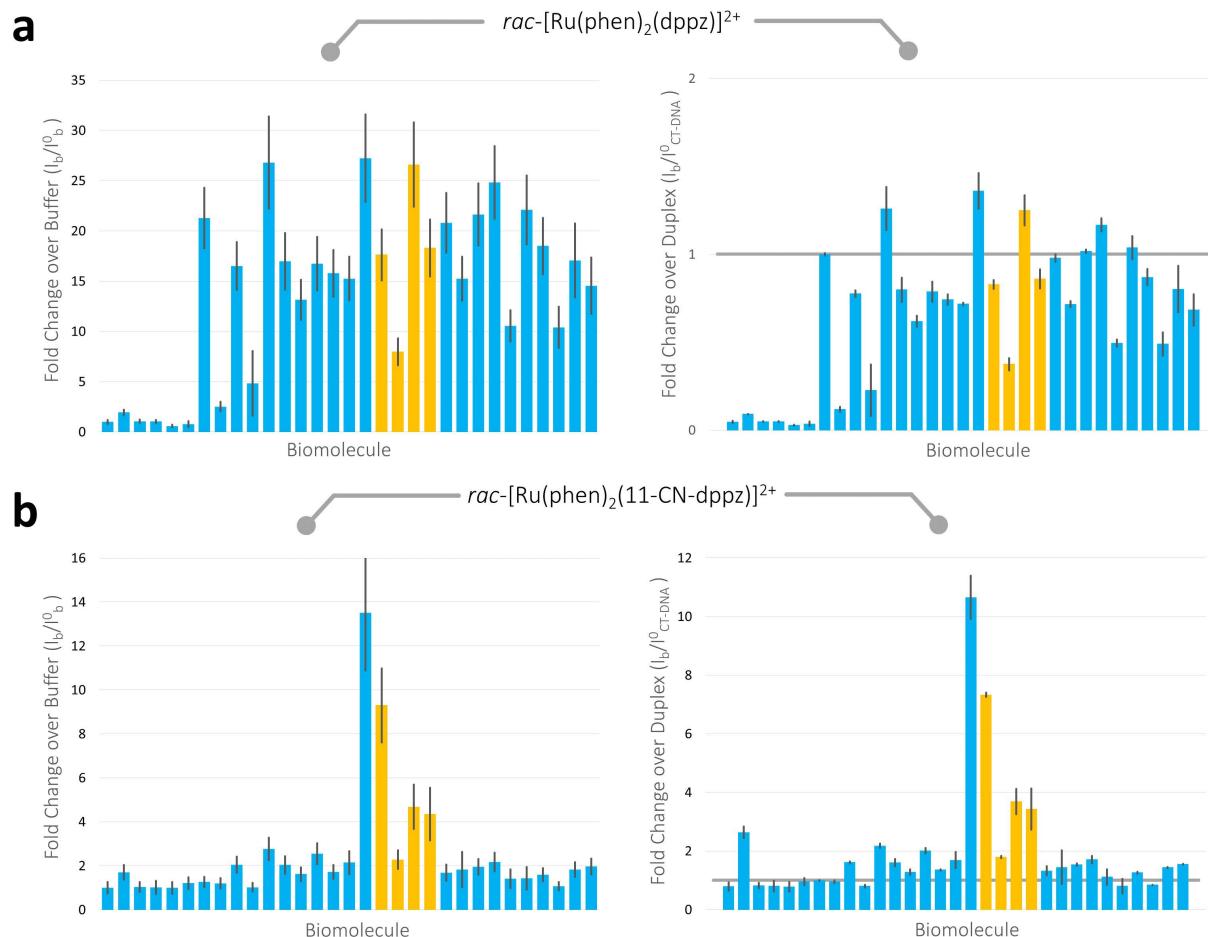
Figure 3.3 – *High-throughput screening results reporting the luminescence fold changes of (a) rac-[Ru(phen)₂(dppz)]²⁺, and (b) rac-[Ru(phen)₂(11-CN-dppz)]²⁺, when compared to the luminescence response of the complexes in the presence of buffer alone or with natural CT-DNA. G-quadruplex forming sequences are highlighted in orange. Fold change equal to one against duplex DNA (Iᵦ/I⁰꜀ₜ₋DNA = 1) is shown by a grey line to highlight specificity against natural DNA. Error bars depict the standard deviations of the triplicate results. Note the generally opposite responses to higher-order DNA between the enantiomers.*

It is evident from these studies that using racemic mixtures decreases any specificity as the two enantiomers are likely to have differing binding modes and/or encapsulation of the chromophore and ancillary ligands. The parent *rac*-[Ru(phen)₂(dppz)]²⁺ exhibits what is effectively an averaging of the two enantiomers luminescence; decreasing the overall motif preference in response. Still, only a few of the biomolecules illicit a light-switch response stronger than with natural CT-DNA. The cyano derivative however exhibits an overall lower intensity in response to the biomolecules but a higher

specificity towards the higher-orders of folded DNA. This is especially true for the triplex forming sequence and the intermolecular G-quadruplex where a 10.7 and 7.4 fold increase is observed compared to calf thymus DNA, respectively. As has been noted across the screens, all complexes tested gave their lowest intensity quadruplex emissions in the presence of the sodium folded quadruplex; with most systems exhibiting >2-fold increase in luminescence when the telomeric sequence was folded with potassium ions. In the tested experimental conditions, the sodium form natively takes on an anti-parallel basket topology, whereas the same sequence takes on a [3+1] mixed-hybrid topology when folded in the presence of a high concentration of potassium ions.[11–13]

Despite the high-throughput investigation being conducted on the racemic mixture, the asymmetric $[Ru(phen)_2(11\text{-}CN\text{-}dppz)]^{2+}$ exhibited a much higher specificity for the higher order motifs than the parent complex. In order to elucidate its origin and to have any discussion of structural rationale for the observation, the enantiomeric preferences in response would need to be evaluated. In this case condensed screening was executed at the University of Reading after successfully separating the enantiomers (see figure A3.1).
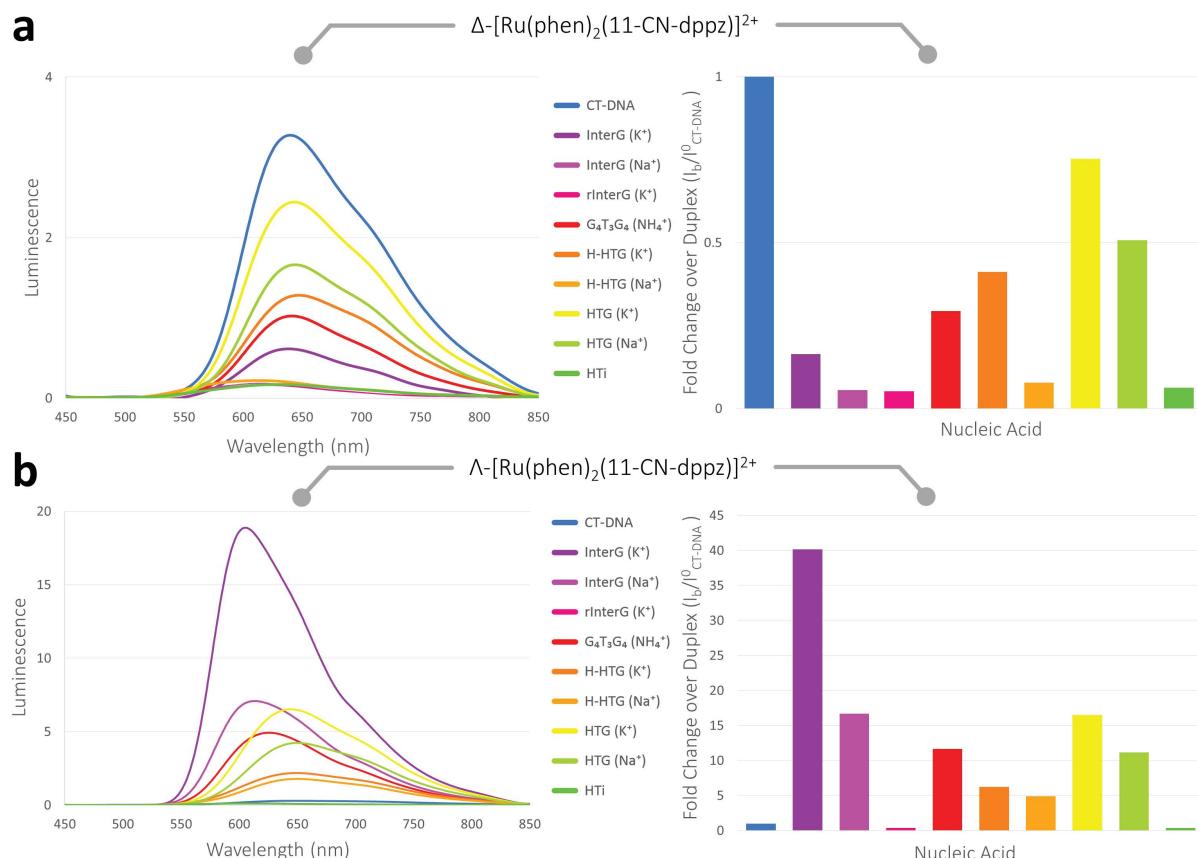
Figure 3.4 – *Luminescence responses of (a) Δ-[Ru(phen)₂(11-CN-dppz)]²⁺, and (b) Λ-[Ru(phen)₂(11-CN-dppz)]²⁺ in the presence of natural duplex CT-DNA and a range of G-quadruplex forming sequences. The raw luminescence profiles are shown on the left and on the right the fold change in integrated peak areas when compared to to CT-DNA; highlighting the motif specificity of the lambda enantiomer.*

Investigation of the separate optical isomers of the nitrile derivative clearly show a further enantiomeric disparity in luminescence response. Similar to the parent compound, the lambda isomer is evidently more luminescent in the presence of the G-quadruplexes studied. This is especially true for the potassium folded InterG sequence, where this preference is pertinent and materialises in a 40.1 times preference in probe response to the tetramolecular sequence over CT-DNA. Conversely, delta is more responsive to the presence of CT-DNA and overall the G-quadruplex response is far lower than that of the lambda counterpart. Interestingly however, besides the two tetramolecular sequences, the same pattern in topological response is observed between the enantiomers; albeit with the delta exhibiting a lower magnitude. This could imply similar binding environments for the two enantiomers when bound to uni and bimolecular sequences, where the depth of interaction potentially provides the disparity in response. As noted in the high-throughput screens and observed here, the sequences folded with sodium all generated a lower response from the complexes than the potassium forms; whether this is a result of the difference in fold or a result of a difference in the yield of quadruplex formation is unknown. The RNA analogue of the InterG sequence in all cases effectively did not provide any light-switch response from the complexes, implying that the 2'-hydroxyl group either blocks direct interaction or prevents the morphological changes required for the interactions. The largest disparity between the optical isomers occurs with the InterG sequences, with the delta exhibiting a far lower luminescence than the lambda and lower in comparison to the other delta systems. In addition, most systems that were measured yielded emission maxima in the range of 635-650 nm (figure 3.4), whereas lambda in the presence of the InterG sequences and with the ammonium stabilised $G_4T_3G_4$ yielded maxima closer to 610 nm. This blueshifting in maxima is indicative of wholly distinctive binding modes or changes in the binding environment of the two isomers and can be attributed to large increases in π-overlap and/or chromophore encapsulation. Such characteristic Stokes shifts have been used alongside changes in the lifetime of the luminescence events as an effective tool towards imaging of *in cellulo* G-quadruplex structure due to the ability to differentiate emission response originating from different DNA stuctures.[14]

3.3.2 Crystallisation of Λ-[Ru(phen)$_2$(11-CN-dppz)]$^{2+}$ with d(TCGGCGCCGA)

3.3.2.1 structure summary
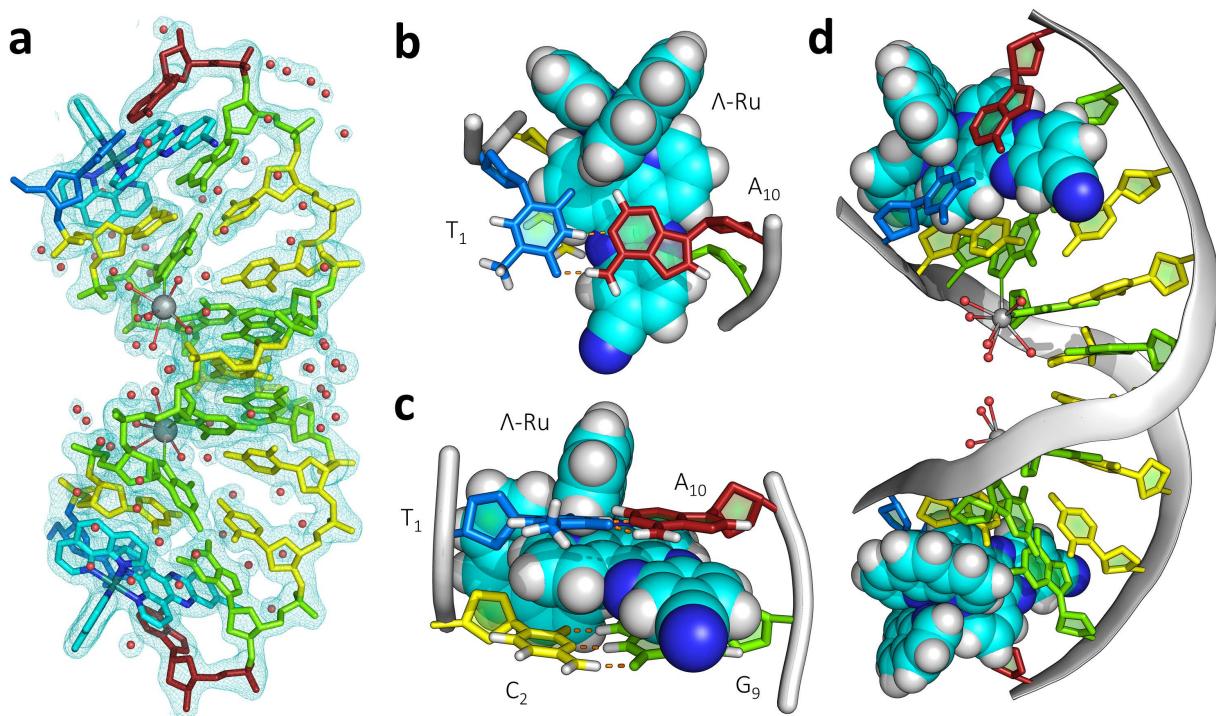
Figure 3.5 – (a) 2F$_o$-F$_c$ map, contoured at 0.29 e Å^{-3}, of 6HWG. Classical intercalation of Λ-[Ru(phen)$_2$(11-CN-dppz)]$^{2+}$ as (b) viewed down the helical axis from the terminal T$_1$·A$_{10}$ base pair side; and (c) the intercalation cavity of the base step T$_1$C$_2$:G$_9$A$_{10}$. (d) the complete model of 6HWG. Nucleic acids are coloured as: adenine in red, cytosine in yellow, guanine in green, and thymine in marine blue. Barium and water are shown as silver and red spheres respectively.

Overall, the structure of Λ-[Ru(phen)$_2$(11-CN-dppz)]$^{2+}$ bound to d(TCGGCGCCGA) is almost isostructural to the analogous system with Λ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ (5NBE) from chapter 2. Superimposition of the two models yields an RMSD of 1.12 Å and highlights the similarity in binding between the isoelectronic phen and TAP complexes (see figure A3.2b). The asymmetric unit contains two crystallographically independent strands of d(TCGGCGCCGA) that accommodate two intercalated Λ-[Ru(phen)$_2$(11-CN-dppz)]$^{2+}$ complexes that are electrostatically associated to neighbouring biological units through semi-intercalative binding modes (figure 3.5). Two hydrated barium cations stabilise this binding mode and a further 72 ordered waters are found in the first and second shells of hydration. Intercalation occurs in a canted/asymmetric manner and is directed by ancillary phen π-stacking interactions with a ribose on the pyramidine side of the intercalation cavity; presumably to increase favourable π-overlap with the purine side of the base step. In addition, as noted before, the terminal base pairs form a complete cavity, such that the A$_{10}$ base is not flipped out as is observed with systems where the dppz does not contain electron withdrawing moieties. In comparison to the 5NBE structure, the terminal T$_1$·A$_{10}$ base pair exhibits far less propeller twist (i.e. the dihedral about an axis

between N1/9 atoms); with the terminal T_1 in 5NBE being slightly disordered throughout the nucleobase (see figure 2.9a). This may be a result of increased ordering of T_1 in the crystal due to preferential 11-CN-dppz/nucleobase LUMO-HOMO localisations in the phen species, or it may be a result simply of a more ordered crystal. This asymmetric classical intercalation means that one of the pyrazine nitrogens is protected from solvent but the other, and the nitrile group, are accessible to water through the major groove. Full conformational analysis of the DNA can be found in table A3.2.

3.3.3 Crystallisation of Λ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ with d(TAGGGTTA)

3.3.3.1 structure summary
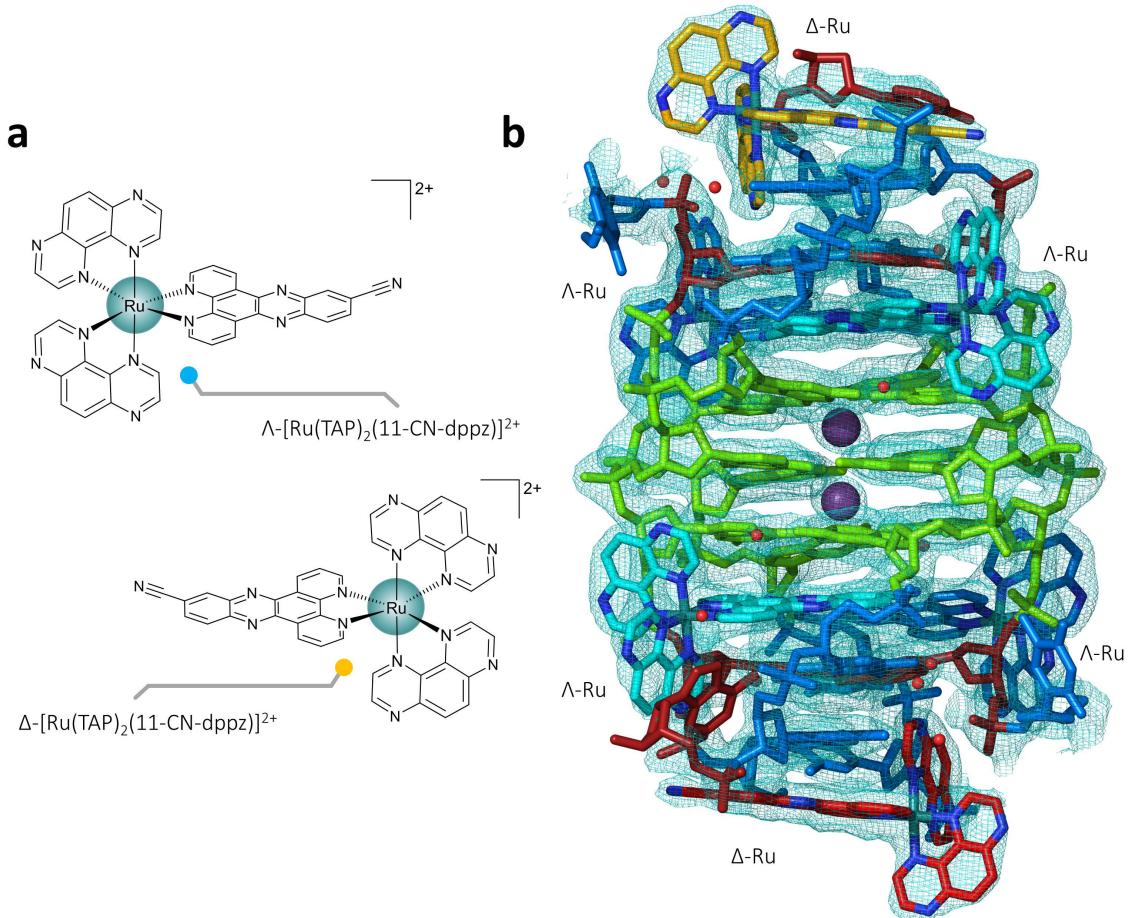
Figure 3.6 – *(a) The two optical isomers of [Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$; and (b) the crystallographic asymmetric unit of the structure showing the 2F$_o$-F$_c$ electron density map. The map is contoured at the 1σ level (0.29 eÅ^{-3}). DNA is coloured as per convention with adenine in red, guanine in green, and thymine in blue. The lambda enantiomers of [Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ are coloured with cyan/marine carbons; with the delta species in orange/red. Different shades highlight different binding environments. Potassium ions and waters are shown as purple and red spheres respectively.*

In the presence of potassium ions, the sequence d(TAGGGTTA) forms an all-parallel tetramolecular G-quadruplex. Here the quadruplex forms an anti-parallel arrangement. Crystallised from a racemic mixture of complexes, the X-ray crystal structure contains both Λ and Δ enantiomers of [Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ bound to an anti-parallel G-quadruplex assembly formed by four strands

of the sequence d(TAGGGTTA) (figure 3.6). The core of the G-tetrad stack is stabilised by two potassium cations which coordinate in square anti-prismatic geometry to the O6 oxygen atoms of the guanosines. The lambda enantiomer, which binds at a 4:1 stoichiometry in relation to the biological unit, intercalates and stacks onto both terminal tetrads of the quadruplex core stack. These interactions stabilise the guanosines in well-defined terminal tetrads, however in the central quartet all four guanosines are disordered at a ratio of 50:50 between the *syn* and *anti*-conformations. The delta isomer however stacks on both the 5' and 3' T/A ends of the biological unit, connecting symmetry mates in the crystallographic lattice and not interacting with the quadruplex core, or intercalating between any bases. In fact, every complex is in contact with a likewise isomer between asymmetric units (figure 3.7a), where the lambda species locks into another by means of stacked TAP ligands. Only a moderate number of 24 waters is observed in the structure, most within the first hydration sphere and are found coordinated to the bases, complexes, and phosphate backbone alike.
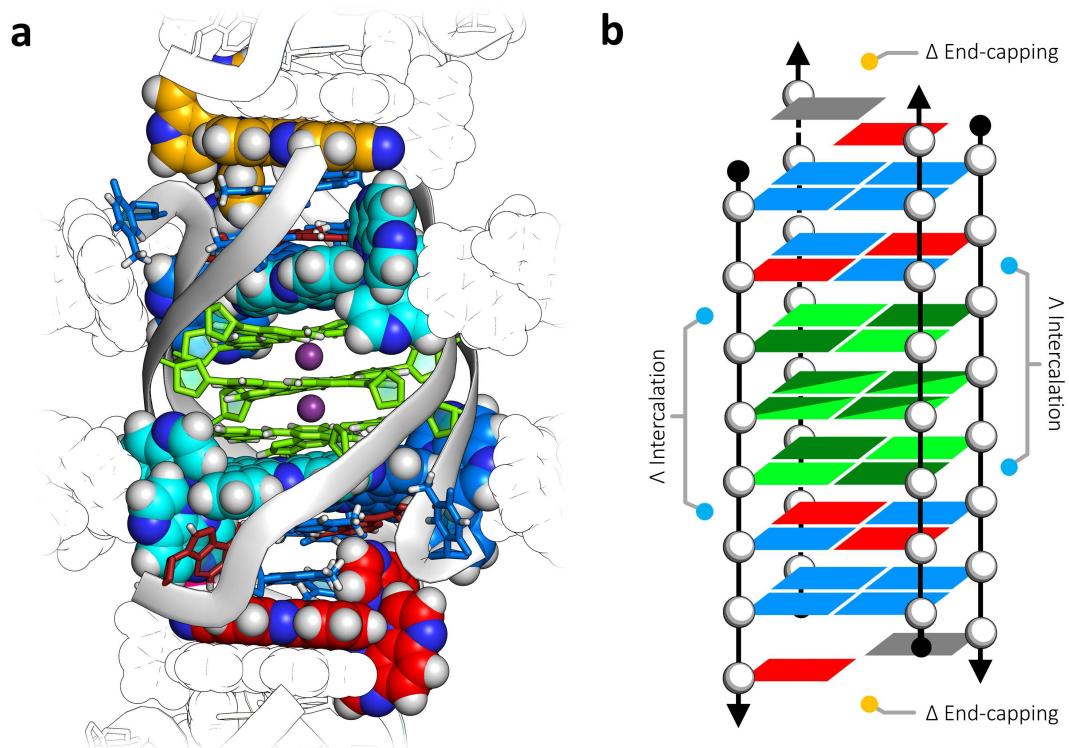
Figure 3.7 – *(a) Crystallographic model illustrating the G-quadruplex unit with all interacting ruthenium complexes. Waters have been omitted for clarity and potassium ions are shown in purple. Interactions with neighbouring units are shown in black and white. (b) Schematic to illustrate the overall structure and the sites of ruthenium interaction. DNA bases are coloured as per convention with adenine in red, guanine in green, and thymine in blue. Syn-guanosine is denoted by dark green and anti-guanosine using light green. Highly disordered bases are coloured in grey. The ruthenium complexes are coloured in cyan/marine for lambda species, and orange/red for delta species.*

3.3.3.2 DNA conformation

Conformational analysis of the DNA in the structure reveals the presence of two disparate forms of the structured DNA; both quadruplex and B-DNA (see table A3.3). The quadruplex core component consists of three guanine quartets stacked laterally and coordinated with two interpolated potassium ions between the three guanine planes. Configurational variability is common in G-quadruplex guanosine conformation; such that strand sequence, polarity, and stabilising salt, all influence the stability about the *N*-glycosidic bond. Here, as observed in other anti-parallel quadruplexes, there are a number of *syn*-guanosines; these are arranged symmetrically such that each strand follows a conformational pattern of 5'-*syn*$_1$-mixed$_2$-*anti*$_3$-3' where mixed denotes a guanosine that is 50 % disordered across both conformations. Due to opposing strand polarities, these are arranged such that if starting at any 5' end, and looking down the helical axis, the guanosines are arranged in a *syn-anti-syn-anti* pattern around both terminal quartets. Such an arrangement is observed in the antiparallel chair conformer of the unimolecular telomeric sequence; and indeed in other symmetrical antiparallel topologies. The central tetrad is disordered across two conformational tetrad arrangements such that 50 % is *syn-anti-syn-anti* and 50 % is *anti-syn-anti-syn* as observed in a clockwise fashion about any one nucleoside (figure 3.8). This type of disorder is not uncommon but here it seems to be a product of the stabilised binding pocket induced upon interaction of the pairs of lambda ruthenium complexes, where ancillary interactions lock the terminal tetrad conformation.
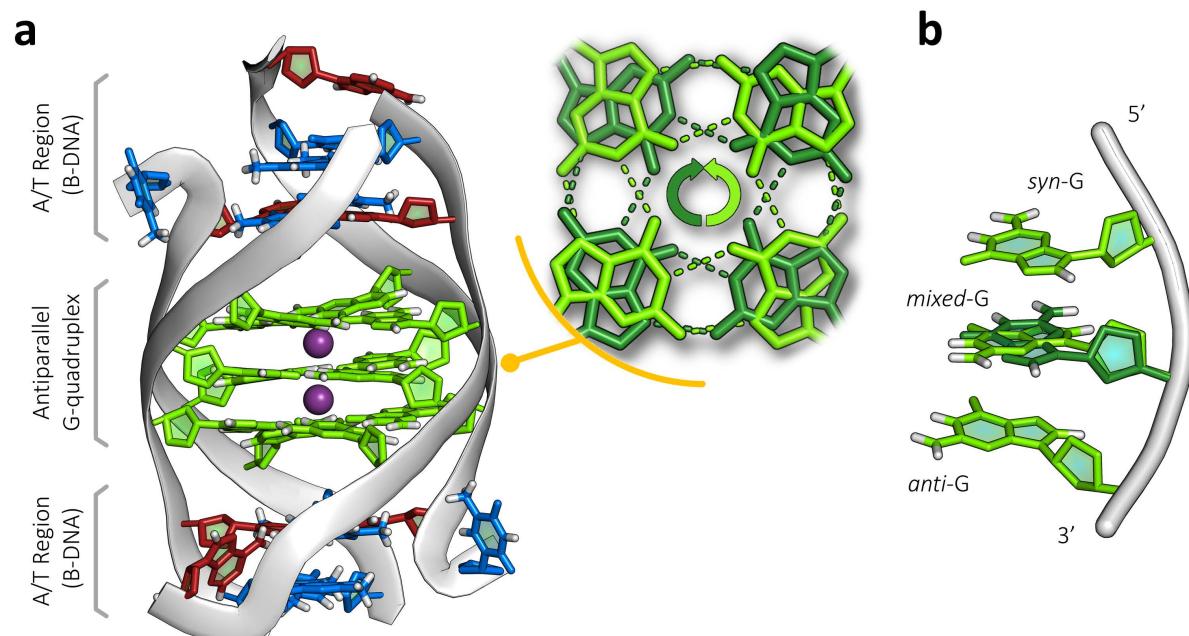
Figure 3.8 – *(a) Crystallographic model illustrating the DNA conformations present in the structure. The central disordered tetrad is highlighted to illustrate how the disorder is a 50:50 split between syn-anti-syn-anti and anti-syn-anti-syn nucleoside conformations. Waters and complexes have been omitted for clarity and potassium ions are shown in purple (b) one strand of the central G-stack showing the overall 5'-syn-mixed-anti-3' N-glycosidic conformations observed by all strands.*

In addition to the quadruplex core, the structure exhibits an additional DNA form in the thymine and adenine containing regions adjacent to the lambda binding sites. Four lambda species intercalate at the interface between the A/T rich regions and the quadruplex core, interacting with both. The four sites consist of two geometrically equivalent pairs, each stacked on the terminal tetrads with each complex in remarkably similar DNA environments and with each pair interacting with every strand. Intercalation through chain A, and transversely, intercalation through chain C by a different complex in the pair, provides subtly distinctive sites however. In each case the local step parameters adjacent to the intercalation cavities possess canonical B-DNA characteristics, albeit considering a guanine γ backbone torsion of 180-183°. This deviation from the gauche form is observably noted in previous intercalatory interactions with duplex DNA and is principally a consequence of the asymmetric binding mode. Adjacent to the intercalation steps involving chain A are T·T wobble pairs between the $T_1 \cdot T_7$ and $T_7 \cdot T_1$ residues between chains A and B respectively. This subtlety is not mirrored in the other intercalation site, where the same terminal thymine (T_1) is flipped out of the base stack in both strands C and D. In addition, only one of the four terminal adenines (A_8) could be identified in the electron density; where it π-stacks with a similarly unpaired T_7 from a crystallographic symmetry mate. Conformational analysis of the adenine and thymine containing steps shows that the majority adopt a B-DNA, even around the seemingly distorted binding pockets. It is unknown whether the all-parallel arrangement would exhibit similar strand pairing and B-DNA characteristics, but the structure here could well be regarded as a model system for a junction between duplex and quadruplex DNA. Full structural analysis can be found in table A3.3.

3.3.3.3 intercalation sites and binding cavities

Crystallisation occurred from a racemic mixture and unlike many previous systems this structure contains both isomers of $[Ru(TAP)_2(11\text{-}CN\text{-}dppz)]^{2+}$; albeit with drastically different binding modes. Six complexes bind in total to the biological unit; four are lambda enantiomers, and two are delta enantiomers (figure 3.9).
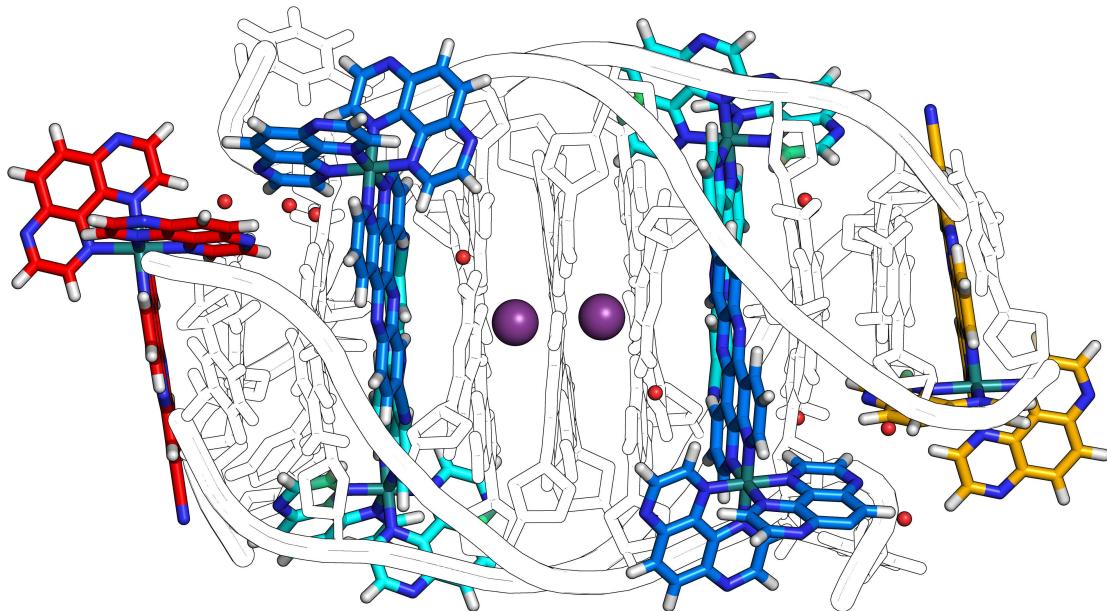
Figure 3.9 – *Crystallographic model of 5LS8 highlighting the non-DNA components; namely the six ruthenium complexes that bind in different ways to the octamer d(TAGGGTTA) biological unit. Ruthenium complexes are coloured in a way that each discrete colour is a different binding environment, where Lambda species are shown in cyan, and marine blue; and the delta species are shown in orange, and red. Potassium ions and ordered waters are shown as purple and red spheres respectively.*

3.3.3.3.1 lambda intercalation sites

Four copies of $\Lambda\text{-}[Ru(TAP)_2(11\text{-}CN\text{-}dppz)]^{2+}$ are seen to have bound, in pairs, to both terminal tetrads of the central quadruplex stack; a stoichiometry of 1 per strand of DNA. In all cases, the complex binds through either wide groove in a pseudo-symmetrical manner about the helical axis, and in terms of the direct stacking environment each is nominally equivalent. The four intercalation cavities are located into the steps of $(X)A_2G_3:(Y)G_5T_6$ and $(X)G_5T_6:(Y)A_2G_3$; where X and Y is either strands A and B, or C and D respectively (figure 3.10a). The dppz ligands of each complex adopts a pseudo intercalatory binding motif between the terminal tetrad surfaces and the adjacent T·A base pairs. The ligand directly π-stacks with two guanines in the tetrad and partially with a third via the extended substituent. The principal π-overlap of the dppz ligand in each account occurs predominantly with the *anti*-guanosine, and secondarily with both *syn*-guanosines through partial overlap and substituent contact. This is a result, not of a discrimination in purine nucleobase π-surface, but because of a secondary interaction between an ancillary TAP ligand and the puckered

deoxyribose of the *syn*-guanosine (figure 3.10b). Such an interaction, not yet seen in ligand-quadruplex structures, compels the complex into a canted binding mode such that the long axis of the dppz is asymmetrically skewed by as much as 50° in relation to the H-bonding of the guanine nucleobases ($G_3 \cdot G_5$).
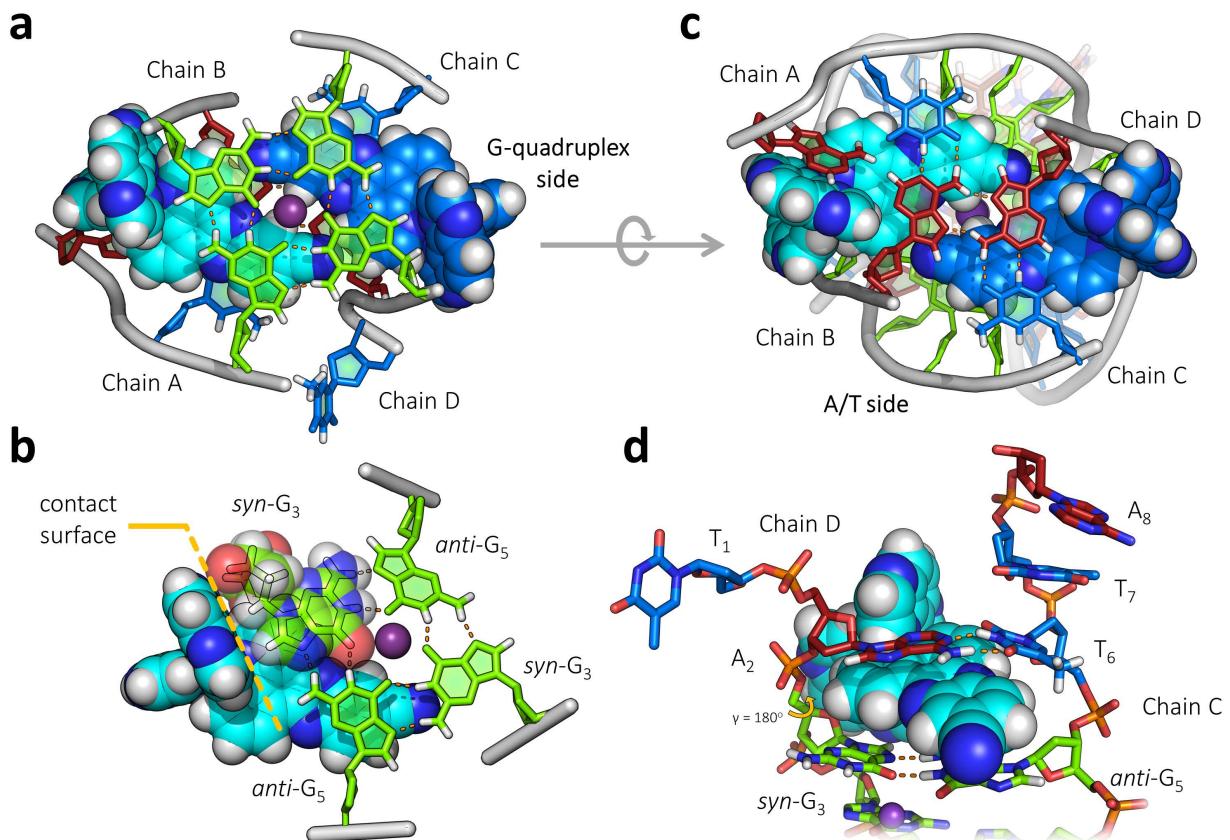
Figure 3.10 *– Crystallographic models illustrating the binding environment of the lambda species. Views of the lambda binding pairs from (a) the G-quadruplex, and (b) the A/T quartet sides. (c) similar view from the G-quadruplex side highlighting the ancillary interaction between a TAP ligand and the syn-guanosine deoxyribose. Note the principal π-stacking interactions with one syn and one anti-guanosine. (d) the ruthenium enters through a 'wide' groove but when only considering the two main interacting strands the binding site looks remarkably similar to minor groove canted intercalation into B-DNA. Waters and complexes have been omitted for clarity and potassium ions are shown in purple. DNA bases are coloured as per convention with adenine in red, guanine in green, and thymine in blue.*

A sheared T·A·A·T quartet lies on the opposite side of the intercalation cavity, however the complexes are not oriented in a manner which promotes favourable π-overlap with these nucleobases, instead running almost parallel across the T·A hydrogen bonds (figure 3.10c). This positioning may be an indication of a preference for increasing overlap with the guanine tetrad, or it is again a consequence of the intercalatory angle dictated by the *syn*-guanosine contact. Regardless, each complex can be thought to interact with one 'side' of the assembly. Removing the other 'side' of the model reveals a binding cavity that is remarkably reminiscent of B-DNA minor groove intercalation; canting upon intercalation and creating an asymmetric base step environment (i.e. large difference in γ torsion between one side of the intercalation cavity and the other)(figure 3.10d).

Interestingly the asymmetric chromophore is found to be intercalated with complete directional selectivity, that is that the nitrile groups are seen to interlock and are directed towards the dppz of the opposing complex with a complete lack of disorder. This interaction and directional preference seem to be determined by the combination of a favourable polar contact with the 2-NH$_2$ group, and an alignment of polarity with the G$_6$ carbonyl lone pair, of an opposing *syn*-guanosine in the flanking quartet.

3.3.3.3.2 delta binding sites

Delta species can be seen in the structure at a stoichiometry of two per asymmetric unit. Interestingly however, these complexes do not meaningfully interact with the nucleic acid components; that is to say they do not intercalate into the A/T rich regions or bind with the quadruplex core. Instead the complexes stack on T·T wobble pairs at the termini of the biological units formed by T$_1$/T$_7$ and T$_7$/T$_1$ in strands A and B. Importantly however, the complexes stack on top of each other, between crystallographically related asymmetric units, acting as a 'molecular glue' holding the biological units together in the crystal (figure 3.11). Each complex is stacked with their nitrile substituents directed towards the opposing TAP ligands with full directional order and in a fashion almost completely analogous to the helical-forming small molecule structures from Chapter 2.
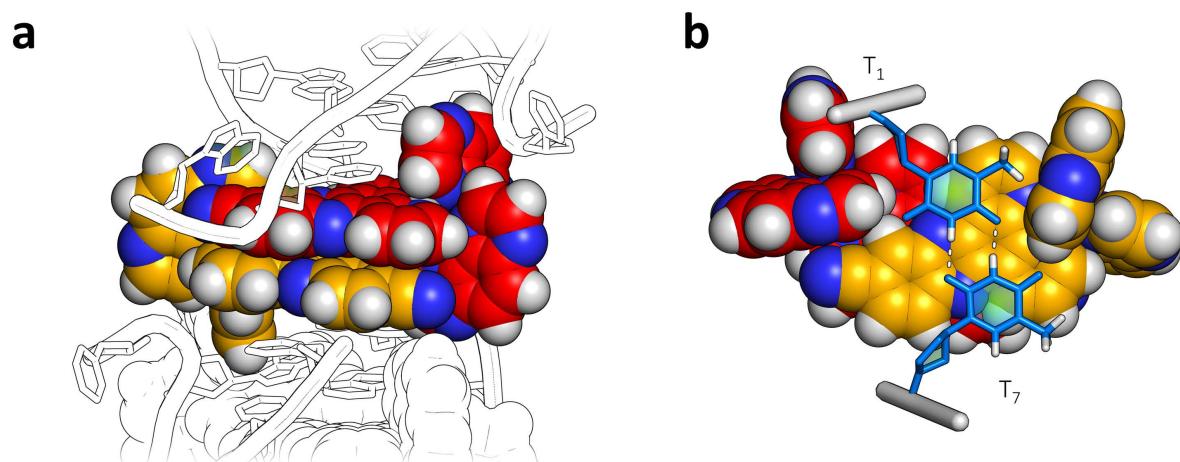
**Figure 3.11** – *Crystallographic models illustrating the binding environment of the delta species. (a) shows how the delta species, shown in red/orange, interlock together and act as an electrostatic glue between the asymmetric units. (b) the stacked delta species each interact symmetrically with a T·T wobble pair formed from T$_1$ and T$_7$ from strands A and B. Thymines are shown in blue.*

3.3.4 Circular dichroism

Synchrotron radiation circular dichroism (SRCD) was employed to aid in the confirmation of the solution state topology of the nucleic acid in the presence of the enantiomers of lambda and delta [Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ (figure 3.12a). Previous studies have described the all-parallel nature of the native assembly and this has been reconfirmed here; however, less trivial is the confirmation of strand polarity in the presence of the complexes due to the strong absorption overlap and subsequent Cotton effects of the enantiopure material, and the comparatively subtle spectral changes between topologies. Using a combination of CD scans and DNA melting analysis the G-quadruplex topology was assigned in the presence of both Λ and Δ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$.

Natively, in potassium buffered solutions, d(TAGGGTTA) forms an all-parallel structure which can be confirmed by indicative peaks in the CD spectra; a negative peak at 241 nm and a positive peak at 260 nm. These same characteristic peaks are present in the classic d(GGGG) parallel quadruplex.[15] Upon addition of Δ, the spectra of d(TAGGGTTA) exhibits a marked increase at 260 nm, and does so further at higher loadings of ruthenium (figure 3.12b). The spectral changes are almost additive, such that the observed spectra are essentially the sum of absorptions of the separate DNA and ligand spectra; implying that little interaction is present, and the native all-parallel topology is conserved. Contrarily, addition of Λ at a stoichiometry of 1:1 to biological unit not only increases the intensity of the 260 nm peak but also redshifts to 264 nm (figure 3.12c). Further addition up to a stoichiometry of 4:1, as observed in the crystal, causes the peak in the region to change drastically from a positive peak (+15 mdeg) to a negative peak (-25 mdeg) coupled with a blueshifting in absorbed light from 264 to 257 nm. The free complex contributes to this region with a negative going peak at 264 nm, however, the magnitude of this peak in the unbound spectrum is lower than is observed in the mixed spectrum (-15 mdeg). In addition, with comparison to the free ligand, the ratio of the spectroscopic bands at 280 and 300 nm shifts from 1:2.8 in the unbound spectrum to 1:1.2 in the bound species spectrum. This change arises from an increase in magnitude of the peak at 280 nm and is indicative of the presence of anti-parallel strand polarities which have maxima at 270 nm and 295 nm, and characteristic minima at 260 nm.
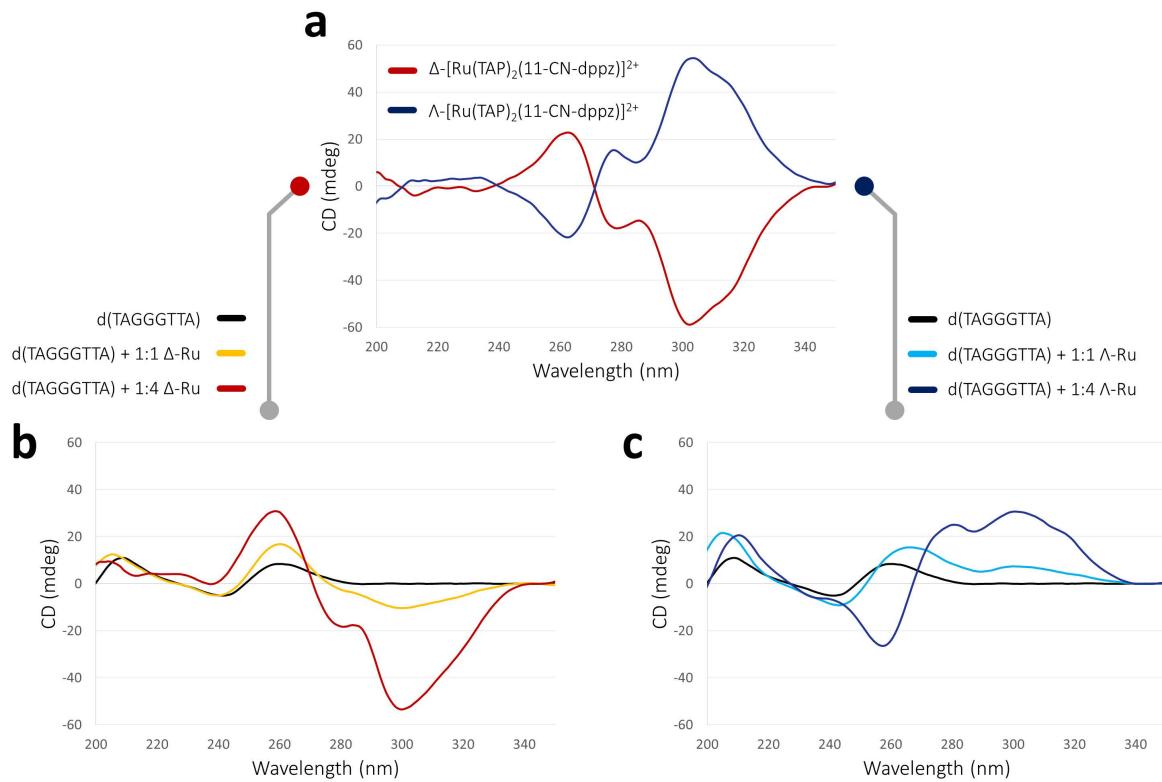
Figure 3.12 *– Circular dichroism spectra showing (a) the CD profile of the separated enantiomers of [Ru(TAP)₂(11-CN-dppz)]²⁺; d(TAGGGTTA) with increasing stoichiometric equivalents of (b) Δ-[Ru(TAP)₂(11-CN-dppz)]²⁺ or (c) Λ-[Ru(TAP)₂(11-CN-dppz)]²⁺.*

Thermal denaturation studies were additionally performed on the same systems containing both enantiomers and d(TAGGGTTA) (figure 3.13). All systems gave canonical melting profiles at 260 nm implying reasonable DNA strand associations. However, only the system containing a high stoichiometry of lambda isomer gave melting curves at higher wavelengths, namely at 280-295 nm. Since the isolated complexes show little to no change under the same temperature conditions, the melting profiles must originate from the denaturing DNA. As such, the presence of said 295 nm profiles is characteristic of antiparallel formations and is further evidence that the addition of the lambda enantiomer promotes the antiparallel formation of d(TAGGGTTA); whereas in the presence of delta species the topology is conserved. Melting profiles for each system are illustrated in figures 3.13 and A3.5. Analysis of the profiles yields a native T_m of 49 °C, where addition of delta yields a T_m of 54 and 59 °C for 1:1 and 4:1, respectively. Modest ΔT_m were to be expected if the complex doesn't intercalate and stabilise the quadruplex core and instead either electrostatically interact with the grooves or TA regions. Interestingly however, the lambda species exhibit melting temperatures of 58 and 59 °C for 1:1 and 4:1 respectively; lower than the equivalent delta systems. Of course, since the stability of the native antiparallel assembly is unknown but presumably a fair bit lower, the true stabilising effect of binding is also unknown.
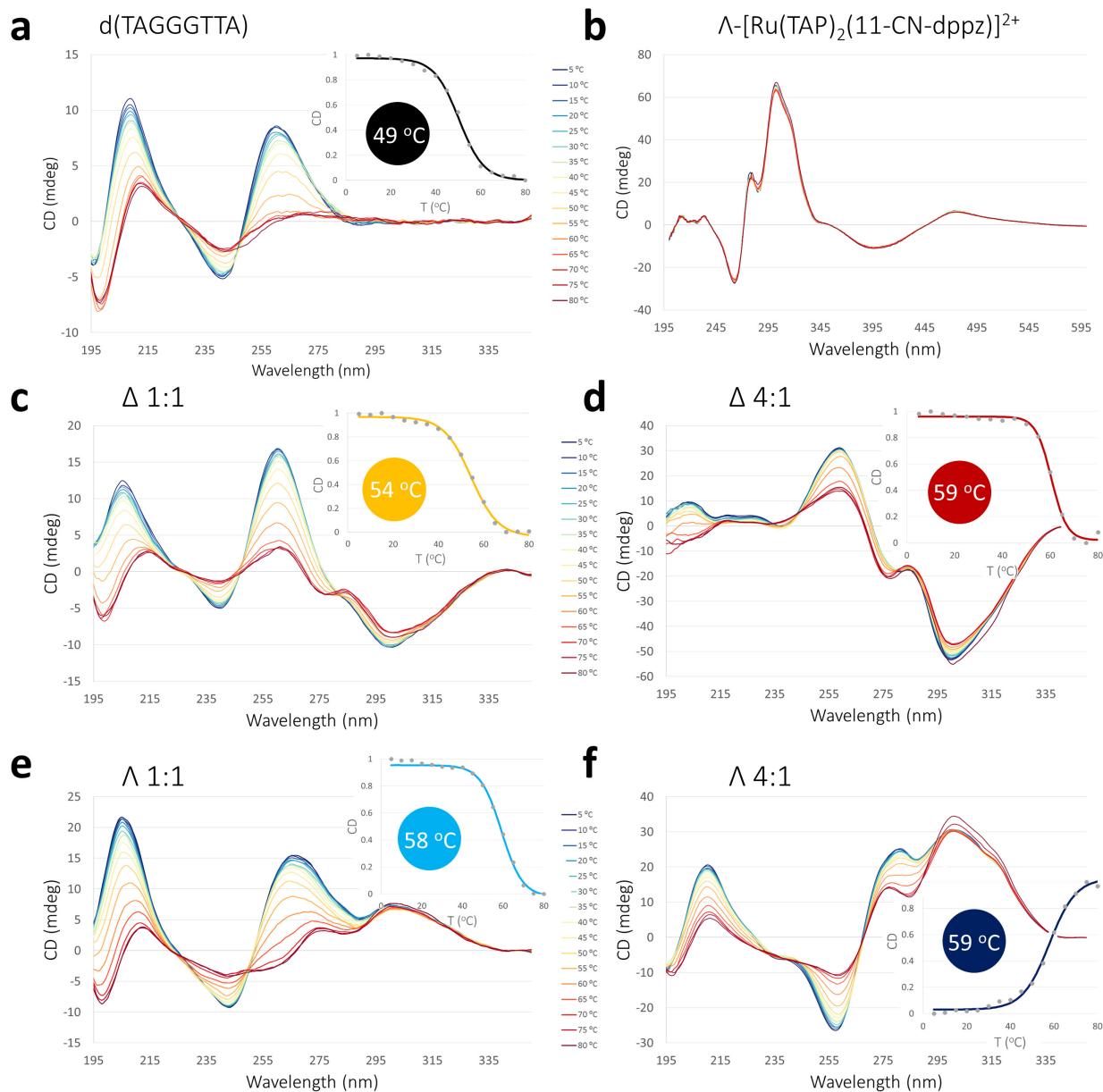
Figure 3.13 – *SRCD melting profiles with subsequent single point melting curves (inset) of: (a) native K⁺ folded d(TAGGGTTA); (b) Λ-[Ru(TAP)₂(11-CN-dppz)]²⁺; d(TAGGGTTA) with stoichiometric ratios of (c) 1:1 and (d) 1:4 Δ-[Ru(TAP)₂(11-CN-dppz)]²⁺; d(TAGGGTTA) with stoichiometric ratios of (e) 1:1 and (f) 1:4 Λ-[Ru(TAP)₂(11-CN-dppz)]²⁺. CD spectra were measured from samples containing 800 μM d(TAGGGTTA) with either 1:1 or 4:1 stoichiometric amounts of complex to biological unit, yielding a final concentration of 200 or 800 μM of respective complex. All samples contained 20 mM K-cacodylate pH 7 and 30 mM KF. Measurements were collected using cells with a 100 μm path length on beamline B23 at Diamond Light Source Ltd.*

3.4 Discussion

One of the core potential utilities for octahedral ruthenium polypyridyls is as selective luminescent probes for *in vivo* nucleic acid structure. The 'light-switch' response, which is sensitive to local hydration around the chromophore and the level of π-interaction with DNA, can suggest a great deal about the local binding environment.[16,17] It is clear from the limited literature and the work presented here on enantiomeric differences, that the lambda enantiomer will almost always

luminesce more brightly than the delta counterpart when bound to G-quadruplexes; this is contrary to binding to duplexes where the enantiomeric disparity is flipped.[18] It is evident then that evaluating the luminescence response of racemic mixtures can yield an unsatisfactory 'dampening' of the motif specificity. Despite this, complexes such as *rac*-[Ru(bpy)$_2$(11-Br-dppz)]$^{2+}$ have been reported to exhibit a 14-fold change in response when bound to d(TAGGGTTA) over natural CT-DNA.[19] However, without knowing the separate enantiomers response and/or binding preferences it is difficult to assess whether this is a result of moderated interaction of the delta enantiomer or indeed, as seen here, a heightened specificity of the lambda enantiomer in the racemic mixture. The reported Λ-[Ru(phen)$_2$(11-CN-dppz)]$^{2+}$ displays a marked selectivity for the same sequence, manifesting as a 40-fold change in peak area against CT-DNA, whilst the delta isomer emitted more brightly with natural duplex DNA than any G-quadruplex tested. This difference is in part due to the almost nil response of the lambda species with duplex DNA, an observation not noted with the parent complex. Crystallisation of said complex with duplex DNA, and of the isoelectronic TAP containing species with G-quadruplex DNA, aids in the potential elucidation of the structural rationale for this divergence.

a

b

Λ-Ru

d(TCGGCGCCGA)

d(TAGGGTTA)
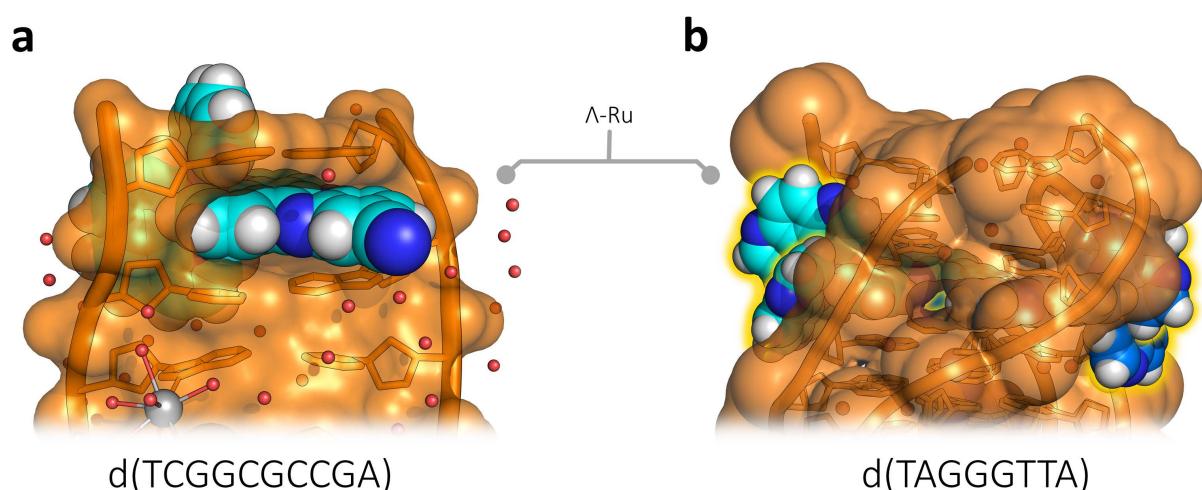
Figure 3.14 – Calculated Van der Waals surfaces from the crystallographic models of 6HWG and 5LS8. Highlighting how when bound to B-DNA (a), the majority of the distal region of the intercalating ligand is exposed in the major groove, allowing polar solvent interactions and subsequent non-radiative relaxation pathways; and that when bound to a G-quadruplex (b) the intercalating ligand is almost entirely encapsulated, protecting the 11-CN-dppz from polar solvent interaction and allowing the complex to luminesce. DNA and Van der Waals surfaces are shown in orange. Lambda complexes are shown as spheres with the carbons coloured cyan; and waters are shown as red spheres.

Whilst it could be predicted that the phenazine nitrogens are at least partially blocked when intercalated into most canonical base steps, the structural investigation of Λ-[Ru(phen)$_2$(11-CN-dppz)]$^{2+}$ bound to d(TCGGCGCCGA) shows how one side of the dppz and the entirety of the nitrile moiety is completely accessible to solvent through the major groove. Potentially then, the nitrile group may be acting as an additional point of H-bonding potential with the polar

media; yielding a 'light-switch' response which discriminates more effectively between duplex and quadruplex assemblies. In contrast, in the reported structure of Λ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ with d(TAGGGTTA), the entirety of the intercalating chromophore is encapsulated by DNA, thus preventing the water-mediated non-radiative decay of the ^{3}MLCT excited state. Calculation of the van der Waals surfaces of the two structures (figure 3.14) depicts how the observed luminescence may be related to the extent of encapsulation of the chromophore and parallels the notion that the delta isomer may not intercalate or may do so in a shallower fashion. Of course, this particular sequence has many advantages as a site of intercalation over other G-quadruplex folds, namely that the what would be loop regions in other systems are replaced by two almost canonical adjacent duplexes that provide two large quartet π-stacking surfaces and a protective hydrophobic pocket. Such sandwich π-stacking with two electron rich quartets is likely to increase the electron density on the intercalating ligand via the parallel-displaced LUMO-HOMO overlap, increasing the energy of the dark ^{3}MLCT$_{dppz}$ dppz-localised photoexcited states and thus favouring population of the previously higher-energy emissive ^{3}MLCT$_{phen}$ excited state lying on the comparatively electron-deficient ancillary ligands.

a Λ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ **b** Λ-[Ru(phen)$_2$(11-CN-dppz)]$^{2+}$ **c** Λ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$

PDB – 5LS8   PDB – 6HWG   PDB – 5NBE
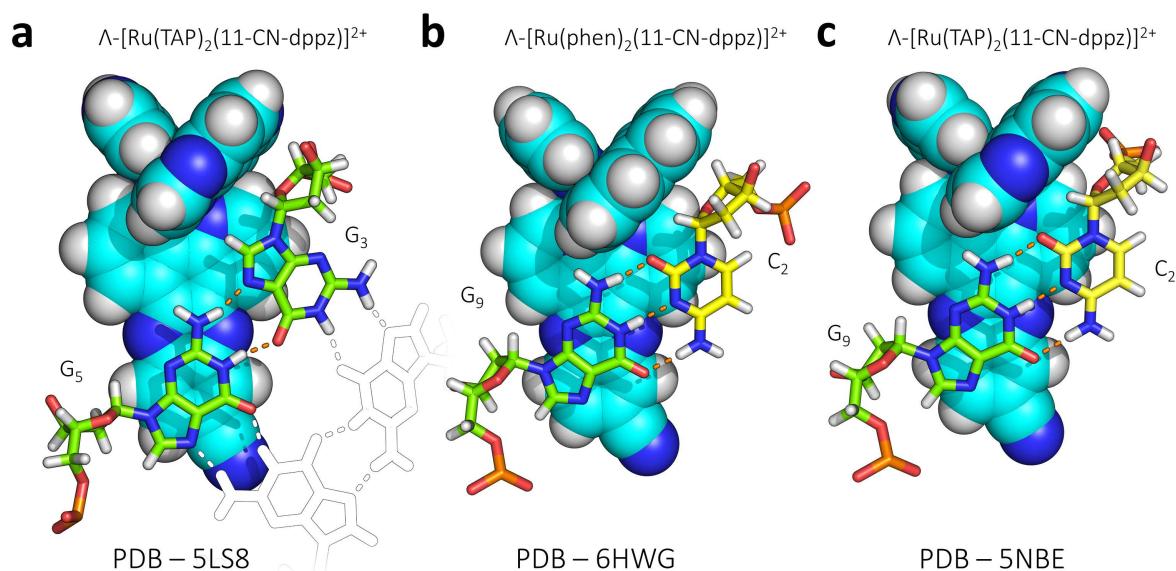
Figure 3.15 – *Crystallographic models highlighting the similarities in binding between the observed G-quadruplex and duplex modes of Λ-[Ru(phen/TAP)$_2$(11-CN-dppz)]$^{2+}$ bound to (a) the guanine tetrad; (b) and (c) B-DNA forming d(TCGGCGCCGA). Each structure exhibits favourable stacking of the ancillary phen/TAP ligand to the 5'-ribose sugar that determines the binding geometry of the intercalative mode. In the case of 5LS8 this stabilises the formation of syn-guanosine; in all cases the adjacent guanosine is anti and its polarity is aligned remarkably similarly in both tetrad and B-DNA interactions, presumably to increase favourable π-orbital overlap. Comparison of (b) and (c) shows the effectively isostructural binding of the phen and TAP complexes alike.*

The parallel quadruplex forming sequence d(TAGGGTTA) had not yet been structurally characterised using crystallography. Initial crystallisation trials were performed using both *rac*-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ and its parent, *rac*-[Ru(TAP)$_2$(dppz)]$^{2+}$, however only the nitrile derivative

crystallised successfully. The resulting structure presents the first structural insight into the interactions of mononuclear ruthenium complexes with quadruplex DNA. Of most note is the observed ability of the lambda enantiomer to induce the formation of an anti-parallel quadruplex by way of specific interactions with *syn*-guanosine. The topological selectivity of binding, which was additionally confirmed in solution by SRCD, is likely a result of a number of key binding features such as enantiospecific ancillary interactions with the deoxyriboses of the *syn*-guanosines; polar contacts between the nitrile substituents and the opposing guanines; and maximisation of the stacking interaction and polarity alignment with the guanine substituents. Interestingly however, when attempted with the RNA analogue r(TAGGGTTA) no such switch in topology was observed (see figure A3.6); this could potentially be explained by the fact that the 2'-OH group would be directly interfering with the *syn*-guanosine interaction with the ancillary ligands. Such binding features are present in previous investigations with duplex DNA, and indeed with the duplex structure presented here between the phen analogue and d(TCGGCGCCGA), suggesting that such features are key for ligand design. In the absence of any lambda species the delta enantiomer would not be able to form such interactions due to unfavourable steric clashes with the backbone, and as such would not be able to intercalate as deeply into the quadruplex core thus preventing these modes of interaction. Similarly, the dinuclear species $\Lambda\Lambda$-[{Ru(bpy)$_2$}$_2$(tpphz)]$^{4+}$ also exhibits such an interaction with the potassium folded wtTel22 sequence in solution, albeit the reduced π-surface of the bpy ligand compared to the TAP ligand leads to less meaningful overlap.[1] Figure A3.7 presents these comparisons and highlights how this interaction is unlikely to occur with delta (and does not with the $\Delta\Delta$ of the above species). With respect to the lambda enantiomers these interactions stabilise the terminal tetrads of the central quartet stack, however, a 50:50 disordering of the central tetrad is observed in the structure. This suggests that the topological requirement for the antiparallel polarity of strands would be unstable in the absence of these stabilising features and may be the reason that in the presence of delta the topology remains parallel. This central tetrad disorder is almost to be expected in an antiparallel system with an odd number of tetrads (of which the flanking tetrads are effectively locked) and such degrees of freedom due to the absence of structurally restraining loop regions. MD simulations and free energy analysis have highlighted the favourable base step stability incurred by 5'-*syn-anti*-3' steps and conversely the instability of 5'-*syn-syn*-3' and 5'-*anti-syn*-3' steps.[20] As a result, this could cause the central step to split between two energetically similar conformations where only the central quartet is free to shift due to the flanking quartets' interaction with the complexes. Interestingly, the 5'-*syn* faces of guanines have been demonstrated to be especially susceptible to chemical and oxidative damage, where concomitant tracts of guanines are more susceptible to oxidative damage than the lone nucleosides.[21,22] In addition, damage and oxidative stress on guanines

in promotor regions have been shown to both independently stall or induce transcription in cancer associated genes such as VEGF-A.[23] Complexes such as parent species [Ru(TAP)$_2$(dppz)]$^{2+}$ have been shown to form covalent adducts with guanine under the correct conditions, causing direct DNA damage. As such, binding with any order of specificity to these regions could be a potential route to selectively damaging distinct topologies of G-quadruplexes in genomic DNA.

Asymmetrical substitution, in respect to the intercalating dppz chromophore, has been explored in relation to binding to G-quadruplexes. One study found that the addition of a single bromine to the dppz of *rac*-[Ru(bpy)$_2$(dppz)]$^{2+}$ yielded a light-switch species that exhibits a 14-fold increase in G-quadruplex selectivity in luminescence response (d(TAGGGTTA) against CT DNA) when compared to the parent complex. Furthermore, it was also reported that the same *rac*-[Ru(bpy)$_2$(11-Br-dppz)]$^{2+}$ exhibits a tenfold increase in binding constant to the same quadruplex, from K_b = 2.1 x 10^5 M^{-1}, to K_b = 19 x 10^5 M^{-1}; highlighting the often profound effect of seemingly minor alterations to the binding environment.[19] Several studies have also noted topological specificity when exploring asymmetric intercalators with G-quadruplex DNA. The complexes, *rac*-[Ru(bpy)$_2$(phenselanazole)]$^{2+}$ and *rac*-[Ru(bpy/phen)$_2$(ptpn)]$^{2+}$ (figure 3.16), have been reported to exhibit distinct selectivities for the potassium folded antiparallel wtTel22 sequence over similar conformations, and can induce the folding of the quadruplex even in the absence of potassium cations.[24,25] Similarly, *rac*-[Ru(bpy)$_2$(itatp)]$^{2+}$ can convert the parallel potassium folded form of wtTel22 into an antiparallel form by titration, and in addition can destroy the folded parallel structure of *c-myc*.[26] However, despite the large interest in tackling topological specificity in binding, very little justification, or indeed structural evidence, is available for the phenomenon. Could an affinity for the *syn*-guanosine's ribose be the key to these topological specificities?
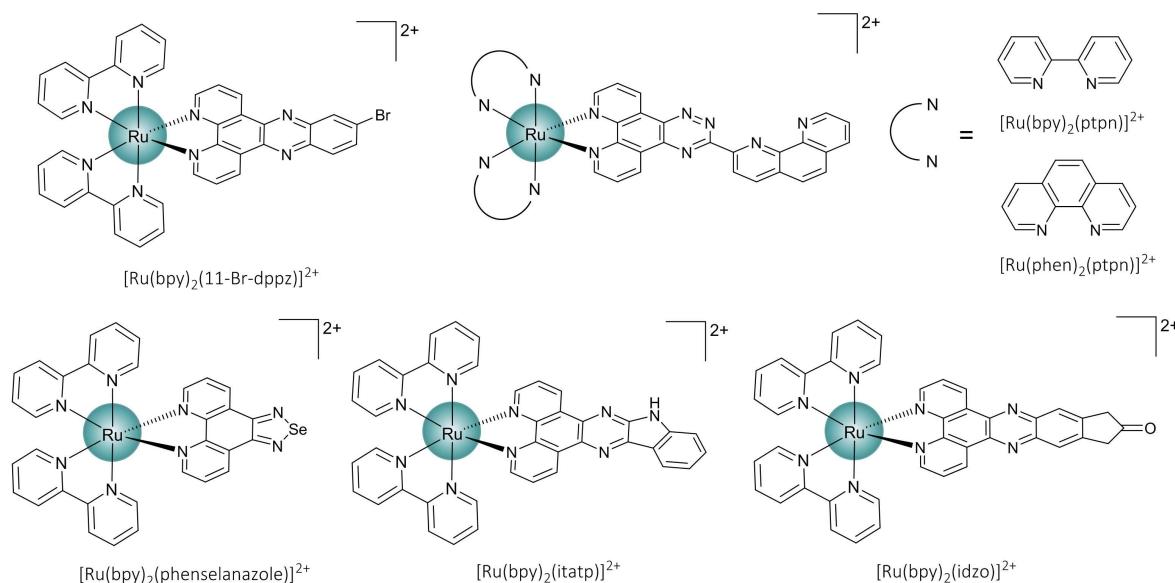
Figure 3.16 – *Ruthenium based complexes discussed that exhibit interesting binding properties towards G-quadruplex DNA.*

The nitrile orientational stability in the reported structure is partly in agreement with previous and present observations from the group of complexes containing substituted dppz ligands. In structural studies of complexes containing asymmetrically substituted methyl groups, the complex was always found to intercalate to B-DNA with complete directional preference, that is, with the methyl groups exclusively orientated towards one side of the groove.[27] As discussed in an earlier chapter however, this has not held true for other substitutions that are less hydrophobic (such as halides and nitriles), where they typically are distributed unequally between two positions.[28] Such certainty in binding orientation can be important if the complexes were to be utilised as photophysical probes where the substitutions could be used as handles for understanding local binding environment; were the orientations to be split, it would be likely that multiple excited state lifetimes, luminescence responses, or IR vibrations, would exist in a single system. Structural insights are however limited to a handful of duplex crystal structures.
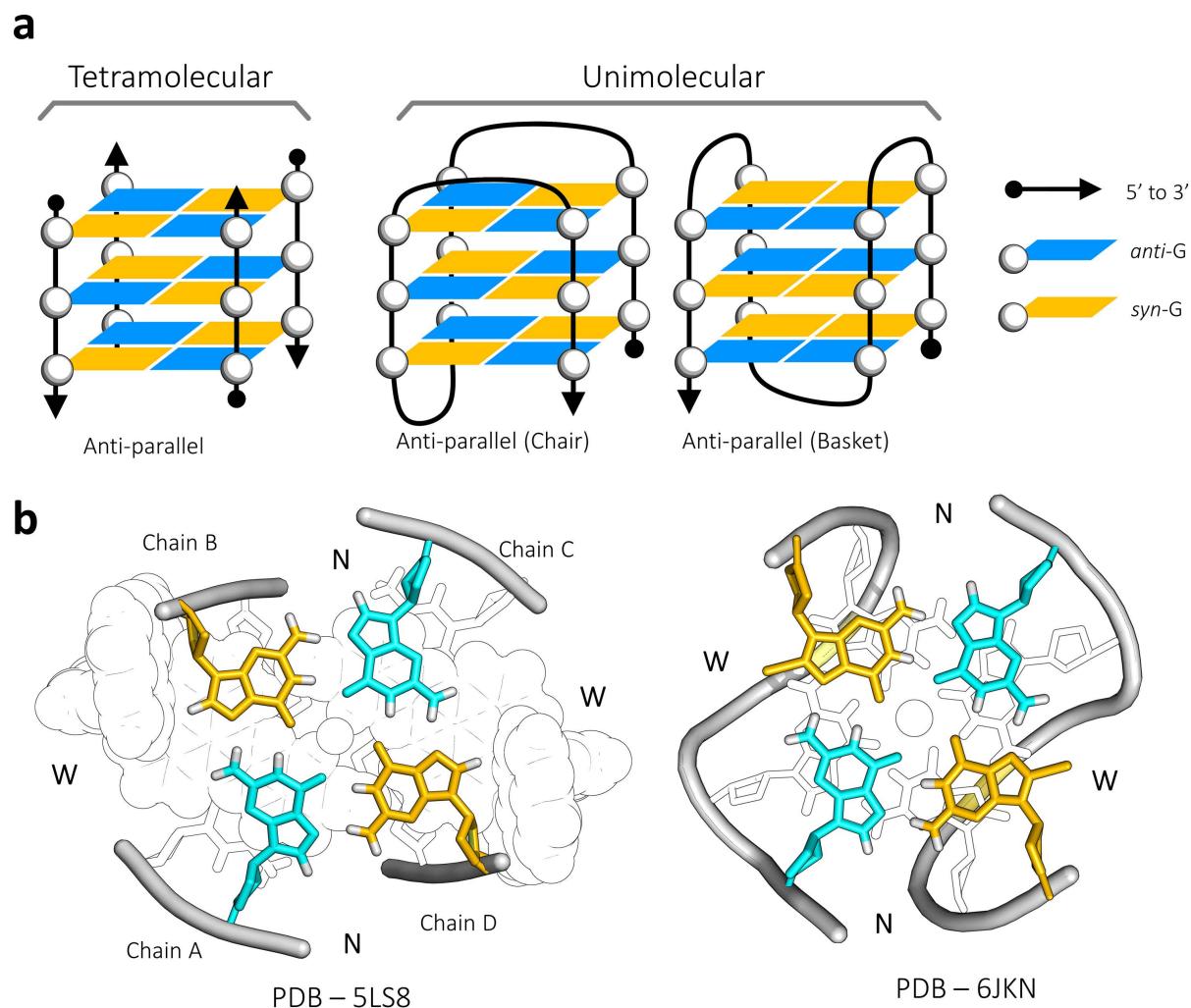
Figure 3.17 – *(a) topologies of the discussed G-quadruplexes; and (b) terminal tetrad conformations of the reported structure (5LS8), and an antiparallel chair forming potassium folded tel21 sequence (6JKN), highlighting the similarity in the groove widths.*

Chiral discrimination is an important feature of binding to DNA, especially so with the topological variation that is prevalent in G-quadruplex assemblies (figure 3.17a). Antagonistically to duplex interaction, the lambda enantiomer is known to bind more strongly to G-quadruplexes than the delta. This notion was echoed in a study of the imidazolone derivatised [Ru(bpy)$_2$(dppz-idzo)]$^{2+}$ with the telomeric quadruplex wtTel22.[29,30] In this investigation, it was found that the complex exhibits a high chiral selectivity when bound to the sodium form of the sequence, with the lambda isomer stabilising and luminescing to a greater extent than the delta. The amphimorphic sequence, which folds natively into an antiparallel basket with sodium but at high potassium concentrations folds into a hybrid-1 mixed topology, conserves its anti-parallel topology in the presence of the complex. Changes in loop morphology, such as the introduction of diagonal loops in the basket form, alter the groove topology. In the basket form these grooves, measured clockwise from the 5' end, follow a width pattern of MNMW, where N = narrow, M = medium, and W = wide, whereas the chair form has an NWNW groove pattern.[31,32] The presented structure here follows a similar groove pattern as the chair form, with the lambda complexes inserting through the wide grooves, widening them further at the expense of further narrowing the narrow grooves (figure 3.17b). The antiparallel chair and the structure here also share the same *N*-glycosidic base conformation pattern of *syn-anti-syn-anti*. As a result of this, the complexes primarily interact with one *syn* and one *anti*-guanosine when binding via the wide grooves. The sodium form basket however follows a *syn-syn-anti-anti* pattern; as such, if Λ-[Ru(bpy)$_2$(dppz-idzo)]$^{2+}$ were to enter through the medium grooves, a similar π-stacking environment could potentially arise with one *anti* and one *syn*-guanosine. Job plots of the interactions imply a 4:1 binding stoichiometry of complex to biological unit, again consistent with what is observed herein. It is difficult to perceive however, with a short diagonal loop partially covering a terminal tetrad, that four complexes could bind substantially. Alternatively, as seen with earlier examples, the complexes may convert the basket assembly to the antiparallel chair topology. This is plausible since the two conformations are notoriously difficult to distinguish apart using circular dichroism, especially so in the presence of bound chiral molecules.[33] In the presence of the potassium folded hybrid-1 quadruplex, a lower stoichiometry of lambda binding was observed in comparison to the sodium form. This disparity to an extent highlights the influence of loop topology on binding cavity formation; where the sodium form incorporates a propeller loop that passes diagonally across the groove. By comparison with the structure here, it is evident that such a loop would prevent proper binding of at least one complex; either by direct steric inhibition or by dissociation of the flanking T/A region.

3.5 Summary

Using high-throughput luminescence screening, the probe response of a range of ruthenium based complexes was examined, where enantiomeric and substitutional effects were investigated. Mirroring previous observations, lambda species were found to luminesce with more specificity to higher order motifs, whereas delta isomers universally luminesced more brightly with duplex DNA. The probe potential of the nitrile derivative, [Ru(phen)$_2$(11-CN-dppz)]·Cl$_2$ was evaluated, and the lambda isomer was found to be quite specific to G-quadruplex sequences, especially the intermolecular d(TAGGGTTA) sequence (40-fold fold change over CT-DNA); albeit at a cost of a lower luminescence yield.

To understand better the interactions of such complexes to quadruplex DNA, the isostructural TAP analogue was successfully crystallised with d(TAGGGTTA) yielding the first structural data of a mononuclear ruthenium complex with G-quadruplex DNA. The structure contains both enantiomers of the complex, bound in quite distinct modes, with the lambda intercalating at a 4:1 stoichiometry at the quadruplex-duplex junctions. The structure revealed many intricacies that have allowed for the better understanding of measurements in solution and confirmed that the lambda isomer binds deeply into the biological unit where it is fully protected from excited state quenching. In addition, the phen analogue was crystallised with the duplex sequence d(TCGGCGCCGA) where it was shown to bind isostructurally to the TAP analogue, allowing for a direct comparison between structures. In the duplex structure over half of the chromophore is accessible through the major groove to the aqueous solvent with the nitrile group unprotected from polar interaction, thus providing a structural rationale for the observed motif specificity in luminescence response.

3.6 References

1   Wilson, T., Costa, P. J., Félix, V., Williamson, M. P. & Thomas, J. A. Structural studies on dinuclear ruthenium(II) complexes that bind diastereoselectively to an antiparallel folded human telomere sequence. *J. Med. Chem.* **56**, 8674–8683 (2013).

2   Hussain, R. *et al.* CDApps: integrated software for experimental planning and data processing at beamline B23, Diamond Light Source. *J. Synchrotron Radiat.* **22**, 465–468 (2015).

3   Waterman, D. G. *et al.* The DIALS framework for integration software. *CCP4 Newslett. Protein Cryst.* **49**, 13–15 (2013).

4   Evans, P. Scaling and assessment of data quality. *Acta. Cryst. D.* **62**, 72–82 (2006).

5   Winter, G., Lobley, C. M. C. & Prince, S. M. Decision making in xia2. *Acta. Cryst. D.* **69**, 1260-1273 (2013).

6   Howell, P. L. & Smith, G. D. Identification of heavy-atom derivatives by normal probability methods. *J. Appl. Cryst.* **25**, 81–86 (1992).

7   Sheldrick, G. M. A short history of SHELX. *Acta. Cryst. A.* **64**, 112–122 (2008).

8   Collaborative Computational Project Number 4. The CCP4 suite: programs for protein crystallography. *Acta. Crystallogr. D.* **50**, 760–763 (1994).

9   Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta. Crystallogr. D.* **66**, 486–501 (2010).

10   Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta. Crystallogr. D.* **53**, 240–255 (1997).

11   Wang, Y. & Patel, D. J. Solution structure of the human telomeric repeat $d[AG_3(T_2AG_3)_3]$ G-tetraplex. *Structure* **1**, 263–282 (1993).

12   Xu, Y., Noguchi, Y. & Sugiyama, H. The new models of the human telomere $d[AGGG(TTAGGG)_3]$ in K^+ solution. *Bioorg. Med. Chem.* **14**, 5584–5591 (2006).

13   Luu, K. N., Phan, A. T., Kuryavyi, V., Lacroix, L. & Patel, D. J. Structure of the human telomere in K^+ solution: an intramolecular (3+1) G-quadruplex scaffold. *J. Am. Chem. Soc.* **128**, 9963–9970 (2006).

14   Gill, M. R. *et al.* A ruthenium(II) polypyridyl complex for direct imaging of DNA structure in living cells. *Nat. Chem.* **1**, 662-667 (2009).

15   Kypr, J., Kejnovská, I., Renciuk, D. & Vorlíčková, M. Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res.* **37**, 1713–1725 (2009).

16   Véry, T. *et al.* Photophysical properties of ruthenium(II) polypyridyl DNA intercalators: Effects of the molecular surroundings investigated by theory. *Chem. Eur. J.* **20**, 12901–12909 (2014).

17   Very, T., Despax, S., Hébraud, P., Monari, A. & Assfeld, X. Spectral properties of polypyridyl ruthenium complexes intercalated in DNA: theoretical insights into the surrounding effects of $[Ru(dppz)(bpy)_2]^{2+}$. *Phys. Chem. Chem. Phys.* **14**, 12496–12504 (2012).

18   Jia, F., Wang, S., Man, Y., Kumar, P. & Liu, B. Recent developments in the interactions of classic intercalated ruthenium compounds: [Ru(bpy)$_2$dppz]$^{2+}$ and [Ru(phen)$_2$dppz]$^{2+}$ with a DNA molecule. *Molecules* **24**, 769 (2019).

19   Wachter, E., Moyá, D., Parkin, S. & Glazer, E. C. Ruthenium complex 'light switches' that are selective for different G-quadruplex structures. *Chem. Eur. J.* **22**, 550–559 (2016).

20   Cang, X., Šponer, J. & Cheatham 3rd, T. E. Explaining the varied glycosidic conformational, G-tract length and sequence preferences for anti-parallel G-quadruplexes. *Nucleic Acids Res.* **39**, 4499–4512 (2011).

21   Fleming, A. M. & Burrows, C. J. G-quadruplex folds of the human telomere sequence alter the site reactivity and reaction pathway of guanine oxidation compared to duplex DNA. *Chem. Res. Toxicol.* **26**, 593–607 (2013).

22   Saito, I. *et al.* Mapping of the hot spots for DNA damage by one-electron oxidation: Efficacy of GG doublets and GGG triplets as a trap in long-range hole migration. *J. Am. Chem. Soc.* **120**, 12686–12687 (1998).

23   Fleming, A. M., Zhu, J., Ding, Y. & Burrows, C. J. Location dependence of the transcriptional response of a potential G-quadruplex in gene promoters under oxidative stress. *Nucleic Acids Res.* **47**, 5049–5060 (2019).

24   Yuan, F. *et al.* Synthesis of a ruthenium(II) polypyridine complex with 1,10-phenanthrolineselenazole as ligand and investigation of its G-quadruplex DNA-binding properties. *J. Coord. Chem.* **65**, 1246–1257 (2012).

25   Chen, X. *et al.* Targeting telomeric G-quadruplexes with the ruthenium(II) complexes [Ru(bpy)$_2$(ptpn)]$^{2+}$ and [Ru(phen)$_2$(ptpn)]$^{2+}$. *Dalt. Trans.* **42**, 4386–4397 (2013).

26   Yu, H., Zhao, Y., Mo, W., Hao, Z. & Yu, L. Ru-indoloquinoline complex as a selective and effective human telomeric G-quadruplex binder. *Spectrochim. Acta A.* **132**, 84–90 (2014).

27   Hall, J. P., Beer, H., Buchner, K., Cardin, D. J. & Cardin, C. J. The structural effect of methyl substitution on the binding of polypyridyl Ru–dppz complexes to DNA. *Organometallics* **34**, 2481–2486 (2015).

28   McQuaid, K., Hall, J. P., Brazier, J. A., Cardin, D. J. & Cardin, C. J. X-ray crystal structures show DNA stacking advantage of terminal nitrile substitution in Ru-dppz complexes. *Chem. Eur. J.* **24**, 15859–15867 (2018).

29   Yao, J.-L., Gao, X., Sun, W., Shi, S. & Yao, T.-M. [Ru(bpy)$_2$dppz-idzo]$^{2+}$: a colorimetric molecular 'light switch' and powerful stabilizer for G-quadruplex DNA. *Dalt. Trans.* **42**, 5661–5672 (2013).

30   Hu, X. *et al.* Regulation of multi-factors (tail/loop/link/ions) for G-quadruplex enantioselectivity of Δ- and Λ- [Ru(bpy)$_2$(dppz-idzo)]$^{2+}$. *Dalt. Trans.* **47**, 5422–5430 (2018).

31   Geng, Y. *et al.* The crystal structure of an antiparallel chair-type G-quadruplex formed by bromo-substituted human telomeric DNA. *Nucleic Acids Res.* **47**, 5395–5404 (2019).

32   Karsisiotis, A. I., O'Kane, C. & Webba da Silva, M. DNA quadruplex folding formalism – A tutorial on quadruplex topologies. *Methods* **64**, 28–35 (2013).

33   Vorlíčková, M., Chládková, J., Kejnovská, I., Fialová, M. & Kypr, J. Guanine tetraplex topology of human telomere DNA is governed by the number of (TTAGGG) repeats. *Nucleic Acids Res.* **33**, 5851–5860 (2005).

4 Structural Study of a Ruthenium Complex bound to a Truncated Telomeric G-quadruplex

Contribution statement

Complex synthesis, enantiomeric purification, characterisation, SR-circular dichroism, crystallisation, phasing, model building, and data analysis were performed by Kane McQuaid. X-ray diffraction data collection was performed by Dr James Hall. Lena Baumgartner and Prof. David Cardin carried out ligand synthesis and offered general synthetic support.

Part of the work presented in this chapter was featured in the following publication:

McQuaid, K., Hall, J.P., Baumgartner, L., Cardin, D.J., and Cardin, C.J. Three thymine-adenine binding modes of the ruthenium complex Λ-[Ru(TAP)$_2$(dppz)]$^{2+}$ to the G-quadruplex d(TAGGGTT) shown by X-ray crystallography. *Chem. Commun.* (2019), **55**, 9116-9119.

The structure listed below was a direct output from the work presented and was published on the Protein Data Bank with the following details:

6RNL - Λ-[Ru(TAP)$_2$(dppz)]$^{2+}$ bound the G-quadruplex forming sequence d(TAGGGTT)

4.1 Introduction

The results obtained in the previous chapter provided useful structural knowledge presenting some of the possible binding modes of ruthenium complexes to G-quadruplex DNA. However, the crystallisation was not easily reproduced and was unsuccessful when screening other complexes. As such, a new set of G-quadruplex forming sequences were screened with a range of different complexes in an attempt to discover a crystal system that could be more easily reproduced. One of the sequences in the screen, d(TAGGGTT), was designed following the observation of disordered terminal adenines in the previous structure. It was conceived that the truncated sequence may provide a better overall packing environment without affecting the topology of the quadruplex or the enantiospecificity of the observed interactions. The sequence, either natively or in the presence of ligands, has never been structurally characterised.

4.2 Methodology

Oligonucleotides were purchased from Eurogentec as HPLC-purified solids and were used without further purification. Unless otherwise stated, all other materials and chemicals were sourced from Sigma Aldrich or Honeywell research chemicals. Most methodology can be found in chapter 7.

4.2.1 Synthesis, characterisation, and enantiomeric separation

Synthesis of the complexes; rac-[Ru(TAP)$_2$(dppz)]$^{2+}$, and rac-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$, were carried out using modifications of previously published methodology which are described in detail in chapter 7. Enantiomeric purification methodology is described in section 7.2.4, with their eluting conditions and the subsequent circular dichroism signals of the optically pure enantiomers shown in figure A3.1.

4.2.2 Synchrotron radiation circular dichroism (SRCD)

SRCD spectra and subsequent CD melting analyses were collected on beamline B23 at Diamond Light Source Ltd. d(TAGGGTT) concentration was kept constant at 800 µM, giving a final concentration of 200 µM of quadruplex unit. The concentration of complex for a 1:1 molar ratio was therefore 200 µM and for a 4:1 molar ratio was 800 µM. All samples contained 20 mM K-cacodylate pH 7, and 30 mM KF. All samples were mixed, heated to 95 ^{o}C for 5 minutes, and then allowed to cool slowly to room temperature before measurement. Scans were measured in quartz plates with cell path lengths of 100 µm; whereas the melting analyses were performed in quartz 'H' cells of path length 100 µm. All spectra were acquired using a 1 second integration time per nm, with a 1 nm slit, between 180-350 nm and then were cut according to an appropriate PMT voltage. This results in a cut-off of 200 nm being applied to the spectra. Final plots were both background and offset corrected using the CDApps suite.[1]

4.2.3 Crystallisation of Λ-[Ru(TAP)$_2$(dppz)]$^{2+}$ with d(TAGGGTT)

4.2.3.1 crystallisation conditions

Crystals containing the oligonucleotide d(TAGGGTT) and the ruthenium complex [Ru(TAP)$_2$(dppz)]$^{2+}$ were grown from sitting drops through the vapour diffusion of water at 18 ^{o}C. Crystals suitable for X-ray diffraction experiments were obtained from various differing conditions, however not all conditions gave rise to well-diffracting samples. The solution forming the sitting drops constituted of two components; 1µL of a pre-annealed mixture of the single stranded oligonucleotide at 250 µM with the complex rac-[Ru(TAP)$_2$(dppz)]·Cl$_2$ at 500 µM in a 100 mM KCl buffer; and 1 µL of a solution containing 35 % v/v Tacsimate^{TM} pH 6.0 and 1 mM spermine; buffered to pH 6.0 using 50 mM

sodium cacodylate trihydrate. The sitting drop was equilibrated against 100 μL of 35 % v/v 2-methyl-2, 4-pentanediol (MPD) in H_2O forming small dark orange/red hexagonal rods within 2 weeks of preparation.

4.2.3.2 data collection, refinement and analysis

The data were collected at Diamond Light Source Ltd., on beamline I03 using radiation with a wavelength of 0.5570 Å from a flash cooled crystal at 100 K. 360° of data were collected with an oscillation of 0.1° per frame, generating 3600 images. The resulting data were processed using DIALS[2] and Aimless[3] through the xia2[4] pipeline and gave an anomalous signal with a mid-slope of anomalous normal probability[5] of 1.246, finding 8849 unique reflections to a resolution of 1.88 Å. The structure was solved using the anomalous scattering of ruthenium by single wavelength anomalous dispersion with the Phaser-EP pipeline in the PHENIX software package.[6,7] The crystallographic model was built using WinCoot[8] and refined using Phenix.refine[9] to give a final R_{work} of 0.1872 and an R_{free} of 0.2145 reserving 5 % of the total reflections for the R_{free} set. Figures were produced using the PyMOL software suite. The structure is deposited in the Protein Data Bank with PDB accession ID: 6RNL. Table 4.1 highlights the main data collection and refinement statistics.

4.2.3.3 data collection statistics

Table 4.1 - *Crystallisation, data collection, and processing parameters/refinement results of the crystal structure.*

Crystallisation Parameters	$[Ru(TAP)_2(dppz)]^{2+}$ with d(TAGGGTT)$_4$
Crystal Morphology	Hexagonal Rod
Growth Temperature (K)	291
Crystal Size (µm)	20x20x300
Growth Time	3 weeks
Data Collection	
Beamline	I03
X-Ray Wavelength (Å)	0.557
Transmission (%)	40.01
Beamsize (µm)	50x20
Exposure Time (s)	0.05
Nº Images/Oscillation (º)	3600/0.10
Space Group	$P\,6_5$
Cell Dimensions a, b, c (Å); α, β, γ (º)	38.53, 38.53, 128.77; 90, 90, 120
Data Processing	
Resolution (Å)	32.29 - 1.88 (1.91 - 1.88)
R_{merge}	0.120 (3.986)
R_{meas}	0.1233 (3.986)
R_{pim}	0.027 (1.003)
Nº Observations	175,231 (7823)
Nº Unique Observations	8849 (465)
$I/\sigma I$	14.3 (0.7)
$CC_{1/2}$	0.999 (0.585)
Completeness (%)	100.00 (100.00)
Multiplicity	19.8 (16.8)
Mid-slope of anom normal probability	1.246
* Outer shell statistics shown in parentheses	
Refinement	
Phase Solution Method	SAD
Resolution	32.3 (1.88)
Nº Reflections	8708
R_{work}/R_{free}	0.1872/0.2145
Nº Atoms	
DNA	576
Metal Complex	204
Water	75
Average B Factors (Å^2)	
DNA	44.16
Metal Complex	42.62
Water	40.21
rmsd	
Bond Lengths (Å)	0.013
Bond Angles (o)	1.0
PDB ID	**6RNL**

4.3 Results

4.3.1 Crystal structure of Λ-[Ru(TAP)$_2$(dppz)]$^{2+}$ with d(TAGGGTT)

4.3.1.1 structure summary
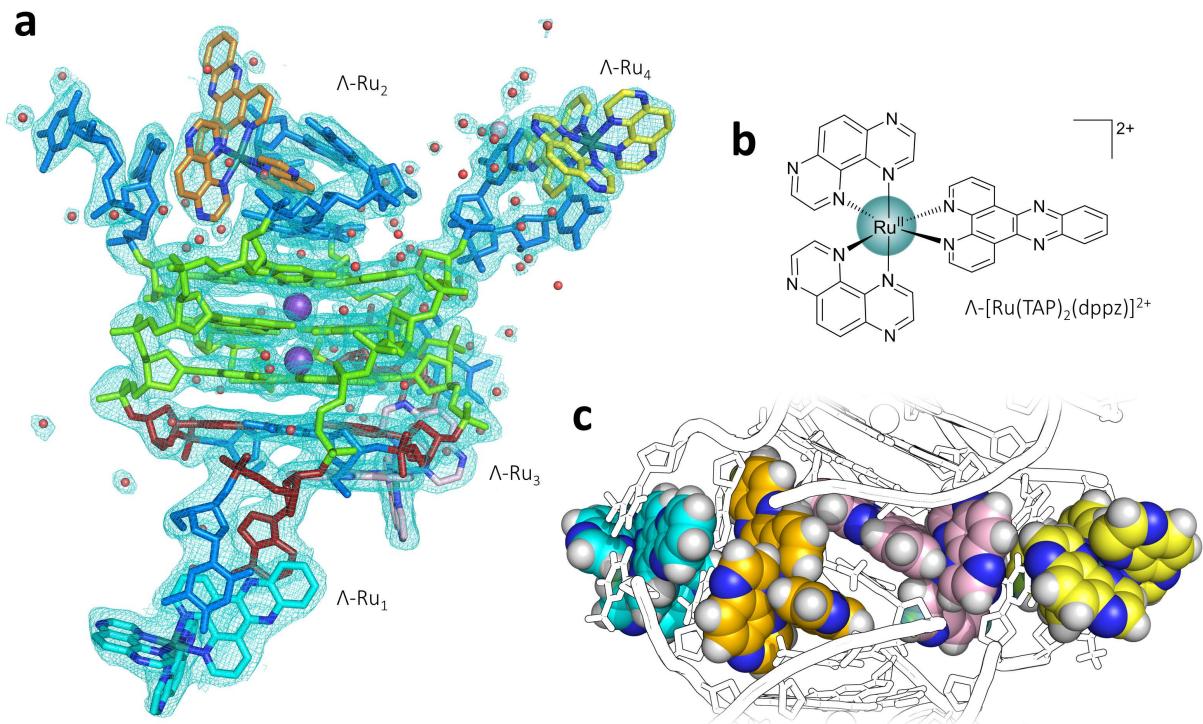
Figure 4.1 – *(a) Crystallographic model and electron density map (contoured at 1σ) of the asymmetric unit containing the G-quadruplex unit with all interacting ruthenium complexes and ordered waters; (b) chemical structure of the crystallised ligand Λ-[Ru(TAP)$_2$(dppz)]$^{2+}$; (c) rows of ruthenium complexes, all with distinct binding modes, observed to interact with two symmetrically related units. DNA bases are coloured as per convention with adenine in red, guanine in green, and thymine in blue. The lambda ruthenium complexes are coloured in cyan/orange/pink/yellow, differing only to portray disparate chemical environments of the complexes. Potassium ions are shown as purple spheres, and water as red spheres.*

Despite being grown in the presence of a racemic mixture, the crystallisation process between [Ru(TAP)$_2$(dppz)]$^{2+}$ and d(TAGGGTT) is, as seen in previous trials, completely enantioselective, containing only the lambda enantiomer bound to the oligomer. In addition, the model contains 75 ordered waters, two potassium ions, and a sodium ion. A 1:1 stoichiometric ratio of ruthenium to DNA chain is observed, equivalent of four per tetrameric quadruplex unit. Each complex is in a chemically distinct environment, and none are directly interacting with the central G-quadruplex unit. In the presence of K$^+$, the oligonucleotide d(TAGGGTT) inherently forms a parallel conformation, and the quadruplex core in the structure similarly retains this topology. What is less known however is the structure of the overhanging A/T regions. The model presented shows three distinct structural features of the A/T regions; 1. T·A/T·A quartet formed from only two strands with base pairing between adjacent residues in the base step; 2. parallel B-DNA duplex region formed entirely of T·T mismatch pairs; and 3. anti-parallel B-DNA duplex regions formed between neighbouring asymmetric

units. Each complex is binding in a unique binding mode to each of these features and although they do not interact directly together they form rows throughout the packing between the DNA assemblies.
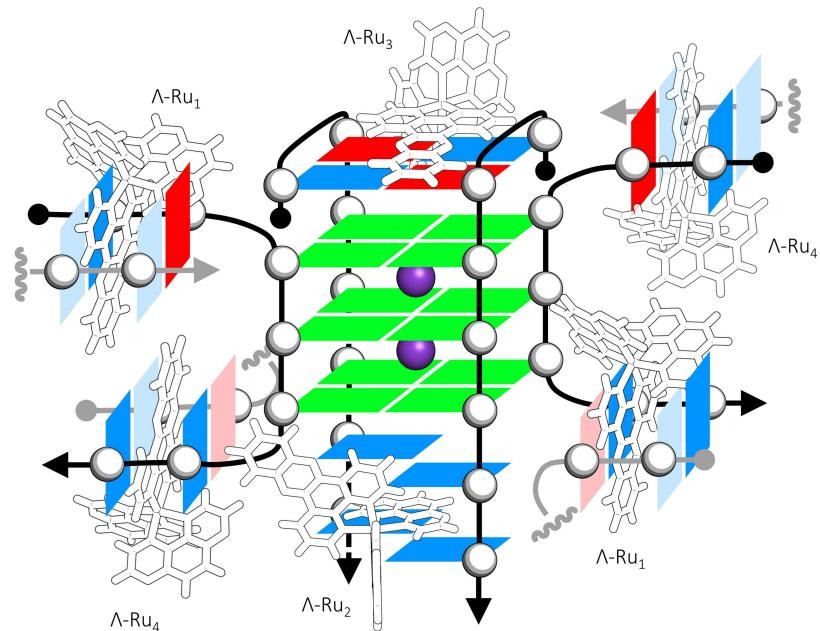
Figure 4.2 – *Schematic to illustrate the multifaceted nature of the structure and to highlight the range of binding modes seen within one biological unit of DNA in the model, Nucleobases are coloured as per standard, with adenine in red, guanine in green, and thymine in blue. Greyed out strands/nucleobases depict pairing interactions with neighbouring asymmetric units. Potassium ions are shown as purple spheres.*

4.3.1.2 DNA structure and conformation

Conformational analysis of the nucleic acid components of the model reveal that despite the odd arrangement of strands, the multitude of non-canonical structural motifs (including a harsh kink), and the number of binding locations, the DNA residues primarily adopt a canonical B-DNA form (mostly B_I). This is similarly true for the majority of loop regions observed in structures of telomeric G-quadruplexes (see table A4.1 for comparison). Phase angle determination and sugar pucker assignments are summarised in table 4.2.

Residue	P	Pucker	Conformer	Residue	P	Pucker	Conformer
A - dT$_1$	184.1	C3'-exo	B (parallel)	C - dT$_1$	169.8	C2'-endo	B (parallel)
A - dA$_2$	164.7	C2'-endo	B (parallel)	C - dA$_2$	140.4	C1'-exo	B-A transition
A - dG$_3$	169.5	C2'-endo	B	C - dG$_3$	23.4	C3'-endo	A-B transition
A - dG$_4$	133.5	C1'-exo	B	C - dG$_4$	149	C2'-endo	B
A - dG$_5$	151.3	C2'-endo	B-A transition	C - dG$_5$	150.4	C2'-endo	B-A transition
A - dT$_6$	164.3	C2'-endo	B	C - dT6	146.7	C2'-endo	B (parallel)
A - dT$_7$	78.7	O1'-endo	B	C - dT$_7$	100.4	O1'-endo	B
B - dT$_1$	199.7	C3'-exo	B	D - dT$_1$	180.7	C3'-exo	B (parallel)
B - dA$_2$	174.5	C2'-endo	B	D - dA$_2$	182.6	C3'-exo	B
B - dG$_3$	173.1	C2'-endo	B	D - dG$_3$	141	C1'-exo	B
B - dG$_4$	162.2	C2'-endo	B	D - dG$_4$	134.2	C1'-exo	B
B - dG$_5$	157.7	C2'-endo	B-A transition	D - dG$_5$	140.6	C1'-exo	B
B - dT$_6$	75.4	O1'-endo	A-B transition	D - dT$_6$	159.9	C2'-endo	B (parallel)
B - dT$_7$	171.8	C2'-endo	B	D - dT$_7$	154.6	C2'-endo	B

The G-quadruplex core of the structure is arranged in a parallel architecture with all guanosines in the stack adopting an *anti* conformation. The strands are arranged such that each groove is on average 16.2 Å wide (M) (figure 4.3); this is almost exactly the same groove width as the 16.1 Å observed in the wtTel22 parallel G-quadruplex structure (PDB: 1KF1).[10]
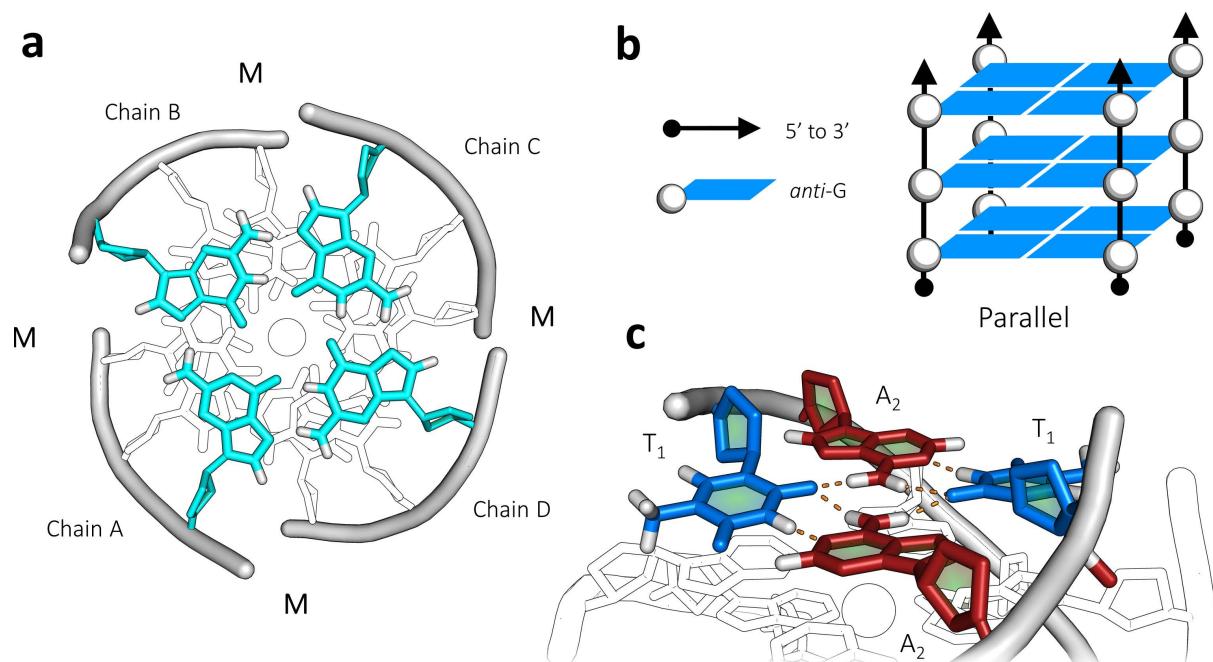
Figure 4.3 – (a) the G-quadruplex core as viewed from the 5' side highlighting the anti conformation of guanines, the pseudo symmetric nature of the parallel stranded motif, and the consistent medium (M) sized grooves; (b) schematic of the parallel polarities of the core tetrad stack. (c) directly above the G-tetrad on the 5' side lies a TA/TA quartet formed by two T·A pairs that originate from only two strands such that adjacent bonded nucleobases are paired equatorially.

On the 5' side of the G-quartet stack lies an additional quartet containing two T·A reverse Watson-Crick base pairs that are associated through additional H-bonding with the O2 of thymine and H62 of adenine to the neighbouring pair (figure 4.4c). Most interestingly, this T_1A_2/T_1A_2 quartet is derived from just two strands arranged in an antiparallel formation, meaning that adjacent T_1 and A_2 bases originating from the same strand are not interacting by π-overlap but by direct equatorial H-bonds (such that the NCCN torsional angle for the step is 357°)(see figure A4.2). Gauche (closed) γ angles are also observed at both the TA steps. Despite the torsional stress at these steps, base pair parameters are within expected ranges except for a little buckling and propeller twist on one of the pairs (buckle = 8.35°, propeller = -14.77°). Additional views of the motif are displayed in figure A4.2.
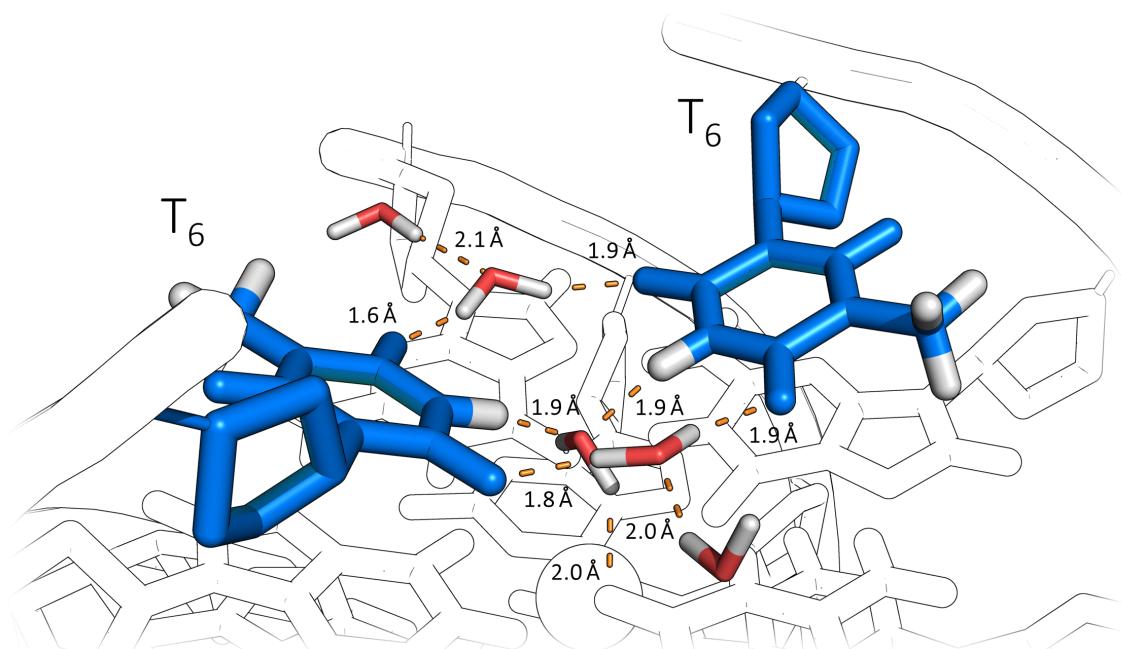
Figure 4.4 – *Water network that bridges two thymine residues, facilitating a head-on pairing mode between the two. Initially the central water was believed to be a disordered metal cation due to its proximity to the central ion channel; geometric analysis of the polar contacts around it proved otherwise and allowed for the placement of hydrogen despite the low resolution of data. Note that the hydrogen positions were not calculated but added manually in plausible positions.*

Prior to crystallisation the DNA was annealed in the presence of 100 mM KCl, however the final crystallisation conditions also contained 25 mM NaCl. As such, determination of the metal sites was done so with care, comparing the difference maps (F_o-F_c) of refined sites, the temperature factors, and using prior chemical knowledge and database statistics. Following an NDB search for all G-quadruplex structures the CheckMyMetal server was used to analyse each structures metal coordination spheres; models were hand curated and the binding environments were assigned by coordination geometry, number, and length.[11] The central two metal sites were determined to be potassium due to their octa-coordinate square antiprismatic geometry and an average donor-metal bond length of 2.81 Å (K--O and Na--O average distance in CSD is 2.8 and 2.4 Å respectively). A third

atom that follows the axis of the previous two ions was initially believed to be a partially occupied potassium. However, this particular atom seems to bridge two thymine residues, coordinating with the accepting N3 hydrogen atoms of each and contains distances/angles characteristic of H-bonding in water and was thus was deemed to be such. There is an elaborate ordering of water around this 5' tetrad and adjacent $T_6 \cdot T_6$ mismatch pair. The central water for example is found to be H-bonding with O6 on two guanines on the adjacent G-tetrad as well as bridging the two thymine via the oxygen atom (figure 4.4). Using this initial confidence in hydrogen placement, the hydrogens of a group of neighbouring water molecules inside the DNA pocket could also be assigned even without atomic resolution data (or neutron diffraction data).

4.3.1.3 binding modes
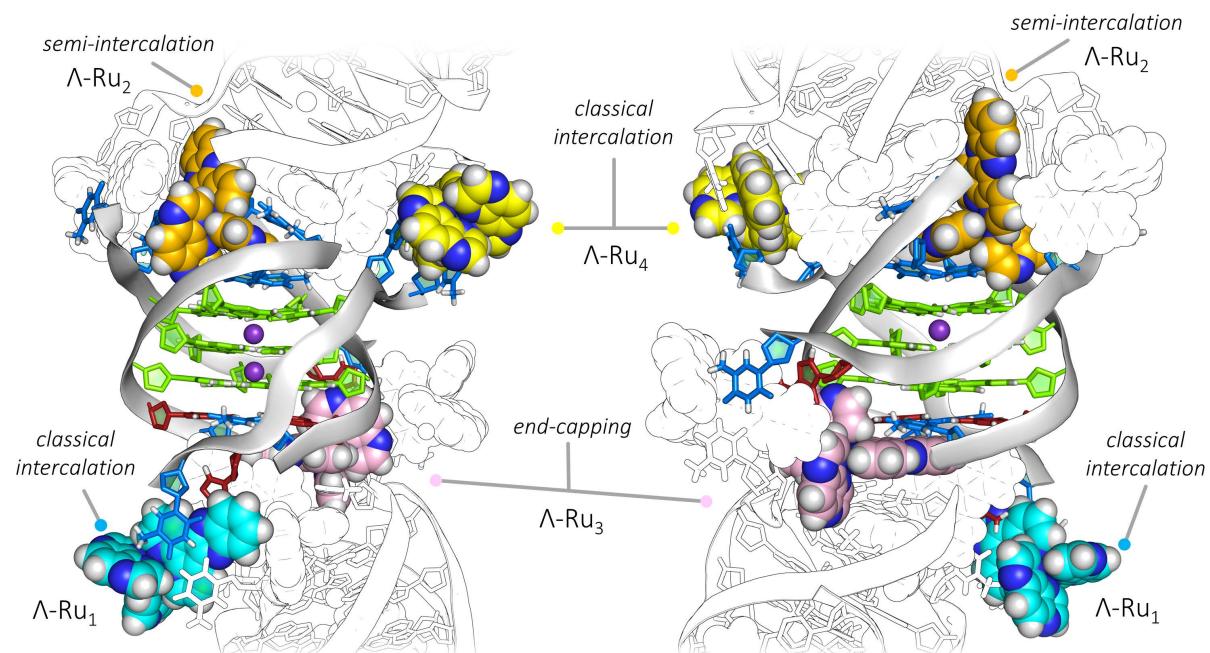
Figure 4.5 – *Crystallographic model illustrating the DNA structure with all interacting ruthenium complexes. Highlighted are the main binding modes exhibited by the lambda complexes in the model. Nucleobases are coloured as per standard, with adenine in red, guanine in green, and thymine in blue. The lambda ruthenium complexes are coloured in cyan/orange/pink/yellow. Potassium ions are shown as purple spheres; waters have been omitted for clarity. Black outline depicts model originating from neighbouring asymmetric units.*

Four separate crystallographically inequivalent lambda complexes are observed in the structure, each with distinct chemical binding modes. None are observed to interact with the G-quadruplex core, instead each complex is found intercalated in some manner to the thymine rich regions of the sequence. Such regions are reminiscent of the single stranded loop regions of natural and higher-order DNA, and as such the structure lends an interesting comparison to the binding modes of metal complexes to these loop regions. The bound species at least partially interact with

every thymine present, suggesting a definite binding preference over the parallel arranged G-quadruplex section. For clarity, each binding mode is described separately.

Classical intercalation (Ru$_1$ and Ru$_4$)
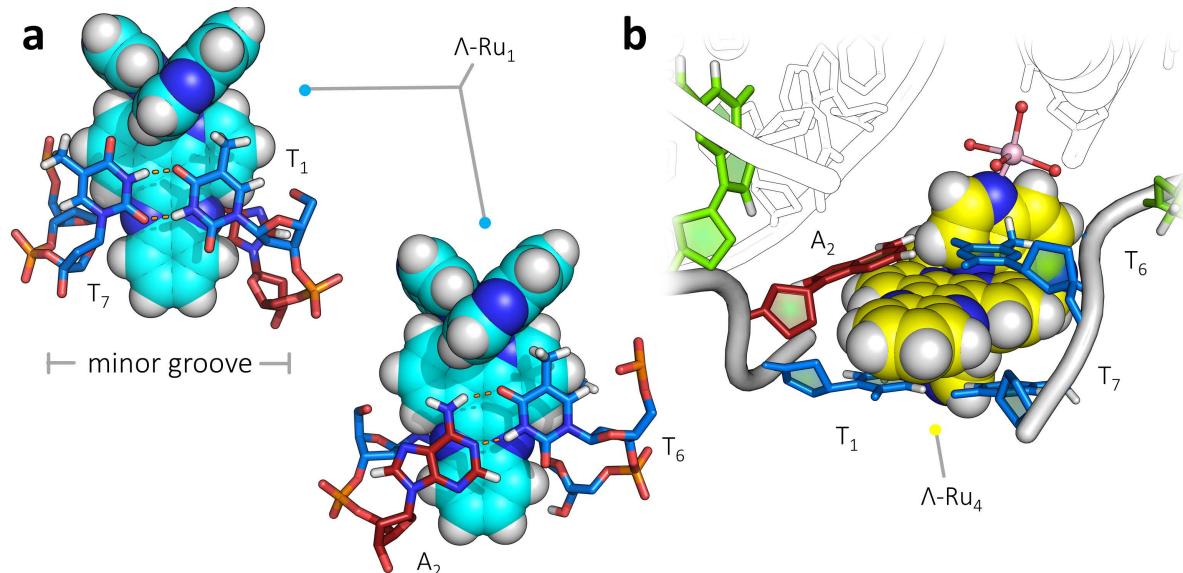
Figure 4.6 *– Crystallographic models illustrating the binding environment and π-stacking overlaps between the similarly bound Ru$_1$ and Ru$_4$. (a) Classical intercalation via the major groove is observed with Ru$_1$, allowing for deep intercalation; the complex overlaps almost completely with all bases in the intercalation cavity and does so perpendicular to the P-P vector. (b) model of Ru$_4$ is almost identical with Ru$_1$ except for the coordination of a Na$^+$ ion to a nitrogen on the TAP ligand. The illustration also shows how tight the cavity is around the intercalating ligand.*

Both of the complexes reversibly bind via classical intercalation where both have very similar binding environments, and both being found at the terminal ends of the overall assembly. The binding cavities consists of a Watson-Crick A·T base pair and T·T mismatch pair formed by T$_1$ and A$_2$ of one strand and T$_6$ and T$_7$ of another strand of opposing polarity. Perhaps most interestingly however is that complexes bind through the major groove. Local step analysis confirms that the terminal regions adopt sugar puckers and torsional conformations characteristic to B-DNA (see table 4.2); if treated as such the structure would be the first to exhibit such binding and could be used as a more general model for major groove binding in duplex DNA. Due to this groove preference no ribose contacts are observed, and the depth of the interaction is controlled by the substituents on the nucleobases, allowing a much deeper intercalation than is observed in common minor groove intercalations. The complex is seen to intercalate in a symmetric manner with the two-fold rotational axis of the dppz effectively perpendicular to the P-P vector (offset by 83°). The two similar sites are distinguished by the coordination of a Na$^+$ ion on one TAP ligand of Ru$_4$ exclusively. The hydrated ion is directly coordinated to one of the TAP ligands of Ru$_4$ but also Ru$_3$ via a water bridge. In both cases the intercalation sites exhibit tight 'closed' cavities with gauche γ angles (43 and 56° for T$_1$/A$_2$ and T$_6$/T$_7$

respectively). The step exhibits an average 7.1 Å rise to accommodate the interaction and the $A_2 \cdot T_6$ pair is buckled by 22.4° despite the $T_6 \cdot T_7$ pair buckling by only 10.9°. Between asymmetric units, the Ru_1 and Ru_4 binding cavities are end-stacked upon each other, generating a quasi-continuous stack that runs orthogonal to the helical axis direction (corresponding as the *b* axial direction in the crystal and in figure A4.3).

Semi-intercalation (Ru₂)
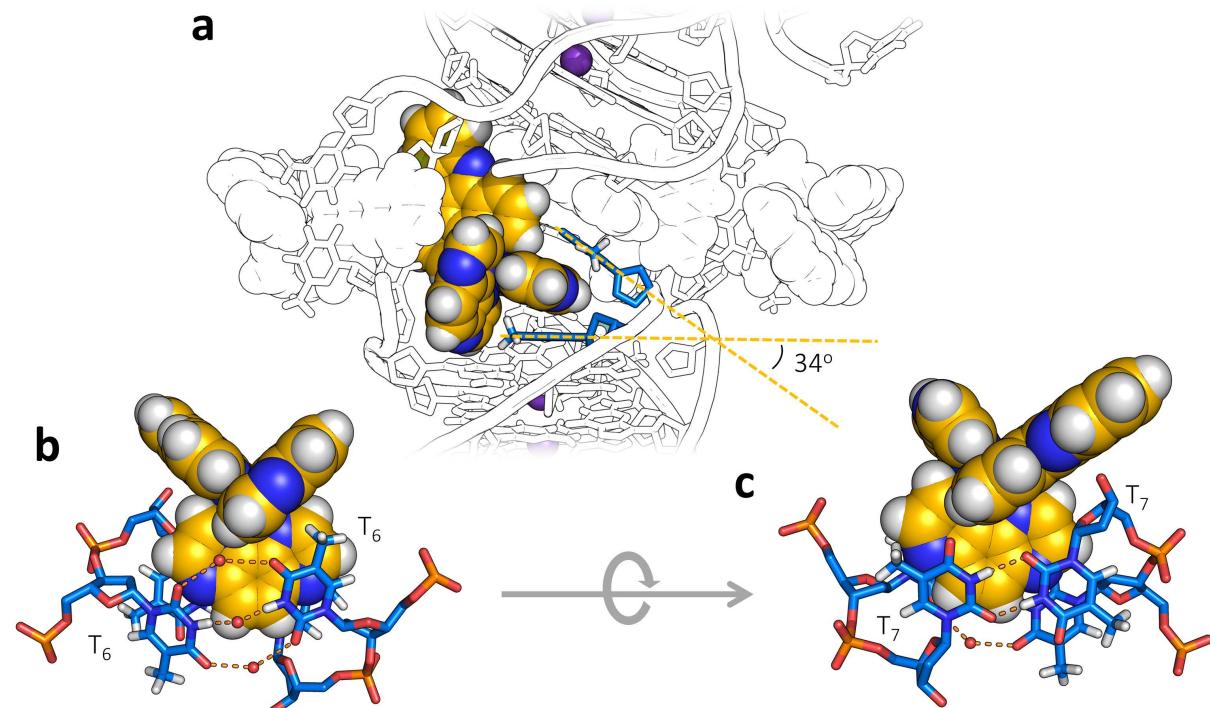
Figure 4.7 *– (a) Semi-intercalation of Ru_2 into T_6T_7/T_6T_7 kinks the backbone locally by 34°, inducing a 25° kink between neighbouring units. Two views of the parallel stranded binding site; (b) from the water mediated symmetric $T_6 \cdot T_6$ mismatch side, and (c) from the $T_7 \cdot T_7$ wobble mismatch pair side. Waters are shown as red spheres.*

A semi-intercalation site is observed between a TAP ancillary ligand of Ru_2 and a T·T mismatch site formed at the terminal base steps of two strands of the same polarity (figure 4.7). The site is composed of two mismatched thymine pairs; one is a wobble reverse pair formed by $T_7 \cdot T_7$ of the two parallel strands, and the other is a 'head-on' pair interaction between $T_6 \cdot T_6$ that is mediated by H-bonded water bridges across all three donor/acceptor atoms on the face of each thymine. Initially the central water was proposed to be a partial occupancy potassium due to its vicinity to the terminal tetrad and ion column however the local bond geometries are characteristic of a water network (see figure 4.4 for a view of the water framework and proposed hydrogen positions). The π-stacking interaction leans in favour of increasing π-overlap with the $T_7 \cdot T_7$ pair. The water bridges naturally increase the P-P distance at the step (up to 16.7 Å) and subsequently stretch the $T_6 \cdot T_6$ pair by 3.2 Å

whilst the head-on approach leads to a shear that is 2.1 Å larger than a canonical WC pair.[12] The partial intercalation kinks the local backbone conformation generating a roll angle of 34° at the $T_6T_7:T_6T_7$ base step; this leads to an overall 28° kink angle between the helical axes of neighbouring units that forms the lattice morphology seen in the crystal (see figure A4.3). Unlike most past examples, this mode is not stabilised by an adjacent large cation (such as Ba^{2+}) and this is the first example of such an interaction with base steps of exclusively pyrimidine content which therefore do not contain a free face with unoccupied donor atoms free for coordination. Unexpectedly, the dppz ligand is free from any major π-stacking interactions so is held entirely by the observed kinking interaction; despite this the complex is very well ordered in all directions.

End-capping (Ru₃)
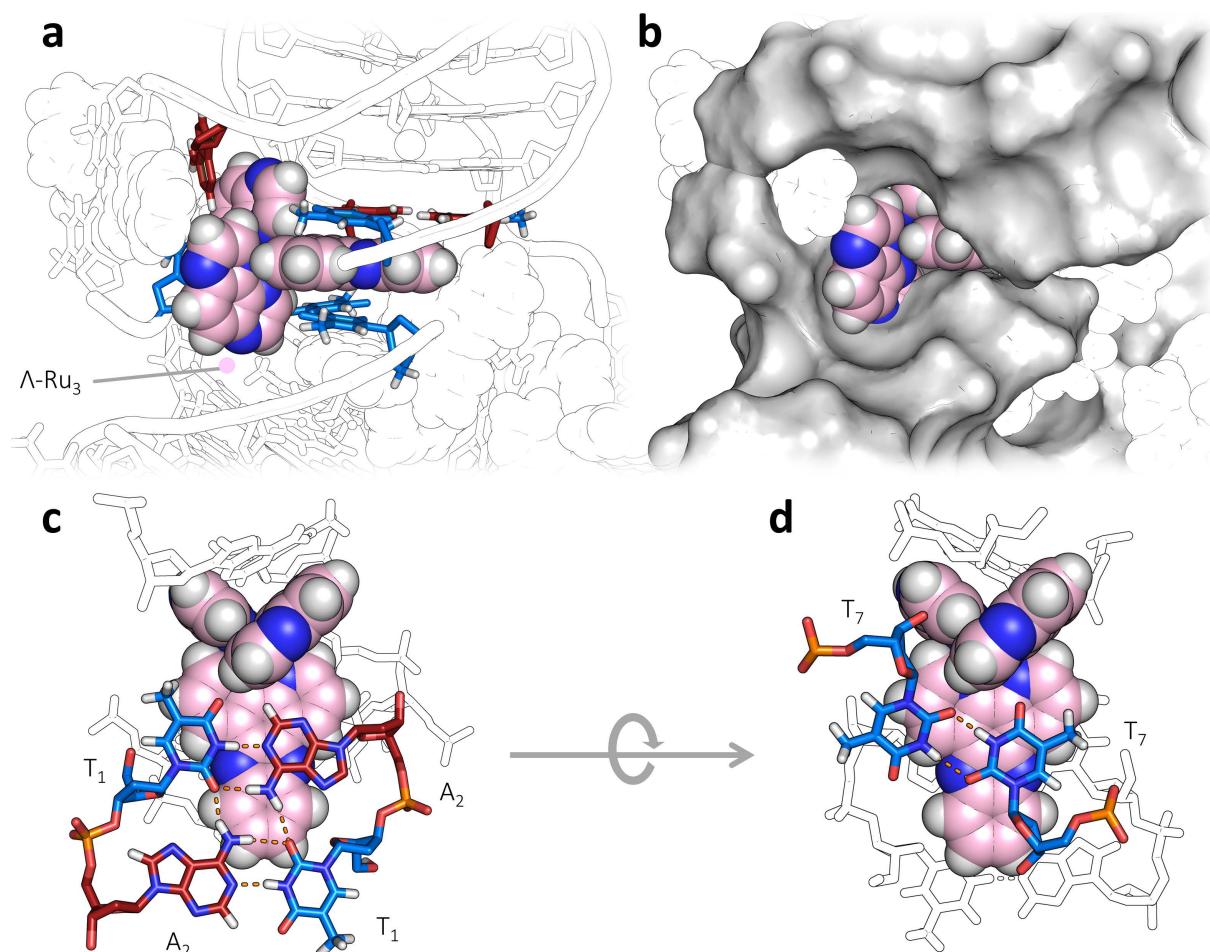
Figure 4.8 *– (a) End-capping of Ru₃ between two adjacent asymmetric units. (b) Binding environment with the calculated Van der Waals surface of the DNA highlight the encapsulation of the complex. Two views of the binding site; (c) from the $T_1A_2:T_1A_2$ tetrad side, and (d) from the $T_7 \cdot T_7$ wobble mismatch pair side. Bases π-stacking with the complex in (a) are highlighted.*

The final mode has the Ru₃ complex almost entirely encapsulated by T·A and T·T base pairs originating from two separate adjacent asymmetric units. As such, the complex acts as a stabiliser

throughout the crystal lattice, exhibiting enantiospecificity in its interactions with DNA which subsequently determine the packing geometry. On one face of the dppz the complex is directly stacking with a T·A/T·A quartet formed by two planar T_1/A_2 terminal steps on one asymmetric unit doing so in a symmetric manner normal to the hydrogen bonds of a $T_1 \cdot A_2$ pair; on the other face it is stacked upon the same $T_7 \cdot T_7$ reverse wobble pair that is kinked by Ru_2 of a different asymmetric unit. In addition, one of the TAP ligands is stacked on an $A_2 \cdot T_6$ base pair, effectively encapsulating the complex in DNA (figure 4.8b). The local A_2/T_6 step is deformed at this location to accommodate the complex; exhibiting a -78.7° tilt, 27.1° of roll, and -31.3° of twist (see table A4.1 for full conformational analysis). Hypothetically, if the delta isomer were to be substituted at this position, this enantiospecific interaction would be unfeasible and would kink the DNA at this step more aggressively therefore breaking the orthogonal interaction with the neighbouring asymmetric unit. The depth of the pseudo-intercalation is regulated by close contacts with the O4 atom on T_1 and H2 on the A_2 of the planar Watson-Crick TA/TA quartet, and by close contacts with the O2 on one T_7 and the O4 of the other T_7 on the parallel stranded mismatch pair.

4.3.2 SR-circular dichroism

Thermal denaturation studies of the truncated quadruplex sequence in the absence/presence of the enantiomers was performed on beamline B23 at Diamond Light Source. Ltd (full melting analysis is presented in figures 4.9 and A4.4). Initially the sequence was investigated in the presence of the nitrile derivative Λ/Δ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ discussed in the previous chapters due to the complex's ability to induce an anti-parallel topology with d(TAGGGTTA) at higher concentrations (such as in the solid state). In all cases melting profiles were observed at 260 nm, however when the DNA is in the presence of a higher stoichiometry of lambda isomer an additional melting profile is observed at 295 nm. Natively, the sequence has a maximum at 260 nm and a minimum at 240 nm, indicative of a parallel conformation, melting at 54 °C (figure 4.8a).
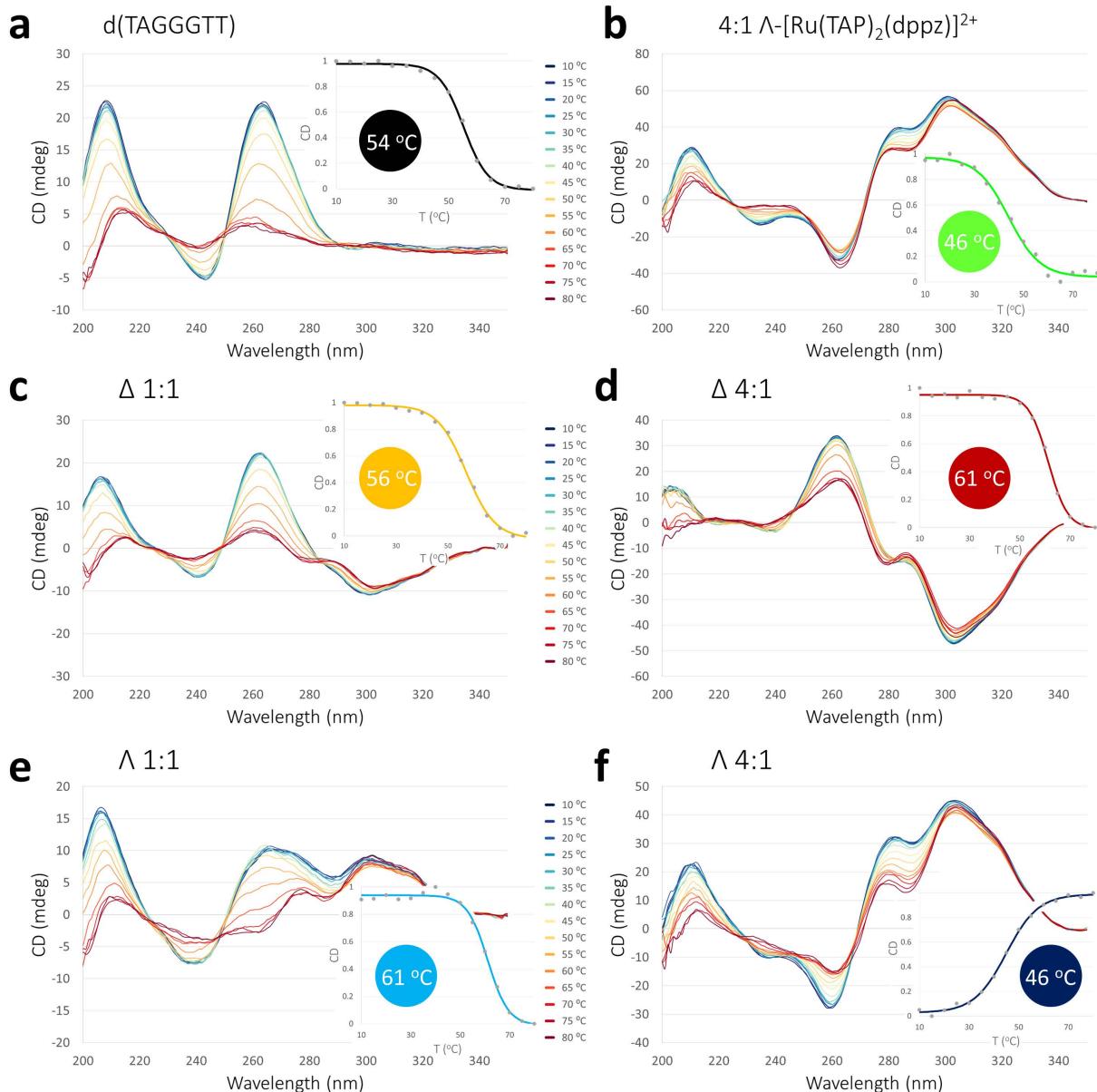
Figure 4.9 – *SRCD melting profiles with subsequent single point melting curves (inset) of: (a) native K⁺ folded d(TAGGGTT); d(TAGGGTT) with stoichiometric ratios of (b) 1:4 Λ-[Ru(TAP)₂(dppz)]²⁺; (c) 1:1 and (d) 1:4 Δ-[Ru(TAP)₂(11-CN-dppz)]²⁺; d(TAGGGTT) with stoichiometric ratios of (e) 1:1 and (f) 1:4 Λ-[Ru(TAP)₂(11-CN-dppz)]²⁺. CD spectra were measured from samples containing 800 μM d(TAGGGTT) with either 1:1 or 4:1 stoichiometric amounts of complex to biological unit, yielding a final concentration of 200 or 800 μM of respective complex. All samples contained 20 mM K-cacodylate pH 7 and 30 mM KF. Measurements were collected using cells with a 100 μm path length on beamline B23 at Diamond Light Source Ltd.*

Addition of either enantiomer of the cyano derivative at a 1:1 ratio stabilises the G-quadruplex by +2 and +7 °C respectively (figures 4.9c and 4.9e). Similarly, at a 4:1 ratio of delta a +7 °C stability increase is observed (figure 4.9d); however, at a 4:1 ratio of lambda, spectral differences occur. At 260 nm CD signal is seen to increase as melting occurs, and conversely the signal at 290 nm is seen to decrease upon heating. These spectral characteristics mirror those seen in the melting analysis of d(TAGGGTTA) (from Chapter 3) at the same stoichiometry and strongly suggest that the DNA has adopted an anti-parallel topology. Interestingly however, when the same

experiment is conducted with the parent species (i.e. no nitrile substituent), at the same stoichiometry as is observed in the crystal structure of the same interaction, no melting endotherm is observed in the 260 nm area. Such a loss of signal at that position could suggest the appearance of a mixed/hybridisation of strand polarities, however, a more in-depth structural probe would be required to definitively describe the topology.

4.4 Discussion

The sequence d(TAGGGTT) used herein was designed to provide a more reproducible crystallisation system when compared to d(TAGGGTTA) used in the previous chapter. Both are natively parallel stranded in architecture and only differ by way of a single terminal adenosine. Structural data (PDB: 5LS8) and CD measurements show that when four complexes of Λ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ are bound to the quadruplex formed by d(TAGGGTTA), the assemblies form an antiparallel arrangement; in addition, two of the A$_8$ adenosines are not ordered in the lattice with the other two stacking on the TAP ligands. As such removal of these bases was predicted to be of little consequence, and presuming the same mode of binding was occurring, it would crystallise in an analogous fashion. The structure presented here contains the same stoichiometric ratio of lambda complexes to assembly as the previous chapter however in this example the quadruplex component is ordered in a parallel topology and the complexes in the presented structure do not interact with the G-quadruplex core. Omitting the truncation of the sequence, the more apparent reasoning for the topological preference in the structures is the presence or absence of the nitrile substituent on the complex, compounding the potential importance of the polar contacts with the nitrile group in forcing the antiparallel fold in the 5LS8 from the previous chapter. Indeed, crystallisation trials were unsuccessful when either the complex or DNA were switched with each other, at least suggesting substantial alterations in the overall packing.

As discussed in chapter one, the definite groove in which the intercalation of ruthenium complexes bind has always been under deep contention. Spectroscopic analyses of [Ru(phen)$_2$(dppz)]$^{2+}$ with B-DNA for example have separately concluded intercalation into either the major or minor groove as the predominant mode.[13,14] However, of the structural studies to date, all exhibit binding of either enantiomer of [Ru(L)$_2$(dppz)]$^{2+}$ through the minor groove side, whether that be to canonical B-DNA, or mismatched sites.[15,16] The structure here presents for the first time an intercalative mode of such complexes that occurs via the major groove, doing so into a terminal T$_1$A$_2$/T$_6$T$_7$ step which is the only part of the structure that exhibits antiparallel polarity of strands. A previously reported structure containing the intercalation of [Ru(phen)$_2$(dppz)]$^{2+}$ into d(CCGGTACCGG) also shows binding to a TA rich site.[15] In that structure the complex binds, via the minor groove, into a

central TA/TA step that is of similar conformation to that seen here, allowing for an appropriate comparison between binding via each groove. The most significant difference observed is the absence of ancillary ligand interactions when bound through the major groove, unlike the favourable ribose interactions seen when bound through the minor groove (figure 4.10). Instead the depth and angle of the major groove intercalation is moderated directly by steric clashes with the exocyclic substituents on the nucleobases, resulting in a shallower intercalation in comparison. These ancillary interactions are enantiospecific to the lambda enantiomer, however, the delta isomer could plausibly intercalate via the major groove in an analogous way to that seen here due to this lack of ancillary ligand interaction/potential steric hindrance. In addition, the P-P separation at the step is reduced to 16.4 Å from the 18.2 Å standard seen in B-DNA, but this is more than the 15.7 Å P-P distance that is induced following intercalation via the minor groove.
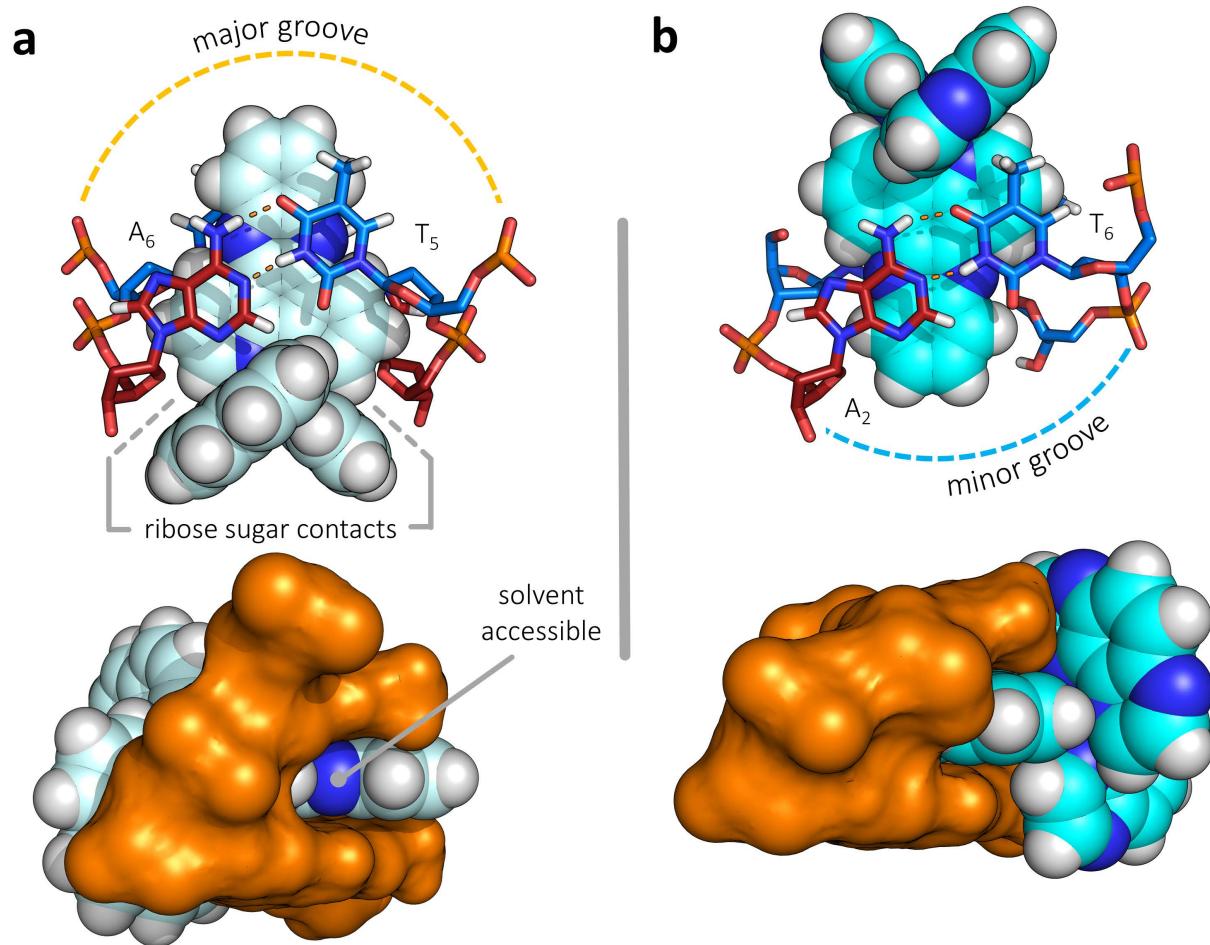
Figure 4.10 – (a) Symmetrical intercalation of [Ru(phen)$_2$(dppz)]$^{2+}$ into the TA/TA step of d(CCGGTACCGG) through the minor groove (PDB: 3U38). (b) Major groove binding in the presented structure of [Ru(TAP)$_2$(dppz)]$^{2+}$ into a TA/TT step (PDB: 6RNL). Calculated Van der Waals surfaces are shown in orange and highlight how binding into the major groove completely protects the pyrazine nitrogens from solvent water. Adjacent bases, waters, and ions have been omitted for clarity.

Assuming the phenanthroline derivative would bind the major groove in the same manner as the TAP variant here, both modes would display very different luminescence properties. When bound into the minor groove in a symmetrical manner, the pyrazine nitrogens on both side of the dppz are accessible to solvent water; conversely the shallower intercalation through the major groove means that the dppz is much better encapsulated and thus protected from the non-radiative relaxation pathways that occur through H-bonding to the pyrazine nitrogens (figure 4.10b). As such it would be expected that symmetrical major groove binding would lead to higher quantum yields and longer excited state lifetimes than previously observed binding modes (sym_{maj}>$canted_{min}$>sym_{min}).

The binding environment of Ru_2 is possibly the most unexpected of the binding sites observed. The overall kinking generated by the semi-intercalation of one of the TAP ligands at a T·T mismatch site is remarkably consistent with previous observations with duplex DNA, however this is the first example at a mismatch site, the first at a step of solely pyrimidine content, and the first kinking seen into parallel stranded DNA.[17] In dilute solutions of B-DNA, semi-intercalative binding modes exhibit relatively weaker thermodynamic constants in comparison to the deeper intercalative modes of dppz.[18,19] In tightly packed environments such as in the crystalline state, a combination of these weaker interactions may force binding modes not predicted from solution state studies. An example of this is this T_6·T_6 'head-on' mismatch pair at the site that has stretched to accommodate the semi-intercalative mode by a TAP ligand but is subsequently stabilised by water bridges across all donor/acceptor atoms. Stabilisation of mispaired bases by bridging water has been observed *in crystallo* before but has only been seen in already partially paired wobble pairs, and never with head-on thymine-thymine interactions.[20,21] The binding mode is a perfect example of the flexibility of nucleic acids when accommodating ligands and is an interaction that could not have been predicted from spectroscopic study alone.
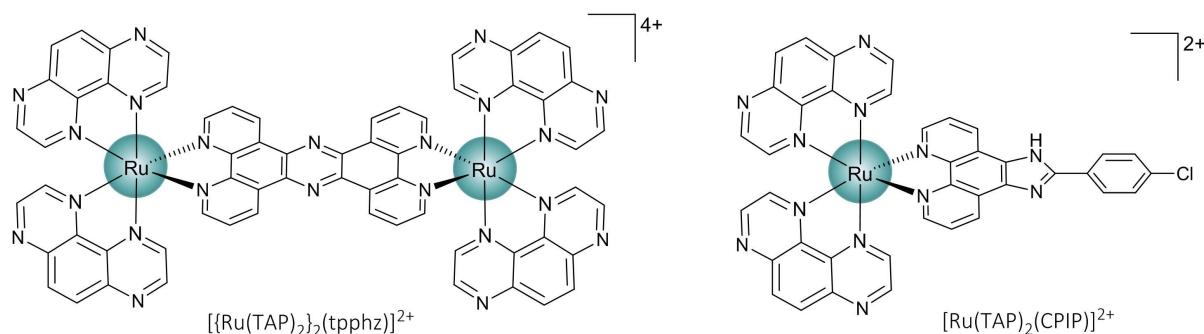

Figure 4.11 – *Chemical structures of the complexes discussed.*

A running theme observed in the structure is a preference for binding to T·T mismatches and/or with semi-intercalative modes in the TA regions. Such interactions may well be akin to the

binding of ruthenium polypyridyl complexes to the loop regions of single stranded G-quadruplexes or i-motifs.[22] Of course most G-quadruplex binders are predominantly planar and are designed to interact with the flat tetrad surface. Octahedral complexes on the other hand hold three-dimensional profiles with more inherent potential for structurally specific interactions, yet less is known about their binding modes.[23] As mentioned before, the only comparable structural coordinates available outside of ones presented in this thesis are those of the dinuclear ruthenium species' ΛΛ/ΔΔ-[{Ru(bpy)$_2$}$_2$(tpphz)]$^{2+}$ bound to an antiparallel basket G-quadruplex.[24] In those solution structures the ΛΛ-Ru cation threaded a diagonal loop in a predominantly end-stacking nature to the terminal tetrad although no distinct ancillary ligand interactions with the adjacent bases were apparent. However, previous investigations show that the bpy ancillary ligand is far less competent at intercalating and kinking than the relatively more expansive π-surfaces of the phen or TAP derivatives, so the absence of such interactions is not surprising.[25] A later study showed that the TAP analogue of the dinuclear complex exhibits efficient *in vivo* phototoxicity towards human melanoma cells (figure 4.11).[26] As with the mononuclear species these photosensitisation properties arise from a Ru^{II} → TAP ^{3}MLCT excited state that oxidises guanosine moieties predominantly. Binding titrations and lifetime measurements of the enantiomeric mixture in the presence of G-quadruplex DNA implicated multiple binding modes that could plausibly be obtained by interactions with the TA regions like the ones presented here. A separate investigation studied the excited state properties of a small range of derivative complexes based on the CPIP ligand.[27] They found that [Ru(TAP)$_2$(CPIP)]$^{2+}$ exhibits sub-micromolar affinity to telomeric G-quadruplexes and could elicit near 100 % mortality in U2OS osteosarcoma cells following irradiation (figure 4.11). Molecular docking calculations of the complex in the presence of the parallel folded telomeric quadruplex wtTel22 yielded two distinct major modes; one with the complex expectedly end-capping the tetrad, and one with the complex threading a TTA propeller loop. Interestingly in the docked loop mode the lambda isomers ancillary TAP ligands are in direct surface contact with the loop residues, effectively dictating the binding mode geometry. Indeed, bio-layer interferometry (BLI) experiments supports the hypothesised importance of these loop-mediated binding modes, where removal of these loop regions gives dissociation constants ten times lower than in the presence of the looped counterpart. Despite the large role of loop regions on the topology, stability, and the processivity of G-quadruplexes, very little is known about how metal complexes bind/may bind to these regions.[28–30]

4.5 Summary

The photooxidising complex Λ-[Ru(TAP)$_2$(dppz)]$^{2+}$ was successfully crystallised with the tetramolecular G-quadruplex d(TAGGGTT). The sequence is a truncated version of d(TAGGGTTA) that was shown to form an anti-parallel quadruplex in the presence of 4:1 Λ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$, with each complex binding adjacent to the terminal tetrads in a manner of pseudo-intercalation. The structure presented here unexpectedly crystallised in an antonymous way, with the quadruplex adopting a parallel formation with the lambda complexes comparatively far from the guanine-rich quadruplex core. A multitude of distinct binding modes of the complexes are observed, including semi-intercalation, mismatch intercalation, and major groove binding; all of which are completely novel and can be conceivably proposed as potential binding modes of such complexes to the often unstructured loop regions of telomeric G-quadruplexes or i-motifs.

4.6 References

1   Hussain, R. *et al.* CDApps: integrated software for experimental planning and data processing at beamline B23, Diamond Light Source. *J. Synchrotron Radiat.* **22**, 465–468 (2015).

2   Waterman, D. G. *et al.* The DIALS framework for integration software. *CCP4 Newslett. Protein Cryst.* **49**, 13–15 (2013).

3   Evans, P. Scaling and assessment of data quality. *Acta. Crystallogr. D.* **62**, 72–82 (2006).

4   Winter, G., Lobley, C. M. C. & Prince, S. M. Decision making in xia2. *Acta. Crystallogr. D.* **69**, 1260–1273 (2013).

5   Howell, P. L. & Smith, G. D. Identification of heavy-atom derivatives by normal probability methods. *J. Appl. Cryst.* **25**, 81–86 (1992).

6   McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).

7   Adams, P. D. *et al.* PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D.* **66**, 213–221 (2010).

8   Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta. Crystallogr. D.* **66**, 486–501 (2010).

9   Adams, P. D. *et al.* The Phenix software for automated determination of macromolecular structures. *Methods* **55**, 94–106 (2011).

10   Parkinson, G. N., Lee, M. P. H. & Neidle, S. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* **417**, 876–880 (2002).

11   Zheng, H. *et al.* CheckMyMetal: a macromolecular metal-binding validation tool. *Acta Crystallogr. Sect. D.* **73**, 223–233 (2017).

12   Olson, W. K. *et al.* A standard reference frame for the description of nucleic acid base-pair geometry. *J. Mol. Biol.* **313**, 229–237 (2001).

13   Jenkins, Y., Friedman, A. E., Turro, N. J. & Barton, J. K. Characterization of dipyridophenazine complexes of ruthenium(II): The light switch effect as a function of nucleic acid sequence and conformation. *Biochemistry* **31**, 10809–10816 (1992).

14   Tuite, E., Lincoln, P. & Nordén, B. Photophysical Evidence That Δ- and Λ-[Ru(phen)$_2$(dppz)]$^{2+}$ Intercalate DNA from the Minor Groove. *J. Am. Chem. Soc.* **119**, 239–240 (1997).

15   Niyazi, H. *et al.* Crystal structures of Λ-[Ru(phen)$_2$dppz]$^{2+}$ with oligonucleotides containing TA/TA and AT/AT steps show two intercalation modes. *Nat. Chem.* **4**, 621–628 (2012).

16   Song, H., Kaiser, J. . & Barton, J. K. Crystal structure of delta-[Ru(bpy)$_2$dppz]$^{2+}$ bound to mismatched DNA reveals side-by-side metalloinsertion and intercalation. *Nat. Chem.* **4**, 615-620 (2012).

17   Hall, J. P. *et al.* Structure determination of an intercalating ruthenium dipyridophenazine complex which kinks DNA by semiintercalation of a tetraazaphenanthrene ligand. *Proc. Natl. Acad. Sci.* **108**, 17610–17614 (2011).

18   Cardin, C. J., Kelly, J. M. & Quinn, S. J. Photochemically active DNA-intercalating ruthenium and related complexes – insights by combining crystallography and transient spectroscopy. *Chem. Sci.* **8**, 4705–4723 (2017).

19   Mihailovic, A. *et al.* Exploring the interaction of ruthenium(II) polypyridyl complexes with DNA using single-molecule techniques. *Langmuir* **22**, 4699–4709 (2006).

20   Cruse, W. B. *et al.* Structure of a mispaired RNA double helix at 1.6-Å resolution and implications for the prediction of RNA secondary structure. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 4160–4164 (1994).

21   Wu, P.-C. *et al.* Cooperative recognition of T:T mismatch by echinomycin causes structural distortions in DNA duplex. *Nucleic Acids Res.* **46**, 7396–7404 (2018).

22   Pages, B. J. *et al.* Stabilization of long-looped i-Motif DNA by polypyridyl ruthenium complexes. *Frontiers in Chemistry* **7**, 744 (2019).

23   Cao, Q. *et al.* G-quadruplex DNA targeted metal complexes acting as potential anticancer drugs. *Inorg. Chem. Front.* **4**, 10–32 (2017).

24   Wilson, T., Costa, P. J., Félix, V., Williamson, M. P. & Thomas, J. A. Structural studies on dinuclear ruthenium(II) complexes that bind diastereoselectively to an antiparallel folded human telomere sequence. *J. Med. Chem.* **56**, 8674–8683 (2013).

25   Hall, J. P. *et al.* Guanine can direct binding specificity of Ru–dipyridophenazine (dppz) complexes to DNA through steric effects. *Chem. Eur. J.* **23**, 4981–4985 (2017).

26   Archer, S. A. *et al.* A dinuclear ruthenium(II) phototherapeutic that targets duplex and quadruplex DNA . *Chem. Sci.* **2**, 3502–3513 (2019).

27   Weynand, J. *et al.* Towards the development of photo-reactive ruthenium(II) complexes targeting telomeric G-quadruplex DNA. *Chem. Eur. J.* **24**, 19216–19227 (2018).

28   Cheng, M. *et al.* Loop permutation affects the topology and stability of G-quadruplexes. *Nucleic Acids Res.* **46**, 9264–9275 (2018).

29   Lerner, L. K. & Sale, J. E. Replication of G-quadruplex DNA. *Genes (Basel).* **10**, 95 (2019).

30   Wang, L., Wang, Q.-M., Wang, Y.-R., Xi, X.-G. & Hou, X.-M. DNA-unwinding activity of Saccharomyces cerevisiae Pif1 is modulated by thermal stability, folding conformation, and loop lengths of G-quadruplex DNA. *J. Biol. Chem.* **293**, 18504–18513 (2018).

5 Biochemical Analysis of Ruthenium Polypyridyl Complexes

Contribution statement

Complex synthesis, characterisation, enantiomeric purification, and DFT/molecular modelling were performed by Kane McQuaid. Klenow fragment replication assays and circular dichroism were performed by Kane McQuaid under the supervision of Dr Shuntaro Takahashi. Immunofluorescence assays were performed by Kane McQuaid and Dr Hisae Tateishi-Karimata. FRET melting analysis was performed at Imperial College London by Kane McQuaid under the supervision of Timothy Kench and Prof. Ramon Vilar. Cell viability studies were performed by Dr Tamaki Endoh. The majority of the work presented here was performed at the FIBER institute at Konan University, Kobe (Japan) as a guest of Prof. Naoki Sugimoto, and made possible by a joint Japan Society for the Promotion of Science (JSPS) and British Council fellowship awarded to Kane McQuaid.

5.1 Introduction

The ability to control or alter the biochemical mechanisms which underpin the fundamental processes of the cell is a frontier aspiration in medical research, where control of these processes opens potential therapeutic pathways. Significant interest has been directed towards controlling the replication and transcription of DNA, especially so in relation to cancer therapies and control of gene expression. Non-canonical forms of DNA have been shown to alter these enzymatic processes and serve as targets to further explore this regulation. Sequence repeats of $d(T_2AG_3)$ found in the telomeres have been shown to form G-quadruplex structures, impeding the activity of telomerase and hindering the elongation of the protective telomeres.[1] *In vivo* and *in vitro* examinations both highlight the ability of the motif to perturb these replication processes, and preceding studies suggest that the thermal stability and topology of the quadruplex is a major contributing factor to this inhibition.[2,3]

In a wide array of cancers the overexpression of telomerase is noted.[4] This enzyme facilitates the polymerisation of the telomeric region and allows for the proper maintenance and protective function of the region to be conserved, protecting the cell from its natural senescence cycle effectively immortalising them. In order for polymerase to replicate DNA, a single strand must be unfolded and as such stabilisation of the secondary structure of the G-quadruplex dense region

represents a hopeful target for the suppression of telomerase action and thus the proliferative behaviour of the diseased cells.

In manner to further amplify the effect of non-canonical structure on the replication process, ligands that are able to selectively target and induce large stabilisation or damage to the tetraplex are desired. Perhaps the most structurally diverse of the DNA architectures, the unimolecular quadruplex can form many topologies; where upwards of 23 arrangements can occur.[5] Although designing selectivity can be an arduous task, we believe three main ideals lie true regarding the development of heteroleptic octahedral complexes with the intent of interacting with G-quadruplexes; that lambda complexes bind with higher efficacy, fundamental tetrad interactions are magnified by extended conjugated ligands; and, that ancillary ligands should be the focal point of enhancing topological specificity due to their interaction with the grooves and loop regions.

Within this chapter, the effect upon binding of a range of ruthenium complexes on the replication of a unimolecular human telomeric G-quadruplex was explored; the sequence can adopt a number of topologies dependent on ionic conditions and external stimuli. Enantiomeric disparities, ancillary ligand effects, substitutional patterns, chromophore extension, and the magnification of responses following MLCT excitation are all investigated as a means of evaluating this class of compounds as arresters of replication and as *in vivo* binders of G-quadruplexes.

5.2 Materials and Methodology

5.2.1 Materials and oligonucleotides

Unless otherwise stated, all materials and chemicals were sourced from Sigma Aldrich or Honeywell research chemicals. All solvents, unless stated in the experimental, were obtained at HPLC grade and used without further purification. Deoxyribonucleotide triphosphates (dNTPs) were purchased from Takara Bio, Japan. cNDI1 has been described previously and was sourced from authors.[6] All ligand manipulations were conducted in MilliQ water. Additional reagents were purchased from Wako Pure Chemicals and used without further purification. Template oligonucleotides and FAM-labelled primer DNAs were purchased from Japan Bio Service as double HPLC purified syntheses. All DNA was checked spectrophotometrically to ensure absence of residual protein.

5.2.2 Synthesis

The heteroleptic ruthenium complexes investigated in these studies are shown in chapter 7 and were prepared from commercially available starting materials. The homoleptic species'

rac-[Ru(phen)₃]·Cl₂ was synthesised in house using literature methods. Full experimental/synthetic details are presented in Chapter 7; chiral purification of [Ru(phen)₂(Aqphen)]·Cl₂ will be discussed further in this chapter.

### 5.2.3 Klenow fragment (exo⁻) preparation

Klenow fragment preparation was conducted by Dr Shuntaro Takahashi

The Klenow Fragment (KF) encoding gene was amplified from *E. coli* genomic DNA (JM109) *via* polymerase chain reaction (PCR). PCR was proceeded using PrimeSTAR DNA polymerase (Takara Bio) and primers (5'-GGGACCATATGGTGATTTCTTATGACAACTACG-3' and 5'-GGGAGAATTCTTAGT-GCGCCTGATCCCAG-3') sourced from Eurofin Genomics (Japan). Following digestion using NdeI and EcoRI, the cloned DNA fragments were cloned into pMal-p5x vector (New England Bio Labs). KF exo⁻ (D355A, E357A) was prepared by mutating the constructed plasmid using a QuikChange mutagenesis kit (Stratagene) followed by using the mutated vector to transform *E. coli* EG2523 (New England Bio Labs). The cells were then cultured in Luria-Bertani medium containing ampicillin and worked up to an A_{600} of roughly 0.5 followed by addition of isopropyl β-*D*-1-thiogalactopyranoside (IPTG). The cultured cells were harvested and lysed via sonication and the soluble fraction was loaded onto an amylose resin packed column (New England Bio Labs). Following treatment with Factor Xa protease to eliminate the MBP-tag, the KF exo- was purified over a Hitrap Heparin column followed by purification through a Hiload Superdex 200 (GE Healthcare). The purified KF exo- was dialysed against a dialysis buffer containing 50 mM Tris-HCl (pH 7.2), 1 mM EDTA, 1 mM DTT, 100 mM NaCl and 50 % glycerol. Concentration was determined spectrophotometrically using an extinction coefficient of 58,790 $M^{-1}cm^{-1}$ at 280 nm. The mutant enzyme was stored at -30 °C until use.

5.2.4 Klenow fragment replication assay

Template (5'-TTAGGGTTAGGGTTAGGGTTAGGGCGGCCCCTATAGTGAGTCGTATTAAGCCGAAGC-ACTAGTATCATCCC-3') and primer (5'-FAM-GGGATGATACTAGTGCTTCGGCTTAATACGACTCACTATA-GGG-3') oligonucleotides were annealed in the following buffer and in the presence, if quoted, of the necessary ligand: 10 mM Tris-HCl (pH 7.5), 8 mM MgCl₂, 1 μM FAM-labelled primer DNA, 1 μM template DNA, 0-20 μM ligand, 250 μM dNTPs and 1/100 μM KCl (as indicated). After annealing, the mixture was incubated at 37 °C and a 100 μM solution of KF exo- was added to the reaction mixture (final concentration of 1 μM) to initiate the enzymatic reaction. At given time intervals during the reaction, aliquots of the mixture were quenched using a stopping solution containing 10 mM EDTA and 80 wt% formamide. Reaction products were separated by denaturing polyacrylamide gel electrophoresis (PAGE) using gel containing 8 M urea in a TBE buffer at 70 °C for 1 hour at 200 V.

Alongside product lanes, a molecular weight marker DNA ladder (10-bp) and a bromophenol blue running aid were run in adjacent and terminal lanes. Gel images were captured using a Fujifilm Fluoreimager FLA-5100 utilising a laser excitation wavelength of 473 nm. Images were collected before and after staining with SYBR Gold (ThermoFisher Scientific) to highlight unlabelled products. Band intensities were analysed using ImageJ2 software package (NIH) by quantifying peak areas after baselining the necessary lanes. The reaction yield of full-length product (P) was quantified by calculating the ratio of intensity of the full-length product bands to the aggregate intensity of all bands. Reaction rate analysis was performed using Dynafit software package (Biokin) after evaluating a global fit of the reaction. This was achieved by applying a kinetic model to the following two-step sequential model;

$$P_0 \overset{k_s}{\to} P_s \overset{k_f}{\to} P_f$$
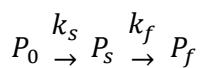
Where P_0 is the starting state of the reaction, P_s represents the state immediately after unwinding of the reaction stall (motif), P_f represents the state after the replication of the full-length product is completed; and K_s and K_f are the rate constants that define the rate of reaction between states.

5.2.5 Sample irradiation

Sample irradiation was performed on pre-annealed mixtures of ligand and DNA in the absence of any polymerase or deoxyribonucleotides and within an optically transparent tube. Samples were placed under a 4 x 4 cm beam of light produced by an Asahi Spectra (Japan) Xenon MAX-302 fitted with a 420 nm bandpass filter and running at 25 % of the 300 W capacity. Samples were turned every 5 mins during irradiation.

5.2.6 UV spectrophotometry

Analysis of the thermal stability of the systems (T_m) was conducted on a Cary 300 UV spectrophotometer (Agilent) with attached Peltier temperature control system running Cary WinUV. Solutions containing 5 µM $(T_2AG_3)_4$, 10 mM Tris-HCl (pH 7.5), 8 mM $MgCl_2$ and 20 µM of respective ligand were prepared for analysis. Samples were slowly annealed from 90 °C to 0 °C at a rate of -1.0 °C min^{-1}, following this the sample temperature was increased at a rate of 0.5 °C min^{-1} to 95 °C; where data were collected at 1 °C intervals. All measurements were conducted under a nitrogen atmosphere. Thermodynamic analysis was performed in the Kaleida Graph software package (Synergy Software) by sinusoidal curve fitting after normalisation of the UV melting curves.

5.2.7 Circular dichroism

Circular dichroism spectra were collected at 37 °C on a JASCO J-1500 CD Spectrophotometer running Spectromanager in quartz cells with path length of 1 cm. CD samples were measured at 5 µM $(T_2AG_3)_4$ DNA, 10 mM Tris-HCl (pH 7.5), 8 mM $MgCl_2$ and 20 µM of respective ligand. Samples were annealed from 95 °C to 20 °C at a rate of -1.0 °C min^{-1} and spectra were collected at 37 °C.

5.2.8 FRET melting assays

Double dyed FAM and TAMRA labelled oligonucleotides (Eurogentec) were used in the FRET melting experiments. DNA was diluted in Milli-Q water to a concentration of 0.4 µM in appropriate buffer and annealed slowly from 95°C. Stock solutions of the complexes were made up to 0, 0.4, 0.8, 2, 4, and 8 µM in the appropriate buffer solutions per DNA. Final samples were prepared by addition of 20 µL of DNA stock, with 20 µL of each respective ligand stock solution. Measurements were performed on an Agilent PCR Stratagene Mx3005P, using an excitation wavelength range of 450-495 nm and an emission detection range of 515-545 nm. Measurements were taken at 0.5 °C increments from 25 to 95°C, holding at each temperature for 30 secs before collection. Measurements were conducted in triplicate. Normalised results were fitted by solving against a five-parameter dose-responsive equation in GraphPad Prism. (LogXb=LogEC50 + (1/Slope)*Log((2^(1/S))-1), where X = dose/concentration, Y = fluorescence response, S = symmetry parameter). Buffer choice depended on DNA sequence/system used. For wtTel21 (10 mM Li cacodylate, 10 mM KCl, and 90 mM LiCl; or, 10 mM Li cacodylate, 10 mM NaCl, and 90 mM LiCl); for c-MYC and ds26 (10 mM Li cacodylate, 1 mM KCl, and 99 mM LiCl).

5.2.9 Immunofluorescence assays

MCF-7 cells (mild breast cancer) were cultivated as per literature standard in Pyrex dishes ($3x10^6$ cells/dish). Addition of 2 % PFA/PBS solution was added to immobilise the cells, followed by membrane permeabilisation treatment with a 0.5 % solution of Triton-X100 in PBS. BSA protein (2 % in PBS) was then added to the culture medium to block nonspecific/electrostatic ligand interactions. Ligands were added in solutions of PBS to a final concentration of 20 µM and incubated at room temperature for 15 mins. Control experiments added the same volume of PBS buffer in the absence of ligand. Immunofluorescence experiments were performed after incubation for 60 mins with BG4 antibody (Absolute Antibody, Oxford, U.K.); Alexa Fluor 546-conjugated goat antirabbit IgG (Molecular probes, OR, U.S.A.) antibody was used as a secondary antibody for the visualisation of the BG4 locations. DAPI (Vector Laboratories, CA, U.S.A.) was used to counterstain the cell nuclei. Confocal fluorescence microscope images were captured using a Zeiss LSM700 confocal microscope

(λ_{ex} = 545 ±25 nm and λ_{em} = 605 ±70 nm for Alexa Fluor 546 visualisation). Fluorescence foci were determined and analysed using the ZEN 2 bioimaging suite (Zeiss).

5.2.10 Cell viability studies

HeLa cells (immortal cervical cancer) were grown as per literature standard. One thousand cells were transferred into each well of a 96-well plate in growth medium (DMEM + 10 % FBS) and left for 24 hours. Following this, the complexes were added at differing concentrations (0, 1.25, 2.5, 5, 10, and 20 µM), and the cells were incubated at 37 °C in the presence of the ligands for 92 hours. After incubation, the number of cells were evaluated using a cell stain/counting kit (tetrazolium salt cell counting kit-8 (Dojindo)). The absorbance was measured following a 90 min incubation with the stain (10 µL), and then normalised against the signal of the cells in the absence of ligand. All measurements are shown as the average of triplicate samples.

5.2.11 Molecular modelling and docking studies

Initial models of each octahedral complex were built in Avogadro and minimised using a steepest descent method with UFF whilst constraining Ru-N geometry where relevant crystallographic input was accessible. Geometry optimisation at a semi-empirical level (PM6) was followed by calculation at the DFT/B3LYP (Becke's three parameter hybrid functional with the Yang-Parr correlation functional method) level using a 6-31G(d') basis set on the carbon, nitrogen, oxygen and hydrogen atoms and a LanL2DZ pseudo potential on the ruthenium atom. Frequency analysis confirmed the local conformational minima on the potential energy surface. All geometry optimisations and subsequent single point energy calculations of the ground state singlet species' were calculated using the Gaussian 09 package where frontier molecular orbital interactions were visualised and rendered in GaussView.

Molecular docking simulations were conducted entirely in the AutoDock 4.2 package. All DNA structural models were obtained from the protein data bank to which all water molecules and ligands (where necessary) were omitted. Gasteiger-Hückel charges were added to the optimised complexes before performing docking calculations and a virtual box grid of dimensions 110 x 100 x 100 Å was centred on the DNA assembly with a 0.5 Å resolution spacing. Free energy coefficients used; desolvation – 0.1159, H-bonding – 0.0974, and electrostatics – 0.1465. Calculated custom parameters to describe Ru were added to the parameter list where VdW radius (r_{ii}) = 2.96 Å, van der Waals well depth (ϵ_{ii}) = 0.056 kcal mol^{-1} and electric charge (q) = +2.0. Rigid molecular docking simulations were performed by application of the Lamarckian Genetic Algorithm (LGA) to search for conformational minima within the boundaries of precalculated atomic affinity grid maps and substrate

electrostatics/desolvation maps. Initial population consisted of 100 randomly placed individuals, a maximum number of 2.5×10^7 energy evaluations, a maximum number of 2.7×10^4 generations, a mutation rate of 0.02 and a crossover rate of 0.8. Docked conformations differing by less than 2 Å in positional RMSD were clustered and expressed as the most energetically favourable conformation dependent on its calculated free binding energy.

5.3 Results

5.3.1 Enantiomeric disparities upon binding of [Ru(phen)$_2$(dppz)]$^{2+}$

The replication of a templated sequence by the polymerase activity of the Klenow fragment exo- was studied in the presence of the light activated ruthenium polypyridyl complex, [Ru(phen)$_2$(dppz)]$^{2+}$, in a time resolved manner. The KF exo- mutant preserves the 5'→3' activity of the KF but lacks the undesirable exonuclease activity (3'→5' or 5'→3') that is used *in vivo* as a repair mechanism.[7] In the following experiments the KF attempts to polymerise a FAM labelled strand from template DNA. This template DNA strand contains an overhang region, downstream to a region complementary to a FAM labelled primer region, consisting of four repeat units of the human telomeric sequence (T$_2$AG$_3$) and a CGGC linker. This sequence is capable of forming a number of topologies dependent on the solution conditions such as cationic concentration and crowding conditions of the medium and in some cases the topology can be determined by the presence of particular ligands.
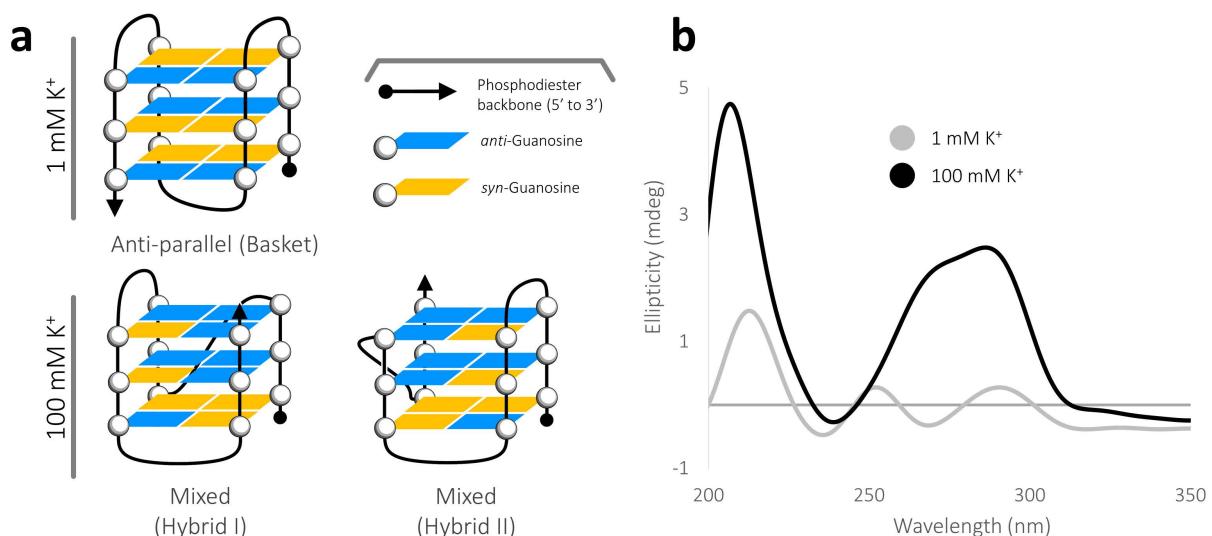
Figure 5.1 *– (a) models of commonly folded topologies formed by telomeric G-quadruplexes at different potassium concentrations; (b) CD spectra of (T$_2$AG$_3$)$_4$ in the presence of either 1 or 100 mM K$^+$ which shows how the sequence folds into either an anti-parallel or mixed hybrid topology respectively. CD was measured at 5 µM DNA in the presence of 10 mM Tris HCl pH 7.5, 8 mM MgCl$_2$, and either 1 or 100 mM KCl at 37 °C in quartz cells with a path length of 10 mm. Relatively low signal is observed for the sequence in the presence of 1 mM KCl due to the low thermodynamic stability of the fold at 37 °C.*

In the presence of low concentrations of K^+ the human telomeric sequence adopts an anti-parallel topology, whereas in higher concentrations (>95 mM) the motif assumes a mixed/hybrid topology. Circular dichroism (CD) spectra of the sequence in a solution containing 10 mM Tris HCl (pH 7.5) and 8 mM $MgCl_2$ and in the presence of either 1 or 100 mM KCl confirms this (see figure 5.1). The lower ionic strength solution presents a CD spectrum with a large positive signal difference around 295 nm, characteristic of an anti-parallel structure, whereas at higher ionic strengths a bimodal peak with maxima at 265 and 295 nm is observed; indicative of a mixed topology. Circular dichroism data from the sequence in the presence of the enantiomers of $[Ru(phen)_2(dppz)]^{2+}$ was collected (figure 5.2), and as expected was particularly difficult to deconvolute. At higher stoichiometric ratios we have shown that similar complexes can induce topological conversions of related G-quadruplex sequences, therefore to conserve parity between CD and replication results a high stoichiometry was maintained. As with previous investigations, subtracting the free complex signal from the bound signal does not simplify the situation since large induced signals arise from the extension of the chiral conjugation.
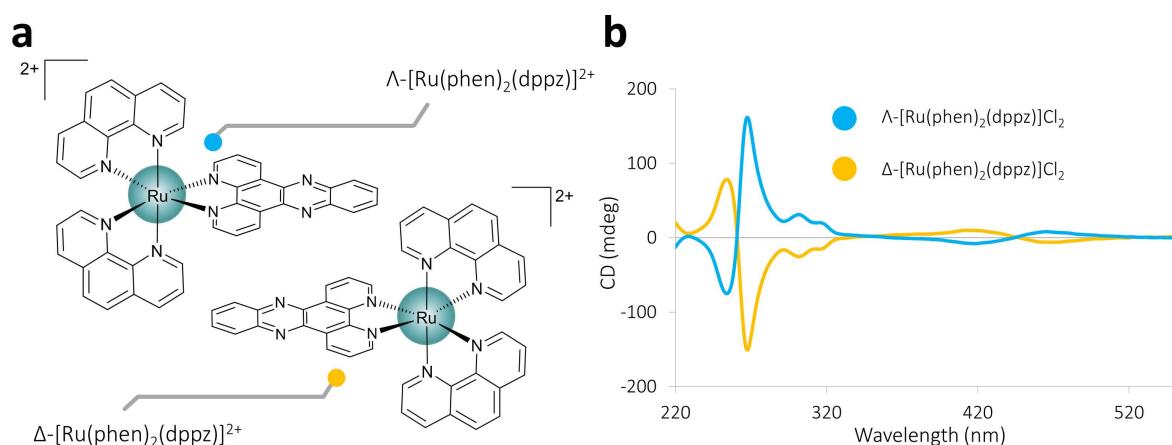
Figure 5.2 – (a) enantiomers of the common 'light switch' complex, $[Ru(phen)_2(dppz)]^{2+}$; (b) CD spectra of the resolved enantiomers. CD spectra measured at 15 μM complex in H_2O in quartz cuvettes with a path length of 10 mm.

In the same conditions, time dependent replication analysis of the same sequence was assessed using the Klenow exo- fragment and dNTPs, in the absence and presence of Λ/Δ-$[Ru(phen)_2(dppz)]\cdot Cl_2$. Experiments were conducted at high (100 mM) and low (1 mM) K^+ concentrations to observe the ligand induced stalling during the replication of mixed and anti-parallel topologies respectively. In the absence of ligand two product bands are detected upon analysis of the FAM-labelled products using denaturing polyacrylamide gel electrophoresis (PAGE; figure 5.3); one corresponding to the full product of replication (71 nts), and one corresponding to a stalled replication product containing the primer + four nucleotides (43 nts) (figure A5.1). This stalled product is a result of the folded quadruplex inhibiting the polymerase action of the Klenow fragment and can be associated with the unfolding kinetics and thermodynamic stability of the quadruplex. Taking aliquots

across the time course of the reaction allows for an assessment of the unfolding kinetics and hence the effects of ligand binding on replication.
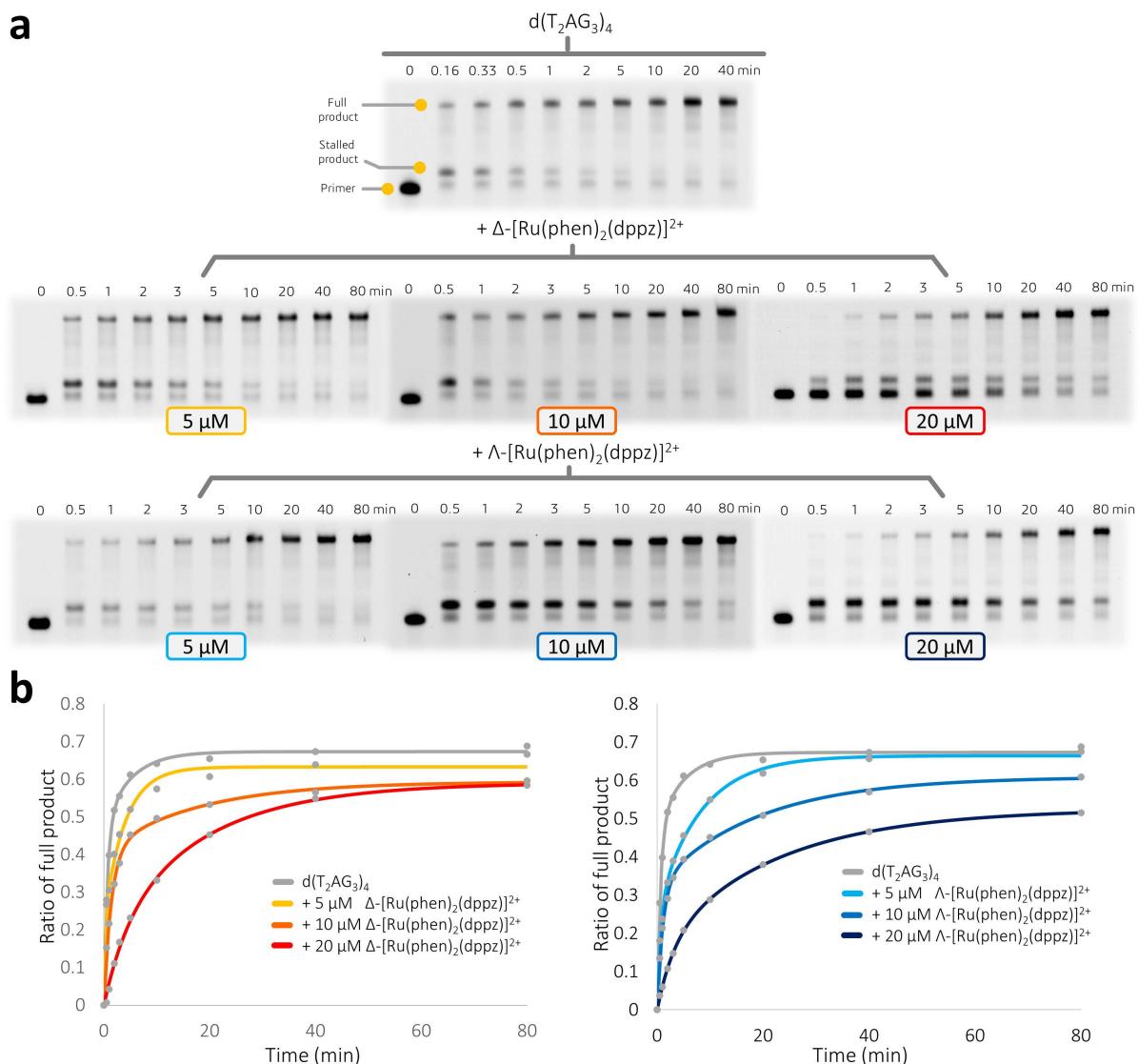
Figure 5.3 – (a) denaturing PAGE gel panels of the time dependent analysis of the KF polymerisation of 1 μM d(T₂AG₃)₄ in the presence of 100 mM KCl and in the absence or presence of the enantiomers of Δ/Λ-[Ru(phen)₂(dppz)]²⁺. Gels were visualised using the fluorescence of the FAM labelled primer strand. (b) Time dependent analysis of the reaction yields fitted to a two-step sequential model.

In the presence of 100 mM KCl and in the absence of any ligand, the folded quadruplex is unfolded and fully replicated within 5 mins with a reaction yield of 70 %; considerably slower than the rapid replication of ssDNA. In the presence of increasing concentrations of the enantiomers of [Ru(phen)₂(dppz)]·Cl₂ the same two product bands obtained in the absence of ligand are observed but the replication is slowed considerably. PAGE analysis shows that at all concentrations, major replication stalling is occurring at the site adjacent to the G-quadruplex folding (figure 5.3a). In addition, the lambda enantiomer stalls more effectively than the delta and, upon inspection of the

gels, is seen to stall primarily at the G-quadruplex whereas at high concentrations the delta stalls the reaction primarily at the B-DNA on the primer end before the G-quadruplex. Fitting of the kinetic data to a two-step rapid process quantifies the observed stalling. The two step sequential model fits a rate of replication before the stall position (k_s), and a rate following the unfolding of the replication stall (k_f). In the absence of any stabilising ligand, the rate constant (k_s) is 0.4 min^{-1}, whereas the secondary rate constant (k_f) is 9.64 min^{-1} (table 5.1). Clearly then, the dissolving of the replication stall is the rate limiting step of the replication process and this also holds true after the introduction of the ligands. In the presence of increasing concentrations of either enantiomer k_s values decrease, down to 0.122 min^{-1} for the delta enantiomer and a 0.0422 min^{-1} for the lambda at 20 µM. Interestingly the k_f rates also drop significantly in the ligand stabilised systems, dropping from 9.64 min^{-1} for the native DNA, down to 0.445 and 0.24 min^{-1} for delta and lambda respectively at 20 µM. This dampening of rate post-unfolding indicates a reduction in processivity of the single stranded template DNA downstream to the initial G-quadruplex fold and implicates a level of nonspecific binding of the ligands to this ssDNA.

Table 5.1 – *Kinetic analysis of the time-dependent replication assays.*

DNA[a] with ligand	Salt[b]	k_s (min^{-1})	k_p (min^{-1})
(T$_2$AG$_3$)$_4$		0.32±0.02	29±1.6
(T$_2$AG$_3$)$_4$ + 10 µM Δ-[Ru(phen)$_2$(dppz)]$^{2+}$		0.196±0.078	5.17±0.61
(T$_2$AG$_3$)$_4$ + 20 µM Δ-[Ru(phen)$_2$(dppz)]$^{2+}$	1 mM KCl	0.305±0.033	2.68±0.32
(T$_2$AG$_3$)$_4$ + 10 µM Λ-[Ru(phen)$_2$(dppz)]$^{2+}$		0.261±0.06	3.11±0.23
(T$_2$AG$_3$)$_4$ + 20 µM Λ-[Ru(phen)$_2$(dppz)]$^{2+}$		0.122±0.026	1.23±0.16
(T$_2$AG$_3$)$_4$		0.417±0.067	9.64±0.54
(T$_2$AG$_3$)$_4$ + 5 µM Δ-[Ru(phen)$_2$(dppz)]$^{2+}$		0.230±0.054	4.4±03
(T$_2$AG$_3$)$_4$ + 10 µM Δ-[Ru(phen)$_2$(dppz)]$^{2+}$		0.26±0.053	2.8±1.1
(T$_2$AG$_3$)$_4$ + 20 µM Δ-[Ru(phen)$_2$(dppz)]$^{2+}$	100 mM KCl	0.0623±0.004	0.445±0.05
(T$_2$AG$_3$)$_4$ + 5 µM Λ-[Ru(phen)$_2$(dppz)]$^{2+}$		0.124±0.016	1.54±0.27
(T$_2$AG$_3$)$_4$ + 10 µM Λ-[Ru(phen)$_2$(dppz)]$^{2+}$		0.0524±0.006	0.903±0.067
(T$_2$AG$_3$)$_4$ + 20 µM Λ-[Ru(phen)$_2$(dppz)]$^{2+}$		0.0422±0.004	0.24±0.03

[a] DNA at 5 µM
[b] All experiments performed in 10 mM Tris-HCl (pH 7.5), 1 or 100 mM KCl, and 8 mM MgCl$_2$ at 37 °C

In addition to the high salt scenario, Klenow fragment stop assays were performed in low salt conditions (1 mM KCl) (figure 5.4)(stained gels are presented in figure A5.2). As mentioned before, in these conditions the quadruplex adopts an anti-parallel topology. In the absence of any stabilising ligands, the replication of the template strand is rapid and lacking in any stalled products, with the reaction running to completeness within the first minute (74 % yield). In the presence of 10 µM delta, very little difference is observed in the rate of reaction however addition up to 20 µM drastically retards the replication of the sequence. Introduction of the lambda species more obviously hinders the development of full product; bands aligning to stalls at the site of G-quadruplex formation are observed at both concentrations. Interestingly, at 20 µM the two enantiomers are shown to stall at

distinct regions of the sequence, with the polymerase stalling at the duplex region in the presence of delta but predominantly stalling at the G-quadruplex site in the presence of lambda.
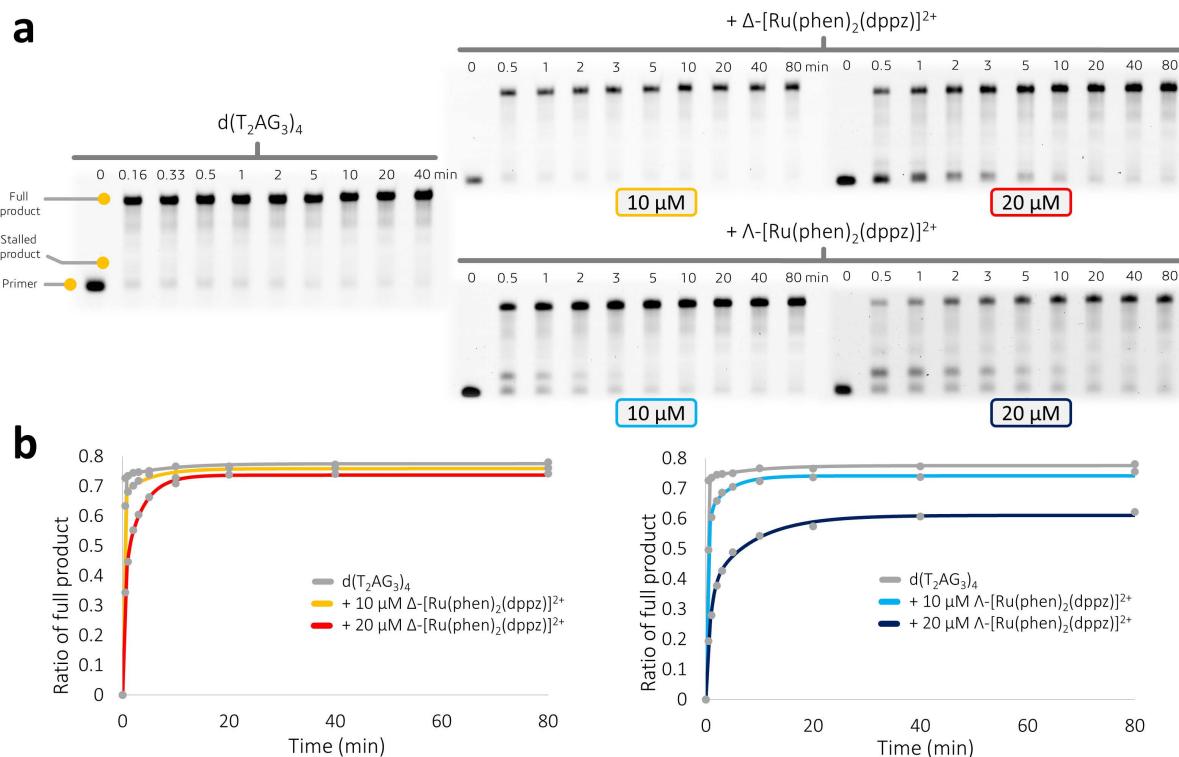

Figure 5.4 – (a) denaturing PAGE gel panels of the time dependent analysis of the KF polymerisation of 1 µM d(T$_2$AG$_3$)$_4$ in the presence of 1 mM KCl and in the absence or presence of the enantiomers of Δ/Λ-[Ru(phen)$_2$(dppz)]$^{2+}$. Gels were visualised using the fluorescence of the FAM labelled primer strand. (b) Time dependent analysis of the reaction yields fitted to a two-step sequential model.

5.3.2 Photoassisted replication stall
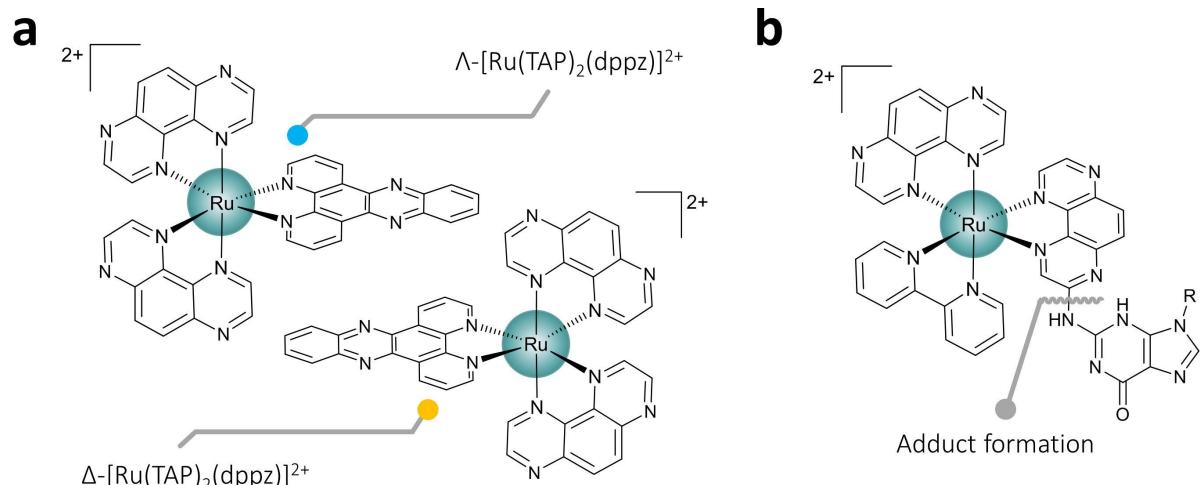
Figure 5.5 – (a) enantiomers of the photooxidising, [Ru(TAP)$_2$(dppz)]$^{2+}$; (b) MLCT excitation of some TAP containing complexes in the presence of guanine containing oligonucleotides poses the possibility of covalent adduct formation. As an example the photoaddition of Λ-[Ru(TAP)$_2$(bpy)]$^{2+}$ to guanine monophosphate is shown, where the adduct is formed between the C3 of TAP and the N15 of guanosine.

Due to the high concentration of electron rich guanines, the abundance of sandwiched π-stacking, and the potential for motif specific binding, the G-quadruplex is considered a prime target for photoassisted damage to DNA. Due to their variable excited-state electronic configurations and MLCT accessibility in the visible-NIR range, metal coordination complexes, and especially ruthenium centred species, have emerged as viable photosensitisers for this task.[8] As discussed before, incorporation of the TAP ligand into such complexes often yields species capable of oxidising and sometimes adducting covalently to DNA upon MLCT excitation (figure 5.5).[9] Here a similar set of experiments to the previous section was performed, however the isoelectronic $[Ru(TAP)_2(dppz)]^{2+}$ was used instead of $[Ru(phen)_2(dppz)]^{2+}$ and the reaction mixtures were irradiated under a Xenon light source (420 nm) prior to enzymatic replication to investigate the possible effects of oxidative stress on the processivity of the quadruplex.
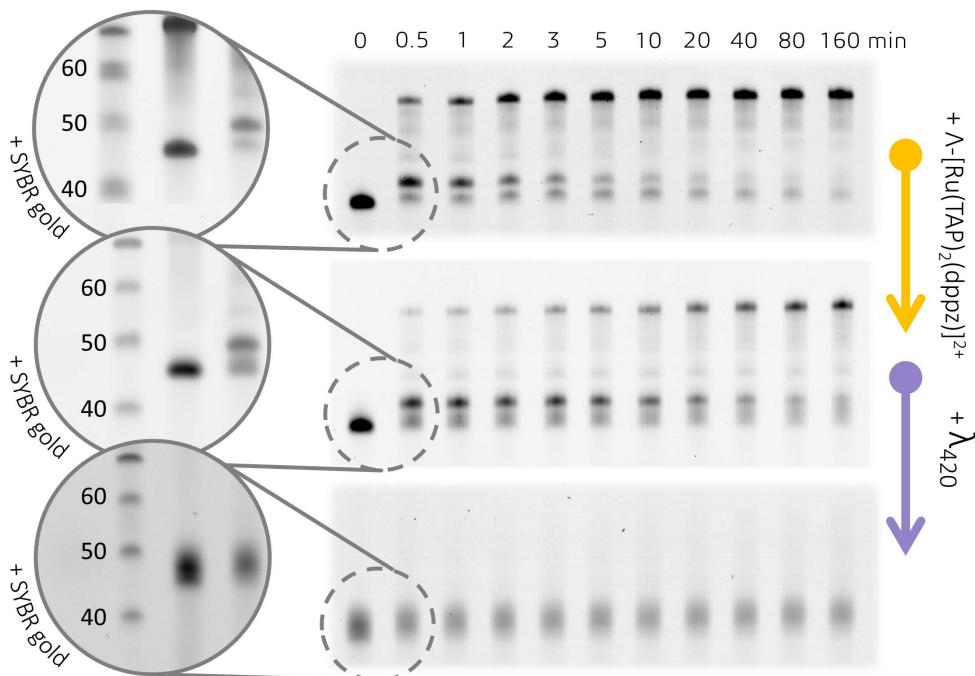
**_Figure 5.6_** – _denaturing PAGE gel panels of the time dependent analysis of the KF polymerisation of 1 μM d(T$_2$AG$_3$)$_4$ in the presence of 100 mM KCl and upon addition of 20 μM Λ-[Ru(TAP)$_2$(dppz)]$^{2+}$ with or without irradiation of the sample at 420 nm for 160 mins. Gels were visualised using the fluorescence of the FAM labelled primer strand. The same gels post-staining with SYBR gold are shown inset, with the DNA ladder lane (shown in units of nts) run alongside the initial reaction aliquot. Note the smearing of the bands and their increase in molecular weight following irradiation._

Electrophoretic analysis of one of the irradiation studies is presented in figure 5.6, showing changes in gel band profile, reaction coordinate, and molecular weight of the FAM-labelled primers. Almost analogously to the phen derivative, the presence of the Λ-Ru-TAP species alone hinders the replication of the sequence, stalling polymerisation at the G-quadruplex until it unfolds and is subsequently processed. If, however, the reaction mixture is irradiated prior to introduction of the Klenow fragment then clear deviations from the expected behaviour of the reaction are observed.

Most apparent is the complete loss of any full-length reaction products even 160 mins after initiation with the Klenow fragment, suggesting that whatever damage is occurring is acutely interrupting the enzymatic process. The exact damage pathway could be up for contention however; $^1O_2$ sensitisation and adduct formation are both potentially viable modes but alternatively it may be possible that the far visible illumination could damage the native DNA even in the absence of complex. As such, additional irradiation experiments were conducted to help narrow the hypotheses; in one case no ruthenium was added, effectively irradiating native DNA for a full 320 mins. In the other case, the experiment was repeated in the presence of ruthenium but was conducted under an argon atmosphere (see figure A5.3 for results). In both cases the altered reaction variable made little to no difference to the original replication kinetics implying that neither direct DNA damage nor $^1O_2$ synthesis is responsible for the polymerase inhibition. Indeed, upon observation, smearing of the formerly discrete bands on the PAGE gels is seen after prolonged exposure (figure 5.6). Inset of the figure shows the same gels following staining with SYBR gold to visualise the DNA ladder lane and this quite clearly suggests that the molecular weight of a percentage of the primer strand population has increased. Scaled, this increase is approximately 2-2.5 nucleotides (nt) in weight which equates to approximately 680-850 Da. As such it is quite possible that this increase is as a result of the covalent binding of Λ-[Ru(TAP)$_2$(dppz)]$^{2+}$ (748 Da) to the DNA following prolonged MLCT excitation; if this were correct it would also imply a 1:1 stoichiometry in adduct formation, at least doing so at the double stranded region upstream to the G-quadruplex. However, the SYBR gold stained gels reveal an additional quirk; the complete loss of template band at higher doses of radiation or higher complex concentrations under a medium dose. Interestingly, when the same experiment was replicated in the presence of a lower K$^+$ concentration (1 mM) similar smearing of the lower Mw bands is observed in the presence of either enantiomer, but the template strand is mostly retained in the stained gels (figure A5.4). The template strand, which is of course unlabelled and so not visible without staining, contains the single stranded G-quadruplex overhang region. As such it may hypothesised that the complexes are sensitising single strand breaks upon excitation, with an efficacy that is dependent on the topology of the folded single strand. It is unknown whether this topological difference in response is a result of a difference in damage vulnerability of the folded DNA (due to site accessibility e.g. *syn* vs *anti*), or because of the unfolding kinetics (i.e. equilibrium between susceptible unfolded strand or protected folded G-quadruplex). Either way further chromatographic/spectrometric study is required to evidence these claims.

In order to evaluate the dependency of sensitiser concentration and irradiation duration on the replication stalling, systematic analysis of those variables was conducted. Experiments relating to exposure duration were all performed at a Λ-[Ru(TAP)$_2$(dppz)]$^{2+}$ concentration of 20 μM, whereas the

reactions in the experiment set that altered concentration were irradiated for 60 mins each. PAGE gels for the subsequent reactions are shown in figure A5.5 whereas the time-dependent kinetic analysis for all experiments is shown in figure 5.7. It is clear from this analysis that either increasing the concentration or increasing the duration of excitation increases the effectiveness of the DNA damage and subsequently the magnitude of replication stalling. Plotting the final reaction yield against the irradiation time/complex concentration shows how both these relationships follow a negative exponential upon a linear increase in dose. It is possible, but certainly not proven here, that this relationship is causally related to the binding constant of the complex, assuming that any DNA damage is only incurred when bound (i.e. no major contribution from non-specific binding). In further studies the salt concentrations could be varied to ascertain the effect of electrostatic potential, and thus the viability of bound states, on the damage response.
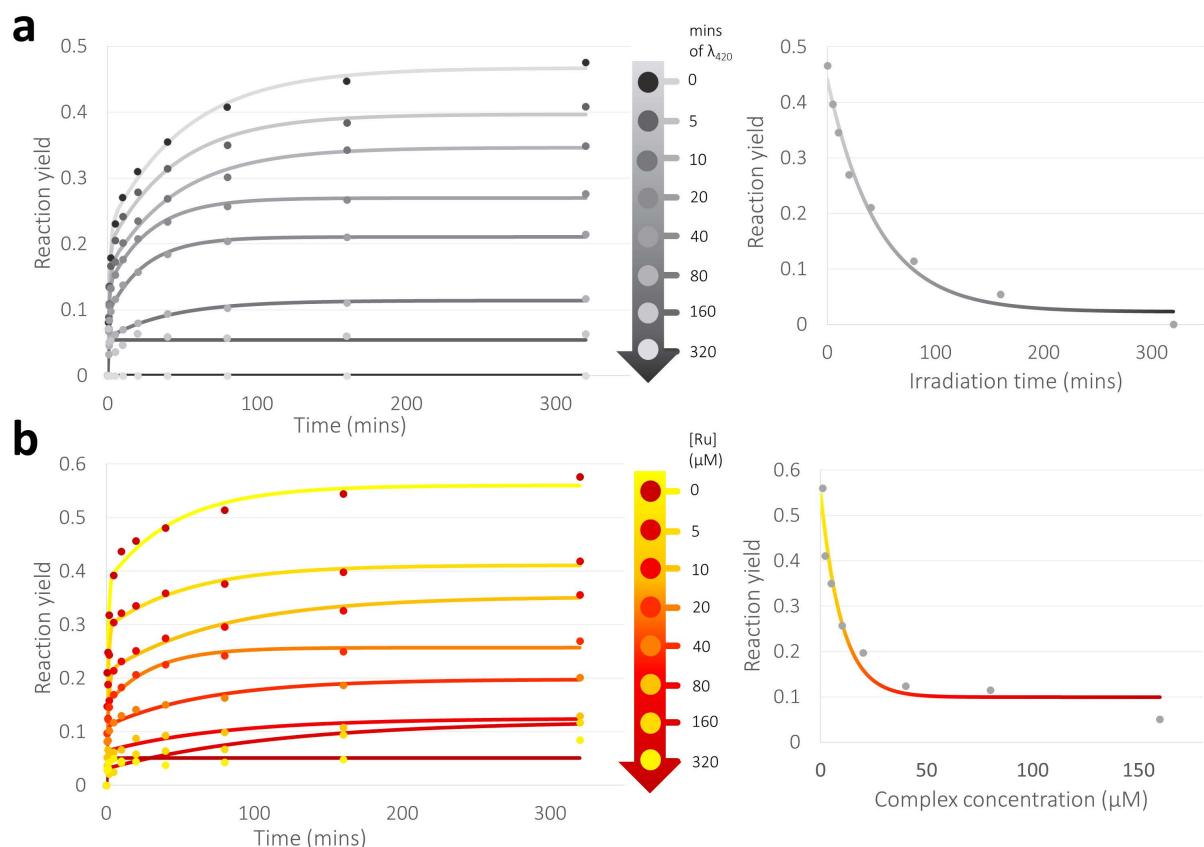
Figure 5.7 –*Systematic kinetic analysis of the replication of the template G-quadruplex d(T$_2$AG$_3$)$_4$ in the presence of Λ-[Ru(TAP)$_2$(dppz)]$^{2+}$ whilst being irradiated at 420 nm. (a) presents the effect of increasing the time that the reaction mixture is irradiated before polymerase introduction whilst keeping the ruthenium concentration constant at 20 μM; (b) presents the effect of increasing ruthenium concentration whilst keeping the irradiation time constant (60 mins). Right of both shows the negative exponential relationship between the reaction yield and the controlled variable by the end of the reaction (320 mins).*

5.3.3 Studies incorporating the Aqphen ligand

Work in this thesis has primarily concentrated on the use of complexes containing the popular intercalating dppz ligand and simple derivatives of such. When binding to a G-quadruplex however, the area of π-surface available (irrespective of backbone clashes) is much larger (roughly 100 Å^2) than is seen in canonical B-DNA base pairs (~9 x 3 Å). As such, using the structural knowledge obtained related to how these complexes may bind to G-quadruplexes, a more expansive ligand was investigated based on its relative size, appropriate geometry and the availability of starting materials. The Aqphen ligand, shown in figure 5.8, consists of a phenanthroline asymmetrically connected to an anthraquinone moiety; by design it fits the contour of the planar G-tetrad but also contains exocyclic oxygens that could potentially interact with the central cationic channel and/or H-bond donors of the nucleobases.
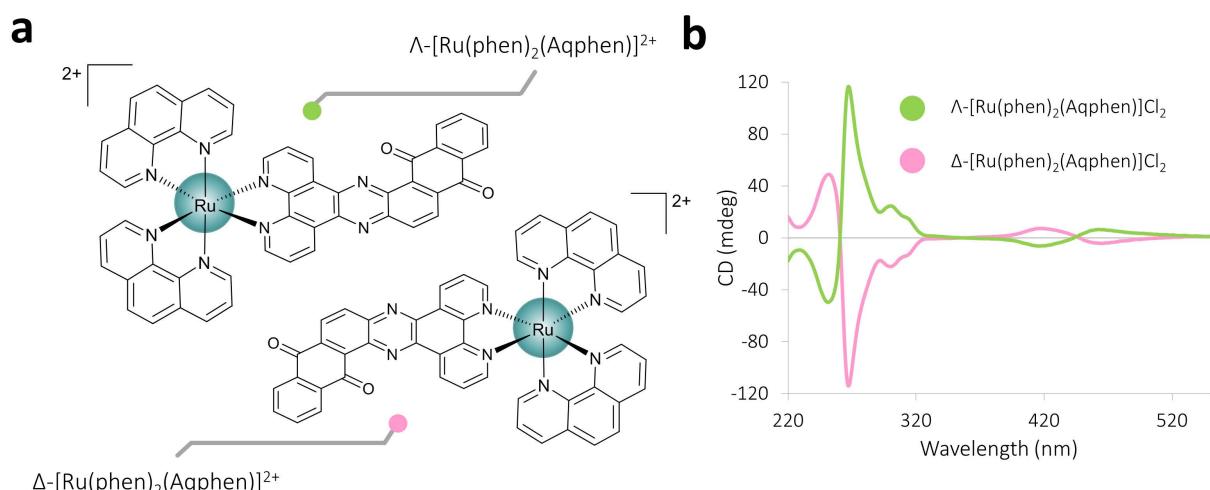
Figure 5.8 – (a) enantiomers of the presented complex [Ru(phen)$_2$(Aqphen)]$^{2+}$; (b) CD spectra of the resolved enantiomers. CD spectra measured at 15 μM complex in H$_2$O in quartz cuvettes with a path length of 10 mm.

5.3.4.1 FRET melting analysis

In order to evaluate the DNA stabilising effect of the extended chromophore a FRET melting assay was employed. The assay, conducted at Imperial College London, contained four oligonucleotides; a duplex 26-mer, a potassium folded C-MYC promoter G-quadruplex, and wtTel22 folded in the presence of either potassium or sodium. Melting profiles were analysed following incubation of the complexes at a range of stoichiometries from 1:1 to 20:1 complex:biological unit using the quenched fluorescence of the FAM-TAMRA labelled oligonucleotides during DNA unfolding. ΔT$_m$ values averaged from triplicate measurements are reported in table 5.2, the full melting curves are shown in figures A5.6 and A5.7.
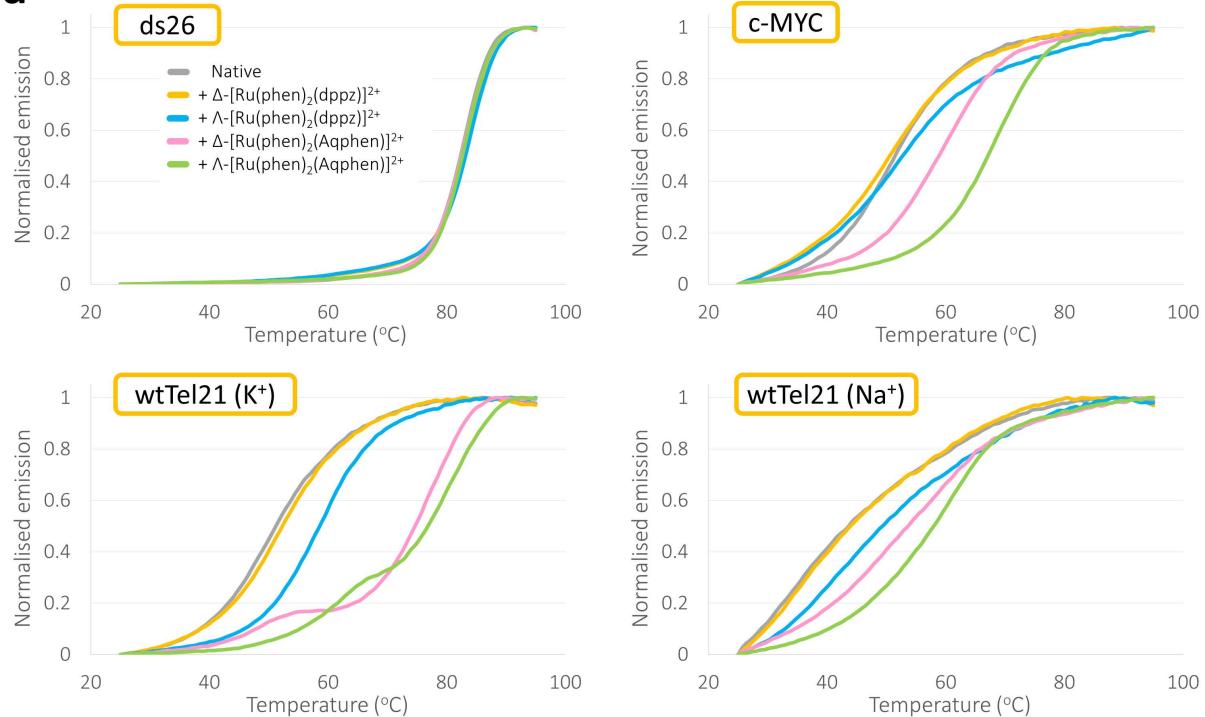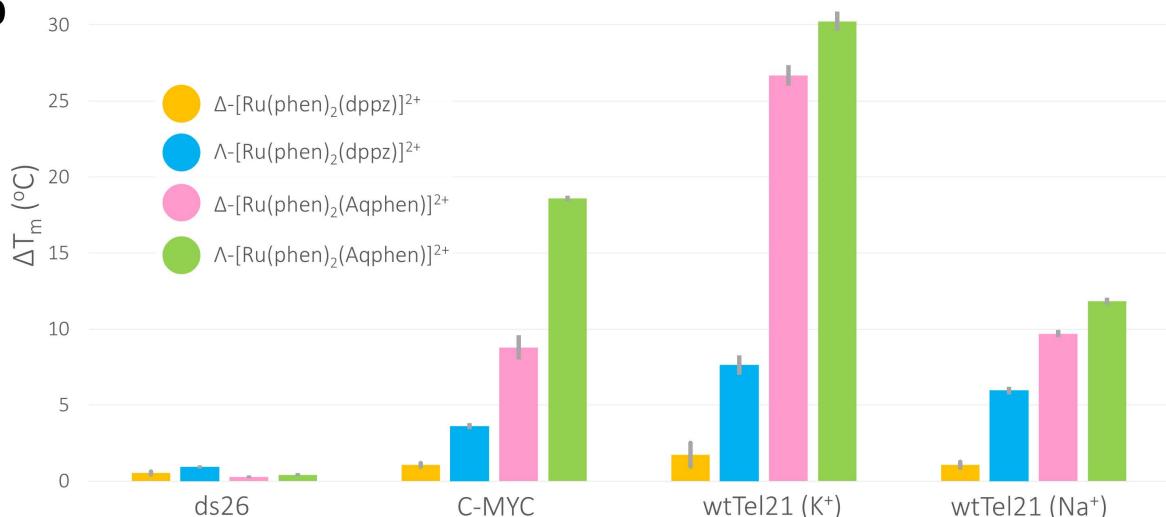
Figure 5.9 – *(a) thermodynamic stability curves from the FRET melting assay of the four DNA motifs in the absence or presence of the enantiomers of [Ru(phen)$_2$(dppz)]$^{2+}$ or [Ru(phen)$_2$(Aqphen)]$^{2+}$. DNA is at 0.4 μM whereas the complexes are at a concentration of 2 μM, therefore at a 5:1 ligand:DNA stoichiometry. C-MYC was measured in the presence of 1 mM KCl, whereas wtTel21 was measured in the presence of 10 mM KCl/NaCl. (b) average ΔT$_m$ values from triplicate measurements for the same systems. In endotherms exhibiting biphasic melt profiles, the major product (determined by absolute hypochromic shift) was used to determine T$_m$.*

Natively the control duplex sequence used has a T$_m$ of 82.4 °C; in the presence of up to 10 eqvs of any of the ligands very little increase in T$_m$ is observed (<2.5 °C), implying an absence of any major intercalative interaction. Previous reports implicate at least modest ΔT$_m$ values for dsDNA binding with Λ/Δ-[Ru(phen)$_2$(dppz)]$^{2+}$, and although only conducted with a racemic mixture, *rac*-[Ru(phen)$_2$(Aqphen)]$^{2+}$ has been shown to stabilise CT-DNA by as much as 6 °C.[10] The

G-quadruplexes were however all stabilised by the presence of ligands, by differing degrees and with differing enantiomeric selectivities. Figure 5.9 shows the melting profiles and subsequent ΔT_m values for each interaction at a 5:1 stoichiometry which is highlighted due to its general preference within the literature as a reference point. Clearly the complexes containing the Aqphen ligand were more effective than dppz at stabilising the G-quadruplexes across the board. Within the enantiomeric pairs, the lambda species was found to stabilise to a higher degree than the delta isomer, which is consistent with previous literature and with previous conclusions in this thesis. Of all the sequences tested, the potassium folded wtTel21 sequence was stabilised the most significantly (+30.2 ^{o}C), by Λ-[Ru(phen)$_2$(Aqphen)]$^{2+}$, but interestingly this mixed hybrid G-quadruplex was also stabilised by +26.7 ^{o}C by the delta enantiomer of the same complex. Interestingly, and exclusively for this set, the melting profiles for both these systems exhibit signs of multiple endotherms, with bimodal denaturation patterns consistent with the existence of multiple binding modes (figure 5.9a). The respectively lower hypochromic shifts exhibited by the lower stability product in these two cases implies, either an equally interactive but minor populated product, or an interactive mode involving less π-π stacking such as loop binding; interestingly this observation does not persist at higher stoichiometries. The largest enantiomeric disparity between the Aqphen complexes is observed with the parallel folded C-MYC sequence. With this sequence the presence of the lambda isomer induced a thermal stability nearly twice as high as the delta (+17 ^{o}C and +8.8 ^{o}C respectively). Conversely, comparatively less difference is observed when the Aqphen complexes bind to the potassium and sodium forms of wtTel21. Binding of the Aqphen complexes to the presumably mixed hybrid and anti-parallel quadruplexes leads to a respective 3.3 and 2.7 ^{o}C enantiomeric difference in T_m ($|\Delta\Delta T_m|$), implying that the ancillary ligand positions play a more important role in the binding of this complex to the parallel stranded quadruplex than to the other topologies.

Table 5.2 – *calculated ΔT_m values derived from FRET melting analysis*

	eqvs	ds26	c-MYC	wtTel21 (K$^+$)	wtTel21 (Na$^+$)		eqvs	ds26	c-MYC	wtTel21 (K$^+$)	wtTel21 (Na$^+$)
Δ-[Ru(phen)$_2$(dppz)]$^{2+}$	1	0.028 ±0.09	-0.25 ±0.23	0.80 ±0.58	-0.76 ±0.32	**Δ-[Ru(phen)$_2$(Aqphen)]$^{2+}$**	1	0.32 ±0.09	2.74 ±0.19	8.00 ±0.68	3.70 ±0.31
	2	0.07 ±0.09	0.27 ±0.33	-0.20 ±1.55	-0.66 ±0.27		2	0.19 ±0.09	3.58 ±0.19	12.70 ±1.77	1.27 ±0.27
	5	0.53 ±0.11	1.08 ±0.13	1.73 ±0.79	1.08 ±0.21		5	0.27 ±0.10	8.79 ±0.79	26.69 ±0.67	9.69 ±0.24
	10	1.31 ±0.11	2.08 ±0.26	3.67 ±0.66	2.13 ±0.30		10	0.40 ±0.14	16.02 ±0.17	30.27 ±0.60	17.84 ±0.31
	20	3.02 ±0.09	4.44 ±0.11	6.63 ±0.55	5.22 ± 0.22		20	5.88 ±0.12	32.40 ±3.81	≈40 ±0.61	37.07 ±0.24
Λ-[Ru(phen)$_2$(dppz)]$^{2+}$	1	0.05 ±0.11	-0.09 ±0.23	3.08 ±0.59	0.84 ±0.24	**Λ-[Ru(phen)$_2$(Aqphen)]$^{2+}$**	1	0.22 ±0.09	7.13 ±0.59	11.04 ±0.17	2.50 ±0.23
	2	0.06 ±0.12	1.34 ±0.12	3.99 ±1.12	2.48 ±0.34		2	0.27 ±0.10	10.51 ±1.20	17.51 ±0.23	6.66 ±0.24
	5	0.93 ±0.11	3.61 ±0.20	7.64 ±0.65	5.949 ±0.248		5	0.41 ±0.11	16.98 ±0.65	30.25 ±0.18	11.83 ±0.24
	10	2.28 ±0.13	7.96 ±0.18	12.09 ±0.64	9.14 ±0.38		10	1.79 ±0.11	25.69 ±0.58	37.41 ±0.12	26.50 ±0.23
	20	4.28 ±0.10	11.29 ±0.15	16.99 ±0.58	13.87 ±0.22		20	6.80 ±0.17	≈45 ±0.72	>50 ±0.50	41.66 ±0.18

5.3.4.2 Klenow fragment polymerase assay

In an analogous fashion to that already described and following successful chiral separation (figure 5.8), the enantiomers of $[Ru(phen)_2(Aqphen)]^{2+}$ were investigated with respect to their ability to stall the replication of the topologically diverse telomeric sequence $d(T_2AG_3)_4$. Replication reactions were followed for 160 mins instead of 80 mins as initial trials suggested a much slower polymerisation in the presence of the extended ligand but also in comparison with the parent species, $[Ru(phen)_2(dppz)]^{2+}$. Denaturing PAGE gels (figure 5.10a) present the reaction progress by following the polymerisation of the fluorescent tagged primer strand; locations of the stall in relation to the position of the primer strand were confirmed following staining the gel with SYBR GOLD by comparison with a DNA ladder lane ran alongside the lanes presented (figure A5.8).

Inspection of the PAGE gels show quite clearly that upon addition of either enantiomer of the Aqphen complex the reaction is significantly stalled. At higher concentrations (20 µM) the replication is effectively arrested, and full product bands become quite difficult to decipher even if the reaction is run for as long as 320 mins (data up to 320 mins not shown). For this reason, as with the excitation reactions, quantitative rate analysis is not presented. The predominant stall at the site of the G-quadruplex fold is the same as seen for previous experiments (at ~49 nucleotides) and persists for the length of each experiment. Plotting the product ratio against the time course gives an idea of the kinetics of the formation of full product (figure 5.10b). In all cases $[Ru(phen)_2(Aqphen)]^{2+}$ outperforms the $[Ru(phen)_2(dppz)]^{2+}$ enantiomers quite drastically; total product yields in the presence of Aqphen are <20 % and <12 % for Δ and Λ respectively whereas in the presence of the parent complex yields of ~60 % and ~50 % were observed for the enantiomers respectively, albeit at a drastically slower rate than in the absence of ruthenium. Interestingly, the same experiments were carried out in the presence of the respective racemic mixtures at the same concentration and on all fronts these mixtures performed mildly better than the separate enantiomers in isolation. In the presence of *rac*-$[Ru(phen)_2(Aqphen)]^{2+}$ only 8 % product was observed across the entire reaction, whereas *rac*-$[Ru(phen)_2(dppz)]^{2+}$ had a final yield of 55 %. This could be a result of the higher stoichiometries of the reaction and can be rationalised by the occurrence of disparate binding modes existing for each enantiomer discretely to each other, or alternatively by cooperativity in binding allowing for more complexes to bind to each biological unit.
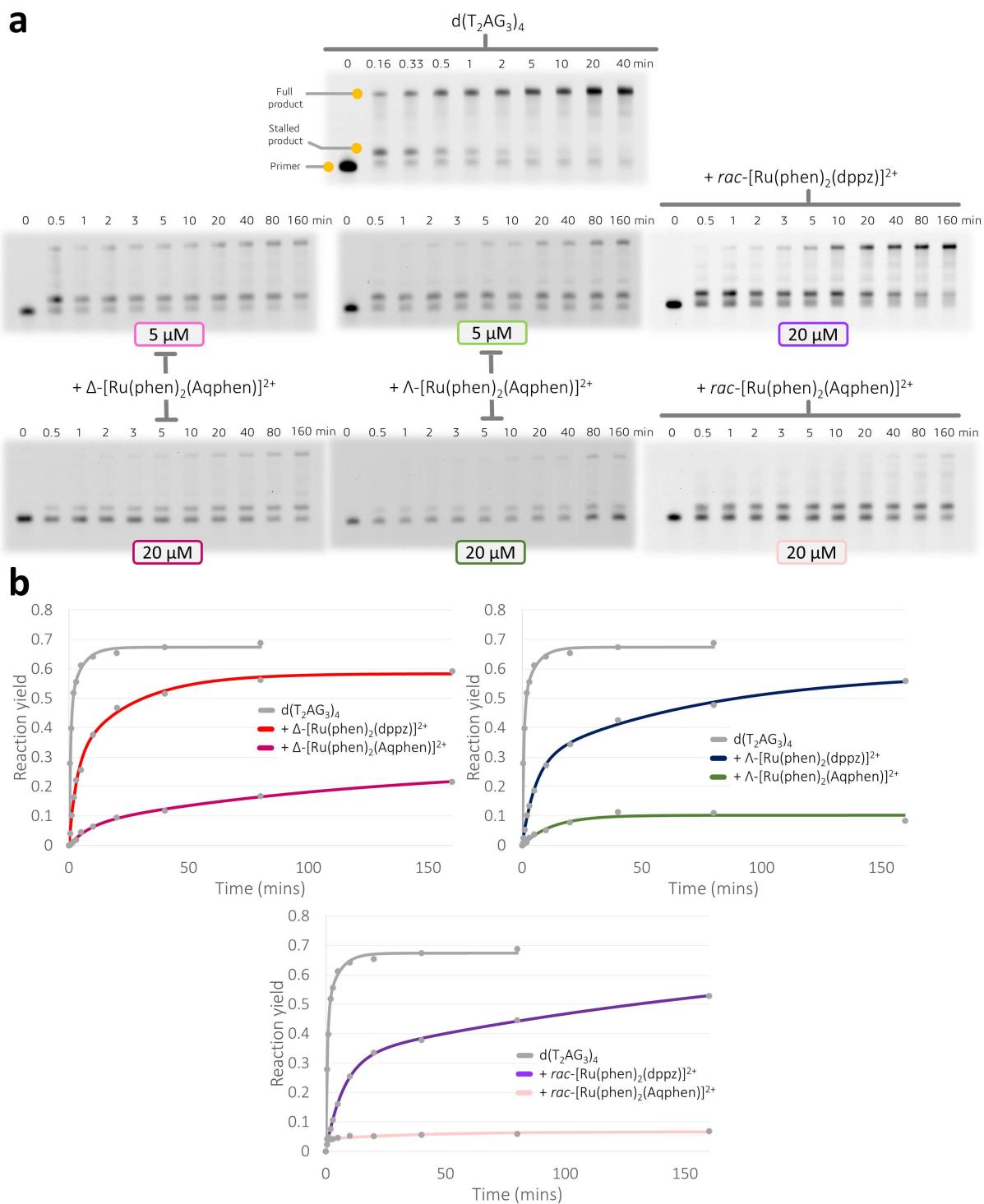
Figure 5.10 – (a) denaturing PAGE gel panels of the time dependent analysis of the KF polymerisation of 1 μM d(T$_2$AG$_3$)$_4$ in the presence of 100 mM KCl and in the absence or presence 5/20 μM of the enantiomers of Δ/Λ/rac-[Ru(phen)$_2$(dppz)]$^{2+}$ or Δ/Λ/rac-[Ru(phen)$_2$(Aqphen)]$^{2+}$. Gels were visualised using the fluorescence of the FAM labelled primer strand. (b) Time dependent analysis of the reaction yields in the presence of the different enantiomers/mixtures (20 μM) fitted to a two-step sequential model. It is evident that the Aqphen derivatives inhibit the polymerisation to a much higher extent than the dppz, and that the lambda species is again the dominant enantiomer in terms of stalling power.

In addition, the same experiment was repeated at low K$^+$ concentrations (1 mM)(figure 5.11).

Natively, this sequence folded at low K$^+$ exists as an anti-parallel basket quadruplex and is replicated

rapidly with an almost 100 % yield within the first minutes of reaction. In similar fashion to what was observed earlier, no distinct stall band is observed, instead when incubated in the presence of the complexes the polymerase action is retarded from the origin of the overhang. Nevertheless, both enantiomers of $[Ru(phen)_2(Aqphen)]^{2+}$ were able to hinder the reaction significantly, with product yields of 30 % and 25 % for Δ and Λ respectively. In comparison the parent enantiomers, which generate similar gel electrographs, impede the reaction rate but inevitably the resulting reaction stall is unfolded considerably faster and with much higher final yields, 74 % and 61 % respectively for Δ and Λ.
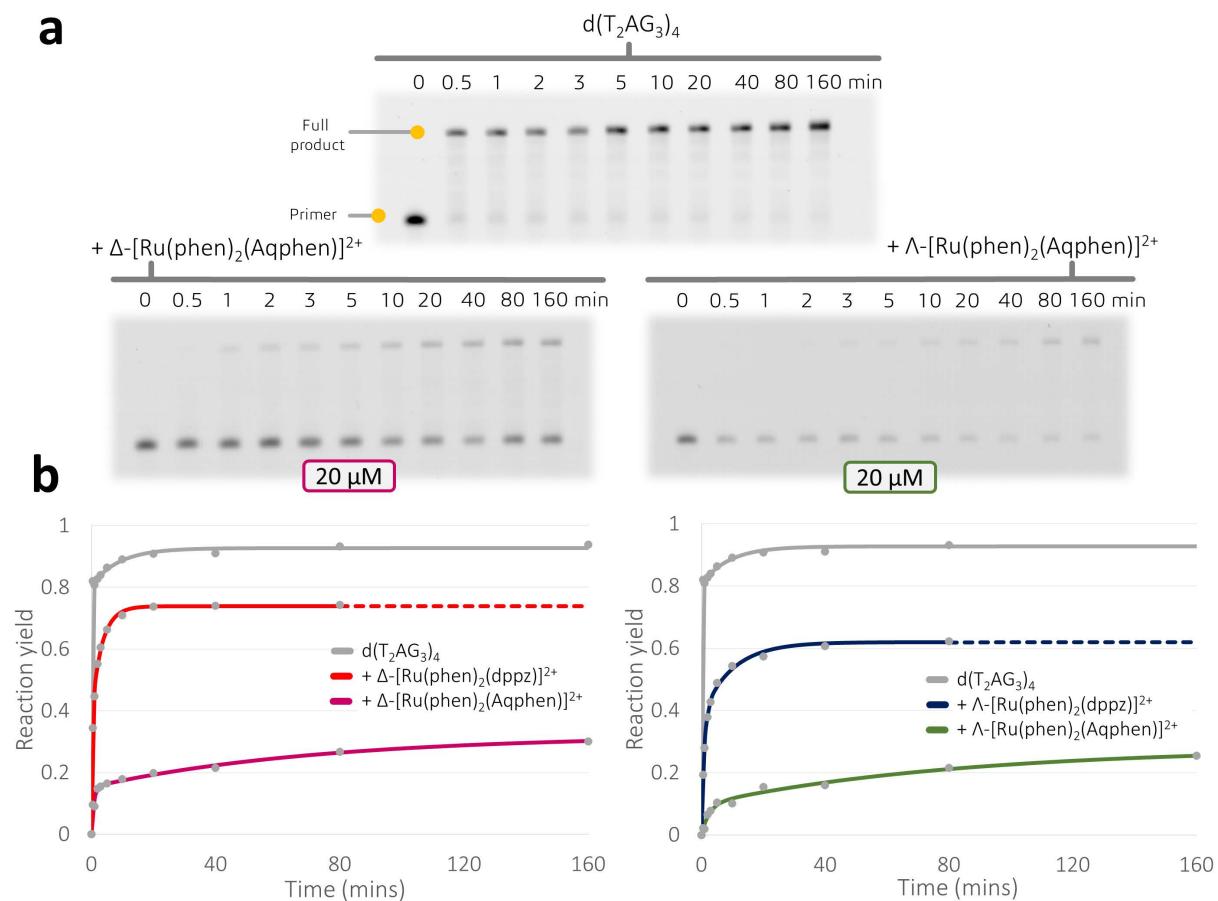
**Figure 5.11** – (a) denaturing PAGE gel panels of the time dependent analysis of the KF polymerisation of 1 μM d(T$_2$AG$_3$)$_4$ in the presence of 1 mM KCl and in the absence or presence 20 μM of the enantiomers of Δ/Λ-[Ru(phen)$_2$(Aqphen)]$^{2+}$. Gels were visualised using the fluorescence of the FAM labelled primer strand. (b) Time dependent analysis of the reaction yields in the presence of the different enantiomers (20 μM) fitted to a two-step sequential model. Dashed lines are an extrapolation on the linear component of the curves.

These results suggest that the Aqphen complexes bind strongly and thus drastically impede the unfolding kinetics of the G-quadruplex, effectively blocking the processability of the sequence. The Aqphen species performed significantly better than the parent dppz species which is consistent with the earlier FRET analysis that evidenced the superior capacity of the Aqphen complexes to stabilise the selected G-quadruplex folds.

5.3.4.3 BG4 immunofluorescence assay

The BG4 antibody binds to intramolecular and intermolecular G-quadruplexes at nanomolar affinities whilst exhibiting no discernible affinity to single or double stranded DNA.[11] When incubated with fixed cells and a fluorescent secondary antibody, the protein allows for the visualisation of G-quadruplexes *in vivo* using confocal microscopy. In a similar fashion to published work using the pyridostatin ligand, MCF-7 breast cancer cells were pre-incubated in the presence of Λ/Δ-[Ru(phen)$_2$(Aqphen)]$^{2+}$ or Λ/Δ-[Ru(phen)$_2$(dppz)]$^{2+}$ before addition of BG4 to assess the effect of the extended ligand on the formation of nuclear G-quadruplexes. Following microscope imaging, fluorescence foci were identified, and their intensities analysed in the ZEN software suite. Figure 5.12 shows the processed photomicrographs and summarises the BG4 foci intensities of the inoculated cells relative to cells in the absence of the metal complexes (unprocessed data is shown in figure A5.9). It is also important to note here that the excitation wavelength of the secondary fluorescent antibody (λ_{ex} = 545 nm) lies predominantly outside of the MLCT absorption range of the ruthenium complexes tested but a small overlap is observed with the tail of the transition ($\varepsilon < 0.15 \times 10^4$ M^{-1}cm^{-1}) (see figure A5.10 for absorption/emission profiles for each molecular species).
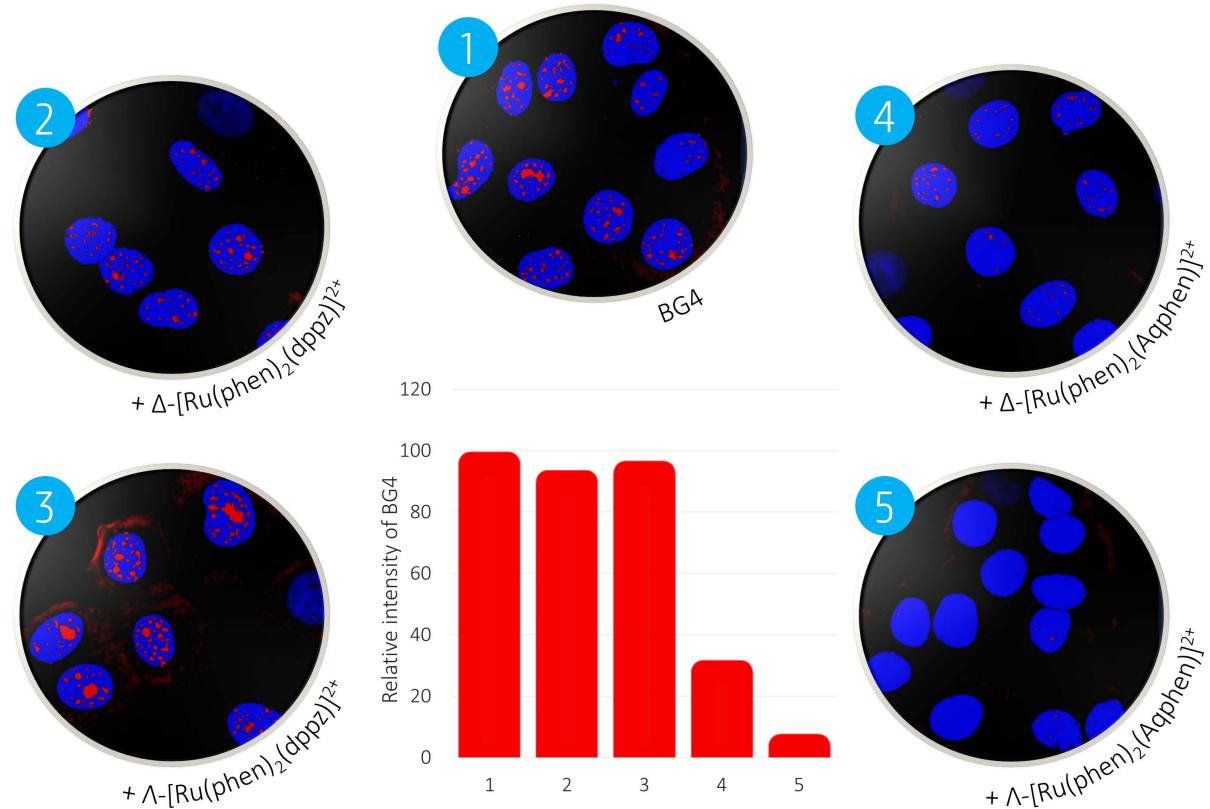
Figure 5.12 – *Foci-counting confocal micrographs of MCF-7 cells inoculated with modified fluorescent BG4 antibody (red) and counterstained with DAPI nuclei stain (blue). Cells were incubated either in the absence (1) or presence of ruthenium complexes; Δ/Λ-[Ru(phen)$_2$(dppz)]$^{2+}$ (2/3) or Δ/Λ-[Ru(phen)$_2$(Aqphen)]$^{2+}$ (4/5). Inset shows the foci fluorescence intensity of each system, relative to the cells incubated in the absence of ruthenium.*

In the absence of any ligand, fluorescent foci are observed within the nucleus where they tend to cluster together with smaller foci more sparsely located in the nucleoplasm. Addition of either lambda or delta [Ru(phen)$_2$(dppz)]$^{2+}$ has little to no effect on the total intensity of the counted foci (95 and 98 % of the native BG4 fluorescence respectively) however the presence of the lambda enantiomer causes fewer but larger and more bright foci than is seen natively or in the presence of delta. This would suggest that the complexes are either not able to effectively stabilise/induce G-quadruplex formation *in vivo*; or the complexes are not localising in the nucleoplasm where the majority of G-quadruplexes are observed. When incubated with the enantiomers of [Ru(phen)$_2$(Aqphen)]$^{2+}$ much more apparent changes are observed. Both optical isomers reduce the intensity of the BG4 response dramatically, with emission in the presence of delta down to 26 % and lambda down to 7 % of the original response. Originally it was expected that the stabilisation of nuclear G-quadruplexes, which would cause an increase in G-quadruplex concentration at any one time, would also cause an increase in BG4 binding and thus fluorescence response. The binding mode of the BG4 antibody to quadruplex DNA is however currently unknown which makes speculation of how they are interacting much more difficult. Such a response could be caused by a number of reasons, but this could imply that the larger Aqphen complexes are displacing the BG4 antibodies from the G-quadruplexes formed *in vivo* subsequently reducing the fluorescence response. Also, of note is the fact that the antibody is non-specific in regard to quadruplex topology and as such no conclusion can be made on the *in vivo* topological specificity of the complexes. However, as a result of the very low antibody fluorescence response it can likely be implied that the complexes also bind to G-quadruplexes with little topological specificity.

5.3.4.4 cell viability

The studies conducted with the BG4 antibody led to the subsequent implication of strong interaction of the Aqphen complexes with G-quadruplexes *in vivo*. Such interactions could affect the normal functions of the cell leading to potential cytotoxicity. As such it is necessary to investigate the viability of living cells in the presence of high doses of the complexes as a means of evaluating the inherent toxicity of the complexes. HeLa cells were incubated for 92 hours in the presence of the complexes at varying concentrations (1.25, 2.5, 5, 10, and 20 µM); cells were then stained and the viable cells counted using the absorbance of the stain. Results for the viability studies are shown in table 5.3 and figure 5.13 and an expanded range of complexes relevant to this thesis were also tested and their results are shown in table A5.1 and figure A5.11. Unfortunately, the viability studies were conducted before the enantiomeric separation of [Ru(phen)$_2$(Aqphen)]$^{2+}$ was successfully performed.
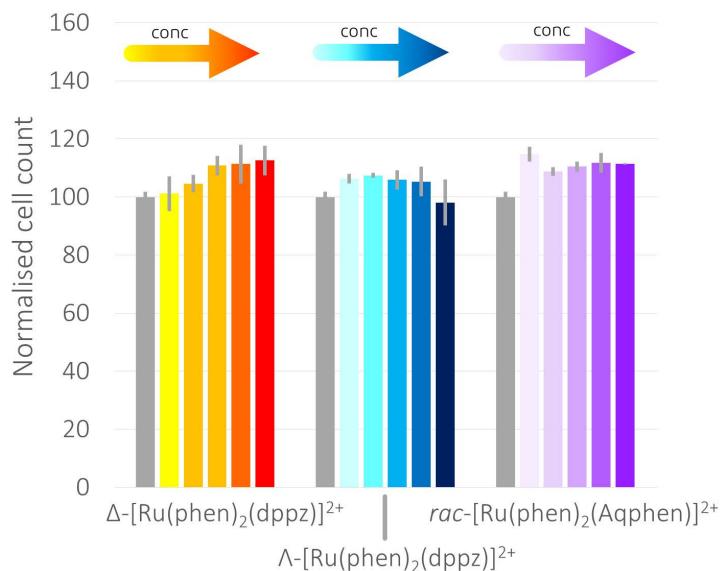
Figure 5.13 – *HeLa cell counting in the absence or presence of a range of ruthenium complexes at a range of concentrations (0, 1.25, 2.5, 5, 10, and 20 µM). Cells were incubated with the ligands for 24 hours in 96-well plates before cell counting was achieved using a commercial absorbance counting kit. Absorbances were normalised in relation to the control cells.*

Table 5.3 – *Normalised absorbance values for the counting of HeLa cells after incubation with a selection of ruthenium complexes.*

Complex	Normalised absorbance					
	0 µM	1.25 µM	2.5 µM	5 µM	10 µM	20 µM
Δ-[Ru(phen)$_2$(dppz)]$^{2+}$	100 ±1.75	101.11 ±5.97	104.55 ±2.99	110.80 ±3.31	111.32 ±6.66	112.55 ±5.12
Λ-[Ru(phen)$_2$(dppz)]$^{2+}$	100 ±1.75	106.26 ±1.60	107.40 ±0.89	105.84 ±5.08	105.26 ±5.08	98.03 ±7.91
rac-[Ru(phen)$_2$(Aqphen)]$^{2+}$	100 ±1.75	114.66 ±2.58	108.70 ±1.48	110.44 ±1.78	111.66 ±3.44	111.34 ±0.24

It is evident then that even high concentrations of ruthenium complex, whether that be the dppz or Aqphen variant, has little negative effect on the viability of the cell line (in fact there is a small viability increase at high concentrations). This also held true for all the ruthenium complexes tested (figure A5.11 and table A5.1). Especially in relation to the G-quadruplex stabilising Aqphen species, this lack of cytotoxicity implies that membrane permeability or cell localisation may be hindering the complexes activity. Alternatively, the complex may localise within the nuclear envelope, as observed in the immunofluorescence studies, but DNA binding may not significantly affect the cellular processing. Of course these experiments were conducted over a relatively short period so do not properly portray the anti-proliferative effect as much as the ability to induce sudden cell apoptosis or autophagy.

5.3.4.5 molecular modelling

DFT calculations at the B3LYP/LAN2LZ level show that the Aqphen ligand is completely planar in the complex of [Ru(phen)$_2$(Aqphen)]$^{2+}$. Initially it was considered that, due to potential bay area repulsions between the ketone oxygen and pyrazine nitrogen on the inside of the pseudo hemicycle,

the ligand would be buckled to accommodate the clash in moieties. Crystallography of the free ligand later confirmed this planarity (data not shown). Calculated frontier molecular orbitals are presented in figure A5.12, and show how the localisation of occupied electronic density is spread predominantly on the Aqphen ligand (with both σ and π character), whereas unoccupied frontier orbitals have a higher distribution across the phen ligands with a small contribution from metal d orbitals.
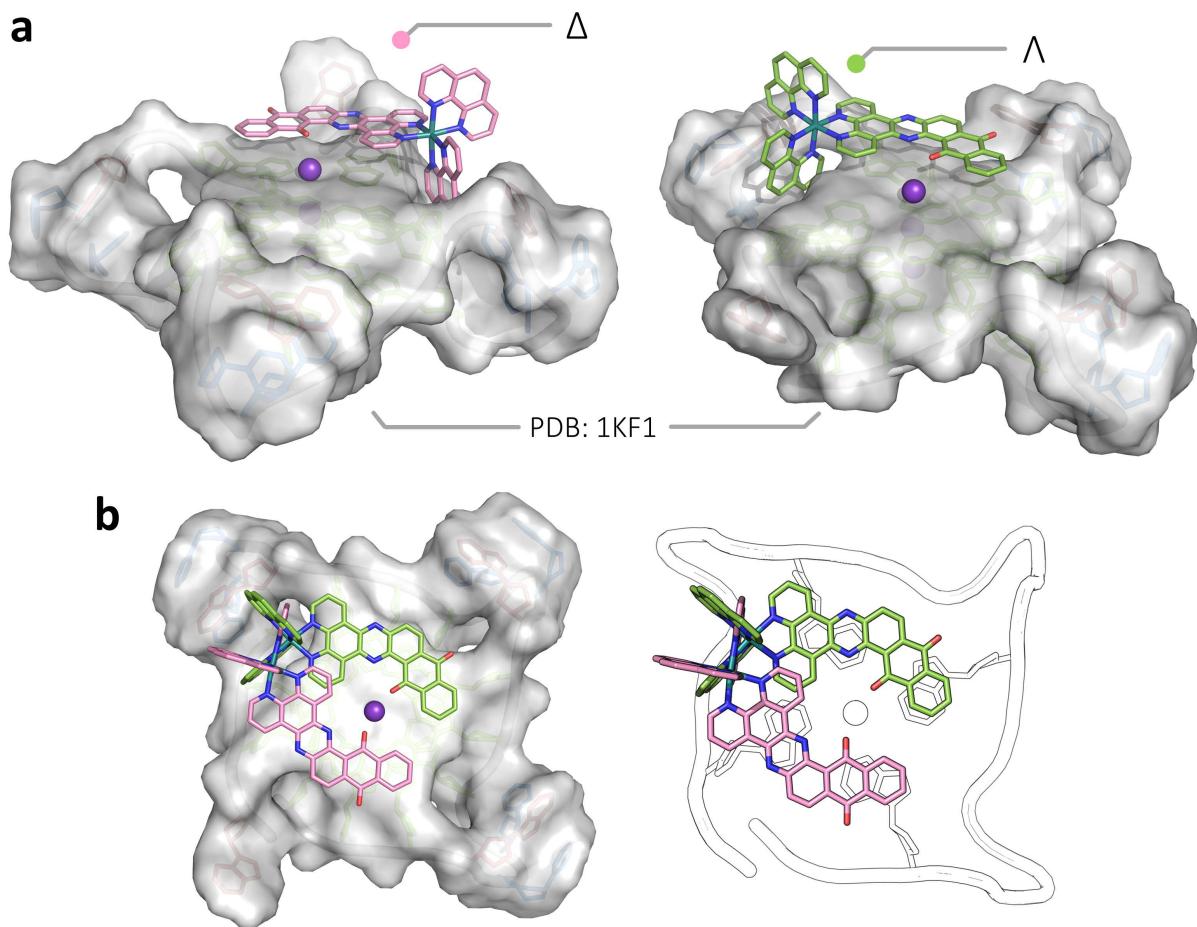
Figure 5.14 – *Top seeded cluster from the rigid molecular docking simulations for both Λ/Δ-[Ru(phen)₂(Aqphen)]²⁺ bound to the wtTel21 sequence (PDB: 1KF1). (a) an overall view of the interaction highlighting how the complexes end-stack onto the 5' terminal tetrad; (b) a birds' eye view of both dockings overlapped to show how the complexes wrap around the tetrad in opposite directions, both stacking with two pairs of distinct guanosines and both forming polar contacts with the distal potassium ion.*

Following DFT calculations, the optimised structures of Δ/Λ-[Ru(phen)₂(dppz)]²⁺ and Δ/Λ-[Ru(phen)₂(Aqphen)]²⁺ were simulated binding electrostatically to the telomeric G-quadruplex wtTel22 (d(AG₃(T₂AG₃)₃). The structural coordinates, which are crystallographically derived, present a parallel quadruplex with three propeller 'fold-back' loops.[12] This structure was chosen for initial calculations due to the availability of the terminal tetrads for electrostatic docking since the docking simulation is based on a rigid target model. Figure 5.14 shows the top clustered conformation for each enantiomer of [Ru(phen)₂(Aqphen)]²⁺ (only one cluster was found for each system implying

unambiguity in calculations). Clearly the binding modes are quite similar between the isomers, with the Aqphen ligand dominating the π-stacking interactions and presumably the predominant forcing condition for the observed docked states. The Aqphen ligand in each case overlaps with two guanines in the terminal 5' tetrad, however, each enantiomer covers different bases, wrapping itself around the tetrad in opposing directions (lambda-clockwise, delta-anticlockwise) with the ruthenium centre originating from the same first groove. In addition, the ligands have arranged such that the endocyclic oxygen from each enantiomer coordinates to the distal potassium ion (K^+--O = 2.6 Å) as first thought. Furthermore, as a result of this directional preference, the lambda complex is stacking with guanines from the beginning of the sequence and from the end, effectively acting as a non-covalent linker between the two ends. If such general traits are true in solution then this could explain to some degree the preference for the lambda isomer to stabilise and inhibit enzymatic processes since it would hinder the initial breathing motions that would initialise full unwinding of the folded motif.

5.4 Discussion

5.4.1 Enantiomeric preferences

Differences in the binding intricacies of the enantiomers of ruthenium polypyridyls to G-quadruplex DNA are fairly well documented. In general, spectroscopic, luminescence, and excited state binding studies have all demonstrated that the lambda enantiomer binds to G-quadruplexes with higher efficacy than the delta, imparting higher DNA stabilisation and quantum yields in their luminescence.[13–15] This is indeed the converse of observations when using dsDNA, where delta predominantly binds stronger and luminesces brighter than the lambda.[16] Studies presented here exhibit similar enantiomeric disparities in the ability to stall the replication of the telomeric G-quadruplex, irrespective of G-quadruplex topology and ionic strength. Enzymatic inhibition via the stabilisation of G-quadruplexes has been previously observed using ruthenium polypyridyls, however these studies have predominantly used racemic mixtures. The complexes rac-[Ru(bpy)$_2$(4idip)]$^{2+}$ and rac-[Ru(phen)$_2$(4idip)]$^{2+}$ (figure 5.15) have for example been shown to effectively arrest the polymerisation of the wtTel21 telomeric sequence in both Telomerase Repeated Amplification Protocol (TRAP) assays, and Polymerase Chain Reaction (PCR) stop assays, with the phen derivative universally being shown to be a more effective binder to the mixed hybrid structure than the bpy variant.[17] Similarly, the complexes rac-[Ru(phen)$_2$(ptpn)]$^{2+}$, rac-[Ru(phen)$_2$(hqpdip)]$^{2+}$, rac-[Ru(phen)$_2$(phenselenazole)]$^{2+}$, rac-[Ru(phen)$_2$(tip)]$^{2+}$, and rac-[Ru(bpy)$_2$(pemitatp)]$^{2+}$ (figure 5.15a),[18–21] have all been shown to exhibit strong G-quadruplex stabilisation and polymerase inhibitions, with some such as the latter complex exhibiting long term proliferative arrest in HeLa cell growth. Investigation of separated enantiomers in the same context has been the exception. Two

such studies, from the same group, incorporated the derivatised imidazophenanthrolines, ρ-MOPIP or ρ-HPIP, into enantiopure complexes of [Ru(phen)$_2$(L)]$^{2+}$ (figure 5.15b).[15,22] In both cases the lambda species outperformed the delta in induced thermal stabilisation, PCR stop assay analysis, and *in vitro* TRAP efficacy, in the presence of the potassium folded mixed hybrid G-quadruplex wtTel21. Both lambda isomers were subsequently found to accumulate in the nucleus of immortal HepG2 human liver cancer cells and exhibit micromolar cytotoxicities roughly twice as effective as the delta counterparts.

a

rac-[Ru(phen)$_2$(4idip)]$^{2+}$

rac-[Ru(phen)$_2$(hqdip)]$^{2+}$

rac-[Ru(phen)$_2$(ptpn)]$^{2+}$

rac-[Ru(bpy)$_2$(pemitatp)]$^{2+}$

b

Λ-[Ru(phen)$_2$(HPIP)]$^{2+}$
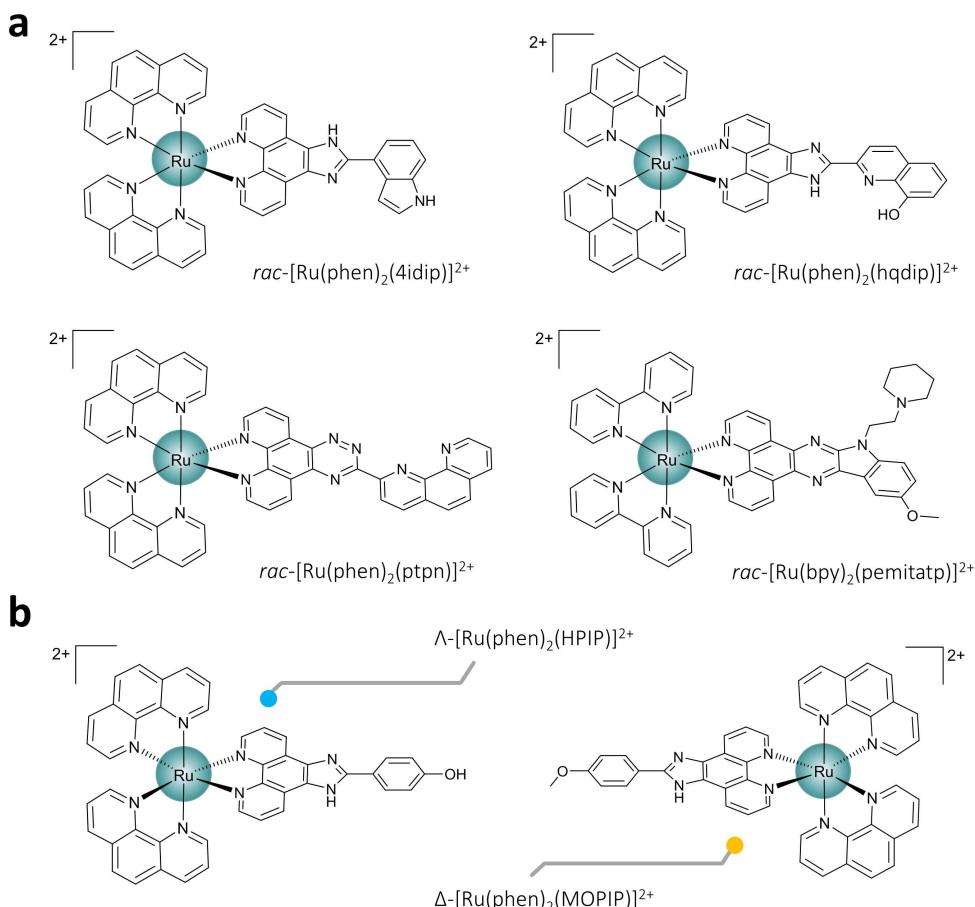
Δ-[Ru(phen)$_2$(MOPIP)]$^{2+}$

Figure 5.15 – *(a) A range of ruthenium centred complexes with studied effects on the enzymatic replication of DNA; (b) the complexes Δ/Λ-[Ru(phen)$_2$(HPIP)]$^{2+}$ and Δ/Λ-[Ru(phen)$_2$(MOPIP)]$^{2+}$ are some of the only enantiopure complexes to be examined in relation to their ability to inhibit polymerase action via G-quadruplex binding.*

Comparison of the FRET melting analyses for the two isomers of [Ru(phen)$_2$(dppz)]$^{2+}$ highlights the preference of the left-handed lambda species which imparts a 7 % to 15 % stabilisation effect to the tested G-quadruplexes in comparison to the 1 % to 3 % imparted by delta binding. Similarly, KF replication of (T$_2$AG$_3$)$_4$ in the presence of Λ-Ru is stalled far more severely, with enzymatic rates around half of that observed in the presence of Δ-Ru at 20 μM complex. It is important to note here that replication arrest tends to be positively correlated to motif thermodynamic stability, but more importantly is related to the unfolding pathways and subsequent susceptibility to polymerase

attachment.[3,23] More thermodynamically favourable intermediates may be present during the unfolding process in the presence of lambda than with delta, such as stabilised triplexes or G-hairpins, that hinder the progress of the replication.[23,24] Structural justification for the generally enantioselective behaviour of such complexes towards G-quadruplexes is still very much in the realm of hypothetics due to scarcity of data and a subsequent survivorship bias due to an absence of detail of delta modes in structural studies. Crystallographic evidence from previous chapters suggests that topologies containing 5'-*syn*-guanosines could be most favourably stabilised by lambda binding as a result of ancillary ribose interactions. Indeed, in regard to the parent $[Ru(phen)_2(dppz)]^{2+}$ species, higher ΔT_ms and larger divergences in enantiomeric response are noted with the folded motifs containing antiparallel polarity character.

5.4.2 Photosensitisation

$[Ru(TAP)_2(dppz)]^{2+}$ is isostructural and isoelectronic to the more famous $[Ru(phen)_2(dppz)]^{2+}$ but differs primarily in its excited state behaviour, exhibiting potential for photo-induced proton-coupled electron transfer with DNA under MLCT excitation as opposed to the light switch phenomenon exhibited by the phen complex.[25] Using reduction potentials of measured complexes, it was found that at least two π-deficient TAP ligands are required to necessitate reaction with guanosines.[26] In the results presented here the lambda species stalls the polymerase action of the Klenow fragment in an analogous fashion to the phenanthroline analogue, however upon irradiation complete cessation of the process was observed. This reaction occurs irrespective of oxygen availability, implying a type I photoreaction caused by photo-induced charge transfer (PET) as opposed to a $^1O_2$-mediated type II reaction. In addition, upward smearing of the aliquot bands on the PAGE gels, even before addition of the Klenow fragment, implies the formation of 1:1 DNA-complex adducts. Addition of the KF polymerase does little to process the template strand and, at either higher [Ru] or higher irradiation durations, there is a distinct lack of G-quadruplex stall, implying that the damage/covalent interaction prevents the polymerase from replicating the short ss linker sequence between the dsDNA and the quadruplex. Legacy work has shown that in the presence of dsDNA, both $[Ru(TAP)_3]^{2+}$ and $[Ru(TAP)_2(bpy)]^{2+}$ can form photo-adduct products by oxidation of guanosines upon MLCT irradiation, whereas in the presence of ssDNA both have been shown to effectively induce single-strand breaks.[27,28] More complex or expansive ligand systems such as the heteroleptic *rac*-$[Ru(bpy)_2(dppn)]^{2+}$ or *rac*-$[Ru(TAP)_2(pdppz)]^{2+}$ have exhibited such photocleavage capabilities to dsDNA, through a combination of direct guanine oxidation and $^1O_2$ manufacture as the principal damage mechanisms.[29,30] Evidence here of ss-cleavage is only observed in the system containing the mixed hybrid G-quadruplex fold (100 mM K$^+$) whereas the system containing the all-antiparallel motif

(1 mM K[+]) does not show this. Two antithetical interpretations could potentially explain this difference: 1. that effective cleavage is contingent on major binding interactions occurring and so the mixed hybrid fold, which is more receptive to the binding of the complexes, is more heavily damaged; or 2. that the efficacy of cleavage is dependent on the unfolding of the motif to a single strand to allow non-specific interaction and thus, in this scenario, the mixed hybrid complex has an equilibrium constant that leans more towards single strand formation than folded quadruplex. Considering all evidence, the former seems more likely since the work presented here, and noted elsewhere, shows that the antiparallel motif is more thermodynamically unstable, and is more rapidly unfolded in comparison to the hybrid mixed conformation, in the absence or presence of ligand.[3]

rac-[Ru(bpy)$_2$(dppn)]$^{2+}$

rac-[Ru(TAP)$_2$(pdppz)]$^{2+}$

rac-[Ru(TAP)$_2$(TPAC)]$^{2+}$

mix-[(TAP)$_2$Ru(TPAC)Ru(TAP)$_2$]$^{2+}$
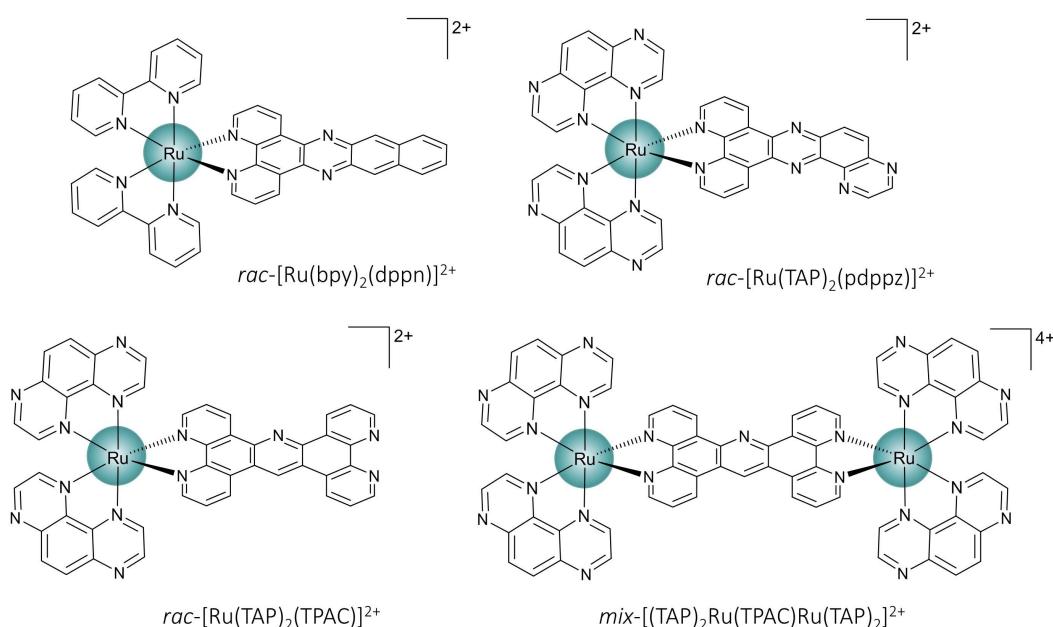
Figure 5.16 – *Small range of ruthenium polypyridyl complexes that have been investigated in respect to their ability to damage replicative processes via photosensitisation. Modes of action include; ROS production, direct guanine oxidation, photoadduct formation, and single strand breakages.*

Despite the conspicuous therapeutic potential of 'trapping' G-quadruplexes through photoadduct or photobridge formation, little published work has explored this property of TAP complexes in the presence of the motif, and so far none have explored how this damage may affect enzymatic processing. One of the few such studies investigated the photoreactivity of the monomer/dimer pair, rac-[Ru(TAP)$_2$(TPAC)]$^{2+}$ and mix-[(TAP)$_2$Ru(TPAC)Ru(TAP)$_2$]$^{4+}$, towards the wtTel24 sequence, d(T$_2$AG$_3$)$_4$ (figure 5.16).[31] Denaturing PAGE analysis following intense [3]MLCT irradiation of the pair separately afforded the generation of complex-DNA adducts whose presence were later confirmed by nano-ESI. The dinuclear variant was proven to be over four times more potent in forming covalent adducts with the G-rich sequence despite the two complexes exhibiting similar reduction potentials. This difference was postulated to be a result of the dinuclear complexes

168

groove binding modes orientating the molecules in perfect distance between the guanosines, as opposed to the monomers tighter binding cavities that may prevent optimal TAP-DNA overlap. Whether this lower general activity of mononuclear complexes imparted by specific binding is simply detrimental to desired therapeutic effect or whether the specificity could impart a greater control in DNA damage is currently unknown.

5.4.3 Effectiveness of the Aqphen ligand

The incorporation of extended ligands, i.e. more expansive π-surface areas than dppz, has been utilised widely in relation to increasing the G-quadruplex specificity and stabilising power of metal complexes. Complexes such as rac-[Ru(bpy)$_2$(icip)]$^{2+}$, rac-[Ru(bpy)$_2$(pdppz)]$^{2+}$, and rac-[Ru(bpy)$_2$(tactp)]$^{2+}$ (figure 5.17) all present as strong inhibitors of telomerase by stabilisation of the wtTel22 mixed hybrid G-quadruplex, and in addition act as potent topoisomerase poisons (human Topo II).[32] Complexes containing the rigid larger ligands, pdppz and tactp, fully quench telomerase action in as low as 100 nM concentrations whilst displaying acute cytotoxicity to HeLA, HepG2, and A549 tumour cell lines (IC$_{50}$ values between 21 – 27 μM). At a 5:1 Ru:DNA the same complexes confer a +3 °C ΔT_m on the telomeric mimic sequence wtTel22 in potassium. Λ-[Ru(phen)$_2$(Aqphen)]$^{2+}$ presented here is especially effective at stabilising the same potassium folded wtTel21 sequence, with the 5:1 Ru:DNA complex exhibiting a ΔT_m of +30 °C, almost 23 °C higher than the dppz parent species and higher than most metal species found in the literature at the same stoichiometry. The asymmetric anthraquinone containing Aqphen ligand was first described photo- and electro-chemically following coordination to rhenium in the species Re(CO)$_3$(Aqphen)Cl and soon after as part of the active light-switch complex rac-[Ru(phen)$_2$(Aqphen)]$^{2+}$, albeit with some focus on DNA binding properties but never as resolved enantiomers, and never in the presence of G-quadruplexes.[10,33–36] FRET melting analysis implies that the Aqphen species', although causing far larger absolute ΔT_ms than the parent species, exhibit less enantiomeric difference in their inferred thermodynamic stability upon binding. Interestingly, this difference in stabilisation between Λ/Δ-[Ru(phen)$_2$(Aqphen)]$^{2+}$ is relatively small for the two telomeric quadruplexes containing anti-parallel character ($|\Delta\Delta T_m|$ = 2.4 and 2.1 °C for K$^+$ mixed hybrid and Na$^+$ anti-parallel respectively), but quite large for the parallel c-MYC promotor G-quadruplex ($|\Delta\Delta T_m|$ = 8.2 °C). Following spectroscopic and crystallographic conclusions from previous chapters it could be postulated that the topologies presenting more accessible 5'-syn-guanosines, i.e. anti-parallel nature, would present as systems exhibiting more enantiomeric disparity. Indeed this is in fact true for the parent complex [Ru(phen)$_2$(dppz)]$^{2+}$ in which lambda stabilises the K$^+$ and Na$^+$ wtTel21 sequence by 6 and 5 °C more than the delta respectively, compared to the 2.5 °C systematic preference for lambda with the all-parallel C-MYC sequence. This may suggest

that the directional propeller loops in the parallel system are better orientated to interact with the phenanthroline ligands in Λ- more so than Δ-[Ru(phen)₂(Aqphen)]²⁺; this may be especially pertinent in this system compared to the parent since the larger ligand is expected to have less translational freedom in its π-stacking with the G-tetrad and therefore may have less ability to align to form such ancillary interactions. Molecular docking simulations of the named complexes to the parallel wtTel22 sequence mirror this sentiment. The spatial arrangement of the Aqphen ligands in the two enantiomers when docked are effectively mirror images in relation to the quadruplex stack, following the geometry of the terminal tetrad in opposite clockwise directions. The lambda isomer ancillary ligands appear a better fit to the left-handed twist of the G-quadruplex, allowing for a more favourable interaction, as opposed to the clashes that would occur if delta were in the same position. It must be said however, that in the analysis of many quadruplex-ligand crystallographic/solution structures, the flexibility of the loop regions to conform to ligand binding is quite evident; as such, analysis of rigid docking studies cannot model the binding of enantiomeric pairs and the ability of the DNA to induce fit.
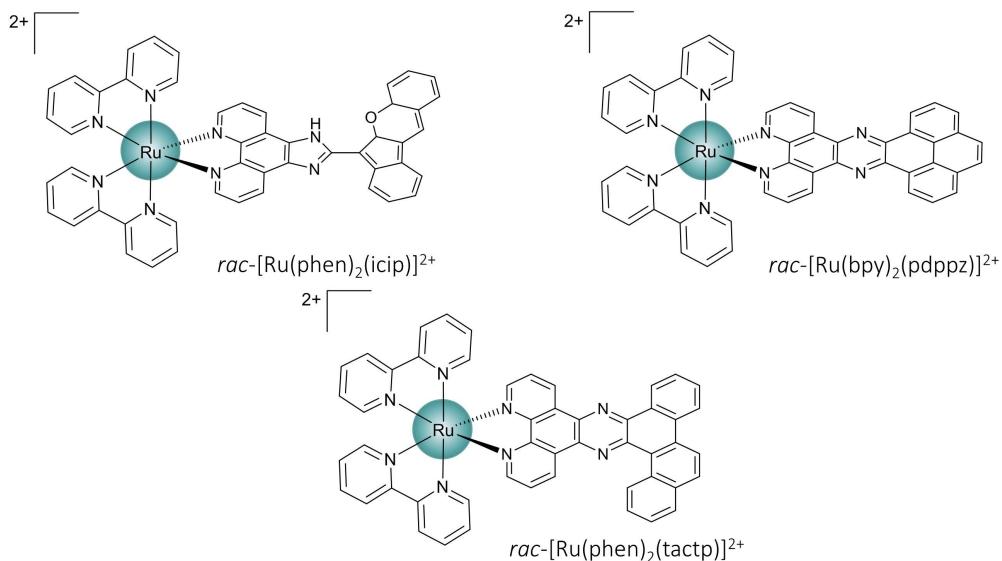
Figure 5.17 – *Three ruthenium polypyridyl complexes, containing intercalating ligands with large π-surfaces, that are capable of strongly inhibiting telomerase action by stabilisation of the wtTel22 G-quadruplex. In addition, all quoted complexes act as topoisomerase poisons.*

The ability of the enantiomers of [Ru(phen)₂(Aqphen)]²⁺ and parent complexes to stabilise G-quadruplexes was investigated *in vivo* using a BG4-conjugated dye antibody pair. The BG4 antibody, which has been shown to bind with near 100 % selectivity to the higher order motif, allows for the visualisation of active G-quadruplexes in the cell. Initial studies utilising the engineered antibody set were the first to present convincing evidence for the biological existence of G-quadruplexes, and furthermore studied the effect on the punctate BG4 staining following incubation with the

G-quadruplex specific ligand, pyridostatin.[11,37] The acyclic compound, which is one of the strongest known stabilisers of telomeric G-quadruplexes (ΔT_m = +35 °C with wtTel21), was shown to actively promote quadruplex formation *in vivo* as evidenced by a 2.9 fold increase in BG4 fluorescence response in the nucleus of MCF-7 breast cancer cells. In comparison, and certainly unexpectedly, incubation with Δ/Λ-[Ru(phen)$_2$(Aqphen)]$^{2+}$ yielded an acute loss of fluorescence signal, whereas in the presence of either parent complex enantiomer, very little change in response was exhibited. The exact reason for this reduction is not absolutely evident but in principle this could be explained by one the three: 1. the presence of ligands causes a systematic destabilisation of G-quadruplexes, 2. relatively high concentrations of binder causes direct irreparable damage and electrostatic overcrowding of nuclear DNA, or 3. the ligand directly obstructs the binding of BG4, either competitively or non-competitively. *In vitro* results presented here certainly show that both Δ/Λ-[Ru(phen)$_2$(Aqphen)]$^{2+}$ exhibit stabilising interactions with G-quadruplexes (if only to the investigated topologies) so it is unlikely that any of the complexes would affect quadruplex formation so unfavourably *in vivo*. In addition, cell viability studies in the presence of the complexes, even at high doses, show that the cells remain completely viable on the timescale of these experiments; if major damage had occurred then the cells would have not proliferated. It is more likely, at least for the Aqphen ligands presented here, that the complexes and BG4 have incompatible binding modes to G-quadruplexes. Similar conclusions were drawn in a recent study where the platinum based self-assembled supramolecular coordination complex, Pt-SCC (figure 5.18), was comprehensively tracked *in cellulo*.[38] In the aforementioned they found, using a systemised co-staining method with BG4, that Pt-SCC disrupted BG4 binding to nucleolar G-quadruplexes, displacing the antibody and subsequently reducing the epitope intensity.
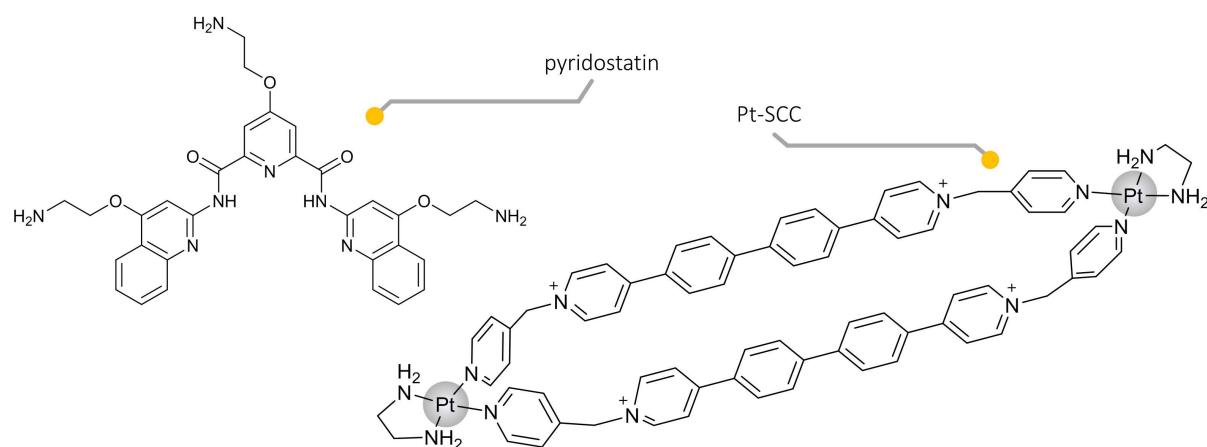

Figure 5.18 – *Pyridostatin and Pt-SCC have both been shown to localise in the nucleus and stabilise G-quadruplexes formed in vivo. Both have been studied upon co-staining with the BG4 antibody set, however, pyridostatin increases epitope intensity, whereas Pt-SCC almost completely destroys the fluorescent response.*

The precise binding mode of the BG4 antibody is currently unknown but its motif specificity is suspected to be a result of the ability of the antibody to recognise and bind to the groove regions of the G-quadruplex. As such a binder with comparable dissociation constant (BG4 $K_d \cong 1$ nm) could outcompete the antibody if its primary binding mode interfered with access to the quadruplex grooves. Figure 5.19 depicts the proposed mechanism of the BG4 immunofluorescence assay following incubation with the stronger binding Aqphen complexes. Unfortunately, due to the competitive nature of the system, no conclusions can be made regarding the stabilisation of putative quadruplexes as was possible with pyridostatin. As mentioned before the antibody does not exhibit any topological specificity towards G-quadruplexes so the fact that very little fluorescence response is noted in the presence of both Δ/Λ-[Ru(phen)$_2$(Aqphen)]$^{2+}$ implies that the complex is also fairly non-specific towards the many folded quadruplexes. More recent advances in G-quadruplex visualisation *in vivo* has generated small molecule probes, such as IMT, GD3 and N-TASQ, and antibodies such as 1H6 and D1, which display more polarity-sensitive topological specificity.[39–43] Such molecules could be utilised as tools to help elucidate further the *in vivo* selectivity for G-quadruplex conformation of this class of complexes.
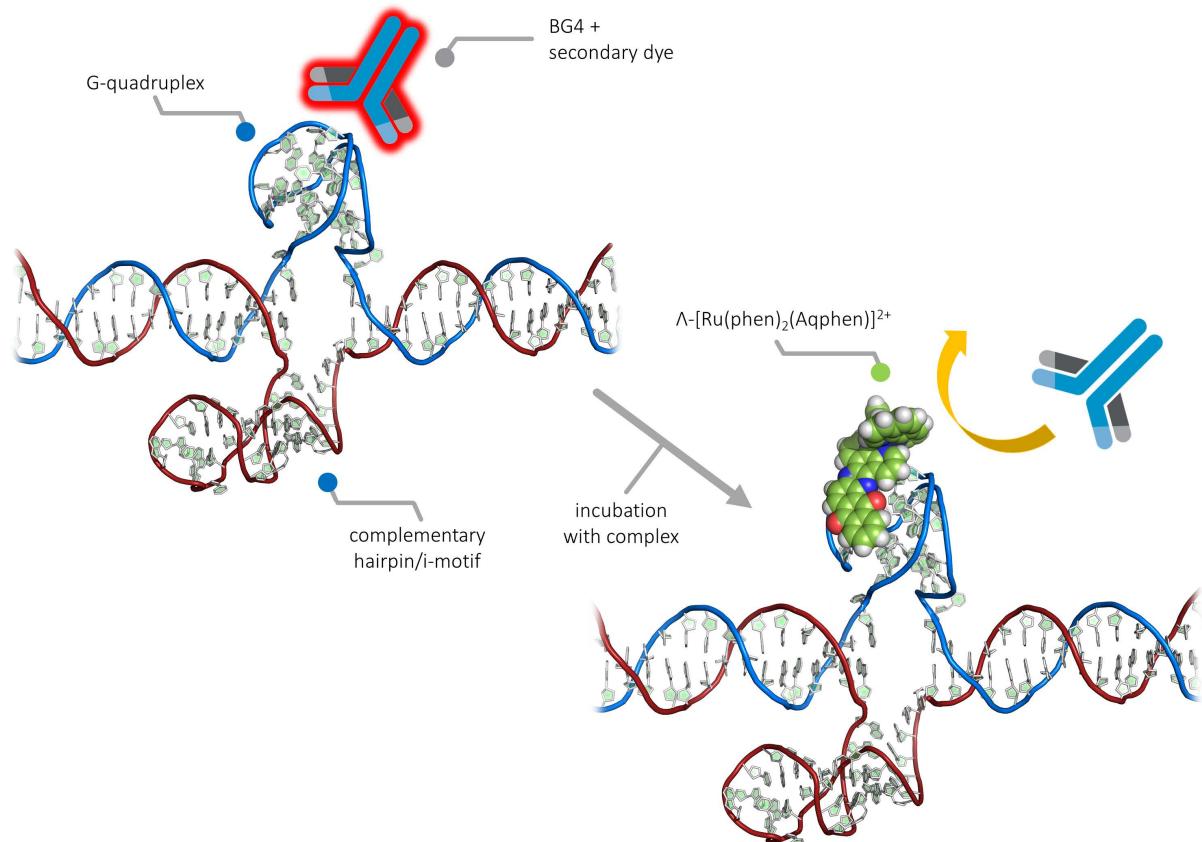

Figure 5.19 – *The BG4 antibody binds to G-quadruplexes in vivo. Incubation with the Λ-[Ru(phen)$_2$(Aqphen)]$^{2+}$ prior to BG4 leads to no fluorescence emission from the antibody set. It is proposed that the complex competitively and preferentially binds to the G-quadruplexes, restricting the antibody from binding.*

5.5 Summary

A number of biochemical assays were performed that probed the ability of a range of enantiopure complexes to stabilise, and stall the replication of, a quadruplex forming DNA. In all measured cases lambda isomers presented as more potent inhibitors than the deltas, supporting previous observations of binding preference and strength, and implying a causation between thermal stability and the replicative processivity of the DNA-ligand complex. In addition to the reversible binding of $[Ru(phen)_2(dppz)]^{2+}$, the isoelectronic counterpart $[Ru(TAP)_2(dppz)]^{2+}$ was examined in irradiated conditions that promote adduct formation and/or photooxidation mechanisms. Replication of these systems is severely stunted under increasing irradiation loads or [Ru], doing so in a dose-dependent logarithmic manner. Synonymous results are observed in the absence of oxygen implying a type-I photoreaction. Signs of adduct formation are noted in the PAGE analysis of these experiments such as the appearance of higher Mw bands that cause a smearing in origin bands. Further work is needed to clarify adduct products and to investigate sequence and/or topological specificity of reactions, and to study the *in vivo* effects of said damage mechanisms.

The G-quadruplex-specific antibody BG4 was utilised to investigate *in vivo* G-quadruplex formations in the presence of the enantiomers of $[Ru(phen)_2(dppz)]^{2+}$ and $[Ru(phen)_2(Aqphen)]^{2+}$. Analysis of the punctate staining of MCF-7 cells following incubation with each complex shows that the Aqphen systems, containing the larger conjugated ligand, almost completely quench the fluorescence response (26 % and 7 % of the original response for Δ/Λ respectively). The complexes, especially $Λ-[Ru(phen)_2(Aqphen)]^{2+}$, were shown to induce impressive thermal stability increases in FRET studies across a range of G-quadruplexes, and furthermore could cause almost complete cessation of the replication of $d(T_2AG_3)_4$ by action of the Klenow Fragment. As such it has been hypothesised that the Aqphen species displaces the BG4 antibody from putative motifs, implying a high *in vivo* binding affinity for nucleolar G-quadruplexes. Subsequent studies should aim to further define the complexes conformational specificity and to incorporate photooxidising capabilities to the Aqphen species as a means to impart photo-mediated damage with greater motif accuracy.

5.6 References

1. Arola, A. and Vilar, R. Stabilisation of G-quadruplex DNA by small molecules. *Curr. Top. Med. Chem* **8**, 1405–1415 (2008).

2. Rhodes, D. & Lipps, H. J. G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Res.* **43**, 8627–8637 (2015).

3. Takahashi, S., Brazier, J. A. & Sugimoto, N. Topological impact of noncanonical DNA structures on Klenow fragment of DNA polymerase. *Proc. Natl. Acad. Sci.* **114**, 9605–9610 (2017).

4. Shay, J. W. & Wright, W. E. Role of telomeres and telomerase in cancer. *Semin. Cancer Biol.* **21**, 349–353 (2011).

5. Brown, W. R. A. *et al.* Structure and polymorphism of human telomere-associated DNA. *Cell* **63**, 119–132 (1990).

6. Esaki, Y., Islam, M. M., Fujii, S., Sato, S. & Takenaka, S. Design of tetraplex specific ligands: cyclic naphthalene diimide. *Chem. Commun.* **50**, 5967–5969 (2014).

7. Imai, M., Tago, Y., Ihara, M., Kawata, M. & Yamamoto, K. Role of the $5' \rightarrow 3'$ exonuclease and Klenow fragment of Escherichia coli DNA polymerase I in base mismatch repair. *Mol. Genet. Genomics* **278**, 211–220 (2007).

8. Liu, J. *et al.* Harnessing ruthenium(II) as photodynamic agents: Encouraging advances in cancer therapy. *Coord. Chem. Rev.* **363**, 17–28 (2018).

9. Jacquet, L., Davies, R. J. H., Kirsch-De Mesmaeker, A. & Kelly, J. M. Photoaddition of [Ru(TAP)$_2$(bpy)]$^{2+}$ to DNA: A new mode of covalent attachment of metal complexes to duplex DNA. *J. Am. Chem. Soc.* **119**, 11763–11768 (1997).

10. Ambroise, A. & Maiya, B. G. Ruthenium(II) complexes of redox-related, modified dipyridophenazine ligands: Synthesis, characterization, and DNA interaction. *Inorg. Chem.* **39**, 4256–4263 (2000).

11. Biffi, G., Tannahill, D., McCafferty, J. & Balasubramanian, S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* **5**, 182–186 (2013).

12. Parkinson, G. N., Lee, M. P. H. & Neidle, S. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* **417**, 876–880 (2002).

13. Park, J. H. *et al.* Enantioselective light switch effect of Δ- and Λ-[Ru(phenanthroline)$_2$ dipyrido[3,2-a:2', 3'-c]phenazine]$^{2+}$ bound to G-quadruplex DNA. *J. Biomol. Struct. Dyn.* **36**, 1948–1957 (2018).

14. Shi, S. *et al.* Interaction of [Ru(bpy)$_2$(dppz)]$^{2+}$ with human telomeric DNA: Preferential binding to G-quadruplexes over i-motif. *Biochimie* **92**, 370–377 (2010).

15. Yu, Q. *et al.* Chiral ruthenium(II) polypyridyl complexes: Stabilization of G-quadruplex DNA, inhibition of telomerase activity and cellular uptake. *PLoS One* **7**, e50902 (2012).

16. Hiort, C., Lincoln, P. & Norden, B. DNA binding of Δ- and Λ-[Ru(phen)$_2$DPPZ]$^{2+}$. *J. Am. Chem. Soc.* **115**, 3448–3454 (1993).

17.   Yu, Q. *et al.* Ruthenium(II) polypyridyl complexes as G-quadruplex inducing and stabilizing ligands in telomeric DNA. *Metallomics* **5**, 222–231 (2013).

18.   Chen, X. *et al.* Targeting telomeric G-quadruplexes with the ruthenium(II) complexes [Ru(bpy)$_2$(ptpn)]$^{2+}$ and [Ru(phen)$_2$(ptpn)]$^{2+}$. *Dalt. Trans.* **42**, 4386–4397 (2013).

19.   Xu, X. *et al.* A hydroxyquinoline-appended ruthenium(II)-polypyridyl complex that induces and stabilizes G-quadruplex DNA. *J. Coord. Chem.* **72**, 201–217 (2019).

20.   Liu, D. *et al.* Polypyridyl complexes of ruthenium(II): Stabilization of G-quadruplex DNA and inhibition of telomerase activity. *Chempluschem* **77**, 551–562 (2012).

21.   Liao, G. *et al.* Novel ruthenium(II) polypyridyl complexes as G-quadruplex stabilisers and telomerase inhibitors. *Dalt. Trans.* **43**, 7811–7819 (2014).

22.   Sun, D. *et al.* Stabilization of G-quadruplex DNA, inhibition of telomerase activity and live cell imaging studies of chiral ruthenium(II) complexes. *Chem. Eur. J.* **18**, 4285–4295 (2012).

23.   Gray, R. D., Trent, J. O. & Chaires, J. B. Folding and unfolding pathways of the human telomeric G-quadruplex. *J. Mol. Biol.* **426**, 1629–1650 (2014).

24.   Hou, X.-M. *et al.* Involvement of G-triplex and G-hairpin in the multi-pathway folding of human telomeric G-quadruplex. *Nucleic Acids Res.* **45**, 11401–11412 (2017).

25.   Ortmans, I., Elias, B., Kelly, J. M., Moucheron, C. & Kirsch-DeMesmaeker, A. [Ru(TAP)$_2$(dppz)]$^{2+}$: a DNA intercalating complex, which luminesces strongly in water and undergoes photo-induced proton-coupled electron transfer with guanosine-5'-monophosphate. *Dalt. Trans.* **4**, 668–676 (2004).

26.   Herman, L., Ghosh, S., Defrancq, E. & Kirsch-DeMesmaeker, A. Ru(II) complexes and light: molecular tools for biomolecules. *J. Phys. Org. Chem.* **21**, 670–681 (2008).

27.   Kelly, J. M., Feeney, M. M., Tossi, A. B., Lecomte, J. P. & Kirsch-DeMesmaeker, A. Interaction of tetra-azaphenanthrene ruthenium complexes with DNA and oligonucleotides. A photophysical and photochemical investigation. *Anticancer. Drug Des.* **5**, 69–75 (1990).

28.   Feeney, M. M., Kelly, J. M., Tossi, A. B., Kirsch-DeMesmaeker, A. & Lecomte, J.-P. Photoaddition of ruthenium(II)-tris-1,4,5,8-tetraazaphenanthrene to DNA and mononucleotides. *J. Photochem. Photobiol. B.* **23**, 69–78 (1994).

29.   Sun, Y., Joyce, L. E., Dickson, N. M. & Turro, C. Efficient DNA photocleavage by [Ru(bpy)$_2$(dppn)]$^{2+}$ with visible light. *Chem. Commun.* **46**, 2426–2428 (2010).

30.   Cloonan, S., Elmes, R. & Erby, M. Detailed biological profiling of a photoactivated and apoptosis inducing pdppz Ruthenium (II) polypyridyl complex in cancer cells. *J. Med. Chem.* **58**, 4494-4505 (2015).

31.   Rickling, S. *et al.* A rigid dinuclear ruthenium(II) complex as an efficient photoactive agent for bridging two guanine bases of a duplex or quadruplex oligonucleotide. *Chem. Eur. J.* **16**, 3951-3961 (2010).

32.   Liao, G. *et al.* Ruthenium(II) polypyridyl complexes as dual inhibitors of telomerase and topoisomerase. *Dalt. Trans.* **44**, 15145–15156 (2015).

33.   Arounaguiri, S., Dattagupta, A. & Maiya, B. G. Redox-activated luminescence and light-induced nuclease activity of a new mixed-ligand ruthenium(II) complex. *Proc. Indian Acad. Sci. - Chem. Sci.* **109**, 155–158 (1997).

34.   Rosa López, B., Bárbara Loeb, L., Boussie, T. & Meyer, T. J. Synthesis of a new phenanthroline derived ligand with acceptor properties. *Tetrahedron Lett.* **37**, 5437–5440 (1996).

35.   López, R. *et al.* Excited-state electron transfer in a chromophore–quencher complex. Spectroscopic identification of a redox-separated state. *Inorg. Chem.* **38**, 2924–2930 (1999).

36.   Whittemore, T. J., White, T. A. & Turro, C. New ligand design provides delocalization and promotes strong absorption throughout the visible region in a Ru(II) complex. *J. Am. Chem. Soc.* **140**, 229–234 (2018).

37.   Biffi, G., Di Antonio, M., Tannahill, D. & Balasubramanian, S. Visualization and selective chemical targeting of RNA G-quadruplex structures in the cytoplasm of human cells. *Nat. Chem.* **6**, 75–80 (2014).

38.   Domarco, O. *et al.* Subcellular duplex DNA and G-quadruplex interaction profiling of a hexagonal PtII metallacycle. *Angew. Chemie Int. Ed.* **58**, 8007–8012 (2019).

39.   Zhang, S. *et al.* Real-time monitoring of DNA G-quadruplexes in living cells with a small-molecule fluorescent probe. *Nucleic Acids Res.* **46**, 7522–7532 (2018).

40.   Laguerre, A. *et al.* Visualization of RNA-quadruplexes in live cells. *J. Am. Chem. Soc.* **137**, 8521-8525 (2015).

41.   Yang, C. *et al.* Visualization of parallel G-quadruplexes in cells with a series of new developed bis(4-aminobenzylidene)acetone derivatives. *ACS omega* **3**, 10487–10492 (2018).

42.   Liu, H.-Y. *et al.* Conformation selective antibody enables genome profiling and leads to discovery of parallel G-quadruplex in human telomeres. *Cell Chem. Biol.* **23**, 1261–1270 (2016).

43.   Henderson, A. *et al.* Detection of G-quadruplex DNA in mammalian cells. *Nucleic Acids Res.* **42**, 860–869 (2013).

6 Summaries, Perspectives, and Future Works

Like looking into a microscope, structural studies, and specifically here, X-ray crystallographic studies, reveal what the eyes (or indeed the light microscope) can't see. The first aim of this thesis was to use this nanoscopic eye to better understand how polypyridyl complexes bind to DNA, more specifically, to develop the current understanding on how derivatisation can affect the intricacies of intercalation. Five different complexes were investigated, in both the absence and presence, of a B-DNA decamer d(TCGGCGCCGA), with an additional structure containing a dicyano complex in the presence of a non-self-complementary B-DNA decamer. The complexes differed by manner of the substituent type and the substitutional pattern. In all of the structures containing complexes with strong electron withdrawing groups ($-NO_2$ or $-C\equiv N$), the terminal T·A base pair forms a canonical base pair that closes the terminal binding cavity in addition to switching the γ angle asymmetry in the backbone. This is in contrast to what is observed in the parent species, or indeed when weaker EWG groups were used, e.g. -Br, -Cl, , -F, -Me, where in these cases the terminal A_{10} flips out to form a reverse Hoogsteen pair with a neighbouring strand.[1,2] This direct observation of stabilisation instigated by the addition of a single group is particularly interesting since in every documented case to date the additional moiety does not directly interact with the bases it is intercalated into. It has therefore been hypothesised that these simple π-withdrawing substituents alter the electronic landscape of the dppz, making it more favourable to stack with the electron-rich guanosines.[3] If such a simple alteration in binder can generate electronic transformations (non-steric) visible in the crystal then perhaps it would be pertinent to engineer next generation binders with calculated electrostatics in mind, instead of design purely based on primary π-surface overlap. Such complexes, along with the added specificity of enantiopure material, could exhibit far greater sequence, or indeed motif, selectivity.

The second aim of this thesis involved providing a structural understanding of ruthenium polypyridyl binding to G-quadruplexes. Two crystallographic structures were products of this work, one of Δ/Λ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ bound to d(TAGGGTTA), and the other of Λ-[Ru(TAP)$_2$(dppz)]$^{2+}$ bound to d(TAGGGTT), yielding the first X-ray data pertaining to this class of compounds with G-quadruplexes. Despite only differing by a single substituent on the complex, and a single base on the sequence, the structures are surprisingly different. In the first case, the nitrile complex switches the conformation of the G-quadruplex from an all-parallel in its native form, to an anti-parallel in the solid state. Although both enantiomers crystallised with the structure at a 4:2 Λ:Δ ratio only the lambda complexes interacted with the core of the G-quadruplex, with the deltas simply stacking upon each other between the biological units. On the contrary the lambda complex is found stacking

directing on the terminal tetrad faces, ancillary interactions between the TAP ligands of the lambda complexes and the ribose sugars of the 5'-guanosines form contacts that could only occur with this enantiomer and as such is hypothesised to be a driving force in the topological switch due to the stabilisation of *syn*-guanosine. In addition, polar contacts between the nitrile groups and 2-NH$_2$ groups on the transverse *syn*-guanosine are apparent, furthering the *syn* conformation stabilisation and thus the 'trapping' of the anti-parallel structure. This topological transformation was also shown to occur in solution at the same 4:1 stoichiometry as is observed in the structure. Interestingly, when the same CD experiment was conducted on the truncated sequence with the nitrile complex in the next chapter, an identical conformational switch was observed, however, when the parent complex is applied the switch is not observed. This implies that the polar contacts are as important in the formation of the anti-parallel fold as the enantiospecific interactions. In addition to elucidating stereospecific and topology dependent binding details, analysis of the structure allows for the interpretation of previous and current luminescence data. The previously reported [Ru(phen)$_2$(11-Br-dppz)]$^{2+}$ and the presented [Ru(phen)$_2$(11-CN-dppz)]$^{2+}$ both exhibit a dramatic specificity in MLCT 'light-switch' response and a particularly high emission when incubated with the intermolecular sequence in comparison to many other G-quadruplex forming sequences.[4,5] Comparing structural data between [Ru(phen)$_2$(11-CN-dppz)]$^{2+}$ bound to duplex, and the isoelectronic [Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ bound to the G-quadruplex, shows that the G-quadruplex bound copies are heavily buried, leading to complete protection from solvent and a thus a much higher quantum yield in response.

The second G-quadruplex structure presented here perfectly encapsulates how unpredictable the crystallisation process can be. Although the structure conserved its native parallel topology, the DNA and indeed the interactions with it, were not typical. Every binding mode observed is a novel mode that has not been witnessed before. The four Λ-Ru complexes that did bind did not do so towards any of the guanines, instead semi-intercalating into parallel stranded-water-bridged mismatched steps, intercalating via the major groove, or end-stacking on T·A/T·A wobble quartets. Such sites are fascinating, and the DNA regions are characteristic of the type of binding pockets available when binding to the loop regions of nuclear G-quadruplexes or i-motifs. Of special contention, is the complex seen to be binding via the major groove. This is the first structural evidence of such an intercalation mode occurring, despite it originally being postulated as the major binding mode of [Ru(phen)$_2$(dppz)]$^{2+}$. In this site the complex is allowed to deeply intercalate as a result of a lack of exocyclic substituents hindering complex insertion.

Despite the slew of interesting binding modes and now future models available to aid in the explanation of solution work, future works should concentrate on the crystallisation of ruthenium complexes with unimolecular sequences of biological context. Of course, although not explicitly discussed in depth here, many efforts were directed towards screening libraries of complexes against putative G-quadruplexes of mainly telomeric origin but none led to useful structural coordinates (although a few close calls were recorded). The biggest hurdles in successfully crystallising such interactions mainly revolve around the observed topological equilibrium of mixed/hybrid sequences, and the inherent flexibility of large loop regions in G-quadruplex DNA. Crystallographic screening should concentrate on sequences with less structural ambiguity and/or with complexes that are structurally specific (that is compounds which drive an equilibrium to one particular quadruplex fold). Another large hurdle in crystallography is the

Utilizing biochemical and *in vivo* immunofluorescence assays, a range of complexes were investigated for their ability to stall the replication of G-quadruplexes by the Klenow fragment. It was observed that Λ-[Ru(phen)$_2$(dppz)]$^{2+}$ is a more effective binder of G-quadruplexes than Δ-[Ru(phen)$_2$(dppz)]$^{2+}$, stalling both anti-parallel and mixed-hybrid structures up to 40 % more than the delta enantiomer. Although these interactions stalled the reactions quite well, they did not cause complete arrest of the replication. To study this further the photooxidising Λ-[Ru(TAP)$_2$(dppz)]$^{2+}$ was used in a similar manner except the DNA-complex solutions were irradiated before introduction of Klenow fragment. It was found that either an appreciable duration of irradiation at 412 nm, or a higher concentration but a lower dose, both caused a complete arrest of replication of the (T$_2$AG$_3$)$_4$ sequence. Damage-mediated arrest has been studied before on a number of different compounds, but never using enantiopure material (not using the Klenow fragment). It was later found that Λ/Δ-[Ru(phen)$_2$(Aqphen)]$^{2+}$ could cause a similar amount of stalling without the need for irradiation of the sample. The Aqphen ligand is much larger than the standard dppz and its asymmetry fits with the geometry of the G-tetrad. Enantiomeric differences were less pronounced, but that was expected due to the relative size of the chiral component in comparison to the achiral. Furthermore, the complex was investigated using an immunofluorescence assay that utilises the BG4 antibody to visualise G-quadruplex formation *in vivo*. It was found that in the absence of complex, or in the presence of Λ/Δ-[Ru(phen)$_2$(dppz)]$^{2+}$, red foci were present that signified the antibodies response to binding to nuclear G-quadruplexes. When the cells were incubated with Λ/Δ-[Ru(phen)$_2$(Aqphen)]$^{2+}$ however, very little response is observed, implying that the complex outcompetes the antibody and binds to the vast majority of the cellular G-quadruplexes.[6] Further investigations need to be conducted to properly assess the binding specificity and *in vivo* localisation of the Aqphen complex, however initial studies imply that the complex is a potent binder of G-quadruplexes, stabilising the potassium folded wtTel22

sequence by +30 °C. In addition, derivatisation of the Aqphen ligand itself and structural analogues of such should be investigated to properly assess what characteristic of the ligand makes it an efficient binder. Of course, X-ray crystallographic studies would be optimal as a way of confirming its major binding modes; screening was conducted on a number of sequences and conditions to no avail.

In conclusion, this body of work incorporates a number of techniques to better understand how ruthenium polypyridyl complexes bind to DNA. It represents new understandings of how these complexes bind to G-quadruplexes, and naturally unveils observations that are not yet understood fully.

References

1   Hall, J. P., Beer, H., Buchner, K., Cardin, D. J. & Cardin, C. J. The structural effect of methyl substitution on the binding of polypyridyl Ru–dppz complexes to DNA. *Organometallics* **34**, 2481–2486 (2015).

2   Hall, J. P., Beer, H., Buchner, K., Cardin, D. J. & Cardin, C. J. Preferred orientation in an angled intercalation site of a chloro-substituted Λ-[Ru(TAP)$_2$(dppz)]$^{2+}$ complex bound to d(TCGGCGCCGA)$_2$. *Philos. Trans. R. Soc. A* **371**, 20120525 (2013).

3   Liu, X.-W. *et al.* Synthesis, characterization, DNA-binding and photocleavage of complexes [Ru(phen)$_2$(6-OH-dppz)]$^{2+}$ and [Ru(phen)$_2$(6-NO$_2$-dppz)]$^{2+}$. *J. Inorg. Biochem.* **99**, 2372–2380 (2005).

4   McQuaid, K. *et al.* Structural studies reveal the enantiospecific recognition of a DNA G-quadruplex by a ruthenium polypyridyl complex. *Angew. Chemie Int. Ed.* **58**, 9881 (2019).

5   Wachter, E., Moyá, D., Parkin, S. & Glazer, E. C. Ruthenium complex 'light switches' that are selective for different G-quadruplex structures. *Chem. Eur. J.* **22**, 550–559 (2016).

6   Biffi, G., Tannahill, D., McCafferty, J. & Balasubramanian, S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* **5**, 182–186 (2013).

7 Experimental

7.1 Experimental Materials

7.1.1 Synthetic reagents and solvents

Unless otherwise stated, all materials and chemicals were sourced from Sigma Aldrich (Merck) or Honeywell research chemicals. Sephedex C-25 anion exchange stationary phase and Dowex 1X2 Chloride form anion exchange resin were purchased from GE Healthcare. All solvents, unless stated in the experimental, were obtained at HPLC grade and used without further purification. Where further purification was needed, protocol from "Purification of Laboratory Chemicals, 4th edition, Armarego *et. al.*" was followed. Deuterated solvents for NMR analysis were purchased either through Sigma-Aldrich or Cambridge Isotope Laboratories.

7.1.2 Oligonucleotide synthesis and manipulations

Unless otherwise stated, all oligonucleotides were purchased as triple HPLC purified syntheses from Eurogentec Ltd or ATDBio Ltd and used without further purification. Stock solutions were prepared using the stated synthetic yields by dissolving the solids in HPLC grade water, vortexing, and centrifuging at 11,000 rpm for 10 mins to remove any insoluble material. DNA concentrations were examined spectrophotometrically; where all samples were denatured at 90°C for 5 mins. Molar absorption coefficients (ε) for the oligonucleotide solutions at 260 nm were calculated using the nearest neighbour method assuming that any secondary structure is fully dissociated (i.e. single stranded). CT-DNA and polynucleotide concentrations are given in units of base pairs ([bp]) and were determined using the molar absorption coefficient (ε_{260}) per nucleotide of 6600 M^{-1}. Unless otherwise stated, all DNA containing experimental samples were annealed prior to use by heating the sample tubes in a 90°C water bath for 15 minutes and allowed to cool to room temperature slowly.

7.1.3 Buffer solutions

All working solutions utilized a buffering salt, where unless stated otherwise, was aqueous sodium or potassium cacodylate, buffered at either pH 7.1 (duplexes and G-quadruplexes) or pH 5.1 (i-motifs) against HCl. Cacodylate buffer solutions were prepared by titrating 500 mM aqueous HCl into a 500 mM solution of the appropriate cacodylate until desired pH is reached, then diluting to stock standard concentration. Additionally, G-quadruplex forming sequences were measured in the presence of differing metal halide salts (NaCl/KCl); where the fluoride alternative (NaF/KF) was used to

reduce lower wavelength absorption in experiments where high PMT voltages clouded higher energy absorption measurements.

7.1.4 Oligonucleotide solutions

All oligonucleotide solutions were prepared using HPLC grade water using dried HPLC-purified solids purchased from Eurogentec Ltd. Stock solutions were concentration checked spectrophotometrically at 95°C using calculated extinction coefficients that utilise the nearest neighbour model. All DNA solutions were confirmed to be protein free by assessing the A_{260}/A_{280} ratio (i.e. around 1.8).

7.1.5 Metal complex solutions

Following dissolution in appropriate solvent; any metal complex solutions were vortexed and filtered through a 0.45 μm PTFE filter, diluted to stock concentration following spectrophotometric determination of concentration using either known or experimentally determined extinction coefficients, and then used fresh. Typically, most experimental solutions were of sufficiently low concentration so corrections for re-adsorption could be avoided. All manipulations of complex solutions were conducted in low light environments and kept in the dark (at -4 °C) when not in use.

7.2 Experimental Techniques and Instrumentation

7.2.1 Nuclear magnetic resonance spectroscopy (NMR)

Unless otherwise stated, all ^{1}H NMR spectra were collected on a Bruker Nanobay 400 MHz instrument, whereas conversely, the majority of ^{13}C NMR spectra were collected on a Bruker DPX 400.1 MHz machine operating at 100.1 MHz. Both machines were calibrated against a tetramethylsilane (TMS) internal standard and have two channels running TOPSPIN 2.4 and ICON NMR 4.2. All J-coupling constants were reported following normalisation against the applied Larmor frequency. Data Manipulations (processing/analysing) and graphical productions were carried out using MestReNova 11.

7.2.2 Infrared spectroscopy

All infrared measurements were collected on a Perkin-Elmer Spectrum 100 FTIR spectrometer fitted with an ATR crystal working platform. Total reflectance was measured on hand ground, vacuum dried solids, and was baseline corrected against atmospheric vapour compensation.

7.2.3 Mass spectrometry

High resolution ESI mass spectra were recorded on a Thermo Scientific LTQ Orbitrap XL running in positive ion mode following elution through a Thermo Hypersil Gold column on an Accela HPLC system. Fragmented Ions were detected on an Orbitrap Ion trap photodiode array detector and

were determined via peak matching against the internally calibrated lock mass for Diisooctyl phthalate (m/z = 413.2662). Data analysis was performed on the Xcalibur Qual Browser software package and all masses are reported within 3 ppm.

7.2.4 High performance liquid chromatography (HPLC) and isomeric resolution

High pressure liquid chromatography (HPLC) analysis and preparative scale enantiomeric separation was performed on a Hitachi Primaide HPLC machine fitted with a CF6 LARIHC cyclofructan based chiral column (internal dimensions; 10 x 250 mm) supplied by AZYP separations; LLC (Arlington, Texas). Baseline separation was achieved using a range of different mobile phases where specificity was largely different for disparate complex systems. Analysis was performed at a flow rate of 5 mL min^{-1} where each preparative injection was 200 µL in volume with a complex concentration of ~3 mg mL^{-1}. Eluent fractions were collected in 2 mL aliquots in case of peak overlap as a result of column overloading; then combined after being peak matched. Due to the small volume of the column, many separations (> 6 runs) were carried out to allow for fluent work-up. Following dialysis, the fractions were combined into 15 mL centrifugal tubes and to each, 2 mL of a 200 mM aqueous solution of potassium hexafluorophosphate was added. The solution was then reduced under pressure using a Genevac miVac DNA concentrator at 40 $^{\circ}$C for 24 hours to remove the organic eluent. After ensuring metathesis had fully concluded, the precipitate was collected via in vacuo filtration and washed with fractions of HPLC-grade water (5 x 5 mL). The water-soluble chloride salts were prepared by dissolving the enantiomers in 60:40 H$_2$O:ACN and subsequently stirring the solution with suspended Amberlite IRA-400 anion exchange resin (Cl$^-$ form) overnight. Complex purity was verified by electrospray mass spectrometry (ES-MS); optical purity was first confirmed by analytical HPLC to quantify the quality of the enantiomeric excess of each work up; in each case EE was >99 % by peak area. Circular dichroism of the aqueous solutions confirmed the success of the separations. Mobile phases used are quoted with the resolved data.

7.2.5 Ultraviolet-Visible absorption spectroscopy

UV-Visible spectroscopic data, including any scans or melting plots, were obtained on Agilent Technologies Cary 100/300 UV-Vis spectrophotometers fitted with a Peltier thermostatted temperature controller, unless stated otherwise. Deuterium (UV) and halogen (Vis) lamps were used as light sources. Quartz cuvettes of path length 1 cm were used throughout any characterization; however, 0.2 mm path length cells were used for high concentration studies using the tetramolecular DNA. Data manipulations and graphical representations were usually conducted in the OriginPro 9.1 software package where, when required, the data was smoothed using a 5 point least-squares Savitzky-Golay filter.

7.2.6 Spectrofluorometry

Emission spectra were measured on a Cary Eclipse fluorescence spectrophotometer fitted with a Peltier thermostatted temperature controller. Scans were measured in triplicate and averaged on samples containing no more than 15 µM complex, in 1 cm path length quartz cells using a 5 nm spectral window for excitation and emission. All samples were pre annealed before collection.

7.2.7 Circular dichroism

CD measurements were recorded on an Applied Photophysics Chirascan V100 spectropolarimeter using sample concentrations with a maximum optical density of 1.0, in quartz cuvettes of varying path length (100 µm to 1 cm). Spectra were typically averaged over 5 cycles and baseline corrected against cell/buffer.

7.2.8 Synchrotron radiation circular dichroism

All SRCD measurements were collected on beamline B23 at Diamond Light Source Ltd. using either 100/200 µm path length quartz cells or using high-throughput quartz 96-well plates with 100/1000 µm path length. Samples were prepared using the fluoride salts instead of chloride salts to reduce high energy absorptions. All samples were mixed, heated to 90 °C and then allowed to cool slowly to room temperature. All spectra were acquired using a 1 second integration time per nm, with a 1 nm slit, between 180-350 nm and then were cut according to an appropriate PMT voltage. This results in a cut-off at approximately 192-200 nm depending on system. Plots have been both background and offset-corrected using the CDApps program suite.

7.3 Synthesis

7.3.1 Ligands

7.3.1.1 1,10-phenanthroline-5,6-dione (phendione)
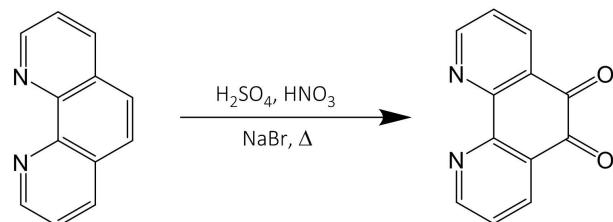
1,10-phenanthroline (phen) (5.0 g, 28 mmol) was stirred, until full dissolution, in anhydrous conc. sulphuric acid (30 mL). To the stirring solution, sodium bromide (NaBr) (2.5 g, 24 mmol) was added, immediately followed by an addition of a ~70 % nitric acid solution (15 mL). The mixture was then heated under reflux conditions for 60 mins. Subsequently, the condenser was removed, whilst continuing to heat, for approximately 15-30 minutes, or until all residual bromine gas had evacuated. The solution was allowed to cool, and then poured directly into a large beaker of analytical grade water (400 mL), cooled on ice, carefully neutralized to pH 7 using a solution of sodium hydroxide

(10M, ~150 mL), and then heated to 80°C. The hot solution was filtered through celite to remove the small amount of brown oil, cooled and then the compound was extracted using dichloromethane (6 x 100 mL). Following this, the combined extracts were washed with water (50 mL), dried over magnesium sulphate, and then evaporated *in vacuo*. The crude compound was further purified by recrystallization from hot ethanol to yield the product as a bright yellow microcrystalline solid (2.51 g, 11.8 mmol, 42 %).

$δ_H$ **(400 MHz, TMS, DMSO-d_6)** – 9.00 (dd, J = 4.8, 1.8 Hz, 2H), 8.40 (dd, J = 7.6, 1.8 Hz, 2H), and 7.68 ppm (dd, J = 7.6, 1.8Hz, 2H). *See figure A7.1.*

$δ_C$ **(101 MHz, DMSO-d_6)** – 117.69, 154.28, 152.21, 135.63, 129.03, and 125.18 ppm.

HRMS-ESI (m/z) - Found ([M+H]$^+$, 211.0508); calc. 211.0502 ($C_{12}N_2O_2H_7^+$)

7.3.1.2 2-cyano-1,10-phenanthroline (2-CN-phen)
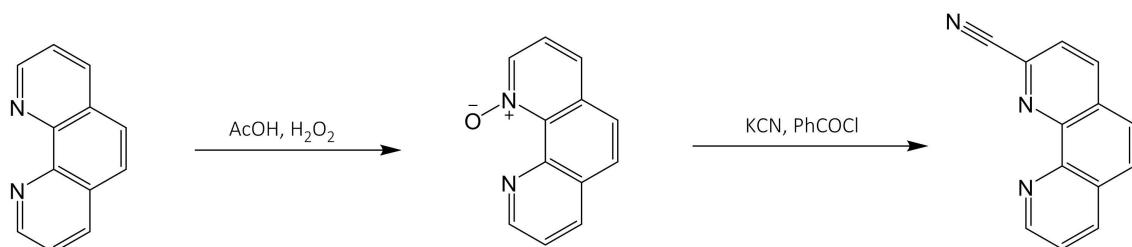
7.3.1.2.1 1,10-phenanthroline-1-oxide

1,10-phenanthroline monohydrate (5 g, 25 mmol) was stirred until full dissolution in glacial acetic acid (15 mL). To this, a solution of hydrogen peroxide (30 %, 30 mL) was added dropwise, ensuring the temperature never exceeded 80 °C. The reaction mixture was maintained at 70-75 °C for 3 hours before an additional portion of hydrogen peroxide (30 %, 30 mL) was added dropwise and the heating was continued; again ensuring the temperature did not exceed 80 °C. After the vessel was allowed to cool, the mixture was neutralised to pH ≈ 10 using a saturated aqueous solution of potassium hydroxide; followed by repeated extraction with chloroform. The combined fractions were dried over anhydrous magnesium sulphate and evaporated to yield the intermediate product as a yellow solid (4.12 g, 21 mmol, 85 %). Mp: 178-181 °C (Lit. 176 – 179 °C)

$δ_H$ **(400 MHz, TMS, CDCl$_3$-d)** – 9.32 (dd, J = 4.4, 1.9 Hz, 1H), 8.75 (dd, J = 6.3, 1.2 Hz, 1H), 8.24 (dd, J = 8.0, 1.9 Hz, 1H), 7.81 (d, J = 8.8 Hz, 1H), 7.75 (dd, J = 8.7, 1.9 Hz, 2H), 7.67 (dd, J = 8.1, 4.4 Hz, 1H), and 7.47 ppm (dd, J = 8.1, 6.3 Hz, 1H). *See figure A7.2.*

### *7.3.1.2.2 2-cyano-1,10-phenanthroline (2-CN-phen)*

1,10-phenanthroline-1-oxide (2.5 g, 13 mmol) was added, with stirring, to an aqueous solution (20 mL) of potassium cyanide (25.0 g). To this mixture, under heavy stirring, benzoyl chloride (2.5 mL) was introduced in a dropwise fashion and the mixture was stirred was an additional 3 hours. The resulting precipitate was collected via suction filtration, washed with water, and dried *in vacuo* to yield the target product as a white solid (0.825 g, 4 mmol, 31 %). Mp: 230-234 °C (Lit. 233 – 234 °C).

$δ_H$ **(400 MHz, TMS, DMSO-d_6)** – 9.24 (dd, J = 4.6, 1.7 Hz, 1H), 8.92-8.78 (m, 2H), 8.41 (d, J = 8.2 Hz, 1H), 8.30 (d, J = 8.9 Hz, 1H), 8.22 (d, J = 8.9 Hz, 1H), and 8.06 ppm (dd, J = 8.2, 4.7 Hz, 1H). *See figure A7.3.*

δ_C (101 MHz, DMSO-d_6) – 150.15, 138.30, 137.57, 132.21, 130.02, 129.77, 129.19, 129.11, 128.51, 126.69, 126.31, 124.55, and 117.74 ppm.

HRMS-ESI (m/z) – Found ([M+H]$^+$, 206.0717); calc. 206.0713 ($C_{13}N_3H_8^+$).

7.3.1.3 1,4,5,8-tetraazaphenanthrene (TAP)
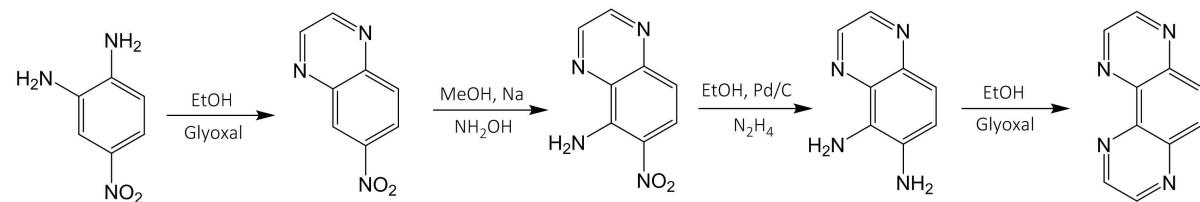
7.3.1.3.1 6-nitroquinoxaline

4-nitrobenzene-1,2-diamine (4.0 g, 27.95 mmol) was suspended, with stirring, in ethanol (100 mL). To the suspension, an aqueous solution of glyoxal (40 % w/w, 6 mL) was added dropwise and the resulting mixture was refluxed (78 °C) for 2 hours. After removing the solvent under reduced pressure, the solid was dissolved in water (200 mL) and the compound was extracted with dichloromethane (DCM) until the organic phase was near colourless (4 x 50 mL). The organic phases were combined and dried over MgSO$_4$ before the solvent was removed under vacuum. The resulting solid was recrystallized twice from propan-2-ol to yield the product as orange crystalline needles (4.11 g, 23.49 mmol, 84 %).

δ_H (400 MHz, TMS, DMSO-d_6) – 9.22 (s, 2H), 8.98 (d, J = 2.5 Hz, 1H), 8.63 (dd, J = 9.2, 2.6 Hz, 1H), and 8.41 ppm (d, J = 9.2 Hz, 1H). *See figure A7.4.*

HRMS-ESI (m/z) – Found ([M+H]$^+$, 176.0455); calc. 176.0455 ($C_8N_3O_2H_6^+$).

7.3.1.3.2 6-nitroquinoxaline-5-amine

Methanol was dried and distilled as per published standards. To this dried methanol (125 mL), freshly cut sodium metal (2.3 g, 100mmol, 3.3 eqv.) was added in a stepwise fashion, waiting for reaction to finish before each addition. Separately, to cooled distilled methanol (50 mL, 0 °C), hydroxylamine hydrochloride (3.13 g, 45 mmol, 1.5 eqv.) was added. The two solutions were slowly combined and the resulting sodium chloride precipitate was allowed to settle before decanting off the solution. The supernatant was added to a well stirred suspension of 6-nitroquinoxaline (5.25 g, 30 mmol, 1 eqv.) in boiling distilled methanol (250 mL). The brown mixture was refluxed for 90 mins before being cooled on ice (0 °C). The resulting yellow precipitate was collected via vacuum filtration and recrystallized from an acetic acid:water mixture (3:1, ~100 mL) yielding bright yellow crystalline needles (3.54 g, 18.6 mmol, 62 %).

δ_H (400 MHz, TMS, DMSO-d_6) – 9.16 (d, J = 1.9 Hz, 1H), 9.00 (d, J = 1.9 Hz, 1H), 8.59 (s, 2H, NH$_2$), 8.36 (d, J = 9.7 Hz, 1H), and 7.25 ppm (d, J = 9.7 Hz, 1H). *See figure A7.5.*

δ_C (101 MHz, DMSO-d_6) – 148.78, 143.15, 126.04, and 114.28 ppm. (4 strong signals).

HRMS-ESI (m/z) – Found ([M+H]$^+$, 191.0569); calc. 191.0564 ($C_8N_4O_2H_7^+$).

7.3.1.3.3 quinoxaline-5,6-diamine

10 % Pd/C (0.2 g) was added to a suspension of 6-nitroquinoxalin-5-amine (1.5 g, 7.86 mmol, 1 eqv.) in ethanol (80 mL) and the mixture was refluxed for 60 mins (78 ^{o}C). To the refluxing mixture, fresh hydrazine monohydrate (50-60 %, 14 mL, >150 mmol, >20 eqv.) was added and the mixture was allowed to reflux for an additional 60 mins. The dark red suspension was filtered hot through a pad of celite and was subsequently washed with dichloromethane (4 x 20 mL).The solvent was removed under reduced pressure and finally dried *in vacuo* to yield a dark red crystalline solid (1.16 g, 7.23 mmol, 92 %)

δ_H (400 MHz, TMS, DMSO-d_6) – 8.59 (d, *J* = 1.9 Hz, 1H), 8.51 (d, *J* = 1.9 Hz, 1H), 7.27 (d, *J* = 8.8 Hz, 1H), 7.20 (d, *J* = 8.8 Hz, 1H), 5.28 (s, 2H, N*H*$_2$), and 5.15 ppm (s, 2H, N*H*$_2$). *See figure A7.6.*

δ_C (101 MHz, DMSO-d_6) – 141.90, 140.08, 136.83, 133.01, 132.38, 125.99, 121.82, and 116.32 ppm.

HRMS-ESI (m/z) – Found ([M+H]$^+$, 161.0825); calc. 161.0822 (C$_8$N$_4$H$_9^+$).

### 7.3.1.3.4 1,4,5,8-tetraazaphenanthrene (TAP)

To a suspension of quinoxaline-5,6-diamine (1 g, 6.24 mmol, 1 eqv.) in ethanol (100 mL), an aqueous solution of oxalaldehyde (40 % w/w, 5 mL) was added dropwise and the subsequent mixture was allowed to reflux for 120 mins. Following this, the solvent was removed under reduced pressure and the solid was redissolved in water (100 mL) before the compound was extracted from the aqueous layer using dichloromethane (5 x 40 mL). The organic phase was dried over MgSO$_4$ before the solvent was removed under reduced pressure. The solid was recrystallized from propan-2-ol to yield the product as dark yellow crystalline needles (285.73 mg, 1.57 mmol, 25 %).

δ_H (400 MHz, TMS, DMSO-d_6) – 9.30 (d, *J* = 2.0 Hz, 2H), 9.27 (d, *J* = 2.0 Hz, 2H), and 8.44 ppm (s, 2H). *See figure A7.7.*

δ_C (101 MHz, DMSO-d_6) – 146.84, 145.61, 143.42, 140.15, and 131.47 ppm

HRMS-ESI (m/z) – Found ([M+H]$^+$, 183.0665); calc. 183.0665 (C$_{10}$N$_4$H$_7^+$).

7.3.1.4 dipyrido[3,2-a:2',3'-c]phenazine (dppz)
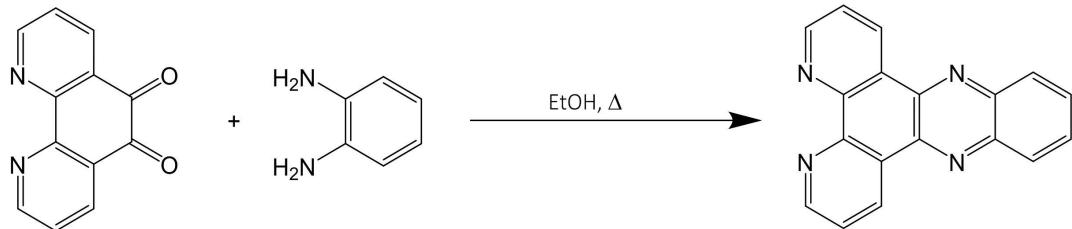
As described in the literature, dipyridophenazine was synthesised via an acid-catalysed condensation of phendione and 1,2-phenylenediamine. A well-mixed ethanolic solution (15 mL) of phendione (0.515 g, 2.45 mmol) was added slowly to an ethanolic solution (10 mL) of phenylenediamine (0.53 g, 4.91 mmol) with a trace amount of *ρ*-toluene sulfonic acid. The suspension was refluxed for 3 hours before removing the condenser and allowing approximately half of the ethanol to evaporate. The remaining suspension was cooled to room temp and a brown precipitate was observed. The solid was collected via suction filtration, washed with cold ethanol (2 x 5 mL) and

recrystallized from aqueous ethanol (1:1) to yield the target product as ochre needle-like crystals (0.569 g, 1.93 mmol, 82 %).

δ_H (400 MHz, TMS, CDCl₃-*d*) – 9.68 (dd, *J* = 8.1, 1.8 Hz, 2H), 9.29 (dd, *J* = 4.5, 1.8 Hz, 2H), 8.39 (dd, *J* = 6.5, 3.4 Hz, 2H), 7.94 (dd, *J* = 6.5, 3.4 Hz, 2H), and 7.82 ppm (dd, *J* = 8.1, 4.5 Hz, 2H). *See figure A7.8.*

δ_C (101 MHz, CDCl₃-*d*) – 152.58, 148.44, 142.51, 141.17, 133.79, 130.68, 129.57, 127.61, and 124.16 ppm.

HRMS-ESI (m/z) – Found ([M+H]⁺, 283.0979); calc. 282.0978 ($C_{18}N_4H_{11}^+$).

7.3.1.5 11-cyano-dipyrido[3,2-a:2',3'-c]phenazine (11-CN-dppz)
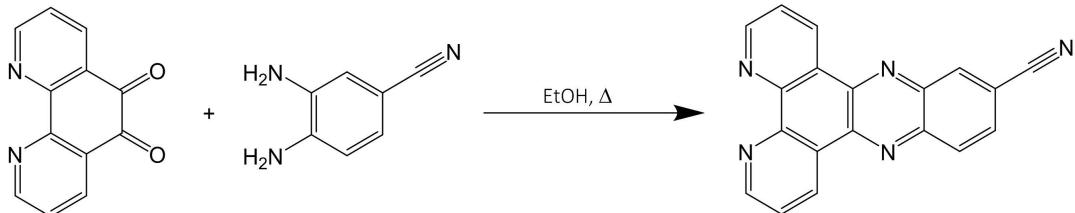
Phendione (0.515 g, 2.45 mmol) was dissolved in dry ethanol (15 mL) with stirring to give a pale yellow solution. Separately, 3,4-diaminobenzonitrile (0.325 g, 2.5 mmol) was dissolved in ethanol (10 mL) and a trace amount of *ρ*-toluene sulfonic acid was added, yielding a dark brown/red solution. The two coloured solutions were combined, slowly, with stirring, and then refluxed (78 °C) for 4 hours. Following this, the condenser was removed and approximately half of the ethanol was allowed to evaporate before allowing the solution to cool to r.t. The insoluble product was collected by suction filtration before being washed with water (3 x 5 mL) and cold ethanol (3 x 5 mL), yielding the target molecule as a light brown/orange microcrystalline solid (685 mg, 2.10 mmol, 86 %).

δ_H (400 MHz, TMS, CDCl₃-*d*) – 9.58 (td, *J* = 6.4, 3.6, 1.6 Hz, 2H), 9.32 (dd, *J* = 2.0, 2.0 Hz, 2H), 8.75 (d, *J* = 1.6 Hz, 1H), 8.45 (d, *J* = 8.4 Hz, 1H), 8.03 (dd, *J* = 1.6, 8.8 Hz, 1H), and 7.83 ppm (dd, *J* = 4.4, 8.0 Hz, 2H). *See figure A7.9.*

δ_C (101 MHz, CDCl₃-*d*) – 153.5, 148.9, 143.3, 142.8, 141.1, 135.8, 134.2, 131.2, 130.7, 127.0, 124.5, 118.0, and 113.9 ppm.

ν_{max}/cm⁻¹ - 2998 (broad, w, Arom. C-H stretch) and 2227 (m, Nitrile -C≡N stretch).

HRMS-ESI (m/z) – Found ([M+H]⁺, 308.0932); calc. 308.0931 ($C_{19}N_5H_{10}^+$).

7.3.1.6 11,12-cyano-dipyrido[3,2-a:2'3'-c]phenazine (11,12-CN-dppz)
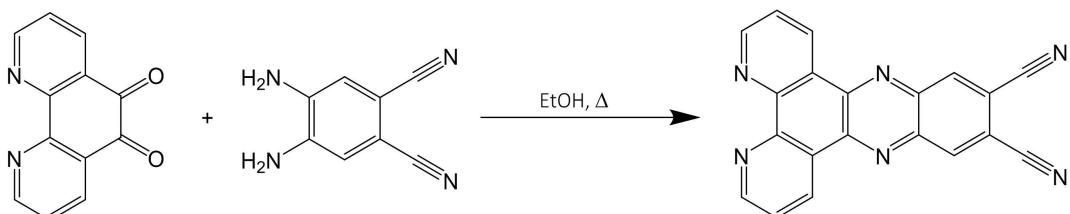
Following similarly to the synthesis of 11-CN-dppz; 11,12-CN-dppz was condensed using the same methodology, however, 4,5-nitrile-1,2-diaminobenzene (0.465 g, 2.5 mmol) was used instead of

3,4-diaminobenzonitrile, yielding the target compound as a beige/light brown solid (702 mg, 2.11 mmol, 87 %).

δ$_H$ (400 MHz, TMS, CDCl$_3$-*d*) – 9.63 (dd, *J* = 8.1, 1.8 Hz, 2H), 9.37 (dd, *J* = 4.4, 1.8 Hz, 2H), 8.90 (s, 2H), and 7.89 ppm (dd, *J* = 8.1, 4.5 Hz). *See figure A7.10.*

δ$_C$ (101 MHz, CDCl$_3$-*d*) – 154.4, 144.8, 142.2, 137.6, 134.7, 126.5, 124.9, 114.9, and 114.4 ppm.

v$_{max}$/cm^{-1} – 3000 (broad, m, Arom. -C-H stretch) and 2239 (m, Nitrile -C≡N stretch).

HRMS-ESI (m/z) – Found ([M+H]$^+$, 333.0884); calc. 333.0883 (C$_{20}$N$_6$H$_9$$^+$).

7.3.1.7 10, 12-dimethyl-dipyrido[3,2-a:2'3'-c]phenazine (10, 12-Me-dppz)

Following similarly to the synthesis of 11-CN-dppz; 10,12-Me-dppz was condensed using the same methodology, however, 1,2-diamino-3,5-dimethylbenzene (341 mg, 2.5 mmol) was used instead of 3,4-diaminobenzonitrile, yielding the target compound as a beige/light brown solid (689 mg, 2.22 mmol, 89 %).

δ$_H$ (400 MHz, TMS, CDCl$_3$-*d*) – 9.65 (ddd, *J* = 9.9, 8.1, 1.8 Hz, 2H), 9.26 (ddd, *J* = 5.0, 3.4, 1.8 Hz, 2H), 7.96 (s, 1H), 7.79 (dd, *J* = 8.1, 4.4 Hz, 2H), 7.60 (s, 1H), 2.99 (s, 3H, -C*H$_3$*), and 2.66 ppm (s, 3H, -C*H$_3$*). *See figure A7.11.*

δ$_C$ (101 MHz, CDCl$_3$-*d*) – 152.29, 152.12, 148,33, 148.10, 142.88, 141.29, 140.55, 140.45, 139.06, 137.29, 133.61, 133.49, 133.01, 128.03, 127.70, 125.88, 124.02, 22.21 and 17.28 ppm.

HRMS-ESI (m/z) – Found ([M+H]$^+$, 311.1289); calc. 311.1291 (C$_{20}$N$_4$H$_{15}$$^+$)
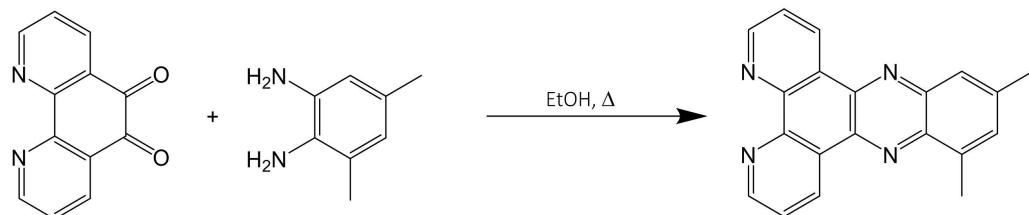
7.3.1.8 11-bromo-dipyrido[3,2-a:2'3'-c]phenazine (11-Br-dppz)
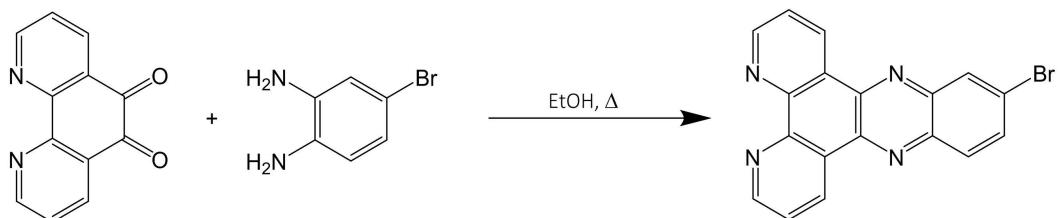
Following similarly to the synthesis of 11-CN-dppz; 11-Br-dppz was condensed using the same methodology, however, 4-bromo-1,2-diaminobenzene (0.465 g, 2.5 mmol) was used instead of 3,4-Diaminobenzonitrile, yielding the target compound as a beige/light brown solid (753 mg, 2.01 mmol, 83 %).

δ_H (400 MHz, TMS, CDCl$_3$-d) − 9.62 (ddd, J = 8.2, 4.2, 1.8 Hz, 2H), 9.29 (d, J = 3.7 Hz, 2H), 8.57 (d, J = 2.2 Hz, 1H), 8.24 (d, J = 9.0 Hz, 1H), 7.99 (dd, J = 9.0, 2.2 Hz, 1H), and 7.82 ppm (ddd, J = 8.2, 4.5, 2.2 Hz, 2H). *See figure A7.12.*

HRMS-ESI (m/z) − Found ([M+H]$^+$, 361.0086); calc. 361.0083 (C$_{18}$N$_4$H$_{10}$Br$^+$).

7.3.1.9 10-nitro-dipyrido[3,2-a:2'3'-c]phenazine (10-NO$_2$-dppz)
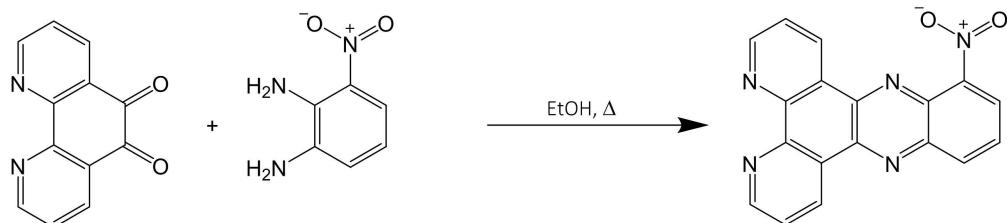
Following similarly to the synthesis of 11-CN-dppz; 10-NO$_2$-dppz was condensed using the same methodology, however, 3-nitro-1,2-phenylenediamine (383 mg, 2.5 mmol) was used instead of 3,4-Diaminobenzonitrile, yielding the target compound as a beige/light brown solid (731 mg, 2.23 mmol, 89 %).

δ_H (400 MHz, TMS, CDCl$_3$-d) − 9.60 (ddd, J = 13.7, 8.1, 1.8 Hz, 2H), 9.32 (td, J = 4.5, 1.8 Hz, 2H), 8.60 (dd, J = 8.6, 1.3 Hz, 1H), 8.35 (dd, J = 7.4, 1.3 Hz, 1H), 8.00 (dd, J = 8.6, 7.5 Hz, 1H), and 7.83 ppm (ddd, J = 8.1, 4.4, 2.8 Hz, 2H). *See figure A7.13.*

δ_C (101 MHz, CDCl$_3$-d) − 153.62, 153.50, 148.96, 148.89, 147.18, 142.72, 142.53, 141.92, 134.64, 134.16, 134.04, 128.57, 126.87, 126.76, 125.21, 124.63, and 124.42 ppm.

v_{max}/cm^{-1} − 3049 (broad, m, Arom. -C-H stretch), 1518 and 1357 (s, Nitro −N=O stretches).

HRMS-ESI (m/z) − Found ([M+H]$^+$, 328.0829); calc. 328.0829 (C$_{18}$N$_5$O$_2$H$_{10}$$^+$)

7.3.1.10 11-nitro-dipyrido[3,2-a:2'3'-c]phenazine (11-NO$_2$-dppz)

Following similarly to the synthesis of 11-CN-dppz, 11-NO$_2$-dppz was condensed using the same methodology, however, 4-nitro-1,2-phenylenediamine (383 mg, 2.5 mmol) was used instead of 3,4-diaminobenzonitrile, yielding the target compound as a beige/light brown solid (702 mg, 2.15 mmol, 86 %).

δ_H (400 MHz, TMS, CDCl$_3$-d) − 9.60 (ddd, J = 2.0, 4.0, 6.0 Hz, 2H), 9.29 (d, J = 9.0 Hz, 2H), 8.57 (d, J = 2.4 Hz, 1H), 8.24 (d, J = 9.2 Hz, 1H), 7.99 (dd, J = 2, 9.2 Hz, 1H), and 7.82 ppm (ddd, J = 2.4, 4.4, 7.2 Hz, 2H). *See figure A7.14.*

δ_C (101 MHz, CDCl$_3$-d) − 153.65, 153.46, 148.98, 148.63, 148.11, 144.09, 143.38, 142.96, 140.74, 134.23, 133.96, 131.17, 126.72, 126.66, 125.94, 124.49, 124.41, and 123.66 ppm.

v_{max}/cm^{-1} – 3059 (broad, m, Arom. -C-H stretch), 1520 and 1339 (s, Nitro –N=O stretches).

HRMS-ESI (m/z) – Found ([M+H]$^+$, 328.0826); calc. 328.0829 (C$_{18}$N$_5$O$_2$H$_{10}$$^+$).
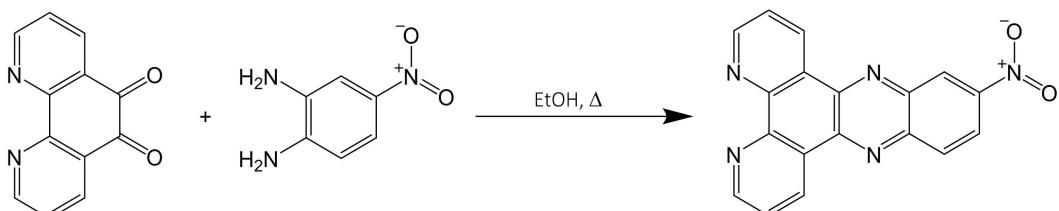
7.3.1.11 12,17-dihydronaphtho[2,3-h]dipyrido[3,2-a:2',3'-c]phenazine-12,17-dione (Aqphen)
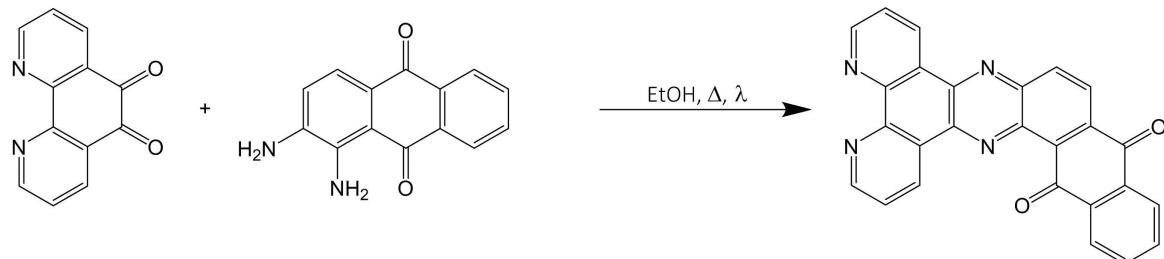
Phendione (50 mg, 0.238 mmol) and 1,2-diaminoanthraquinone (56.7 mg, 0.238 mmol) were both suspended together in an ethanolic solution (7 mL) containing a trace amount of ρ-toluene sulfonic acid within a CEM microwave tube (10 mL). The violet coloured solution was degassed/evacuated with Ar for 15 minutes before being fully sealed and installed into the synthesis microwave. The sample was irradiated with 150W at 140°C for 20 minutes, yielding a deep red/violet solution which was ensured to be cool and then filtered by suction to collect the black precipitate. The powder was suspended in hot chloroform (100 mL) in the presence of powdered charcoal and filtered through a glass frit, yielding a brown solution. The filtrate was reduced to approximately 5 mL in volume (mixture of purple and yellow coloured precipitation is noted) and diethyl ether (50 mL) was added to complete the precipitation. The powder was collected via suction filtration and washed with diethyl ether (3 x 10 mL), yielding the target product as a yellow-ochre powder (52 mg, 0.129 mmol, 54 %).

δ$_H$ (400 MHz, TMS, CDCl$_3$-d) – 9.90 (dd, J = 8.1, 1.6 Hz, 1H), 9.66 (dd, J = 7.6, 1.8 Hz, 1H), 9.34 (d, J = 4.7 Hz, 2H), 8.84 (d, J = 8.8 Hz, 1H), 8.72 (d, J = 8.8 Hz, 1H), 8.44 (d, J = 7.6 Hz, 1H), 8.35 (d, J = 7.6 Hz, 1H), 7.96-7.88 (m, 2H), and 7.90-7.82 ppm (m, 2H). *See figure A7.15.*

HRMS-ESI (m/z) – Found ([M+H]$^+$, 413.1032); calc. 413.1033 (C$_{26}$N$_4$O$_2$H$_{13}$$^+$).

7.3.2 Precursors

7.3.2.1 ruthenium (η^4-cycloocta-1,5-diene) dichloride (Ru(COD)Cl$_2$)

RuCl$_3$·xH$_2$O (8.0 g) was dried overnight in an oven (110 ^{o}C) to remove any residual H$_2$O of hydration. RuCl$_3$·3H$_2$O (5.0 g, 19.13 mmol) was suspended in a solution of ethanol (40 mL) and cycloocta-1,5-diene (5 mL, 40.8 mmol) and refluxed for 24 hours in an argon atmosphere. The cooled suspension was filtered by suction and the precipitate washed with diethyl ether (3 x 5 mL) yielding the target complex as a brown powdered solid (2.18 g, 7.90 mmol, 41 %).

7.3.2.2 ruthenium tetrakis-(dimethyl sulfoxide) dichloride (cis-Ru(DMSO)₄Cl₂)

$$RuCl_3 \cdot 3H_2O \xrightarrow{\text{DMSO, } \Delta} $$

RuCl₃·xH₂O (3.0 g) was dried overnight in an oven (110 °C) to remove any residual H₂O of hydration. Separately, DMSO (15 mL) was added to a round bottom flask and degassed by bubbling with Ar. To the degassed liquid, dry RuCl₃·3H₂O (2.5 g, ~9.6 mmol) was added, and the mixture was heated to 160 °C with condenser attached. After 10 minutes, the condenser was removed and the brown-orange coloured solution was reduced to roughly half original volume. Following this, the solution was allowed to cool and then transferred *via* Pasteur pipette to an Erlenmeyer flask; where, with stirring, acetone (50 mL) was added. Subsequently, the flask was cooled on ice and left to stand for 30 mins, yielding a yellow precipitate which was collected by suction filtration. The complex was washed with acetone (3 x 5 mL) and diethyl ether (3 x 5 mL), then allowed to dry *in vacuo* yielding the target complex as a bright yellow microcrystalline solid (3.10 g, 6.4 mmol, 67 %).

δ$_H$ (400 MHz, TMS, H₂O-d_2) – 3.44 (d, *J* = 8.4 Hz, *cis*-12H), and 3.35 ppm (s, *trans*-12H).

### 7.3.2.3 ruthenium *bis*-(phenanthroline) dichloride (*cis*-Ru(phen)₂Cl₂)

Ru(COD)Cl₂ (578 mg, 1.2 mmol) and anhydrous 1,10-phenanthroline (450 mg, 2.5 mmol) were dissolved in dry dimethylacetamide (DMA, 20 mL). After full dissolution, lithium chloride (4.2 g) was added, with stirring, to the solution and then refluxed (150 °C) for 4 hours, or until a hint of orange/red can be seen in the black solution. The condenser was then removed and approximately ½ of the DMA was evaporated and the slurry was cooled to r.t. Following this, with heavy stirring, acetone (40 mL) was added and the mixture was allowed to stand overnight in a fridge (-20 °C). The product was filtered off by vacuum, and washed with cold acetone (3 x 5 mL) and cold water (3 x 10 mL); yielding the crude complex as a black precipitate. Further purification via recrystallization was achieved by suspending the solid in hot ethanol (50 mL); adding water (50 mL) with heavy stirring and boiled until complete dissolution. The solution was filtered whilst still hot, followed by a careful addition of lithium chloride (12 g). The mixture was then heated (80 °C) until roughly ½ of the original volume had evaporated, then allowed to cool following product separation from the mother liquor via *vacuo* filtration. The retentate was washed with cold water (3 x 15 mL), diethyl ether (3 x 10 mL) and finally allowed to dry in air to yield the target complex as a black/dark violet crystalline solid (281 mg, 0.53 mmol, 44 %).

δ$_H$ (400 MHz, TMS, DMSO-d_6) – 10.29 (d, *J* = 5.3 Hz, 2H), 8.73 (d, *J* = 8.0 Hz, 2H), 8.30 (d, *J* = 8.9 Hz, 2H), 8.27 - 8.19 (m, 4H), 8.15 (d, *J* = 8.9 Hz, 2H), 7.76 (d, *J* = 5.3 Hz, 2H), and 7.34 ppm (dd, *J* = 8.1, 5.3 Hz, 2H). *See figure A7.16.*

HRMS-ESI (m/z) – Found (M⁺, 531.9787); calc. 531.9790 (RuC$_{24}$N$_4$H$_{16}$Cl$_2^{2+}$).

## 7.3.2.4 ruthenium bis-(2-cyano-phenanthroline) dichloride (cis-Ru(2-CN-phen)₂Cl₂)

1,10-phenanthroline-2-carbonitrile (0.13 g, 0.63 mmol), lithium chloride (0.10 g, 2.38 mmol, 4 eqvs.), ruthenium(II) (η⁴-cycloocta-1,5-diene) dichloride (0.09 g, 0.32 mmol, 0.5 eqvs.) were stirred together until in DMA in a CEM microwave tube. The suspension was degassed with Ar for 10 mins. The solution was irradiated in a CEM microwave at 150W and 140°C for 20 mins, forming a dark purple solution. The solution was quickly suction filtered, added to acetone (150 mL), and left overnight in an Erlenmeyer in a freezer overnight (-20 °C). The resultant solution was filtered and the solid product retained, washed with cold water (3 x 15 mL), and diethyl ether (3 x 10 mL). The final product was yielded as a dark purple crystalline solid (0.10 g, 0.17 mmol, 53 %).

δ$_H$ (400 MHz, TMS, DMSO-d_6) – 10.45 (d, J = 5.4 Hz, 2H), 8.83 (d, J = 8.1 Hz, 2H), 8.71 (d, J = 8.3 Hz, 2H), 8.50 (d, J = 8.9 Hz, 2H), 8.33 (d, J = 8.9 Hz, 2H), 8.29 (dd, J = 8.1, 5.4 Hz, 2H), and 8.05 ppm (d, J = 8.3 Hz, 2H). *See figure A7.17.*

HRMS-ESI (m/z) – Found (M⁺, 581.9704); calc. 581.9695 (RuC$_{26}$N$_6$H$_{14}^+$)

7.3.2.5 ruthenium bis-tetraazaphenanthrene dichloride (cis-Ru(TAP)₂Cl₂)

Ru(DMSO)$_4$Cl$_2$ (334 mg, 1.2 mmol) and anhydrous TAP (460 mg, 2.5 mmol) were dissolved in dry dimethylacetamide (DMA, 20 mL). After full dissolution, lithium chloride (420 mg) was added, with stirring, to the solution and then refluxed (150 °C) for 40 minutes, or until a hint of orange/red can be seen in the black solution. The condenser was then removed and approximately ½ of the DMA was evaporated and the slurry was cooled to r.t. Following this, with heavy stirring, acetone (40 mL) was added and the mixture was allowed to stand overnight in a fridge (0 °C). The product was filtered off under vacuum, and washed with cold acetone (3 x 5 mL) and cold water (3 x 10 mL); yielding the crude complex as a black precipitate. Further purification via recrystallization was achieved by suspending the solid in hot ethanol (50 mL); adding water (50 mL) with heavy stirring and boiled until complete dissolution. The solution was filtered whilst still hot, followed by a careful addition of lithium chloride (12 g). The mixture was then heated (80 °C) until roughly ½ of the original volume had evaporated, then allowed to cool following product separation from the mother liquor via suction filtration. The retentate was washed with cold water (3 x 15 mL), Diethyl ether (3 x 10 mL) and finally allowed to dry in air to yield the target complex as a black/dark violet crystalline solid (340 mg, 0.073 mmol, 61 %).

δ_H (400 MHz, TMS, DMSO-d_6) – 10.25 (d, J = 2.8 Hz, 2H), 9.55 (d, J = 2.8 Hz, 1H), 8.70 (d, J = 3.5 Hz, 2H), 8.69 (d, J = 2.8 Hz, 1H), 8.56 (d, J = 9.3 Hz, 2H), and 8.40 ppm (d, J = 3.0 Hz, 2H). *See figure A7.18.*

HRMS-ESI (m/z) – Found (M$^+$, 535.9603); calc. 535.9600 (RuC$_{20}$N$_8$H$_{12}^+$).

7.3.2.6 ruthenium *bis*-bathophenanthroline dichloride (*cis*-Ru(BPhen)$_2$Cl$_2$)

RuCl$_3$·xH$_2$O (1.0 g) was dried overnight in an oven (110 °C) to remove any residual H$_2$O of hydration. RuCl$_3$·3H$_2$O (20.5 mg, 0.078 mmol) and bathophenanthroline (52 mg, 0.28 mmol) were suspended in a solution of DMA (7 mL) in a CEM microwave tube (10 mL). The black suspension was degassed with Argon and lightly stirred for 15 minutes before being sealed and installed in the synthesis microwave. The sample was irradiated with 150 W at 140 °C for 20 minutes, yielding a dark violet solution which was cooled and then evaporated *in vacuo*. The black powder was washed with water (4 x 5 mL) and acetone (2 x 5 mL) yielding the target product as a black-violet microcrystalline solid (0.11 g, 0.12 mmol, 43 %).

δ_H (400 MHz, TMS, DMSO-d_6) – 8.15 (d, *J* = 9.3 Hz, 2H), 7.88 (d, *J* = 9.3 Hz, 2H), 7.78 (s, 2H), 7.72 (d, *J* = 7.4 Hz, 4H), 7.60 (t, *J* = 7.4 Hz, 4H), 7.55 (d, *J* = 7.0 Hz, 4H), 7.52-7.46 (m, 6H), 7.41 (d, *J* = 7.6 Hz, 4H), and 7.41-7.31 ppm (m, 4H). *See figure A7.19.*

HRMS-ESI (m/z) – Found (M$^+$, 836.1026); calc. 836.1011 (^{100}RuC$_{48}$N$_4$H$_{32}$Cl^{37}Cl$^+$).

7.3.2.7 ruthenium *bis*-(4,5-diazafluoren-9-one) dichloride (*cis*-Ru(dafo)$_2$Cl$_2$)

Ru(COD)Cl$_2$ (38.4 mg, 0.14 mmol) and 4,5-diazafluoren-9-one (58 mg, 0.28 mmol) were suspended in a solution of DMA (7 mL) containing LiCl (50 mg) in a CEM microwave tube (10 mL). The brown suspension was degassed with Ar and light stirring for 15 minutes before being sealed and installed in the synthetic microwave. The sample was irradiated at 150 W at 140 °C for 20 minutes, yielding a dark violet-blue solution which was ensured to be cool and then evaporated *in vacuo*. The black powder was washed with water (4 x 5 mL) and acetone (2 x 5 mL) yielding the target product as a black-violet microcrystalline solid (42.0 mg, 0.078 mmol, 56 %).

δ_H (400 MHz, DMSO-d_6) – 9.36 (d, *J* = 5.6 Hz, 2H), 8.23 (d, *J* = 7.5 Hz, 2H), 8.01 (d, *J* = 5.5 Hz, 2H), 7.97 (d, *J* = 7.5 Hz, 2H), 7.89 (dd, *J* = 7.5, 5.6 Hz, 2H), and 7.37 ppm (t, *J* = 6.7 Hz, 2H). *See figure A7.20.*

HRMS-ESI (m/z) – Found (M$^+$, 535.9375); calc. 535.9375 (RuC$_{22}$N$_4$O$_2$H$_{12}^+$).

7.3.3 Homoleptic ruthenium polypyridyl complexes

7.3.3.1 ruthenium *tris*-(phenanthroline) dichloride (*rac*-[Ru(phen)₃]·Cl₂)

Ru(DMSO)₄Cl₂ (530 mg, 1.1 mmol) was dissolved in ethylene glycol (15 mL) in a 50 mL round bottomed flask; separately, 1,10-phenanthroline (612 mg, 3.4 mmol) was added to ethylene glycol (10 mL) and stirred until full dissolution. The dull yellow and orange solutions were carefully combined, with stirring, and heated under reflux conditions (70 °C) in an argon environment for 20 hours. Colour change was noted as the solution turned a transparent wine red from a dull yellow/brown. Isolation of the product was achieved via rotary evaporation, followed by collection by vacuum filtration and finally washed with toluene (3 x 5 mL). The crude material can then be further purified by means of flash chromatography on neutral alumina, using ethanol as the mobile phase, to yield the target complex as a red crystalline solid (721 mg, 1.0 mmol, 92 %).

Note that where required, the hexafluorophosphate form (PF₆⁻) can be generated through metathesis by dropwise addition of a warm supersaturated solution of potassium hexafluorophosphate, followed by cooling overnight (0°C), filtering the substrate *in vacuo* and subsequently washing the compound with water (5 mL).

δ_H (400 MHz, TMS, H₂O-*d₂*) – 8.51 (d, *J* = 8.4 Hz, 6H), 8.16 (s, 6H), 8.04 (d, *J* = 5.2 Hz), and 7.53 ppm (dd, *J* = 5.2, 8.4 Hz, 6H). *See figure A7.21.*

δ_C (101 MHz, H₂O-*d₂*) – 152.25, 147.89, 136.56, 130.74, 127.80, and 125.31 ppm.

HRMS-ESI (m/z) – Found (M⁺, 321.0552); calc. 321.0547 (RuC₃₆N₆H₂₄²⁺).

7.3.3.2 ruthenium *tris*-(tetraazaphenanthrene) dichloride (*rac*-[Ru(TAP)₃]·Cl₂)

Following the same procedure (but on a quarter scale) as for *rac*-[Ru(phen)₃]·Cl₂, *rac*-[Ru(TAP)₃]·Cl₂ was prepared using tetraazaphenanthrene (155 mg, 0.85 mmol) instead of phenanthroline yielding the target product as a red crystalline solid (181 mg, 0.252 mmol, 89 %).

δ_H (400 MHz, TMS, CH_3CN-d_3) – 9.01 (d, J = 2.7 Hz, 6H), 8.64 (s, 6H), and 8.27 ppm (d, J = 2.7 Hz, 6H).

δ_C (101 MHz, CH_3CN-d_3) – 149.27, 148.99, 145.27, 141.47, and 132.54 ppm.

HRMS-ESI (m/z) – Found (M$^+$, 323.5411); calc. 323.5411 ($^{101}RuC_{36}N_6H_{24}^{2+}$).

7.3.3.3 ruthenium *tris*-(dipyridophenazine) dichloride (*rac*-[Ru(dppz)$_3$]·Cl$_2$)

Following the same procedure as for *rac*-[Ru(phen)$_3$]·Cl$_2$, *rac*-[Ru(dppz)$_3$]·Cl$_2$ was prepared using dipyridophanazine (960 mg, 3.4 mmol) instead of phenanthroline yielding the target product as a red crystalline solid (953 mg, 0.935 mmol, 82 %).

δ_H (400 MHz, TMS, CH_3CN-d_3) – 9.73 (dd, J = 8.3, 1.3 Hz, 6H), 8.66 (dd, J = 5.5, 1.3 Hz, 6H), 8.47 (dd, J = 6.6, 3.4 Hz, 6H), 8.07 (dd, J = 6.6, 3.4 Hz, 6H), and 8.03 ppm (dd, J = 8.3, 5.5 Hz, 6H).

δ_C (101 MHz, CH_3CN-d_3) – 154.23, 150.35, 142.37, 139.58, 133.40, 132.19, 130.51, 129.24, and 127.03 ppm.

HRMS-ESI (m/z) – Found (M$^+$, 473.5869); calc. 473.5881 ($^{101}RuC_{54}N_{12}H_{30}^{2+}$).

7.3.4 Heteroleptic ruthenium polypyridyl complexes

7.3.4.1 ruthenium *bis*-(phenanthroline) dipyridophenazine dichloride (*rac*-[Ru(phen)$_2$(dppz)]·Cl$_2$)

Ru(phen)$_2$Cl$_2$ (80 mg, 0.15 mmol) and dppz (42.5 mg, 0.15 mmol) were both suspended together in an aqueous ethanol solution (7 mL, 1:1) within a CEM microwave tube (10 mL). The violet coloured solution was degassed/evacuated with Ar for 15 minutes before being fully sealed and installed into the synthetic microwave. The sample was irradiated at 150 W at 140 °C for 40 minutes,

yielding a deep red/brown solution which was ensured to be cool and then filtered by suction filtration. Subsequent precipitation of the target compound from the filtrate was achieved by metathesis via dropwise addition of a saturated solution of aqueous potassium hexafluorophosphate (KPF$_6$). Isolation of the PF$_6^-$ salt by suction filtration yielded a dark orange/brown solid, which, after washing with cold water (2 x 2 mL) was allowed to dry in air. Conversion to the compounds chloride form through the dissolution of the crude material in a minimal amount of acetonitrile (~5 mL) and after addition of HPLC grade water (10 mL) was added to a beaker of dry, washed, Amberlite ion exchange resin (IRA-400, Cl$^-$ form, 2.4 g), covered and lightly stirred for 20 hours. Following removal of the resin by gravity filtration, the complex was isolated via rotary evaporation and purified on an aqueous Sephadex C-25 column using 0.2M NaCl as the mobile phase (eluting as a deep red/orange band). Finally, the compound was isolated as the chloride form after anionic exchange via treatment with Amberlite resin (IRA-400, Cl$^-$ form, 2.4 g), to yield the complex as a deep red/brown microcrystalline solid (111 mg, 0.137 mmol, 91 %).

δ_H (400 MHz, TMS, CH$_3$CN-d_3) — 9.54 (dd, J = 8.2, 1.3 Hz, 2H), 8.57 (ddd, J = 8.3, 4.8, 1.3 Hz, 4H), 8.38 (dt, J = 6.5, 3.3 Hz, 2H), 8.21 (s, 4H), 8.16 (dd, J = 5.4, 1.3 Hz, 2H), 8.08-8.02 (m, 4H), 7.96 (dd, J = 5.2, 1.3 Hz, 2H), 7.70 (dd, J = 8.2, 5.4 Hz, 2H), and 7.59 ppm (ddd, J = 13.7, 8.3, 5.2 Hz, 4H). *See figure A7.22.*

δ_C (101 MHz, CH$_3$CN-d_3) — 155.94, 153.89, 152.28, 150.53, 147.91, 147.72, 141.87, 139.24, 137.07, 136.91, 132.87, 132.52, 130.93, 129.80, 129.50, 129.27, 128.63, 128.08, 127.91, 126.65, and 125.56 ppm.

HRMS-ESI (m/z) — Found (M$^+$, 372.0656); calc. 372.0656 (RuC$_{42}$N$_8$H$_{26}^{2+}$).

7.3.4.2 ruthenium *bis*-(phenanthroline) 11-cyano-dipyridophenazine dichloride (*rac*-[Ru(phen)$_2$(11-CN-dppz)]·Cl$_2$)

The synthesis of *rac*-[Ru(phen)$_2$(11-CN-dppz)]·Cl$_2$ is commensurate to the methodology implemented for the parent complex, however, 11-CN-dppz (44.3 mg, 0.15 mmol) was used instead of dppz, yielding the target complex as a deep red microcrystalline solid (103 mg, 0.12 mmol, 82 %).

δ_H (400 MHz, TMS, CH$_3$CN-d_3) — 9.65 (q, J = 8.4 Hz, 2H), 9.01-8.97 (m, 1H), 8.59 (t, J = 7.0 Hz, 4H), 8.62 (d, J = 9.0 Hz, 1H), 8.32 (s, 1H), 8.31 (s, 4H), 8.26 (d, J = 4.9 Hz, 2H), 8.17 (d, J = 5.7 Hz, 2H), 8.05 (d, J = 4.9 Hz, 2H), 7.82 (dd, J = 8.2, 5.0 Hz, 2H), and 7.74-7.64 ppm (m, 4H). *See figure A7.23.*

v_{max}/cm^{-1} — 3000 (broad, m, Arom. -C-H stretch) and 2230 (m, Nitrile -C≡N stretch).

HRMS-ESI (m/z) — Found (M$^+$, 384.5643); calc. 384.5633 (RuC$_{43}$N$_9$H$_{25}^{2+}$).

7.3.4.3 ruthenium *bis*-(phenanthroline) 11,12-dicyano-dipyridophenazine dichloride (*rac*-[Ru(phen)₂(11,12-CN-dppz)]·Cl₂)

The synthesis of *rac*-[Ru(phen)₂(11,12-CN-dppz)]·Cl₂ is commensurate to the methodology implemented for the parent complex, however, 11,12-CN-dppz (50 mg, 0.15 mmol) was used instead of dppz, yielding the target complex as a red/brown microcrystalline solid (116 mg, 0.135 mmol, 90 %).

δH (400 MHz, TMS, H₂O-d_2) – 9.57 (d, J = 8.3 Hz, 2H), 8.94 (s, 2H), 8.59 (t, J = 7.6 Hz, 4H), 8.31-8.21 (m, 4H), 8.22 (s, 4H), 8.06 (d, J = 5.1 Hz, 2H), 7.77 (dd, J = 8.3, 5.5 Hz, 2H), 7.65 (dd, J = 8.3, 5.3 Hz, 2H), and 7.59 ppm (dd, J = 8.4, 5.3 Hz, 2H). *See figure A7.24.*

ν$_{max}$/cm⁻¹ – 3020 (broad, m, Arom. -C-H stretch) and 2234 (m, Nitrile -C≡N stretch).

HRMS-ESI (m/z) – Found (M⁺, 396.5616); calc. 396.5615 (^{101}RuC₄₄N₁₀H₂₄$^{2+}$).

7.3.4.4 ruthenium *bis*-(phenanthroline) 10-nitro-dipyridophenazine dichloride (*rac*-[Ru(phen)₂(10-NO₂-dppz)]·Cl₂)

The synthesis of *rac*-[Ru(phen)₂(10-NO₂-dppz)]·Cl₂ is commensurate to the methodology implemented for the parent complex, however, 10-NO₂-dppz (49 mg, 0.15 mmol) was used instead of dppz, yielding the target complex as a red microcrystalline solid (117 mg, 0.137 mmol, 91 %).

δH (400 MHz, CH₃CN-d_3) – 9.55 (dd, J = 8, 2 Hz, 1H), 9.44 (dd, J = 8, 2 Hz, 1H), 8.65 (dd, J = 8, 2 Hz, 1H), 8.57 (t, J = 8 Hz, 4H), 8.48 (dd, J = 8, 2 Hz, 1H), 8.21 (s, 2H), 8.23-8.12 (m, 3H), 8.08 (t, J = 4 Hz, 2H), 7.98-7.93 (m, 2H), 7.76-7.66 (m, 2H), and 7.65-7.54 ppm (m, 4H). *See figure A7.25.*

HRMS-ESI (m/z) – Found (M⁺, 394.0589); calc. 394.0588 (^{101}RuC₄₄N₁₀H₂₄$^{2+}$).

7.3.4.5 ruthenium *bis*-(phenanthroline) 11-nitro-dipyridophenazine dichloride (*rac*-[Ru(phen)$_2$(11-NO$_2$-dppz)]·Cl$_2$)


The synthesis of *rac*-[Ru(phen)$_2$(11-NO$_2$-dppz)]·Cl$_2$ is commensurate to the methodology implemented for the parent complex, however, 11-NO$_2$-dppz (49 mg, 0.15 mmol) was used instead of dppz, yielding the target complex as a red microcrystalline solid (105 mg, 0.122 mmol, 82 %).

δ$_H$ (400 MHz, TMS, CH$_3$CN-d_3) – 9.27-9.18 (m, 2H), 8.63 (d, J = 8.4 Hz, 1H), 8.61-8.52 (m, 3H), 8.44 (d, J = 5.2 Hz, 1H), 8.36 (d, J = 8.7 Hz, 2H), 8.30 (d, J = 9.7 Hz, 1H), 8.20 (m, 7H), 8.06 (dd, J = 13.6, 5.2 Hz, 2H), 7.76 (dd, J = 8.3, 5.6 Hz, 1H), 7.69 (dd, J = 8.3, 5.6 Hz, 1H), and 7.66-7.53 ppm (m, 4H). *See figure A7.26.*

HRMS-ESI (m/z) – Found (M$^+$, 394.0587); calc. 394.0588 (^{101}RuC$_{42}$N$_{19}$O$_2$H$_{25}$$^{2+}$).
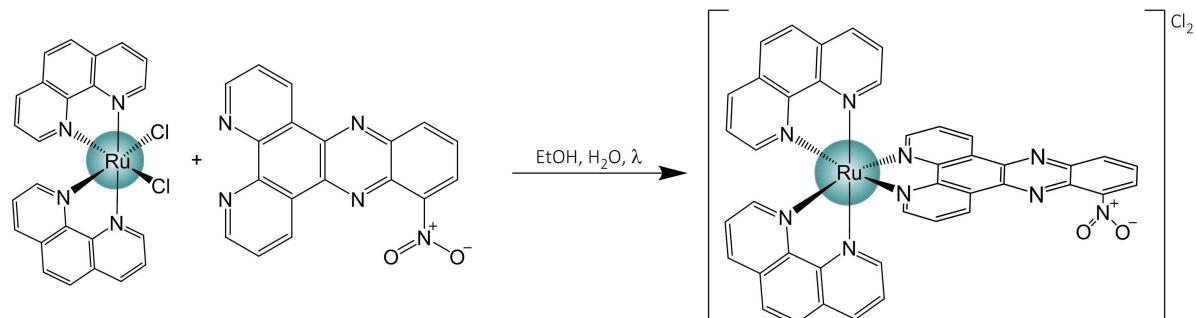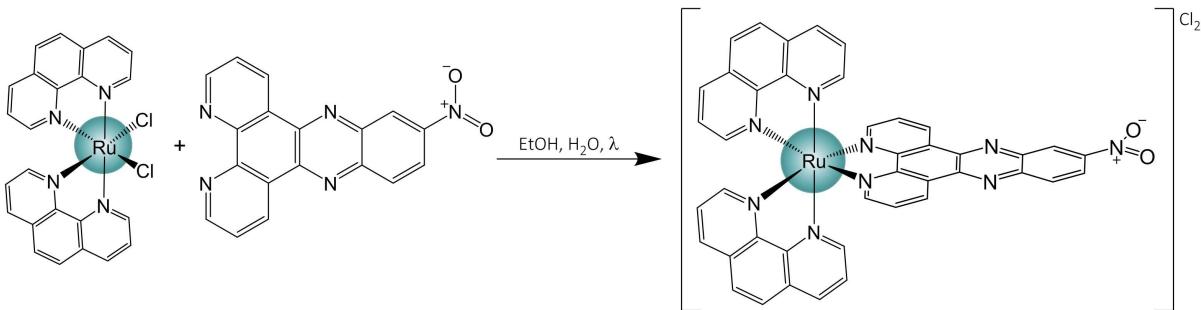
7.3.4.6 ruthenium bis-(phenanthroline) 12,17-dihydro-naphtho-dipyridophenazine-12,17-dione dichloride (*rac*-[Ru(phen)$_2$(Aqphen)]·Cl$_2$)
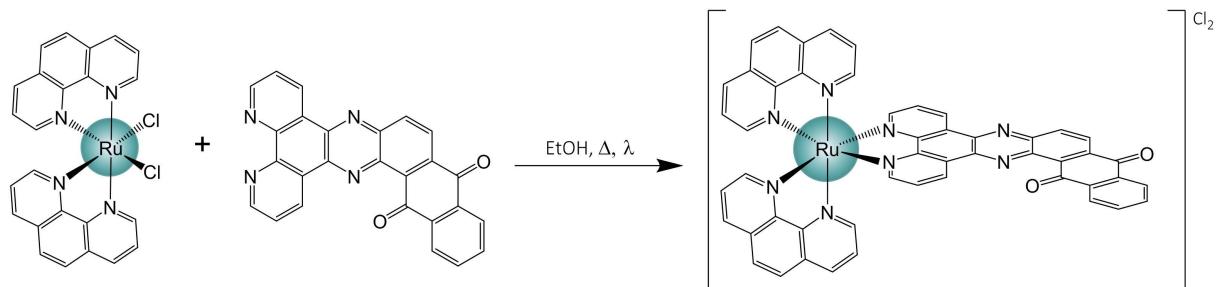
The synthesis of *rac*-[Ru(phen)$_2$(Aqphen)]·Cl$_2$ is commensurate to the methodology implemented for the parent complex, however, Aqphen (62 mg, 0.15 mmol) was used instead of dppz. Purification of the species was done so on a silica column using an eluent of 80:20:0.5 ACN:H$_2$O:sat.KNO$_3$, yielding the target complex as a dark red microcrystalline solid (100 mg, 0.106 mmol, 91 %).

δ$_H$ (400 MHz, TMS, CH$_3$CN-d_3) – 9.84 (d, J = 8.2 Hz, 1H), 9.64 (d, J = 8.2 Hz, 1H), 8.90 (d, J = 8.9 Hz, 1H), 8.80-8.71 (m, 3H), 8.69 (dd, J = 5.4, 1.3 Hz, 1H), 8.64 (dd, J = 5.4, 1.3 Hz, 1H), 8.52 (m, 6H), 8.34 (dt, J = 7.0, 2.6 Hz, 2H), 8.16 (s, 4H), 8.08 (ddd, J = 7.8, 5.4, 1.7 Hz, 2H), and 7.99-7.85 ppm (m, 6H). *See figure A7.27.*

HRMS-ESI (m/z) – Found (M$^+$, 437.0691); calc. 437.0684 (RuC$_{50}$N$_8$O$_2$H$_{28}$$^{2+}$).

7.3.4.7 ruthenium *bis*-(tetraazaphenanthrene) dipyridophenazine dichloride (*rac*-[Ru(TAP)$_2$(dppz)]·Cl$_2$)

Ru(TAP)$_2$Cl$_2$ (81 mg, 0.15 mmol) and dppz (42.5 mg, 0.15 mmol) were both suspended together in an aqueous ethanol solution (7 mL, 1:1) within a CEM microwave tube (10 mL). The violet coloured solution was degassed/evacuated with Ar for 15 minutes before being fully sealed and installed into the synthetic microwave. The sample was irradiated at 150 W at 140 °C for 40 minutes, yielding a deep red/brown solution which was ensured to be cool and then collected by suction filtration. Subsequent precipitation of the target compound from the filtrate was achieved by metathesis via dropwise addition of a saturated solution of aqueous potassium hexafluorophosphate (KPF$_6$). Isolation of the PF$_6^-$ salt by suction filtration yielded a dark orange/brown solid, which, after washing with cold water (2 x 2 mL) was allowed to dry in air. Conversion to the compounds chloride form through the dissolution of the crude material in a minimal amount of acetonitrile (∼5 mL) and after addition of HPLC grade water (10 mL) was added to a beaker of dry, washed, Amberlite ion exchange resin (IRA-400, Cl$^-$ form, 2.4 g), covered and lightly stirred for 20 hours. Following removal of the resin by gravity filtration, the complex was isolated via rotary evaporation and purified on an aqueous Sephadex C-25 column using 0.2M NaCl as the mobile phase (eluting as a deep red/orange band). Finally, the compound was isolated as the chloride form after anionic exchange via treatment with Amberlite resin (IRA-400, Cl$^-$ form, 2.4 g), to yield the complex as a deep red/brown microcrystalline solid (96 mg, 0.12 mmol, 80 %).

δ$_H$ (400 MHz, TMS, CH$_3$CN-d_3) – 9.67 (dd, *J* = 8.2, 1.3 Hz, 2H), 8.90 (dd, *J* = 9.6, 2.8 Hz, 4H), 8.55 (s, 4H), 8.42 (dd, *J* = 6.6, 3.4, 2H), 8.23 (d, *J* = 2.8 Hz, 2H), 8.17 (d, *J* = 2.8 Hz, 2H), 8.13-8.06 (m, 4H), and 7.77 ppm (dd, *J* = 8.3, 5.4 Hz, 2H). *See figure A7.28.*

δ$_C$ (101 MHz, CH$_3$CN-d_3) – 151.85, 147.79, 146.81, 146.62, 146.34, 143.00, 140.55, 140.48, 137.50, 133.45, 130.84, 130.63, 130.58, 128.89, 126.95, 125.41, and 124.93 ppm.

HRMS-ESI (m/z) — Found (M$^+$, 374.0559); calc. 374.0561 (RuC$_{38}$N$_{12}$H$_{23}^{2+}$).

7.3.4.8 ruthenium *bis*-(tetraazaphenanthrene) 11-cyano-dipyridophenazine dichloride (*rac*-[Ru(TAP)₂(11-CN-dppz)]·Cl₂)

In an analogous fashion to the synthesis of *rac*-[Ru(TAP)₂(dppz)]·Cl₂, *rac*-[Ru(TAP)₂(11-CN-dppz)]·Cl₂ was synthesised using 11-CN-dppz (46 mg, 0.15 mmol) instead of dppz; yielding the target product as an orange/red microcrystalline solid (94 mg, 0.11 mmol, 74 %).

δ_H (400 MHz, TMS, H₂O-*d₂*) – 9.74 (d, *J* = 8.4 Hz, 2H), 9.02 (dd, *J* = 3.2, 6.8 Hz, 2H), 8.99 (d, *J* = 3.2 Hz, 2H), 8.73 (s, 1H), 8.65 (s, 4H), 8.53 (dd, *J* = 7, 12.4 Hz, 2H), 8.47 (d, *J* = 8.8 Hz, 1H), 8.38 (t, *J* = 2.0 Hz, 2H), 8.24 (d, *J* = 5.2 Hz, 2H), 8.07 (d, *J* = 8.8 Hz, 1H), and 7.92 ppm (dd, *J* = 5.2, 8.8 Hz, 2H). *See figure A7.29.*

δ_C (101 MHz, H₂O-*d₂*) – 154.8, 154.7, 150.5, 149.1, 148.6, 145.2, 143.5, 142.6, 141.2, 135.9, 132.8, 132.2, 131.0, 130.6, 127.9, and 114.4 ppm.

ν_max/cm⁻¹ - 3000 (broad, m, Arom. -C-H stretch) and 2232 (m, Nitrile -C≡N stretch).

HRMS-ESI (m/z) – Found (M+H⁺, 386.5536); calc. 386.5538 (RuC₃₉N₁₃H₂₁²⁺)

7.3.4.9 ruthenium *bis*-(tetraazaphenanthrene) 11,12-dicyano-dipyridophenazine dichloride (*rac*-[Ru(TAP)₂(11,12-CN-dppz)]·Cl₂)

In an analogous fashion to the synthesis of *rac*-[Ru(TAP)₂(dppz)]·Cl₂, *rac*-[Ru(TAP)₂(11,12-CN-dppz)]·Cl₂ was synthesised using 11,12-CN-dppz (50 mg, 0.15 mmol) instead of dppz; yielding the target product as an orange/red microcrystalline solid (100 mg, 0.115 mmol, 77 %).

δ_H (400 MHz, TMS, H₂O-*d₂*) – 9.80 (d, *J* = 8.3 Hz, 2H), 9.07-8.90 (m, 6H), 8.68 (s, 4H), 8.54 (d, *J* = 3.0 Hz, 2H), 8.40 (d, *J* = 3.0 Hz, 2H), 8.28 (d, *J* = 5.4 Hz, 2H), and 7.95 ppm (dd, *J* = 8.4, 5.4 Hz, 2H). *See figure A7.30.*

δ_C (101 MHz, H$_2$O-d_2) – 150.06, 150.63, 149.24, 149.08, 148.86, 148.54, 145.32, 145.25, 142.71, 142.28, 141.98, 141.85, 137.71, 135.22, 132.60, 132.51, 129.83, 127.72, 115.18 and 114.90 ppm.

ν_{max}/cm^{-1} – 3051 (broad, m, Arom. -C-H stretch) and 2231 (m, Nitrile -C≡N stretch).

HRMS-ESI (m/z) – Found (M$^+$, 398.5519); calc. 398.5520 (Ru^{101}C$_{40}$N$_{14}$H$_{20}$$^{2+}$)
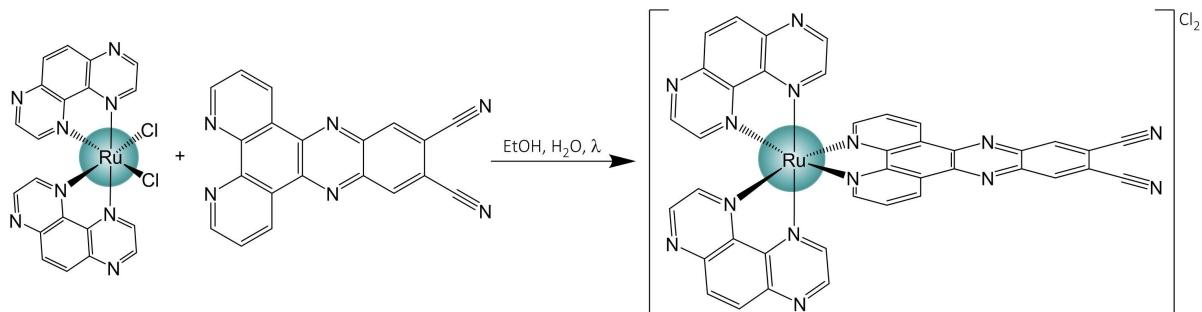
7.3.4.10 ruthenium *bis*-(tetraazaphenanthrene) 11-chloro-dipyridophenazine dichloride (*rac*-[Ru(TAP)$_2$(11-Cl-dppz)]·Cl$_2$)
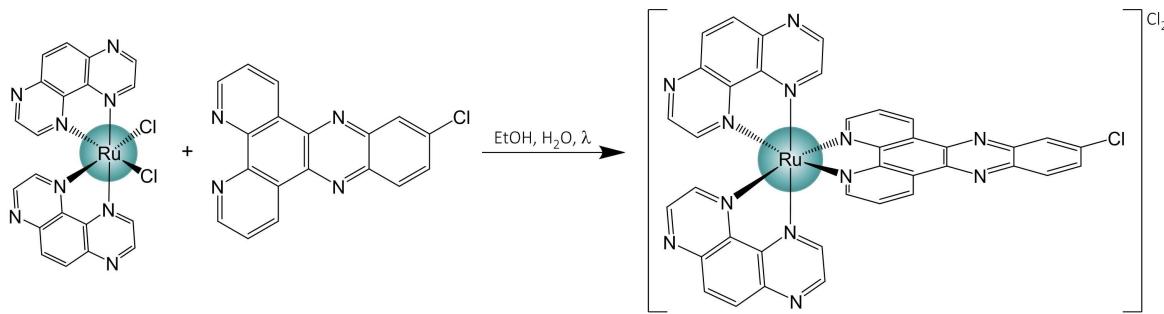
In an analogous fashion to the synthesis of *rac*-[Ru(TAP)$_2$(dppz)]·Cl$_2$, *rac*-[Ru(TAP)$_2$(11-Cl-dppz)]·Cl$_2$ was synthesised using 11-Cl-dppz (47.5 mg, 0.15 mmol) instead of dppz; yielding the target product as an orange/red microcrystalline solid (86 mg, 0.106 mmol, 67 %).

δ_H (400 MHz, TMS, H$_2$O-d_2) – 9.69 (dd, J = 8.4, 2.8 Hz, 2H), 9.02 (dd, J = 8.7, 3.0 Hz, 2H), 9.00 (d, J = 2.9 Hz, 2H), 8.66 (s, 4H), 8.53 (d, J = 2.9 Hz, 1H), 8.49 (d, J = 2.9 Hz, 1H), 8.38 (t, J = 3.0 Hz, 2H), 8.35 (d, J = 9.3 Hz, 1H), 8.31 (d, J = 2.1 Hz, 1H), 8.21 (t, J = 5.8 Hz, 2H), 7.95 (dd, J = 9.5, 2.1 Hz, 1H), and 7.89 ppm (ddd, J = 7.8, 5.4, 1.9 Hz, 2H). *See figure A7.31.*

7.3.4.11 ruthenium *bis*-(tetraazaphenanthrene) 11-bromo-dipyridophenazine dichloride (*rac*-[Ru(TAP)$_2$(11-Br-dppz)]·Cl$_2$)
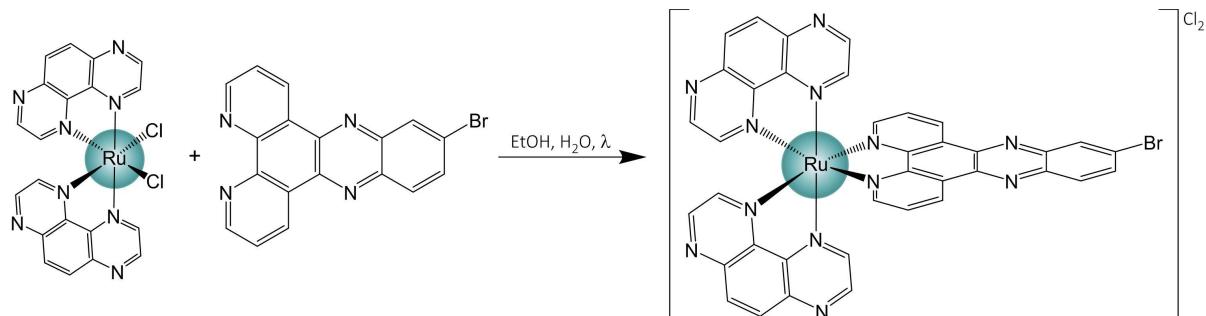
The synthesis of *rac*-[Ru(TAP)$_2$(11-Br-dppz)]·Cl$_2$ is commensurate to the methodology implemented for the chloride alternative, however, 11-Br-dppz (54 mg, 0.15 mmol) was used instead of 11-CN-dppz, yielding the target complex as a deep red/brown solid (101 mg, 0.11 mmol, 75 %).

δ_H (400 MHz, TMS, H$_2$O-d_2) – 9.71 (d, J = 8.3 Hz, 2H), 9.05-8.96 (m, 4H), 8.66 (s, 4H), 8.57 (s, 1H), 8.52 (d, J = 2.9 Hz, 1H), 8.47 (d, J = 2.9 Hz, 1H), 8.37 (t, J = 2.9 Hz, 2H), 8.32 (d, J = 9.1 Hz, 1H), 8.23-8.17 (m, 2H), 8.13 (dd, J = 9.1, 2.1 Hz, 1H), and 7.88 ppm (ddd, J = 8.2, 5.4, 2.7 Hz, 2H). *See figure A7.32.*

HRMS-ESI (m/z) – Found (M$^+$, 413.0119); calc. 413.0128 (Ru^{101}C$_{38}$N$_{12}$BrH$_{21}$$^{2+}$).

7.3.4.12 ruthenium *bis*-bathophenanthroline dipyridophenazine (*rac*-[Ru(BPhen)₂(dppz)]·Cl₂)
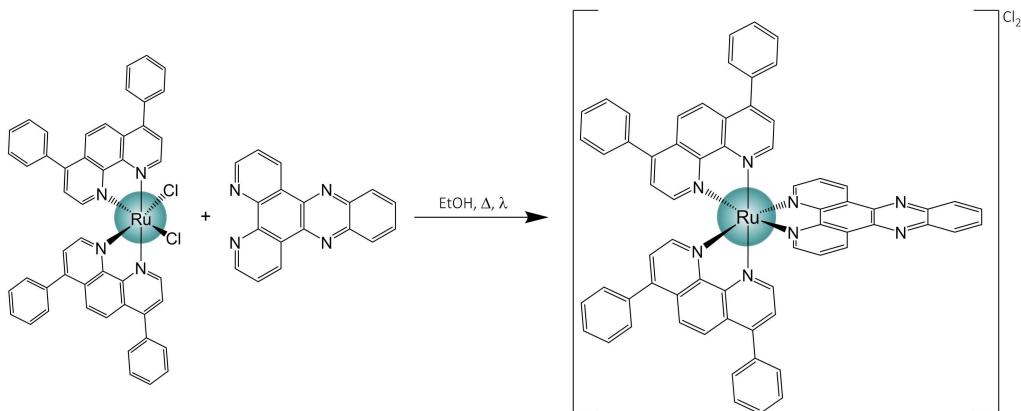
The synthesis of *rac*-[Ru(BPhen)₂(dppz)]·Cl₂ is commensurate to the methodology implemented for [Ru(phen)₂(dppz)]·Cl₂, however, *cis*-Ru(BPhen)₂Cl₂ (125.5 mg, 0.15 mmol) was used instead of *cis*-Ru(phen)₂Cl₂, yielding the target complex as a deep red/brown solid (136 mg, 0.122 mmol, 81 %).

δ_H (400 MHz, TMS, DMSO-d_6) – 9.69 (dd, J = 8.2, 1.3 Hz, 2H), 8.56 (dd, J = 6.6, 3.4 Hz, 2H), 8.40 (d, J = 5.5 Hz, 2H), 8.38-8.34 (m, 4H), 8.29 (m, 4H), 8.23 (dd, J = 6.6, 3.4 Hz, 2H), 8.03 (dd, J = 8.3, 5.5 Hz, 2H), 7.88-7.81 (m, 4H), and 7.71-7.61 ppm (m, 20H). *See figure A7.33.*

HRMS-ESI (m/z) – Found (M⁺, 524.1278); calc. 524.1282 (RuC₆₆N₈H₄₂²⁺).

7.3.4.13 ruthenium *bis*-diazafluoren-9-one dipyridophenazine (*rac*-[Ru(dafo)₂(dppz)]·Cl₂)
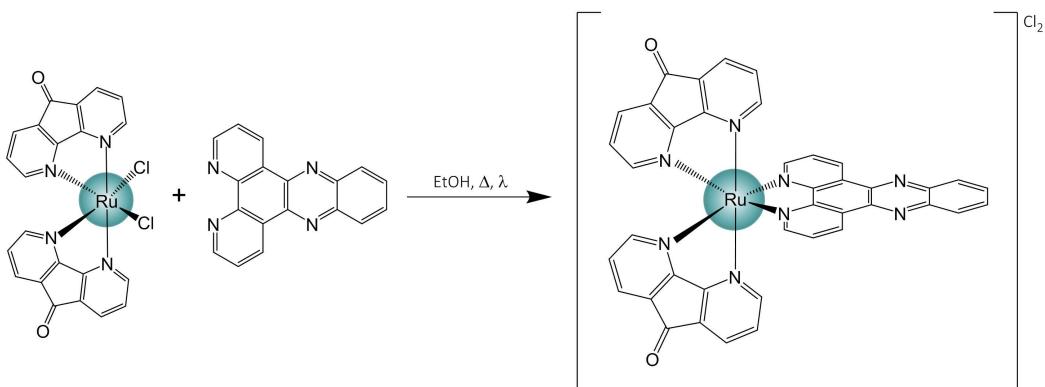
The synthesis of *rac*-[Ru(dafo)₂(dppz)]·Cl₂ is similar to the methodology implemented for [Ru(phen)₂(dppz)]·Cl₂, however, *cis*-Ru(dafo)₂Cl₂ (75.6 mg, 0.15 mmol) was used instead of *cis*-Ru(phen)₂Cl₂. The complex was purified on a silica column eluting with an 80:20:0.5 ACN:H₂O:sat.KNO₃ mixture, yielding the target complex as a deep red/brown solid (136 mg, 0.122 mmol, 81 %).

δ_H (400 MHz, TMS, CH₃CN-d_3) – 9.73 (dd, J = 8.3, 1.3 Hz, 2H), 8.74 (dd, J = 5.4, 1.2 Hz, 2H), 8.52 (dd, J = 6.5, 3.4 Hz, 2H), 8.25 (dd, J = 7.7, 0.9 Hz, 2H), 8.23 (dd, J = 5.6, 0.9 Hz, 2H), 8.18 (dd, J = 6.6, 3.4 Hz, 2H), 8.12 (dd, J = 5.7, 0.9 Hz, 2H), 8.03 (dd, J = 8.2, 5.4 Hz, 2H), 7.77 (dd, J = 5.7, 0.9 Hz, 2H), 7.69 (dd, J = 7.7, 5.6 Hz, 2H), and 7.44 ppm (dd, J = 7.6, 5.7 Hz, 2H).

8 Appendix

8.1 Chapter 1

Table A1.1 – *Summary of all publicly available G-quadruplex-ligand solution NMR structures.*

PDB	Date	Experiment	DNA Sequence	Salt	Topology	Ligand	Ligand Class	Ligand Types	Binding Interactions
1NZM	2003	Solution NMR	TTAGGGT	100 mM K$^+$	Parallel	Fluorinated pentacyclic quinoacridinium cation	Organic Macrocycle (non-planar)	P/nP/Hc/HA/1+	Tetrad π-stacking/GpT intercalation
2A5R	2005	Solution NMR	TGAGGGTGGIGAGGGTGGGGAAGG	90 mM K$^+$	Parallel (Foldback)	TMPyP4 (Substituted Porphyrin)	Organic Macrocycle (Planar)	P/nP/Hc/HA/HD/4+	Tetrad π-stacking and Loop threading
2I7T	2008	Solution NMR	TGGGGT	80 mM K$^+$	Parallel	Distamycin-A (Pyrrole-amidine)	Polyamide (Flexible chain)	P/nP/Hc/HA/HD/1+	Groove binding
2JWQ	2008	Solution NMR	TTAGGGT	110 mM K$^+$	Parallel	MMQ1 (meta-quinacridine n-propylamine)	Organic Heterocycle w/ saturated chain substituents (charged)	P/nP/Hc/HA/HD	Tetrad pi-stacking/GpA intercalation
2KVY	2010	Solution NMR	TGGGGT	80 mM K$^+$	Parallel	Distamycin A Derivative (Pyrrole-amidine)	Polyamide (Flexible chain)	P/nP/Hc/HA/HD	Groove binding
2L7V	2011	Solution NMR	TGA(GGGT)$_2$A(GGGT)$_2$AA	100 mM K$^+$	Parallel	Quindoline derivative	Organic Heterocycle (Planar)	P/nP/Hc/HA/HD	Tetrad pi-stacking/Side-on nucleobase H-bonding
2MB3	2013	Solution NMR	TTGGG(TTAG3)$_3$A	90 mM K$^+$	Hybrid-1	L2H2-6M(2)OTD (Telomestatin Dev/macrocyclic hexaoxazole)	Organic Macrocycle (Planar)	nP/Hc/HA/HD/2+	Tetrad π-stacking
2MCC	2013	Solution NMR	AGGG(TTAG3)$_3$	50 mM Na$^+$	Antiparallel (basket)	Ruthenium polypyridyl (ΔΔ-Dinuclear)	Metal-Organic (Octahedral/charged)	P/nP/Hc/HD/4+	Tetrad π-stacking
2MCO	2013	Solution NMR	AGGG(TTAG3)$_3$	50 mM Na$^+$	Antiparallel (basket)	Ruthenium polypyridyl (ΛΛ-Dinuclear)	Metal-Organic (Octahedral/charged)	P/nP/Hc/HD/4+	Tetrad π-stacking/Loop threading
2MGN	2014	Solution NMR	TGAGGGTGGGTGAGGGTGGGAAGG	45 mM K$^+$	Parallel	Phen-DC3 (Bisquinolinium)	Alkyl-linked macrocycles	P/nP/Hc/HA/HD/2+	Tetrad π-stacking
2MS6	2014	Solution NMR	TTAGGGT	110 mM K$^+$	Parallel	Quercetin (Flavonoid)	Organic Heterocycle	P/nP/Hc/HA/HD	Tetrad π-stacking and T-T Intercalation
5LIG	2016	Solution NMR	TAGGGAGGGTAGGGAGGGT	120 mM K$^+$	Parallel	DAOTA-M2 (Triangulenium Derviative)	Organic Heterocycle w/ saturated chain substituents (charged)	P/nP/Hq/HD/1+	Tetrad π-stacking
5MVB	2017	Solution NMR	(TTAGGG)$_4$TT	95 mM K$^+$	Hybrid-2	Au-oxo6 (Gold(III) oxo complex)	Metal-Organic (Square-Planar)	P/nP/Hc/HD/2+	Tetrad π-stacking
5W77	2018	Solution NMR	TGA(GGGT)$_2$A(GGGT)$_2$AA	50 mM K$^+$	Parallel	DC-34 (Linked heterocycles with p-trifluoro-methylbenzene)	Organic Heterocycle	P/nP/Hc/HA/HD	Tetrad π-stacking
5Z80	2018	Solution NMR	AAAGGG(TTAG3)$_3$AA	95 mM K$^+$	Hybrid-1	Pt(II)-based Tripod	Metal-Organic (Square-planar/charged)	P/nP/Hc/HA/6+	Tetrad π-stacking
5Z8F	2018	Solution NMR	AAAGGG(TTAG3)$_3$AA	95 mM K$^+$	Hybrid-1	Pt(II)-based Tripod	Metal-Organic (Square-planar/charged)	P/nP/Hc/HA/HD/6+	Tetrad π-stacking and H-bond with O1P
6CCW	2018	Solution NMR	(TTAGGG)$_4$TT	100 mM K$^+$	Hybrid-2	Berberine (benzyltetrahydroisoquinoline ...)	Organic Heterocycle (Non-planar)	P/nP/Hc/HD	Tetrad π-stacking/Side-on nucleobase H-bonding

8.2 Chapter 2

Table A2.1 – *Crystal data and data refinement statistics for the four small molecule crystals containing; rac-[Ru(TAP)₂(11-CN-dppz)]2PF₆, Δ-[Ru(TAP)₂(11,12-CN-dppz)]·(PF₆)₂, rac-[Ru(phen)₂(10-NO₂-dppz)]·(PF₆)₂, or rac-[Ru(phen)₂(11-NO₂-dppz)]·(PF₆)₂.*

Crystal Data	*rac*-[Ru(TAP)₂(11-CN-dppz)]·(PF₆)₂	Δ-[Ru(TAP)₂(11,12-CN-dppz)]·(PF₆)₂
Empirical formula	$RuC_{39}N_{13}H_{21}P_2F_{12}$	$RuC_{40}N_{14}H_{20}P_2F_{12}$
Formula weight	1062.67	1087.71
Temperature (K)	150(2)	150(2)
Crystal system	Triclinic	Tetragonal
Space group	*P -1*	*P* $4_3 2_1 2$
a (Å)	13.2508(5)	16.3465(2)
b (Å)	17.5538(5)	16.3465(2)
c (Å)	23.2434(10)	15.5227(2)
α (°)	83.429(3)	90
β (°)	78.654(4)	90
γ (°)	89.764(3)	90
Volume (Å)	5265.0(4)	4147.79(9)
Z	4	4
ρ_{calcd} (g cm⁻³)	1.366	1.742
μ (mm⁻¹)	2.679	4.751
F_{000}	1560	2160
Radiation wavelength (Å)	1.54184	1.54184
R_{int}	0.0511	0.0499
Goodness of fit on F^2	1.021	1.070
Data Refinement		
Nº reflections	37534	32013
Nº unique reflections	19553	4323
Largest diff. peak and hole (eÅ⁻³)	1.1 and -0.7	1.2 and -0.5
Final *R* indices ($I > 2\sigma(I)$)	$R_1 = 0.0942$	$R_1 = 0.0560$
R indices (all data)	w$R_2 = 0.3024$	w$R_2 = 0.1617$

Crystal Data	*rac*-[Ru(phen)₂(10-NO₂-dppz)]·(PF₆)₂	*rac*-[Ru(phen)₂(11-NO₂-dppz)]·(PF₆)₂
Empirical formula	$RuC_{42}N_9O_2H_{25}P_2F_{12}$	$RuC_{42}N_9O_2H_{25}P_2F_{12}$
Formula weight	1078.72	1078.72
Temperature (K)	100(2)	100(2)
Crystal system	Triclinic	Triclinic
Space group	*P -1*	*P -1*
a (Å)	12.2726(6)	9.1834(2)
b (Å)	13.8702(4)	12.7142(2)
c (Å)	14.7863(2)	21.8132(3)
α (°)	95.302(2)	92.351(10)
β (°)	96.781(3)	99.504(2)
γ (°)	116.098(4)	108.925(2)
Volume (Å)	2214.94(15)	2363.92(8)
Z	2	2
ρ_{calcd} (g cm⁻³)	1.617	1.573
μ (mm⁻¹)	4.442	4.192
F_{000}	1076	1120
Radiation wavelength (Å)	1.54184	1.54184
R_{int}	0.0535	0.0525
Goodness of fit on F^2	1.051	1.050
Data Refinement		
Nº reflections	33180	35484
Nº unique reflections	8909	9529
Largest diff. peak and hole (eÅ⁻³)	1.2 and -2.5	0.9 and -1.1
Final *R* indices ($I > 2\sigma(I)$)	$R_1 = 0.0589$	$R_1 = 0.0511$
R indices (all data)	w$R_2 = 0.1676$	w$R_2 = 0.1428$
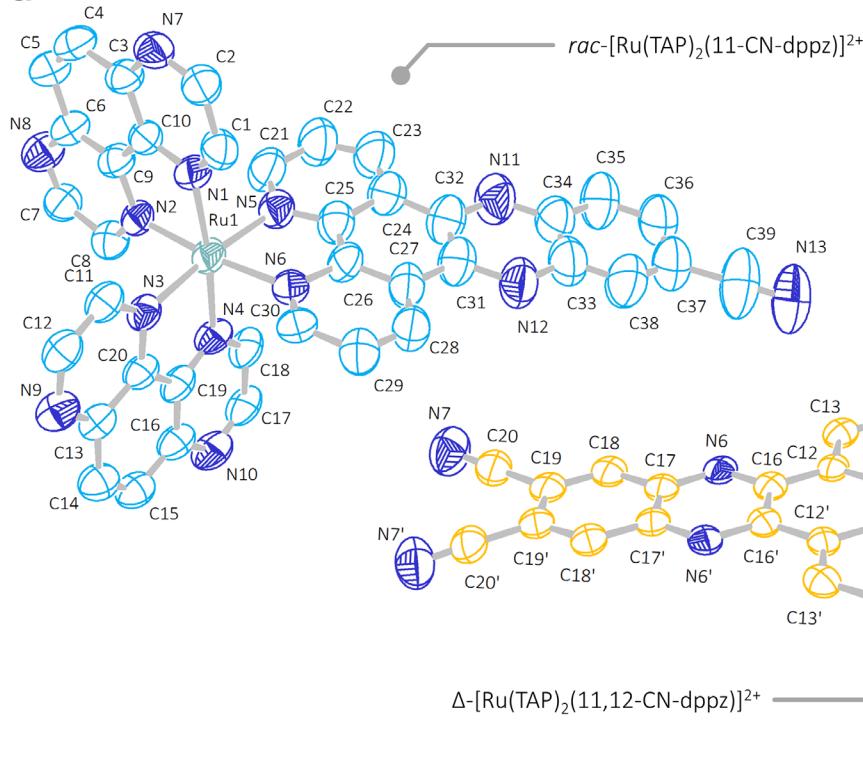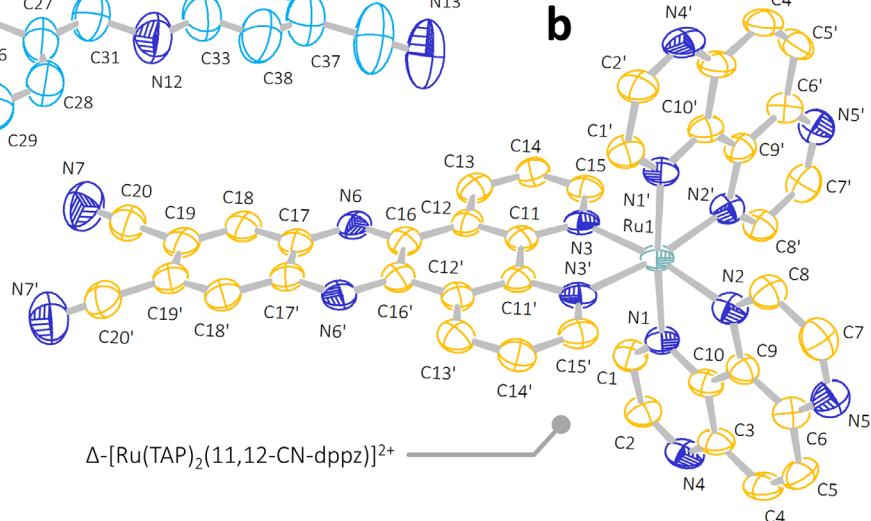
Figure A2.1 _– Crystal structures of two nitrile derivatives; (a) rac-[Ru(TAP)₂(11-CN-dppz)]·(PF₆)₂ and (b) Δ-[Ru(TAP)₂(11,12-CN-dppz)]·(PF₆)₂. Thermal ellipsoids shown at 50 % density probability._

Table A2.2 _– Selected bond lengths, angles, and torsions from the crystal structures of two nitrile derivatives; (a) rac-[Ru(TAP)₂(11-CN-dppz)]·(PF₆)₂ and (b) Δ-[Ru(TAP)₂(11,12-CN-dppz)]·(PF₆)₂._

a — rac-[Ru(TAP)$_2$(11-CN-dppz)]·(PF$_6$)$_2$

Selected Bond Lengths

Ru-N1	2.074(7)	Ru-N4	2.051(8)	C37-C39	1.50(3)
Ru-N2	2.035(8)	Ru-N5	2.080(8)	C39-N13	1.12(2)
Ru-N3	2.063(7)	Ru-N6	2.071(1)		

Selected Bond Angles

N1-Ru-N2	80.3(3)	N2-Ru-N4	95.6(3)	N3-Ru-N6	95.8(3)
N1-Ru-N3	95.4(3)	N2-Ru-N5	93.7(3)	N4-Ru-N5	95.6(3)
N1-Ru-N4	173.7(3)	N2-Ru-N6	173.5(3)	N4-Ru-N6	87.3(3)
N1-Ru-N5	89.4(3)	N3-Ru-N4	79.7(3)	N5-Ru-N6	80.2(3)
N1-Ru-N6	97.3(3)	N3-Ru-N5	174.1(3)	C37-C39-N13	175.0(2)
N2-Ru-N3	90.5(3)	N4-Ru-N6	87.3(3)		

Selected Torsion Angles

N1-C10-C9-N2	3(1)	N3-Ru-N4-C19	-1.6(6)	N6-Ru-N5-C25	3.8(8)
N1-Ru-N2-C9	2.2(6)	N4-Ru-N3-C20	1.4(6)	C38-C37-C39-N13	139.0(23)
N2-Ru-N1-C10	-0.3(7)	N5-C25-C26-N6	-1.0(2)	C36-C37-C39-N13	-46(24)
N3-C20-C19-N4	0(1)	N5-Ru-N6-C26	-4.0(7)		

b — Δ-[Ru(TAP)$_2$(11,12-CN-dppz)]·(PF$_6$)$_2$

Selected Bond Lengths

Ru-N1	2.048(6)	Ru-N1'	2.048(6)	C19-C20	1.443(15)
Ru-N2	2.071(7)	Ru-N2'	2.071(7)	C20-N7	1.139(15)
Ru-N3	2.056(6)	Ru-N3'	2.056(6)		

Selected Bond Angles

N1-Ru-N2	80.3(3)	N2-Ru-N1'	95.9(3)	N3-Ru-N3'	80.0(4)
N1-Ru-N3	93.0(3)	N2-Ru-N2'	90.1(4)	N1'-Ru-N2'	80.3(3)
N1-Ru-N1'	174.7(8)	N2-Ru-N3'	95.3(3)	N1'-Ru-N3'	93.0(3)
N1-Ru-N2'	95.9(3)	N3-Ru-N1'	91.1(3)	N2'-Ru-N3'	171.8(3)
N1-Ru-N3'	91.1(3)	N3-Ru-N2'	95.3(3)	C19-C20-N7	175.3(2)
N2-Ru-N3	171.8(3)				

Selected Torsion Angles

|					
N1-C10-C9-N2	-1.9(10)	N3-Ru-N3'-C11'	-1.6(6)	C18-C19-C20-N7	-129(10)
N1-Ru-N2-C9	-0.9(6)	N4-Ru-N3-C20	1.4(6)	C19'-C19-C20-N7	49(11)
N2-Ru-N1-C10	-0.1(6)	N3'-Ru-N3-C11	0.2(4)	C20'-C19'-C19-C2	6.2(9)
N3-C11-C11'-N3'	1(1)				
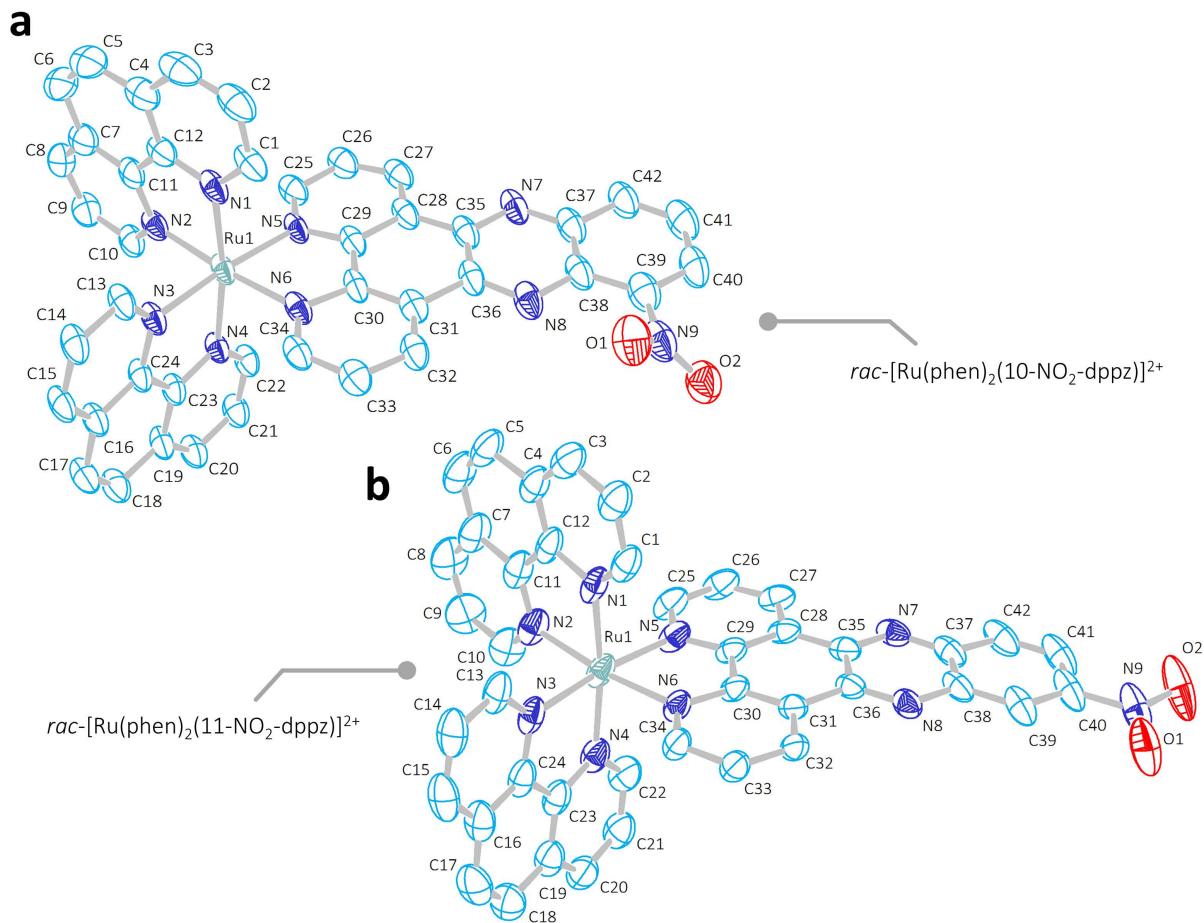
Figure A2.2 – Crystal structures of two nitro derivatives; (a) rac-[Ru(phen)₂(10-NO₂-dppz)]·(PF₆)₂ and (b) rac-[Ru(phen)₂(11-NO₂-dppz)]·(PF₆)₂. Thermal ellipsoids shown at 50 % density probability.

Table A2.3 – Selected bond lengths, angles, and torsions from the crystal structures of the two nitro derivatives; (a) rac-[Ru(phen)₂(10-NO₂-dppz)]·(PF₆)₂ and (b) rac-[Ru(phen)₂(11-NO₂-dppz)]·(PF₆)₂.

a   *rac*-[Ru(phen)₂(10-NO₂-dppz)]·(PF₆)₂

Selected Bond Lengths

Ru-N1	2.062(5)	Ru-N4	2.064(4)	C39-N9	1.448(7)
Ru-N2	2.075(4)	Ru-N5	2.079(4)	N9-O1	1.211(8)
Ru-N3	2.076(4)	Ru-N6	2.075(4)	N9-O2	1.249(9)

Selected Bond Angles

N1-Ru-N2	79.9(2)	N2-Ru-N4	93.8(2)	N4-Ru-N5	99.1(2)
N1-Ru-N3	94.1(2)	N2-Ru-N5	96.8(2)	N4-Ru-N6	89.9(2)
N1-Ru-N4	171.1(2)	N2-Ru-N6	175.3(2)	N5-Ru-N6	79.8(2)
N1-Ru-N5	87.8(2)	N3-Ru-N4	79.5(2)	C39-N9-O1	119.5(5)
N1-Ru-N6	96.7(2)	N3-Ru-N5	173.6(2)	C39-N9-O2	118.3(5)
N2-Ru-N3	89.6(2)	N3-Ru-N6	93.9(2)	O1-N9-O2	122.2(5)

Selected Torsion Angles

N1-C12-C11-N2	0.9(7)	N4-Ru-N3-C24	-0.1(3)	C38-C39-N9-O1	-41.6(9)
N1-Ru-N2-C11	-8.7(4)	N5-C29-C30-N6	-3.7(7)	C38-C39-N9-O2	140.8(6)
N2-Ru-N1-C12	9.0(4)	N5-Ru-N6-C30	2.3(3)	C40-C39-N9-O1	137.4(6)
N3-C24-C23-N4	3.3(7)	N6-Ru-N5-C29	-4.2(3)	C40-C39-N9-O2	-40.2(9)
N3-Ru-N4-C23	1.8(3)				

b   *rac*-[Ru(phen)₂(11-NO₂-dppz)]·(PF₆)₂

Selected Bond Lengths

Ru-N1	2.060(4)	Ru-N4	2.072(4)	C40-N9	1.472(6)
Ru-N2	2.083(3)	Ru-N5	2.072(4)	N9-O1	1.213(6)
Ru-N3	2.067(4)	Ru-N6	2.075(2)	N9-O2	1.224(6)

Selected Bond Angles

N1-Ru-N2	80.2(1)	N2-Ru-N4	96.6(1)	N4-Ru-N5	97.0(1)
N1-Ru-N3	94.0(1)	N2-Ru-N5	95.3(1)	N4-Ru-N6	88.6(1)
N1-Ru-N4	172.6(1)	N2-Ru-N6	173.0(1)	N5-Ru-N6	79.3(1)
N1-Ru-N5	89.9(1)	N3-Ru-N4	79.4(1)	C40-N9-O1	118.4(4)
N1-Ru-N6	95.2(1)	N3-Ru-N5	172.8(1)	C40-N9-O2	117.8(4)
N2-Ru-N3	91.4(1)	N3-Ru-N6	94.2(1)	O1-N9-O2	123.7(5)

Selected Torsion Angles

N1-C12-C11-N2	1.2(6)	N4-Ru-N3-C24	5.8(3)	C39-C40-N9-O1	9.6(7)
N1-Ru-N2-C11	-1.3(3)	N5-C29-C30-N6	-0.4(5)	C39-C40-N9-O2	-169.9(5)
N2-Ru-N1-C12	1.9(3)	N5-Ru-N6-C30	0.9(3)	C41-C40-N9-O1	-172.1(5)
N3-C24-C23-N4	1.4(6)	N6-Ru-N5-C29	-1.2(3)	C41-C40-N9-O2	8.3(7)
N3-Ru-N4-C23	-5.0(3)				
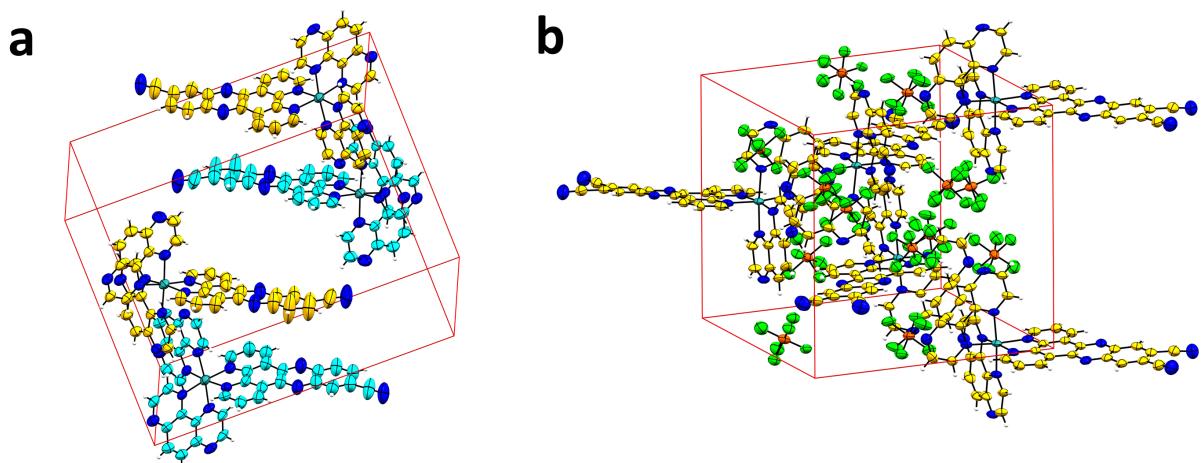
Figure A2.3 – *Perspective view of the unit cells of the crystal structures of the nitrile derivatives (a) rac-[Ru(TAP)$_2$(11-CN-dppz)]·(PF$_6$)$_2$ in SG P-1; (b) Δ-[Ru(TAP)$_2$(11,12-CN-dppz)]·(PF$_6$)$_2$ in SG P4$_3$2$_1$2.*
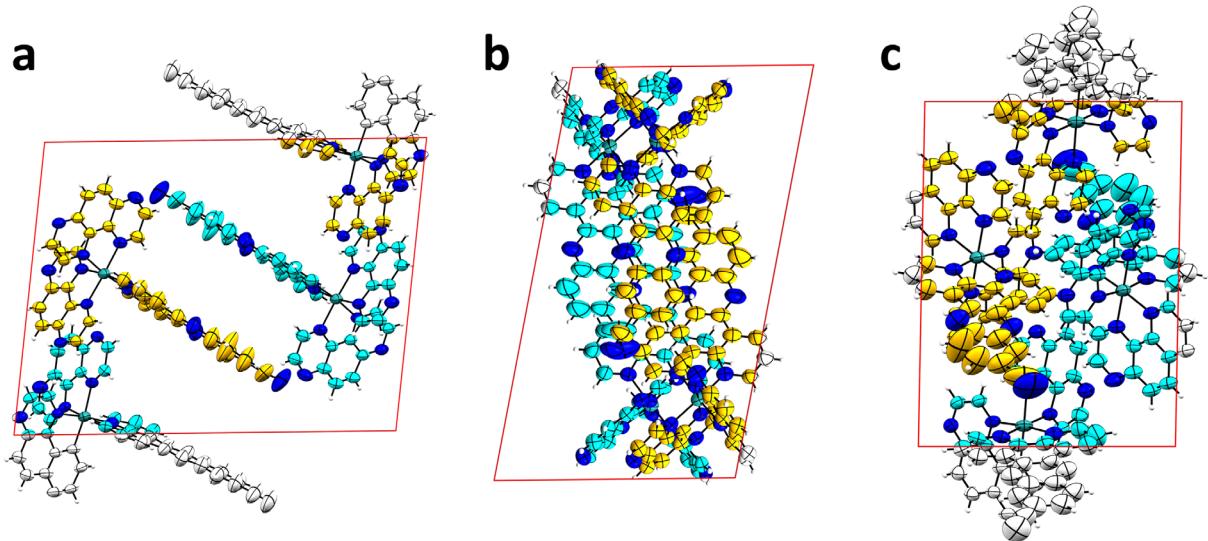
Figure A2.4 – *Unit cells of the crystal structure of rac-[Ru(TAP)$_2$(11-CN-dppz)]·(PF$_6$)$_2$ looking down (a) the a axis; (b) the b axis; and (c) the c axis.*
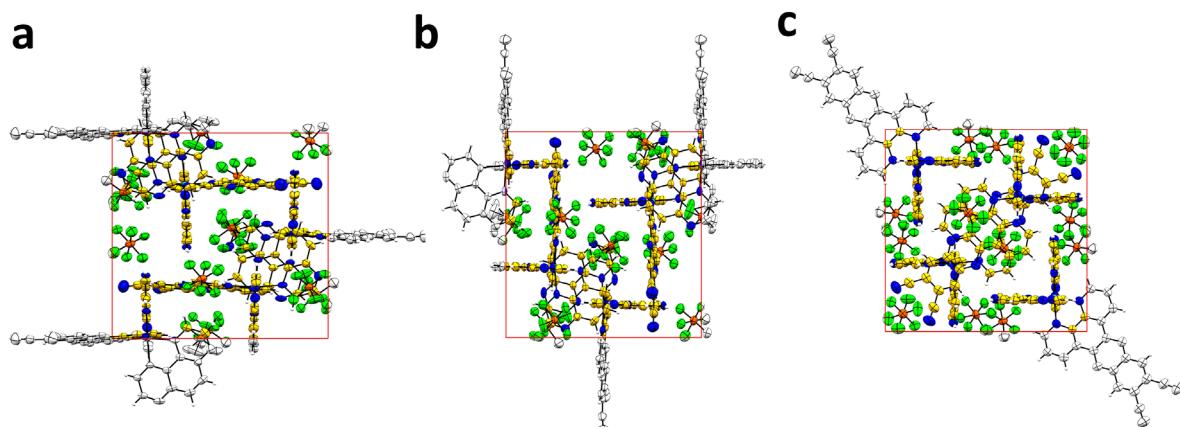
Figure A2.5 – *Unit cells of the crystal structure of Δ-[Ru(TAP)$_2$(11,12-CN-dppz)]·(PF$_6$)$_2$ looking down (a) the a axis; (b) the b axis; and (c) the c axis.*
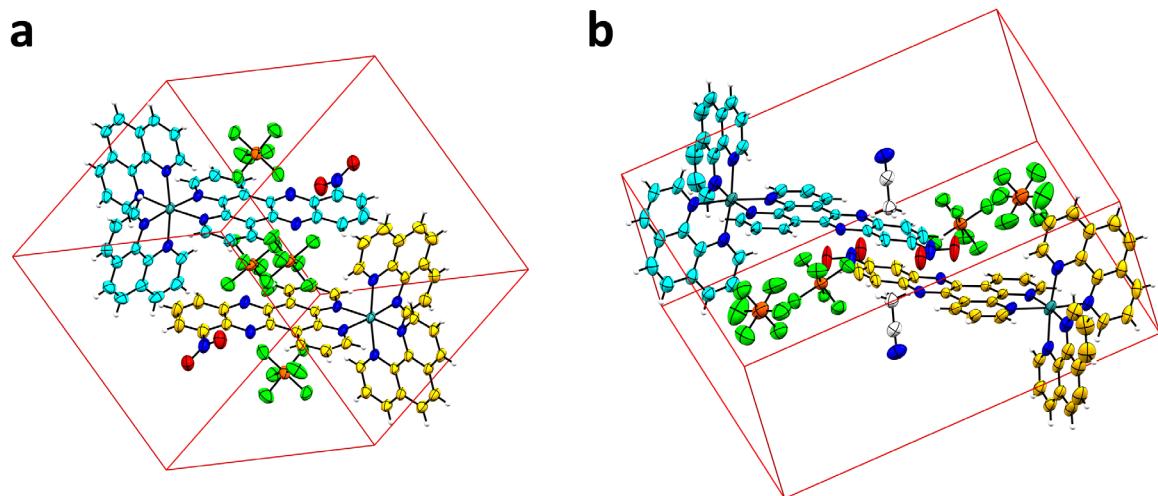
Figure A2.6 *– Perspective view of the unit cells of the crystal structures of the nitro derivatives (a) rac-[Ru(phen)$_2$(10-NO$_2$-dppz)]·(PF$_6$)$_2$ in SG P-1; (b) rac-[Ru(phen)$_2$(11-NO$_2$-dppz)]·(PF$_6$)$_2$ in SG P-1.*
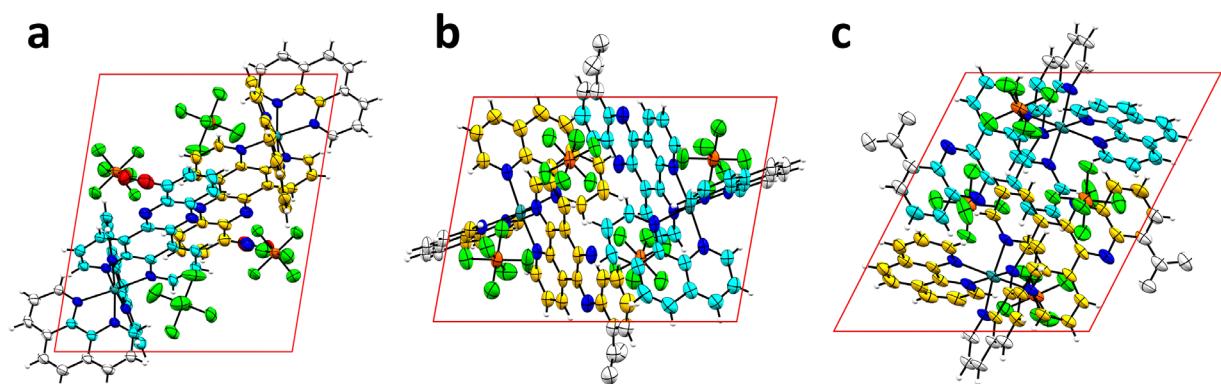
Figure A2.7 *– Unit cell of the crystal structure of rac-[Ru(phen)$_2$(10-NO$_2$-dppz)]·(PF$_6$)$_2$ looking down (a) the a axis; (b) the b axis; and (c) the c axis.*
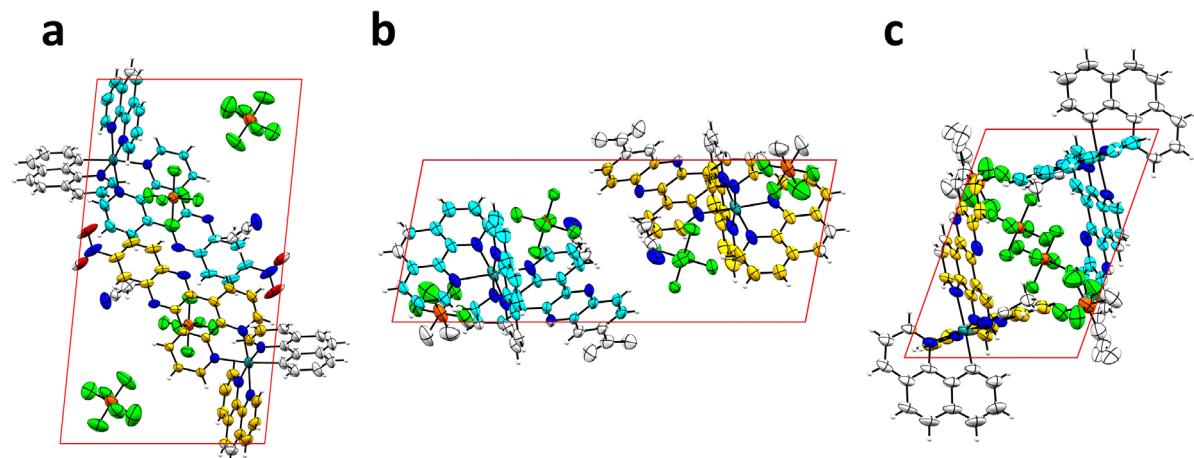
Figure A2.8 *– Unit cell of the crystal structure of rac-[Ru(phen)$_2$(11-NO$_2$-dppz)]·(PF$_6$)$_2$ looking down (a) the a axis; (b) the b axis; and (c) the c axis.*

Table A2.4 – *Crystallisation, data collection, and processing parameters/refinement results of the crystal structures of Δ/Λ-[Ru(TAP)₂(dppz)]·Cl₂ collected, solved, and refined in a manner analogous to macromolecular techniques.*

Crystallisation Parameters	Δ-[Ru(TAP)$_2$(dppz)]·Cl$_2$	Λ-[Ru(TAP)$_2$(dppz)]·Cl$_2$
Crystal Morphology	Rod	Rod
Growth Temperature (K)	291	291
Crystal Size (μm)	30x30x130	30x30x130
Growth Time	12 weeks	12 weeks

Data Collection

Beamline	I03	I03
X-Ray Wavelength (Å)	0.557	0.557
Transmission (%)	49.99	35.01
Beamsize (μm)	80x20	80x20
Exposure Time (s)	0.05	0.05
Nº Images/Oscillation (º)	7200/0.1	7200/0.1
Space Group	$I\,4_1\,2\,2$	$I\,4_1\,2\,2$
Cell Dimensions a, b, c (Å); α, β, γ (º)	37.05, 37.05, 52.78; 90, 90, 90	37.07, 37.07, 52.79; 90, 90, 90

Data Processing

Resolution (Å)	26.39 - 1.42 (1.47 - 1.42)*	26.36 - 1.46 (1.49 - 1.46)
R_{merge}	0.085 (2.831)	0.079 (2.273)
R_{meas}	0.086 (2.858)	0.080 (2.297)
R_{pim}	0.012 (0.387)	0.012 (0.328)
Nº Observations	179842 (9362)	165341 (8360)
Nº Unique Observations	3674 (174)	3420 (174)
I/σI	31.5 (2.2)	30.9 (2.3)
CC$_{1/2}$	0.999 (0.906)	1.000 (0.845)
Completeness (%)	99.6 (100.00)	100.0 (99.4)
Multiplicity	49.0 (53.8)	48.3 (48.0)
Mid-slope of anom normal probability	1.903	1.811

* Outer Shell Statistics Shown in Parentheses

Refinement

Phase Solution Method	SAD	SAD
Resolution	32.3 (1.88)	26.2 (1.46)
No. of Reflections	3654 (357)	3387 (320)
R_{work}/R_{free}	0.1702/0.2105	0.1635/0.1865
No. of Atoms		
DNA	-	-
Metal Complex	106	106
Water	12	13
Average B Factors (Å^2)		
DNA	-	-
Metal Complex	30.94	29.89
Water	32.38	45.09
rmsd		
Bond Lengths (Å)	0.025	0.015
Bond Angles (o)	1.95	2.41

● +[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$　　● +[Ru(TAP)$_2$(11,12-CN-dppz)]$^{2+}$　　● +[Ru(phen)$_2$(10-NO$_2$-dppz)]$^{2+}$　　● +[Ru(phen)$_2$(11-NO$_2$-dppz)]$^{2+}$
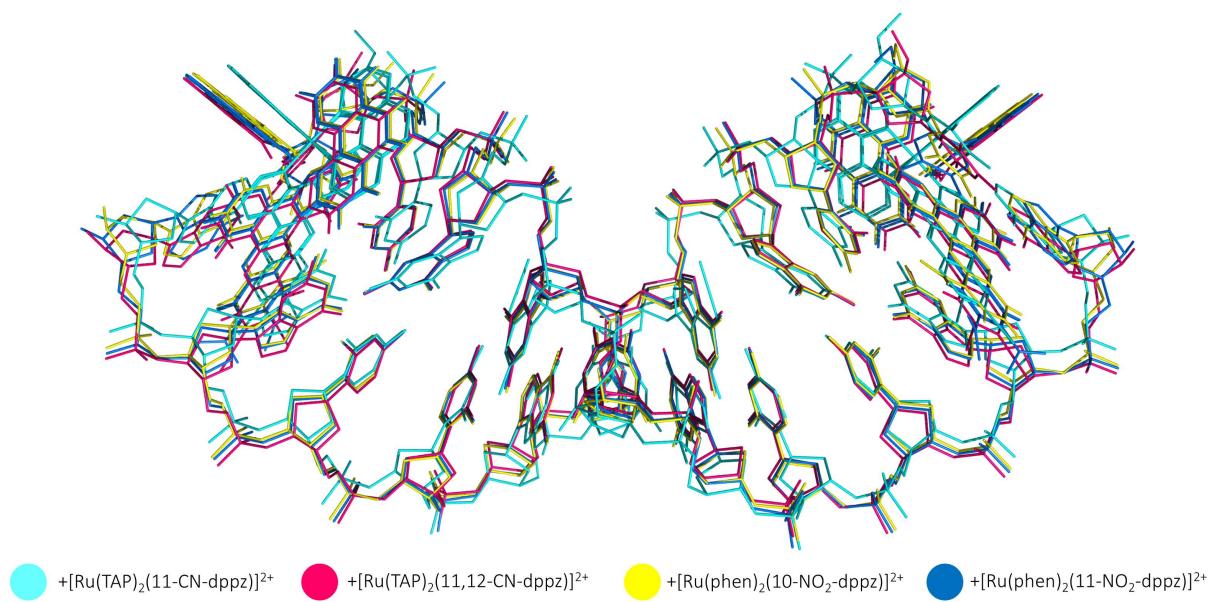
Figure A2.9 *– Superimposition of the four structures containing complexes with electron withdrawing moieties to highlight the similarities in atomic*

Table A2.5 – *Conformational analysis and local base pair parameters for structure 5NBE.*

[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ with d(TCGGCGCCGA)

Local base-pair parameters

| base pair	shear (Å)	stretch (Å)	stagger (Å)	buckle (°)	propeller (°)	opening (°)
T$_1$·A$_{10}$	0.20	-0.13	-0.13	-10.02	24.45	9.54
C$_2$·G$_9$	0.08	-0.11	-0.03	-7.40	3.10	-0.33
G$_3$·C$_8$	-0.23	-0.01	0.25	24.77	-5.13	0.09
G$_4$·C$_7$	-0.17	-0.10	-0.17	-6.32	5.28	-1.19
C$_5$·G$_6$	0.25	-0.10	-0.11	14.49	-7.72	-0.25
G$_6$·C$_5$	-0.25	-0.10	-0.11	-14.49	-7.72	-0.25
C$_7$·G$_4$	0.17	-0.10	-0.17	6.32	5.28	-1.19
C$_8$·G$_3$	0.23	-0.01	0.25	-24.77	-5.13	0.09
G$_9$·C$_2$	-0.08	-0.11	-0.03	7.40	3.10	-0.33
A$_{10}$·T$_1$	-0.20	-0.13	-0.13	10.02	24.45	9.54

Local base step parameters

base step	shift (Å)	slide (Å)	rise (Å)	tilt (°)	roll (°)	twist (°)
T$_1$/C$_2$	-2.67	0.24	6.70	7.8	-31.6	19.4
C$_2$/G$_3$	-0.25	1.75	2.74	15.0	2.9	22.5
G$_3$/G$_4$	0.08	1.15	4.94	-14.7	55.4	14.6
G$_4$/C$_5$	-0.51	0.46	2.91	10.5	-6.9	35.3
C$_5$/G$_6$	0.20	1.40	5.19	-12.7	49.5	14.2
G$_6$/C$_7$	0.88	0.15	2.90	10.1	3.1	34.1
C$_7$/C$_8$	0.29	0.48	5.29	-14.6	42.1	23.4
C$_8$/G$_9$	-0.14	1.24	2.69	16.9	7.6	23.0
G$_9$/A$_{10}$	1.02	0.76	6.96	-12.4	-6.4	28.8

Local base-pair step parameters

base pair step	shift (Å)	slide (Å)	rise (Å)	tilt (°)	roll (°)	twist (°)
T$_1$C$_2$:G$_9$A$_{10}$	-1.84	0.47	6.89	10.6	-19.1	24.5
C$_2$G$_3$:C$_8$G$_9$	-0.05	1.49	2.73	-0.9	5.3	22.5
G$_3$G$_4$:C$_7$C$_8$	-0.09	0.82	5.13	0.0	48.8	19.1
G$_4$C$_5$:G$_6$C$_7$	-0.70	0.30	2.90	0.2	-2.0	34.5
C$_5$G$_6$:C$_5$G$_6$	0	1.40	5.19	0.0	49.4	13.6
G$_6$C$_7$:G$_4$C$_5$	0.70	0.30	2.90	-0.2	-2.0	34.5
C$_7$C$_8$:G$_3$G$_4$	0.09	0.82	5.13	0	48.8	19.1
C$_8$G$_9$:C$_2$G$_3$	0.05	1.49	2.73	0.9	5.3	22.5
G$_9$A$_{10}$:T$_1$C$_2$	1.84	0.47	6.89	-10.6	-19.1	24.5

Local base step torsional angles

step	δ	ε	ζ	α	β	γ	δ	χ	NtC	CANA
T$_1$/C$_2$	150.6	281.7	201.5	87.1	182.1	188.9	96.5	236.8	NANT	NAN
C$_2$/G$_3$	96.5	181.4	277.0	290.9	170.0	51.3	88.7	250.4	AA02	AAA
G$_3$/G$_4$	88.7	212.5	287.0	296.7	185.7	58.1	147.5	250.3	AB01	A-B
G$_4$/C$_5$	147.5	203.1	236.7	291.4	167.5	36.3	88.8	261.4	BA01	B-A
C$_5$/G$_6$	88.8	215.2	284.9	294.4	187.2	55.7	137.3	245.8	AB01	A-B
G$_6$/C$_7$	137.3	187.4	264.3	291.6	169.2	39.8	78.0	267.6	BA05	B-A
C$_7$/C$_8$	78.0	209.4	274.5	300.4	172.2	47.4	85.3	243.2	AA02	AAA
C$_8$/G$_9$	85.3	201.8	278.1	287.4	169.0	56.4	93.0	226.8	AA02	AAA
G$_9$/A$_{10}$	93.0	191.1	275.9	300.0	173.4	63.2	80.4	241.3	AA02	AAA

Λ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$

T$_1$
C$_2$
G$_3$
G$_4$
C$_5$
G$_6$
C$_7$
C$_8$
G$_9$
A$_{10}$
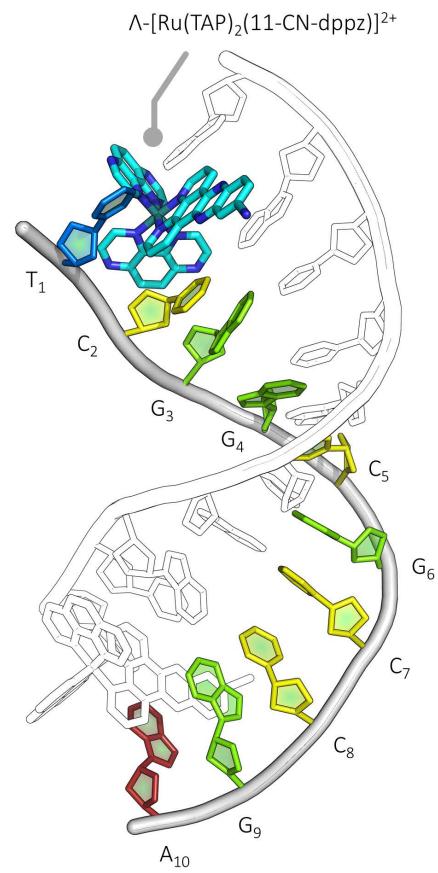
Table A2.6 – *Conformational analysis and local base pair parameters for structure 6R6D.*

[Ru(TAP)$_2$(11,12-CN-dppz)]$^{2+}$ with d(TCGGCGCCGA)
Local base-pair parameters

base pair	shear (Å)	stretch (Å)	stagger (Å)	buckle (°)	propeller (°)	opening (°)
T$_1$·A$_{10}$	0.18	-0.15	-0.18	-14.7	21.0	9.0
C$_2$·G$_9$	0.06	-0.08	-0.17	-4.8	4.4	1.5
G$_3$·C$_8$	-0.27	-0.03	0.36	26.0	-5.7	0.9
G$_4$·C$_7$	-0.21	-0.15	-0.04	-9.3	6.5	0.2
C$_5$·G$_6$	0.32	-0.11	0.05	8.4	-2.5	1.0
G$_6$·C$_5$	-0.32	-0.11	0.05	-8.4	-2.5	1.0
C$_7$·G$_4$	0.21	-0.15	-0.04	9.3	6.5	0.2
C$_8$·G$_3$	0.27	-0.03	0.36	-26.0	-5.7	0.9
G$_9$·C$_2$	-0.06	-0.08	-0.17	4.8	4.4	1.5
A$_{10}$·T$_1$	-0.18	-0.15	-0.18	14.7	21.0	9.0

Local base step parameters

base step	shift (Å)	slide (Å)	rise (Å)	tilt (°)	roll (°)	twist (°)
T$_1$/C$_2$	-2.47	0.50	6.52	6.2	-29.2	16.9
C$_2$/G$_3$	-0.14	1.88	2.78	13.3	5.6	20.5
G$_3$/G$_4$	0.03	1.02	5.11	-17.8	56.0	13.0
G$_4$/C$_5$	-0.54	0.49	3.05	7.6	-5.9	39.4
C$_5$/G$_6$	-0.08	1.22	4.51	-7.2	36.8	17.6
G$_6$/C$_7$	0.87	0.36	3.02	8.0	3.0	38.6
C$_7$/C$_8$	0.42	0.43	5.38	-15.1	42.3	20.5
C$_8$/G$_9$	-0.26	1.34	2.56	17.4	10.7	22.9
G$_9$/A$_{10}$	0.95	1.15	6.77	-3.0	-8.8	26.0

Local base-pair step parameters

base pair step	shift (Å)	slide (Å)	rise (Å)	tilt (°)	roll (°)	twist (°)
T$_1$C$_2$:G$_9$A$_{10}$	-1.70	0.79	6.71	5.0	-18.9	21.6
C$_2$G$_3$:C$_8$G$_9$	0.06	1.59	2.69	-2.1	8.3	21.4
G$_3$G$_4$:C$_7$C$_8$	-0.18	0.73	5.26	-1.4	49.0	16.9
G$_4$C$_5$:G$_6$C$_7$	-0.70	0.43	3.03	-0.2	-1.4	38.8
C$_5$G$_6$:C$_5$G$_6$	0	1.22	4.51	0	36.7	17.5
G$_6$C$_7$:G$_4$C$_5$	0.70	0.43	3.03	0.2	-1.4	38.8
C$_7$C$_8$:G$_3$G$_4$	0.18	0.73	5.26	1.4	49.0	16.9
C$_8$G$_9$:C$_2$G$_3$	-0.06	1.59	2.69	2.1	8.3	21.4
G$_9$A$_{10}$:T$_1$C$_2$	1.70	0.79	6.71	-5.0	-18.9	21.6

Local base step torsional angles

step	δ	ε	ζ	α	β	γ	δ	χ	NtC	CANA
T$_1$/C$_2$	147.6	212.4	237.2	261.8	189.8	70.4	81.9	239.6	NANT	NAN
C$_2$/G$_3$	81.9	172.3	276.0	301.2	170.6	48.6	92.0	255.5	AA02	AAA
G$_3$/G$_4$	92.0	216.1	286.4	292.6	180.5	62.6	139.4	253.3	NANT	NAN
G$_4$/C$_5$	139.4	189.5	239.7	304.1	178.8	28.8	96.0	250.9	NANT	NAN
C$_5$/G$_6$	96.0	202.9	283.2	301.3	182.5	53.1	145.5	245.9	AB01	A-B
G$_6$/C$_7$	145.5	189.4	255.8	298.0	173.5	34.3	87.6	260.2	BA05	B-A
C$_7$/C$_8$	87.6	213.0	282.1	292.8	174.8	50.6	90.0	242.3	AA02	AAA
C$_8$/G$_9$	90.0	196.8	272.5	279.1	169.1	67.3	101.1	229.1	AA02	AAA
G$_9$/A$_{10}$	101.1	193.4	274.0	278.6	177.6	100.1	102.9	241.0	NANT	NAN

Λ-[Ru(TAP)$_2$(11,12-CN-dppz)]$^{2+}$

213
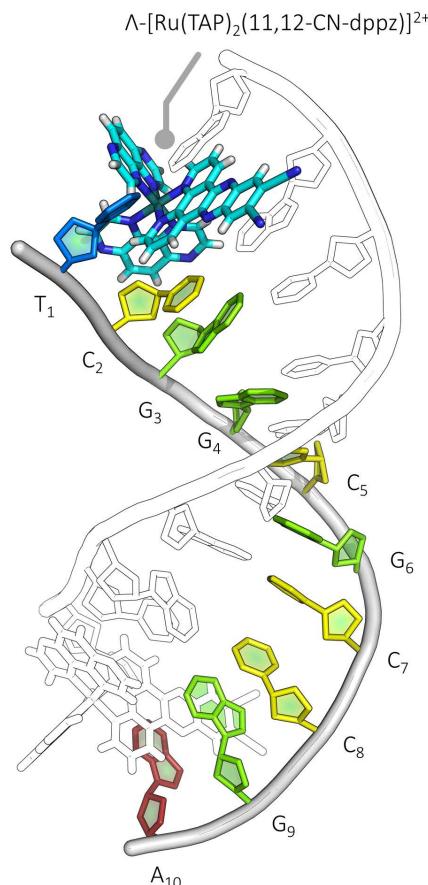
Table A2.7 – *Conformational analysis and local base pair parameters for structure 6RSO.*

$[Ru(phen)_2(10\text{-}NO_2\text{-}dppz)]^{2+}$ with d(TCGGCGCCGA)

Local base-pair parameters

base pair	shear (Å)	stretch (Å)	stagger (Å)	buckle (°)	propeller (°)	opening (°)
$T_1 \cdot A_{10}$	0.18	-0.23	0.12	2.4	-11.8	11.0
$C_2 \cdot G_9$	0.24	-0.18	0.03	-13.9	9.0	2.8
$G_3 \cdot C_8$	-0.32	-0.15	0.50	26.4	-0.7	-3.8
$G_4 \cdot C_7$	-0.20	-0.14	-0.25	-14.4	1.1	-2.0
$C_5 \cdot G_6$	0.38	-0.07	0.25	5.0	-6.5	1.6
$G_6 \cdot C_5$	-0.50	-0.12	0.23	-4.7	-7.0	0.8
$C_7 \cdot G_4$	0.29	-0.12	-0.14	10.8	1.8	-0.4
$C_8 \cdot G_3$	0.44	-0.16	0.54	-27.1	0.1	-2.1
$G_9 \cdot C_2$	-0.32	-0.15	0	12.2	9.0	4.0
$A_{10} \cdot T_1$	-0.56	-0.02	-0.16	1.6	-1.2	17.0

Local base step parameters

base step	shift (Å)	slide (Å)	rise (Å)	tilt (°)	roll (°)	twist (°)
A - T_1/C_2	0.49	0.29	7.03	-1.4	14.6	16.9
A - C_2/G_3	-0.52	1.64	2.58	17.1	1.8	19.3
A - G_3/G_4	0.14	0.88	5.31	-18.7	53.6	14.8
A - G_4/C_5	-0.37	0.62	3.16	6.1	-4.7	41.2
A - C_5/G_6	-0.08	1.35	4.30	-2.8	35.0	14.6
A - G_6/C_7	1.10	0.66	2.87	10.2	5.6	39.4
A - C_7/C_8	0.40	0.20	5.66	-18.1	50.7	20.8
A - C_8/G_9	-0.40	1.37	2.29	19.8	6.2	25.4
A - G_9/A_{10}	0.34	1.17	6.94	-9.4	-6.3	24.7
B - T_1/C_2	-0.79	0.36	7.18	-2.6	1.2	11.3
B - C_2/G_3	-0.61	1.69	2.61	16.8	1.1	18.6
B - G_3/G_4	0.13	0.99	5.27	-18.2	53.5	14.6
B - G_4/C_5	-0.27	0.64	3.26	6.1	-4.8	40.6
B - C_5/G_6	0.03	1.41	4.29	-3.5	35.5	15.5
B - G_6/C_7	0.96	0.67	2.63	13.9	5.8	37.7
B - C_7/C_8	0.38	0.30	5.82	-21.6	51.0	19.0
B - C_8/G_9	-0.31	1.32	2.30	20.5	8.3	26.6
B - G_9/A_{10}	0.08	1.14	7.07	-13.1	-3.3	24.8

Local base-pair step parameters

base pair step	shift (Å)	slide (Å)	rise (Å)	tilt (°)	roll (°)	twist (°)
$T_1C_2:G_9A_{10}$	0.23	0.73	7.07	5.3	5.2	20.5
$C_2G_3:C_8G_9$	-0.09	1.47	2.47	-1.8	4.9	22.2
$G_3G_4:C_7C_8$	-0.09	0.61	5.58	1.5	52.6	16.7
$G_4C_5:G_6C_7$	-0.66	0.65	2.89	-3.9	0.6	39.5
$C_5G_6:C_5G_6$	-0.05	1.38	4.28	0.4	35.2	15.0
$G_6C_7:G_4C_5$	0.68	0.66	3.07	2.0	0.4	40.0
$C_7C_8:G_3G_4$	0.11	0.61	5.48	0.0	52.2	17.6
$C_8G_9:C_2G_3$	0.09	1.51	2.48	1.5	3.8	21.2
$G_9A_{10}:T_1C_2$	0.52	0.75	7.08	-3.0	-3.1	17.9

Local base step torsional angles

step	δ	ε	ζ	α	β	γ	δ	χ	NtC	CANA
A - T_1/C_2	88.0	244.6	265.8	286.4	180.0	69.2	79.2	214.2	NANT	NAN
A - C_2/G_3	79.2	174.5	279.4	287.7	168.1	64.9	91.2	246.8	AA02	AAA
A - G_3/G_4	91.2	223.9	284.7	298.7	174.2	61.8	132.7	246.6	NANT	NAN
A - G_4/C_5	132.7	188.3	248.4	289.5	185.8	36.2	96.6	249.7	NANT	NAN
A - C_5/G_6	96.6	207.8	280.5	301.7	182.9	54.8	144.2	245.6	AB01	A-B
A - G_6/C_7	144.2	188.0	256.0	297.1	170.0	34.9	85.7	267.5	BA05	B-A
A - C_7/C_8	85.7	213.2	284.2	289.6	180.2	54.0	87.4	244.6	NANT	NAN
A - C_8/G_9	87.4	192.6	275.9	295.0	175.6	44.4	89.8	228.2	AA02	AAA
A - G_9/A_{10}	89.8	161.2	290.0	193.3	181.1	178.9	155.1	249.7	NANT	NAN
B - T_1/C_2	87.6	211.5	261.7	290.7	183.9	65.2	80.4	228.6	NANT	NAN
B - C_2/G_3	80.4	174.8	277.7	295.2	167.2	57.7	92.1	248.4	AA02	AAA
B - G_3/G_4	92.1	216.8	289.7	289.8	181.2	64.1	136.0	249.1	AB01	A-B
B - G_4/C_5	136.0	192.0	242.9	292.0	179.0	38.3	93.9	253.4	BA01	B-A
B - C_5/G_6	93.9	207.2	280.5	297.4	182.9	58.5	143.8	245.7	AB01	A-B
B - G_6/C_7	143.3	188.3	255.4	299.3	170.3	34.0	85.7	266.0	BA05	B-A
B - C_7/C_8	85.7	213.5	283.9	289.4	177.6	56.4	90.2	248.8	NANT	NAN
B - C_8/G_9	90.2	195.1	271.5	296.9	175.2	42.4	91.8	225.2	AA02	AAA
B - G_9/A_{10}	91.8	159.5	284.8	191.4	182.1	179.7	154.5	251.1	NANT	NAN

$\Lambda\text{-}[Ru(phen)_2(10\text{-}NO_2\text{-}dppz)]^{2+}$

(B)A_{10}, (B)G_9, (B)C_8, (B)C_7, (B)G_6, (A)C_5, (A)G_6, (A)C_7, (A)C_8, (A)G_9, (A)A_{10}, (B)G_4, (B)G_3, (B)C_2, (B)T_1, (A)G_4, (A)G_3, (A)C_2, (A)T_1
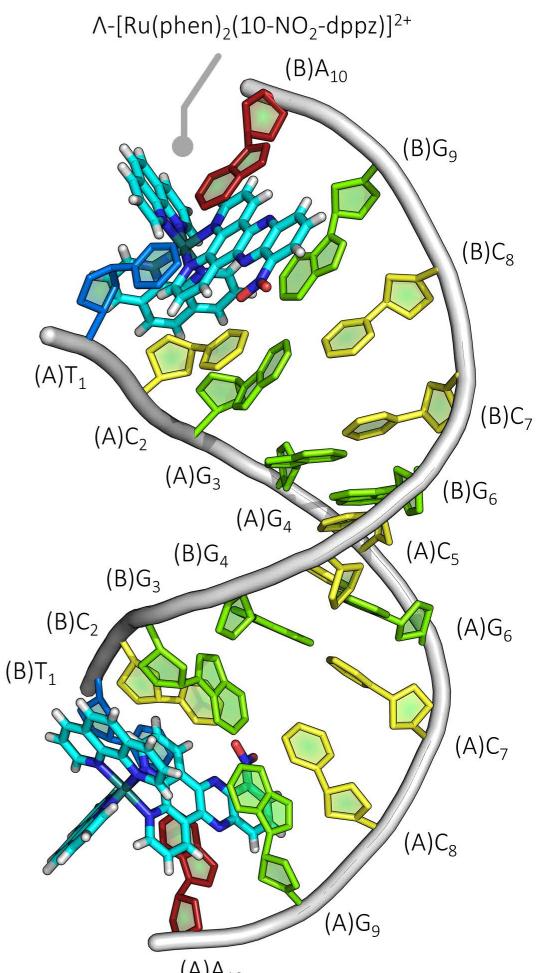
Table A2.8 *– Conformational analysis and local base pair parameters for structure 6RSP.*

$[Ru(phen)_2(11\text{-}NO_2\text{-}dppz)]^{2+}$ with d(TCGGCGCCGA)

Local base-pair parameters

base pair	shear (Å)	stretch (Å)	stagger (Å)	buckle (°)	propeller (°)	opening (°)
$T_1{\cdot}A_{10}$	1.09	1.89	-1.13	0.4	-24.9	-84.5
$C_2{\cdot}G_9$	0.23	-0.20	0.20	-14.2	8.1	0.1
$G_3{\cdot}C_8$	-0.30	-0.08	0.40	25.0	-2.4	-3.2
$G_4{\cdot}C_7$	-0.11	-0.14	-0.22	-12.9	1.1	-0.8
$C_5{\cdot}G_6$	0.47	-0.04	-0.03	8.1	-4.2	-0.3
$G_6{\cdot}C_5$	-0.36	-0.01	-0.10	-9.1	-2.7	0.6
$C_7{\cdot}G_4$	0.18	-0.14	-0.15	11.4	2.3	-0.5
$C_8{\cdot}G_3$	0.25	-0.08	0.51	-25.7	-2.1	-2.4
$G_9{\cdot}C_2$	-0.18	-0.15	0.08	12.9	6.8	1.4
$A_{10}{\cdot}T_1$	-1.08	-1.86	1.15	-0.2	22.1	87.2

Local base step parameters

base step	shift (Å)	slide (Å)	rise (Å)	tilt (°)	roll (°)	twist (°)
A - T_1/C_2	-0.02	1.14	8.13	-23.4	18.7	7.2
A - C_2/G_3	-0.42	1.67	2.54	19.4	-1.3	19.7
A - G_3/G_4	0.35	0.89	5.13	-18.3	55.4	15.1
A - G_4/C_5	-0.45	0.63	2.97	9.3	-3.4	38.9
A - C_5/G_6	-0.21	1.61	4.48	-6.8	36.8	16.0
A - G_6/C_7	0.93	0.46	2.91	9.6	3.6	37.7
A - C_7/C_8	0.24	0.33	5.57	-18.0	48.9	18.7
A - C_8/G_9	-0.35	1.35	2.43	19.2	6.4	24.0
A - G_9/A_{10}	-6.94	-3.13	-1.39	-130.4	109.4	-177.8
B - T_1/C_2	0.04	1.10	8.01	-20.9	15.7	9.2
B - C_2/G_3	-0.30	1.73	2.67	17.9	0.1	19.8
B - G_3/G_4	0.28	0.91	5.07	-17.5	55.2	13.9
B - G_4/C_5	-0.53	0.63	2.91	9.9	-1.8	38.8
B - C_5/G_6	-0.22	1.62	4.56	-8.7	35.5	14.8
B - G_6/C_7	0.95	0.49	2.80	11.5	3.2	38.5
B - C_7/C_8	0.28	0.41	5.58	-18.2	50.2	18.5
B - C_8/G_9	-0.37	1.30	2.51	18.1	7.1	23.3
B - G_9/A_{10}	7.01	3.18	-1.41	131.6	-106.8	177.7

Local base-pair step parameters

base pair step	shift (Å)	slide (Å)	rise (Å)	tilt (°)	roll (°)	twist (°)
$T_1C_2{:}G_9A_{10}$	-	-	-	-	-	-
$C_2G_3{:}C_8G_9$	-0.02	1.49	2.55	0.6	2.6	20.9
$G_3G_4{:}C_7C_8$	0.05	0.67	5.36	0.0	53.0	16.6
$G_4C_5{:}G_6C_7$	-0.70	0.56	2.88	-1.1	-0.2	38.6
$C_5G_6{:}C_5G_6$	0.01	1.61	4.51	0.9	36.1	15.3
$G_6C_7{:}G_4C_5$	0.73	0.54	2.90	-0.1	0.9	38.1
$C_7C_8{:}G_3G_4$	-0.05	0.64	5.33	-0.3	52.2	16.2
$C_8G_9{:}C_2G_3$	-0.03	1.53	2.57	0.7	3.1	21.3
$G_9A_{10}{:}T_1C_2$	-	-	-	-	-	-

Local base step torsional angles

step	δ	ε	ζ	α	β	γ	δ	χ	NtC	CANA
A - T_1/C_2	95.3	211.9	269.1	311.2	181.5	52.9	86.3	256.4	NANT	NAN
A - C_2/G_3	86.3	175.0	274.0	295.0	170.4	53.6	90.6	245.9	AA02	AAA
A - G_3/G_4	90.6	216.7	289.5	291.3	180.1	64.6	131.8	251.9	AB01	A-B
A - G_4/C_5	131.8	186.5	250.1	287.1	184.8	42.8	93.5	248.5	NANT	NAN
A - C_5/G_6	93.5	206.6	278.0	300.9	180.6	58.4	142.8	248.0	AB01	A-B
A - G_6/C_7	142.8	188.0	262.6	294.1	171.8	33.4	86.2	265.3	BA05	B-A
A - C_7/C_8	86.2	213.6	278.4	293.0	175.4	50.8	89.1	245.6	NANT	NAN
A - C_8/G_9	89.1	193.8	272.9	290.4	171.9	54.8	95.6	227.5	AA02	AAA
A - G_9/A_{10}	95.6	180.7	286.9	206.3	157.4	170.1	130.3	247.0	NANT	NAN
B - T_1/C_2	87.2	203.1	268.2	313.9	183.5	53.1	86.8	255.3	NANT	NAN
B - C_2/G_3	86.8	177.6	275.5	294.3	171.1	52.0	90.4	246.9	AA02	AAA
B - G_3/G_4	90.4	212.6	289.9	295.4	180.9	60.1	136.6	252.4	AB01	A-B
B - G_4/C_5	136.6	188.0	246.8	295.7	187.2	32.8	97.9	251.7	NANT	NAN
B - C_5/G_6	97.9	201.3	280.8	296.6	183.0	60.7	145.4	254.7	AB01	A-B
B - G_6/C_7	145.4	192.2	255.8	299.7	166.2	31.3	83.9	266.0	BA05	B-A
B - C_7/C_8	83.9	215.1	280.8	290.1	176.4	53.6	88.2	246.8	NANT	NAN
B - C_8/G_9	88.2	193.4	274.3	292.6	172.0	50.9	95.0	228.6	AA02	AAA
B - G_9/A_{10}	95.0	178.9	286.6	211.8	158.9	166.6	131.9	246.7	NANT	NAN

$\Lambda\text{-}[Ru(phen)_2(11\text{-}NO_2\text{-}dppz)]^{2+}$

(B)A_{10}, (B)G_9, (B)C_8, (A)T_1, (A)C_2, (B)C_7, (A)G_3, (A)G_4, (B)G_6, (A)C_5, (B)C_5, (B)G_4, (A)G_6, (B)G_3, (B)C_2, (A)C_7, (B)T_1, (A)C_8, (A)G_9, (A)A_{10}
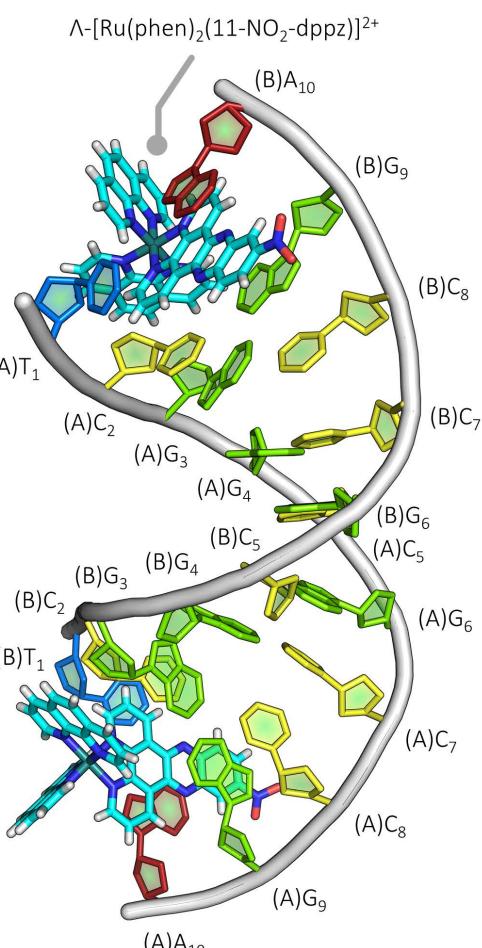
Table A2.9 – *Conformational analysis and local base pair parameters for structure 6GLD.*

[Ru(TAP)$_2$(11-Br-dppz)]$^{2+}$ with d(TCGGCGCCGA)

Local base-pair parameters

base pair	shear (Å)	stretch (Å)	stagger (Å)	buckle (°)	propeller (°)	opening (°)
T$_1$·A$_{10}$	-0.14	-1.15	0.16	-13.7	-8.5	-170.3
C$_2$·G$_9$	0.07	-0.08	0.27	-13.3	7.0	1.0
G$_3$·C$_8$	-0.24	-0.03	0.26	24.7	-5.9	-0.8
G$_4$·C$_7$	-0.20	-0.18	-0.07	-9.1	2.1	-1.1
C$_5$·G$_6$	0.25	-0.07	0.30	-3.6	-5.2	0.4
G$_6$·C$_5$	-0.25	-0.07	0.30	3.6	-5.2	0.4
C$_7$·G$_4$	0.20	-0.18	-0.07	9.1	2.1	-1.1
C$_8$·G$_3$	0.24	-0.03	0.26	-24.7	-5.9	-0.8
G$_9$·C$_2$	-0.07	-0.08	0.27	13.3	7.0	1.0
A$_{10}$·T$_1$	0.14	1.15	-0.16	13.7	8.5	170.3

Local base step parameters

base step	shift (Å)	slide (Å)	rise (Å)	tilt (°)	roll (°)	twist (°)
T$_1$/C$_2$	0.19	1.57	6.81	-19.9	22.0	19.5
C$_2$/G$_3$	-0.76	1.81	2.45	19.7	-1.9	19.0
G$_3$/G$_4$	0.13	1.05	5.11	-18.1	55.7	12.7
G$_4$/C$_5$	-0.58	0.10	3.42	-0.1	-9.7	39.9
C$_5$/G$_6$	-0.02	0.97	3.51	4.7	25.3	25.7
G$_6$/C$_7$	1.10	0.24	3.02	6.3	1.5	38.7
C$_7$/C$_8$	0.43	0.53	5.33	-13.6	45.5	19.7
C$_8$/G$_9$	0.31	1.50	2.69	17.5	8.8	21.2
G$_9$/A$_{10}$	-9.96	-1.71	-1.79	103.0	-79.1	-65.9

Local base-pair step parameters

base pair step	shift (Å)	slide (Å)	rise (Å)	tilt (°)	roll (°)	twist (°)
T$_1$C$_2$:G$_9$A$_{10}$	1.4	1.0	6.7	0.4	15.7	-67.5
C$_2$G$_3$:C$_8$G$_9$	-0.5	1.7	2.6	1.1	3.4	19.6
G$_3$G$_4$:C$_7$C$_8$	-0.1	0.8	5.2	-2.2	50.7	16.1
G$_4$C$_5$:G$_6$C$_7$	-0.8	0.2	3.2	-3.2	-4.1	39.4
C$_5$G$_6$:C$_5$G$_6$	0.0	1.0	3.5	0.0	25.3	25.8
G$_6$C$_7$:G$_4$C$_5$	0.8	0.2	3.2	3.2	-4.1	39.4
C$_7$C$_8$:G$_3$G$_4$	0.1	0.8	5.2	2.2	50.7	16.1
C$_8$G$_9$:C$_2$G$_3$	0.5	1.7	2.6	-1.1	3.4	19.6
G$_9$A$_{10}$:T$_1$C$_2$	-1.4	-1.0	-6.7	-0.4	-15.7	67.5

Local base step torsional angles

step	δ	ε	ζ	α	β	γ	δ	χ	NtC	CANA
T$_1$/C$_2$	85.3	185.9	278.4	299.9	188.9	55.6	93.6	262.0	NANT	NAN
C$_2$/G$_3$	93.6	186.2	274.9	290.5	165.0	56.5	91.6	246.5	AA02	AAA
G$_3$/G$_4$	91.6	213.0	290.2	288.9	183.1	59.4	143.3	252.3	AB01	A-B
G$_4$/C$_5$	143.3	191.4	226.3	299.4	163.3	50.1	95.3	257.7	BA08	B-A
C$_5$/G$_6$	95.3	202.5	286.8	292.8	190.8	54.0	143.0	227.0	AB01	A-B
G$_6$/C$_7$	143.0	183.0	262.5	299.9	168.4	42.6	85.2	265.7	BA01	B-A
C$_7$/C$_8$	85.2	220.4	285.8	292.6	176.4	57.8	108.8	236.9	NANT	NAN
C$_8$/G$_9$	108.8	200.7	271.2	293.3	170.6	51.4	144.6	242.0	AB03	A-B
G$_9$/A$_{10}$	144.6	251.0	74.7	227.7	167.4	61.2	124.4	285.1	NANT	NAN

Λ-[Ru(TAP)$_2$(11-Br-dppz)]$^{2+}$
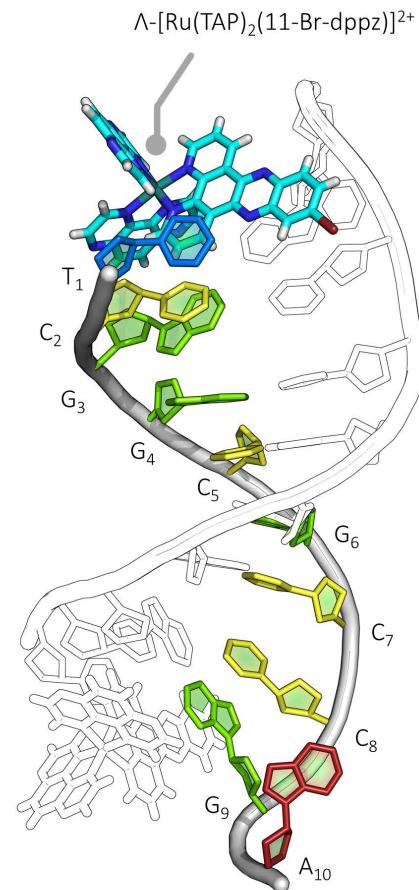
Table A2.10 – *Conformational analysis and local base pair parameters for structure 6HWG.*

$[Ru(TAP)_2(11,12\text{-}CN\text{-}dppz)]^{2+}$ with (CCGGACCCGG/CCGGGTCCGG)

Local base-pair parameters

base pair	shear (Å)	stretch (Å)	stagger (Å)	buckle (°)	propeller (°)	opening (°)
$T_1 \cdot A_{10}$	0.01	2.29	-0.41	7.3	4.4	-85.1
$C_2 \cdot G_9$	0.22	-0.18	-0.16	-4.9	5.9	-0.7
$G_3 \cdot C_8$	-0.19	-0.04	0.38	27.2	-6.8	0.7
$G_4 \cdot C_7$	-0.26	-0.16	-0.14	-13.1	0.9	0.1
$C_5 \cdot G_6$	-0.33	-0.30	0.27	-3.9	-4.6	2.6
$G_6 \cdot C_5$	0.11	-0.35	0.24	2.3	-5.1	0.2
$C_7 \cdot G_4$	0.28	-0.11	-0.22	14.3	-0.3	0.9
$C_8 \cdot G_3$	0.26	-0.08	0.41	-27.7	-6.0	-0.6
$G_9 \cdot C_2$	-0.36	-0.19	-0.23	3.8	5.7	-0.8
$A_{10} \cdot T_1$	0.06	-2.16	1.25	-15.7	-12.1	84.5

Local base step parameters

base step	shift (Å)	slide (Å)	rise (Å)	tilt (°)	roll (°)	twist (°)
A - C_1/C_2	-1.49	1.70	6.84	-3.0	-7.3	17.7
A - C_2/G_3	-0.20	2.06	2.77	13.7	2.5	20.7
A - G_3/G_4	0.26	1.20	5.12	-17.7	57.5	10.5
A - G_4/C_5	-0.93	0.04	3.44	-4.9	3.8	33.0
A - A_5/G_6	0.21	-0.60	3.15	4.3	8.1	28.6
A - C_6/C_7	1.29	0.78	2.85	16.3	13.2	32.5
A - C_7/C_8	0.43	0.68	5.56	-20.1	48.8	16.8
A - C_8/G_9	-0.36	1.35	2.55	18.5	9.6	20.3
A - G_9/G_{10}	6.97	3.50	-2.39	148.0	-92.1	142.4
B - C_1/C_2	-2.14	2.24	7.07	-6.6	-16.4	19.9
B - C_2/G_3	-0.35	2.08	2.81	12.5	2.1	19.1
B - G_3/G_4	0.28	1.25	5.12	-17.8	55.9	11.8
B - G_4/C_5	-0.95	0.05	3.43	-3.2	4.1	33.0
B - G_5/G_6	0.62	-0.46	3.16	4.3	9.0	30.9
B - T_6/C_7	1.04	0.71	2.89	15.0	13.4	29.4
B - C_7/C_8	0.39	0.70	5.51	-18.4	48.2	17.7
B - C_8/G_9	-0.28	1.34	2.58	18.0	11.0	20.8
B - G_9/G_{10}	7.18	3.39	-1.76	143.6	-94.7	159.1

Local base-pair step parameters

base pair step	shift (Å)	slide (Å)	rise (Å)	tilt (°)	roll (°)	twist (°)
$C_1C_2{:}G_9G_{10}$	-0.72	3.28	6.79	-0.6	-6.2	-24.4
$C_2G_3{:}C_8G_9$	0.05	1.69	2.70	-2.2	6.8	20.3
$G_3G_4{:}C_7C_8$	-0.04	0.97	5.33	0.3	52.7	13.8
$G_4A_5{:}T_6C_7$	-0.98	0.39	3.18	-9.9	8.7	31.4
$A_5C_6{:}G_5T_6$	-0.21	-0.52	3.17	0.0	8.6	29.8
$C_6C_7{:}G_4G_5$	1.12	0.43	3.16	9.7	8.7	32.9
$C_7C_8{:}G_3G_4$	0.05	0.98	5.35	-1.2	52.2	13.9
$C_8G_9{:}C_2G_3$	-0.01	1.69	2.71	3.0	5.8	19.3
$G_9G_{10}{:}C_1C_2$	-	-	-	-	-	-

Local base step torsional angles

step	δ	ε	ζ	α	β	γ	δ	χ	NtC	CANA
A - C_1/C_2	161.3	210.9	240.6	133.6	155.0	190.7	108.8	268.1	NANT	NAN
A - C_2/G_3	108.8	186.9	275.7	299.8	168.4	46.6	95.8	258.2	AA02	AAA
A - G_3/G_4	95.8	222.8	283.9	302.3	176.6	51.1	148.6	253.4	NANT	NAN
A - G_4/A_5	148.6	205.4	210.6	311.9	157.1	38.8	137.2	261.1	BB04	B12
A - A_5/C_6	137.2	197.3	264.2	300.0	168.8	39.8	106.0	243.2	BB01	BBB
A - C_6/C_7	106.0	202.3	271.0	304.3	168.1	30.5	84.4	239.8	AA02	AAA
A - C_7/C_8	93.6	197.6	265.7	303.9	163.2	59.0	137.2	234.6	AB01	A-B
A - C_8/G_9	93.6	196.8	272.6	271.1	171.1	67.9	95.5	234.6	AA02	AAA
A - G_9/G_{10}	95.5	192.1	259.0	99.8	183.9	248.5	96.5	242.0	NANT	NAN
B - C_1/C_2	159.3	245.9	201.1	80.0	180.7	221.4	114.0	270.8	NANT	NAN
B - C_2/G_3	114.0	187.7	272.5	303.8	166.5	43.7	96.1	260.9	NANT	NAN
B - G_3/G_4	96.1	218.2	284.9	299.5	178.7	52.6	147.9	255.2	NANT	NAN
B - G_4/G_5	147.9	198.3	218.0	308.4	164.7	41.7	134.3	261.0	BB04	B12
B - G_5/T_6	134.3	194.6	268.9	303.7	169.2	38.4	104.7	244.0	BB01	BBB
B - T_6/C_7	104.7	202.9	271.7	304.9	166.3	31.3	84.1	239.3	AA02	AAA
B - C_7/C_8	84.1	211.2	280.2	295.8	177.3	47.2	93.9	252.0	AA02	AAA
B - C_8/G_9	93.9	195.6	268.5	303.3	167.6	57.0	137.4	235.1	AB01	A-B
B - G_9/G_{10}	137.4	189.2	282.2	177.9	120.3	166.2	86.6	264.7	NANT	NAN

$\Lambda\text{-}[Ru(TAP)_2(11,12\text{-}CN\text{-}dppz)]^{2+}$

LUMO+2
-7. 0280 eV

LUMO+1
-7. 0515 eV

LUMO
-7.1257 eV

HOMO
-10.6004 eV

HOMO-1
-10.7829 eV

HOMO-2
-10.8525 eV
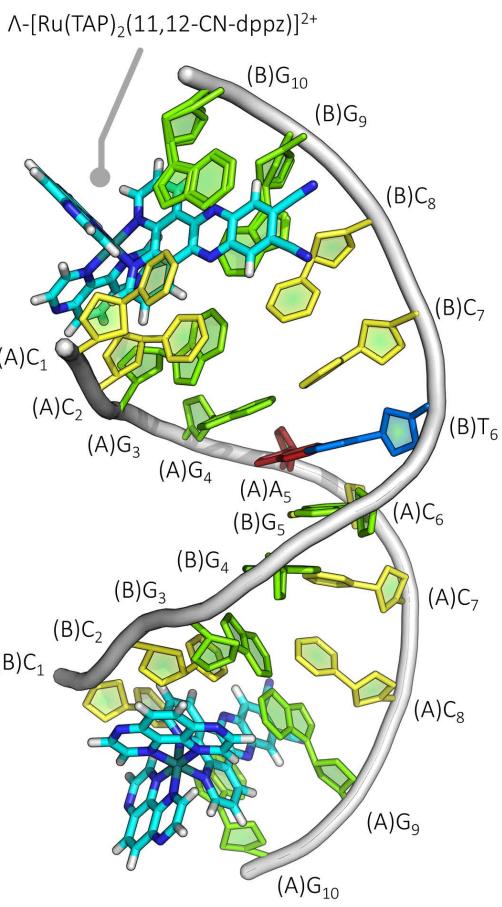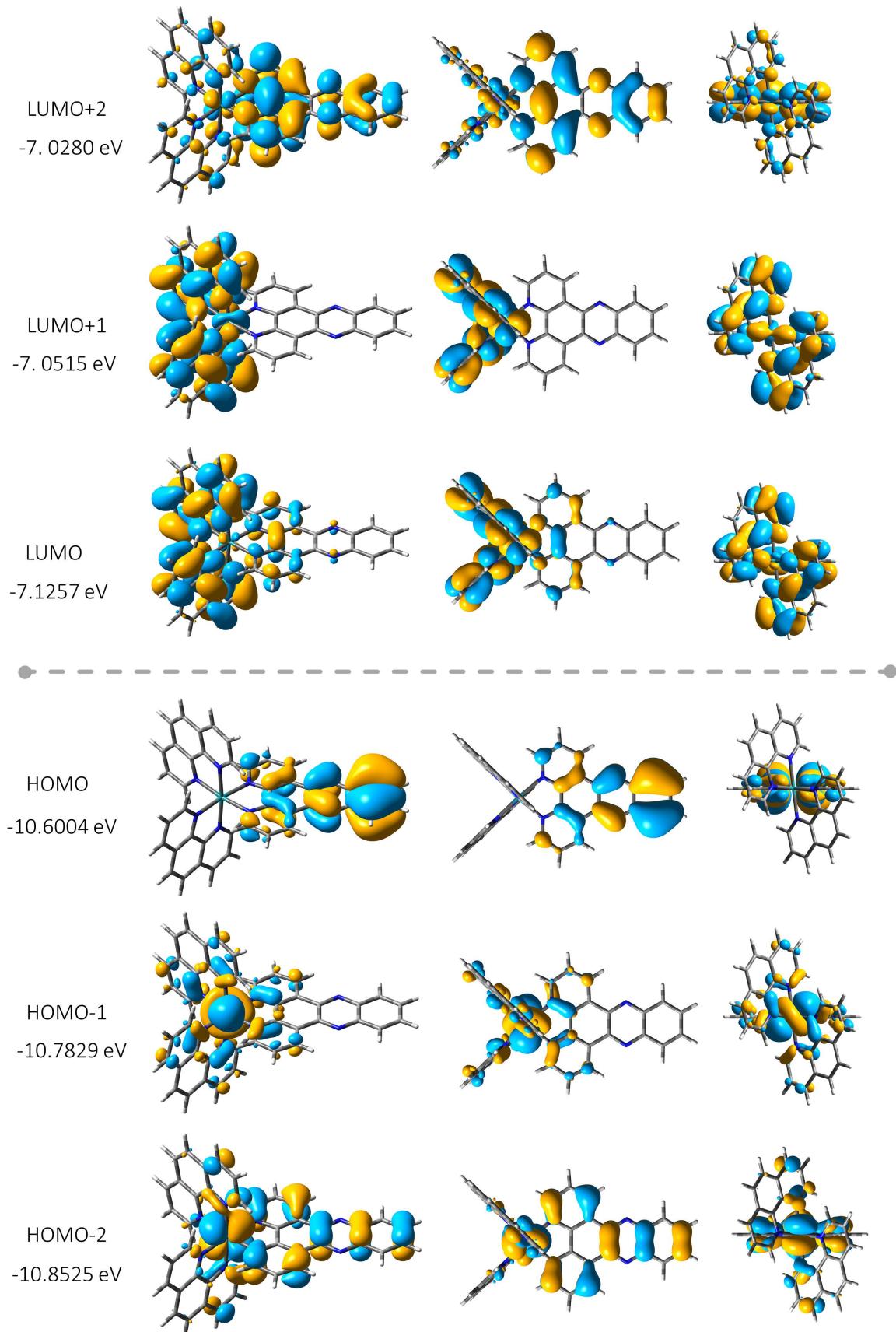
**Figure A2.10** – *Selected frontier molecular orbitals of [Ru(phen)₂(dppz)]²⁺ calculated at a DFT B3LYP/LAN2LZ level.*
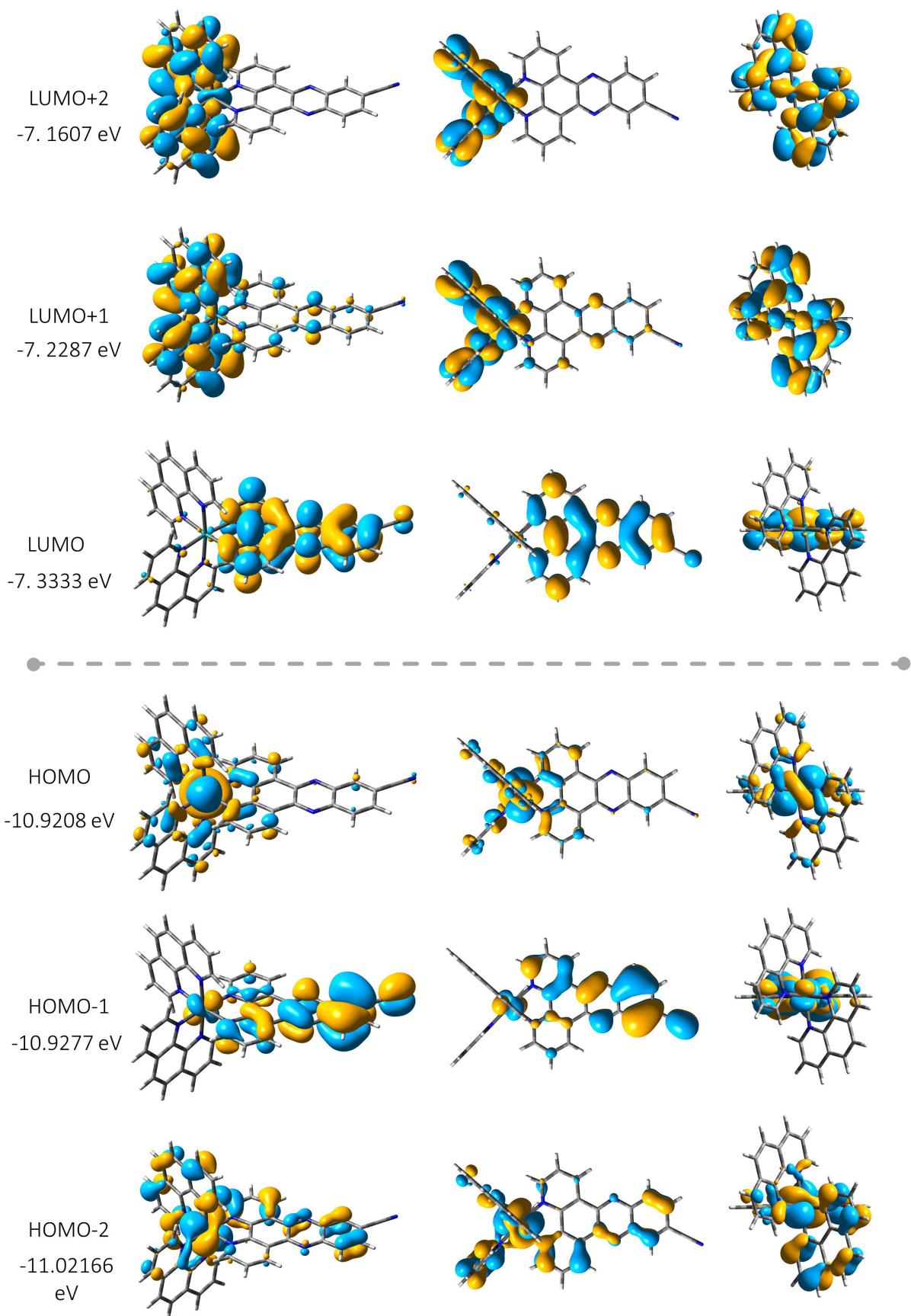
Figure A2.11 – *Selected frontier molecular orbitals of [Ru(phen)₂(11-CN-dppz)]²⁺ calculated at a DFT B3LYP/LAN2LZ level.*

The orbital labels and energies shown in the figure:

LUMO+2
-7. 1607 eV

LUMO+1
-7. 2287 eV

LUMO
-7. 3333 eV

HOMO
-10.9208 eV

HOMO-1
-10.9277 eV

HOMO-2
-11.02166 eV
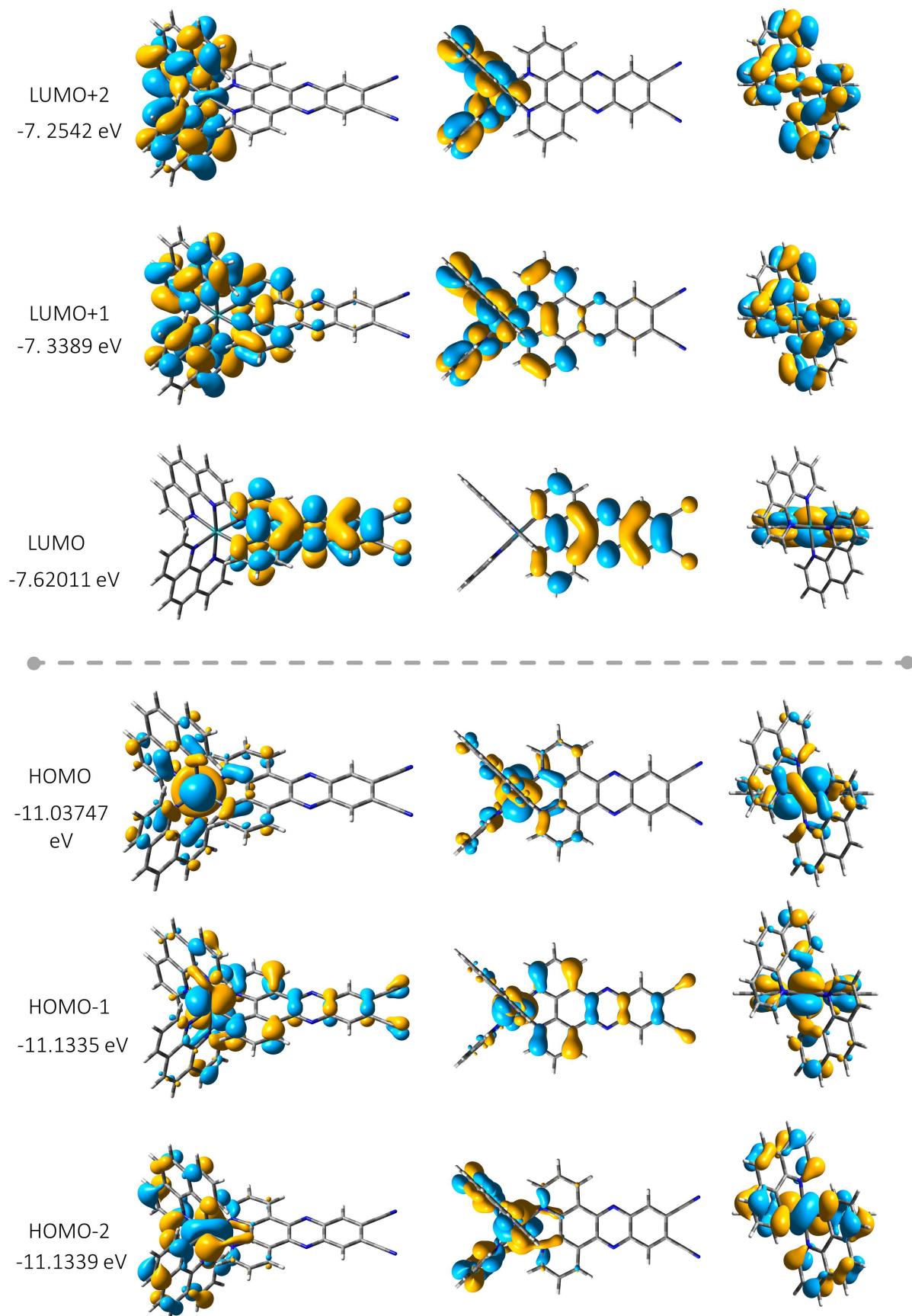
Figure A2.12 – *Selected frontier molecular orbitals of [Ru(phen)$_2$(11,12-CN-dppz)]$^{2+}$ calculated at a DFT B3LYP/LAN2LZ level.*

LUMO+2
-7. 1293 eV

LUMO+1
-7. 1974 eV

LUMO
-7. 3286 eV

HOMO
-10.8938 eV

HOMO-1
-10.9669 eV

HOMO-2
-10.9888 eV
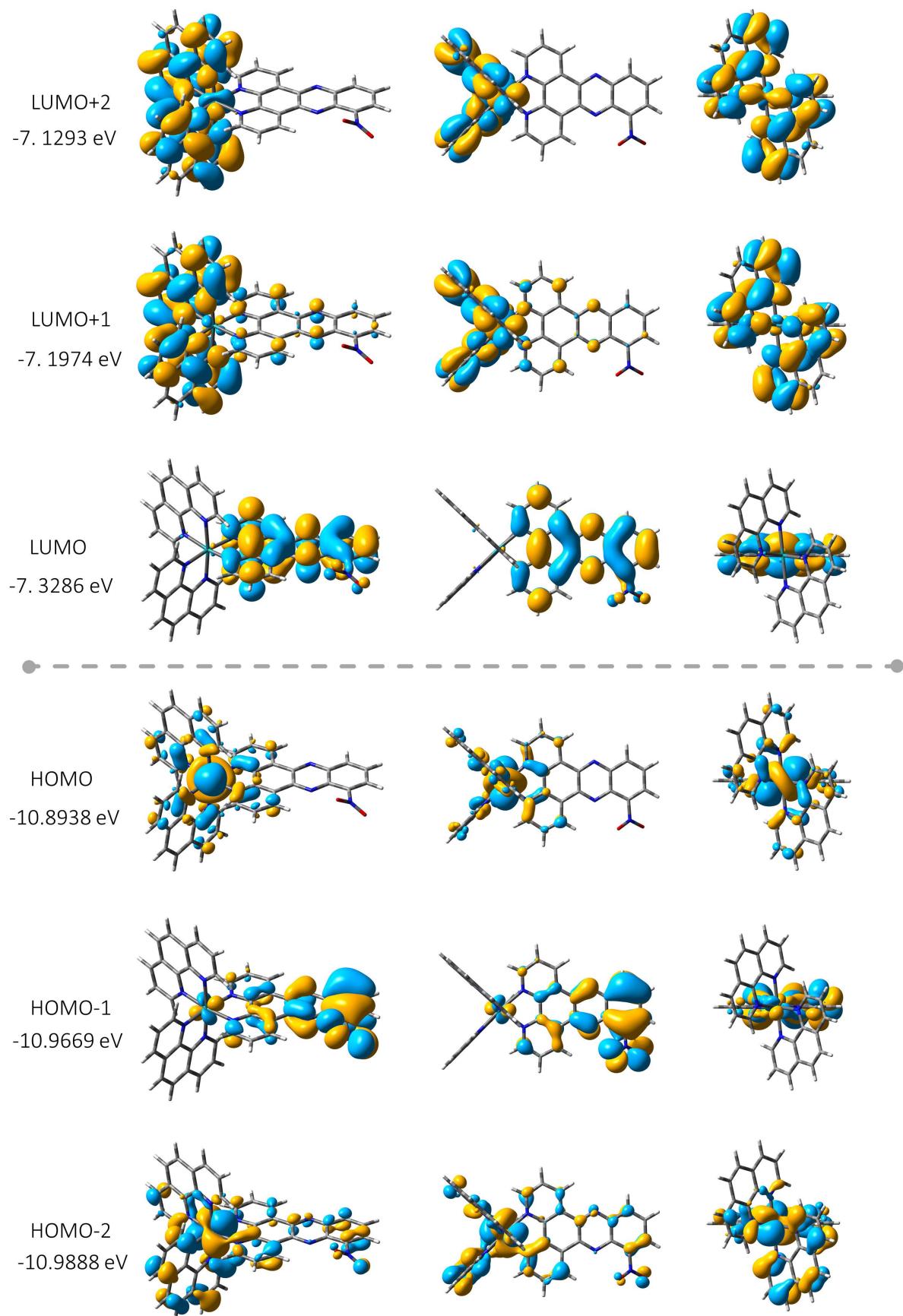
Figure A2.13 – *Selected frontier molecular orbitals of [Ru(phen)$_2$(10-NO$_2$-dppz)]$^{2+}$ calculated at a DFT B3LYP/LAN2LZ level.*

LUMO+2
7. 1856 eV

LUMO+1
-7. 2560 eV

LUMO
-7. 4068 eV

HOMO
-10.9458 eV

HOMO-1
-11.0390 eV

HOMO-2
-11.0400 eV
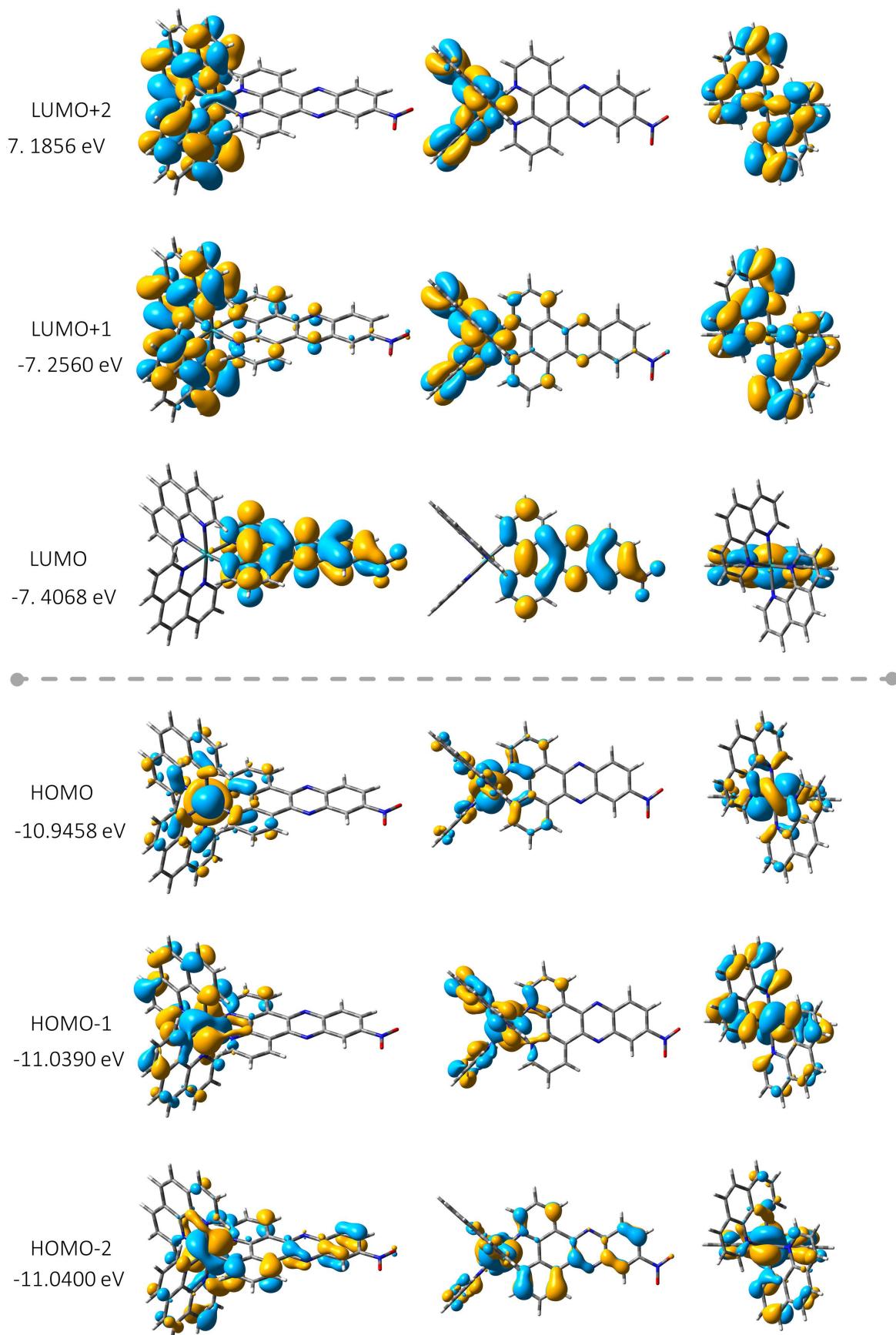
Figure A2.14 – *Selected frontier molecular orbitals of [Ru(phen)$_2$(11-NO$_2$-dppz)]$^{2+}$ calculated at a DFT B3LYP/LAN2LZ level.*

8.3 Chapter 3

Table A3.1 *– Proteins, nucleotides, and nucleic acid sequences used in the high throughput fluorescence screen.*

	biomolecule	conditions	annealing	sequence 1	sequence 2
1	Buffer	50 mM NaCl, 5 mM Tris, pH 7.0	-	-	-
2	BSA	50 mM NaCl, 5 mM Tris, pH 7.0	-	-	-
3	dA	50 mM NaCl, 5 mM Tris, pH 7.0	-	-	-
4	dG	50 mM NaCl, 5 mM Tris, pH 7.0	-	-	-
5	dT	50 mM NaCl, 5 mM Tris, pH 7.0	-	-	-
6	dC	50 mM NaCl, 5 mM Tris, pH 7.0	-	-	-
7	CT DNA	50 mM NaCl, 5 mM Tris, pH 7.0	-	-	-
8	poly G/poly C	50 mM NaCl, 5 mM Tris, pH 7.0	90 °C / 10 min	GGG GGG GGG GGG GGG GGG GG	CCC CCC CCC CCC CCC CCC CC
9	poly GC	50 mM NaCl, 5 mM Tris, pH 7.0	90 °C / 10 min	GCG CGC GCG CGC GCG CGC GC	GCG CGC GCG CGC GCG CGC GC
10	poly GC (Z form)	20 mM NaCl, 4 µM Co(NH$_3$)$_6$, 5 mM Tris, pH 7.0	90 °C / 10 min	GCG CGC GCG CGC GCG CGC GC	GCG CGC GCG CGC GCG CGC GC
11	poly A/poly T	50 mM NaCl, 5 mM Tris, pH 7.0	90 °C / 10 min	AAA AAA AAA AAA AAA AAA AA	TTT TTT TTT TTT TTT TTT TT
12	poly AT	50 mM NaCl, 5 mM Tris, pH 7.0	90 °C / 10 min	ATA TAT ATA TAT ATA TAT AT	ATA TAT ATA TAT ATA TAT AT
13	TGGT	50 mM NaCl, 5 mM Tris, pH 7.0	90 °C / 10 min	CCT CTC TGG TTC TTC	GAA GAA CCA GAG AGG
14	Cruciform	50 mM NaCl, 5 mM Tris, pH 7.0	90 °C / 10 min	TCG CAA CTG CTA TGA CAG TTG CGA	TCG CAA CTG TCA TAG CAG TTG CGA
15	Hairpin Match	50 mM NaCl, 5 mM Tris, pH 7.0	90 °C / 5 min	GTC ACG AGA GCC TCA AAT CTC GTG AC	-
16	Hairpin Mismatch	50 mM NaCl, 5 mM Tris, pH 7.0	90 °C / 5 min	GTC ACG AGA GCC TCA AAT CTC ATG AC	-
17	Triplex	50 mM NaCl, 10 µM MgCl$_2$, 5 mM Tris, pH 7.0	90 °C / 5 min	TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT	AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA
18	Inter G-quad (K$^+$)	50 mM KCl, 5 mM Tris, pH 7	90 °C / 5 min	TAGGGTTA	-
19	HT G-quad (Na$^+$)	10 mM Napi, 100 mM NaCl, pH 7.0	90 °C / 5 min	A GGG TTA GGG TTA GGG TTA GGG	-
20	HT G-quad (K$^+$)	10 mM Kpi, 100 mM KCl, pH 7.0	90 °C / 5 min	A GGG TTA GGG TTA GGG TTA GGG	-
21	c-myc G-quad (K$^+$)	10 mM Kpi, 100 mM KCl, pH 7.0	90 °C / 10 min	TGG GGT GGG TGG GGT GGG TGG GGA AGG	-
22	i-motif	10 mM Kpi, 100 mM KCl, pH 5.5	90 °C / 5 min	CCC TTA ACC CTT ACC TTA CCC T	-
23	AB1-M[a]	50 mM NaCl, 5 mM Tris, pH 7.0	90 °C / 15 min	GAC TTA TCT AGG GGT GAT AAG CTG GTC	GAC CAG CTT ATC ACC CCT AGA TAA GTC
24	AB1-MM[a]	50 mM NaCl, 5 mM Tris, pH 7.0	90 °C / 15 min	GAC TTA TCT AGG GCT GAT AAG CTG GTC	GAC CAG CTT ATC ACC CCT AGA TAA GTC
25	AB1-X (A)[a]	50 mM NaCl, 5 mM Tris, pH 7.0	90 °C / 15 min	GAC TTA TCT AGG GRT GAT AAG CTG GTC	GAC CAG CTT ATC AAC CCT AGA TAA GTC
26	AB1-X (C)[a]	50 mM NaCl, 5 mM Tris, pH 7.0	90 °C / 15 min	GAC TTA TCT AGG GRT GAT AAG CTG GTC	GAC CAG CTT ATC ACC CCT AGA TAA GTC
27	AB1-X (G)[a]	50 mM NaCl, 5 mM Tris, pH 7.0	90 °C / 15 min	GAC TTA TCT AGG GRT GAT AAG CTG GTC	GAC CAG CTT ATC AGC CCT AGA TAA GTC
28	AB1-X (T)[a]	50 mM NaCl, 5 mM Tris, pH 7.0	90 °C / 15 min	GAC TTA TCT AGG GRT GAT AAG CTG GTC	GAC CAG CTT ATC ATC CCT AGA TAA GTC
29	B1-X (A)[a]	50 mM NaCl, 5 mM Tris, pH 7.0	90 °C / 15 min	GAC TTA TCT AGG GT GAT AAG CTG GTC	GAC CAG CTT ATC AAC CCT AGA TAA GTC
30	B1-X (C)[a]	50 mM NaCl, 5 mM Tris, pH 7.0	90 °C / 15 min	GAC TTA TCT AGG GT GAT AAG CTG GTC	GAC CAG CTT ATC ACC CCT AGA TAA GTC
31	B1-X (G)[a]	50 mM NaCl, 5 mM Tris, pH 7.0	90 °C / 15 min	GAC TTA TCT AGG GT GAT AAG CTG GTC	GAC CAG CTT ATC AGC CCT AGA TAA GTC
32	B1-X (T)[a]	50 mM NaCl, 5 mM Tris, pH 7.0	90 °C / 15 min	GAC TTA TCT AGG GT GAT AAG CTG GTC	GAC CAG CTT ATC ATC CCT AGA TAA GTC

a [Ru(phen)₂(dppz)]²⁺

b [Ru(phen)₂(11-CN-dppz)]²⁺

c [Ru(phen)₂(11,12-CN-dppz)]²⁺

d [Ru(TAP)₂(dppz)]²⁺

e [Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$

Λ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$
Δ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$
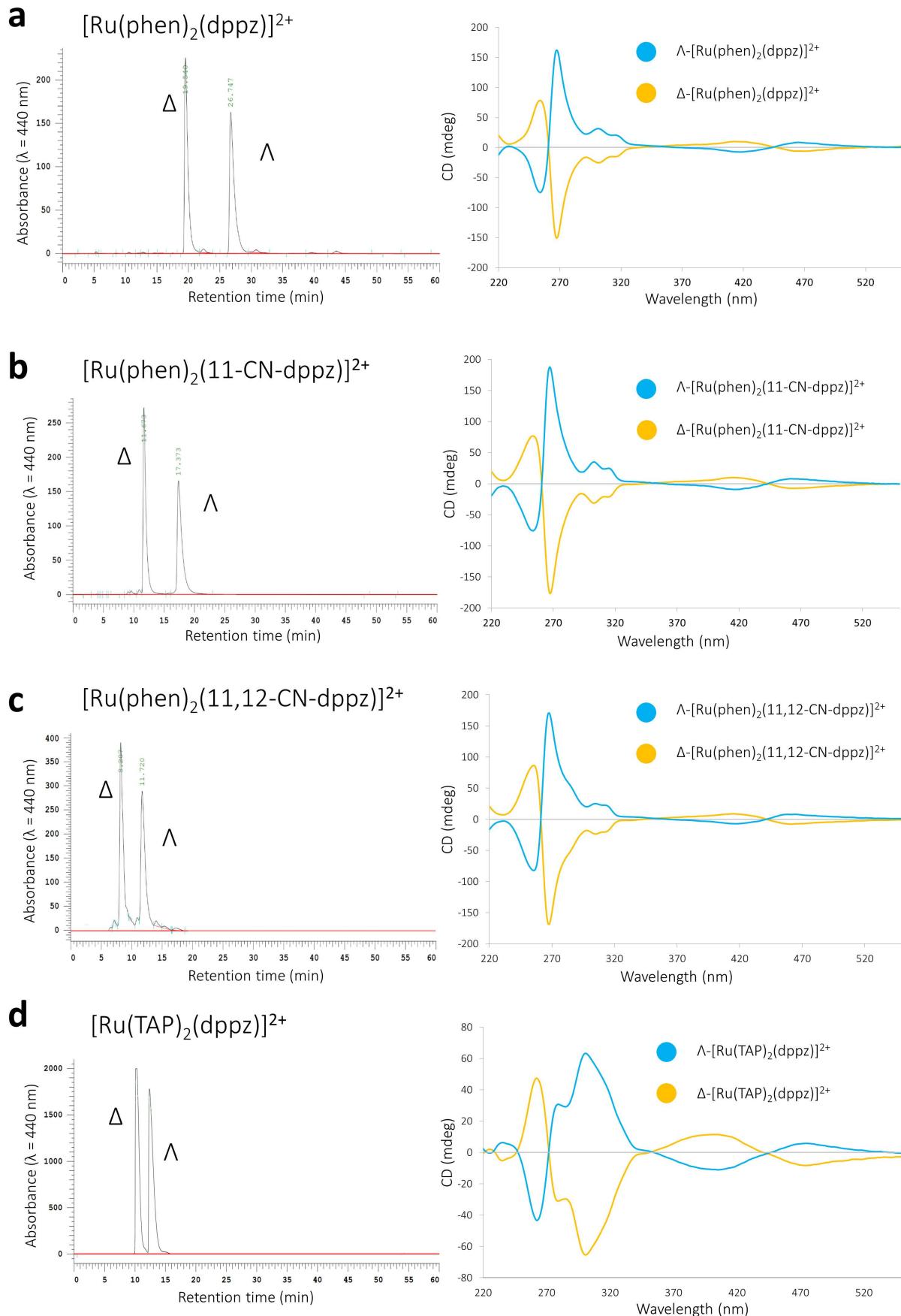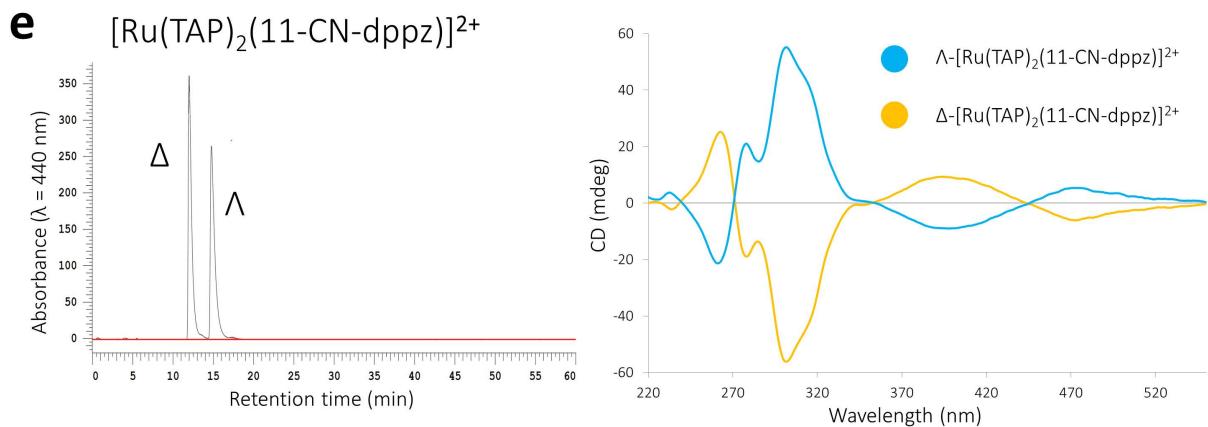
Figure A3.1 *– HPLC chromatograms (left) and the subsequent circular dichroism spectra (right) of the separated enantiomers of; (a) [Ru(phen)$_2$(dppz)]$^{2+}$, (b) [Ru(phen)$_2$(11-CN-dppz)]$^{2+}$, (c) [Ru(phen)$_2$(11,12-CN-dppz)]$^{2+}$, (d) [Ru(TAP)$_2$(dppz)]$^{2+}$, and (e) [Ru(TAP)$_2$(11,12-CN-dppz)]$^{2+}$. A&B were eluted using a 70:30:4:1.6 v/v ACN:MeOH:TEA:AA mixture, whereas C was eluted a 85:15:4:1.6 mixture of the same constituents; D&E were eluted in a buffer composed of 60:40:2:0.8 v/v MeOH:ACN:TEA:AA mixture. All CD spectra were collected at a complex concentration of 15 mM that was confirmed photometrically just before measurement.*
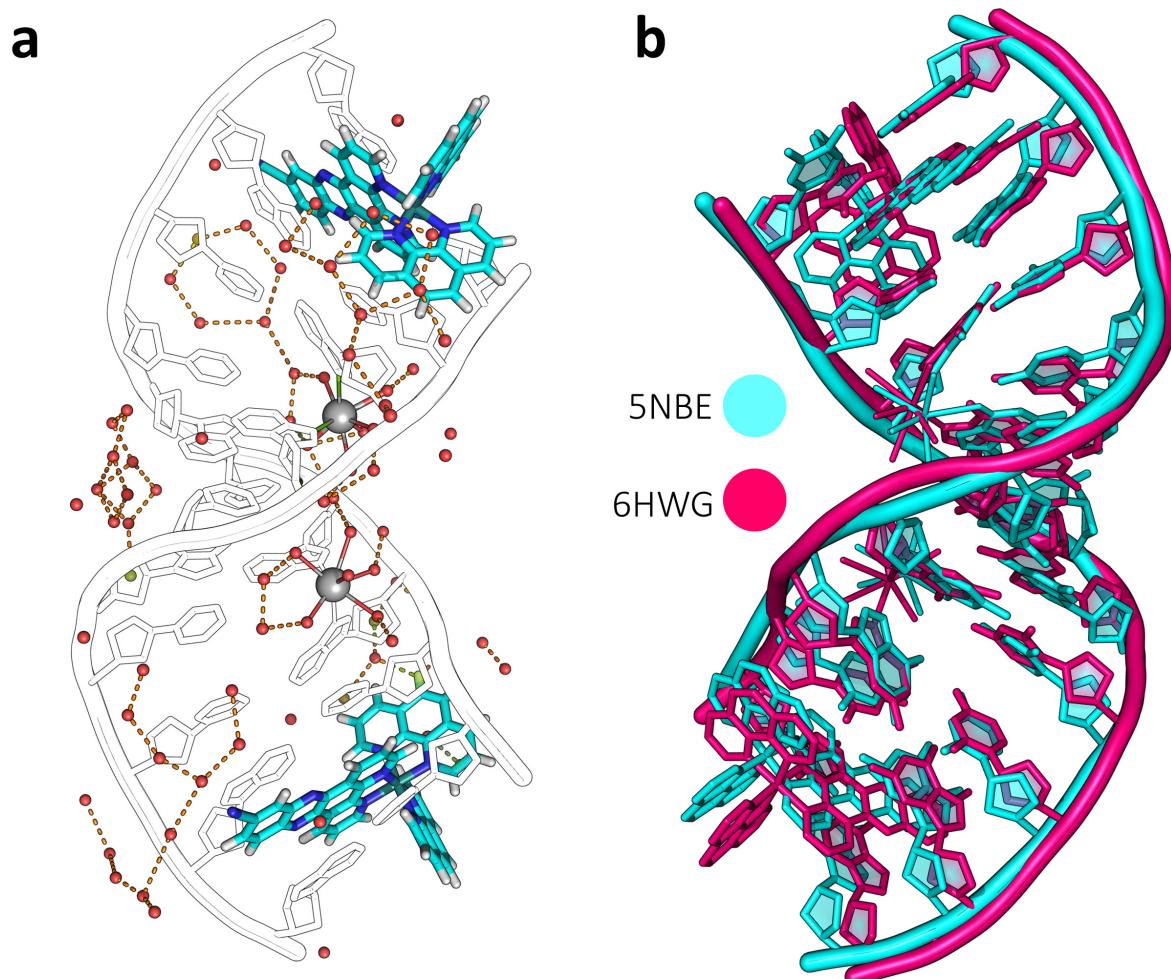
Figure A3.2 *– (a) ordered water network and complex placement around d(TCGGCGCCGA) in 6HWG; (b) superimposition of the two structures of d(TCGGCGCCGA) with either Λ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ (5NBE - blue), or Λ-[Ru(phen)$_2$(11-CN-dppz)]$^{2+}$ (6HWG - pink), highlighting a reoccurring observation that the isoelectric and isostructural complexes bind very similarly to dsDNA (structural RMSD = 1.12 Å).*

Table A3.2 – *Conformational analysis and local base pair parameters for structure 6HWG.*

[Ru(phen)$_2$(11-CN-dppz)]$^{2+}$ with d(TCGGCGCCGA)

Local base-pair parameters

base pair	shear (Å)	stretch (Å)	stagger (Å)	buckle (°)	propeller (°)	opening (°)
T$_1$·A$_{10}$	0.06	-0.32	0.35	7.9	-21.0	-0.1
C$_2$·G$_9$	0.21	-0.12	0.15	-11.7	6.4	2.3
G$_3$·C$_8$	-0.17	-0.06	0.32	22.1	-5.3	-2.3
G$_4$·C$_7$	-0.30	-0.17	-0.14	-10.9	2.1	-0.7
C$_5$·G$_6$	-0.04	-0.12	0.16	-3.3	-13.3	-3.8
G$_6$·C$_5$	0.01	-0.08	0.25	2.3	-12.1	-3.1
C$_7$·G$_4$	0.26	-0.14	-0.27	13.1	1.1	-2.2
C$_8$·G$_3$	0.17	-0.09	0.17	-21.4	-5.9	-3.1
G$_9$·C$_2$	-0.21	-0.13	0.16	11.5	4.7	0.8
A$_{10}$·T$_1$	-0.37	-0.42	-0.18	0.3	11.0	-0.2

Local base step parameters

base step	shift (Å)	slide (Å)	rise (Å)	tilt (°)	roll (°)	twist (°)
A - T$_1$/C$_2$	0.42	0.23	6.73	-7.6	19.7	22.5
A - C$_2$/G$_3$	-0.50	1.69	2.64	17.3	-2.0	20.9
A - G$_3$/G$_4$	-0.04	0.99	5.15	-17.2	59.2	9.7
A - G$_4$/C$_5$	-0.92	0.33	3.49	1.3	-6.4	32.6
A - C$_5$/G$_6$	0.45	0.61	3.31	4.2	17.0	27.2
A - G$_6$/C$_7$	1.37	0.75	2.84	13.0	11.0	33.4
A - C$_7$/C$_8$	0.41	0.80	5.41	-16.0	51.3	12.3
A - C$_8$/G$_9$	-0.12	1.36	2.73	13.8	8.2	25.4
A - G$_9$/A$_{10}$	0.89	0.33	6.83	-6.7	-5.5	18.1
B - T$_1$/C$_2$	-1.78	0.09	7.08	-7.2	-13.4	18.2
B - C$_2$/G$_3$	-0.36	1.68	2.57	18.5	0.2	21.3
B - G$_3$/G$_4$	-0.12	1.12	5.15	-17.9	58.3	9.1
B - G$_4$/C$_5$	-0.89	0.38	3.50	0.6	-6.0	33.1
B - C$_5$/G$_6$	0.34	0.60	3.22	5.9	16.0	26.1
B - G$_6$/C$_7$	1.49	0.68	3.05	9.4	12.4	35.2
B - C$_7$/C$_8$	0.36	0.58	5.41	-14.3	51.4	13.7
B - C$_8$/G$_9$	-0.02	1.34	2.65	15.6	7.0	25.6
B - G$_9$/A$_{10}$	0.90	0.50	6.87	-8.9	-6.6	19.7

Local base-pair step parameters

base pair step	shift (Å)	slide (Å)	rise (Å)	tilt (°)	roll (°)	twist (°)
T$_1$C$_2$:G$_9$A$_{10}$	-0.23	0.37	6.83	0.5	6.5	20.5
C$_2$G$_3$:C$_8$G$_9$	-0.24	1.51	2.66	0.8	2.5	22.8
G$_3$G$_4$:C$_7$C$_8$	-0.18	0.79	5.28	-1.4	55.4	11.6
G$_4$C$_5$:G$_6$C$_7$	-1.22	0.50	3.28	-4.2	2.9	34.2
C$_5$G$_6$:C$_5$G$_6$	0.05	0.59	3.29	-0.9	16.3	27.0
G$_6$C$_7$:G$_4$C$_5$	1.14	0.57	3.18	6.4	2.3	33.6
C$_7$C$_8$:G$_3$G$_4$	0.25	0.96	5.28	0.9	55.0	10.5
C$_8$G$_9$:C$_2$G$_3$	0.12	1.52	2.66	-2.3	4.0	23.1
G$_9$A$_{10}$:T$_1$C$_2$	1.32	0.22	6.97	0.3	-9.5	18.4

Local base step torsional angles

step	δ	ε	ζ	α	β	γ	δ	χ	NtC	CANA
A - T$_1$/C$_2$	82.8	212.2	269.2	303.6	191.6	52.0	92.9	225.7	NANT	NAN
A - C$_2$/G$_3$	92.9	182.9	271.2	297.7	164.8	53.6	92.0	248.4	AA02	AAA
A - G$_3$/G$_4$	92.0	216.8	293.6	283.2	190.4	60.9	148.5	250.2	AB01	A-B
A - G$_4$/C$_5$	148.5	207.5	206.8	310.1	151.8	50.6	129.1	262.5	BB04	B12
A - C$_5$/G$_6$	129.1	206.9	270.7	292.6	165.9	43.4	115.1	248.5	BB00	BBB
A - G$_6$/C$_7$	115.1	191.4	262.2	308.9	165.7	31.4	84.2	262.7	BA05	B-A
A - C$_7$/C$_8$	84.2	209.9	292.2	279.6	180.4	54.0	89.5	252.7	NANT	NAN
A - C$_8$/G$_9$	89.5	190.6	268.0	287.3	165.2	57.6	92.7	240.2	AA02	AAA
A - G$_9$/A$_{10}$	92.7	200.9	279.4	306.7	174.1	60.1	79.7	238.7	NANT	NAN
B - T$_1$/C$_2$	92.1	180.2	249.5	300.3	191.8	59.4	91.7	237.9	NANT	NAN
B - C$_2$/G$_3$	91.7	185.0	268.9	304.5	165.6	48.2	92.7	246.3	AA02	AAA
B - G$_3$/G$_4$	92.7	213.7	295.8	280.3	190.0	61.9	149.2	253.0	NANT	NAN
B - G$_4$/C$_5$	149.2	208.1	205.6	306.7	154.9	50.8	131.9	263.5	BB04	B12
B - C$_5$/G$_6$	131.9	204.0	270.5	291.7	169.4	43.8	116.9	249.3	BB00	BBB
B - G$_6$/C$_7$	116.9	189.3	263.7	299.5	167.1	37.6	83.6	265.0	BA05	B-A
B - C$_7$/C$_8$	83.6	211.7	290.4	275.4	180.3	59.2	86.7	246.9	NANT	NAN
B - C$_8$/G$_9$	86.7	189.8	273.0	287.3	168.5	56.2	92.2	236.4	AA02	AAA
B - G$_9$/A$_{10}$	92.2	196.5	278.7	310.6	175.3	52.8	81.7	241.7	NANT	NAN

Λ-[Ru(phen)$_2$(11-CN-dppz)]$^{2+}$

(B)A$_{10}$
(B)G$_9$
(B)C$_8$
(A)T$_1$
(B)C$_7$
(A)C$_2$
(A)G$_3$
(A)G$_4$
(B)G$_6$
(B)C$_5$ (A)C$_5$
(B)G$_4$ (A)G$_6$
(B)G$_3$
(B)C$_2$
(B)T$_1$
(A)C$_7$
(A)C$_8$
(A)G$_9$
(A)A$_{10}$
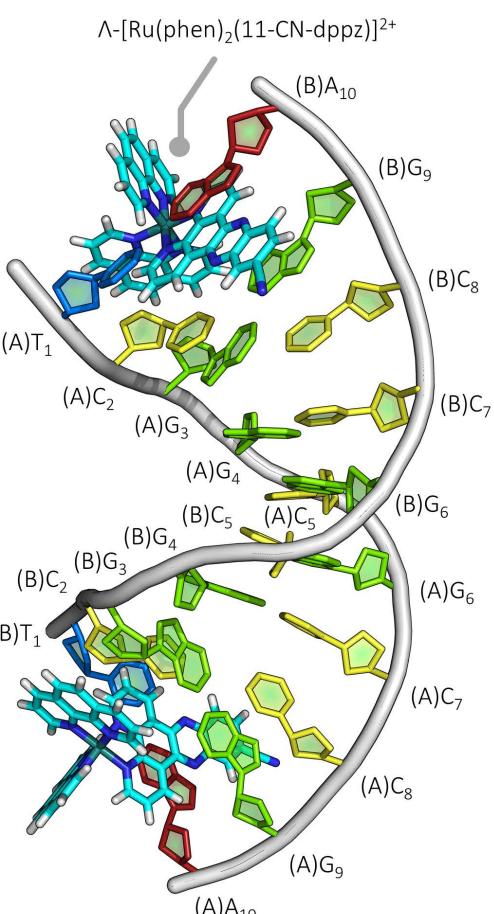
Table A3.3 – *Conformational analysis and local base pair parameters, and chain/base numbering for structure 5LS8.*

$[Ru(TAP)_2(11\text{-}CN\text{-}dppz)]^{2+}$ with d(TAGGGTTA)

Local base-pair parameters

base pair	shear (Å)	stretch (Å)	stagger (Å)	buckle (°)	propeller (°)	opening (°)
(A)T_1·T_7(C)	-2.08	-1.80	0.19	-7.6	-4.3	9.4
(A)A_2·T_6(C)	-0.16	-0.12	0.51	14.5	0.1	8.3
(A)G_3·G_4(C)	-1.75	-3.63	-0.51	7.5	6.4	87.9
(A)G_4·G_4(C)	1.80	3.38	0.45	-7.8	-7.4	-88.5
(A)G_5·G_3(C)	1.78	3.52	-0.09	9.2	3.8	-88.6
(A)T_6·A_2(C)	0.03	-0.05	0.54	18.0	0.4	7.6
(A)T_7·T_1(C)	-2.15	-1.84	0.11	-11.7	-7.0	10.2
(B)A_2·T_6(D)	0.07	-0.14	0.66	25.3	4.1	7.3
(B)G_3·G_5(D)	-1.66	-3.36	-0.66	9.4	11.8	90.8
(B)G_4·G_4(D)	1.70	3.30	-0.42	3.3	4.2	-92.2
(B)G_5·G_3(D)	-1.77	-3.45	-0.65	7.9	10.0	90.0
(B)T_6·A_2(D)	-0.15	-0.01	0.88	-29.5	1.3	5.9

Local base step parameters

base step	shift (Å)	slide (Å)	rise (Å)	tilt (°)	roll (°)	twist (°)
A - T_1/A_2	1.66	1.14	3.06	9.2	7.7	30.3
A - A_2/G_3	2.58	6.31	-4.33	142.8	-97.2	31.4
A - G_3/G_4	3.19	1.29	3.30	-79.9	156.1	13.2
A - G_4/G_5	-0.75	-0.58	3.10	7.1	7.7	22.2
A - G_5/T_6	-0.67	0.79	7.82	-26.0	-7.1	32.4
A - T_6/T_7	1.05	-0.03	2.34	17.8	-1.4	33.3
B - T_1/A_2	-14.68	-1.81	12.00	-97.3	-8.9	-101.4
B - A_2/G_3	3.63	6.09	-5.04	135.2	-90.5	58.3
B - G_3/G_4	1.95	3.34	2.40	-78.5	155.1	-75.9
B - G_4/G_5	-0.75	-0.22	3.11	7.1	10.2	24.4
B - G_5/T_6	-1.10	0.73	8.38	-33.8	-7.7	29.1
B - T_6/T_7	2.66	0.28	3.39	3.9	8.5	54.2
C - T_1/A_2	1.64	1.18	2.88	13.1	7.9	30.5
C - A_2/G_3	2.70	6.24	-4.50	140.7	-96.3	34.0
C - G_3/G_4	0.41	1.42	-3.32	1.5	-4.1	-23.9
C - G_4/G_5	-1.47	-4.14	-0.77	79.2	-152.9	164.1
C - G_5/T_6	-0.55	0.83	7.74	-24.0	-6.7	32.3
C - T_6/T_7	0.78	0.03	2.61	14.0	1.2	31.8
D - T_1/A_2	-16.89	-1.66	10.19	-125.4	-40.8	-117.6
D - A_2/G_3	3.22	6.35	-4.86	134.5	-88.5	48.6
D - G_3/G_4	0.19	1.40	-3.26	-0.1	-3.4	-26.6
D - G_4/G_5	-1.05	-4.45	-0.55	80.8	-151.0	179.1
D - G_5/T_6	-1.09	0.75	8.15	-27.4	-5.6	30.3
D - T_6/T_7	3.80	0.79	3.63	6.2	-14.0	61.8
D - T_7/A_8	2.43	-0.10	2.90	10.7	16.1	17.4

Local base-pair step parameters

base pair step	shift (Å)	slide (Å)	rise (Å)	tilt (°)	roll (°)	twist (°)
(A)T_1A_2:T_6T_7(C)	0.40	0.67	2.84	-2.6	5.3	31.1
(A)A_2G_3:G_5T_6(C)	-1.06	-7.11	2.73	-69.0	26.7	19.7
(A)G_3G_4:G_4G_5(C)	-2.40	-3.01	-1.30	81.7	-158.6	99.8
(A)G_4G_5:G_3G_4(C)	-0.09	-0.55	3.29	3.2	1.0	22.9
(A)G_5T_6:A_2G_3(C)	0.32	2.78	7.25	-8.6	1.8	-18.0
(A)T_6T_7:T_1A_2(C)	-0.25	0.67	2.61	2.7	4.4	31.9
(B)A_2G_3:G_5T_6(D)	-	-	-	-	-	-
(B)G_3G_4:G_4G_5(D)	2.06	3.44	0.23	-82.2	157.0	-180.0
(B)G_4G_5:G_3G_4(D)	0.08	-0.36	3.27	5.0	2.4	25.4
(B)G_5T_6:A_2G_3(D)	-	-	-	-	-	-

Local base step torsional angles

step	δ	ε	ζ	α	β	γ	δ	χ	NtC	CANA
A - T$_1$/A$_2$	163.3	198.5	270.9	300.3	166.0	33.2	140.8	263.1	BB00	BBB
A - A$_2$/G$_3$	140.8	232.9	186.9	136.6	175.3	180.2	144.9	283.8	NANT	NAN
A - G$_3$/anti-G$_4$	144.9	185.0	266.4	291.7	162.9	51.8	76.7	77.3	NANT	NAN
A - G$_3$/syn-G$_4$	144.9	173.3	275.7	296.1	188.4	31.1	89.5	77.3	NANT	NAN
A - anti-G$_4$/G$_5$	76.7	68.4	68.5	218.0	178.2	51.3	133.7	217.1	AB02	A-B
A - syn-G$_4$/G$_5$	89.5	63.2	63.6	220.5	178.2	51.3	133.7	58.2	NANT	NAN
A - G$_5$/T$_6$	133.7	241.3	201.1	310.7	153.7	51.4	141.6	252.1	NANT	NAN
A - T$_6$/T$_7$	141.6	199.7	251.0	312.3	161.9	27.9	116.6	251.5	BB00	BBB
B - T$_1$/A$_2$	93.1	192.7	85.9	80.4	169.7	33.3	141.4	241.7	NANT	NAN
B - A$_2$/G$_3$	141.4	223.6	182.5	142.5	190.6	181.6	142.9	293.3	NANT	NAN
B - G$_3$/anti-G$_4$	142.9	185.5	263.2	309.2	164.9	48.9	112.7	72.2	BBS1	SYN
B - G$_3$/syn-G$_4$	142.9	172.4	275.4	304.1	188.8	34.9	114.4	72.2	NANT	NAN
B - anti-G$_4$/G$_5$	112.7	195.7	255.4	301.4	169.7	48.0	130.6	236.3	BB00	BBB
B - syn-G$_4$/G$_5$	114.4	191.3	265.8	304.6	169.7	48.0	130.6	60.3	NANT	NAN
B - G$_5$/T$_6$	130.6	237.2	197.0	322.6	149.4	45.5	142.3	251.3	NANT	NAN
B - T$_6$/T$_7$	142.3	190.3	274.3	295.9	182.7	22.2	86.3	246.0	BA05	B-A
C - T$_1$/A$_2$	153.8	199.2	271.7	296.3	168.4	38.0	138.7	259.9	BB00	BBB
C - A$_2$/G$_3$	138.7	237.8	187.6	128.3	180.1	183.2	142.9	284.4	NANT	NAN
C - G$_3$/anti-G$_4$	142.9	170.0	272.8	306.1	187.0	39.1	109.7	75.4	NANT	NAN
C - G$_3$/syn-G$_4$	142.9	186.0	257.4	309.6	154.2	58.2	100.6	75.4	NANT	NAN
C - anti-G$_4$/G$_5$	109.7	187.0	270.3	316.9	170.1	37.7	135.2	54.7	NANT	NAN
C - syn-G$_4$/G$_5$	100.6	196.3	263.3	314.3	170.1	37.7	135.2	218.9	AB03	A-B
C - G$_5$/T$_6$	135.2	242.9	206.7	314.9	153.0	46.7	129.7	250.0	NANT	NAN
C - T$_6$/T$_7$	129.7	190.0	253.9	328.4	166.7	20.6	128.3	250.9	BB00	BBB
D - T$_1$/A$_2$	146.0	205.5	91.9	72.4	183.4	40.7	145.0	273.1	NANT	NAN
D - A$_2$/G$_3$	145.0	224.8	179.4	142.6	189.0	180.7	142.6	298.4	NANT	NAN
D - G$_3$/anti-G$_4$	142.6	171.9	275.6	304.7	189.3	34.9	111.1	70.3	NANT	NAN
D - G$_3$/syn-G$_4$	142.6	183.6	265.2	302.0	163.6	56.2	99.9	70.3	NANT	NAN
D - anti-G$_4$/G$_5$	111.1	182.6	270.9	309.0	174.9	43.3	137.8	55.0	NANT	NAN
D - syn-G$_4$/G$_5$	99.9	190.9	266.9	307.6	174.9	43.3	137.8	227.7	AB01	A-B
D - G$_5$/T$_6$	137.8	264.7	194.7	65.6	206.6	235.5	157.4	251.4	NANT	NAN
D - T$_6$/T$_7$	157.4	175.5	267.2	149.7	202.5	165.2	88.5	241.9	NANT	NAN
D - T$_7$/A$_8$	88.5	261.2	257.3	81.7	268.7	239.8	124.4	211.4	NANT	NAN
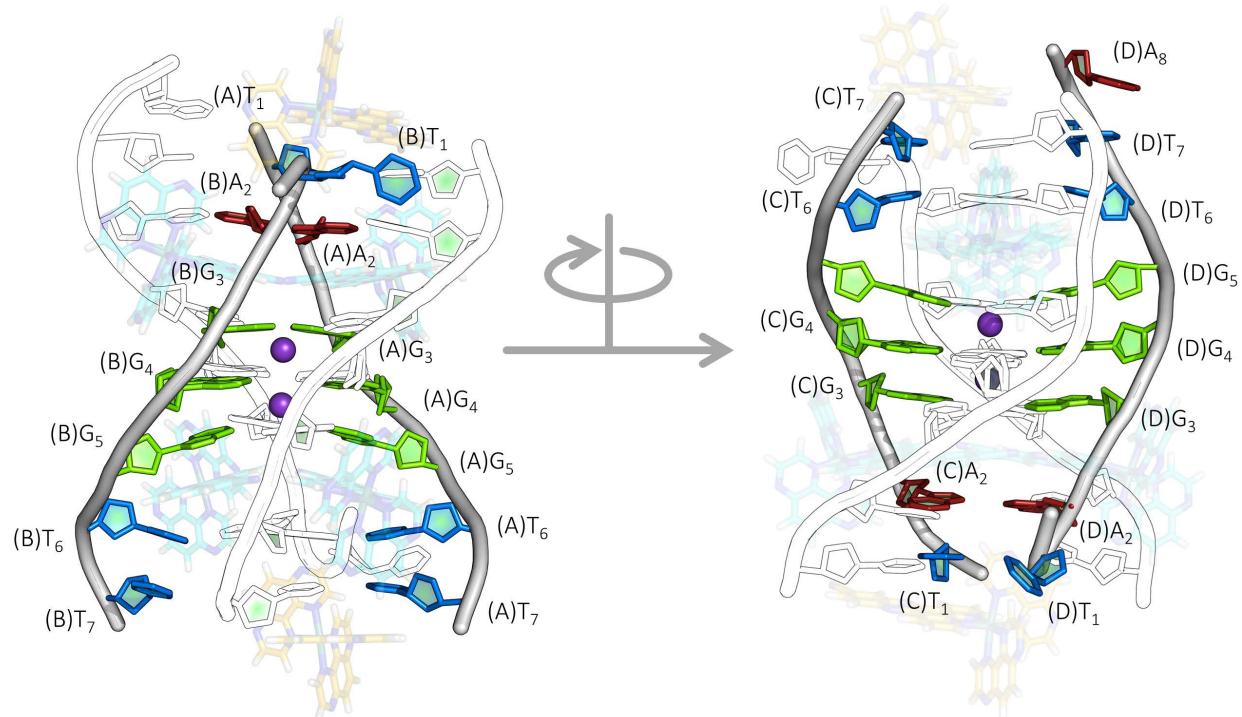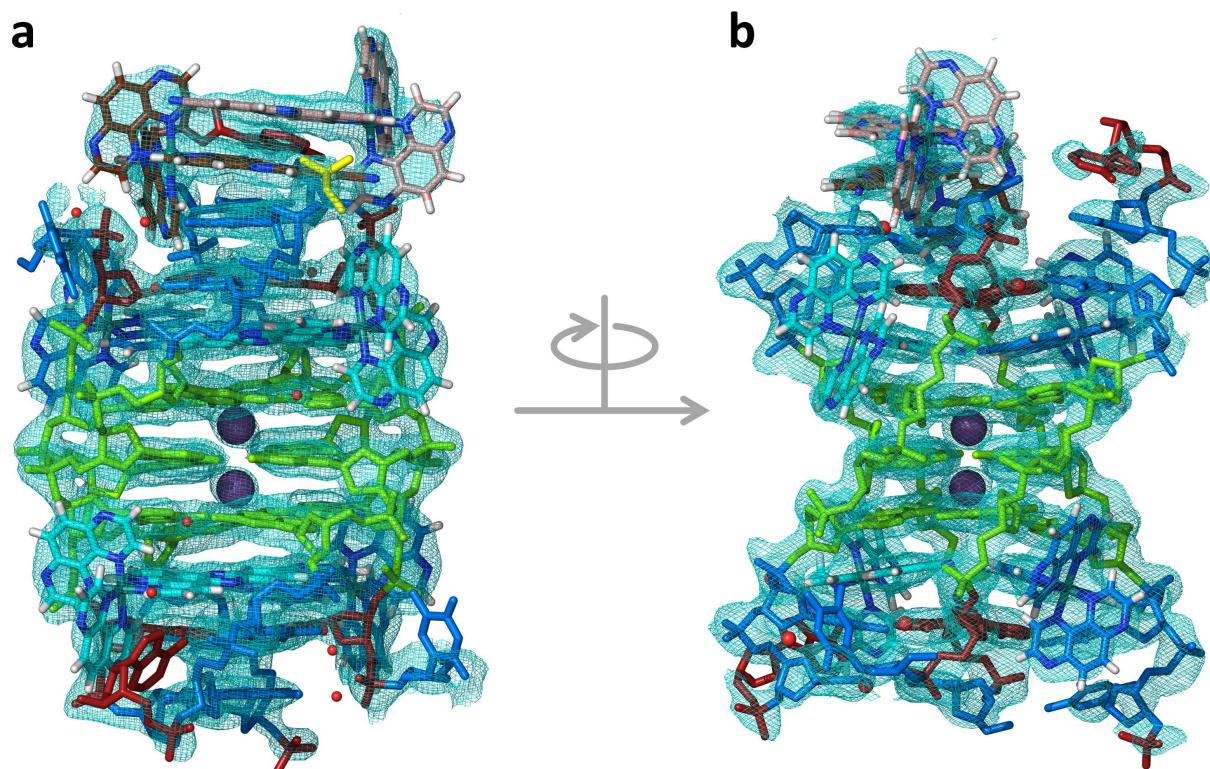
Figure A3.3 *– Two rotated views ((b) is a 90° rotation of (a) around the helical axis) of the model and F_o-F_c map for 5LS8. Density maps contoured at 1σ. Adenines are coloured red, cytosines are blue, and guanines are green; potassium ions are shown as purple spheres whereas waters are shown as red spheres.*
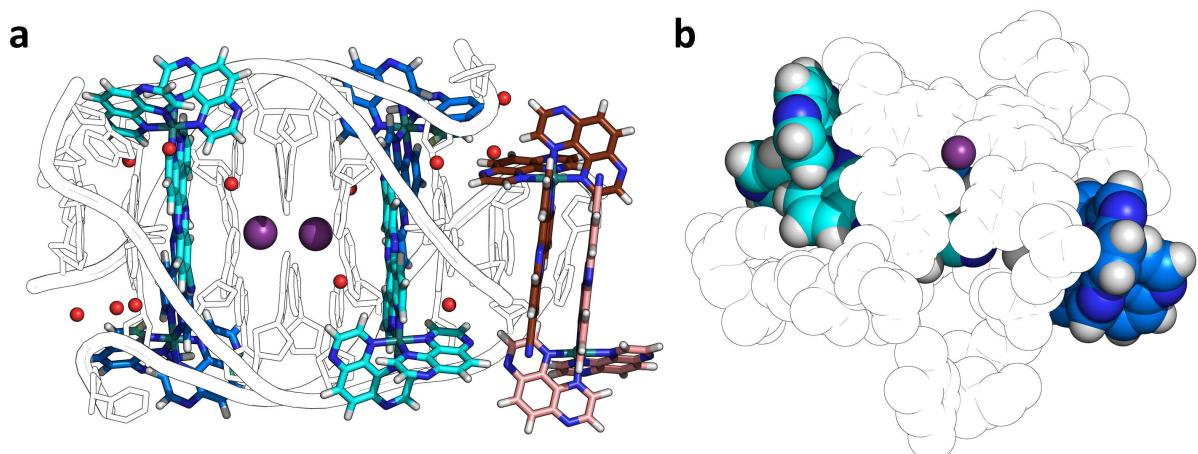
Figure A3.4 *– (a) water framework and complex placement in the 5LS8 structure. Complexes are shown to intercalate in 'pairs'; each lambda complex is coloured light blue or dark blue, and delta as brown and salmon, where each colour is signifies a distinct binding mode. (b) Sphere depiction of the lambda binding modes highlighting how enveloped the intercalating ligands by the G-tetrad and neighbouring T·A·A·T quartet. Note that bases from neighbouring symmetry mates are omitted so ancillary ligands incorrectly look exposed.*
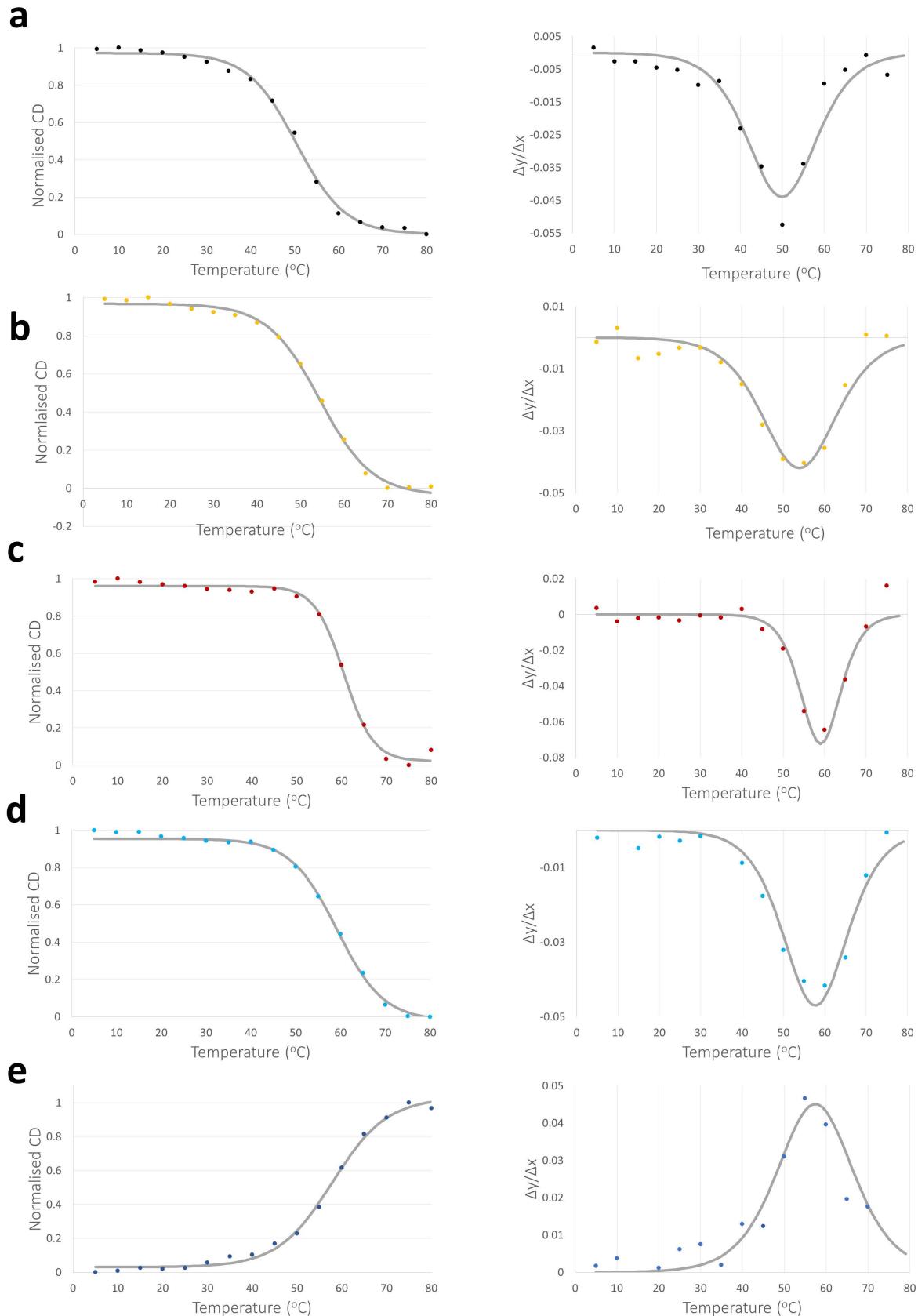
Figure A3.5 – *Normalised SRCD melting profiles (left) and derivative curves (right) for the melting of d(TAGGGTTA) (800 µM) in, (a) the absence of ligand; or in the presence of (b) 200 µM or (c) 800 µM Δ-[Ru(TAP)₂(11-CN-dppz)]²⁺; (d) 200 µM or (e) 800 µM Λ-[Ru(TAP)₂(11-CN-dppz)]²⁺.*
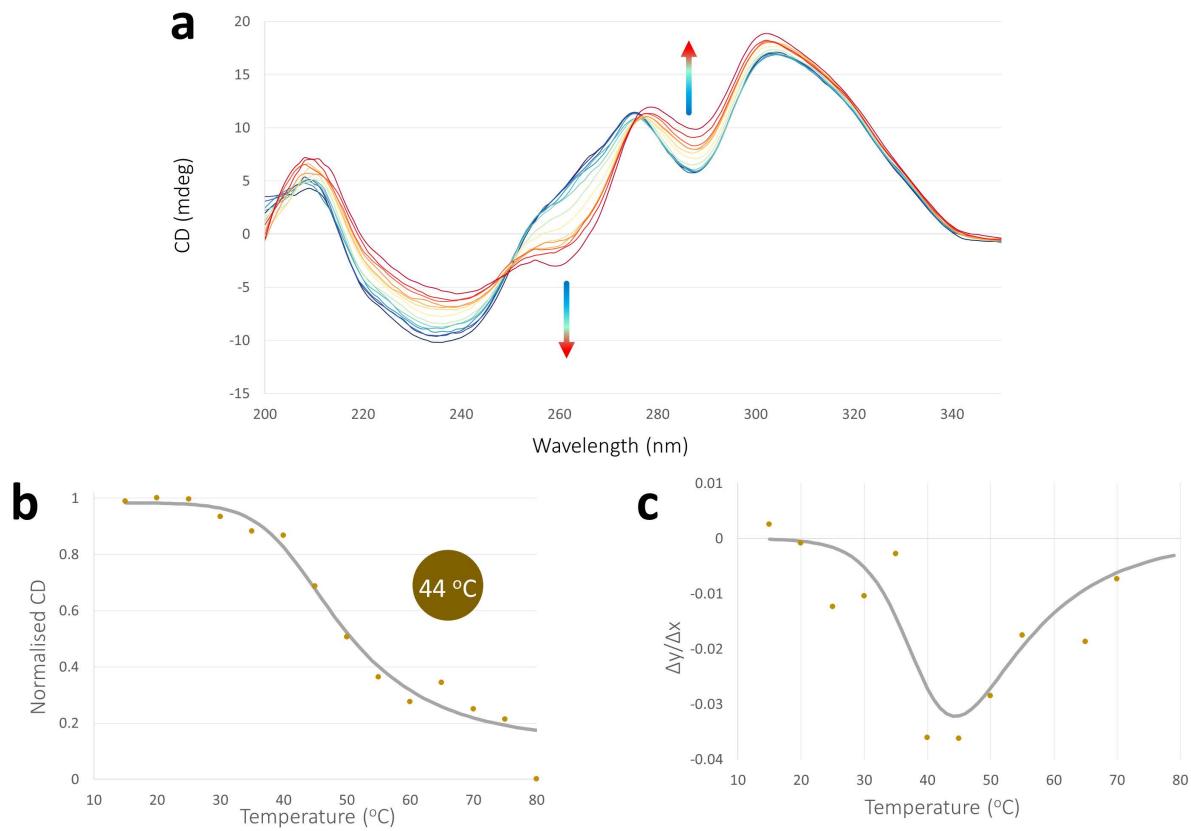
Figure A3.6 *– Circular dichroism melting analysis of r(TAGGGTTA) (800 µM) in the presence of Λ-[Ru(TAP)₂(11-CN-dppz)]²⁺ (800 µM) and 20 mM K-cacodylate pH 7 and 30 mM KF. Melting analysis implies that the parallel topology is conserved in the presence of the complex unlike d(TAGGGTTA) which converts to an anti-parallel in the same conditions.*
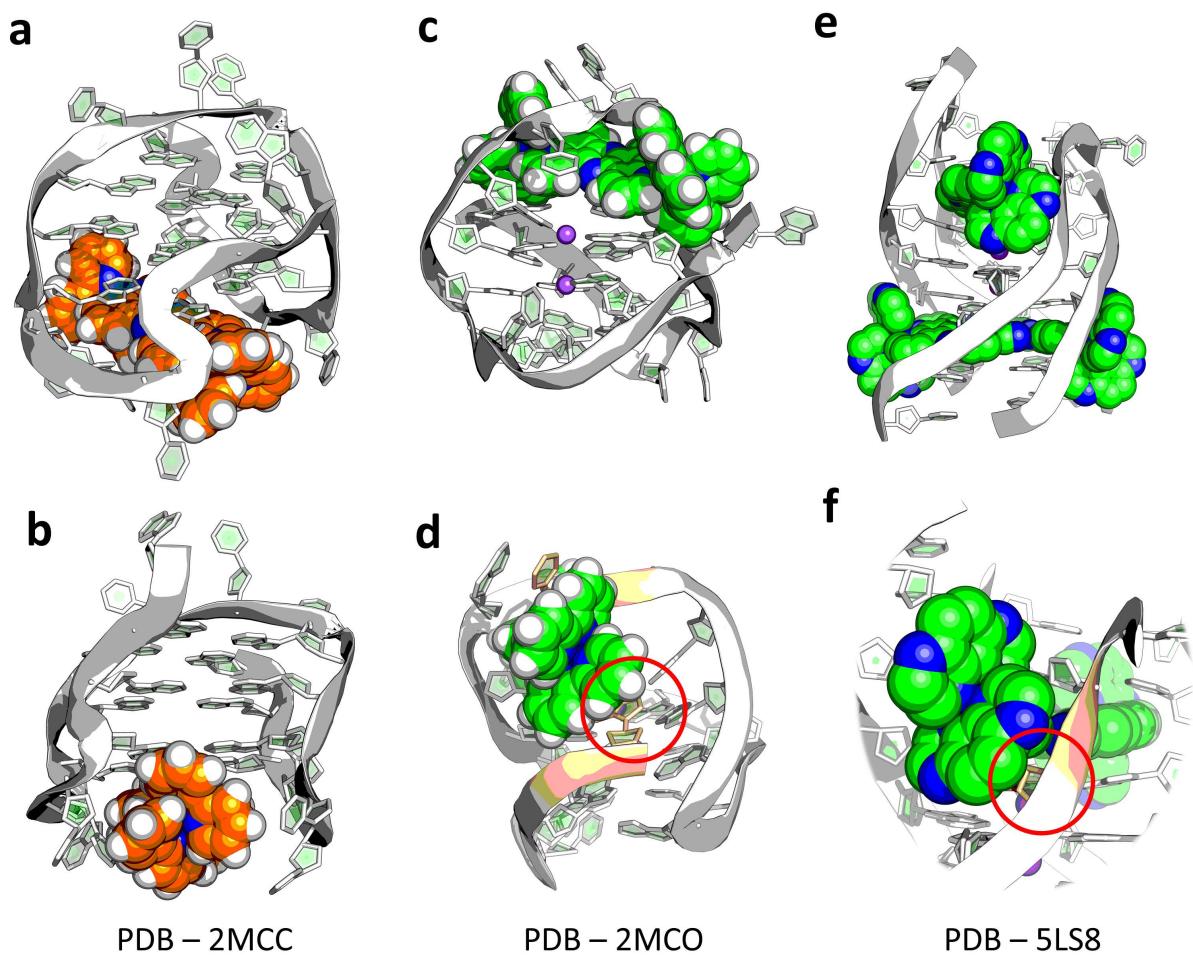
Figure A3.7 – *Structural models of: (a-b) ΔΔ-[{Ru(bpy)₂}₂(tpphz)]⁴⁺ or (c-d) ΛΛ-[{Ru(bpy)₂}₂(tpphz)]⁴⁺ with the potassium folded wtTel22 sequence, and (e-f) Λ-[Ru(TAP)₂(11-CN-dppz)]²⁺ with the potassium folded d(TAGGGTTA). The red circles highlight the ancillary interactions between the respective bpy/TAP ligands and the adjacent ribose of guanosine on the 5' side. This is a reoccurring feature seen in structures that as of yet seems to be completely enantiospecific.*

8.4 Chapter 4
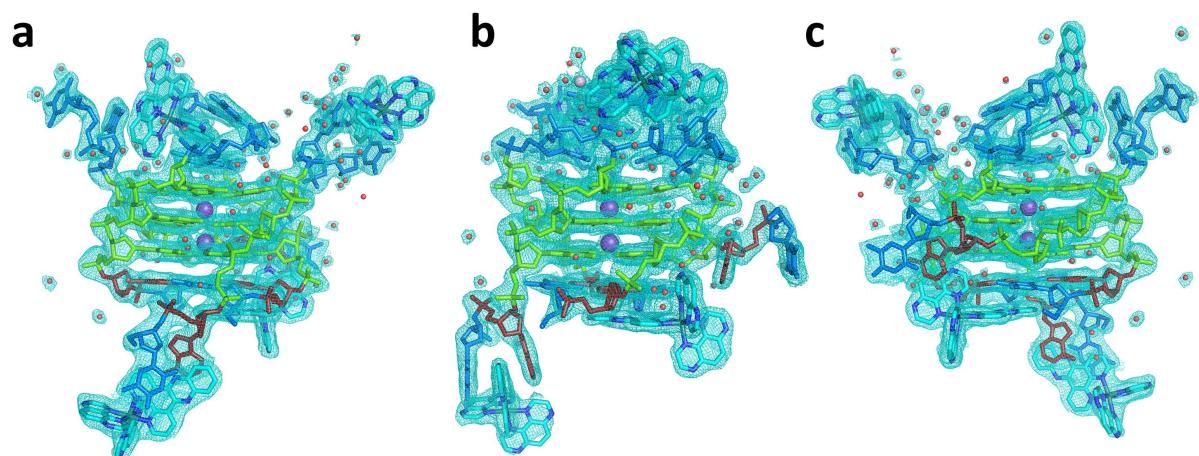
Figure A4.1 – *2Fₒ-Fᵤ electron density maps surrounding the asymmetric unit of the crystal structure of Λ-[Ru(TAP)₂(dppz)]²⁺ bound to d(TAGGGTT), as viewed from three angles perpendicular to the helical axis.*

Table A4.1 *– Conformational analysis, local base pair parameters, and chain/base numbering for structure 6RNL.*

[Ru(TAP)$_2$(dppz)]$^{2+}$ with d(TAGGGTT)

Local base-pair parameters

base pair	shear (Å)	stretch (Å)	stagger (Å)	buckle (°)	propeller (°)	opening (°)
(A)T$_1$·T$_7$(C)	-2.56	-1.97	-0.06	11.9	0.0	10.8
(A)A$_2$·T$_6$(C)	0.50	-0.17	0.06	-10.9	6.2	2.6
(A)G$_3$·G$_3$(B)	-1.67	-3.27	0.01	-6.2	10.1	91.5
(A)G$_4$·G$_4$(B)	1.62	3.30	0.12	1.2	-2.3	-90.8
(A)G$_5$·G$_5$(B)	-1.63	-3.38	0.09	-1.7	5.1	88.7
(A)G$_3$·G$_3$(D)	1.55	3.45	-0.36	14.9	-12.2	-88.5
(A)G$_4$·G$_4$(D)	-1.69	-3.28	0.06	-9.1	3.8	90.2
(A)G$_5$·G$_5$(D)	1.82	3.27	-0.06	1.4	-1.5	-89.5
(A)T$_6$·A$_2$(C)	-0.35	-0.10	0.16	22.4	5.3	8.5
(A)T$_7$·T$_1$(C)	2.74	-2.02	0.32	-6.3	-4.7	11.8
(B)T$_1$·A$_2$(B)	7.13	-0.15	-0.41	0.7	-2.9	8.1
(B)T$_1$·A$_2$(D)	-0.09	-0.98	0.24	-5.4	5.6	-174.0
(B)A$_2$·T$_1$(D)	0.21	1.01	-0.09	8.3	-14.8	168.6
(B)G$_3$·G$_3$(C)	-1.68	-3.19	0.66	-16.7	6.6	89.6
(B)G$_4$·G$_4$(C)	-1.65	-3.46	0.12	-9.8	1.8	86.9
(B)G$_5$·G$_5$(C)	-1.59	-3.17	0.00	-1.7	-4.6	93.3
(B)T$_6$·T$_6$(D)	-	-	-	-	-	-
(B)T$_7$·T$_7$(D)	2.06	3.20	-0.20	9.9	-11.6	173.1
(C)G$_3$·G$_3$(D)	-1.80	-3.31	-0.55	-0.3	8.0	88.0
(C)G$_4$·G$_4$(D)	1.37	3.34	0.04	-0.7	-1.6	-92.0
(C)G$_5$·G$_5$(D)	-1.84	-3.09	0.05	4.6	4.8	88.7
(D)T$_1$·A$_2$(D)	7.17	-0.62	-0.11	-8.6	-5.8	9.1

Local base step parameters

base step	shift (Å)	slide (Å)	rise (Å)	tilt (°)	roll (°)	twist (°)
A - T$_1$/A$_2$	0.42	1.14	7.01	-13.8	-8.1	43.5
A - A$_2$/G$_3$	-3.08	3.03	13.95	-70.7	-6.7	-1.7
A - G$_3$/G$_4$	-0.31	-1.11	3.29	-2.5	4.8	28.1
A - G$_4$/G$_5$	-0.92	-0.97	3.06	3.4	5.3	25.5
A - G$_5$/T$_6$	4.62	2.51	15.18	-78.7	-82.3	68.8
A - T$_6$/T$_7$	3.21	0.98	7.12	-9.5	-8.0	44.2
B - T$_1$/A$_2$	-7.13	0.15	0.41	-0.7	2.9	-8.1
B - A$_2$/G$_3$	-0.17	-0.36	3.05	6.6	4.9	34.3
B - G$_3$/G$_4$	-0.62	-1.31	3.09	1.4	1.1	30.4
B - G$_4$/G$_5$	-0.83	-0.54	3.07	4.0	6.1	26.7
B - G$_5$/T$_6$	1.87	-0.50	2.92	6.8	6.2	36.7
B - T$_6$/T$_7$	2.07	1.01	6.36	-17.4	29.2	45.7
C - T$_1$/A$_2$	1.18	0.95	7.09	-18.6	-3.6	40.2
C - A$_2$/G$_3$	-10.51	-3.21	8.32	-74.6	27.1	-31.3
C - G$_3$/G$_4$	-1.15	-1.52	3.26	-1.5	6.3	24.2
C - G$_4$/G$_5$	-0.47	-1.12	3.34	-1.3	0.9	29.7
C - G$_5$/T$_6$	3.61	1.00	12.48	-56.9	-50.2	61.4
C - T$_6$/T$_7$	3.52	0.26	7.09	-10.9	-14.3	51.8
D - T$_1$/A$_2$	-7.17	0.62	0.11	8.6	5.8	-9.1
D - A$_2$/G$_3$	0.90	-0.42	3.22	2.9	2.1	33.7
D - G$_3$/G$_4$	-0.85	-0.63	3.24	1.3	3.8	27.2
D - G$_4$/G$_5$	-1.03	-0.70	3.02	3.3	8.0	25.7
D - G$_5$/T$_6$	2.13	-1.06	3.11	5.0	0.2	44.3
D - T$_6$/T$_7$	2.41	0.18	7.67	-24.5	-24.2	56.4

Local base-pair step parameters

base pair step	shift (Å)	slide (Å)	rise (Å)	tilt (°)	roll (°)	twist (°)
(A)T$_1$A$_2$:T$_6$T$_7$(C)	-1.57	0.80	7.07	-1.7	-11.9	47.5
(A)G$_3$G$_4$:G$_4$G$_3$(B)	0.76	0.50	3.05	-2.3	-4.2	-26.1
(A)G$_4$G$_5$:G$_5$G$_4$(B)	0.26	0.96	3.17	-0.9	-3.5	-29.2
(A)G$_3$G$_4$:G$_4$T$_3$(D)	0.59	0.54	3.06	-3.3	-4.8	-25.6
(A)G$_4$G$_5$:G$_5$T$_4$(D)	0.24	0.42	3.28	0.8	-1.7	-27.7
(A)T$_6$T$_7$:T$_1$A$_2$(C)	1.01	1.14	7.17	4.7	-7.0	41.8
(B)T$_1$A$_2$:A$_2$T$_1$(D)	-7.16	0.39	0.15	3.9	4.5	-179.6
(B)G$_3$G$_4$:G$_4$T$_3$(C)	0.26	0.45	3.23	-2.8	-0.7	-28.3
(B)G$_4$G$_5$:G$_5$T$_4$(C)	0.55	1.18	3.23	1.9	-1.6	-27.2
(B)T$_6$A$_7$:T$_7$T$_6$(D)	-	-	-	-	-	-
(C)G$_3$G$_4$:G$_4$T$_3$(D)	0.69	0.84	3.16	1.8	-4.7	-27.7
(C)G$_4$G$_5$:G$_5$T$_4$(D)	1.02	0.65	3.22	-0.8	-4.5	-25.7

Local base step torsional angles

step	δ	ε	ζ	α	β	γ	δ	χ	NtC	CANA
A - T_1/A_2	151.8	227.7	288.5	297.4	188.1	34.0	143.8	200.9	IC06	ICL
A - A_2/G_3	143.8	253.5	305.5	163.9	75.7	172.8	143.5	273.2	NANT	NAN
A - G_3/G_4	143.5	174.0	261.0	297.9	183.5	47.6	124.5	258.8	BB01	BBB
A - G_4/G_5	124.5	189.0	255.0	303.0	185.2	35.7	137.1	237.8	BB00	BBB
A - G_5/T_6	137.1	198.0	212.2	55.6	183.1	52.5	143.5	251.6	NANT	NAN
A - T_6/T_7	143.5	206.5	268.0	303.9	182.6	42.4	93.5	258.1	NANT	NAN
B - T_1/A_2	155.3	246.1	174.9	250.0	164.4	43.2	145.0	276.6	NANT	NAN
B - A_2/G_3	145.0	178.9	276.2	175.8	148.9	163.4	152.4	245.7	BB03	BB1
B - G_3/G_4	152.4	181.6	265.4	290.6	200.8	37.6	142.8	235.4	BB00	BBB
B - G_4/G_5	142.8	187.1	251.1	291.9	188.6	39.3	141.8	256.5	BB00	BBB
B - G_5/T_6	141.8	179.5	265.2	298.7	169.7	49.3	89.8	265.4	BA01	B-A
B - T_6/T_7	89.8	198.9	279.6	292.6	194.4	55.6	145.5	229.5	AB04	A-B
C - T_1/A_2	150.2	237.4	289.9	308.8	169.2	21.1	132.2	204.4	IC06	ICL
C - A_2/G_3	132.2	215.9	185.6	250.3	235.4	48.2	84.3	267.6	NANT	NAN
C - G_3/G_4	84.3	73.4	67.4	218.5	185.9	49.4	132.0	203.5	AB02	A-B
C - G_4/G_5	132.0	191.5	250.0	313.9	172.0	36.4	134.8	238.4	BB00	BBB
C - G_5/T_6	134.8	202.5	209.1	40.8	137.6	45.0	135.1	243.6	NANT	NAN
C - T_6/T_7	135.1	196.7	281.7	292.3	188.9	59.0	105.3	232.9	NANT	NAN
D - T_1/A_2	143.0	233.2	184.7	252.8	167.1	46.2	149.9	270.3	NANT	NAN
D - A_2/G_3	149.9	183.8	266.7	309.2	175.1	34.7	130.7	251.3	BB00	BBB
D - G_3/G_4	130.7	195.1	227.1	316.4	157.2	42.7	126.7	250.7	BB04	B12
D - G_4/G_5	126.7	196.9	244.5	301.1	168.5	46.4	131.2	240.4	BB00	BBB
D - G_5/T_6	131.2	160.5	276.9	184.4	171.5	171.4	143.5	243.6	BB03	BB1
D - T_6/T_7	143.5	205.7	268.7	304.3	181.7	54.8	137.0	221.1	NANT	NAN
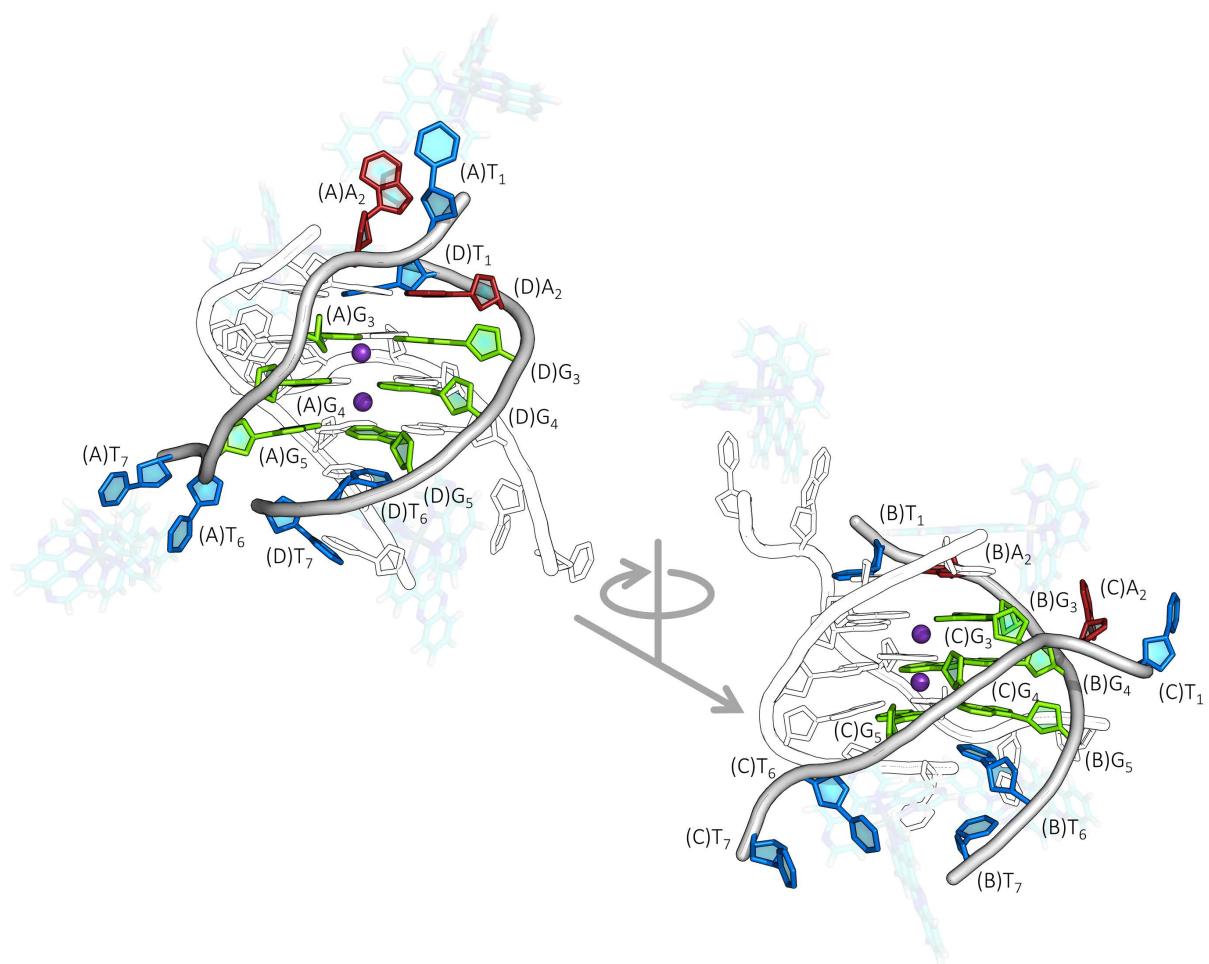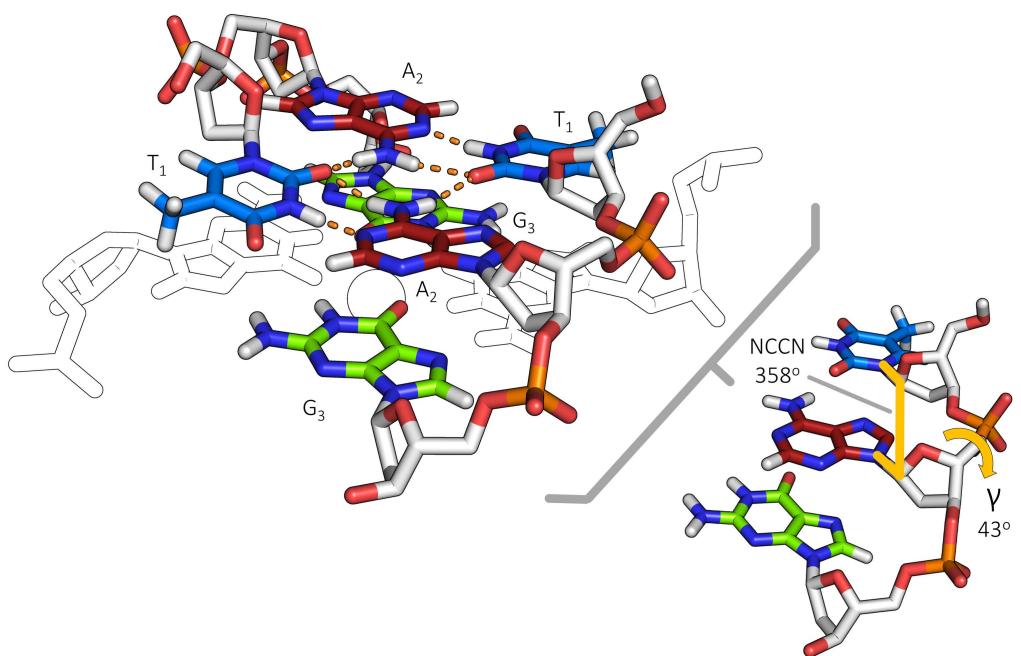
Figure A4.2 – *A T·A:T·A quartet is formed on the 5' side of the G-quadruplex stack. The quartet is formed from two strands (coloured) with each TA step being effectively planar; this can be seen in the torsional angles of the step with a 358° NCCN dihedral, and gauche (or closed) 43° γ angle. Inset shows how these torsional angles are measured.*
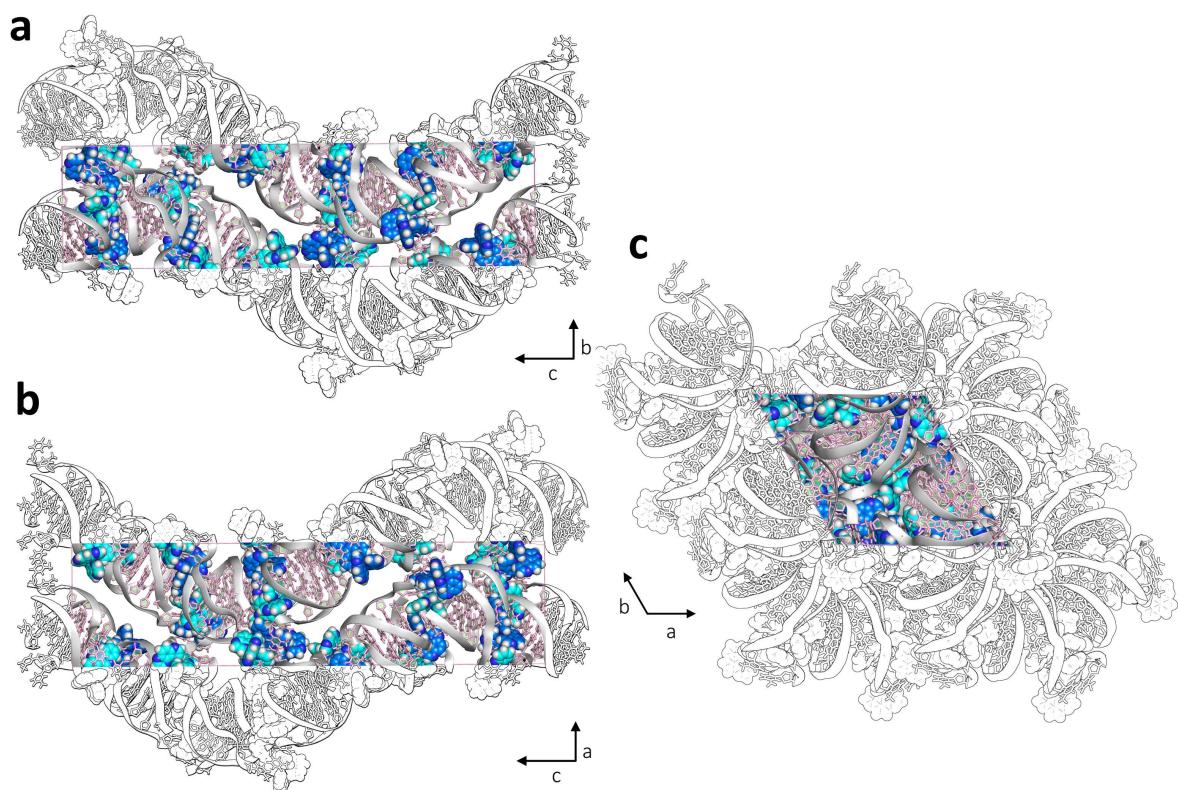
Figure A4.3 – *Multiple views of the crystallographic unit cell of 6RNL; looking down (i.e. perpendicular to the plane of the paper) the (a) a axis, (b) b axis, and (c) c axis. DNA is coloured grey and light pink, ruthenium complexes are coloured light or dark blue. Coloured areas represent the unit cell.*
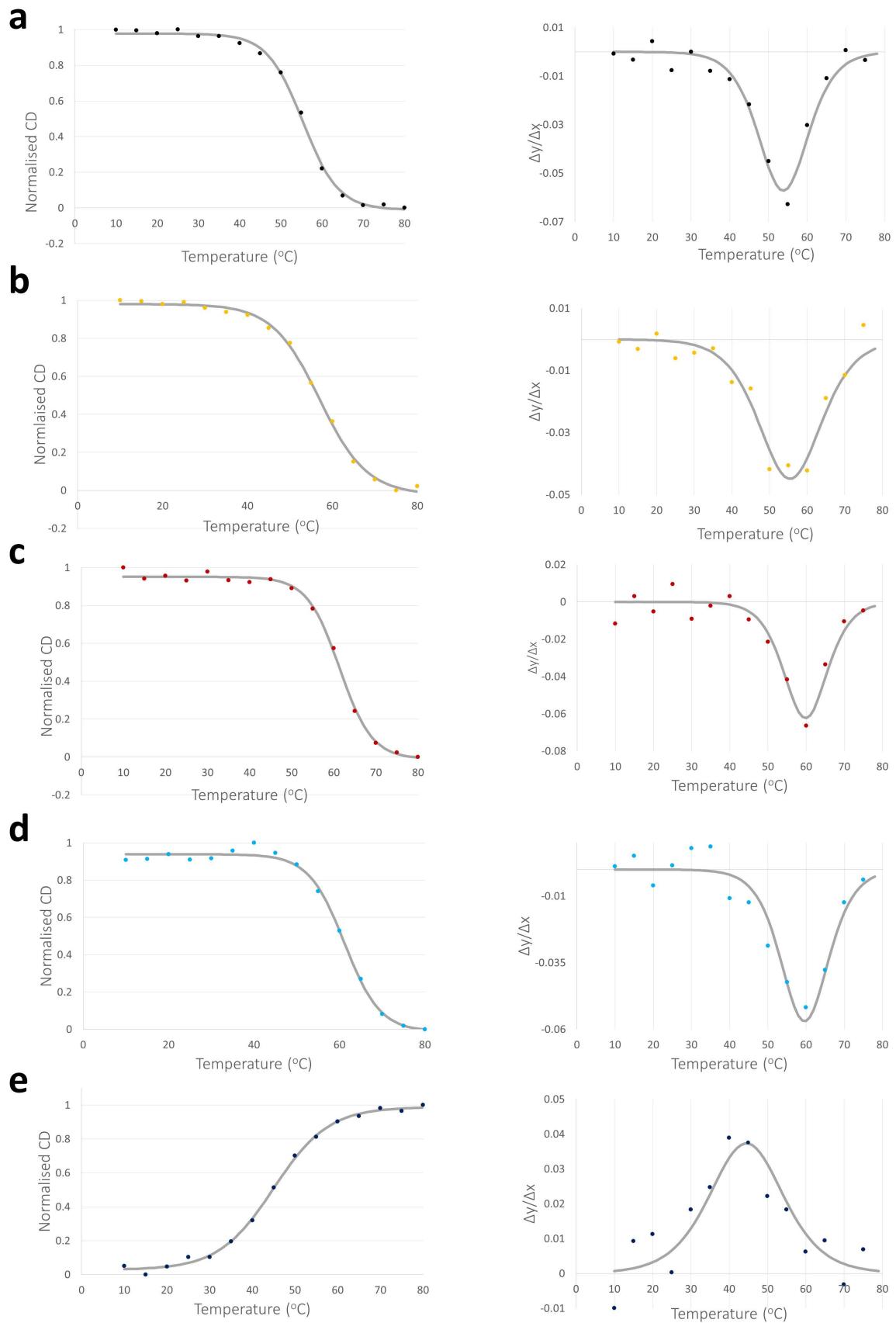
Figure A4.4 – Normalised SRCD melting profiles (left) and derivative curves (right) for the melting of d(TAGGGTTA) (800 μM) in, (a) the absence of ligand; or in the presence of (b) 200 μM or (c) 800 μM Δ-[Ru(TAP)₂(dppz)]²⁺; (d) 200 μM or (e) 800 μM Λ-[Ru(TAP)₂(dppz)]²⁺.
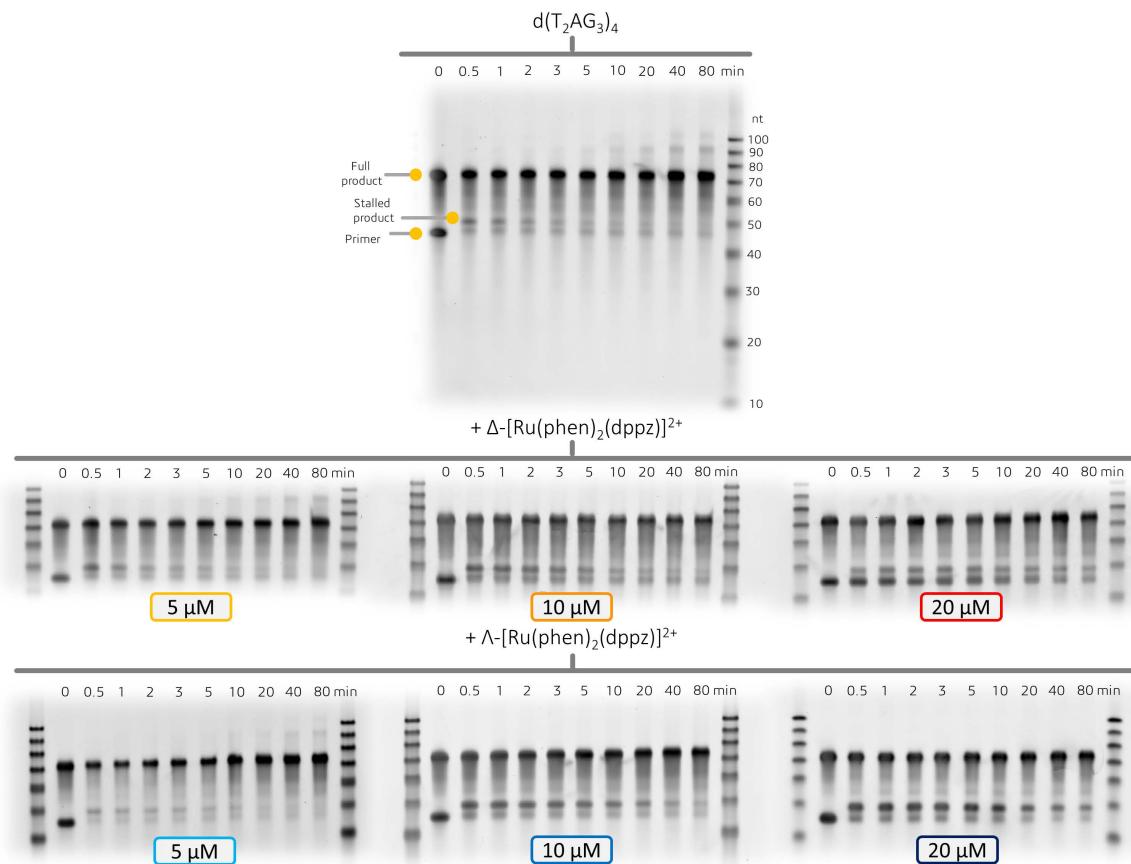
Figure A5.1 – *Denaturing PAGE electrograms that each show the progression of the replication of the G-quadruplex formed by d(T$_2$AG$_3$)$_4$ in the overhang region of a template strand whilst in the presence of 100 mM KCl and 5/10/20 µM Δ/Λ/rac-[Ru(phen)$_2$(dppz)]$^{2+}$. Each gel has been stained with SYBR gold so that non-FAM-labelled DNA (such as the DNA ladders) could be visualised.*
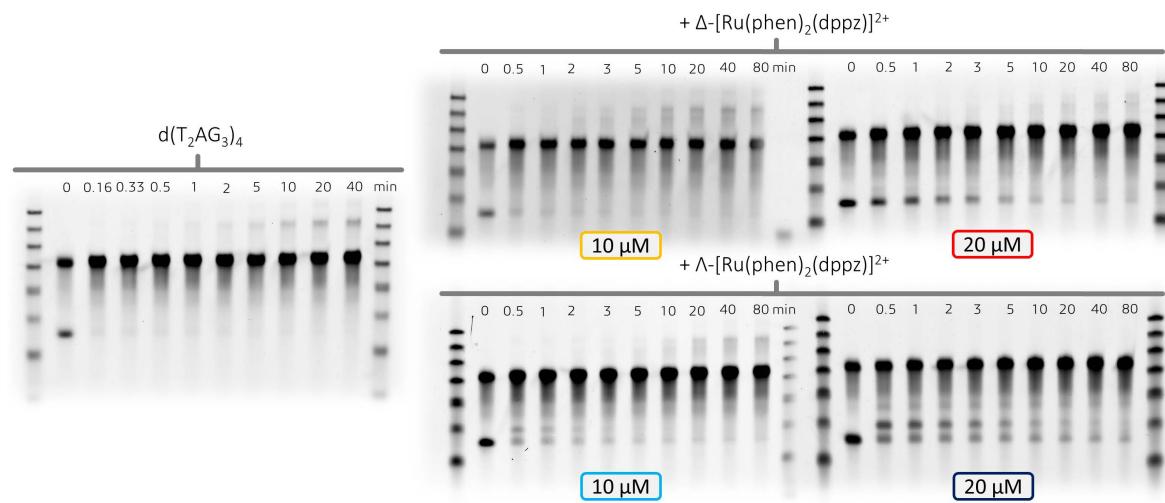
Figure A5.2 – *Denaturing PAGE electrograms that each show the progression of the replication of the G-quadruplex formed by d(T$_2$AG$_3$)$_4$ in the overhang region of a template strand whilst in the presence of 1 mM KCl and 5/10/20 µM Δ/Λ/rac-[Ru(phen)$_2$(dppz)]$^{2+}$. Each gel has been stained with SYBR gold so that non-FAM-labelled DNA (such as the DNA ladders) could be visualised.*
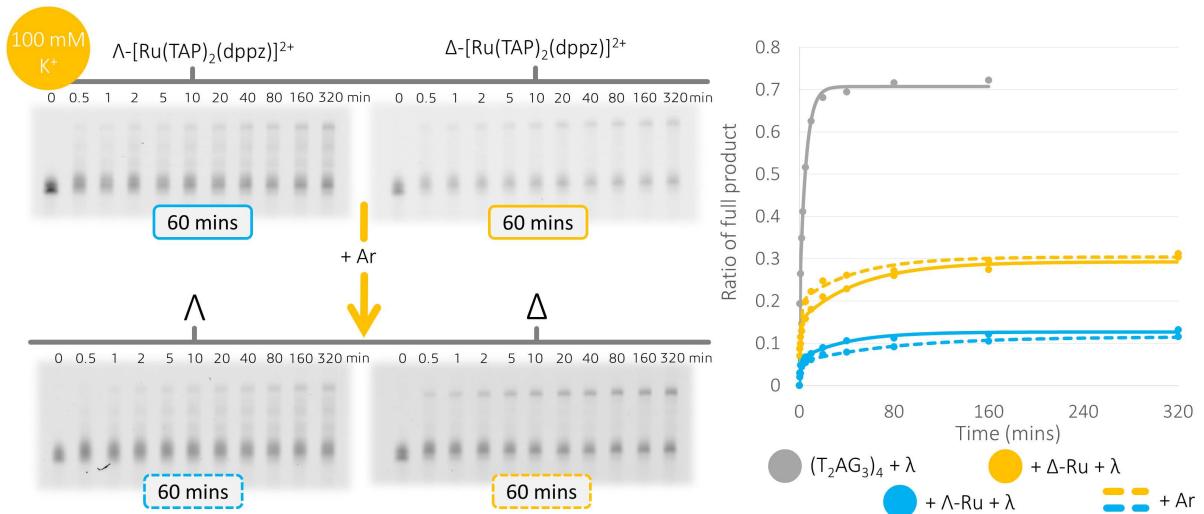
Figure A5.3 – *Denaturing PAGE electrograms that each show the progression of the replication of the G-quadruplex formed by d(T₂AG₃)₄ in the overhang region of a template strand whilst in the presence of __100 mM__ KCl and 20 μM Δ/Λ/rac-[Ru(TAP)₂(dppz)]²⁺. Each polymerase experiment was conducted following 60 mins of irradiation at 420 nm (except DNA by itself which was irradiated for 320 mins to ensure direct irradiation damage was absent). Inset shows the rate analysis for each of the reactions. Both experiments with either enantiomer were repeated in the presence of an argon atmosphere (dashed lines) to ensure ¹O₂ production was not the determining mode of action.*
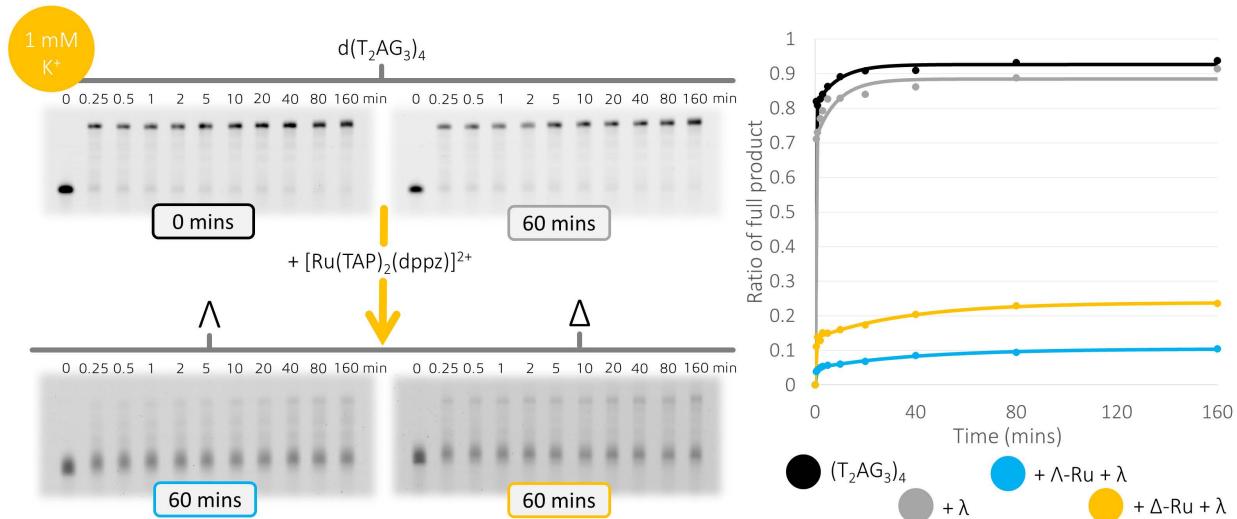
Figure A5.4 – *Denaturing PAGE electrograms that each show the progression of the replication of the G-quadruplex formed by d(T₂AG₃)₄ in the overhang region of a template strand whilst in the presence of __1 mM__ KCl and 20 μM Δ/Λ/rac-[Ru(TAP)₂(dppz)]²⁺. Each polymerase experiment was conducted following 60 mins of irradiation at 420 nm (except the first electrograph which was used as a comparison to ensure that direct irradiation of the DNA had little to no effect on subsequent processivity). Inset shows the rate analysis for each of the reactions. Both experiments with either enantiomer were repeated in the presence of an argon atmosphere (dashed lines) to ensure ¹O₂ production was not the determining mode of action.*
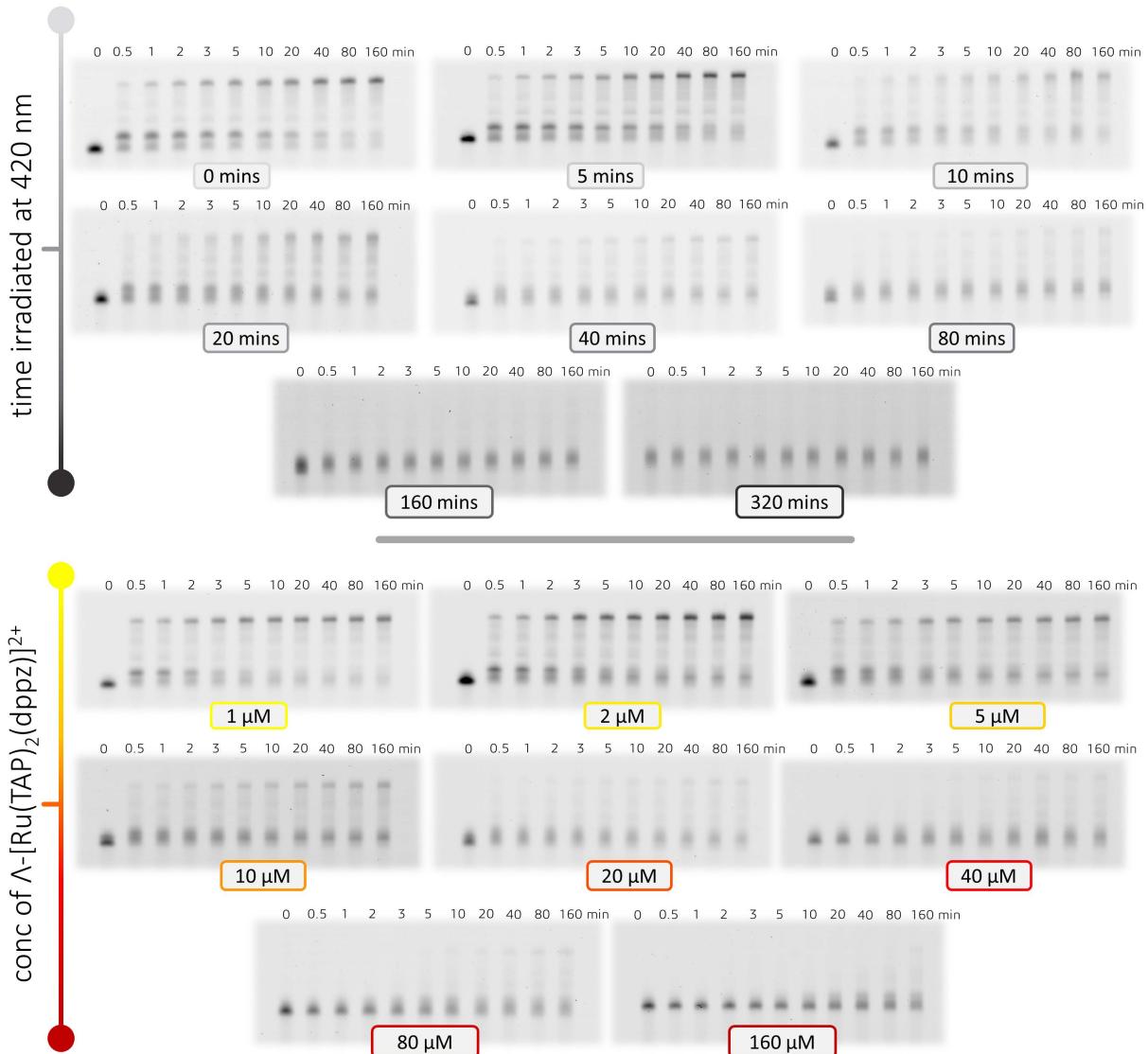
Figure A5.5 – *Denaturing PAGE gels that each show the progression of the replication of the G-quadruplex formed by d(T$_2$AG$_3$)$_4$ (1 μM) in the overhang region of a template strand whilst in the presence of Λ-[Ru(TAP)$_2$(dppz)]$^{2+}$. Each gel differs in the fact that the reaction mixtures in each case were either irradiated at 412 nm for the displayed amount of time before introduction of the KF exo polymerase (0, 5, 10, 20, 40, 80, 160, and 320 mins) and the concentration was kept constant at 20 μM; or for the second set, was irradiated for 60 mins each but the concentration of ruthenium was altered (1, 2, 5, 10, 20, 40, 80, or 160 μM) . Note the large reduction in the formation of product on either an increase in conc or irradiation time. In addition, the stalled phases become less discrete (see first lanes especially) and an increase in the molecular weight of the primer strand is observed which implies adduct formation.*

239
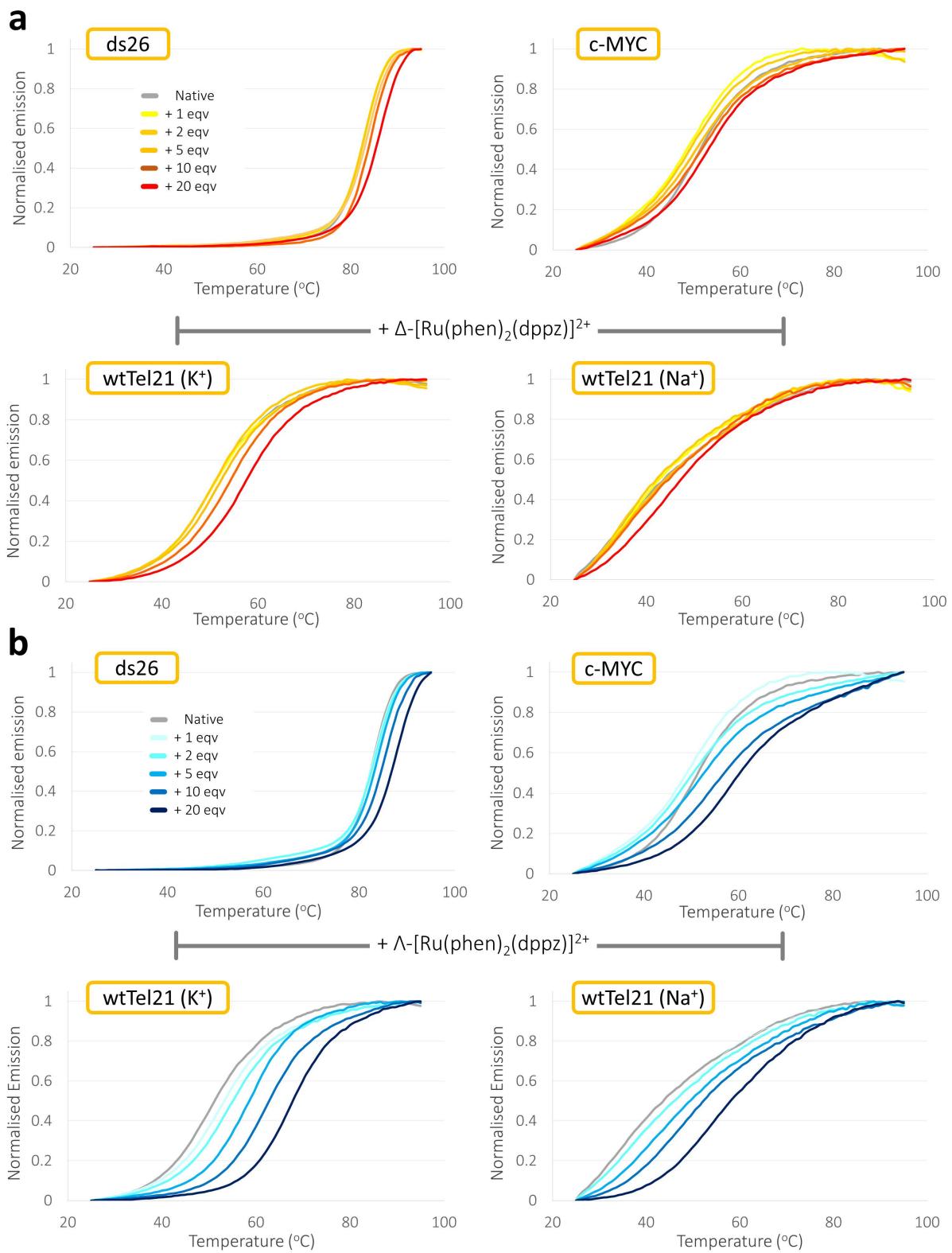
Figure A5.6 – *FRET melting endotherms for (a) Δ-[Ru(phen)₂(dppz)]²⁺, and (b) Λ-[Ru(phen)₂(dppz)]²⁺ at varying concentrations equating to 0,1,2,5,10, and 20 equivalents to biological unit, in the presence of either ds26, c-MYC, wtTel21 (K⁺), or wtTel21(Na⁺). For full experimental details, buffer etc, see section 5.2.6).*
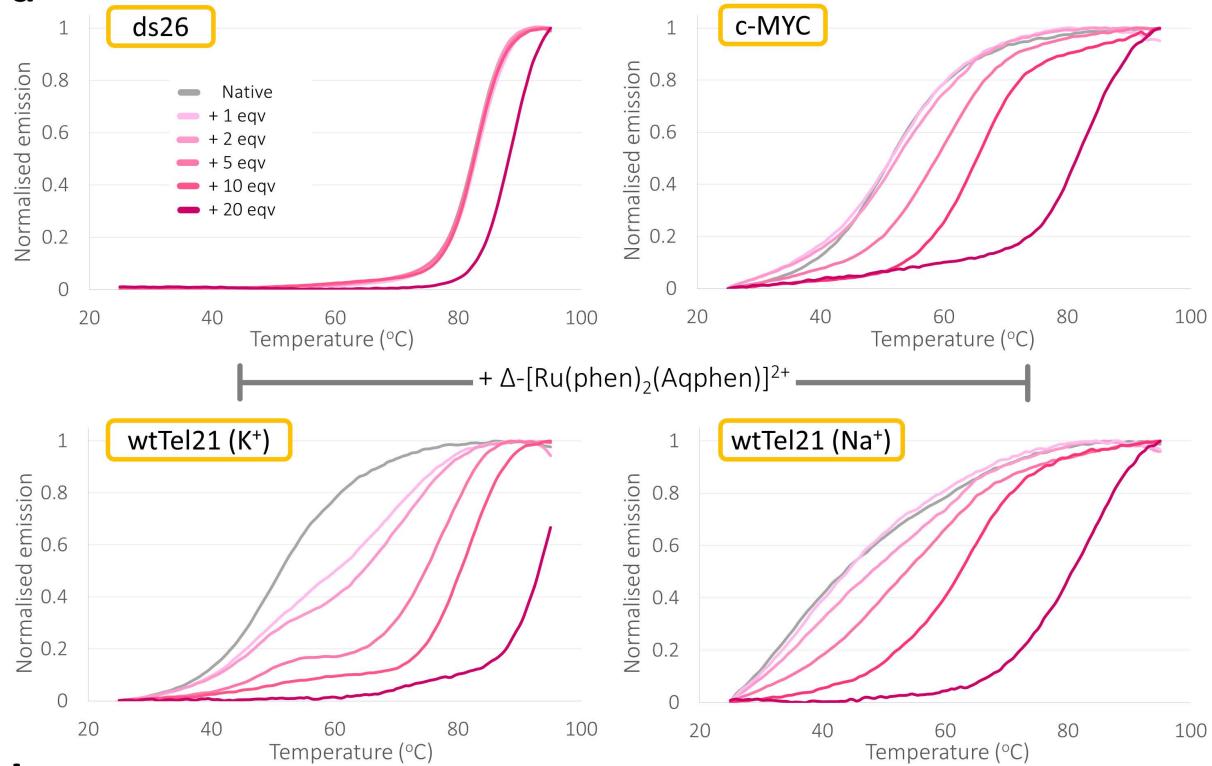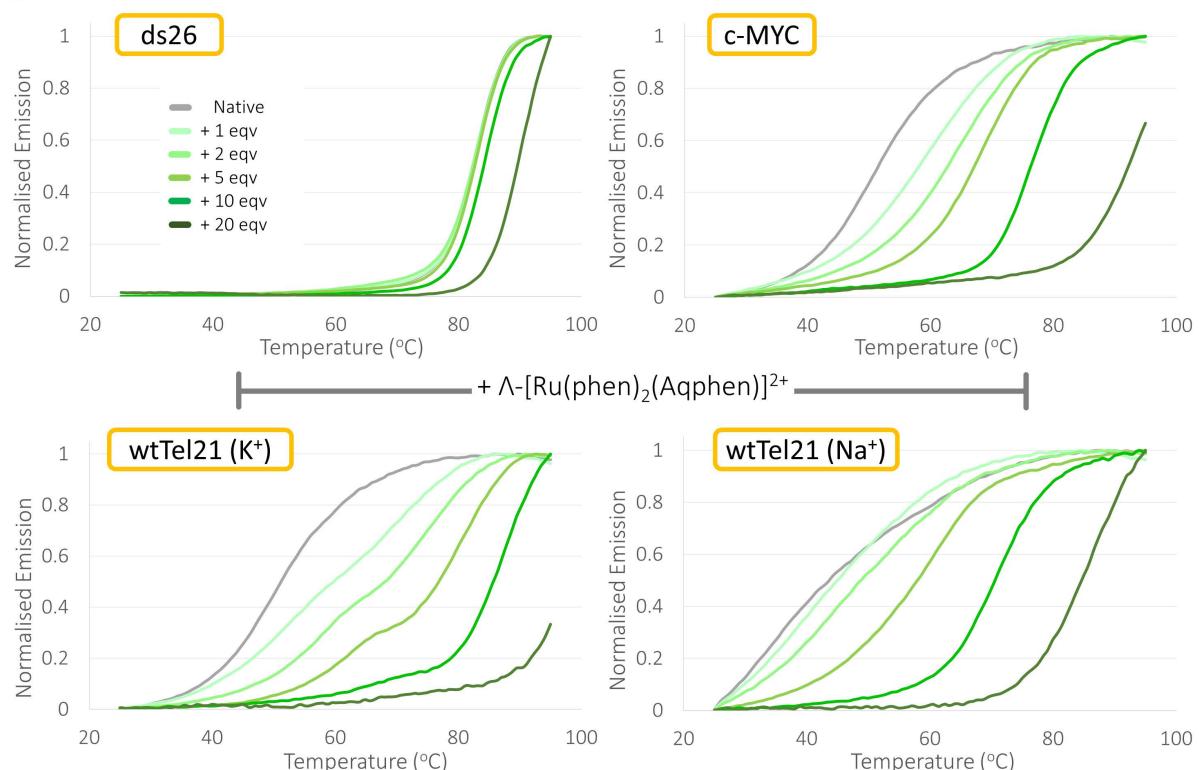
Figure A5.7 — *FRET melting endotherms for (a) Δ-[Ru(phen)₂(Aqphen)]²⁺, and (b) Λ-[Ru(phen)₂(Aqphen)]²⁺ at varying concentrations equating to 0,1,2,5,10, and 20 equivalents to biological unit, in the presence of either ds26, c-MYC, wtTel21 (K⁺), or wtTel21(Na⁺). For full experimental details, buffer etc, see section 5.2.6). Note that for endotherms that did not reach Δy/Δx=0, their normalised intensities were estimated.*
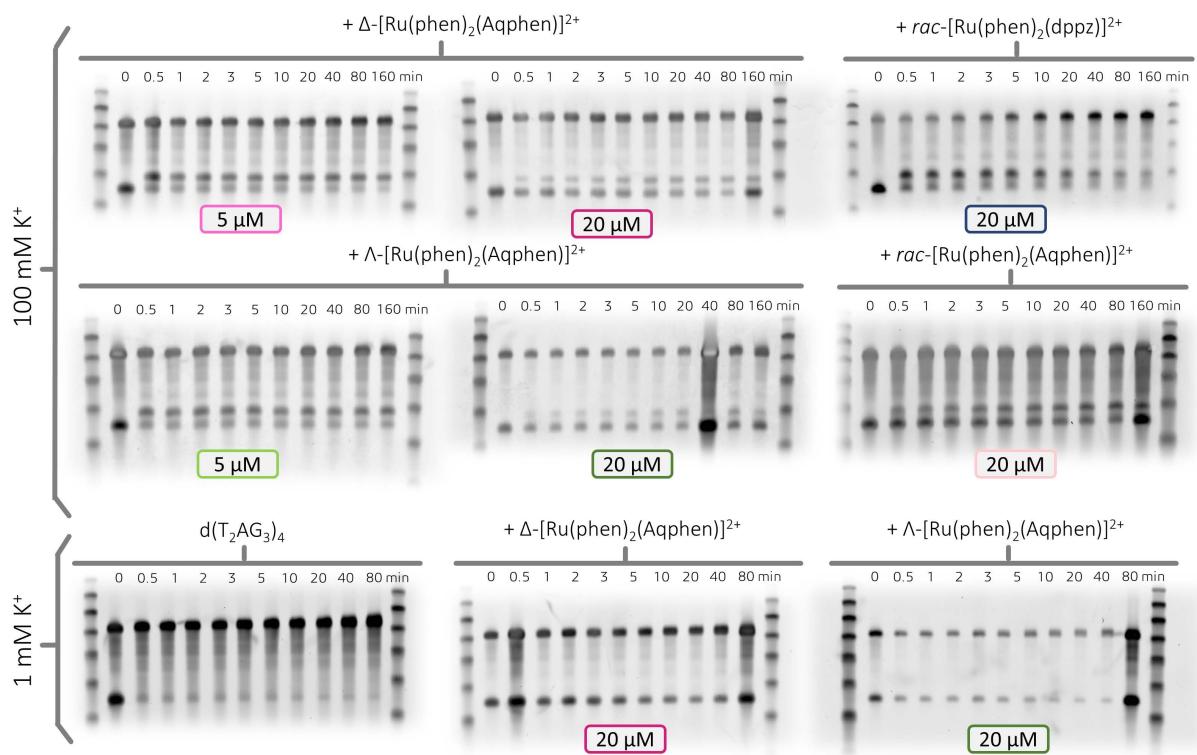
Figure A5.8 – *Denaturing PAGE electrograms that each show the progression of the replication of the G-quadruplex formed by d(TA$_2$G$_3$)$_4$ in the overhang region of a template strand whilst in the presence of 100 mM or 1 mM K$^+$ and 5/20 μM Δ/Λ/rac-[Ru(phen)$_2$(Aqphen)]$^{2+}$. Each gel has been stained with SYBR gold so that non-FAM-labelled DNA (such as the DNA ladders) could be visualised.*
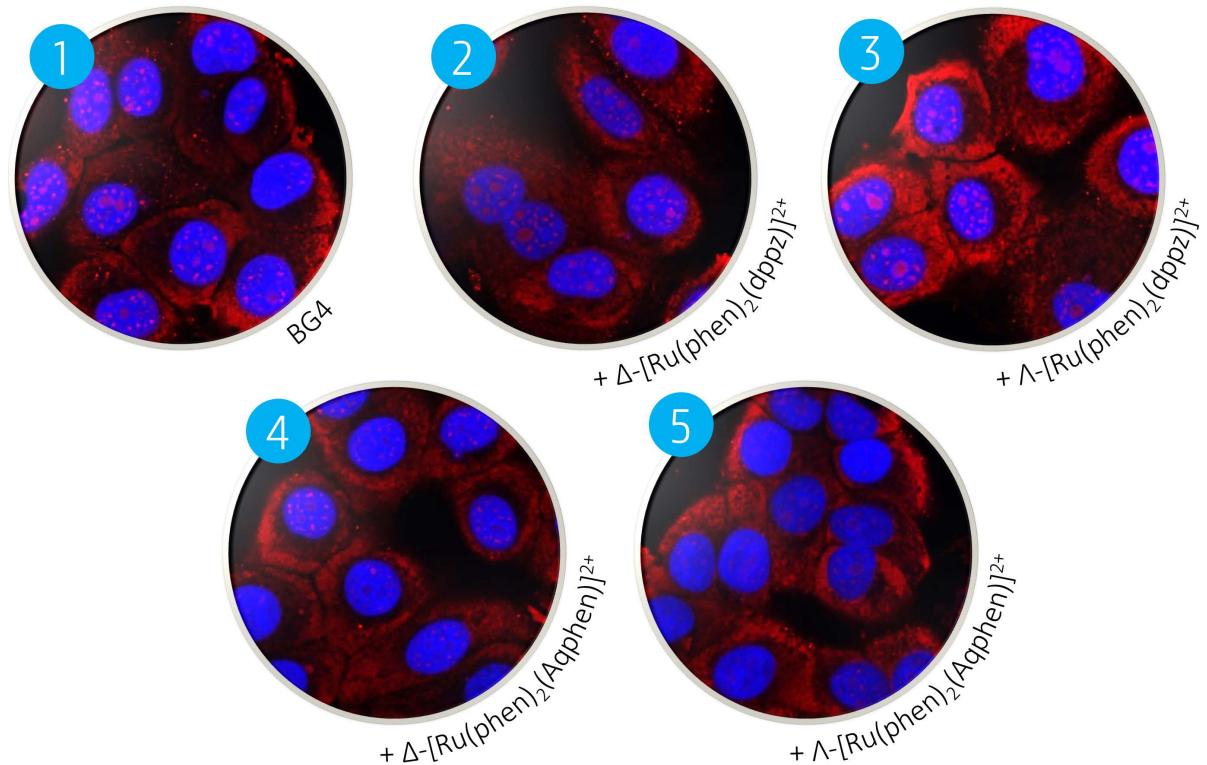
Figure A5.9 – *Confocal micrographs (before foci counting with ZEN) of MCF-7 cells inoculated with modified fluorescent BG4 antibody (red) and counterstained with DAPI nuclei stain (blue). Cells were incubated either in the absence (1) or presence of ruthenium complexes; Δ/Λ-[Ru(phen)$_2$(dppz)]$^{2+}$ (2/3), or Δ/Λ-[Ru(phen)$_2$(Aqphen)]$^{2+}$ (4/5). Note the stark reduction in foci intensity in the nucleus of cells incubated with the Aqphen complexes in comparison to the dppz variants.*
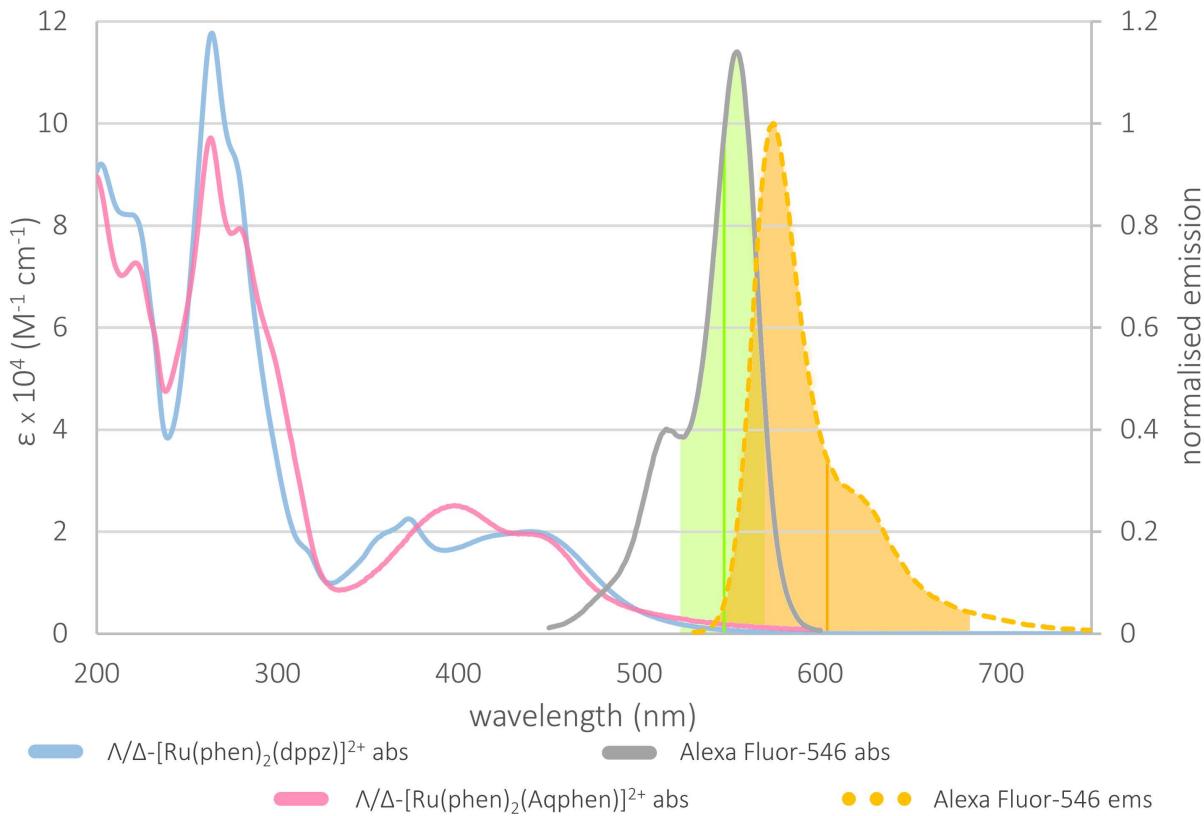
Figure A5.10 – *UV-Vis absorbance or emission profiles for the molecular species utilised in conjunction with BG4 in the immunofluorescence assays. Continuous lines represent absorbance profiles whereas the dotted line represents the emission profile. Shaded areas represent the respective excitation (λ = 545 ±25 nm) or emission (λ = 605 ±70 nm) bandwidth filters applied in the confocal microscopy experiments. Note the very small (but apparent) overlap of the experimental excitation band and the absorption spectra of the complexes.*
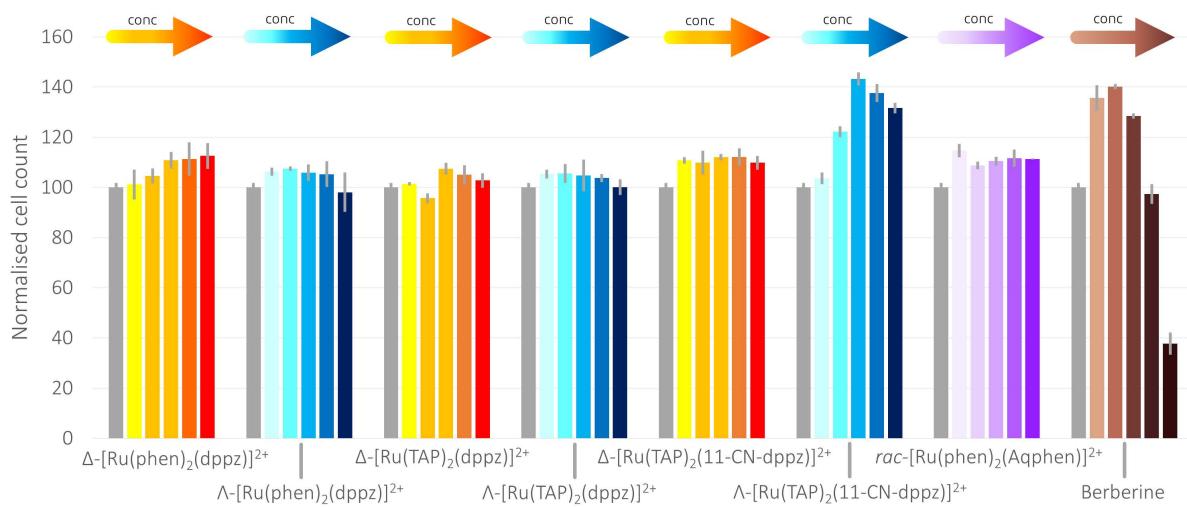
Figure A5.11 – *HeLa cell counting in the absence or presence of a range of ruthenium complexes at a range of concentrations (0, 1.25, 2.5, 5, 10, and 20 µM). Cells were incubated with the ligands for 24 hours in 96-well plates before cell counting was achieved using a commercial absorbance counting kit. Absorbances were normalised in relation to the control cells. Note that even at the highest concentrations the ruthenium complexes are non-toxic, whereas the at the higher concentrations the berberine begins to heavily affect the cells.*

Table A5.1 – *Normalised absorbance values for the counting of HeLa cells after incubation with a selection of ruthenium complexes.*

Complex	Normalised absorbance					
	0 µM	1.25 µM	2.5 µM	5 µM	10 µM	20 µM
Δ-[Ru(phen)$_2$(dppz)]$^{2+}$	100 ±1.75	101.11 ±5.97	104.55 ±2.99	110.80 ±3.31	111.32 ±6.66	112.55 ±5.12
Λ-[Ru(phen)$_2$(dppz)]$^{2+}$	100 ±1.75	106.26 ±1.60	107.40 ±0.89	105.84 ±5.08	105.26 ±5.08	98.03 ±7.91
Δ-[Ru(TAP)$_2$(dppz)]$^{2+}$	100 ±1.75	101.35 ±0.65	95.71 ±1.94	107.45 ±2.30	104.96 ±3.91	102.74 ±2.87
Λ-[Ru(TAP)$_2$(dppz)]$^{2+}$	100 ±1.75	105.32 ±1.79	105.50 ±1.79	104.72 ±6.40	103.71 ±1.61	100.06 ±3.13
Δ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$	100 ±1.75	110.75 ±1.33	109.83 ±4.71	112.04 ±1.28	112.08 ±3.51	109.81 ±2.74
Λ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$	100 ±1.75	103.57 ±2.34	122.21 ±2.12	143.199 ±2.69	137.61 ±3.61	131.57 ±2.09
rac-[Ru(phen)$_2$(Aqphen)]$^{2+}$	100 ±1.75	114.66 ±2.58	108.70 ±1.47	110.44 ±1.78	111.66 ±3.44	111.34 ±0.24
Berberine	100 ±1.75	135.69 ±5.04	140.15 ±1.11	128.43 ±0.97	97.35 ±3.92	37.73 ±4.44
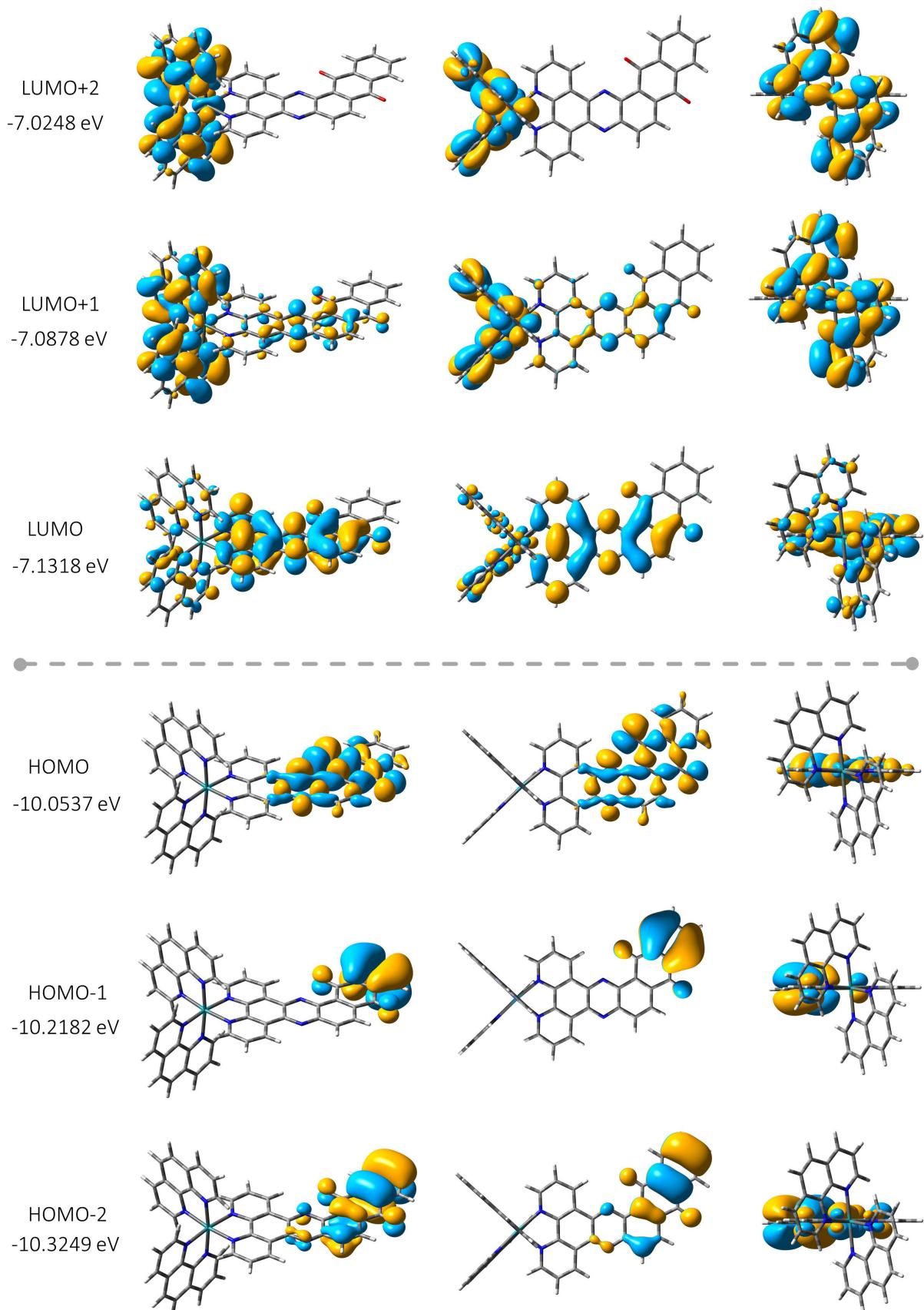
Figure A5.12 – *Selected frontier molecular orbitals of [Ru(phen)₂(Aqphen)]²⁺ calculated at a DFT B3LYP/LAN2LZ level.*

LUMO+2
-7.0248 eV

LUMO+1
-7.0878 eV

LUMO
-7.1318 eV

HOMO
-10.0537 eV

HOMO-1
-10.2182 eV

HOMO-2
-10.3249 eV
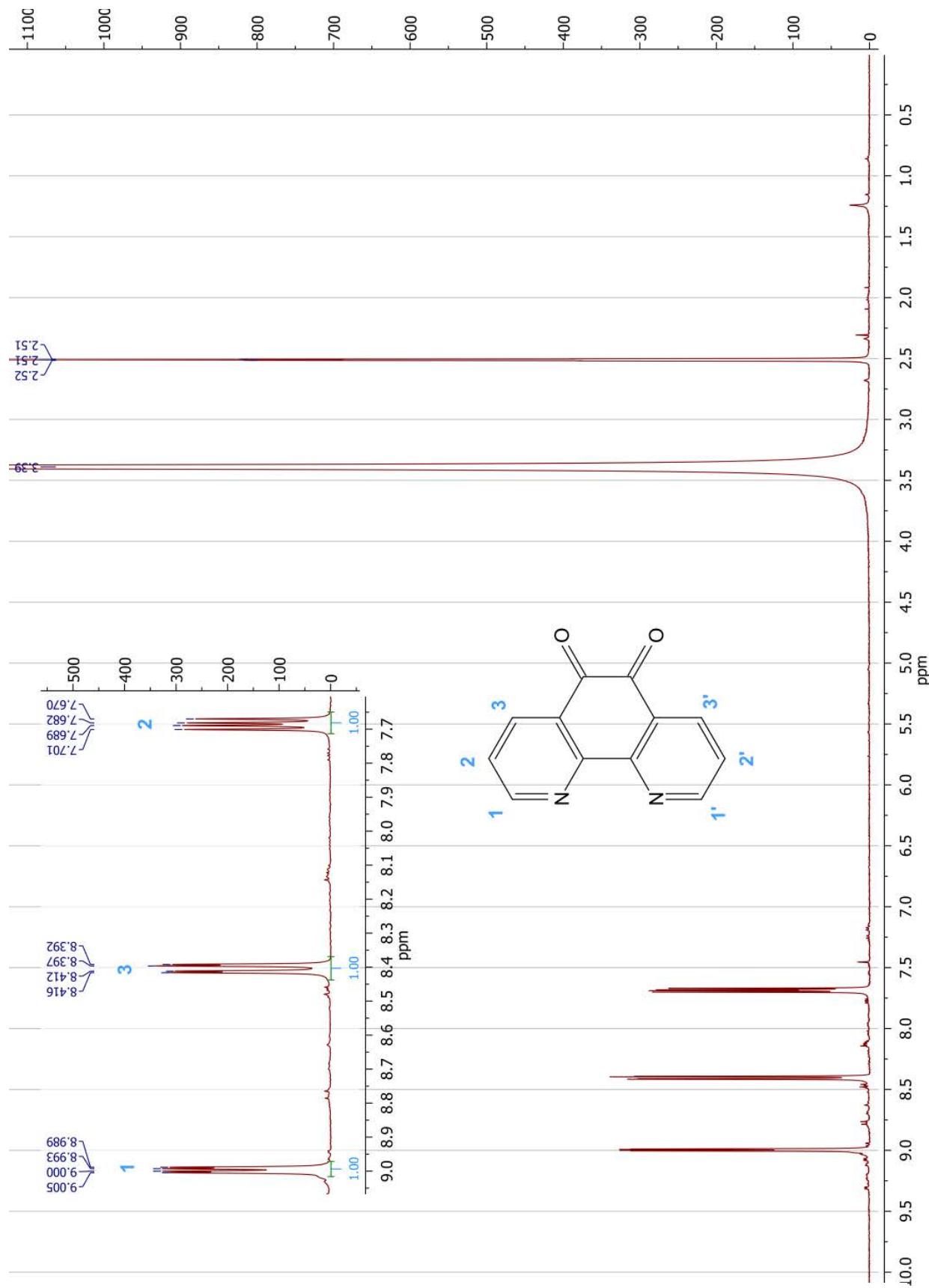
Figure A7.1 – ^{1}H NMR of phendione.
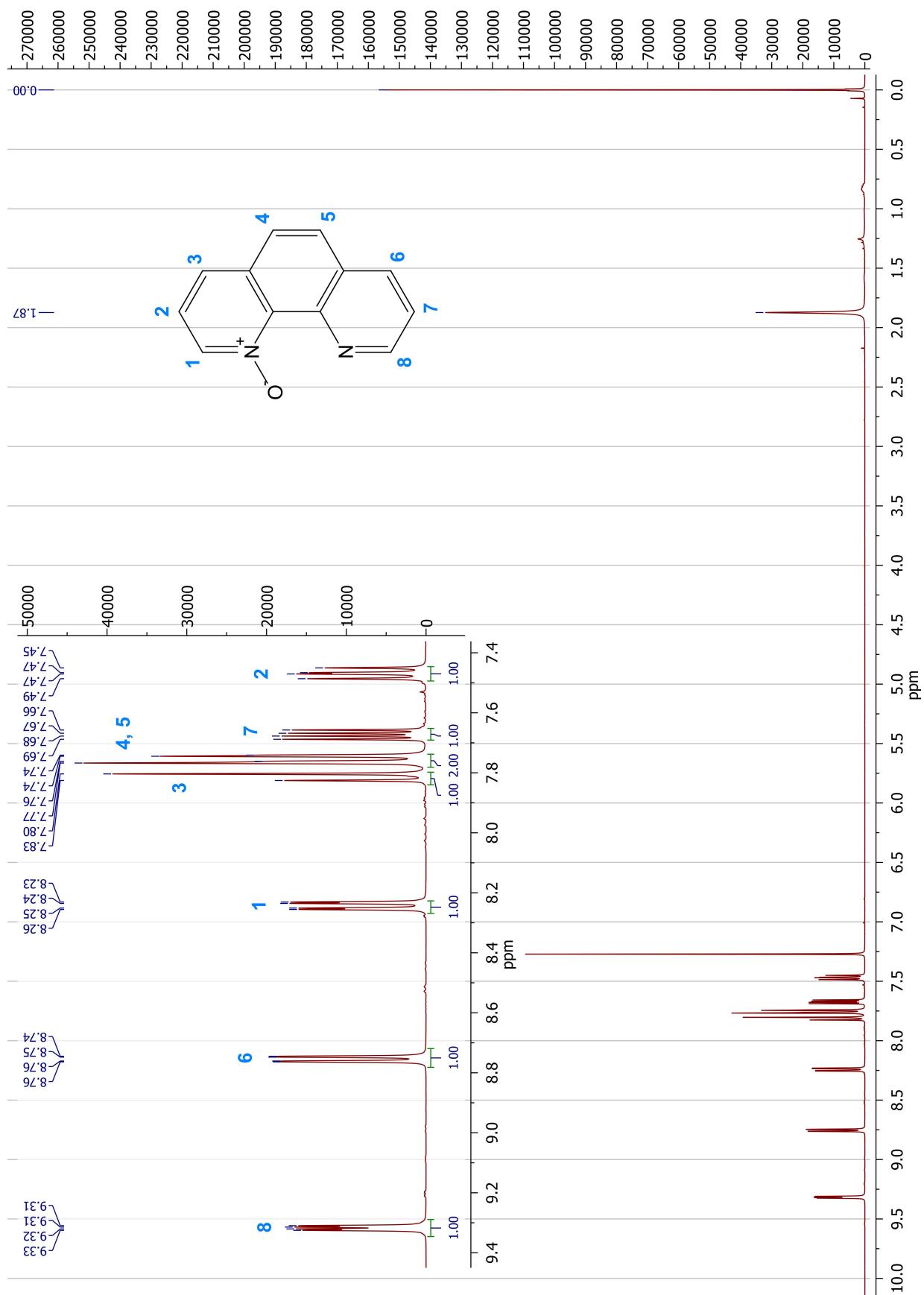
Figure A7.2 – ^1H NMR 1,10-phenanthroline-1-oxide.
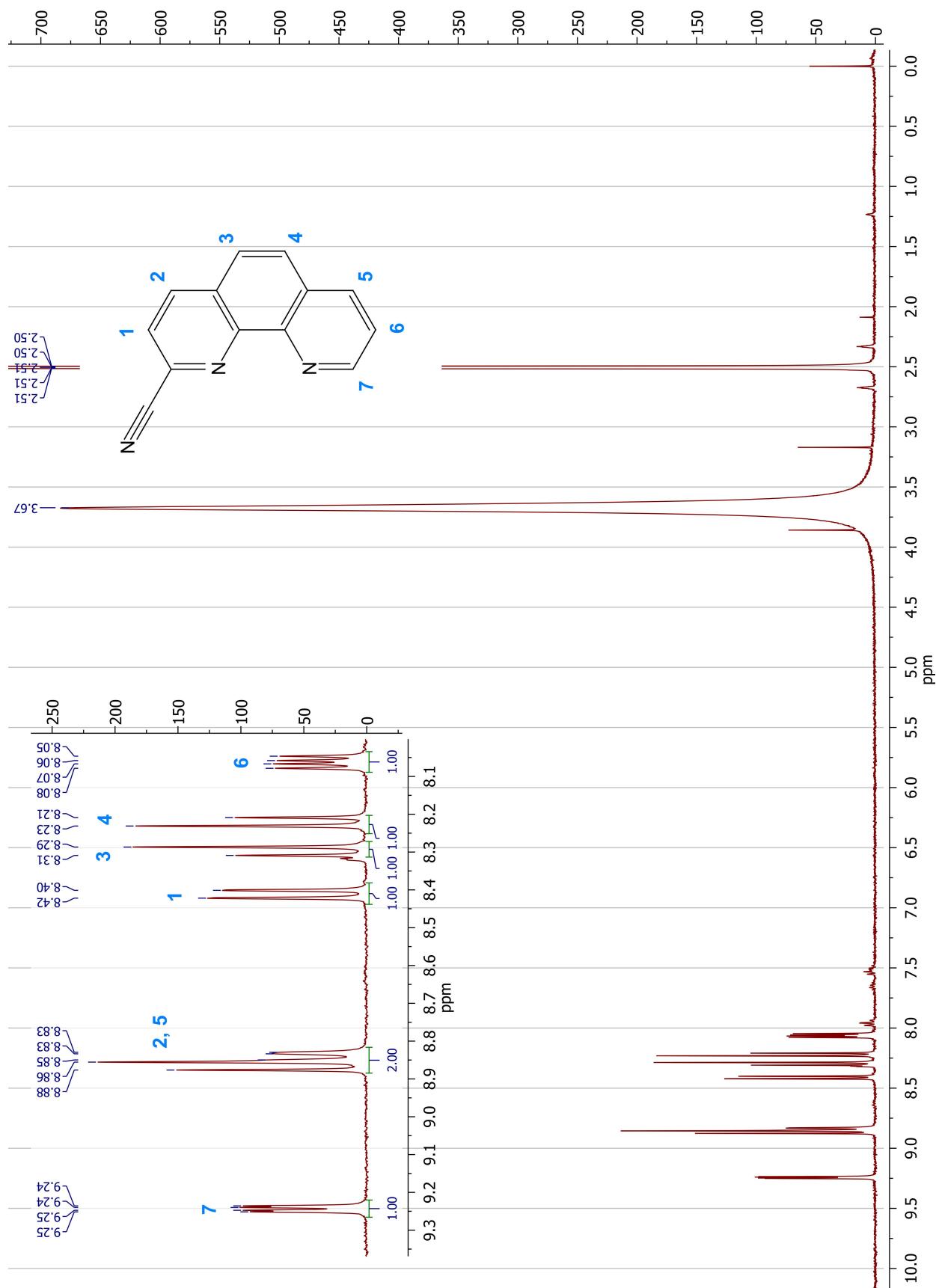
Figure A7.3 – ¹H NMR 2-cyana-1,10-phanthroline (2-CN-phen).
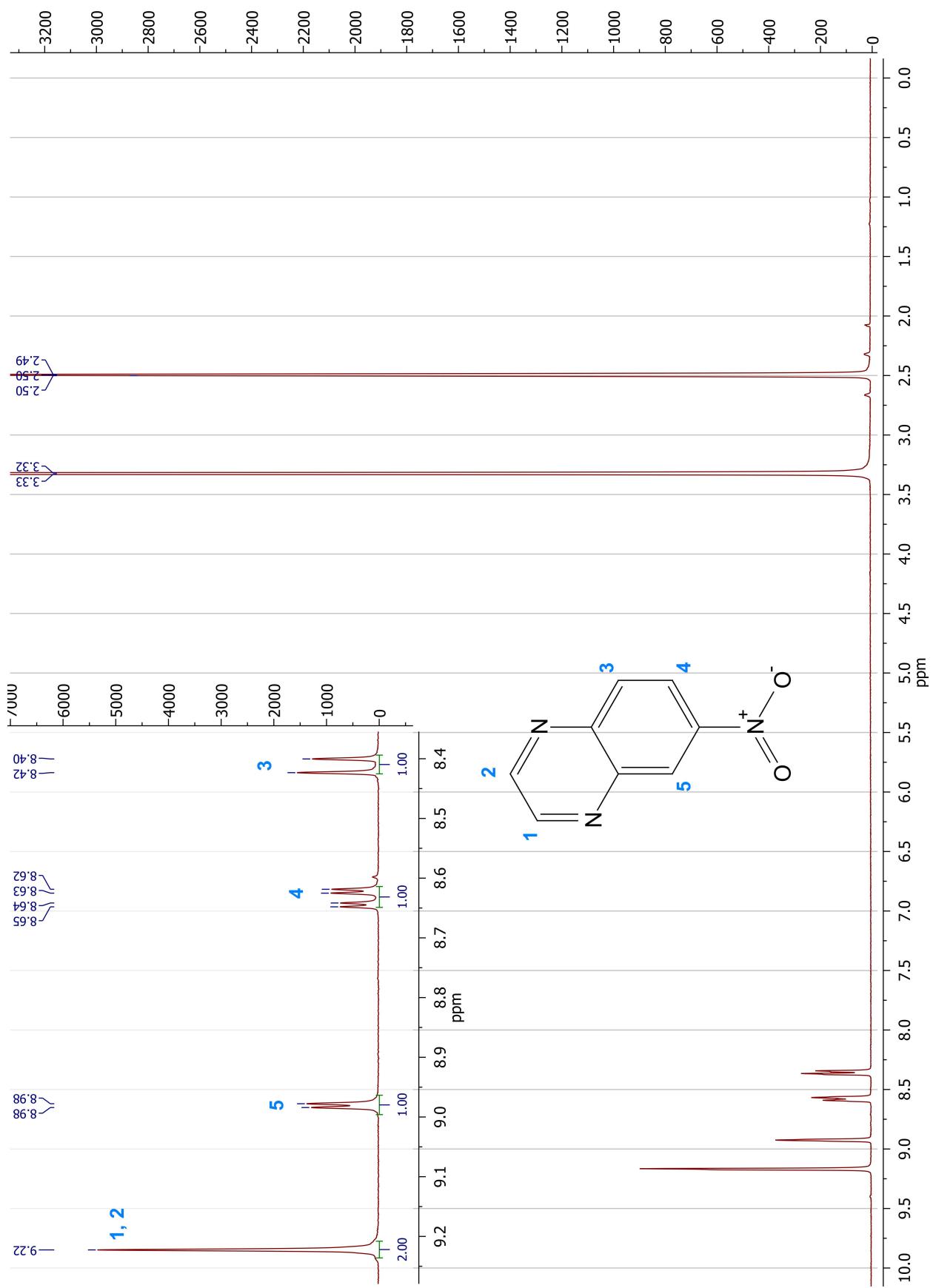
Figure A7.4 – ¹H NMR of 6-nitroquinoxaline.
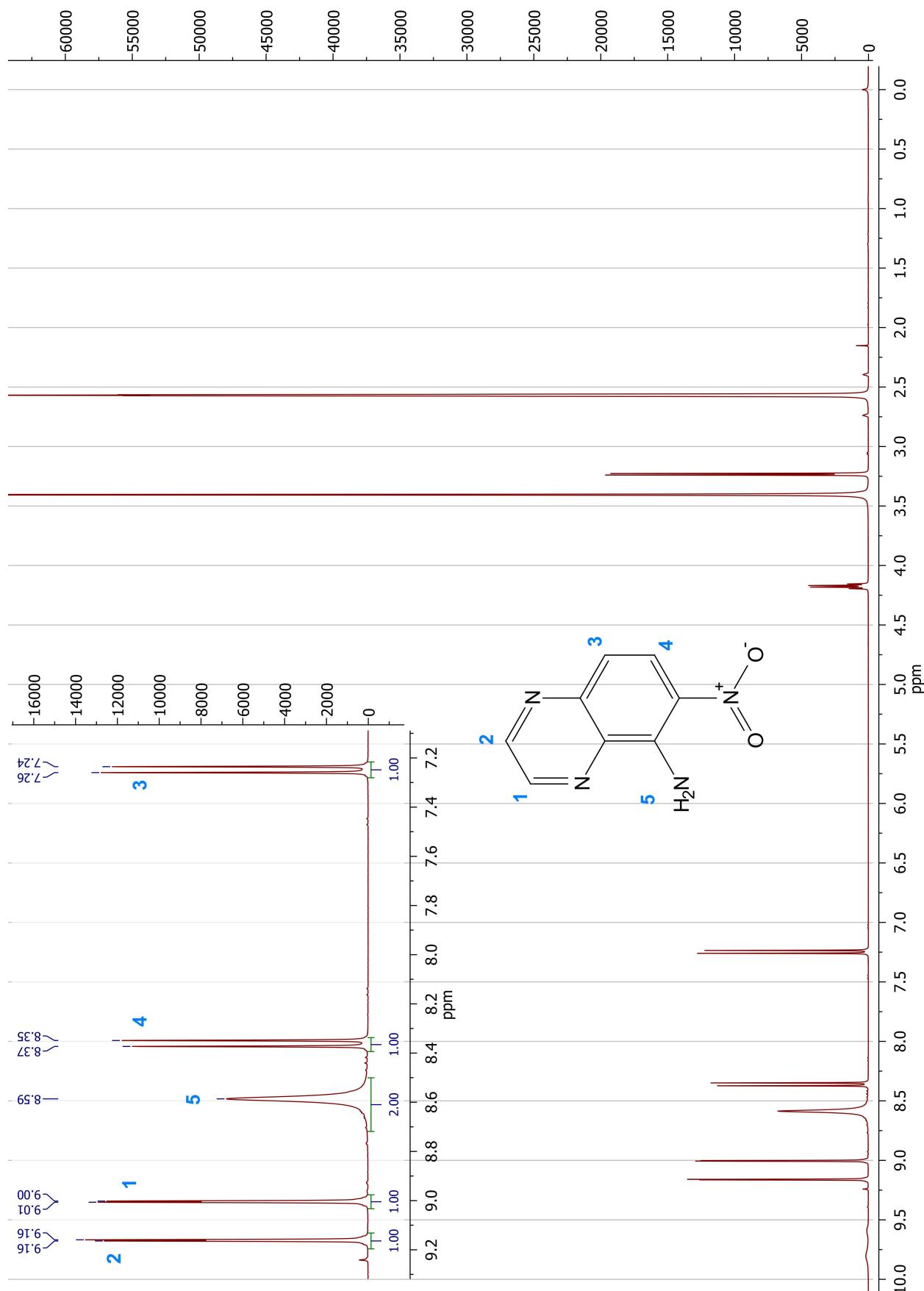
Figure A7.5 – ¹H NMR of 6-nitroquinoxaline-5-amine.
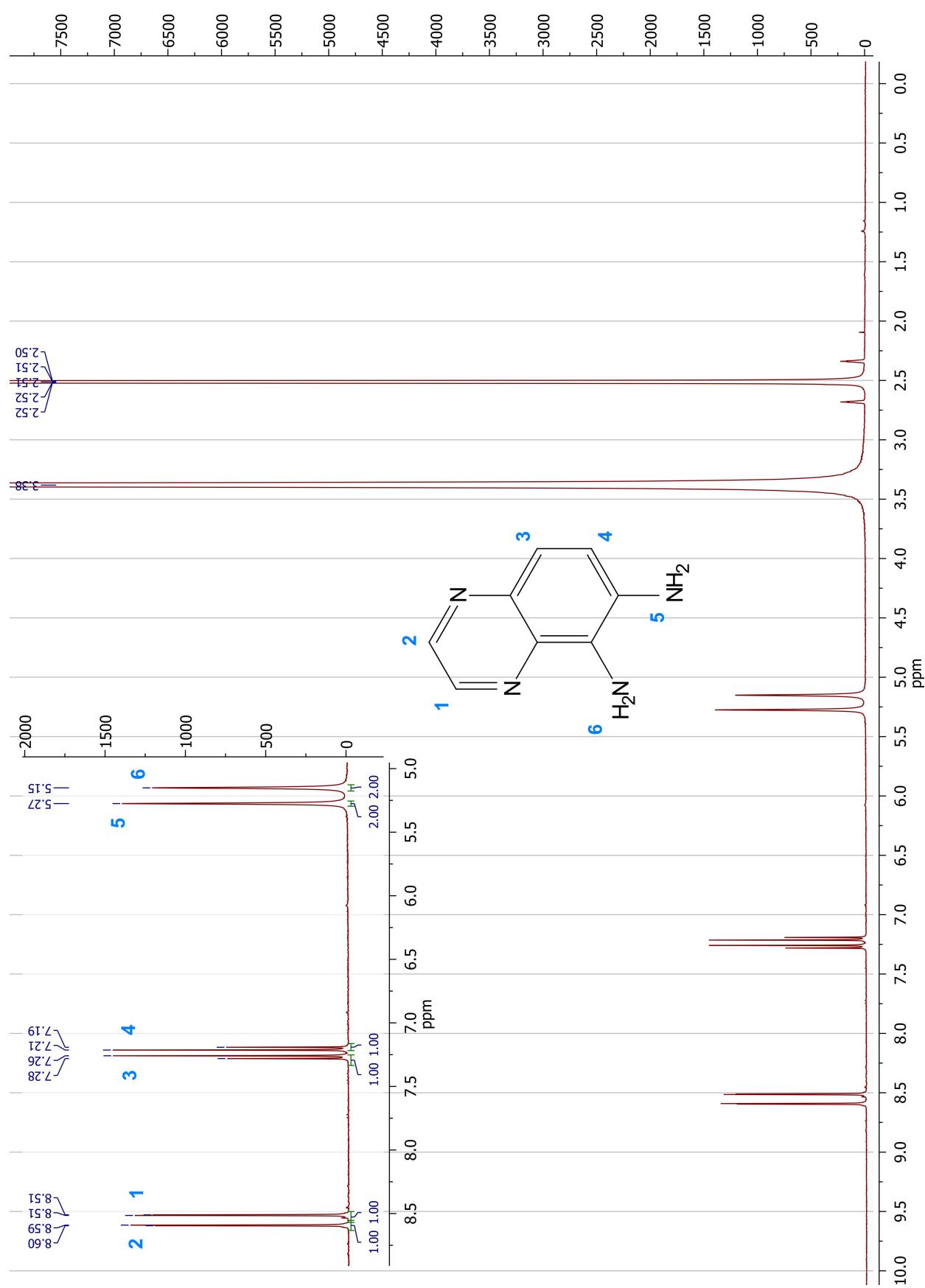
Figure A7.6 — *^1H NMR of quinoxaline-5,6-diamine.*
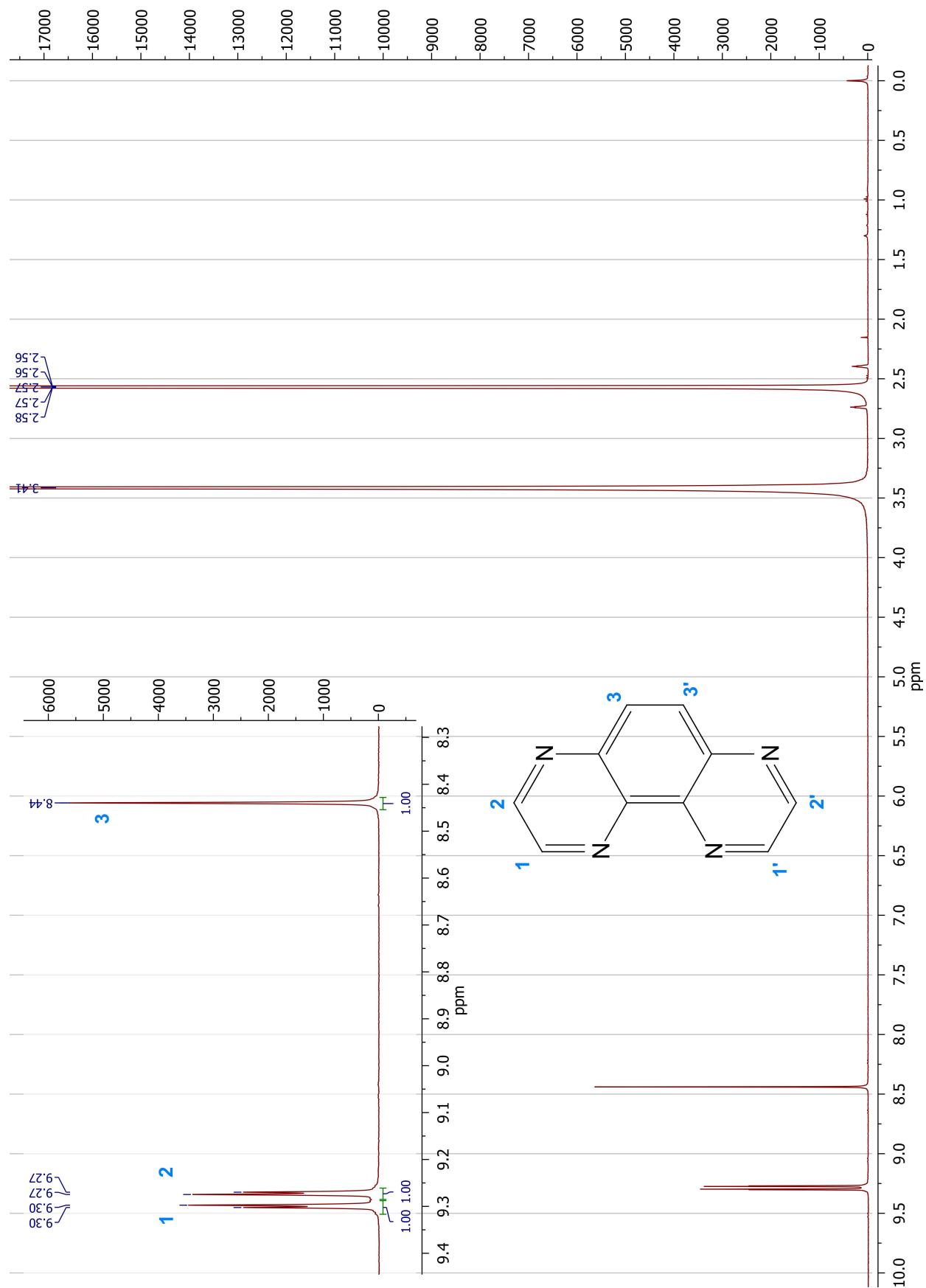
Figure A7.7 – ^1H NMR of 1,4,5,8-tetraazaphenanthrene (TAP).
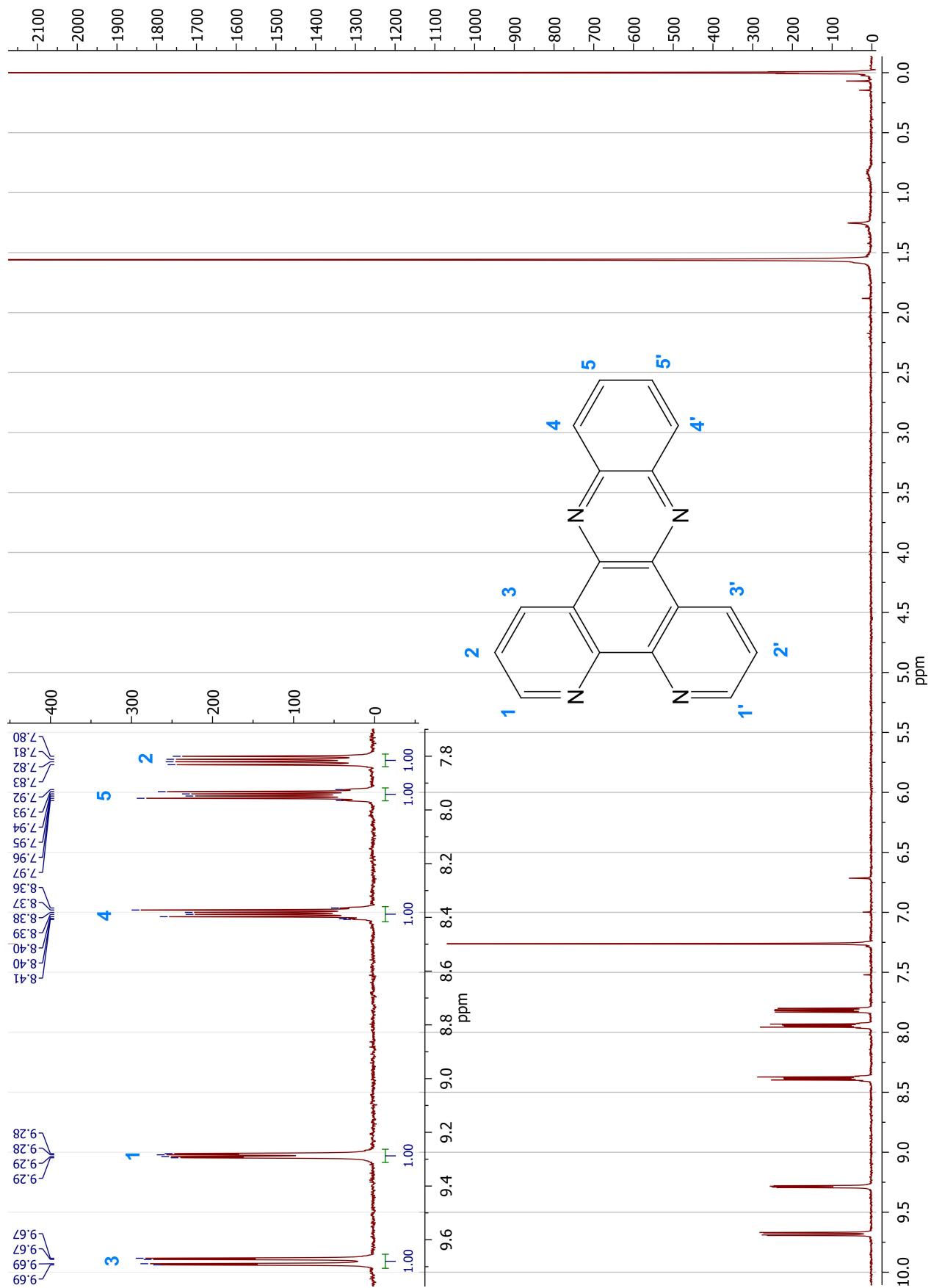
Figure A7.8 *– ¹H NMR of dipyrido[3,2-a:2',3'-c]phenazine.*
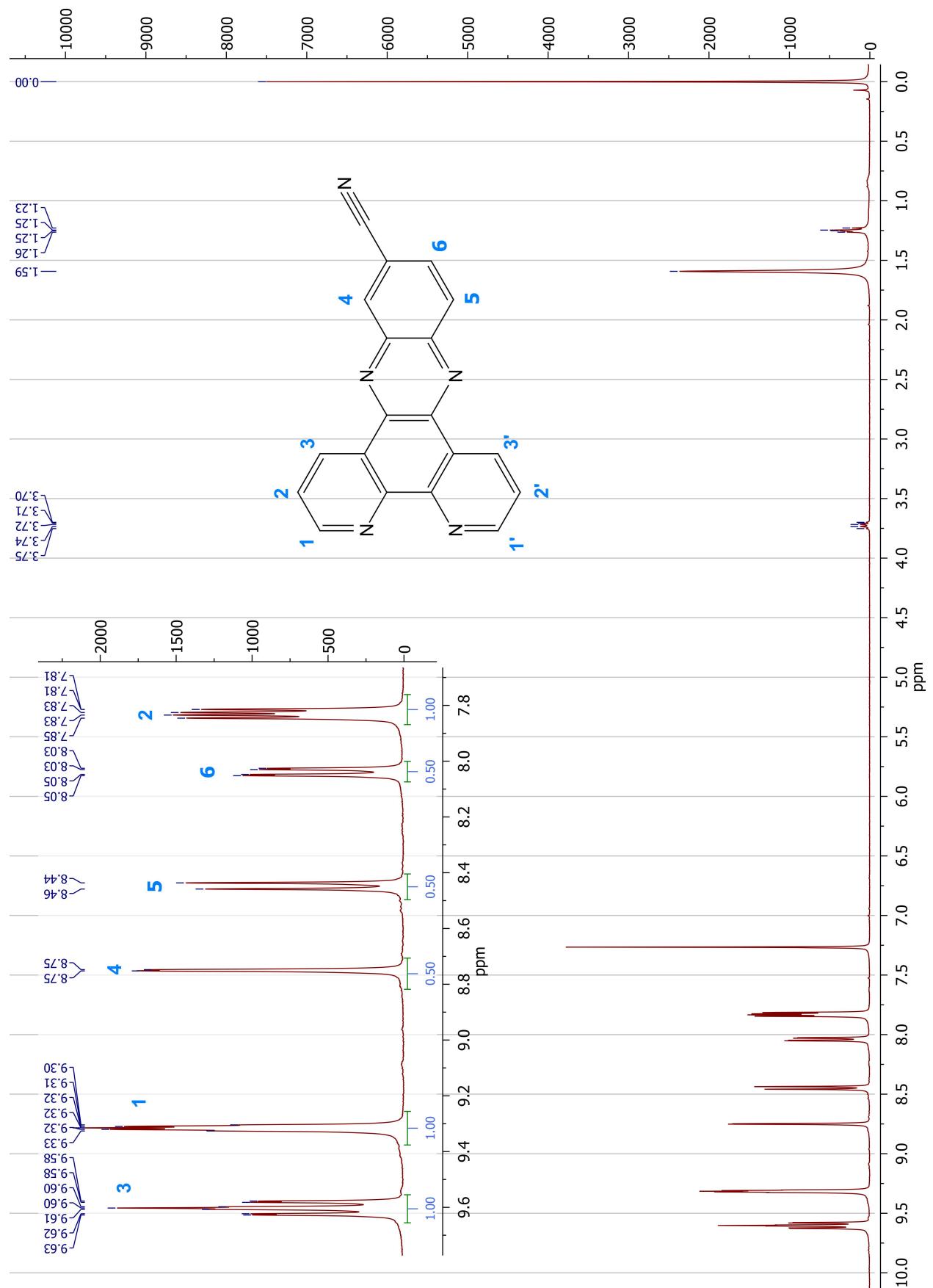
Figure A7.9 — ¹H NMR of 11-cyano-dipyridophenazine (11-CN-dppz).
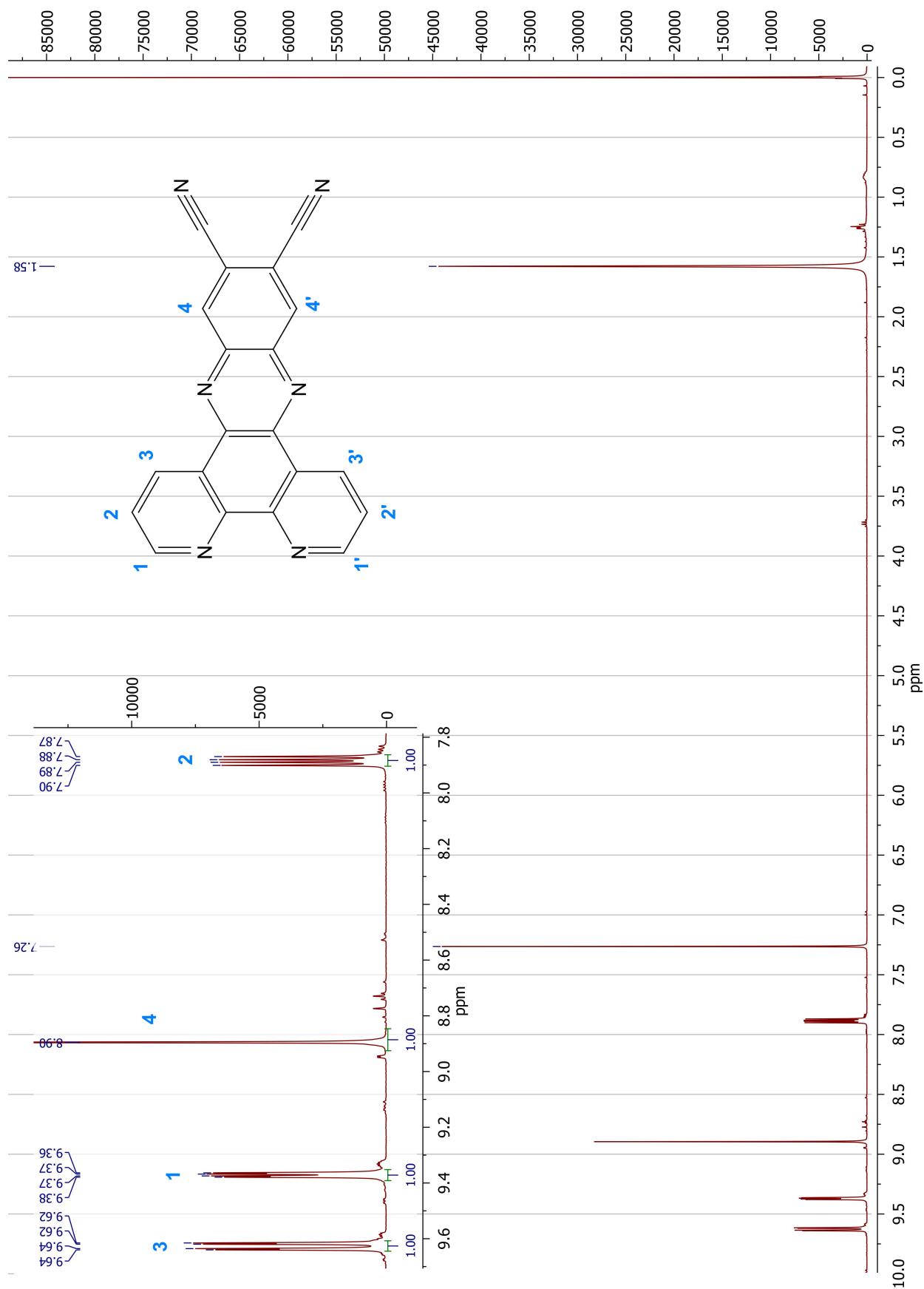
Figure A7.10 – ^{1}H NMR of 11,12-dicyano-dipyridophenazine (11,12-CN-dppz).
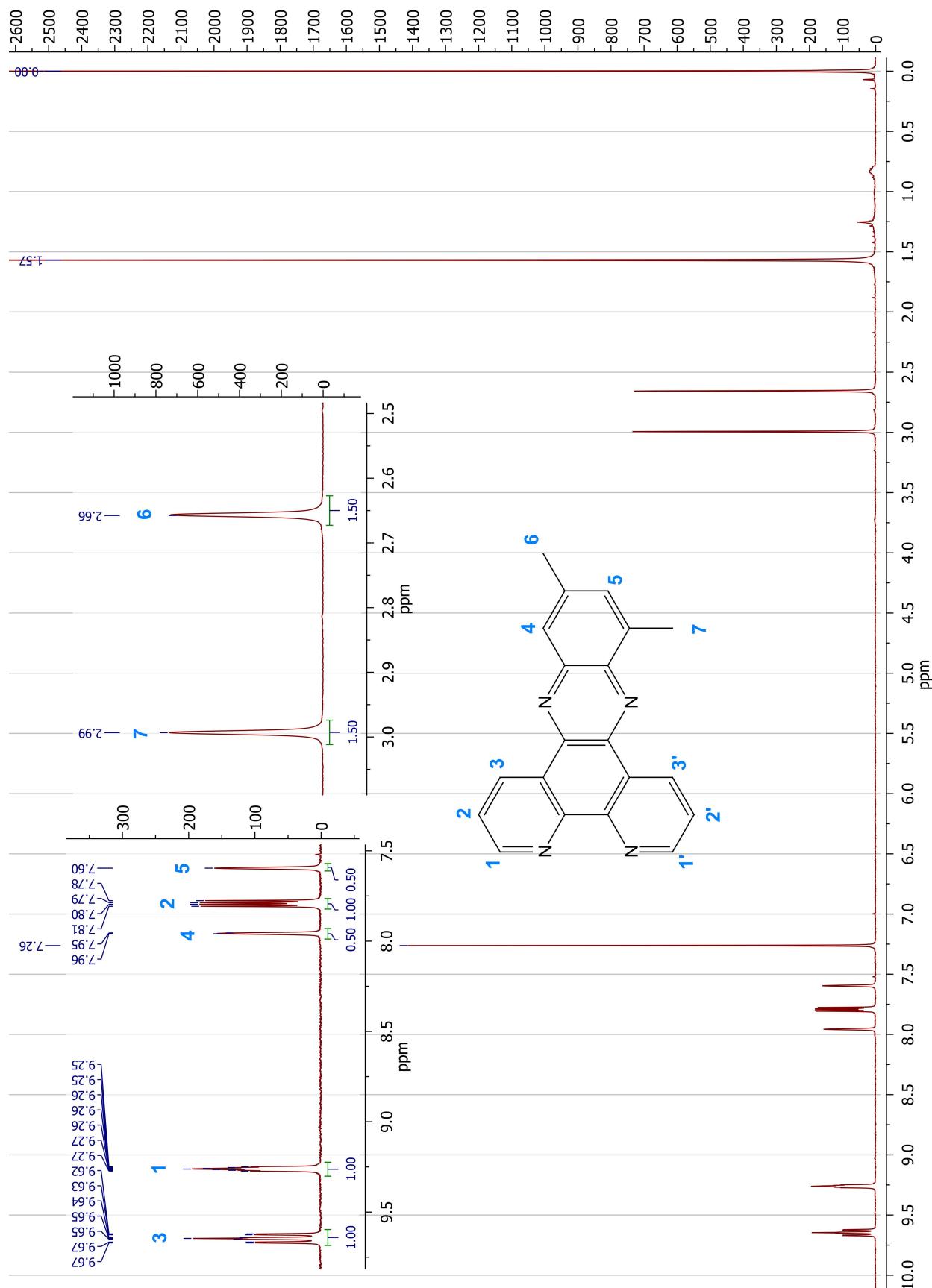
Figure A7.11 – *^{1}H NMR of 10,12-dimethyl-dipyridophenazine (10,12-Me-dppz).*
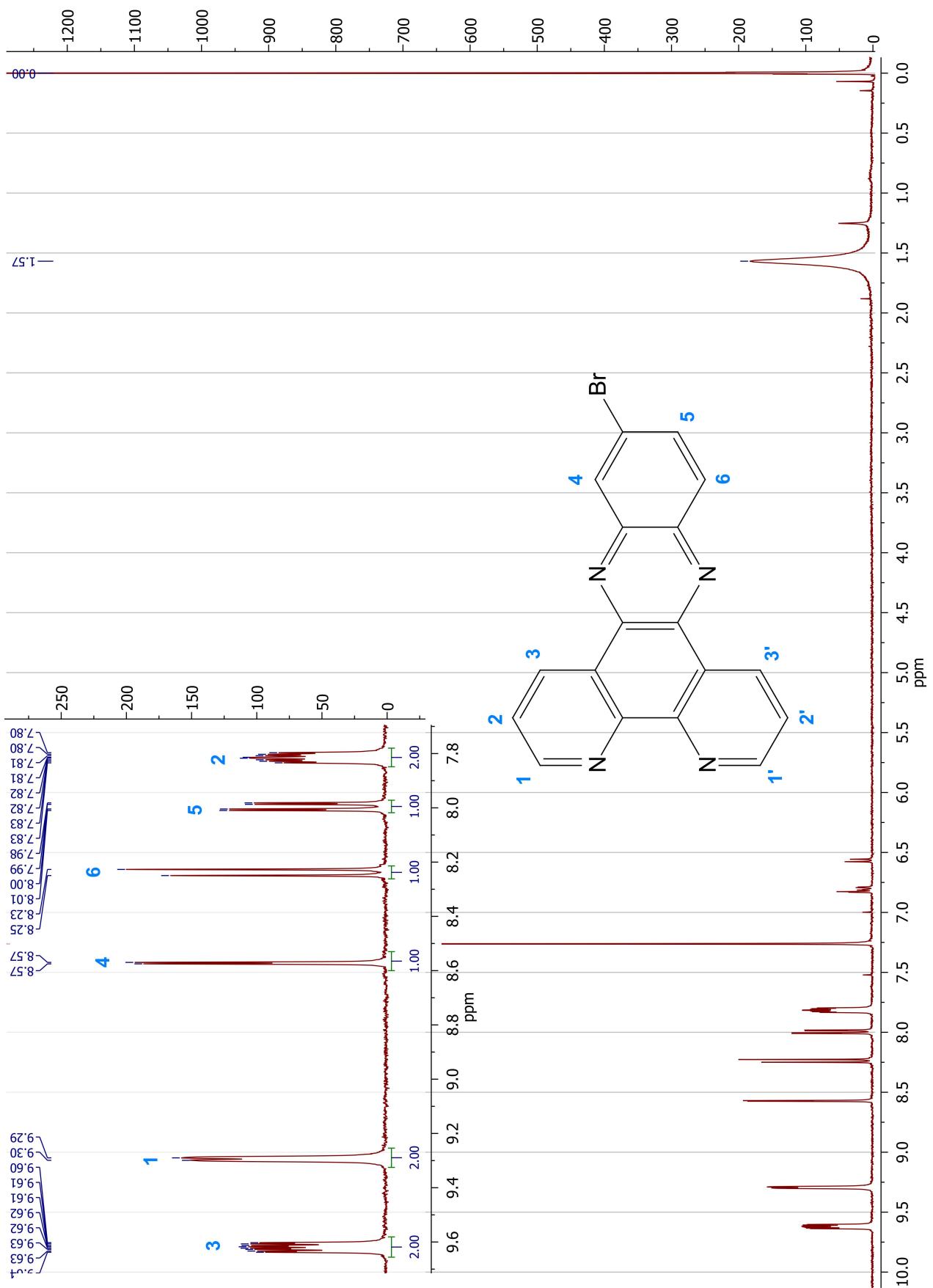
Figure A7.12 – ¹H NMR of 11-bromo-dipyridophenazine (11-Br-dppz).
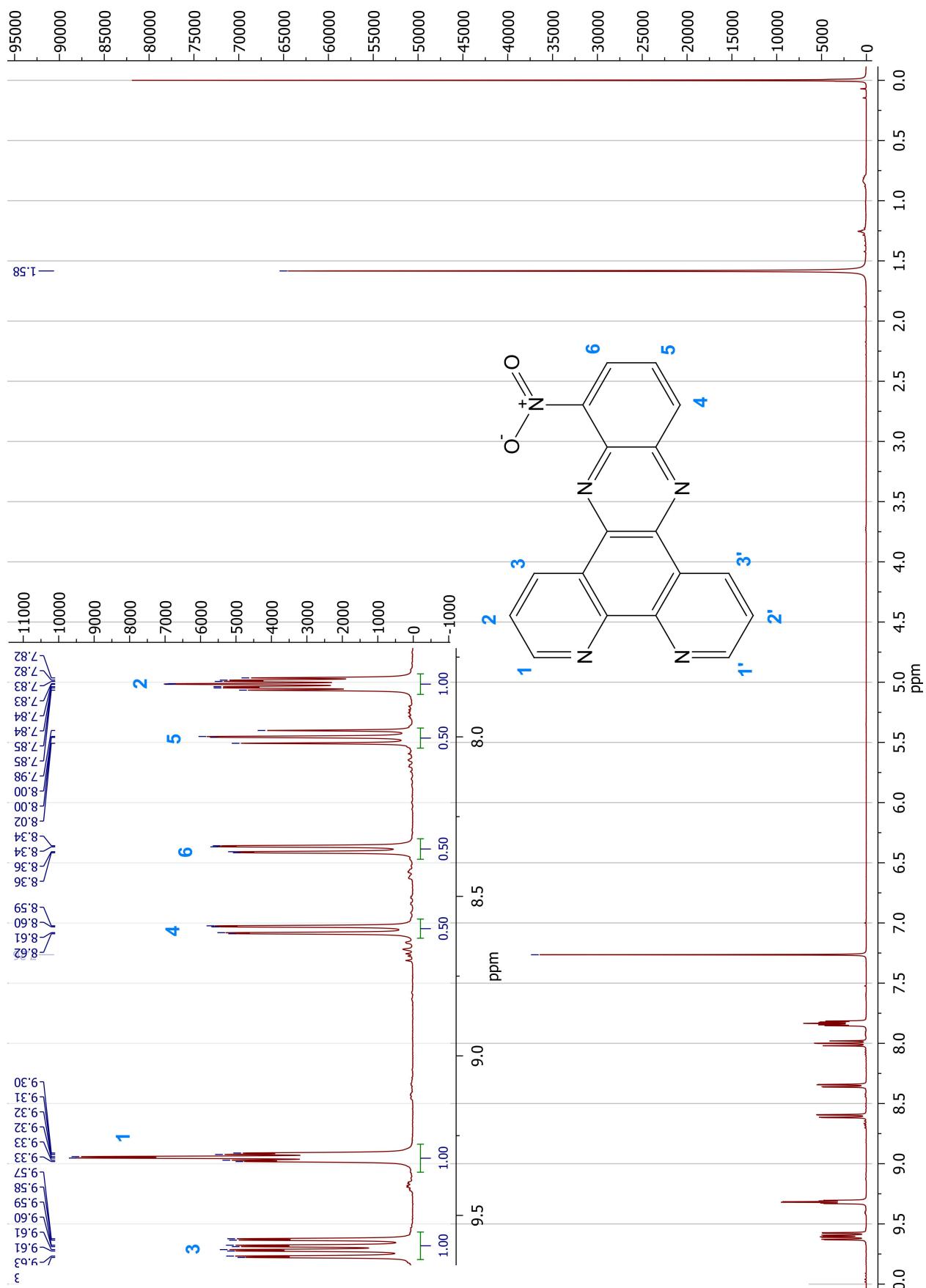
Figure A7.13 – ¹H NMR of 10-nitro-dipyridophenazine (10-NO₂-dppz).
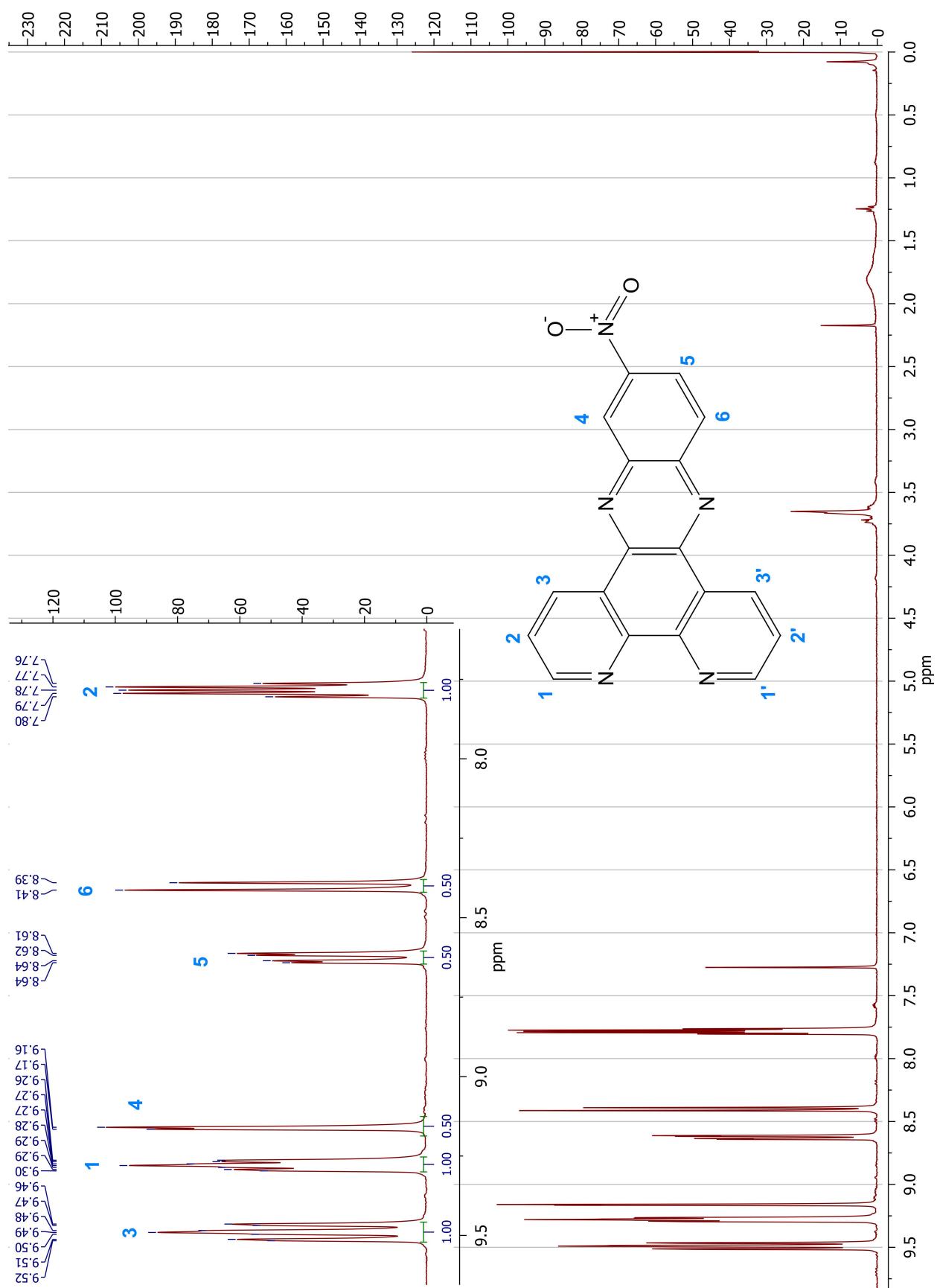
Figure A7.14 – ^{1}H NMR of 11-nitro-dipyridophenazine (11-NO$_2$-dppz).
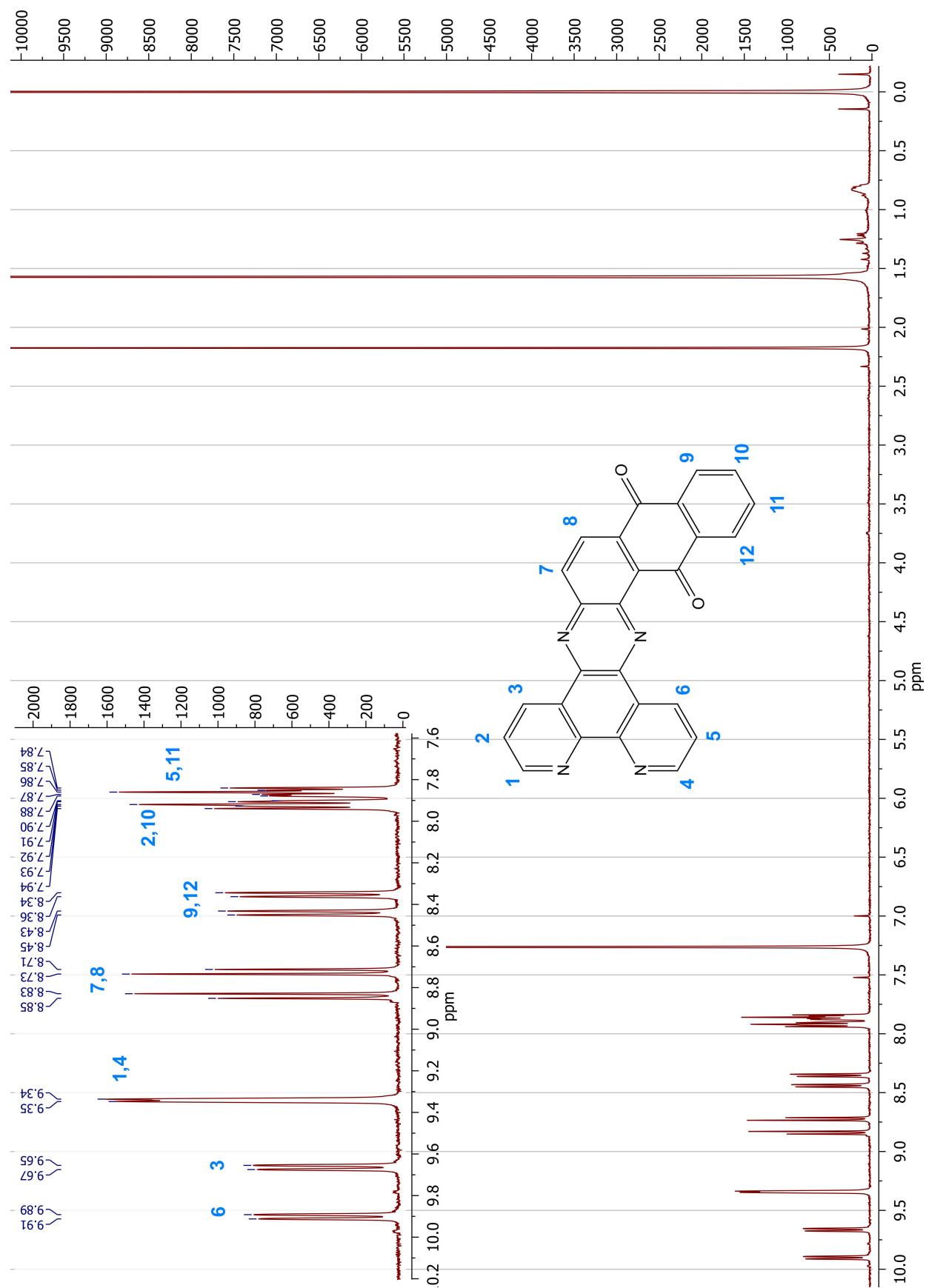
Figure A7.15 – ^{1}H NMR of 12,17-dihydronaphtho[2,3-h]dipyrido[3,2-a:2',3'-c]phenazine-12,17-dione (Aqphen)
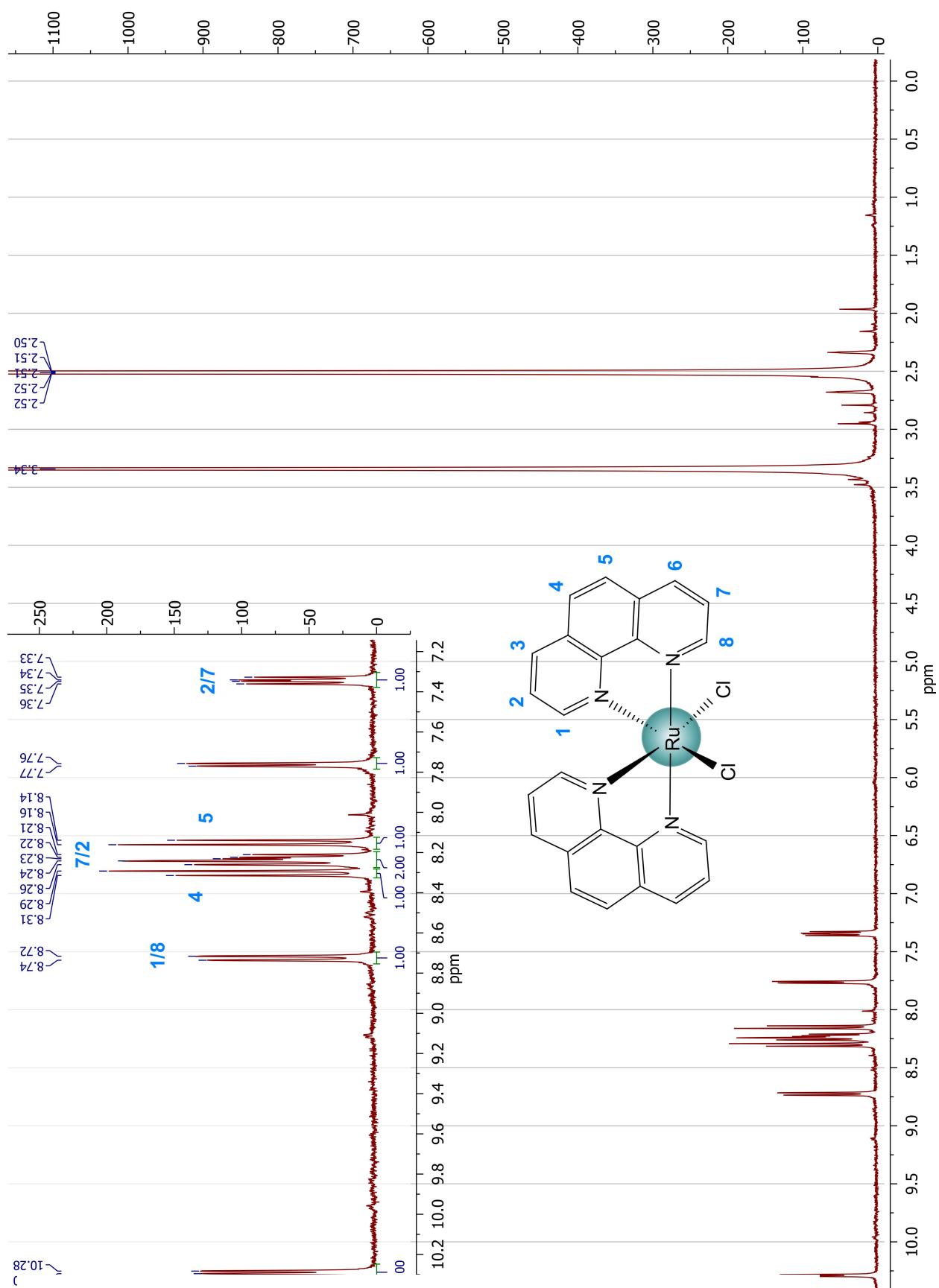
Figure A7.16 – *¹H NMR of cis-Ru(phen)₂Cl₂.*
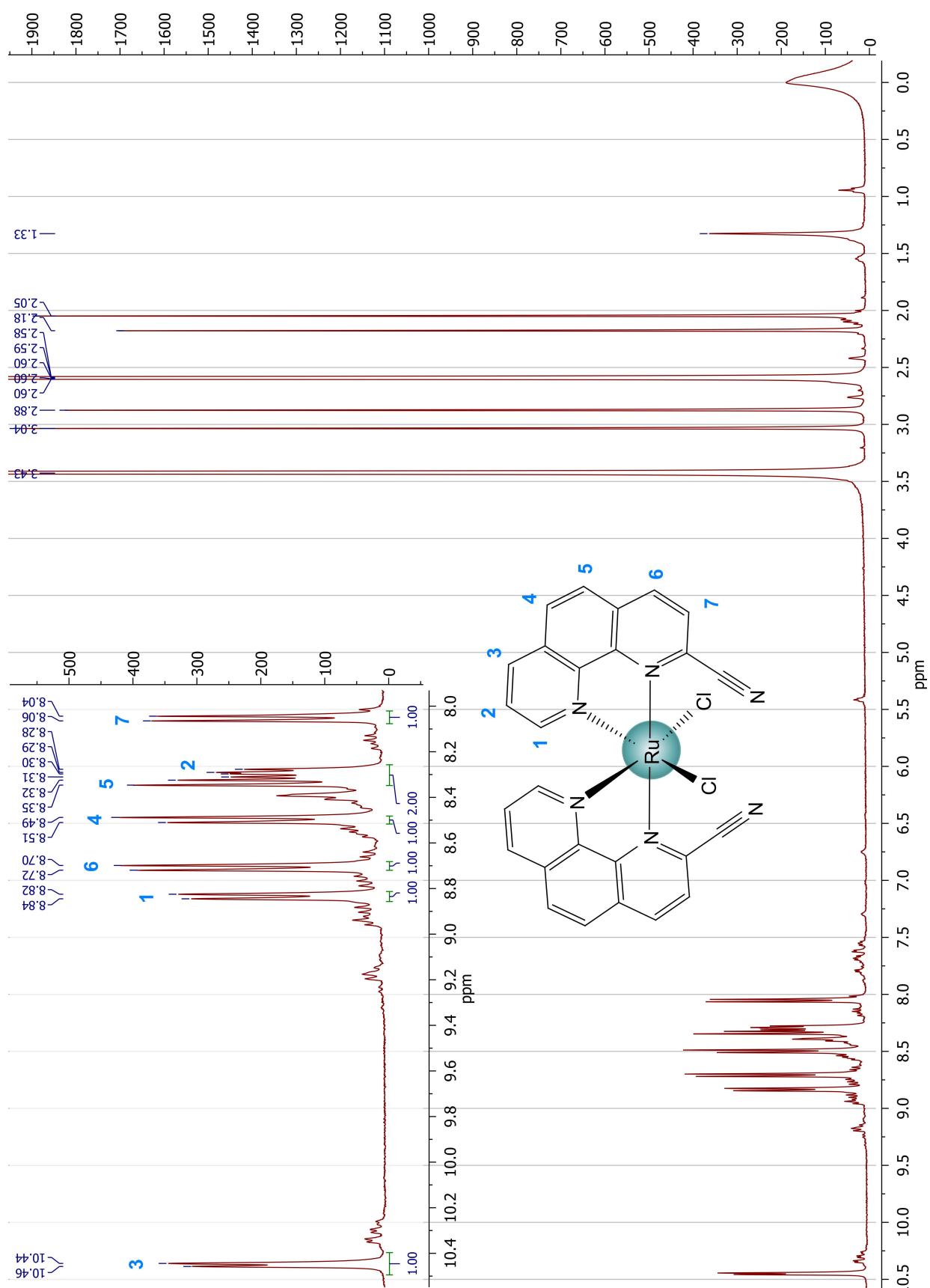
Figure A7.17 – ¹H NMR of cis-Ru(2-CN-phen)₂Cl₂.
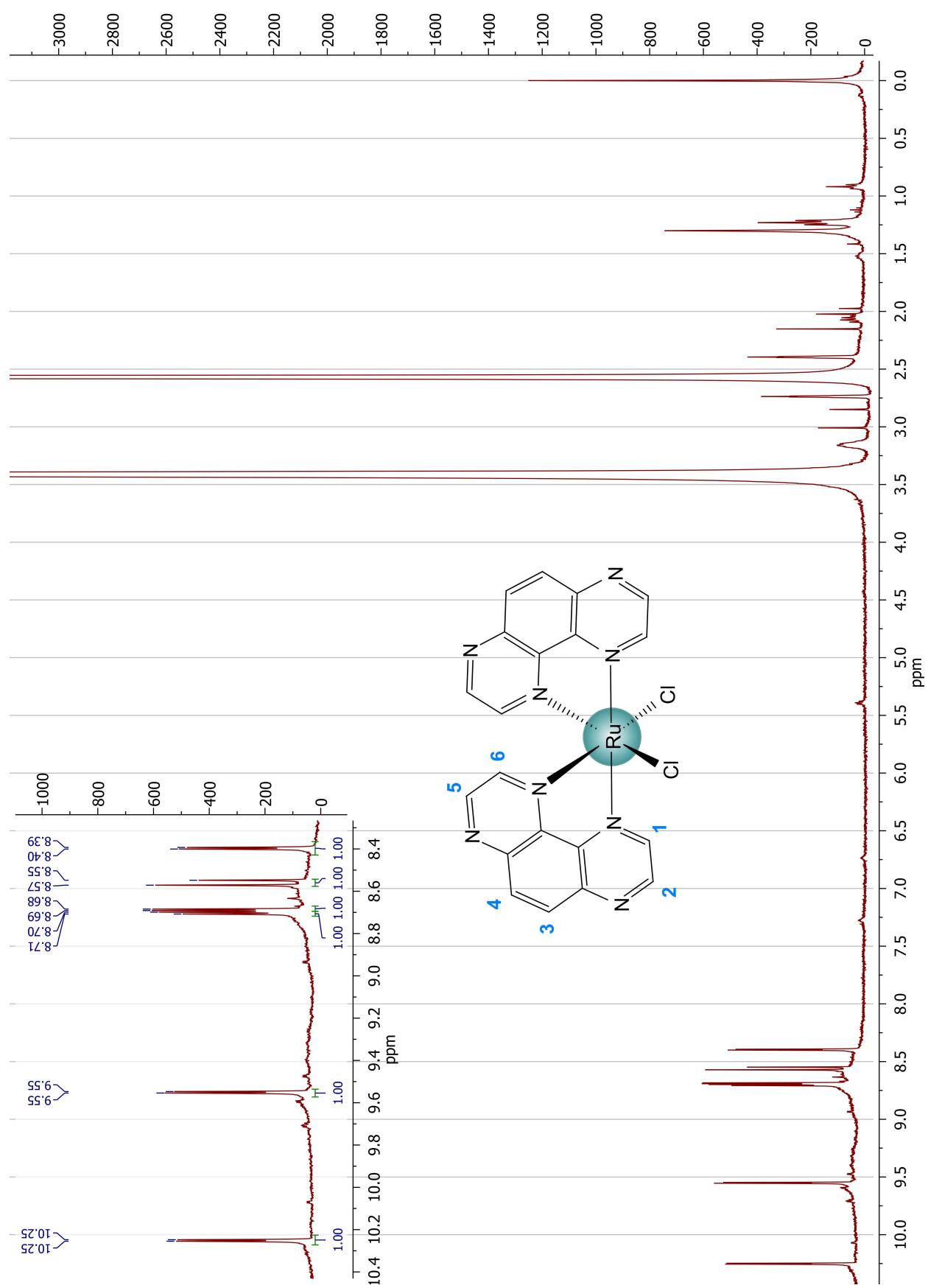
Figure A7.18 – *^{1}H NMR of cis-Ru(TAP)$_2$Cl$_2$.*
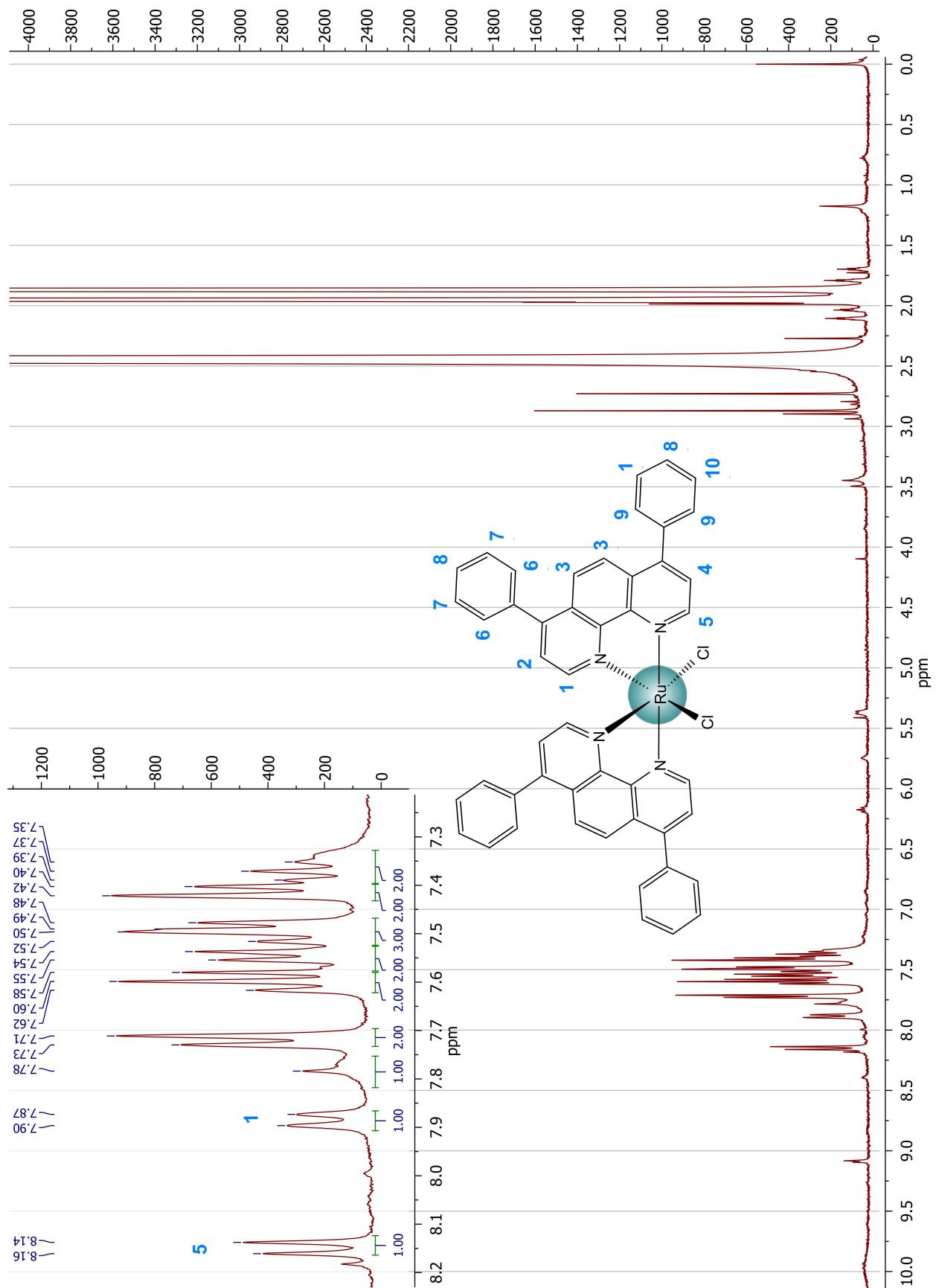
Figure A7.19 – ¹H NMR of cis-Ru(BPhen)₂Cl₂.
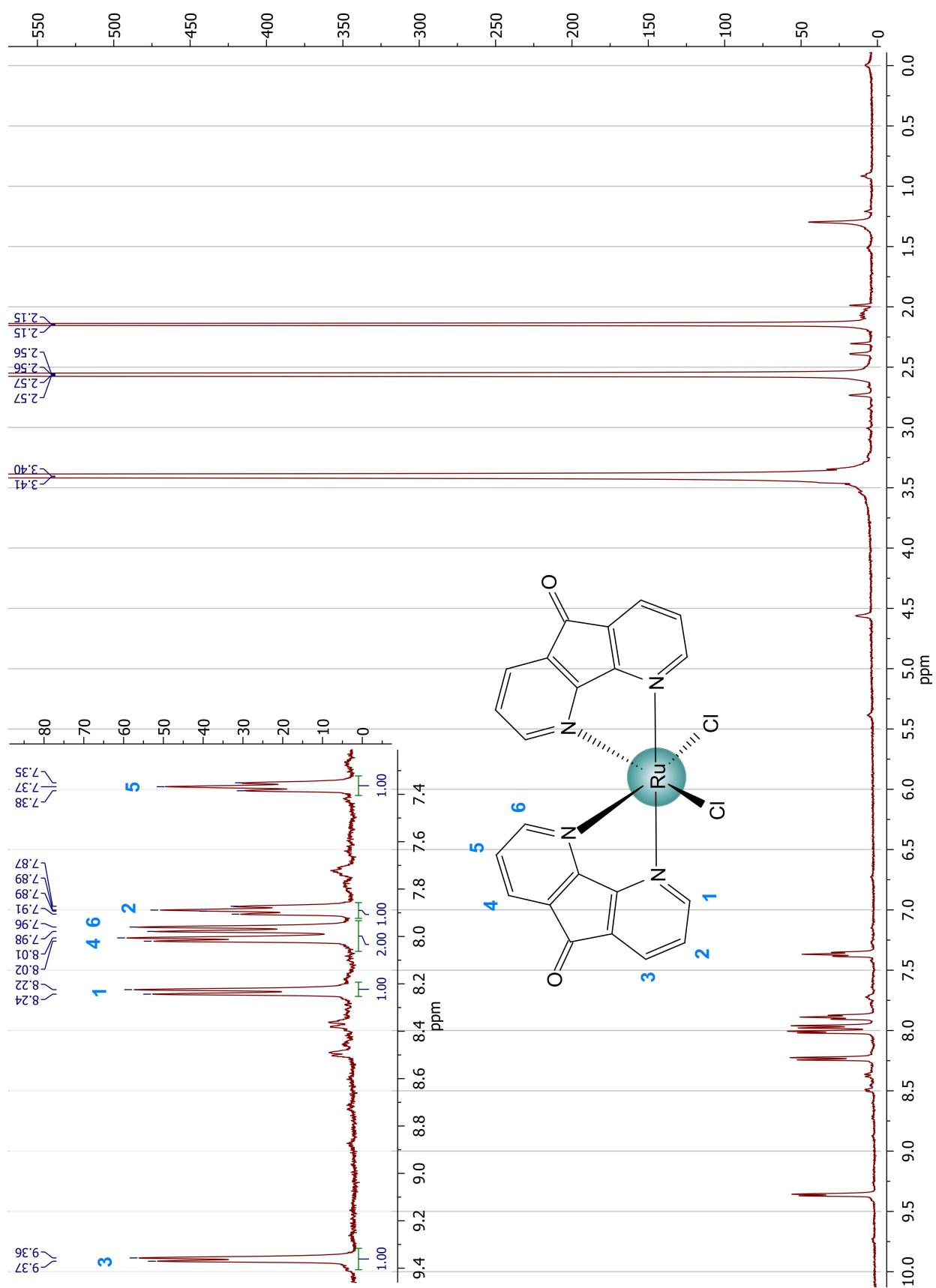
Figure A7.20 – ^1H NMR of cis-Ru(dafo)$_2$Cl$_2$.
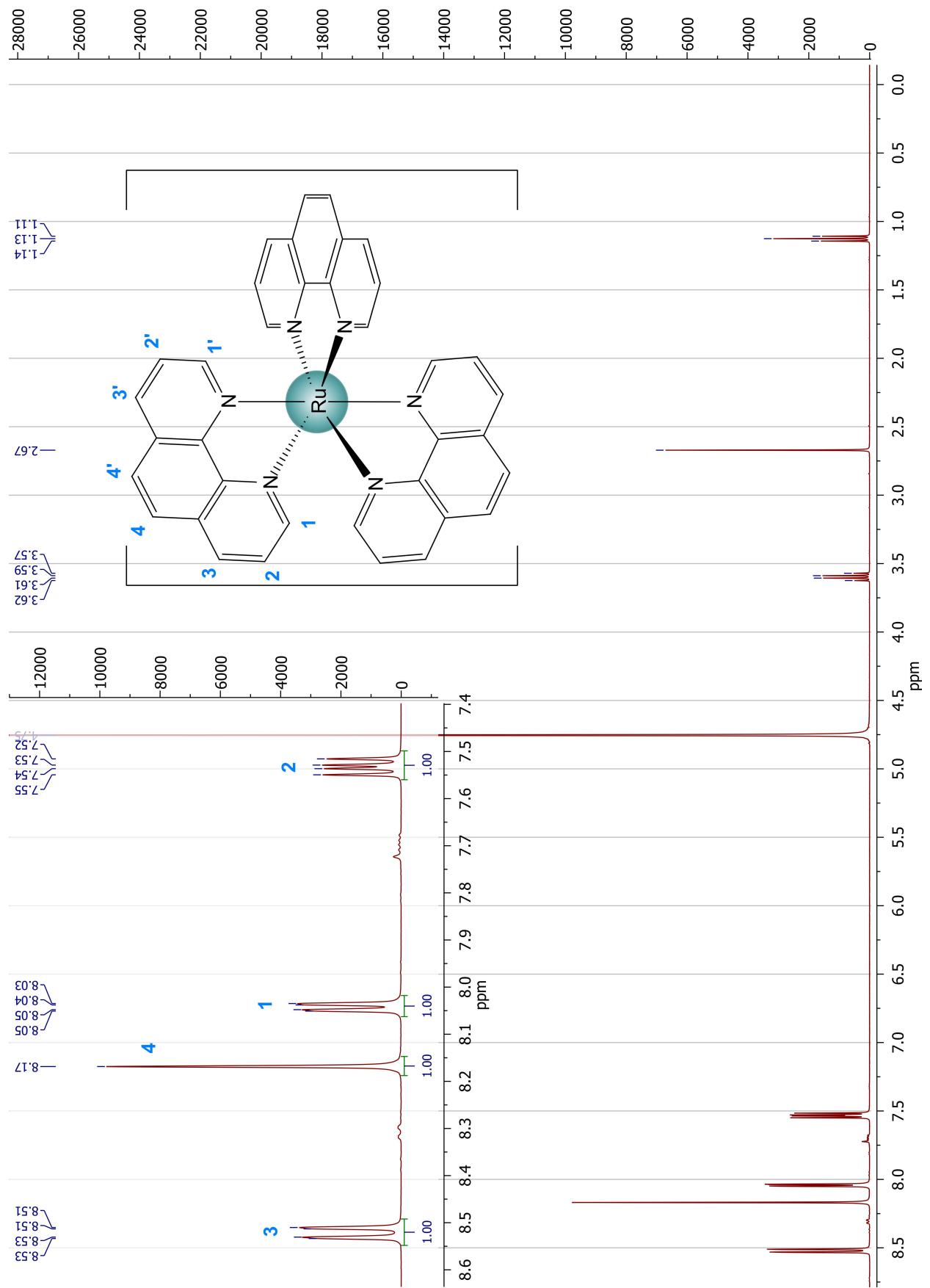
Figure A7.21 – *^{1}H NMR of rac-[Ru(phen)$_3$]·Cl$_2$.*
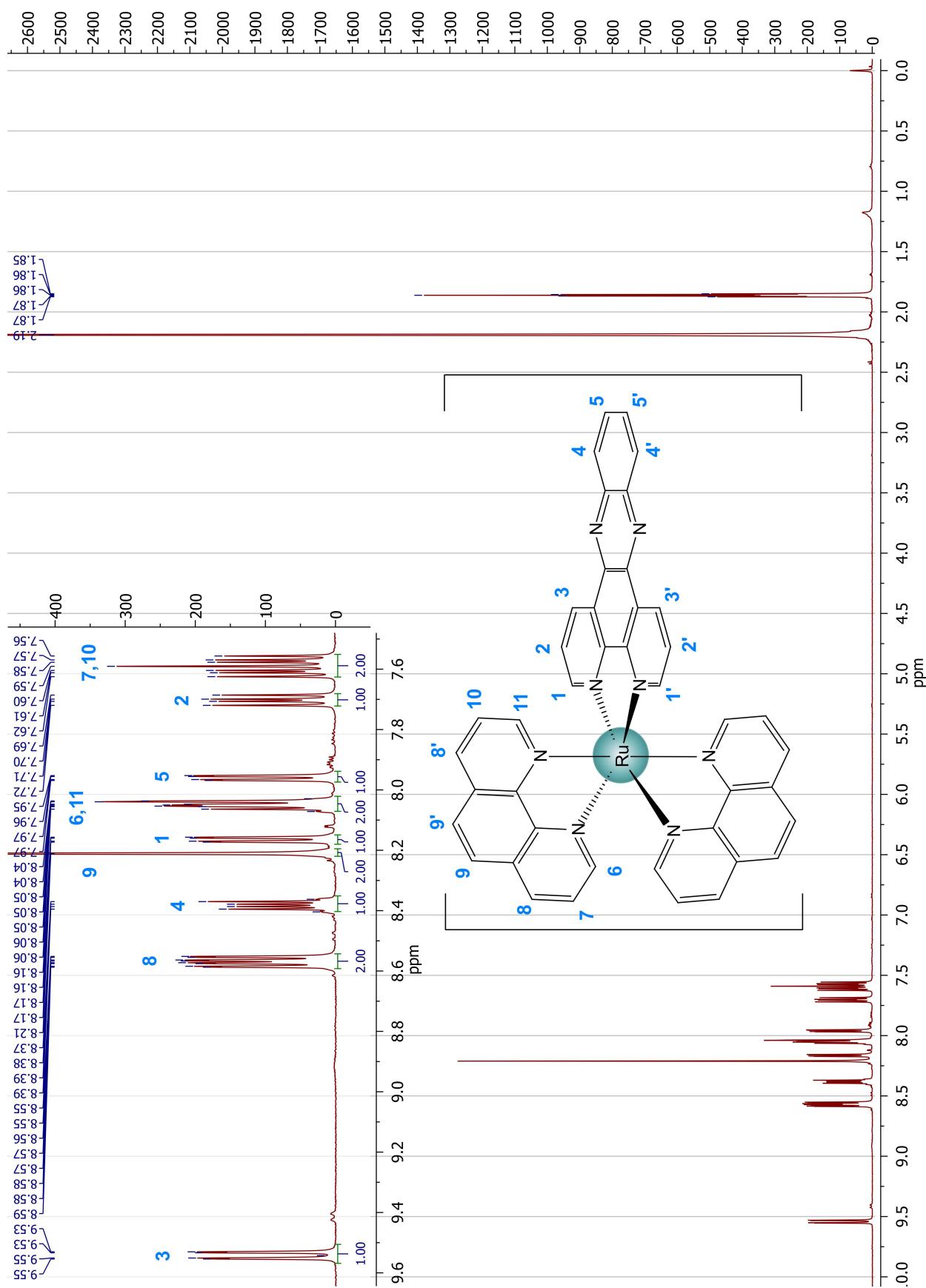
Figure A7.22 – *1H NMR of rac-[Ru(phen)₂(dppz)]·Cl₂.*
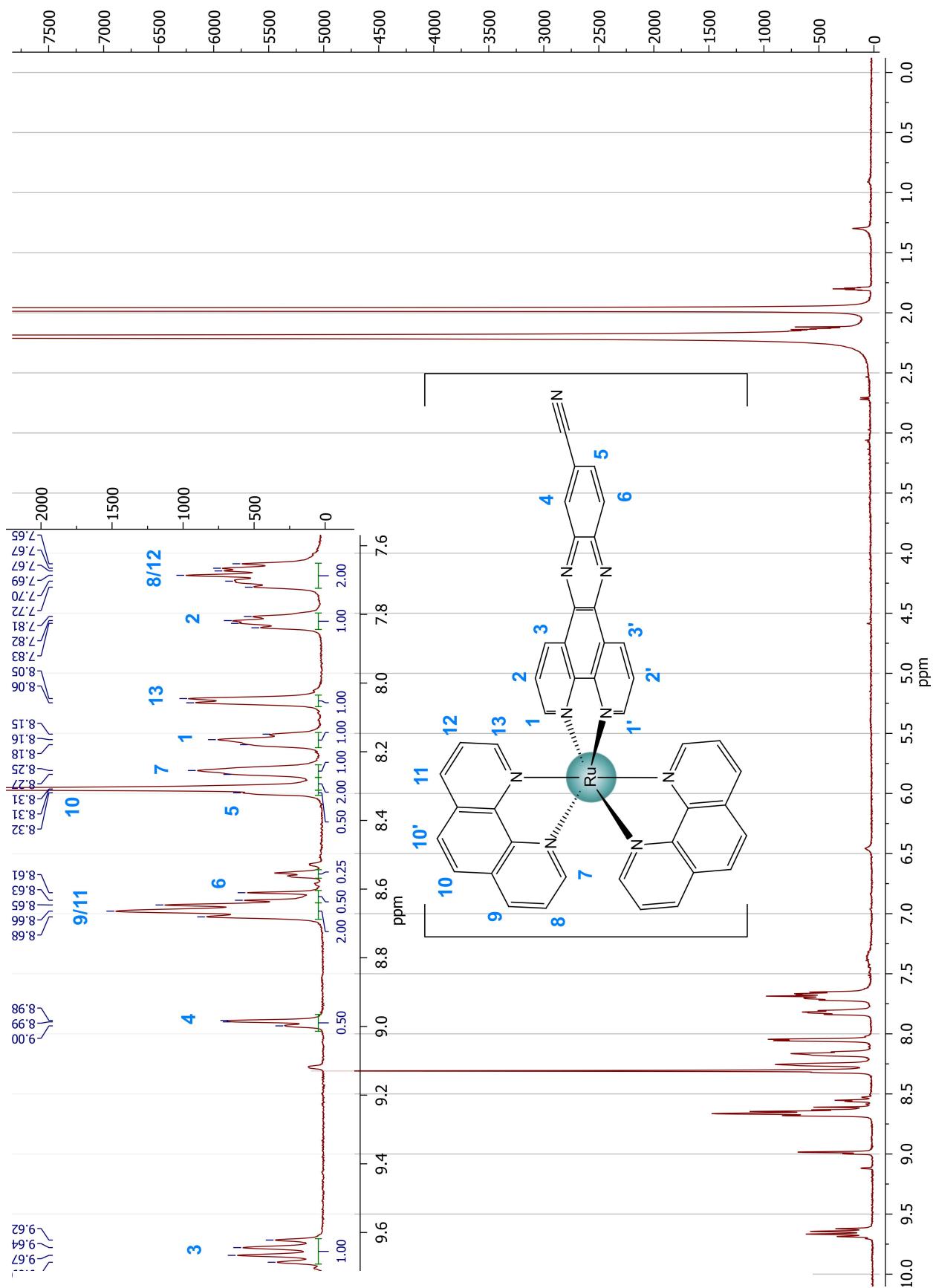
Figure A7.23 — ¹H NMR of rac-[Ru(phen)₂(11-CN-dppz)]·Cl₂.
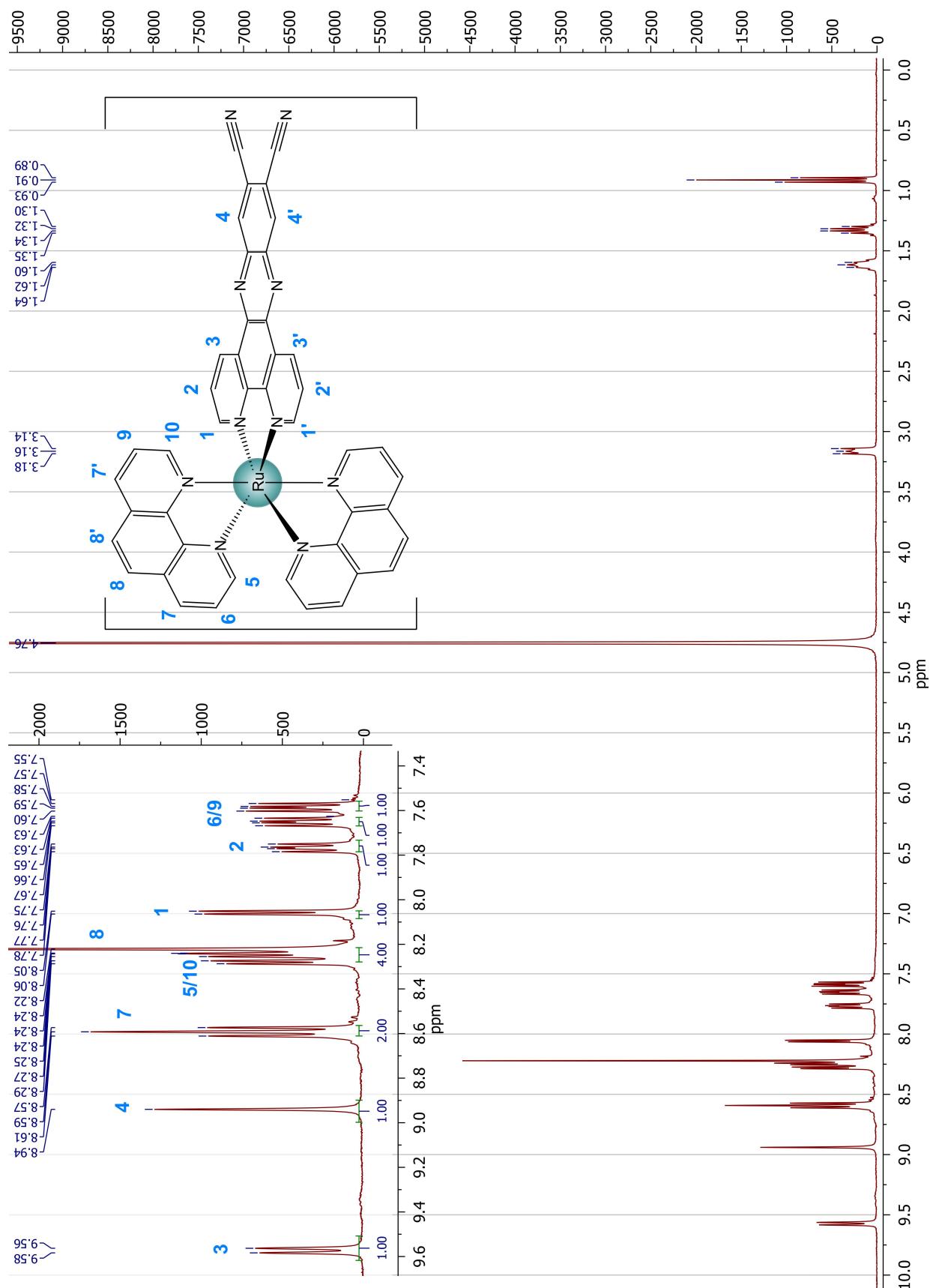
Figure A7.24 – ^1H NMR of rac-[Ru(phen)$_2$(11,12-CN-dppz)]·Cl$_2$.

Figure A7.25 − *¹H NMR of rac-[Ru(phen)₂(10-NO₂-dppz)]·Cl₂*
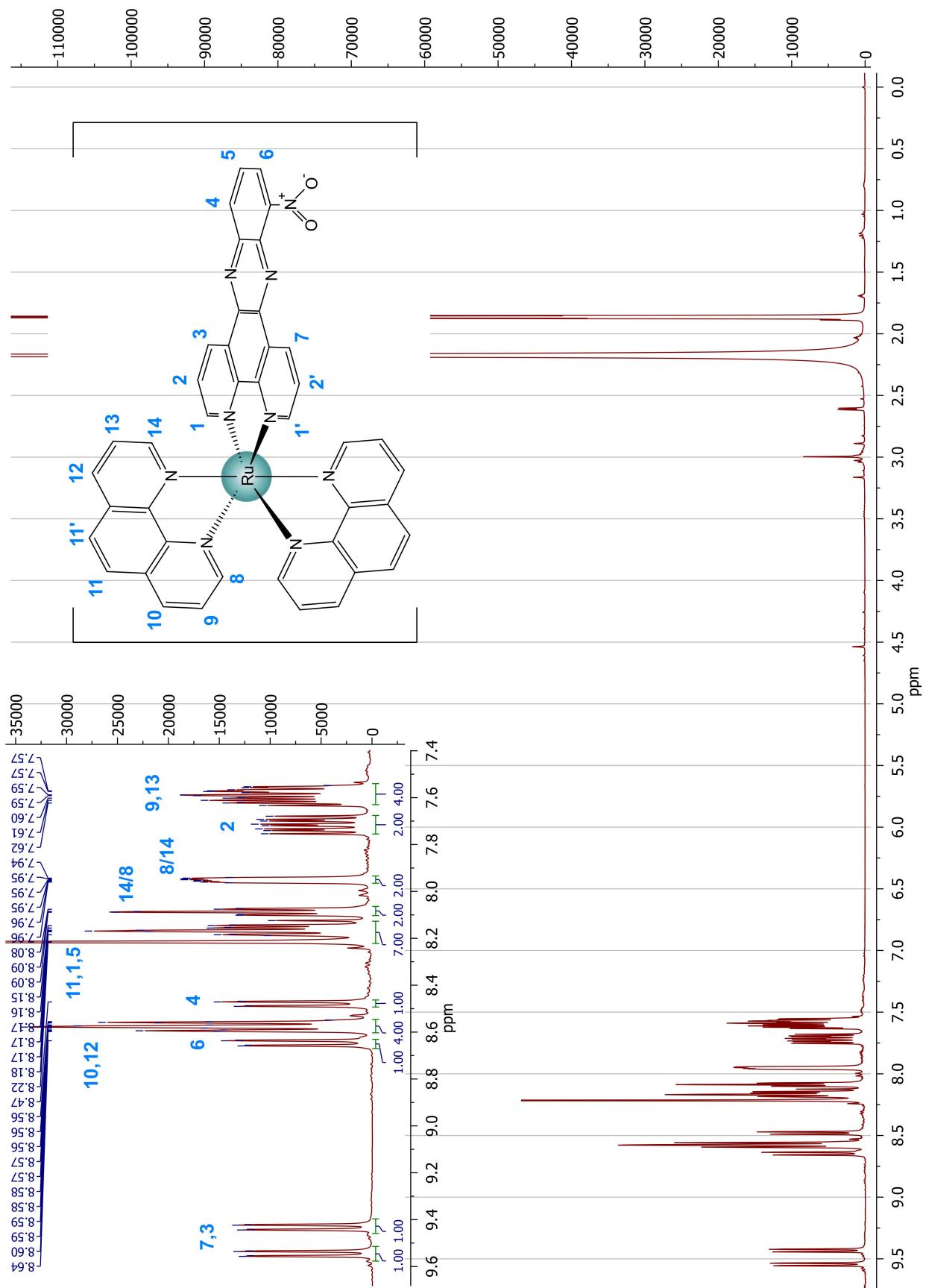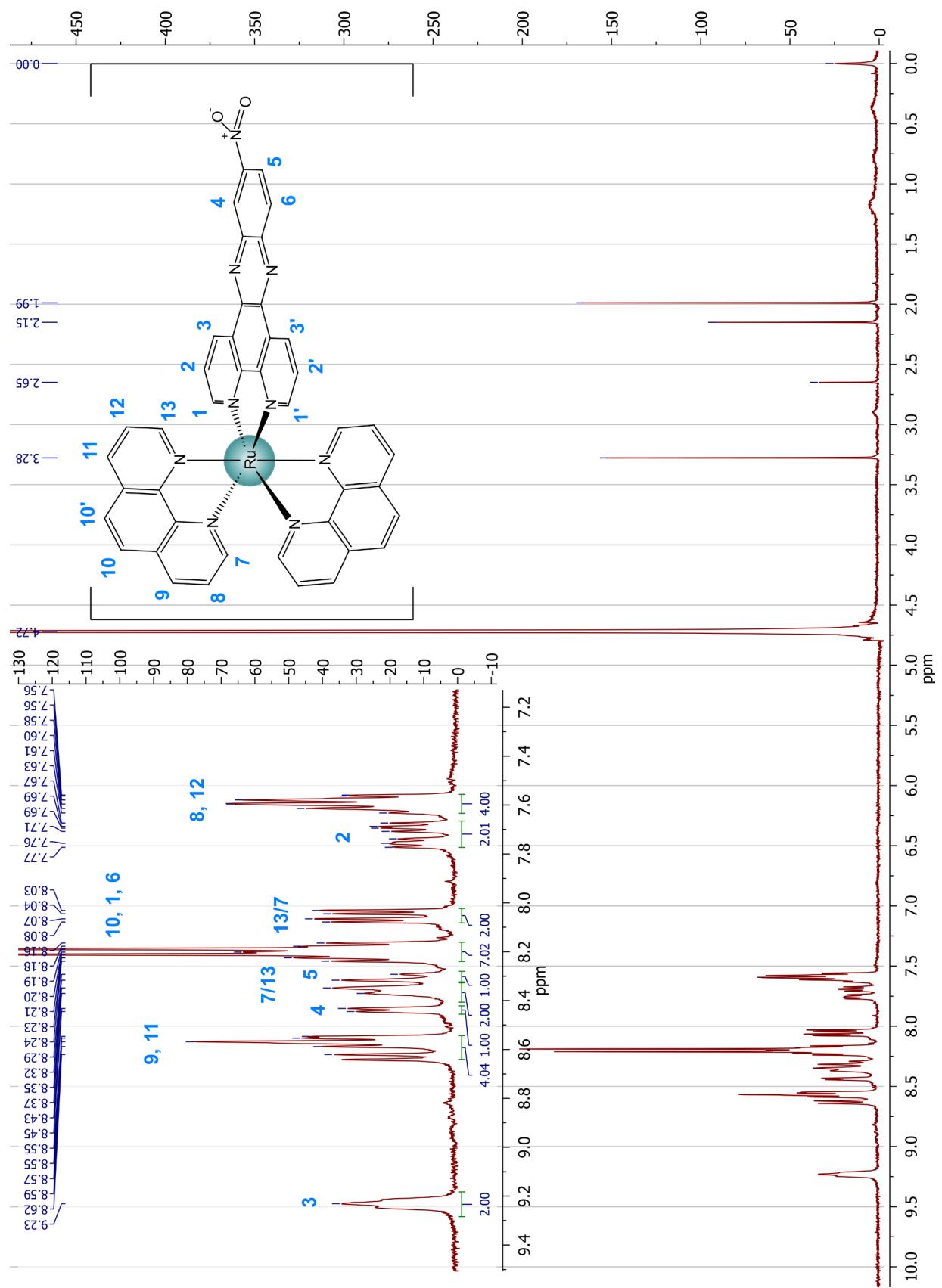
Figure A7.26 — ¹H NMR of rac-[Ru(phen)₂(11-NO₂-dppz)]·Cl₂.
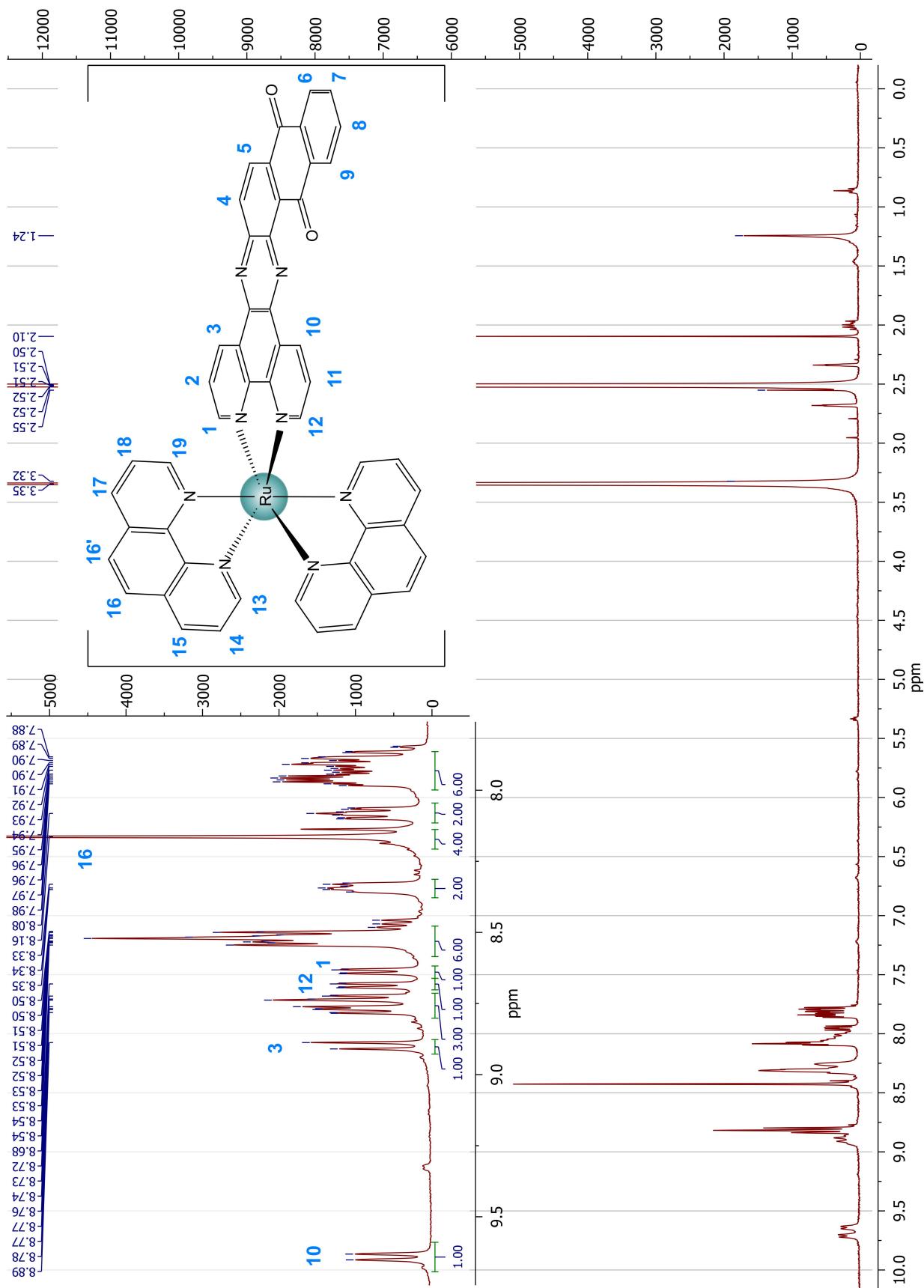
Figure A7.27 – ¹H NMR of rac-[Ru(phen)₂(Aqphen)]·Cl₂.
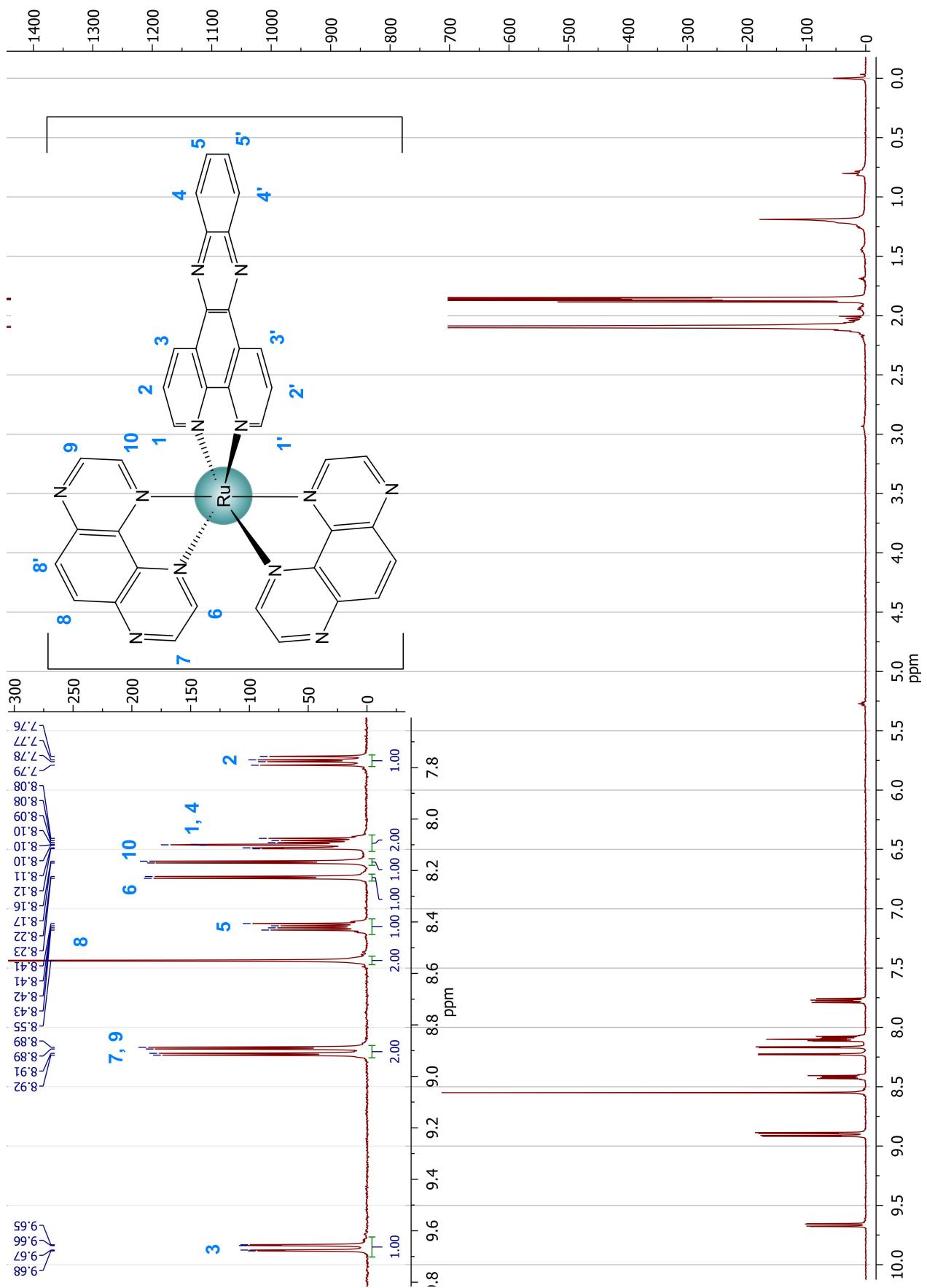
Figure A7.28 – ¹H NMR of rac-[Ru(TAP)₂(dppz)]·Cl₂.
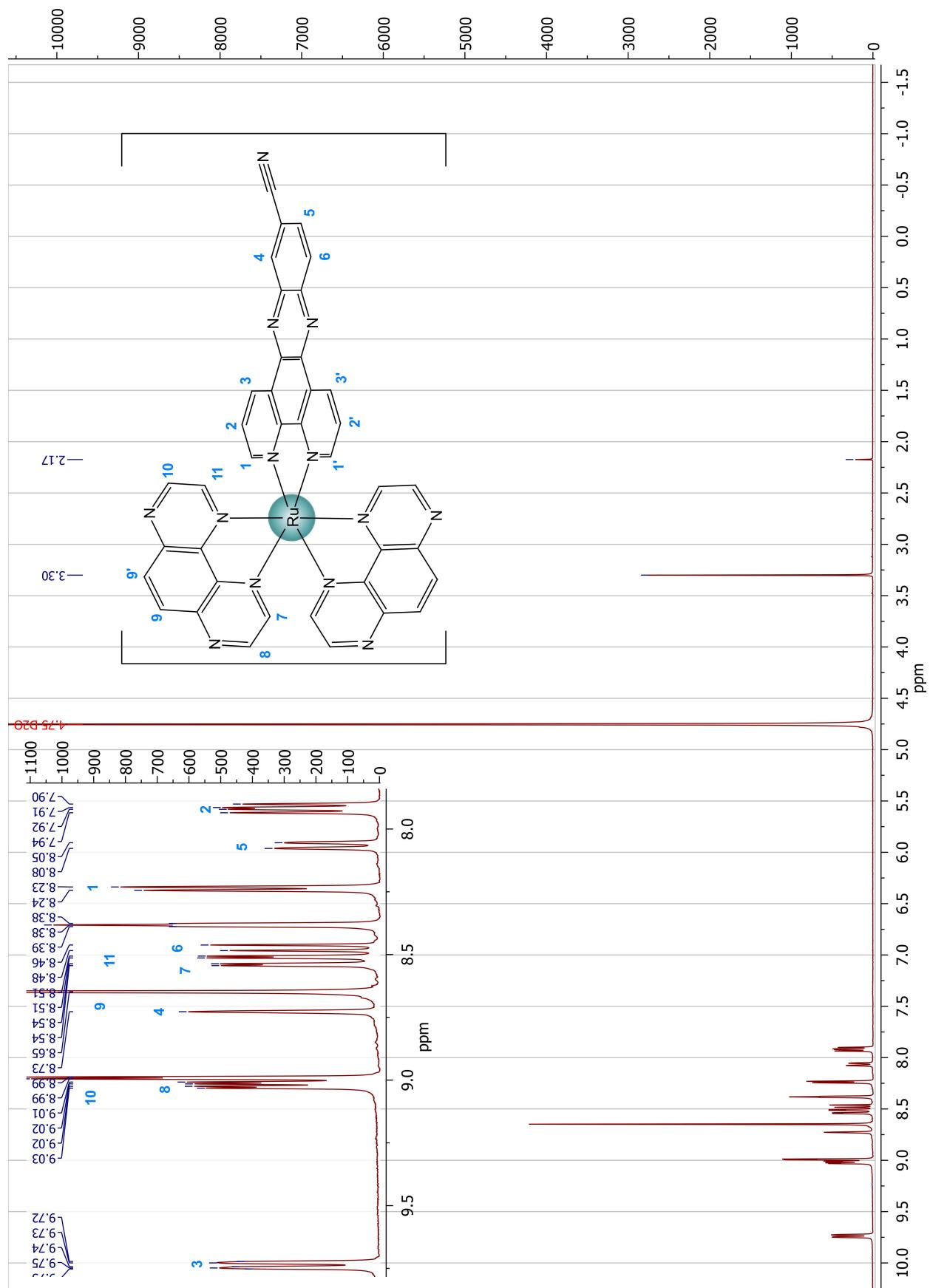
Figure A7.29 – ¹H NMR of rac-[Ru(TAP)₂(11-CN-dppz)]·Cl₂.
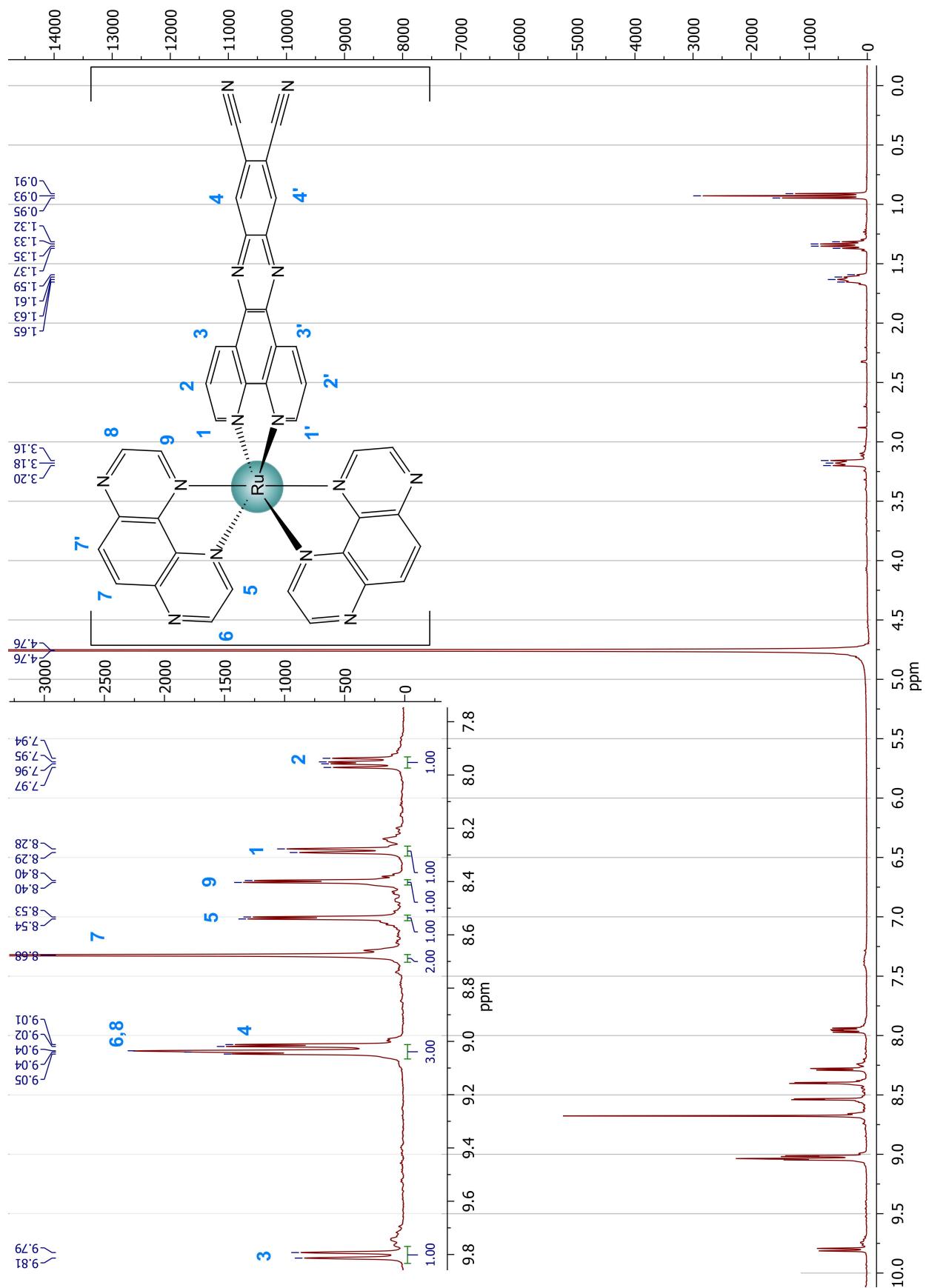
Figure A7.30 – ¹H NMR of rac-[Ru(TAP)₂(11,12-CN-dppz)]·Cl₂.
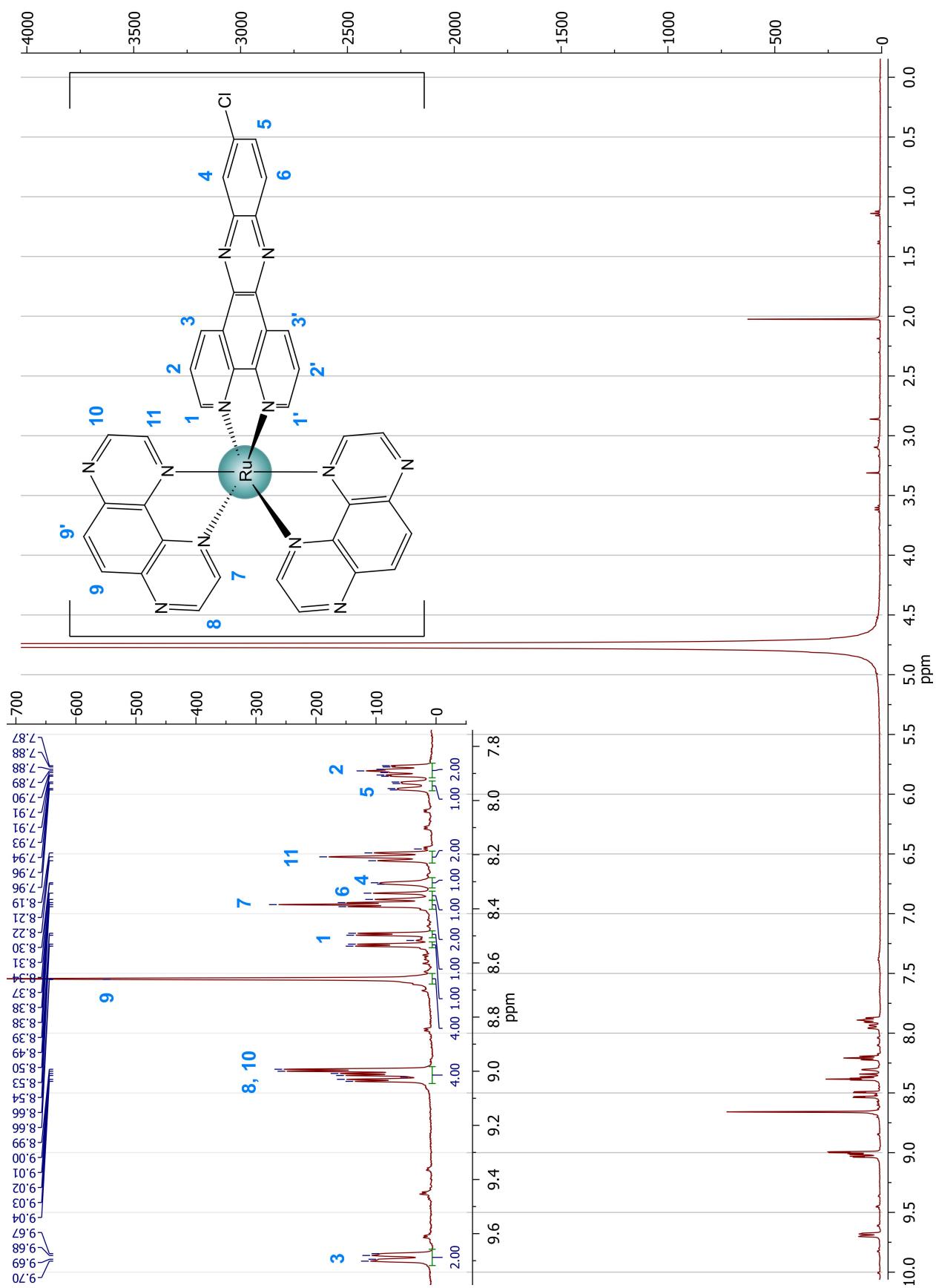
Figure A7.31 — ^{1}H NMR of rac-[Ru(TAP)$_2$(11-Cl-dppz)]·Cl$_2$.
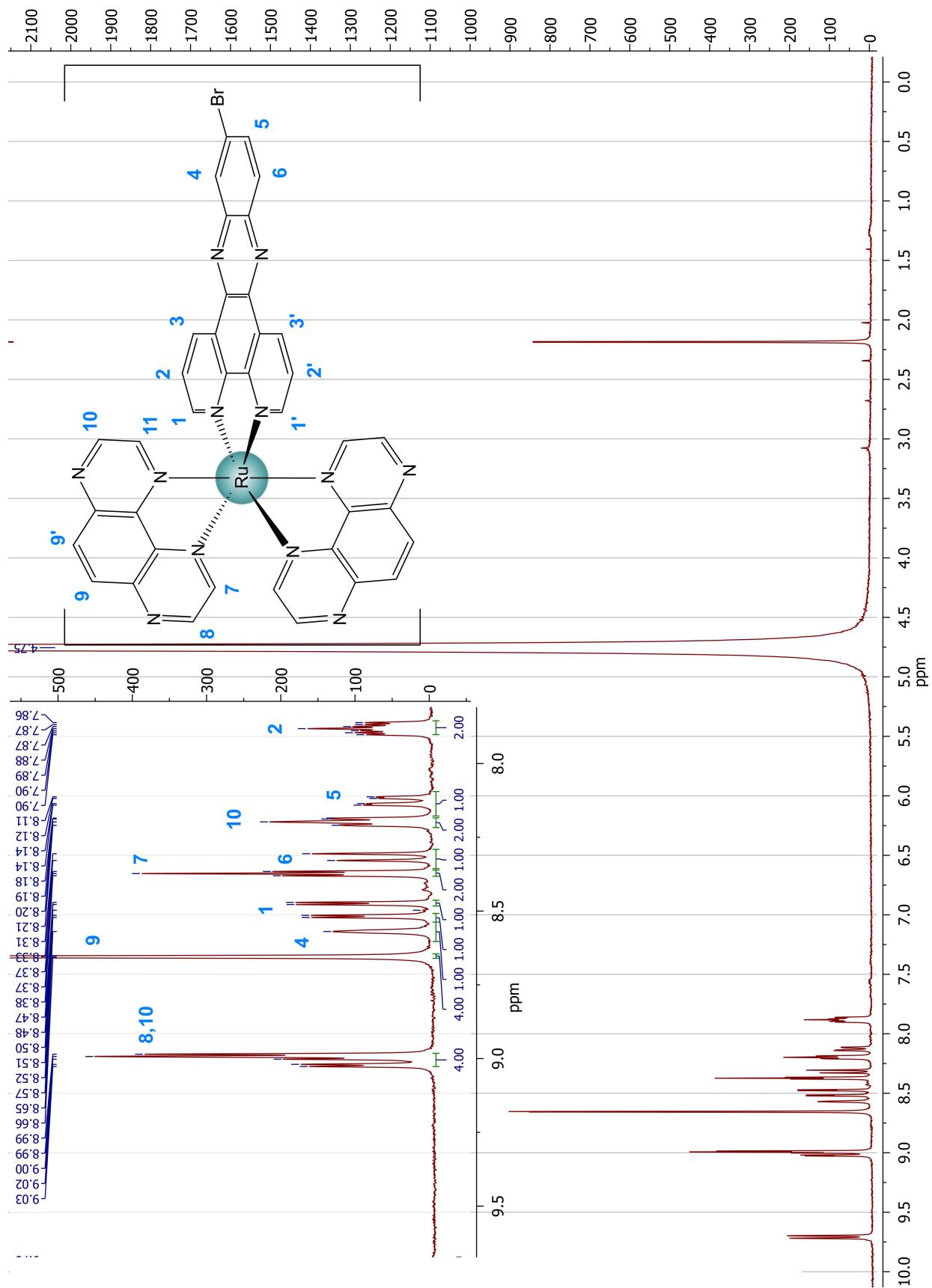
Figure A7.32 − *¹H NMR of rac-[Ru(TAP)₂(11-Br-dppz)]·Cl₂.*
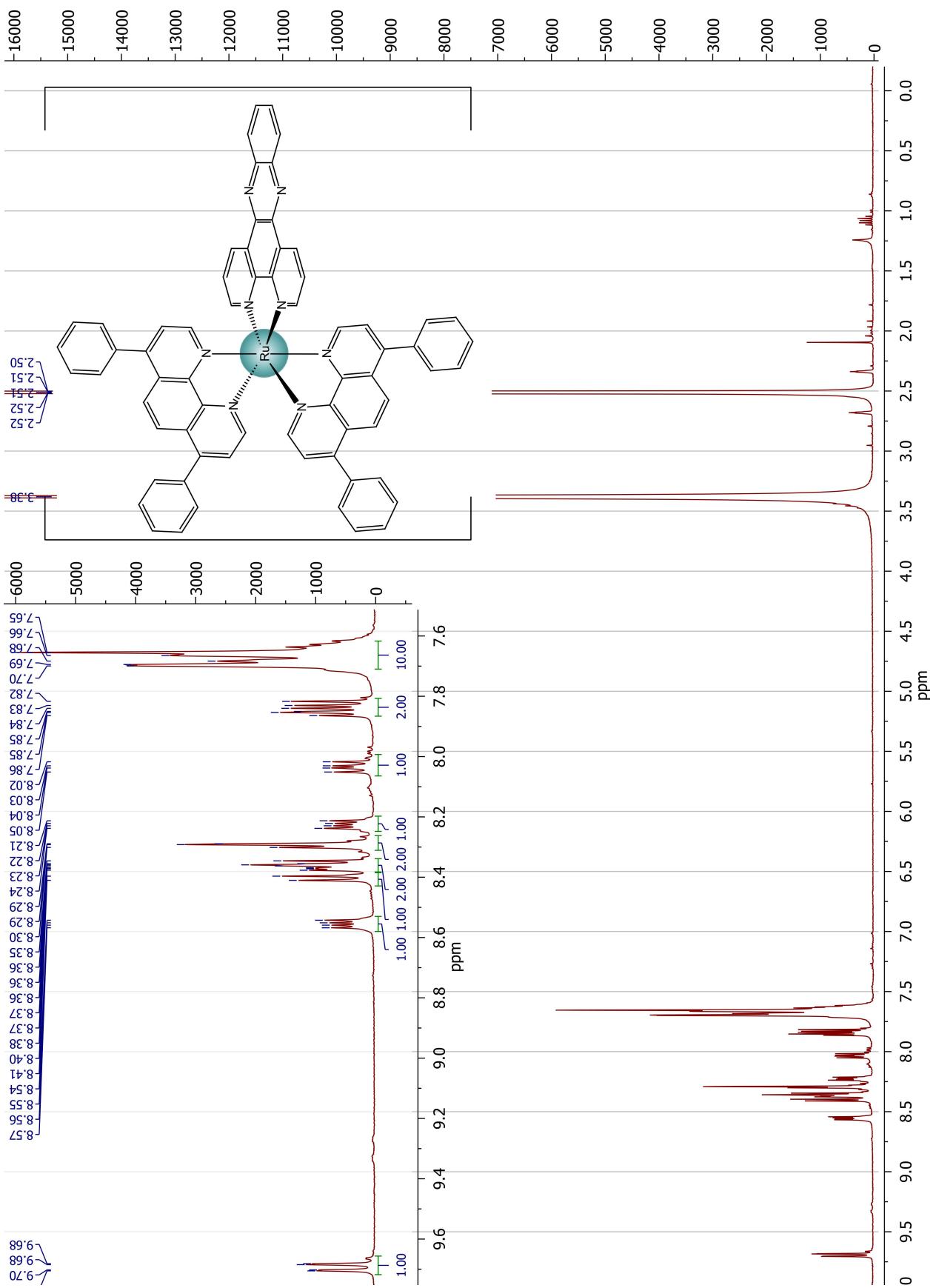
Figure A7.33 – *^{1}H NMR of rac-[Ru(BPhen)$_2$(dppz)]·Cl$_2$.*

Publications from the presented work

- McQuaid, K., Hall, J.P., Brazier, J.A., Cardin, D.J., and Cardin, C.J. X-ray crystal structures show DNA stacking advantage of terminal nitrile substitution in Ru-dppz complexes. *Chem. Eur. J.* **24**, 15859-15867, (2018).

- McQuaid, K., Abell., Gurung, S., Allan, D.R., Winter, G., Sorensen, T., Cardin, D.J., Brazier, J.A., Cardin, C.J., and Hall, J.P. Structural studies reveal enantiomeric recognition of a DNA G-quadruplex by a ruthenium polypyridyl complex. *Angew. Chem. Int. Ed.* **58**, 9881-9885, (2019).

- McQuaid, K., Hall, J.P., Baumgaertner, L., Cardin, D.J., and Cardin, C.J. Three thymine-adenine binding modes of the ruthenium complex Λ-[Ru(TAP)$_2$(dppz)]$^{2+}$ to the G-quadruplex d(TAGGGTT) shown by X-ray crystallography. *Chem. Commun*. **55**, 9116-9119, (2019).

FULL PAPER

X-ray crystal structures show DNA stacking advantage of terminal nitrile substitution in Ru-dppz complexes

Kane McQuaid[a,c], James P. Hall[a,b,c], John A. Brazier[b], David J. Cardin[a] and Christine J. Cardin[a]*

Abstract: The new complexes [Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$, [Ru(TAP)$_2$(11-Br-dppz)]$^{2+}$and [Ru(TAP)$_2$(11,12-diCN-dppz)]$^{2+}$ are reported. The addition of nitrile substituents to the dppz ligand of the DNA photooxidising complex [Ru(TAP)$_2$(dppz)]$^{2+}$ promote π-stacking interactions and ordered binding to DNA, as shown by X-ray crystallography.

The structure of Λ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ with the DNA duplex d(TCGGCGCCGA)$_2$ shows, for the first time with this class of complex, a closed intercalation cavity with an AT base pair at the terminus. The structure obtained is compared to that formed with the 11-Br and 11,12-dinitrile derivatives, highlighting the stabilization of *syn* guanine by this enantiomer when the terminal basepair is GC. In contrast the AT basepair has the normal Watson-Crick orientation, highlighting the difference in charge distribution between the two purine bases and the complementarity of the dppz-purine interaction. The asymmetry of the cavity highlights the importance of the purine-dppz-purine stacking interaction.

Introduction

The complex [Ru(phen)$_2$(dppz)]$^{2+}$ and its derivatives have been extensively studied since the original demonstration of the DNA 'light-switch' effect using [Ru(bpy)$_2$(dppz)]$^{2+}$ (phen = phenanthroline, dppz = dipyridophenazine and bpy = bipyridine).[1] These complexes, and close derivatives, have been shown to be able to act as structure specific luminescent probes for mismatched DNA,[2,3] and for G-quadruplexes.[4,5] Related complexes can oxidise guanine upon irradiation,[6–9] explicitly directing DNA damage.[10] Such damage pathways are utilized in the study of anti-cancer photodynamic therapies (PDT) [11–16] for which ruthenium polypyridyl complexes are considered promising candidates for the next generation of photosensitizers. [13,16,17] The extension of the dppz ligand with additional rings for example, leads to powerful anti-tumour properties.[18–20] Gaining insight into the manner in which these

complexes bind to their target is therefore of importance, where a deeper structural knowledge allows for the superior design of more specific DNA binders. In one recent report, for example, a rapid screening approach was used to identify structural selectivity using racemic mixtures of a range of halide derivatives of [Ru(bpy)$_2$dppz]$^{2+}$, and including the 11-Br analogue of the complex, studied in this work.[21]

[Ru(TAP)$_2$(bpy)]$^{2+}$ (TAP = 1,4,5,8-tetraazaphenanthrene) has been shown to covalently link to guanine upon irradiation[6] and the TAP ligand was subsequently used to form the photooxidising and DNA intercalating [Ru(TAP)$_2$(dppz)]$^{2+}$ complex.[22] That complex is the parent compound of the three derivatives reported here (Figure 1a). A detailed review of the photooxidation kinetics of [Ru(TAP)$_2$(dppz)]$^{2+}$ with sequence specific guanine oligonucleotides has recently been published.[23] The biophysical and solution behaviour of these complexes was intensively investigated in the absence of a structural model, until our publication in 2011[24] highlighting the binding of Λ-[Ru(TAP)$_2$(dppz)]$^{2+}$ to the DNA decamer sequence d(TCGGCGCCGA)$_2$, and, the following year, the binding of Λ-[Ru(phen)$_2$(dppz)]$^{2+}$ to the sequence d(CCGGTACCGG)$_2$.[25] The

Figure 1 – Illustrations of (a) structures of the reported complexes Λ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ (**I**), Λ-[Ru(TAP)$_2$(11-Br-dppz)]$^{2+}$ (**II**) and Λ-[Ru(TAP)$_2$(11,12-diCN-dppz)]$^{2+}$ (**III**); (b) Stick plot of (**I**) showing the numbering scheme used throughout the text. Carbon atoms are coloured cyan, nitrogen – dark blue, ruthenium – teal and hydrogen – white; (c) oligonucleotides used in the study highlighting the standard nucleobase colouring used throughout.

complexes were shown to intercalate from the minor groove of the B-DNA, as confirmed by several further studies.[26] The primary stabilising interaction was shown to be the stacking between the DNA bases and the dppz ligand. The angle of the (canted) intercalation was subsequently shown to be determined by a secondary stacking interaction between the phen or TAP ancillary ligand and the 2'-deoxyribose sugar of one of the four bases forming the intercalation cavity, thus altering the

[a] K. McQuaid, Dr J. P. Hall, Prof D.J. Cardin, Prof C.J. Cardin
Department of Chemistry, University of Reading, Whiteknights, Reading, RG6 6AD, UK.
E-mail: c.j.cardin@reading.ac.uk

[b] Dr J. P. Hall and Dr J. A. Brazier
Department of Pharmacy, University of Reading, Whiteknights, Reading, RG6 6AD, UK.

[c] K. McQuaid and Dr J. P. Hall
Diamond Light Source, Harwell Science and Innovation Campus, Didcot, Oxfordshire, OX11 0DE, UK

Supporting information for this article is given via a link at the end of the document.

photophysical properties of the bound complexes.[27] This binding mode is now seen to be the dominant binding mode for both enantiomers of [Ru(L)$_2$(dppz)]$^{2+}$ (where L = bidentate *N*-heterocycle). The only exception recorded so far is the symmetrical binding of the lambda enantiomer to a central TA/TA step, which we now believe can only be seen at this site.[25] In this binding mode there is a high twist angle of nearly 40°, with a secondary interaction with the two symmetrically equivalent phen ligands.

We were able to show that the angled intercalation mode is also seen for [Ru(bpy)$_2$(dppz)]$^{2+}$ and [Ru(phen)$_2$(11,12-Me$_2$-dppz)]$^{2+}$ using a different DNA sequence, and also hypothesised that the angled (canted) intercalation mode seen at the symmetrical CG/CG step is a structural consequence of the projection of the 2-NH$_2$ group of guanine into the minor groove.[28] More recently these structural observations have been used to interpret longstanding discussions in the literature about observations of multiple luminescence lifetimes even in apparently homogeneous Ru-DNA model systems.[29] The realisation that the typical binding mode for these complexes is angled intercalation prompted the investigation of the effect of asymmetric substitution of the distal ring of the dppz ligand[30] (positions 10-13 in Figure 1b).
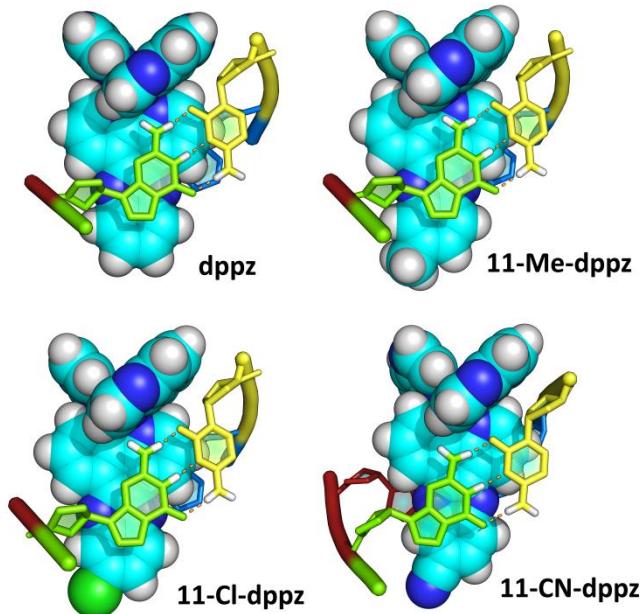

Figure 2 – Asymmetric angled binding by the lambda enantiomers of substituted [Ru(TAP)$_2$(dppz)]$^{2+}$ complexes viewed from the C2-G9 base pair side.[17] Of particular note is the similarity in binding between all complexes besides the 11-CN-dppz derivative.

Here we report the crystal structure of *rac*-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ (**I**), and the structural consequences of binding to the DNA duplex d(TCGGCGCCGA)$_2$, with a comparison to the CG/CG terminal step (Figure 1c). The phen analogue of this compound was recently described for potential photodynamic therapy applications.[14] The duplex is that used in our original publication from 2011,[24] and our intention was to investigate the effect of electron withdrawing substituents with useful infrared reporting groups,[31] and which could be accommodated in the solvent space of the well understood crystal packing.[32] We make two comparisons, using two more new complexes – with the Λ-11-Br analogue (**II**) bound to the same sequence, and with the symmetric disubstituted Λ-[Ru(TAP)$_2$(11,12-diCN-dppz)]$^{2+}$ (**III**) to an asymmetric decamer sequence.

We have previously observed with different derivatives of Λ-[Ru(TAP)$_2$(dppz)]$^{2+}$ that intercalation into the terminal TC/GA base step can force the terminal adenine (A10) to flip out and form a reverse Watson-Crick base pair with a symmetry related strand.[27] Therefore, as illustrated in Figure 2, the 'purine side' of the intercalation cavity was incomplete, even for an unsubstituted dppz. Perhaps most strikingly, we found that reversible dehydration of our original crystal exhibited a remarkable reversal, with the dehydrated form showing that the 'purine side' of the intercalation cavity was now intact, but the 'pyrimidine side' instead was the one which had flipped out.[33] In this work we report that Λ-11-nitrile substitution gives us the first example of a complete TC/GA cavity at this terminal intercalation step, with the asymmetry of the cavity suggesting that the addition of a nitrile moiety can generate additional favourable π orbital overlap, perhaps leading to enhanced specificity for DNA binding.

Results

Synthesis and crystal structure of (I) with d(TCGGCGCCGA)$_2$. [Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ (**I**) was synthesized using a variation of a previously published method as both the PF$_6^-$ and Cl$^-$ salts. The racemic PF$_6^-$ salt was recrystallized from acetonitrile via the vapour diffusion of diethyl ether to give crystals suitable for X-ray diffraction. The Cl$^-$ salt was purified by chromatography and was then suitable for crystallization with the DNA sequence d(TCGGCGCCGA)$_2$.

In some previous work we have found that well diffracting crystals could only be obtained with some DNA sequences by starting with the pure enantiomeric complex and the desired DNA oligonucleotide. In this case, the crystallisation was completely enantioselective and, on crystallization of the chloride salt of the racemic complex of the cation with the DNA decamer sequence d(TCGCCGCCGA)$_2$, red crystals were obtained. Diffraction data to 1.5 Å resolution were collected on beamline I02 at Diamond Light Source Ltd. Data collection and refinement statistics are given in Table S1.
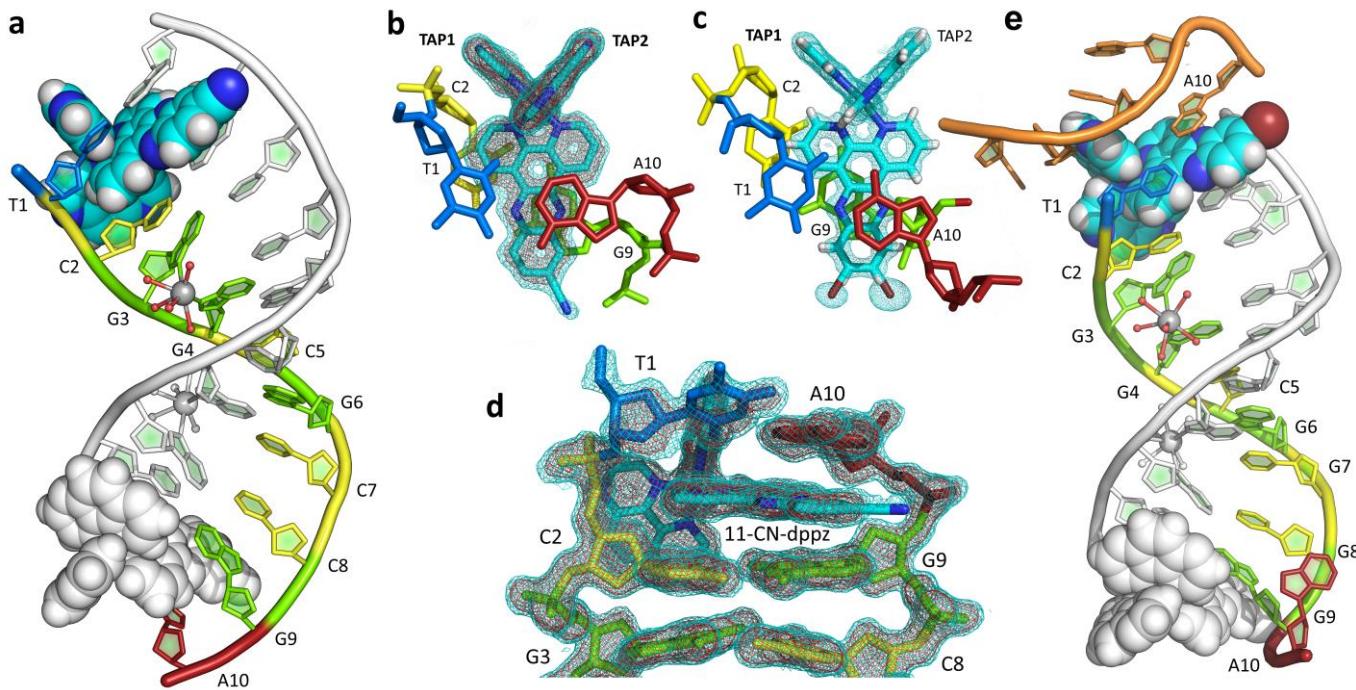
Figure 3 – The lambda enantiomers of (**I**) and (**II**) crystallised with the DNA sequence d(TCGGCGCCGA)$_2$. Any hydrogen atoms are shown in calculated positions. (a) The complete duplex assembly containing two asymmetric units (one coloured, one off white), omitting the symmetry generated complexes that interact at the G3G4/C7C8 step. The standard nucleic acid colour scheme of the Nucleic Acid Database is used for the bases, and Ba^{2+} ions have been coloured silver; (b) 2F$_o$-F$_c$ electron density map for (**I**) showing the major orientation of the dppz ligand. The cyan map is countoured at 0.29 e/Å^3 and the red map at 0.44 e/Å^3 ; (c) 2F$_o$-F$_c$ electron density map for (**II**) showing the major and minor Br orientations, with map contoured at 0.29 e/Å^3 ; (d) The ordered complete intercalation cavity with (**I**), contour levels as in (b); (e) The complete duplex assembly of (**II**) bound to d(TCGGCGCCGA)$_2$ containing two asymmetric units and a symmetry related strand (shown in orange) that forms a reverese Hoogsteen base pair to complete the intercalation cavity.

What was quite unexpected, and previously unobserved, was the efficiency of the nitrile substituent in anchoring the terminal adenine base (A10), with the creation of a complete intercalation cavity at the terminal T1C2/G9A10 step of the duplex (Figure 3 a,b,d). The complete assembly (Figure 3a) has twofold symmetry about the central steps. The orientation of the dppz ligand is determined by the contact between TAP1 and the 2′-deoxyribose sugar of cytosine (Figure 3b). The bound dppz moiety is still not aligned with the Ru-N square plane (Figure S3b). The level of detail visible at 1.5 Å resolution (Figure S2) is sufficient to show partial disorder of the 11-CN substituent in the resulting structure. Electron density fitting revealed the presence of a minor component, giving the best fit at 0.33 occupancy. As with previously reported structures using this sequence, the crystals contained only the lambda enantiomer of the complex, at a binding stoichiometry of 1:1 complex to single DNA strand. Further details of backbone conformation and water structure are shown in Figures S3 and S4.

The structure shows a ~50° kink at the central G5C6/G5C6 step, previously seen in our reversible crystal dehydration study for the less hydrated form.[33] In that work, the dehydration produced a remarkable switch from an open purine cavity, in the hydrated crystal, to an open pyrimidine cavity in the dehydrated crystal. In the present example, both sides of the cavity are complete, and the crystal form was obtained without the use of a humidity controlled environment. The asymmetry of the cavity is shown by the γ backbone dihedral angles, which are 189° on the 'pyrimidine side' and 63° on the 'purine side' (see Discussion section and Table 1 for comparisons). Table S2 lists the derived parameters for all the structures discussed here, and Figures S5 and S6 show the extent of the stacking between the dppz ligand and the G9 base. The N7 of the guanine base is close to the 11-CN position. A barium cation originating from the crystallization solution is present in the major groove at the G3G4/C7C8 step, where TAP2 semi-intercalates, forming a 50° kink as we have previously observed, and which accounts for the complete enantioselectivity of the crystal packing. The Ba-Ba distance is 8.0 Å, suggesting an intermediate degree of hydration, also supported by the short c axial direction (*a* = 47.88 Å, *c* = 29.14 Å) consistent with that previously reported.[33] The orientation of the dppz chromophore is determined by the cytosine C2 sugar ring face contacts to the deoxyribose hydrogen atoms of C1′, C2′ and C4′ with the TAP2 ligand, as described in our previous paper concerning the effect of the orientation of inosine substitution at this position.[27] The major orientation of the nitrile substituent is on the G9-A10 (the 'purine side') of the complete intercalation cavity.

WILEY-VCH
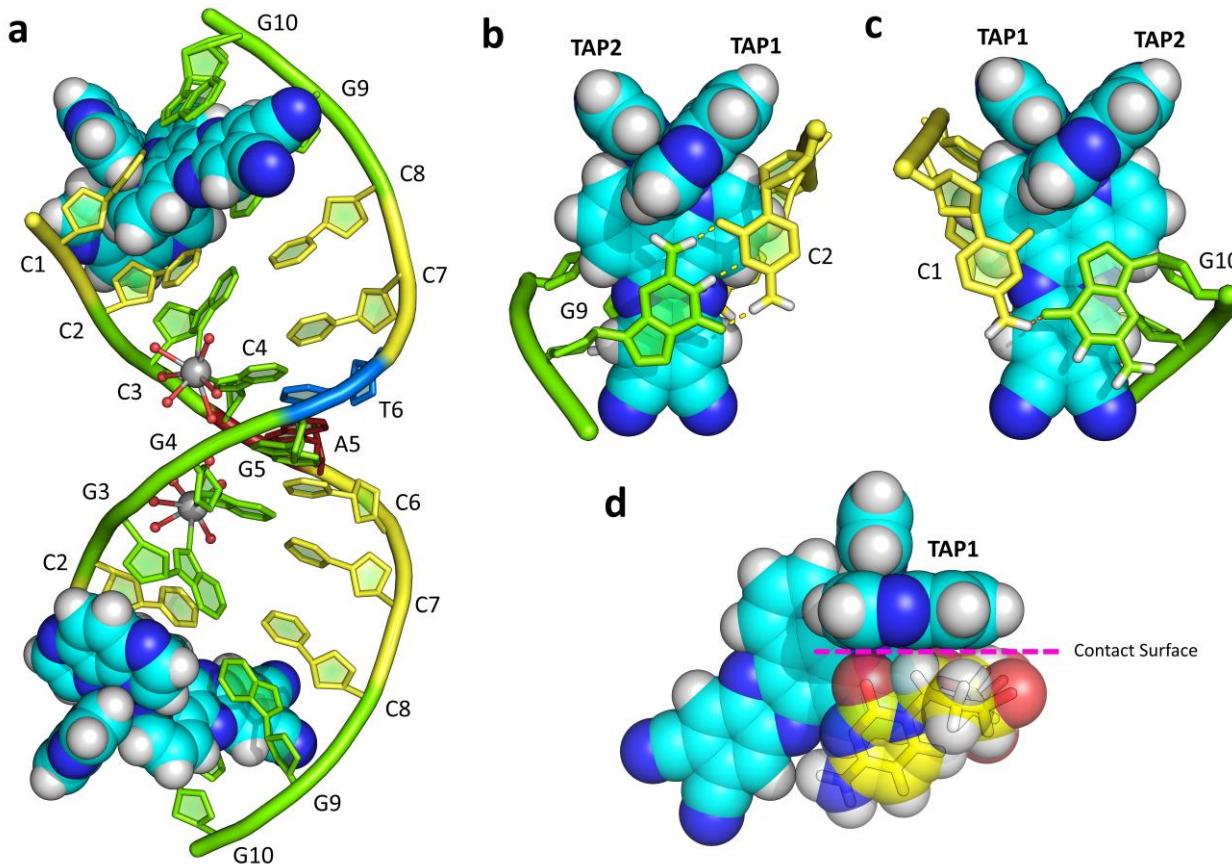
Figure 4 – The lambda enantiomer of **(III)** crystallised with the DNA sequences d(CCGGACCCGG) and d(CCGGGTCCGG), where both strands combine stoichiometrically. (a) shows the complete duplex assembly (one asymmetric unit) omitting the symmetry generated complexes that interact at the G3G4/C7C8 step. The structure was solved/refined in space group *P4*₃ where 50:50 disorder at the central step is observed (not shown). Standard nucleic acid colour scheme is used and Ba^{2+} ions have been coloured silver with red oxygen atoms (depicting water). Projections of the (b) C2-G9 and (c) C1-G10 base pair, onto the dppz ligand plane, omitting other residues. (d) Space filling representation of the C2-deoxyribose sugar contact with the ligand TAP1, which determines the angle of intercalation of the dppz ligand in the cavity.

Crystal structure of (II) with d(TCGGCGCCGA)₂. [Ru(TAP)₂(11-Br-dppz)]$^{2+}$ (**II**) was then synthesised using a variation of a previously published method as both the PF₆$^-$ and Cl$^-$ salts, with the latter used for crystallisation with DNA. The crystal structure of Λ-(**II**) with the sequence d(TCGGCGCCGA)₂ was determined to better than 1.1 Å resolution (Table S1), giving an extremely clear map (Figure 3c). The structure obtained (Figure 3e) was isomorphous to that seen with the Λ-11-Cl analogue with a flipped-out A10 stacked on a symmetry equivalent dppz ligand and 2:1 disorder of the 11-Br-dppz ligand.[30] The major Br orientation is on the 'purine side' of the open intercalation cavity. Unlike the effect of 11-CN substitution, there is no water network around the Br positions, despite the high data resolution and the location of 88 water molecules per DNA strand (Figures S7 and S8). The Br atom projects directly into the disordered part of the solvent space. The Ba-Ba distance in the major groove is 9.40 Å, as previously observed for the Λ-11-Cl analogue, and is associated with the fully hydrated form of this structure. A comparison of Figures 3b and 3c highlights the difference in adenine (A10) orientation resulting from the different intercalation cavities.

Crystal structure of (III) with d(CCGGACCCGG/CCGGGTCCGG).
The symmetrical compound [Ru(TAP)₂(11,12-diCN-dppz)]$^{2+}$ (**III**) was then synthesised for comparison. Attempted crystallisation

of the racemic chloride salt with d(TCGGCGCCGA)₂ was unsuccessful, but crystals were obtained with closely related decamer sequences. The best diffracting crystals, to 1.67 Å resolution (Table S1), were obtained with the asymmetric central step A5C6/G5T6 generated by sequence combination d(CCGGACCCGG/CCGGGTCCGG) (Figure 1c and Figure 4). This crystal structure also showed complete cavity formation, and because of the central asymmetry, was solved and refined in the lower symmetry space group P4₃, in which the asymmetric unit is the full duplex, as shown in Figure 4a and Figures S9 and S10. The CN substituents protrude into the solvent cavity from the major groove of the DNA, as shown in Figures 4b and 4c, and do not overlap directly with the DNA bases. Figure 4d shows that the orientation of the dppz is determined by the TAP1-cytosine C2 contact, so that the nitrile groups are not contained within the DNA base stack. At the intercalation cavity there is evidence of backbone disorder, and the phosphate group of residue G10 was modelled as a mixture of B$_I$ and B$_{II}$ conformations.[34] As we have previously seen with the unsubstituted dppz and this sequence,[25] there is a reversal in the conformation of the terminal guanine base G10, so that it stacks in a *syn* conformation with the negative, and normally major groove face of the guanine stacked over the pyrazine ring, and aligning the 2-NH₂ vector direction of the guanine with one

of the ligand -CN groups. Further details of map quality and water structure are shown in Figures S9 and S10.

Discussion

This first study of the comparative effect of nitrile and bromo substitution on dppz binding has highlighted both the effect of electron withdrawal and of an additional lone pair donor on the distal ring of the dppz. The structural evidence is that it affects both the base stacking and the solvent water interactions.
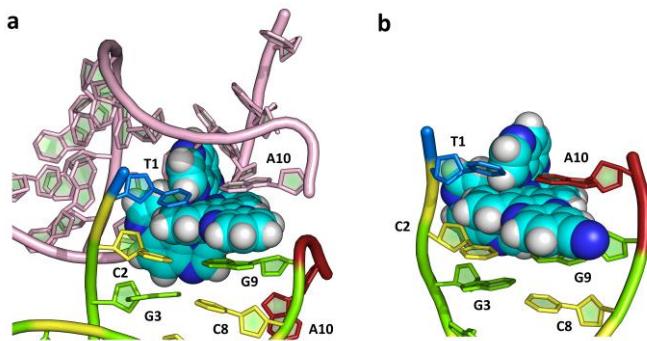

Figure 5 – Crystal structures of (a) Λ-[Ru(TAP)$_2$(dppz)]$^{2+}$ and (b) Λ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ bound to DNA sequence d(TCGGCGCCGA)$_2$ highlighting the differences in intercalation cavity. The symmetry related strand is shown as light pink showing how the terminal adenine (A10) flips out, whereas the A10 in (b) is anchored, forming a closed cavity and *gauche* γ torsion angle.

Cavity stabilisation/base stacking. The structural work reported here shows the unexpected effect of a substituent nitrile on the dppz ligand in stabilising the intercalation cavity formed by a TA basepair (Figure 5). It causes the T1-C2 side of the intercalation cavity to adopt the expanded backbone conformation with γ dihedral angle of 189° (Table 1 and Figure S3). Our previous work on the structure of *rac*-[Ru(phen)$_2$dppz]$^{2+}$ with the hexamer duplex d(ATGCAT)$_2$ also showed asymmetric complete cavities, for both enantiomers of the complex, and in that work could be directly related back to the luminescence behaviour of the enantiomers.[26] The crucial effect of backbone expansion on luminescence behaviour is the increased exposure of the dppz ligand to luminescence quenching, due to the additional hydrogen bonding by solvent water. X-ray crystallography has been uniquely useful in interpretation of such differences.

Table 1 – Selected γ dihedral angles (°) for crystal structures of d(TCGGCGCCGA)$_2$ with a range of derivatised dppz complexes.

Complex	PDB Accession Nº	γ(T1/C2)	γ(G9/A10)
Λ-[Ru(TAP)$_2$(10-Me-dppz)]$^{2+}$	4MJ9	58.4	187.2
Λ-[Ru(TAP)$_2$(11-Me-dppz)]$^{2+}$	4X18	59.7	189.5
Λ-[Ru(TAP)$_2$(10,12-Me$_2$-dppz)]$^{2+}$	4X1A	59.9	190.7
Λ-[Ru(TAP)$_2$(11,12-Me$_2$-dppz)]$^{2+}$	4E8S	59	186.5
Λ-[Ru(TAP)$_2$(11-Cl-dppz)]$^{2+}$	4III	57.7	194
Λ-[Ru(TAP)$_2$(11-Br-dppz)]$^{2+}$	6GLD	55.6	61.2
Λ-[Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$	5NBE	188.9	63.2

Table 1 compares the dihedral angles derived in the present work for the T1C2/G9A10 step with those from our previous work on -Cl and -Me substitution, which in every case gave γ dihedral angles in the *gauche* range of 57-60° on the pyrimidine side.[30,35] In all these cases the cavity was incomplete. Interestingly, for the first time with the 11-Br substitution we see a *gauche* dihedral angle at the purine G9-A10 as well, which on closer inspection reveals that, what in the other structures is an unfavourable interaction, may instead be an attractive interaction between the C3′-OH and the bromine atom, with an O3′-Br separation of 4.2 Å.

The closure of the intercalation cavity with the 11-CN substituent suggest both strong electron withdrawal and a favourable stacking interaction. Since intercalation is favoured by the π-π stacking interactions of the intercalating moiety with the surrounding base pairs, altering the π quadrupole of the interacting complex should modify the binding affinity. Nucleobases, especially guanine, are electron rich, and due to the electrostatic repulsions caused by direct overlap of π orbitals, sandwich and parallel-displaced stacking formations favour less negative π density. Therefore the electron withdrawing nitrile group on the ligand presumably polarises the dppz, relocating π electron density away from the interacting π orbitals, rationalising the favourable π-π stacking interactions seen here.

This observation suggests a design lead for more specific binding agents, by the direct modification of the π framework. Perhaps in this manner much weaker bound base steps or base mismatch/mutations could be targeted by such fine tuning of electronic properties. The targeting of mismatches probably requires enantiomer separation, since that is (structurally) a property of the delta enantiomer. [36,37] Such distal ring derivatisation has been shown to increase cellular uptake and heighten the potency of the proposed PDT photosensitizers.[13]

Nitrile group orientation and water structure. The major orientation of the 11-CN substituent, shown in both Figures 3 and 5, and for the 11-Br substituent in Figure 3, corresponds to that previously seen for 11-Me and 11-Cl substitution.[30,35] The effect on the water structure is different, however. Here we observe direct hydrogen bonding to the nitrile N atom when bound to DNA (Figure S4), also present with the 11,12-diCN ligand (Figure S10), creating additional water ordering. Strikingly, the methyl group substitution seen with the asymmetric 10-Me-, 11-Me- and 10,12-Me$_2$-dppz[35] are the most strongly directional (where total ordering was seen with X-ray data to 0.9 Å resolution in each of these cases), and here the ordered water creates a cage around the ligand, but there is no ordered water structure around the methyl groups in the major groove (Figure S11). The 11-Br structure studied in this work is a substituent with a steric effect very similar to that of methyl group, and also does not generate any water ordering (Figure S8b). The methyl

group is also the most hydrophobic and electron donating of those studied, but nevertheless did not stabilise the intercalation cavity. This difference suggests that we can distinguish hydrophobic, steric, and electronic effects on orientation, and propose that the alternative orientation for methyl substituents is strongly disfavoured by the preference of the more hydrophobic substituent to remain within the intercalation cavity rather than project into the major groove. Intercalatory interactions are energetically favourable, especially so when intercalation disturbs the hydration sphere of the binding site such as with the methylated complexes. Such logic has been used to explain why the binding constant for analogous methylated complexes is larger than that of the unsubstituted parent. [38] Therefore we hypothesise that the effects on orientation of asymmetrically substituted moeties are not only a consequence of a balance between the attractive polar contacts/Van der Waals forces and the increased entropy of hydration on binding, but additionally the hydrophobicity of the dppz substituent. Further systematic studies will be needed to confirm this proposal.

Stabilisation of *syn* guanine. A final point of note is the role of purines in the stabilising of stacking interactions, particularly the different adenine and guanine orientations consistent with their differing polarities, and previously seen with the unsubstituted dppz ligand.[25] Figure 6 summarises this difference, comparing both the two purine bases as well as the introduction of the electron withdrawing substituent. The adenine base stacks on the 11-CN dppz ring to give a normal AT basepair, with the 6-amino substituent pointing towards the major groove and the depth of intercalation determined by the hydrogen atom at the 2-position. In contrast, the guanine base stacks directly over the 11,12-diCNdppz ligand, with the 6-carbonyl group directed away from the nitrile substituents and over one pyrazine N atom. The amino substituent, in the 2-position, is again directed toward the major groove and the nitrile substituents. The guanine is thus stabilised in a *syn* conformation, even though there is no stabilisation due to additional hydrogen bond formation.
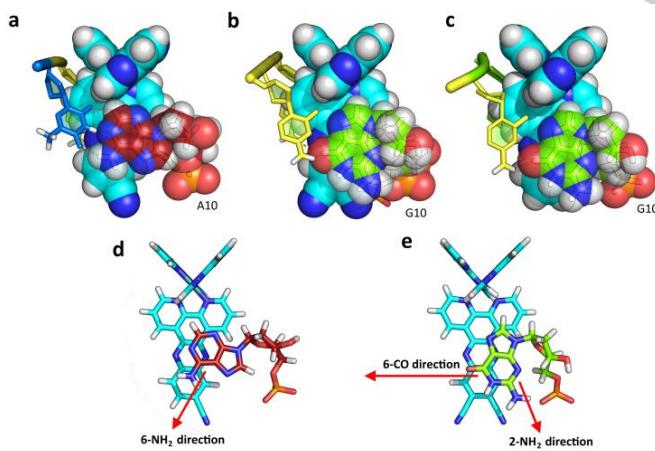

Figure 6 – Adenine and guanine stacking with the dppz ligand, showing the connection to base polarity. Adenine adopts the anti (Watson-Crick) conformation, wheras guanine adopts the syn (Hoogsteen) conformation. The difference can be related to the polarity of the bases. (a) The stacking of A10 on the 11-CN substituted dppz, seen for the first time in this work; (b) the stacking of G10 on the 11,12-diCN-dppz chromophore, reported here. (c) Comparison with the corresponding dppz-G10 stacking seen previously. (d) and (e) show stick representations of (a) and (b) highlighting the directionality of the polarity of the bases.

A new conclusion which can be drawn is that the guanine alignment is unaffected by the introduction of the electron withdrawing substituent. Comparing Figures 6b and 6c, the guanine orientation is almost identical, with no contact to the ancillary TAP or phen ligand, but a precise alignment of the carbonyl group over one pyrazine N. The adenine comparison is now possible because of the effect of the 11-CN substitution in figure 6a. Here the adenine H2 is in contact with the TAP2 ligand and may therefore influence the depth of intercalation. Figures 6d and 6e show the greater stacking of the guanine compared to adenine, and the differing amino and carbonyl orientations.

Substituted dppz ligands bound to platinum have shown interesting G-quadruplex binding properties, [39] and this area merits further study, as *syn*-guanine stabilisation could be key to the conformation adopted. [5]

Conclusions

In our crystallographic studies of ruthenium polypyridyl complex/oligonucleotide interactions, we have sought to provide a rationale for the design of new, more structurally specific compounds, and a means of interpreting biophysical data. The present work evolved from this approach, where the new compounds were chosen for their probable crystallisability and desirable TRIR reporting abilities. They have shown the orienting effect of asymmetric dppz substitution but also the unexpected additional stability conferred on the intercalation cavity by a single nitrile group, a property not shared by halogen or methyl substitution. We propose that by altering the electronic quadrupole moment of the dppz by substituent effects we create a preferable π-stacking arrangement for intercalation, and can infer that by controlling the electronic properties further we may see a higher specificity for electron rich or deficient base steps. Further, based on previous structures we suggest that the orientation of asymmetrically derivatised intercalators will depend on the hydrophobicity of the substituents, an observation which will aid the design of better targeted intercalators. A systematic understanding of such interactions is indispensable in improving our comprehension of the excited state photophysics of these systems, where structural knowledge helps to elucidate observations in solution phase.

Further studies with other substituents will allow us to further understand the balance between electronic, hydrophobic and steric effects. Further comparative studies are also required to establish whether the *syn* guanine stabilisation is enantiospecific. To date it has not been observed for the delta enantiomer.

Experimental Section

Synthetic Reagents and Materials

Unless otherwise stated, all materials and chemicals were sourced from Sigma Aldrich (Merck) or Honeywell research chemicals. Sephadex C-25 stationary phase anion exchange resin was purchased from GE Healthcare. All solvents, unless otherwise stated, were obtained at HPLC grade and used without further purification. Where further purification was needed, protocol from "Purification of Laboratory Chemicals, 4th edition, Armarego *et. al.*" was followed. Deuterated solvents for NMR analysis were purchased either through Sigma-Aldrich or Cambridge Isotope Laboratories.

Synthesis of Ruthenium Complexes (I) - (III).

Syntheses of the novel nitrile derivatives of [Ru(TAP)$_2$(dppz)]$^{2+}$ were carried out by modifying existing literature methods.[22] Both were prepared via the condensation of the relevant aromatic diamine with Ru(TAP)$_2$Cl$_2$, all of which were also synthesised by modifying our previously published literature methods.[30,35] Ru(TAP)$_2$Cl$_2$ (81 mg, 0.15 mmol) and 11-CN-dppz (49 mg, 0.16 mmol) (or 11,12-diCN-dppz (53 mg, 0.16 mmol)/11-Br-dppz (58 mg, 0.16 mmol)) were suspended together in an aqueous ethanol solution (7 mL, 1:1) within a CEM microwave tube (10 mL). The violet coloured solution was degassed/evacuated with Ar for 15 minutes before being fully sealed and installed into a microwave synthesiser. The sample was irradiated at 140 W at 60 ºC for 40 minutes, yielding a deep red/brown solution which was allowed to cool and then filtered *in vacuo*. Subsequent precipitation of the target compound from the filtrate was achieved by metathesis via dropwise addition of a saturated solution of aqueous potassium hexafluorophosphate (KPF$_6$). Isolation of the PF$_6^-$ salt by suction filtration yielded a dark orange/brown solid, which, after washing with cold water (2 x 2 mL) was allowed to dry in air. Conversion to the chloride form was achieved by the dissolution of the crude material in a minimal amount of acetonitrile (~5 mL), addition of HPLC grade water (10 mL), followed by dry, washed, Amberlite ion exchange resin (IRA-400, Cl- form, 2.4 g), covering, and lightly stirring for 20 hours. Following removal of the resin by gravity filtration, the complex was isolated via rotary evaporation and purified on an aqueous Sephadex C-25 column using 0.2 M NaCl as the mobile phase (eluting as a deep red/orange band). The compound was isolated as the chloride form, after anionic exchange via treatment with Amberlite resin (IRA-400, Cl- form, 2.4 g), to yield the complex as a deep red/brown microcrystalline solid. NMR spectra are shown in Figure S1.

(I) *[Ru(TAP)$_2$(11-CN-dppz)]Cl$_2$*. (94 mg, 0.11 mmol, 74 %). δ$_H$ (400 MHz, H$_2$O-d_2) – 9.74 (d, *J* = 8.4 Hz, 2H), 9.02 (dd, *J* = 3.2, 6.8 Hz, 2H), 8.99 (d, *J* = 3.2 Hz, 2H), 8.73 (s, 1H), 8.65 (s, 4H), 8.53 (dd, *J*= 7, 12.4 Hz, 2H), 8.47 (d, *J* = 8.8 Hz, 1H), 8.38 (t, *J* = 2.0 Hz, 2H), 8.24 (d, *J* = 5.2 Hz, 2H), 8.07 (d, *J* = 8.8 Hz, 1H) and 7.92 ppm (dd, *J*= 5.2, 8.8 Hz, 2H). δ$_C$ (101 MHz, H$_2$O-d_2) – 154.8, 154.7, 150.5, 149.1, 148.6, 145.2, 143.5, 142.6, 141.2, 135.9, 132.8, 132.2, 131.0, 130.6, 127.9 and 114.4 ppm. FT-IR - 3000 (broad, m, Arom. v(-C-H)) and 2232 cm^{-1} (m, Nitrile v(-C≡N)). HRMS-ESI (m/z) – Found (M+H$^+$, 386.5536); calc. 386.5538 (RuC$_{39}$N$_{13}$H$_{21}^{2+}$) (σ < 3 ppm).

(II) *[Ru(TAP)$_2$(11,Br-dppz)]Cl$_2$*. (86 mg, 0.10 mmol, 66 %). δ$_H$ (400 MHz, H$_2$O-d_2) – 9.71 (d, *J* = 8.3 Hz, 2H), 9.05-8.96 (m, 4H), 8.66 (s, 4H), 8.57 (s, 1H), 8.52 (d, *J* = 2.9 Hz, 1H), 8.47 (d, *J* = 2.9 Hz, 1H), 8.37 (t, *J* = 2.9 Hz, 2H), 8.32 (d, *J* = 9.1 Hz, 1H),

8.23-8.17 (m, 2H), 8.13 (dd, *J* = 9.1, 2.1 Hz, 1H) and 7.88 ppm (ddd, *J* = 8.2, 5.4, 2.7, 2H). HRMS-ESI (m/z) – Found (M$^+$, 413.0119); calc. 413.0128 (Ru^{101}C$_{38}$N$_{12}$BrH$_{21}^{2+}$) (σ < 3 ppm).

(II) *[Ru(TAP)$_2$(11,12-diCN-dppz)]Cl$_2$*. (86 mg, 0.10 mmol, 66 %). δ$_H$ (400 MHz, H$_2$O-d_2) – 9.80 (d, *J* = 8.3 Hz, 2H), 9.07-8.90 (m, 6H), 8.68 (s, 4H), 8.54 (d, *J* = 3.0 Hz, 2H), 8.40 (d, *J* = 3.0 Hz, 2H), 8.28 (d, *J* = 5.4 Hz, 2H) and 7.95 ppm (dd, *J* = 8.4, 5.4 Hz, 2H). δ$_C$ (101 MHz, H$_2$O-d_2) – 150.06, 150.63, 149.24, 149.08, 148.86, 148.54, 145.32, 145.25, 142.71, 142.28, 141.98, 141.85, 137.71, 135.22, 132.60, 132.51, 129.83, 127.72, 115.18 and 114.90 ppm. FT-IR – 3051 (broad, m, Arom. v(-C-H)) and 2231 cm^{-1} (m, Nitrile v(-C≡N)). HRMS-ESI (m/z) – Found (M$^+$, 398.5519); calc. 398.5520 (Ru^{101}C$_{40}$N$_{14}$H$_{20}^{2+}$) (σ < 3 ppm).

Macromolecular Crystallography

The oligonucleotides d(TCGGCGCCGA), d(CCGGACCCGG) and d(CCGGGTCCGG) were purchased as HPLC purified solids from Eurogentec Ltd and used without further purification.

Crystallisation, data collection and analysis of Λ-(1) with d(TCGGCGCCGA)$_2$. Crystals containing the oligonucleotide d(TCGGCGCCGA)$_2$ and ligand Λ-[Ru(TAP)$_2$(11-CN-dppz)]Cl$_2$ were grown via the vapour diffusion method from sitting drops at 291 K. Crystallisation was observed in a number of conditions from the Nucleic Acid Mini-Screen from Hampton Research, where the best diffracting example came from a 8µL drop containing; 125 µM d(TCGGCGCCGA)$_2$, 750 µM *rac*-[Ru(TAP)$_2$(11-CN-dppz)]Cl$_2$, 7.5% v/v MPD, 30 mM pH 7 sodium cacodylate, 9 mM spermine tetrahydrochloride, 60 mM KCl and 15 mM BaCl$_2$, all equilibrated against 500 µL of 35% v/v MPD. Orange/red rods grew following roughly 3 weeks of incubation at 291 K.

Diffraction data were collected from single, nitrogen flash-cooled crystal fragments at 100 K on beamline I02 at Diamond Light Source, Ltd. Data were automatically processed with xia2[40], using XDS[41] and XSCALE to integrate and merge peaks from all collected images; yielding 5452 unique reflections. The structure was solved using single-wavelength anomalous dispersion (SAD), using the anomalous diffraction of barium, with the SHELXC/D/E package[42]. The model was built by hand, using Wincoot[43], and refined against the original data using Refmac 5.0[44] in the CCP4 suite[45]. Ligand restraints were calculated using eLBOW[46] from the phenix [47] package. As a result of the asymmetry of the intercalating ligand, two sites of non-integer occupancy are observed. This was fitted by refining the occupancy of two complete complexes in the two mirrored orientations, with the sum adding to 100% occupancy This procedure preserves the correct restraints, and also is appropriate to the actual situation in the crystal used. 5% of reflections were reserved for the R$_{free}$ set. The final model has an R$_{cryst}$/R$_{free}$ of 0.16/0.19 and has been deposited in the Protein Data Bank with ID 5NBE.

Crystallisation, data collection and analysis of Λ-(II) with d(TCGGCGCCGA)$_2$. Crystals containing the oligonucleotide d(TCGGCGCCGA)$_2$ and ligand Λ-[Ru(TAP)$_2$(11-Br-dppz)]Cl$_2$ were grown via the vapour diffusion method from sitting drops at 291 K. Crystallisation was observed in a number of conditions

from the Nucleic Acid Mini-Screen from Hampton Research, where the best diffracting example came from a 8µL drop containing; 125 µM d(TCGGCGCCGA)$_2$, 625 µM *rac*-[Ru(TAP)$_2$(11-Br-dppz)]Cl$_2$, 7.5% v/v MPD, 30 mM pH 7 sodium cacodylate, 9 mM spermine tetrahydrochloride, 60 mM KCl and 15 mM BaCl$_2$, all equilibrated against 500 µL of 35% v/v MPD. Orange/red rods grew following roughly 3 weeks of incubation at 291 K.

Diffraction data were collected from single, nitrogen flash-cooled crystal fragments at 100 K on beamline I03 at Diamond Light Source, Ltd. Data were automatically processed with xia2[40], using DIALS[48] and Aimless[49] to integrate and merge peaks from all collected images; yielding 16,081 unique reflections. The structure was solved using single-wavelength anomalous dispersion (SAD), using the anomalous diffraction of barium, with the SHELXC/D/E package[42]. The model was built by hand, using Wincoot[43], and refined against the original data using Phaser[50] in the Phenix software package[51]. Ligand restraints were calculated using eLBOW [46] from the phenix[47] package. As a result of the asymmetry of the intercalating ligand, two sites of non-integer occupancy are observed. This was fitted by refining the occupancy of two complete complexes in the two mirrored orientations, with the sum adding to 100% occupancy This procedure preserves the correct restraints, and also is appropriate to the actual situation in the crystal used. 5% of reflections were reserved for the R$_{free}$ set. The final model has an R$_{cryst}$/R$_{free}$ of 0.15/0.16 and has been deposited in the Protein Data Bank with ID 6GLD.

Crystallisation, data collection and analysis of Λ-(III) with asymmetric decamer d(CCGGACCCGG)/d(CCGGGTCCGG).
Crystals were grown from sitting drops by vapour diffusion at 291 K. The drop contained the premixed and annealed duplex DNA at a concentration of 125 µM d(CCGGACCCGG)/d(CCGGGTCCGG), 125 µM *rac*-[Ru(TAP)$_2$(11,12-diCN-dppz)]Cl$_2$, 7.5% v/v MPD, 30 mM pH 7 sodium cacodylate, 9 mM spermine tetrahydrochloride, 60 mM KCl and 15 mM BaCl$_2$, equilibrated against 500 µL of 35% v/v MPD. Orange crystals grew after several weeks. A large crystalline fragment was obtained from a larger sample using a microloop with an elongated aperture to give crystal dimensions of 60 µm x 30 µm x 100 µm).

Diffraction data were collected from a single, nitrogen flash-cooled crystal at 100 K on beamline I03 at Diamond Light Source, Ltd. Data were processed with xia2, using DIALS[48] and Aimless[49] to integrate and merge peaks from all collected images. The structure was solved using MR-SAD, using the anomalous diffraction of barium, with Phaser[50] in the Phenix software package[51]. Ligand restraints were calculated using eLBOW [46] from the phenix [47] package. 5% of reflections were reserved for the R$_{free}$ set. The final model has an R$_{cryst}$/R$_{free}$ of 0.19/0.20 and has been deposited in the Protein Data Bank with ID 6G8S. Analysis of DNA dihedral angles was performed using DNATCO[34]. Structural diagrams were created using PyMol (Schrödinger).

Acknowledgements

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

[2]   H. Song, J. T. Kaiser, J. K. Barton, *Nat. Chem.* **2012**, *4*, 615–620.

[3]   A. N. Boynton, L. Marcélis, A. J. McConnell, J. K. Barton, *Inorg. Chem.* **2017**, *56*, 8381–8389.

[4]   S. Shi, X. Gao, H. Huang, J. Zhao, T. Yao, *Chem. - A Eur. J.* **2015**, *21*, 13390–13400.

[5]   E. Wachter, D. Moyá, S. Parkin, E. C. Glazer, *Chem. - A Eur. J.* **2016**, *22*, 550–559.

[6]   L. Jacquet, R. J. H. Davies, A. Kirsch-De Mesmaeker, J. M. Kelly, *J. Am. Chem. Soc.* **1997**, *119*, 11763–11768.

[7]   L. Ghizdavu, F. Pierard, S. Rickling, S. Aury, M. Surin, D. Beljonne, R. Lazzaroni, P. Murat, E. Defrancq, C. Moucheron, et al., *Inorg. Chem.* **2009**, *48*, 10988–10994.

[8]   S. Kajouj, L. Marcélis, V. Lemaur, D. Beljonne, C. Moucheron, *Dalt. Trans.* **2017**, *46*, 6623–6633.

[9]   M. Rebarz, L. Marcélis, M. Menand, D. Cornut, C. Moucheron, I. Jabin, A. Kirsch-De Mesmaeker, *Inorg. Chem.* **2014**, *53*, 2635–2644.

[10]   S. Le Gac, M. Foucart, P. Gerbaux, E. Defrancq, C. Moucheron, A. Kirsch-De Mesmaeker, *Dalton Trans.* **2010**, *39*, 9672–83.

[11]   S. G. Bown, *Philos. T. R. Soc. A.* **2013**, *371*, 20120371.

[12]   V. Pierroz, R. Rubbiani, C. Gentili, M. Patra, C. Mari, G. Gasser, S. Ferrari, *Chem. Sci.* **2016**, *7*, 6115–6124.

[13]   F. Heinemann, J. Karges, G. Gasser, *Acc. Chem. Res.* **2017**, *50*, 2727–2736.

[14]   C. Mari, R. Rubbiani, G. Gasser, *Inorganica Chim. Acta* **2017**, *454*, 21–26.

[15]   J. Hess, H. Huang, A. Kaiser, V. Pierroz, O. Blacque, H. Chao, G. Gasser, *Chem. - A Eur. J.* **2017**, *23*, 9888–9896.

[16]   F. E. Poynton, S. A. Bright, S. Blasco, D. C. Williams, J. M. Kelly, T. Gunnlaugsson, *Chem. Soc. Rev.* **2017**, *46*, 7706–7756.

[17]   C. Mari, V. Pierroz, S. Ferrari, G. Gasser, *Chem. Sci.* **2015**,

6, 2660–2686.

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

[19] B. C. Poulsen, S. Estalayo-Adrián, S. Blasco, S. A. Bright, J. M. Kelly, D. C. Williams, T. Gunnlaugsson, B. la Cour Poulsen, S. Estalayo-Adrian, S. Blasco, et al., *Biophys. Chem.* **2016**, *45*, 18208–18220.

[20] S. M. Cloonan, R. B. P. Elmes, M. Erby, S. A. Bright, F. E. Poynton, D. E. Nolan, S. J. Quinn, T. Gunnlaugsson, D. C. Williams, *J. Med. Chem.* **2015**, *58*, 4494–4505.

[21] E. Wachter, D. Moyá, E. C. Glazer, *ACS Comb. Sci.* **2017**, *19*, 85–95.

[22] I. Ortmans, B. Elias, J. M. Kelly, C. Moucheron, A. Kirsch-DeMesmaeker, *Dalton Trans.* **2004**, *2*, 668–676.

[23] P. M. Keane, J. M. Kelly, *Coord. Chem. Rev.* **2018**, *364*, 137–154.

[24] J. P. Hall, K. O'Sullivan, A. Naseer, J. A. Smith, J. M. Kelly, C. J. Cardin, *Proc. Natl. Acad. Sci.* **2011**, *108*, 17610–17614.

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

[26] 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, et al., *J. Am. Chem. Soc.* **2013**, *135*, 12652–12659.

[27] P. M. Keane, J. P. Hall, F. E. Poynton, B. C. Poulsen, S. P. Gurung, I. P. Clark, I. V. Sazanovich, M. Towrie, T. Gunnlaugsson, S. J. Quinn, et al., *Chem. - A Eur. J.* **2017**, *23*, 10344–10351.

[28] J. P. Hall, S. P. Gurung, J. Henle, P. Poidl, J. Andersson, P. Lincoln, G. Winter, T. Sorensen, D. J. Cardin, J. A. Brazier, et al., *Chem. - A Eur. J.* **2017**, *23*, 4981–4985.

[29] C. J. Cardin, J. M. Kelly, S. J. Quinn, *Chem. Sci.* **2017**, *8*, 4705–4723.

[30] J. P. Hall, H. Beer, K. Buchner, D. J. Cardin, C. J. Cardin, *Philos. Trans. A. Math. Phys. Eng. Sci.* **2013**, *371*, 20120525.

[31] P. M. Keane, F. E. Poynton, J. P. Hall, I. V. Sazanovich, M. Towrie, T. Gunnlaugsson, S. J. Quinn, C. J. Cardin, J. M. Kelly, *Angew. Chemie - Int. Ed.* **2015**, *54*, 8364–8368.

[32] J. P. Hall, F. E. Poynton, P. M. Keane, S. P. Gurung, J. A. Brazier, D. J. Cardin, G. Winter, T. Gunnlaugsson, I. V. Sazanovich, M. Towrie, et al., *Nat. Chem.* **2015**, *7*, 961–967.

[33] J. P. Hall, J. Sanchez-Weatherby, C. Alberti, C. H. Quimper, K. O'Sullivan, J. A. Brazier, G. Winter, T. Sorensen, J. M. Kelly, D. J. Cardin, et al., *J. Am. Chem. Soc.* **2014**, *136*, 17505–17512.

[34] B. Schneider, P. Božíková, I. Nečasová, P. Čech, D. Svozil, J. Černý, *Acta Crystallogr. Sect. D Struct. Biol.* **2018**, *74*, 52–64.

[35] J. P. Hall, H. Beer, K. Buchner, D. J. Cardin, C. J. Cardin, *Organometallics* **2015**, *34*, 2481–2486.

[36] A. N. Boynton, L. Marcélis, J. K. Barton, *J. Am. Chem. Soc.* **2016**, *138*, 5020–5023.

[37] J. P. Hall, P. M. Keane, H. Beer, K. Buchner, G. Winter, T. L. Sorensen, D. J. Cardin, J. A. Brazier, C. J. Cardin, *Nucleic Acids Res.* **2016**, *44*, 9472–9482.

[38] A. K. F. Mårtensson, P. Lincoln, *Phys. Chem. Chem. Phys.* **2018**, *20*, 7920–7930.

[39] M. Dik-Lung, C. M. Che, S. C. Yan, *J. Am. Chem. Soc.* **2009**, *131*, 1835–1846.

[40] G. Winter, *J. Appl. Crystallogr.* **2010**, *43*, 186–190.

[41] W. Kabsch, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 125–132.

[42] G. M. Sheldrick, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 479–485.

[43] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 486–501.

[44] G. N. Murshudov, P. Skubák, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, A. A. Vagin, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67*, 355–367.

[45] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, et al., *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67*, 235–242.

[46] N. W. Moriarty, R. W. Grosse-Kunstleve, P. D. Adams, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2009**, *65*, 1074–1080.

[47] P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, et al., *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 213–221.

[48] J. M. Parkhurst, G. Winter, D. G. Waterman, L. Fuentes-Montero, R. J. Gildea, G. N. Murshudov, G. Evans, *J. Appl. Crystallogr.* **2016**, *49*, 1912–1921.

[49] P. R. Evans, G. N. Murshudov, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2013**, *69*, 1204–1214.

[50] A. J. McCoy, R. J. Read, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 458–469.

[51] P. D. Adams, R. W. Grosse-Kunstleve, L. W. Hung, T. R. Ioerger, A. J. McCoy, N. W. Moriarty, R. J. Read, J. C. Sacchettini, N. K. Sauter, T. C. Terwilliger, et al., *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2002**, *58*, 1948–54.

COMMUNICATION

WILEY-VCH

Structural studies reveal the enantiospecific recognition of a DNA G-quadruplex by a ruthenium polypyridyl complex

Kane McQuaid,[a,c] Holly Abell,[a] Sarah P. Gurung,[a,c] David R. Allan,[c] Graeme Winter,[c] Thomas Sorensen,[c] David J. Cardin,[a] John A. Brazier,*[b] Christine J. Cardin,*[a] and James P. Hall*[a,b,c]

Abstract: Using X-ray crystallography, we show an enantiospecificity in DNA G-quadruplex binding, using the complexes Λ/Δ-[Ru(TAP)$_2$(dppz-11-CN)]$^{2+}$ (TAP=1,4,5,8-tetraazaphenanthrene) containing the dppz (dipyridophenazine) ligand, paralleling the specificity of the complexes with duplex DNA. The Λ complex crystallises with the normally parallel stranded d(TAGGGTTA) tetraplex to give the first such antiparallel strand assembly in which *syn*-guanosine is adjacent to the complex at the 5' end of the quadruplex core. SRCD measurements confirm that the same conformational switch occurs in solution. The Δ enantiomer, by contrast, is present in the structure but stacked at the ends of the assembly. In addition, we report the structure of Λ-[Ru(phen)$_2$(11-CN-dppz)]$^{2+}$ bound to d(TCGGCGCCGA), a duplex forming sequence, and use both structural models to aid in the elucidation of the motif-specific luminescence response of the isostructural phen analogue enantiomers.

Guanine quadruplexes are four-stranded nucleic acid structures, formed by G-rich DNA and RNA sequences. They have been shown to play an important role in gene expression,[1] regulation[2] and have been visualised in human cells.[3] Targeting the G-quadruplex, by small-molecule binders, is an area of significant interest as stabilisation of the structure is an effective method of inducing apoptosis in cancer cells.[4] Development has focussed on compounds able to discriminate between duplex and quadruplex-forming DNA to favour the four stranded structure,[5] with examples including metallo-porphyrins,[6] acridines,[7] naphthalene-based compounds[8] and Pt-terpyridines.[9] The development of quadruplex-binding compounds as luminescent probes offers an attractive way of visualising such structures *in-vivo*. Octahedral polypyridyl ruthenium complexes are not only able to bind to and stabilise DNA G-quadruplexes but also possess a range of useful photophysical properties. For example, complexes containing the 1,10-phenanthroline (phen) and dipyridophenazine (dppz)

ligand can act as "light-switch" complexes, luminescing strongly when bound to DNA due to protection against excited state quenching via H-bonding with aqueous media.[10] Some examples have luminescence visible to the naked eye and are specific for G-quadruplexes.[11] Others can stabilise specific conformations of the human telomeric G-quadruplex sequence. [12] In contrast, ruthenium complexes containing the 1,4,5,8-tetraazaphenanthrene (TAP) ligand cause direct DNA damage, by guanine photooxidation, when exposed to visible radiation.[13] The absorption of light by the complex localises the damage to within several Å of the metal centre, as we have shown in both solution[14] and crystalline states,[15] and G-quadruplexes are particularly vulnerable to damage.[16]

Recently we explored the structural effect of substitution on the distal ring of [Ru(TAP)$_2$(dppz)]$^{2+}$.[17] We found that the addition of a nitrile substituent to give [Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ (**1**)(Figure 1a) caused (so far uniquely) the formation of a complete intercalation cavity when the lambda complex (Λ-**1**) bound to the d(TCGGCGCCGA) duplex,[18] showing that even this small modification of the dppz ligand strengthened the stacking interaction. Based on this finding we explored the binding of this compound **1** with a G-quadruplex forming sequence, reasoning that a G-quartet has a larger surface area available for π-stacking and therefore could accommodate the full footprint of the derivatised dppz group. This has led not only to the first crystal structure showing a mononuclear ruthenium polypyridyl complex bound to a DNA G-quadruplex but has also shed light on the structure-selective luminescence behaviour of the isostructural analogue, [Ru(phen)$_2$(11-CN-dppz)]$^{2+}$ (**2**). Most published structures of ligand binding to quadruplexes show extensive stacking (end-pasting) by a flexible planar ligand on the G-quartet surface, whereas the work reported here shows the additional feature of interaction with a ribose sugar. A summary of all metal containing ligands bound to G-quadruplexes that have been structurally characterised by X-ray or NMR is provided as Table S2. More generally, most of the compounds under development for specific targeting of key G-quadruplex containing genes are flat and often angled chromophores with short flexible substituents. [19]

[a]   K. McQuaid, H. Abell, Dr S. Gurung, Prof D.J. Cardin, Prof C.J. Cardin and Dr J. P. Hall
Department of Chemistry, University of Reading, Whiteknights, Reading, RG6 6AD, UK.
E-mail: james.hall@reading.ac.uk and c.j.cardin@reading.ac.uk
[b]   Dr J. A. Brazier and Dr J. P. Hall
Department of Pharmacy, University of Reading, Whiteknights, Reading, RG6 6AH, UK.
Email: j.a.brazier@reading.ac.uk
[c]   K. McQuaid, Dr D. R. Allan, Prof T. Sorensen, Dr G. Winter and Dr J. P. Hall
Diamond Light Source, Harwell Science and Innovation Campus, Didcot, Oxfordshire, OX11 0DE, UK.

Supporting information for this article is given via a link at the end of the document.
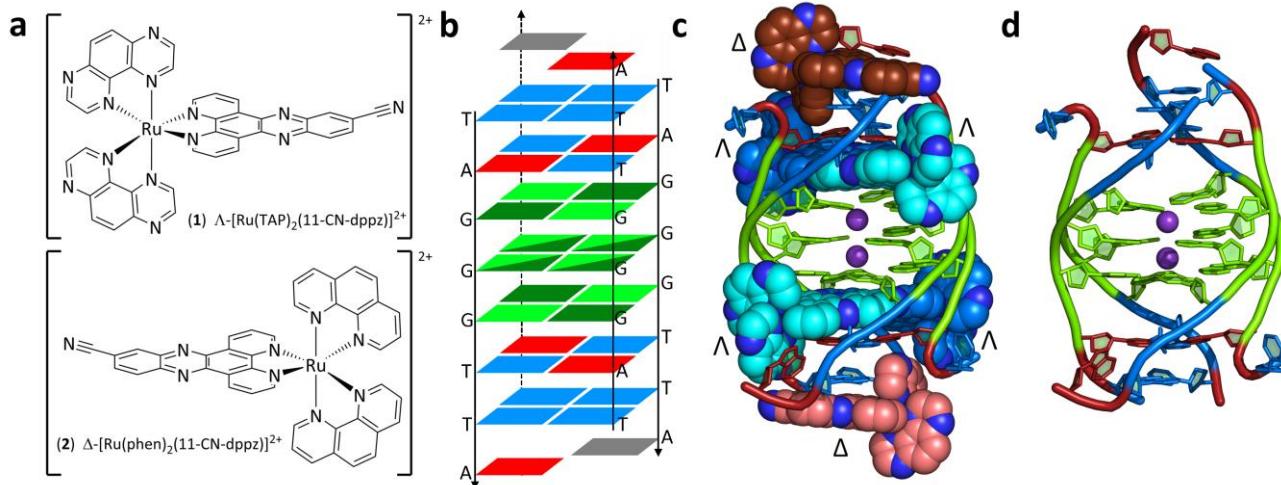
Figure 1 – (a) Skeletal formulae of the complexes, [Ru(TAP)$_2$(11-CN-dppz)]$^{2+}$ (**1**) and [Ru(phen)$_2$(11-CN-dppz)]$^{2+}$ (**2**). (b) Graphical representation of the crystallographic DNA assembly of d(TAGGGTTA) where adenine, guanine and thymine are coloured in red, green and blue respectively. Syn/Anti conformations of guanosine are highlighted using dark and light green, respectively. Grey marks a disordered base. (c) and (d) Crystallographic models (PDB:5LS8) showing Λ/Δ-**1** crystallised with the tetramolecular G-quadruplex d(TAGGGTTA) with and without the complex coordinates included, respectively. Λ-**1** complexes have been shown in teal or marine blue whereas Δ-**1** is shown in brown or salmon pink. Barium ions are shown in silver and potassium in purple. The oligonucleotide and the metal complex were annealed before crystallisation, and the resolution of the final dataset is 1.78 Å.

The parallel quadruplex forming sequence d(TAGGGTTA) has been widely studied but never previously crystallised.[20] Crystallization screening was performed using *rac*-**1** and this sequence, yielding crystals from which an X-ray structure was obtained (Figure 1b-d and S1-2). Crystallisation details and refinement results can be found in SI (Section S1.5 and Table S1). Structural analysis of binding between **1** and d(TAGGGTTA) reveals the direct interaction of Λ-**1** with the G-tetrad stack (Figure 2a-b), with the quadruplex unexpectedly adopting an antiparallel topology. The core of the quadruplex is stabilised by two potassium cations which coordinate to the O6

oxygen atoms in the three G-quartets. On either side of the central G-quartet stack, two Λ-**1** complexes are intercalated, giving an overall stoichiometry of one Λ-**1** per d(TAGGGTTA) strand, or four molecules bound to the tetraplex all bound with the same geometry. Inspection of the refined electron density shows that in the central quartet the guanine nucleosides are disordered, with all four bases split 50:50 between the *syn* and
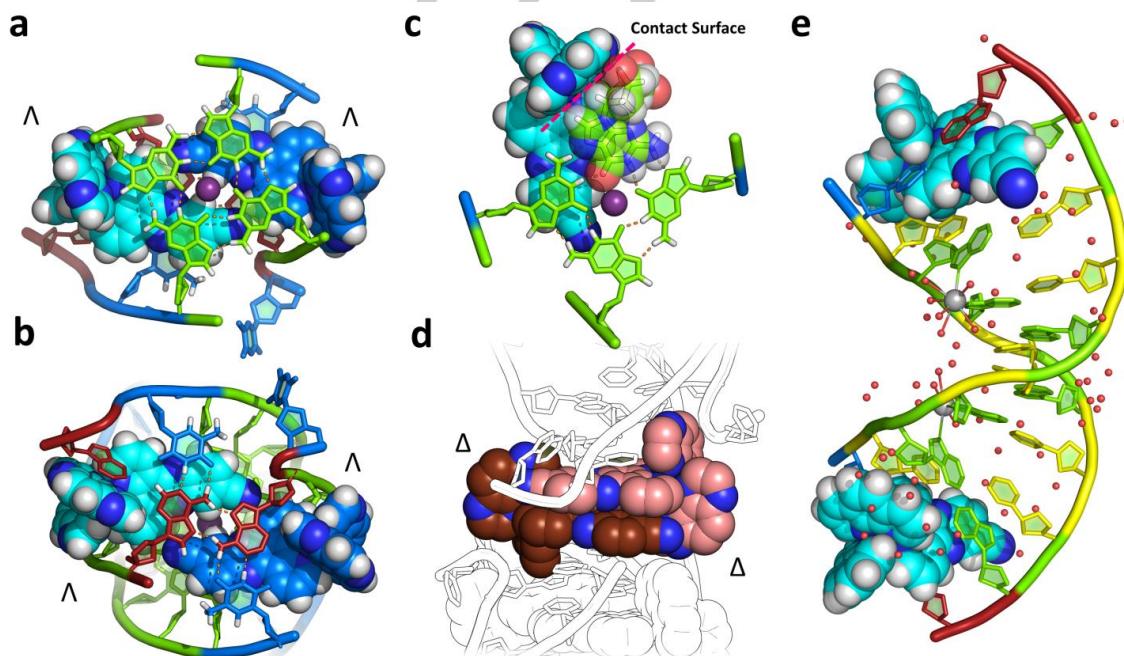

Figure 2 – (a-b) Crystallographic models showing the stacking environment of Λ-**1** to the adjacent B-DNA TA-TA side and the G-quadruplex tetrad respectively; (c) the overall guanine interaction of Λ-**1** highlighting the contact surface that we hypothesise determines intercalation angle and depth; (d) Δ-**1** stacking on the ends of the DNA assembly providing crystal packing between the biological units. (e) The complete duplex assembly of Λ-**2** crystallised with the duplex forming d(TCGGCGCCGA) PDB: 6HWG (see SI for further analysis).

anti-conformation (Figure S3), and that in the flanking quartets, *syn* and *anti* conformations alternate with neighbouring strands. The disorder means that all four strands exhibit the same ordering of conformations in a 5'-*syn-mix-anti*-3' manner, hence generating a π-stacking environment which is the same for each bound Λ-**1** complex. The 5'-*syn* guanosine residue interacts directly with one ancillary TAP ligand of Λ-**1**, with the face of the

to the duplex is isostructural to that previously reported for the TAP analogue and highlights the similarity between duplex and quadruplex intercalation angles; in both cases determined by the enantiospecific contact between ancillary ligand and the adjacent sugar.

Enantiospecificity in the photooxidation of $d(G_5C_5)$ by $[Ru(TAP)_2(dppz)]^{2+}$ has been previously explored, and we found
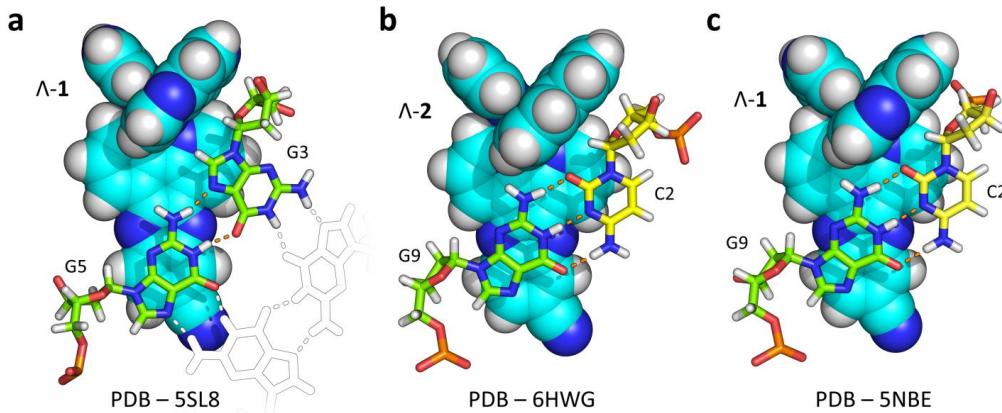
Figure 3– X-ray crystallographic models highlighting the similarities in binding between Λ-**1**/**2** to (a) the guanine tetrad; (b) and (c) B-DNA forming d(TCGGCGCCGA). Each structure shows the feature that favourable stacking of the ancillary ligand to the 5'- sugar determines the binding geometry. In the case of (a) this stabilises the formation of *syn* guanosine; in all the cases the adjacent guanosine is *anti* and is aligned remarkably similarly in both tetrad and B-DNA interactions, presumably to increase favourable π orbital overlap. Comparison of (b) and (c) shows the striking similarity of the binding orientations of Λ-**1** and Λ-**2**.

deoxyribose sugar contacting the face of the TAP ligand (Figure 2c). This *syn* arrangement results in a maximal stacking interaction between the surface of the 11-CN-dppz ligand and the purine bases. Each 11-CN-dppz ligand contacts three out of four of the guanine bases, stacking fully on two of them with the nitrile substituent forming a polar contact to the 2-NH$_2$ substituent of a third guanine. The combination of these interactions, with the nitrile group perfectly aligned to the G6 carbonyl lone pair, provides an optimal fit between complex and G-quartet surface where all four lambda complexes show the same set of interactions with the G-quartet. Δ-**1** is present in the structure where it stacks between the terminal T-T wobble pairs, bridging neighbouring biological units in the crystal packing, and showing no direct interaction with any G-quartet (Figure 2d). This structure shows how mononuclear ruthenium polypyridyl complexes can bind to DNA G-quadruplexes, demonstrating the complexities of ligand intercalation and providing potential binder design leads for next generation DNA probes and damage agents. The same Λ-**1** isomer bound in a duplex cavity also showed the features of *syn* guanosine stabilisation, maximisation of the stacking interaction and polarity alignment with the guanine substituents,[18] suggesting that these are key features for ligand design. Such characteristics are also observed in the structure reported here, of the isostructural 'light-switch' complex Λ-[Ru(phen)₂(11-CN-dppz)]²⁺ (**2**) bound to the same d(TCGGCGCCGA) duplex (Figure 2e and S4-6). This structure further reinforces observations to date that the spatial binding modes of phen derived complexes are comparable to those of the analogues containing TAP.[18] In all these cases, the stabilisation of *syn*-guanosine, but not *syn*-adenosine, is notable. Figure 3 presents how the binding cavity of the interaction of Λ-**2**

that the Λ enantiomer is more efficient than the Δ and also gives a higher quantum yield compared to the oxidation of $d(GC)_5$.[21] We interpreted this effect of sequence as due to the favourability of electron transfer through a stack of guanine bases and would expect the present assembly also to be a hotspot for photodamage. We suggest the observed enantiomeric disparities in photooxidation may be explained by a difference in proximity of the photoactive metal centre to the nucleotide. Favourable contact with the 5' sugar on the terminal guanosine allows the Λ complex to stack efficiently on the G-quartet surface in a way not possible if the chirality is reversed. The vulnerability of the 5'-guanosine of a G-quadruplex to chemical damage has been established in a detailed study by Burrows *et al.*[22]

Only two NMR structures of binuclear ruthenium complexes bound to G-quadruplexes have previously been reported,[23] with the Λ,Λ-enantiomer of the binuclear ruthenium complex threaded through a diagonal loop of an antiparallel quadruplex, but the Δ,Δ-enantiomer end-stacked to the lateral loop end of the same conformation. In the Λ,Λ- enantiomer model there are stacking interactions between an ancillary bpy ligand and a thymine residue in the loop, and it is estimated to bind about 40 times more strongly (Figure S10). For mononuclear complexes there are modelling studies [24] rather than NMR or crystal structure evidence. There are only four other published crystal structures showing metal complex binding to quadruplexes – and these are structurally unrelated planar species - two salphen complexes,[25] and two gold complexes[26,27] and with all these giving parallel-stranded assemblies in the crystal, in contrast to the antiparallel arrangement reported here. The most recently reported second gold complex structure shows a disordered end-pasted binding mode (Figure S11). [27]
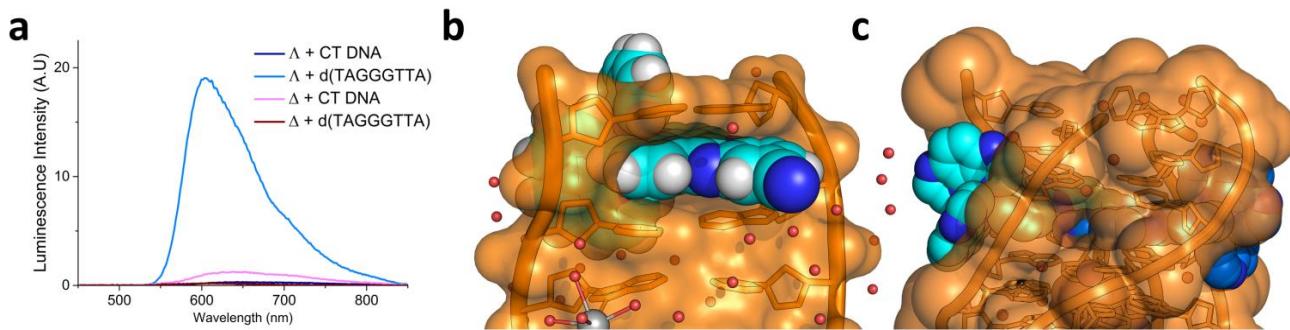
Figure 4 – (a) Emission spectra highlighting the luminescence intensity of Λ/Δ-**2** in the presence of CT DNA and G-quadruplex forming DNA d(TAGGGTTA), where λ_{ex}=440 nm. Spectra were run after annealing of the DNA samples in the presence of complex. (b), Calculated Van der Waals surfaces from the crystallographic models, highlighting how when bound to B-DNA a large proportion of the distal region of Λ-**2** is exposed in the major groove, allowing non-radiative relaxation pathways via H-bonding; and that when bound to a G-quadruplex (c) almost the entirety of the stacked ligand is encapsulated, giving protection from aqueous solvent.

We have previously used structural data to relate luminescence intensity of the 'light-switch' complex [Ru(phen)₂(dppz)]²⁺ to the degree of encapsulation of the dppz ligand.[28] While we would predict that the phenazine nitrogens are at least partially blocked when bound to most duplex DNA base steps,[29] the crystal structure of Λ-**2** bound to d(TCGGCGCCGA) shows that the nitrile moiety protrudes into the major groove and is thus accessible to solvent. Λ-**2** could be expected to be luminescent with d(TAGGGTTA) if the entirety of the intercalating ligand were to be encapsulated within the tetra-stranded motif as observed in the reported structure. In contrast it would be non-emissive when bound to duplex DNA as a result of the exposed substituent.

To investigate this, the enantiomers of **2** were separated by preparative chiral HPLC and the luminescence selectivity of the separated optical isomers was assessed using fluorescence spectroscopy. Figure 4a highlights a disparity in the observed luminescence intensity between the binding modes of **2** to a G-quadruplex and to calf thymus DNA (CT DNA), where the lambda enantiomer (Λ-**2**) is essentially non-emissive in the presence of B-DNA but luminesces brightly when bound to d(TAGGGTTA). The delta enantiomer (Δ-**2**) in contrast, exhibits little luminescence when bound to CT DNA and is non-emissive in the presence of the G-quadruplex. Figures 4b and 4c show how this luminescence enhancement can be related to the extent of encapsulation of the chromophore, and parallels the implication that the delta either does not intercalate into the quadruplex or does so but not deeply.

SRCD melting experiments were performed to examine the conformation of the quadruplex in solution, both in the presence and absence of the complex (Figure S8). All DNA samples were annealed in the presence of complex, and measured within 24 hours. The structure of the native tetrameric d(TAGGGTTA) assembly was assumed to be a parallel form, based on NMR data for the d(TGGGGT) quadruplex assembly.[30] Subsequently this parallel stranded model has been widely used to interpret ligand binding with this sequence.[20,31]. SRCD spectra that we obtained confirmed this topology assignment and this parallel conformation was maintained in the presence of a 4:1 ratio of the delta enantiomer (Δ-**1**). With the Λ enantiomer (Λ-**1**) the quadruplex was found to adopt an antiparallel topology,

consistent with the crystal structure reported here (Figure S9). For full details of the SRCD assignment see sections S1.6 and 2.8 in the supplementary information.

The structural evidence reported here is the first showing how a mononuclear ruthenium polypyridyl complex can bind to a DNA G-quadruplex, part of a larger programme of work which included studies towards a unimolecular G-quadruplex structure with these compounds. Unexpectedly, the Λ-enantiomer was shown to direct the formation of the quadruplex into an antiparallel assembly, an observation not mirrored by the Δ isomer; we postulate that this is a consequence of the increased stabilisation of *syn*-guanosine by the derivatised isomer. The rationalisation of the DNA structure-selective luminescence behaviour demonstrated here will not only allow for the systematic design of new complexes with increased luminescence response selectivity between duplex and higher-order DNA forms, but will allow us to extend this understanding to the design of new photooxidising agents to specifically damage the G-quadruplex, potentially with topological precision.

Experimental Section

For experimental please see supporting information.

Acknowledgements

[1]   D. Rhodes, H. J. Lipps, *Nucleic Acids Res.* **2015**, *43*, 8627–8637.

[2]   S. Kendrick, L. H. Hurley, *Pure Appl. Chem.* **2010**, *82*, DOI 10.1351/PAC-CON-09-09-29.

[3]   G. Biffi, D. Tannahill, J. McCafferty, S. Balasubramanian, *Nat. Chem.* **2013**, *5*, 182–186.

[4]   M. Düchler, *J. Drug Target.* **2012**, *20*, 389–400.

[5]   R. Vilar, in *Met. Dev. Action Anticancer Agents*, De Gruyter, Berlin, Boston, **2018**, pp. 325–350.

[6]   Q. Cao, Y. Li, E. Freisinger, P. Z. Qin, R. K. O. Sigel, Z.-W. Mao, *Inorg. Chem. Front.* **2017**, *4*, 10–32.

[7]   S. M. Haider, S. Neidle, G. N. Parkinson, *Biochimie* **2011**, *93*, 1239–1251.

[8]   S. Neidle, *Curr. Opin. Struct. Biol.* **2009**, *19*, 239–250.

[9]   D. L. Ang, B. W. J. Harper, L. Cubo, O. Mendoza, R. Vilar, J. Aldrich-Wright, *Chem. - A Eur. J.* **2016**, *22*, 2317–2325.

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

[11]   H.-L. Huang, X.-H. Lu, J.-L. Yao, T.-M. Yao, S. Shi, X. Gao, *J. Inorg. Biochem.* **2014**, *140*, 64–71.

[12]   H. J. Yu, L. Yu, Z. F. Hao, Y. Zhao, *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.* **2014**, *124*, 187–193.

[13]   B. Elias, C. Creely, G. W. Doorley, M. M. Feeney, C. Moucheron, A. Kirsch-DeMesmaeker, J. Dyer, D. C. Grills, M. W. George, P. Matousek, et al., *Chem. A Eur. J.* **2008**, *14*, 369–75.

[14]   P. M. Keane, F. E. Poynton, J. P. Hall, I. V. Sazanovich, M. Towrie, T. Gunnlaugsson, S. J. Quinn, C. J. Cardin, J. M. Kelly, *Angew. Chemie - Int. Ed.* **2015**, *54*, 8364–8368.

[15]   J. P. Hall, F. E. Poynton, P. M. Keane, S. P. Gurung, J. A. Brazier, D. J. Cardin, G. Winter, T. Gunnlaugsson, I. V. Sazanovich, M. Towrie, et al., *Nat. Chem.* **2015**, *7*, 961–967.

[16]   S. Takahashi, K. T. Kim, P. Podbevšek, J. Plavec, B. H. Kim, N. Sugimoto, *J. Am. Chem. Soc.* **2018**, *140*, 5774–5783.

[17]   J. P. Hall, H. Beer, K. Buchner, D. J. Cardin, C. J. Cardin,

[18]   K. McQuaid, J. P. Hall, J. A. Brazier, D. J. Cardin, C. J. Cardin, *Chem. - A Eur. J.* **2018**, *24*, 15859–15867.

[19]   S. Asamitsu, S. Obata, Z. Yu, T. Bando, H. Sugiyama, *Molecules* **2019**, *24*, 429.

[20]   E. Wachter, D. Moyá, S. Parkin, E. C. Glazer, *Chem. - A Eur. J.* **2016**, *22*, 550–559.

[21]   P. M. Keane, F. E. Poynton, J. P. Hall, I. P. Clark, I. V. Sazanovich, M. Towrie, T. Gunnlaugsson, S. J. Quinn, C. J. Cardin, J. M. Kelly, *J. Phys. Chem. Lett.* **2015**, *6*, 734–738.

[22]   A. M. Fleming, C. J. Burrows, *Chem. Res. Toxicol.* **2013**, *26*, 593–607.

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

[24]   J. Rubio-Magnieto, S. Kajouj, F. Di Meo, M. Fossépré, P. Trouillas, P. Norman, M. Linares, C. Moucheron, M. Surin, *Chem. - A Eur. J.* **2018**, DOI 10.1002/chem.201802147.

[25]   N. H. Campbell, N. H. A. Karim, G. N. Parkinson, M. Gunaratnam, V. Petrucci, A. K. Todd, R. Vilar, S. Neidle, *J. Med. Chem.* **2012**, *55*, 209–222.

[26]   C. Bazzicalupi, M. Ferraroni, F. Papi, L. Massai, B. Bertrand, L. Messori, P. Gratteri, A. Casini, *Angew. Chemie - Int. Ed.* **2016**, *55*, 4256–4259.

[27]   F. Guarra, T. Marzo, M. Ferraroni, F. Papi, C. Bazzicalupi, P. Gratteri, G. Pescitelli, L. Messori, T. Biver, C. Gabbiani, *Dalt. Trans.* **2018**, *47*, 16132–16138.

[28]   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, et al., *J. Am. Chem. Soc.* **2013**, *135*, 12652–12659.

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

[30]   O. Y. Fedoroff, M. Salazar, H. Han, V. V. Chemeris, S. M. Kerwin, L. H. Hurley, *Biochemistry* **1998**, *37*, 12367–12374.

[31]   M. Dik-Lung, C. M. Che, S. C. Yan, *J. Am. Chem. Soc.* **2009**, *131*, 1835–1846.

Check for updates

Three thymine/adenine binding modes of the ruthenium complex Λ-[Ru(TAP)₂(dppz)]²⁺ to the G-quadruplex forming sequence d(TAGGGTT) shown by X-ray crystallography†

Kane McQuaid, [ID] [ab] James P. Hall, [ID] [bc] Lena Baumgaertner,[a] David J. Cardin[a] and Christine J. Cardin [ID] *[a]

Λ-[Ru(TAP)₂(dppz)]²⁺ was crystallised with the G-quadruplex-forming heptamer d(TAGGGTT). Surprisingly, even though there are four unique binding sites, the complex is not in contact with any G-quartet surface. Two complexes stabilise cavities formed from terminal T·A and T·T mismatched pairs. A third shows kinking by a TAP ligand between T·T linkages, while the fourth shows sandwiching of a dppz ligand between a T·A/T·A quadruplex and a T·T mismatch, stabilised by an additional T·A base pair stacking interaction on a TAP surface. Overall, the structure shows an unexpected affinity for thymine, and suggests models for G-quadruplex loop binding.

Currently there are no structural models for the interactions of monomeric ruthenium polypyridyl complexes with the loop regions of nucleic acid assemblies such as the G-quadruplex and the i-motif. The biological importance of the G-quadruplex has become clear in the last few years,[1] and it has become an important drug target.[2] The DNA at the ends of human chromosomes, in the single stranded telomeric region, has upwards of 2000 repeats of the sequence 5′-AGGGTT-3′, and various versions of the sequence have been widely studied due to the potential therapeutic significance of G-quadruplex forming regions.[2] These single stranded assemblies are often polymorphic in solution and probably for this reason have so far resisted nearly all attempts to crystallise them with metal complexes,[3,4] although NMR has been successfully used to provide elegant binding models for diruthenium complexes.[5] In that work, the binding mode of the diruthenium complex was clearly enantioselective, with only the Λ,Λ-enantiomer able to interact convincingly with the diagonal loop. This solution model is still the only one to define how this important class of metal complexes can interact with a unimolecular G-quadruplex, though strong 'light-switch' effects

have been seen with related complex and known quadruplex-forming sequences.[6,7]

The binding mode of Λ-[Ru(TAP)₂(11-CN-dppz)]²⁺ to duplex DNA was recently described by us and showed that the inclusion of the 11-CN substituent in the dppz ligand resulted in the closing of the T·C/G·A terminal step, in contrast to a wealth of previous observations showing that the T·A base pair was readily flipped out when adjacent to dppz at a terminal step.[8,9] When racemic [Ru(TAP)₂(11-CN-dppz)]²⁺ was crystallised with d(TAGGGTTA), a G-quadruplex assembly was formed in which two Λ-enantiomers were bound at each end of the G-quartet stack.[10] In that work, two Δ-enantiomers (not the stoichiometric equivalent of four) were sandwiched between adjacent quadruplex assemblies and made little interaction with the DNA component, and two of the four 3′ terminal adenine bases were not visible at all due to disorder, leading us to believe that this base was unimportant. The enantiomeric difference shown by this study reinforces the Λ preference previously reported by Thomas et al.[5] The structure also showed the stabilisation of the 5′-syn-guanine residues in the quadruplex assembly, leading to an overall antiparallel conformation. The disorder of the terminal adenine suggested an investigation of the truncated telomeric repeat sequence d(TAGGGTT) might give a more reproducible crystallisation. Unexpectedly, in this work we show for the first time that the parent Λ complex (without the 11-CN dppz substituent) can stabilise a T·T mismatch pair as part of a T·A/T·T cavity, and also for this first time that semi-intercalation (kinking) can be seen between thymine residues, whereas up till now we have only seen such kinking induced between two guanine residues, at a G·G/C·C step.[11] These are structural features which are most readily observed by X-ray crystallography, although very probably detectable in solution experiments and by single molecule approaches.[12,13] The lack of interaction with the parent dppz chromophore strongly supports our previous observation of the surprisingly powerful effect of 11-CN substitution in these Λ enantiomers on DNA binding.

The complex rac-[Ru(TAP)₂(dppz)]²⁺ crystallised with the d(TAGGGTT) sequence and K⁺ ions to give crystals containing

[a] Department of Chemistry, University of Reading, Whiteknights, Reading, RG6 6AD, UK. E-mail: c.j.cardin@reading.ac.uk
[b] Diamond Light Source Ltd., Harwell Science and Innovation Campus, Didcot, Oxfordshire, OX11 0DE, UK
[c] Department of Pharmacy, University of Reading, Whiteknights, Reading, RG6 6AP, UK
† Electronic supplementary information (ESI) available. See DOI: 10.1039/c9cc04316k
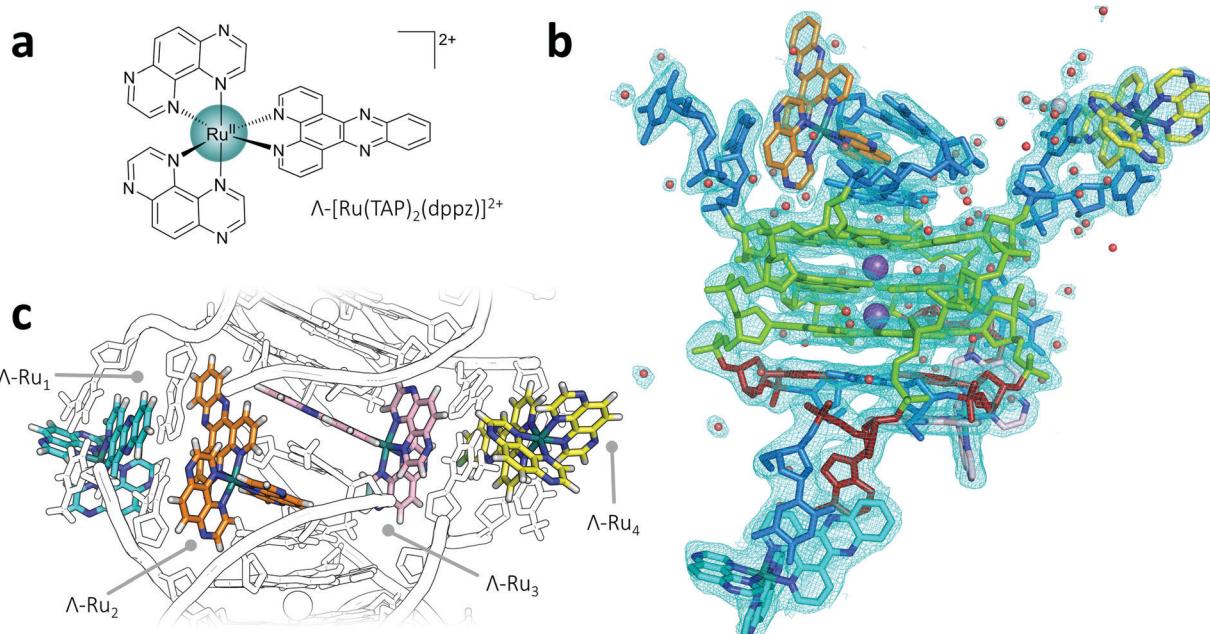
Fig. 1 (a) The Λ-[Ru(TAP)₂(dppz)]²⁺ complex used in this study; (b) overall view of the parallel stranded asymmetric unit of the structure reported here (PDB code 6RNL). Four strands of the sequence d(TAGGGTT) assembled with two K⁺ ions and four crystallographically independent Λ-[Ru(TAP)₂(dppz)]²⁺ cations. Colour code for residues throughout: guanine – green; adenine – red; and thymine – blue. Ruthenium complexes are coloured separately in cyan, orange, pink, and yellow for carbons; teal for ruthenium; and dark blue and white for nitrogen and hydrogen respectively. Alternative views are illustrated in Fig. S1 and S2 (ESI†). The map is contoured at 0.29 e Å⁻³; (c) generation of the four ruthenium complex environments at the interface between two nucleic acid assemblies. The numbering of the four ruthenium complexes corresponds to that used in the text. The kink in the DNA stack is generated by one of the TAP ligands of complex Ru₂, between a T·T mismatch and a water–bridged T·T mismatch.

only the lambda enantiomer (Fig. 1a). The structure was phased using SAD data measured above the Ru absorption edge at 22.26 keV. Data collection and refinement parameters are given in Table S1 (ESI†). The stoichiometric ratio in the resulting crystal is 1:1, giving four complexes per tetrameric assembly. This is the same ratio as in the previous study but giving an entirely different outcome.[10] In our study of [Ru(TAP)₂(11-CN-dppz)]²⁺ crystallised with d(TAGGGTTA), each of the four crystallographically independent lambda complexes had an almost identical nucleic acid environment. In this crystal structure, each complex has a distinct environment within the crystal, and none makes contact with the central G-quadruplex unit (Fig. 1b and c). Here, we observe a parallel-stranded assembly, illustrated schematically in Fig. 2, held together by two K⁺ ions. A Na⁺ ion can also be identified (Fig. 3a and b). What was unexpected is the overall bend (Fig. 1c) introduced into an otherwise parallel stack by a semi-intercalative kinking motif similar to that we have we have previously observed in duplex structures.[9,11,14] The ∼50° kink seen previously was always at GG steps of the sequence. Here, the kink is formed by a T·T mismatch pair and a second pair of thymine bases linked by water bridges (Fig. 4b). The motif generates an overall kink angle of about 28°, as can be seen by looking at the angle generated between the G-quartets of the assembly in Fig. 1c, with a local kink of 34°, measured from the thymine base planes shown in Fig. 3c. The packing diagrams of the structure viewed perpendicular to the long axis show the overall effect of this kinking on the assembly. The asymmetric units are packed
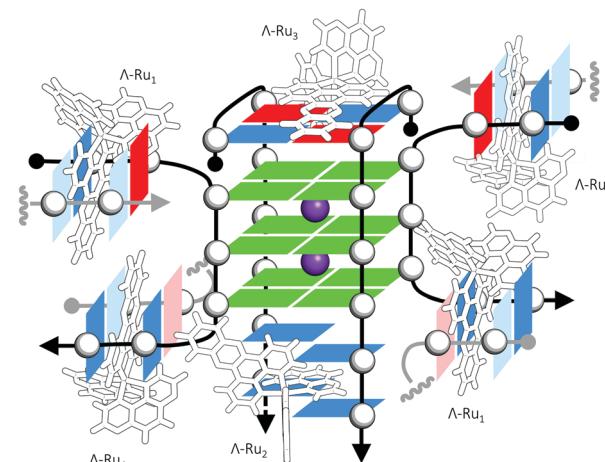

Fig. 2 Structural schematic. The four strands of the sequence d(TAGGGTT) are shown with arrows in the chain direction 5′–3′. The T·A and T·T base pairs formed with bases from symmetry related strands have thymine as pale blue and adenine as pale red. Potassium ions are shown as purple spheres.

together in spirals about the z direction in space group P6₅ (Fig. S4, ESI†), giving head to tail stacking and generating the four ruthenium environments observed at the interface between the units. All four crystallographically independent complexes are bound in thymine-rich environments and hence suggest comparisons with the binding of metal complexes to loop regions in single stranded DNA, as thymine–thymine mismatched base pairs are situated adjacent to, and possibly stabilised by, all of

Chem. Commun.
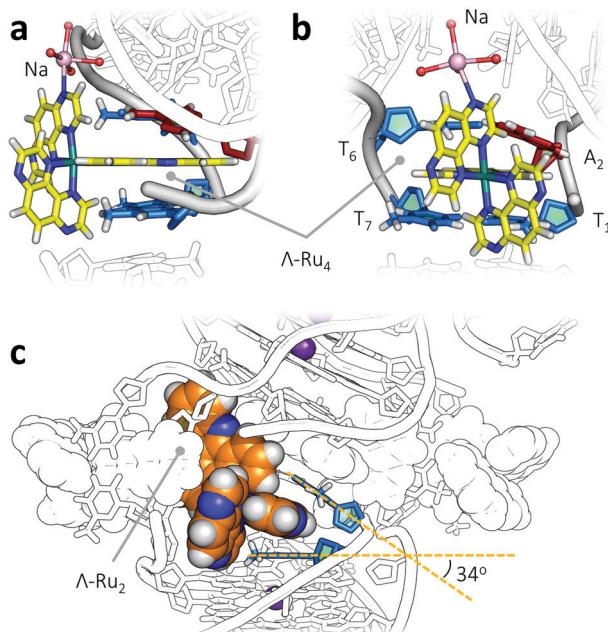
Fig. 3 Structural details of TAP ligands. (a) and (b) coordination of a Na$^+$ ion (pink) to a TAP ligand of Ru$_4$; and (c) the local 34° kink introduced by a TAP ligand of Ru$_2$. Potassium ions are shown as purple spheres, sodium in pink, and coordinated waters as red spheres.

the complexes, and stacking with both the TAP and the dppz ligands. For clarity, each will be described separately.

Two of the four complexes (Ru$_1$ and Ru$_4$) have almost the same environments, at the ends of the overall assembly. T$_1$ and A$_2$ of one strand pair with T$_6$ and T$_7$ of another strand, as shown in the schematic of Fig. 2, to generate an intercalation cavity created by a standard A·T base pair and a T·T mismatch pair (Fig. 4a). These binding sites also provide a model for what would be major groove binding in duplex DNA. The sites are distinguished by coordination of a Na$^+$ ion to one TAP ligand of Ru$_4$ only. The ion is directly coordinated to one of the TAP

ligands, and through a water bridge to a TAP ligand of Ru$_3$. The TAP ligands in these structures have previously been observed to be hydrated,[9] and this provides a convenient rationale for the sometimes relative ease of crystallisation when compared to the more hydrophobic but isosteric and isoelectronic phen analogue, the well-known 'light-switch' complex [Ru(phen)$_2$(dppz)]$^{2+}$.[15,16] In the overall assembly the Ru$_1$ and Ru$_4$ cavities are end-stacked on each other, generating a quasi-continuous stack running orthogonal to the main helix axis direction, which corresponds to the b axial direction in Fig. S4 (ESI†). The additional charge neutralisation by Na$^+$ is possibly an additional stabilising factor for this assembly. This monovalent ion coordination may also account for the asymmetry introduced by the differing environments of Ru$_2$ and Ru$_3$, since there is no corresponding ion linking Ru$_1$ and Ru$_2$.

The environments of the two central ruthenium complexes, Ru$_2$ and Ru$_3$, are distinctly different, thus generating the overall lack of quasi-twofold symmetry in this structure. Ru$_3$ appears almost completely surrounded by T·A and T·T base pairs, and the two faces of the complex are shown in Fig. 4c and d. One dppz face contacts a T·A/T·A quartet surface formed by two T$_1$/A$_2$ ends, shown in the same figure. The other dppz surface contacts a T·T mismatch formed from two T$_7$ residues, shown in Fig. 4d. A further A$_2$·T$_6$ base pair contacts the TAP ligand, almost completely encapsulating it, shown in Fig. 4c.

The environment of Ru$_2$ is perhaps the most unexpected and as already stated generates the kink in the overall $P6_5$ packing shown in Fig. 3c and Fig. S4 (ESI†). The kink is generated at one of the TAP ligands, with a T$_7$·T$_7$ mismatch on one side of the TAP and two thymine residues, T$_6$ and T$_6$, with two water bridges on the other side of the TAP (Fig. 4b). Unexpectedly, the dppz is free, so that this complex is only held in place by this kinking interaction. This kinking site shows a remarkable overall resemblance to that seen in the original Λ-[Ru(TAP)$_2$(dppz)]$^{2+}$ structure, with the DNA duplex sequence d(TCGGCGCCGA) and in many structurally isomorphous examples since then.[9,11] In dilute solutions of B-DNA the thermodynamic binding
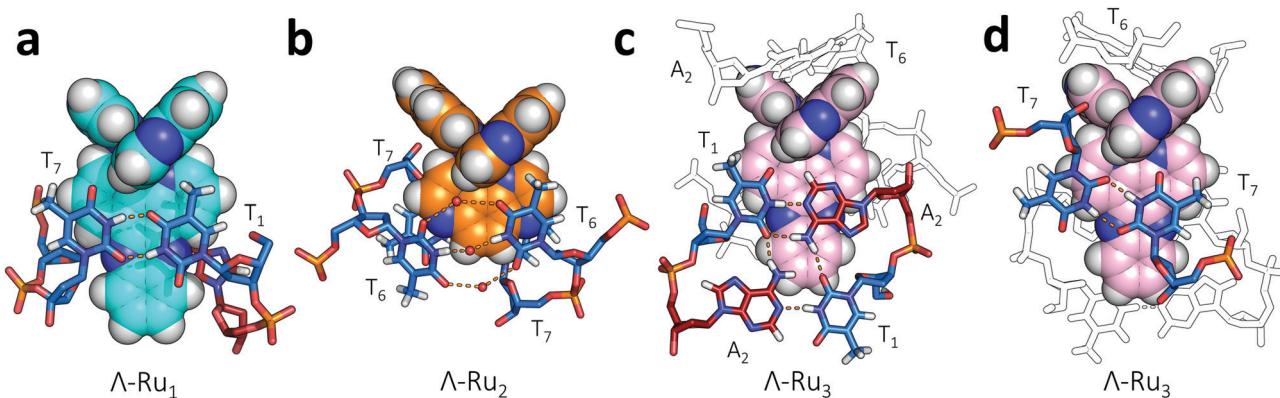

Fig. 4 The ruthenium complex environments: (a) T·T mismatch and A·T match form the cavity for Ru$_1$ and Ru$_4$ similarly; (b) kinking (semi-intercalation) at Ru$_2$, also generated by T·T mismatching. This complex is not bound by intercalation of the dppz chromophore, but just by the interaction of one of the TAP ligands as shown here; (b) and (c) illustrate the T·T mismatch cavity and additional T stacking and hydrogen bonding around Ru$_3$. Note that all ruthenium complex environments feature T·T mismatched base pairs. Coordinated waters are shown as red spheres. For clarity, interacting residues from neighbouring units are not coloured.

constants show a relatively weak interaction compared with dppz intercalation.[9] In crystals and in other tightly packed environments, a combination of weak interactions can lead to environments which could not be predicted from any solution study, and what we are seeing here is perhaps a model for such cases.

Both T·T mismatched base pairs and kinking by phen and TAP ligands may be important components of the binding of ruthenium polypyridyl complexes to higher-order DNA structures containing loop regions. Octahedral complexes have an inherently greater potential for specific interactions than square planar ones but not much is known about their binding modes.[17] So far the only structural evidence is provided by the work of the Thomas group as already stated.[5] In that work (in which the ancillary ligands were bpy) the diruthenium cation threaded through a diagonal loop, with the principal interactions being with the central thymine residue of the loop. We have previously shown that, of the well-known ancillary ligands in these systems, bpy is less likely to cause kinking and stacking than either phen or TAP.[18] More recently the TAP analogue of this diruthenium compound has been shown to have a range of useful properties in cell systems.[19] It was studied as an enantiomeric mixture and gave spectroscopic results clearly indicative of several binding modes. The specificity of these complexes does not just arise from end-stacking to the G-quadruplex chromophore but plausibly also includes the sort of thymine interactions revealed by the present work. There are several examples of ruthenium polypyridyl complexes which are luminescent when bound to what may well be thymine–adenine loop regions of G-quadruplexes, but there is no structural data for any of these. The binding modes seen in this work, which have no counterpart in duplex DNA, and would not be predictable from any modelling calculation, provide a useful springboard for understanding luminescence and other spectroscopic behaviour. Each of the binding modes shown in Fig. 4 would have different luminescence behaviour if it were the phen analogue, based on our previous work. Ru_3, being almost enclosed, would be most luminescent, with Ru_1 and Ru_4 expected to be similar, and Ru_2 the most exposed to quenching of the excited state *via* non-radiative pathways. A previous paper from our laboratory has considered the delta enantiomer/duplex DNA case in detail.[14]

In future we aim to provide a comparable interpretation of the binding of lambda complexes to G-quadruplex loop regions. We would also like to understand the crucial factor which determines whether the G-quadruplex is parallel, or antiparallel as in ref. 10. It is not clear how much of the switch can be ascribed to the modification of the dppz ligand and how much the crystallisation is sequence dependent.

Conflicts of interest

There are no conflicts to declare.

References

1 G. Biffi, D. Tannahill, J. McCafferty and S. Balasubramanian, *Nat. Chem.*, 2013, **5**, 182–186.
2 S. Asamitsu, S. Obata, Z. Yu, T. Bando and H. Sugiyama, *Molecules*, 2019, 24.
3 N. H. Campbell, N. H. A. Karim, G. N. Parkinson, M. Gunaratnam, V. Petrucci, A. K. Todd, R. Vilar and S. Neidle, *J. Med. Chem.*, 2012, **55**, 209–222.
4 C. Bazzicalupi, M. Ferraroni, F. Papi, L. Massai, B. Bertrand, L. Messori, P. Gratteri and A. Casini, *Angew. Chem., Int. Ed.*, 2016, **55**, 4256–4259.
5 T. Wilson, P. J. Costa, V. Félix, M. P. Williamson and J. A. Thomas, *J. Med. Chem.*, 2013, **56**, 8674–8683.
6 L. Xu, X. Chen, J. Wu, J. Wang, L. Ji and H. Chao, *Chem. – Eur. J.*, 2014, **21**, 4008–4020.
7 Q. Yu, Y. Liu, C. Wang, D. Sun, X. Yang, Y. Liu and J. Liu, *PLoS One*, 2012, 7(12), e50902.
8 K. McQuaid, J. P. Hall, J. A. Brazier, D. J. Cardin and C. J. Cardin, *Chem. – Eur. J.*, 2018, **24**, 15859–15867.
9 C. J. Cardin, J. M. Kelly and S. J. Quinn, *Chem. Sci.*, 2017, **8**, 4705–4723.
10 K. McQuaid, H. Abell, S. P. Gurung, D. Allan, G. Winter, T. Sorensen, D. J. Cardin, J. A. Brazier, C. J. Cardin and J. P. Hall, *Angew. Chem., Int. Ed.*, 2019, **58**, 9881–9885.
11 J. Hall, K. O'Sullivan, A. Naseer, J. Smith, J. Kelly and C. J. Cardin, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 17610–17614.
12 P. Lincoln and B. Nordén, *J. Phys. Chem. B*, 1998, **102**, 9583–9594.
13 W. Vanderlinden, P. J. Kolbeck, W. Frederickx, S. F. Konrad, T. Nicolaus, C. Lampe, A. S. Urban, C. Moucheron and J. Lipfert, *Chem. Commun.*, 2019, DOI: 10.1039/C9CC02838B.
14 J. P. Hall, P. M. Keane, H. Beer, K. Buchner, G. Winter, T. L. Sorensen, D. J. Cardin, J. A. Brazier and C. J. Cardin, *Nucleic Acids Res.*, 2016, **44**, 9472–9482.
15 J. P. Hall, D. Cook, S. R. Morte, P. McIntyre, K. Buchner, H. Beer, D. J. Cardin, J. A. Brazier, G. Winter and J. M. Kelly, *et al.*, *J. Am. Chem. Soc.*, 2013, **135**, 12652–12659.
16 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.
17 Q. Cao, Y. Li, E. Freisinger, P. Z. Qin, R. K. O. Sigel and Z.-W. Mao, *Inorg. Chem. Front.*, 2017, **4**, 10–32.
18 J. P. Hall, S. P. Gurung, J. Henle, P. Poidl, J. Andersson, P. Lincoln, G. Winter, T. Sorensen, D. J. Cardin, J. A. Brazier and C. J. Cardin, *Chem. – Eur. J.*, 2017, **23**, 4981–4985.
19 S. A. Archer, A. Raza, F. Dröge, C. Robertson, A. J. Auty, D. Chekulaev, J. A. Weinstein, T. Keane, A. J. H. M. Meijer, J. W. Haycock, S. MacNeil and J. A. Thomas, *Chem. Sci.*, 2019, **2**, 3502–3513.