Supramolecular Assembly and Metal-Coordination in G-quadruplex Structures

A thesis presented in partial fulfilment
of the requirements for the degree of

Doctor of Philosophy
in
Chemistry

Massey University, Manawatu, New Zealand

Jamie Michael Withers

2014

Abstract

G-quadruplexes are supramolecular structures formed from the self association of guanine-rich oligonucleotides. They consist of stacks of G-quartets, each of which is made of four guanine bases arranged in a cyclic, planar array and held together through hydrogen bonding. Recent evidence has suggested G-quadruplexes have a biological role. Telomeric DNA, which has been linked to cancer formation and ageing, is capable of forming a variety of G-quadruplex structures. This observation has led to the development of a number of drug candidates that are currently undergoing clinical trials. Although G-quadruplex structure can be controlled to a certain extent by the careful selection of oligonucleotide sequence, this strategy alone is not always sufficient to allow predictable assembly to occur. To address this problem, we investigated metal-coordinating nucleotides as a method of controlling the predictable assembly of G-quadruplexes. Our research was focused on the use of 1,2,4-triazole and cytosine nucleotides in the formation of base pairs and base quartets; their ability to coordinate transition metal ions was examined.

We investigated a variety of intramolecular and intermolecular G-quadruplex assemblies modified with triazole nucleotides, each providing a different environment for metal coordination. We found that 1,2,4-triazole nucleotides were not capable of forming either base pairs or quartets. We were not able to find evidence of a direct interaction between triazole nucleotides and metal-ions under the conditions studied.

The stabilisation of a DNA duplex by coordination of silver(I) ions to cytosine nucleotides was observed. However, this interaction was not useful in directing the assembly of a G-quadruplex structure. In most cases we found that the presence of transition metal ions had a negative impact on the stability of G-quadruplex structures, and that destabilisation occurred before metal coordination could be observed.

Our investigations into the metal-coordinating properties of a 3+1 G-quadruplex based on the 3*htel* sequence have highlighted the significance of the Watson-Crick base pair formed in this assembly. In the absence of this base pair, the G-quadruplex is completely destabilised, however, evidence suggests the formation of a G-triplex instead.

Although we did not observe G-quadruplex assembly directed by metal-coordinating nucleotides, the contents of this thesis may be considered as a basis for the further development of supramolecular oligonucleotide assemblies.

Acknowledgements

I would like to thank my supervisors, Vyacheslav Filichev and Shane Telfer, for their support, assistance and enthusiasm throughout this project.

I am grateful for the assistance provided by Pat Edwards whose assistance with NMR spectroscopy has been invaluable. I would also like to thank my friends, and colleagues, Osman Doluca and Adam Stephenson.

Contents

1 Introduction

DNA is capable of forming a variety of different supramolecular assemblies. The DNA duplex, for example, displays all the hallmark properties of a supramolecular construct, i.e. directed self-assembly of discrete subunits into a well-defined array through non-covalent bonding. Although the duplex is the most widely known, DNA forms structures as diverse as triplexes, quadruplexes and i-motifs.[1] In all cases DNA retains the same fundamental structure: a polymer of different nucleotide subunits which only differ in the identity of the attached nucleobase. Whereas most supramolecular structures are assembled from a large number of separate subunits which, collectively, assume one conformation, DNA assemblies are typically formed from a small number of individual DNA strands which can be directed into a number of different structures by selecting an appropriate sequence of nucleotides. The ease with which an oligonucleotide can be directed into a particular structure, and the diversity of known DNA structures, makes DNA a flexible platform to study supramolecular assembly. DNA assembly is primarily mediated through hydrogen bonding and the hydrophobic effect. However, these interactions do not offer a great deal of dynamic control. The goal of this thesis is to explore the use of metal-ligand coordination for the control of DNA assemblies with particular focus on G-quadruplexes.

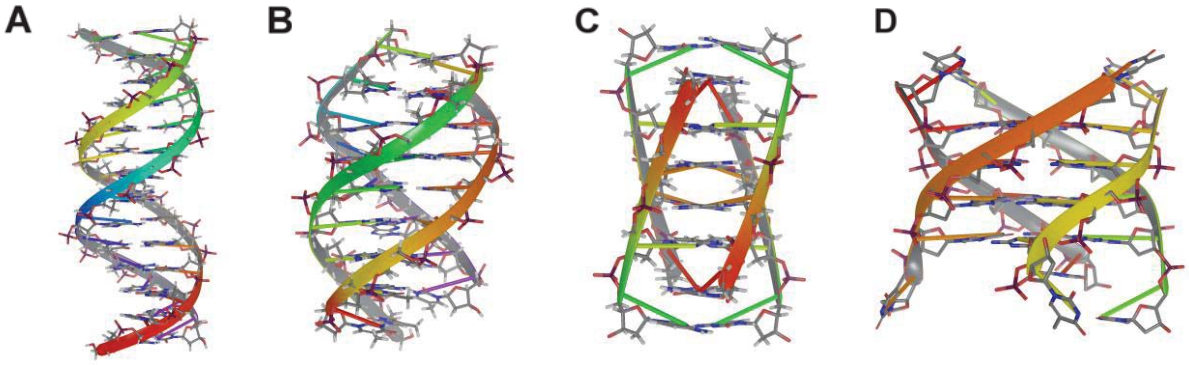


Figure 1.1 Several nucleic acid structures. Duplex, A. Triplex, B. i-Motif, C. G-quadruplex, D.

1.1 Deoxyribonucleic Acid

DNA, a type of nucleic acid, is a biopolymer that consists of long chains of nucleotide subunits, each of which consists of a phosphate group and a 2′-deoxyribose sugar with an attached nucleobase; the sugar rings and phosphates alternate along the backbone of the DNA strand, the bases, attached through the 1′ position of the sugar, are oriented perpendicular to the backbone.[2] In biological systems, DNA is composed of four types of nucleobases: adenine, guanine, cytosine and thymine. Each base possesses a unique arrangement of hydrogen-bond donors and acceptors. It is these arrangements that lead DNA to self-associate into a wide variety of supramolecular structures. The type of

conformation a given oligonucleotide will assume is largely dictated by its nucleobase content. Several other nucleic acids exist: ribonucleic acid (RNA) is naturally occurring while examples of artificial nucleic acids include peptide nucleic acids (PNA),[3] locked nucleic acids (LNA),[4] glycol nucleic acids (GNA)[5] and the L-enantiomer of DNA (L-DNA),[6] see Figure 1.2. Most artificial nucleic acids substitute the sugar ring for another moiety, although changes to the phosphate group have been explored.[3,7-10]

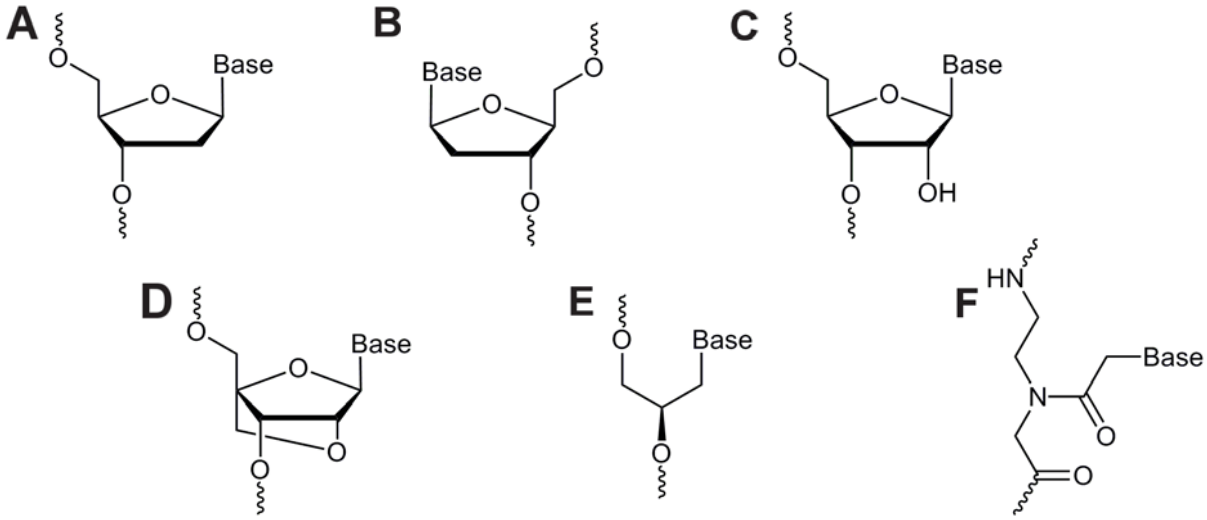


Figure 1.2. 2′-deoxy-D-ribosenucleic acid (DNA), A. 2′-deoxy-L-ribosenucleic acid (L-DNA), B. Ribonucleic acid (RNA), C. Locked nucleic acid (LNA), D. Glycol nucleic acid (GNA), E. Peptide nucleic acid (PNA), F.

DNA is typically thought of in a biological context but there is growing interest in the use of DNA in nanotechnology and supramolecular chemistry where its regular structure, diverse H-bonding abilities and well-established chemistry make it suitable for a 'bottom-up' approach to macromolecular assembly, in which simple building blocks are used to create larger complexes.[11,12] The basic building blocks of DNA structure could be considered the motifs created from the hydrogen bonding of nucleobases; the behaviour of these motifs can be described by the polygon self-assembly principle, described below. The hydrogen bonding pattern of the base pairs A/T and G/C are well known, producing a linear subunit that gives rise to a duplex structure. The slight twist between stacked base pairs results in the characteristic helix. The H-bond donors and acceptors in the major groove of duplex DNA allow a third strand to bind, resulting in a three base subunit with approximately three-fold symmetry which forms the basis of a DNA triplex. i-Motifs and G-quadruplexes both form structures with four-fold symmetry through different means. i-Motifs form base pairs, Figure 1.3A, similar to those found in duplex DNA. However, base pairs in i-motifs stack perpendicular to their neighbours. The result is an assembly with four-fold symmetry that resembles two intermeshed duplexes. G-quadruplexes are composed of stacks of G-quartets, each of which consists of four guanine bases hydrogen bonded in a planar, cyclic arrangement. The production of synthetic nucleobases with different geometries and

bonding properties, or the use of natural nucleobases in new ways, should allow the exploration of novel DNA structures.

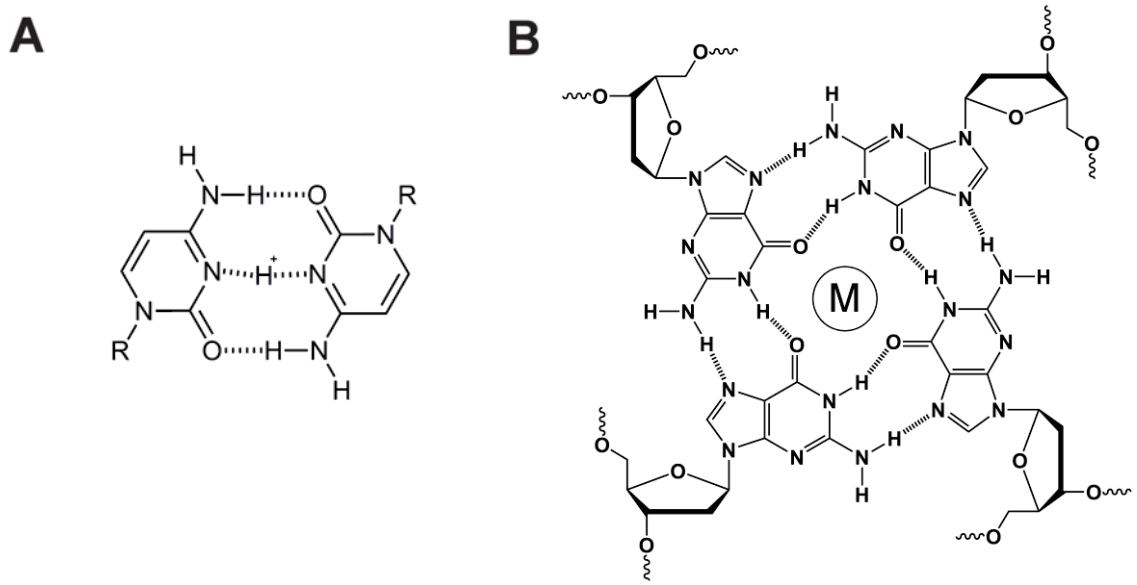


Figure 1.3. A protonated cytosine/cytosine base pair found in i-motifs, A. The hydrogen bonding patterns of a G-quartet. The central cavity between guanines is occupied by a metal cation, typically Na^+ or K^+, which reduces electrostatic repulsion between guanine O6 atoms that line the cavity, B.

1.1.1 Duplex DNA

A number of different types of DNA duplex exist. The most familiar form, B-DNA, possesses two oligonucleotide strands orientated in an anti-parallel fashion. The DNA strands assume a distinctive, right-handed double-helix conformation surrounding a central core of base pairs.[2] The helical shape is the result of electrostatic repulsion and the hydrophobic effect where the twist of the helix maximises contact between base pairs and increases the distance between negatively charged phosphate groups. The exclusion of water is the primary driving force with Van der Waals contacts between the aromatic surfaces of the base pairs playing only a small role.[13,14] There is a 36 ° twist and 3.4 Å of space between adjacent base pairs, thus a full turn of a B-helix contains 10 base pairs and is 34 Å in length. Both sets of base pairs, A/T and G/C (Figure 1.4), are identical in length. This means that the double helix has a constant width of 23.7 Å. A large number of other duplex DNA forms exist. A brief description follows of the most common alternative duplexes, A-DNA and Z-DNA. A-DNA is formed in dehydrated DNA samples where a lack of stabilising water causes changes in structure, most significantly in the twist between base pairs– only 33.6 ° – resulting in a more compact helix with ~11 base pairs per turn.[15]

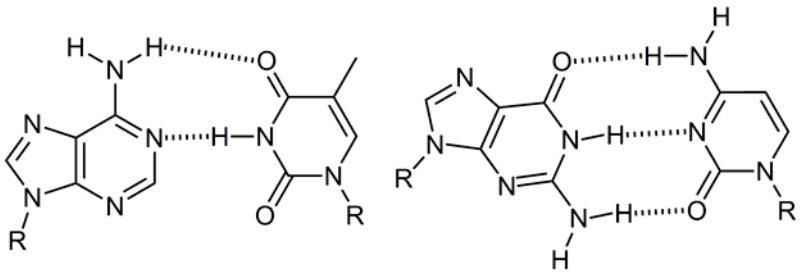


Figure 1.4. Base pairs of DNA. Adenine/Thymine, left. Guanine/Cytosine, right.

In contrast to the right-handed helix of B-type DNA, Z-DNA forms a left-handed helix in which the base pairs are located nearly perpendicular to the phosphate backbone.[16,17] The formation of Z-DNA is predominately limited to alternating purine and pyrimidine sequences, $d(CG)_n$ is the most favoured sequence. High ionic strength, negative supercoiling, and protein binding can convert B-DNA into Z-DNA.[18] Z-DNA is characterised by the zigzag pattern of its phosphate backbone, in which negatively charged phosphates are closer to each other than in B-DNA (8 Å as opposed to 11.7 Å). This means that increased salt concentration promotes the formation of Z-DNA as a result of decreased electrostatic repulsion. For the most extensively studied Z-DNA sequence, d(CG)n, guanosine assumes a conformation with guanine in a syn position (glycosidic angles close to 60°), Figure 1.6. Conversely, cytosine assumes anti conformations. The stacking of the bases in B- and Z-DNA are different: discrete clusters of interstrand four-base π stacks of G–C/C–G exist in Z-DNA, while there is continuous stacking along each strand in B-DNA. The net result of these differences is that the minor groove in Z-DNA is very deep and narrow and the major groove is a convex surface where cytosine C5 and guanine N7 and C8 are exposed to the environment. The biological functions of Z-DNA and related Z-DNA binding proteins are not fully understood. A body of evidence is accumulating that suggests that Z-DNA can be involved in the regulation of gene expression particularly in actively transcribed regions and potentially enhances the frequencies of recombination, deletion, and translocation events in cells.[18]

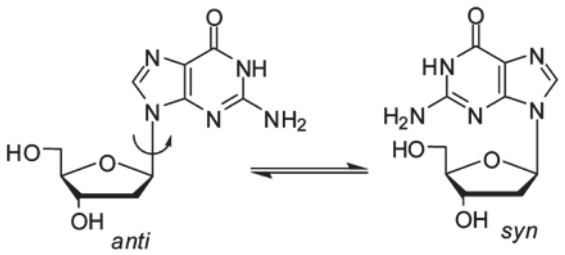


Figure 1.5. Syn and anti conformations of 2′-deoxyguanosine.

1.1.2 Triplex DNA

Triple helices are formed when a single-stranded triplex forming oligonucleotide (TFO) binds in the major groove of double-stranded DNA (dsDNA).[19-21] Depending on the orientation of the TFO relative to the homopurine strand of the dsDNA, triplexes can be classified as either parallel or antiparallel. In parallel triplexes a homopyrimidine TFO will bind to dsDNA (a thymine from the TFO will bind to an adenine in the duplex, but cytosine bases in the TFO must be protonated at the N3 atom to form Hoogsteen base-pairing with guanines in dsDNA), see Figure 1.6. This makes the formation of parallel triplexes at neutral pH problematic. The vast majority of studies on DNA triple helices have been focused on the improvement of pH-sensitive parallel triplexes. Antiparallel TFOs, in contrast, feature reverse Hoogsteen GT/GA-TFOs and do not require acidic pH. In these triplexes, adenine binds to an A–T base pair and a guanine to a G–C pair. GT-rich TFOs can form both parallel and antiparallel triplexes depending on the sequence, Figure 1.6.[22] However, applications of G-rich TFOs are limited because of their tendency to form highly stable aggregates such as G-quadruplexes in preference to triplexes. This behavior creates uncertainty as biological effects may be due to the self aggregation of TFOs into G-quadruplexes, rather than due to the formation of triplexes.

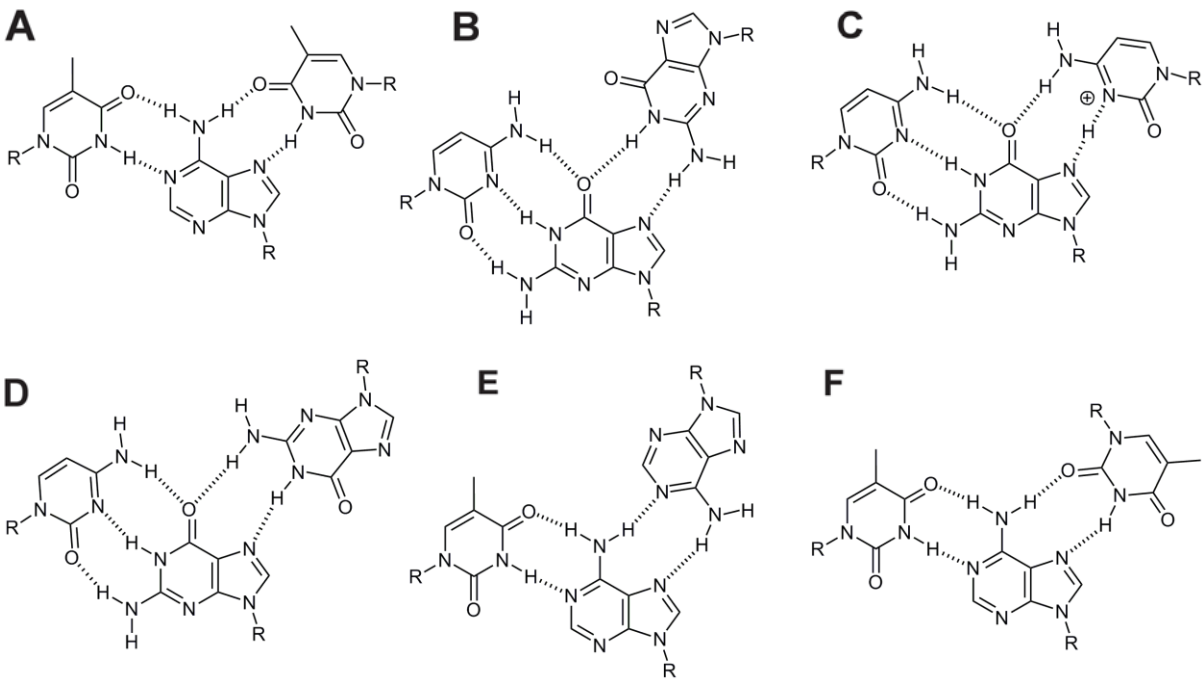


Figure 1.6. Hydrogen bonding patterns of bases in DNA triplexes. A,B and C occur in parallel triplexes. D, E and F occur in antiparallel triplexes.

1.1.3 G-quadruplex DNA

Four guanines can self-assemble through the formation of eight Hoogsteen hydrogen bonds and form a guanine- or G-quartet, Figure 1.3B. G-quartets can stack with each other and form a G-quadruplex. The central cavity of G-quadruplexes is occupied by cations which neutralise the electrostatic repulsion between guanine O6 oxygens and thus stabilise the overall structure. G-quadruplexes can be *inter*molecular or *intra*molecular and can display a wide variety of topologies based on variations in strand polarity, loop length and geometry, the presence of metal ions, the oligonucleotide concentration in solution, etc. In parallel G-quadruplexes each G-tract is aligned in the same direction, 5′ – 3′. If one or two G-tracts are oriented in opposition to the others, then a G-quadruplex is classified as antiparallel. The sections of oligonucleotide connecting the G-tracts of multimeric quadruplexes are divided into lateral, diagonal or propeller loops, see Figure 1.7.

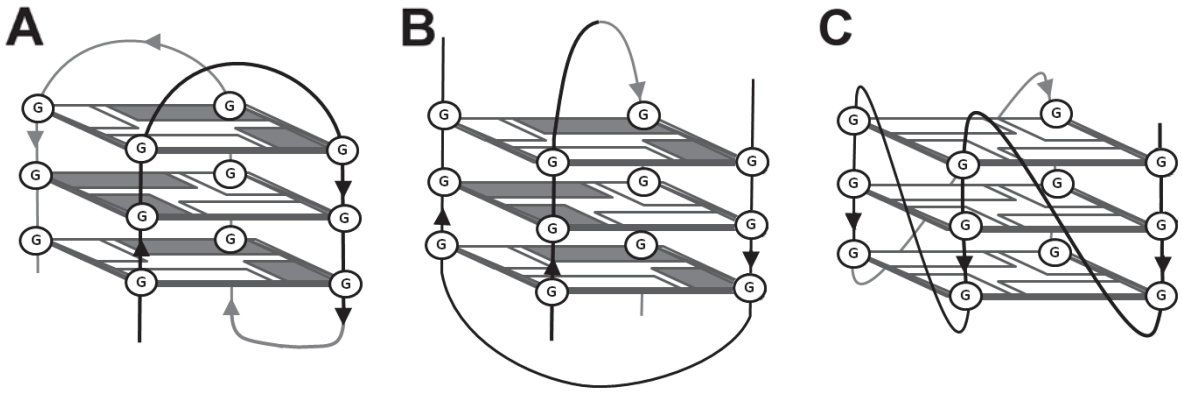


Figure 1.7. Antiparallel G-quadruplex featuring three lateral loops, A. Antiparallel G-quadruplex featuring two diagonal loops, B. Parallel G-quadruplex featuring three propeller loops, C.

The influence of nucleotide sequences of different lengths has been studied most extensively for G-quadruplexes due to the potential biological relevance of G-rich sequences in the human genome. The insertion of a single (non-G) nucleotide into G-tracts (3–6 guanines) generally disrupts G-quadruplex assembly in Na^+ containing conditions.[23] However, self-assembled structures, which are not necessarily G-quadruplexes, can still be detected in the presence of K^+ ions, especially at elevated concentrations of the cation (≥100 mM).[24] Increasing the length of the loops between G-tracts decreases the thermal stability of the resulting G-quadruplexes. This, however, may be associated with a topological change in the complex.[25] A single nucleotide can form a loop to bridge a quadruplex stem and can also be accommodated as a bulge between G-quartets.[26] For RNA GROs the topology of G-quadruplexes seems to be less polymorphic than for DNA G-quadruplexes: the folding topology is predominantly parallel, and the loop length only slightly affects the thermodynamic stability. The improved stability of RNA over DNA quadruplexes is mediated by the increased H-bonding afforded by the 2′ hydroxyl group of ribose.[27] Stable multimeric

RNA G-quadruplexes have also been detected using gel electrophoresis and mass spectrometry.[28]

1.2 Supramolecular Chemistry

Supramolecular chemistry is an area of research covering a broad range of topics from metal-organic frameworks, rotaxanes and catenanes to ion-pair receptors and biological macromolecules. It has been described as the chemistry of molecular assemblies and of the intermolecular bond;[29] from this perspective, supramolecular chemistry is concerned with the use of non-covalent interactions such hydrogen bonds, Van der Waals forces, π-π interactions and metal-ligand bonds. The following paragraphs describe aspects of supramolecular chemistry that are most relevant to the design of DNA assemblies.

1.2.1 Self-Assembly

Self-assembly refers to the spontaneous arrangement of molecules into larger, ordered complexes, usually driven by non-covalent interactions. It is the goal of the supramolecular chemist to understand self-assembly processes and use them to direct the formation of assemblies in a predictable and reproducible manner. Because the assembly of subunits into a complex takes place without human intervention (a 'bottom up' approach) all the information required to create the complex must be contained within each of the subunits themselves. This information is stored in the shape of the subunits and their composition – that is, the location and directionality of H-bond donor/acceptors, hydrophobic or aromatic surfaces, dipoles, ionic charges or other moieties that participate in non-covalent interactions. These non-covalent interactions result in reversible binding between subunits – a property that allows an assembly to correct incorrectly bound subunits and ultimately leads an assembly to assume the thermodynamically most stable conformation. In the absence of reversible binding, or in situations where the lability of components is low, assemblies will accumulate improperly aligned subunits, leading to a reduced overall stability and malformed structures. In DNA, the information required to direct an oligonucleotide to assemble, for example, into a duplex rather than a G-quadruplex is contained within the sequence of bases rather than the bases themselves – nucleobases are flexible in their H-bonding arrangements, but their incorporation into an oligonucleotide restricts the number of available H-bonding conformations and thereby determines the most thermodynamically stable structure. It is the number of nucleobases, and the relative alignment that H-bonding holds them in, that gives rise to the overall structure of a DNA assembly – this behavior is described by the concept of polygonal libraries.

1.2.2 Polygonal Libraries

An integral part of directed self-assembly and supramolecular design is the concept of polygon libraries.[30] Given the specific position of the groups on a molecule that are capable of binding another molecule, it is possible to predict the structure of the assembly formed from these components. It is often helpful to think of individual components as geometric building blocks, as in Figure 1.8. Predictable assembly is possible because, given the reversibility of binding of subunits, the tendency towards maximum entropy will lead to the maximum number of individual particles or assemblies in solution – favouring discrete assemblies over polymers or larger assemblies. Thermodynamics direct subunits to move towards the maximum occupancy of binding sites. It is a balance between thermodynamics and entropy that determine the final structure of an assembly – this can often lead to some uncertainty in predicting supramolecular assembly; using metal-organic-frameworks (MOFs) as an example, thermodynamics lead to maximum occupancy of all binding sites ultimately creating large single crystals where entropy would favour the creation of large numbers of individual assemblies (this example is simplified, and ignores any contribution towards increased entropy of solvent liberation and release of coordinated ions that frequently occurs in the creation of new metal-ligand bonds).

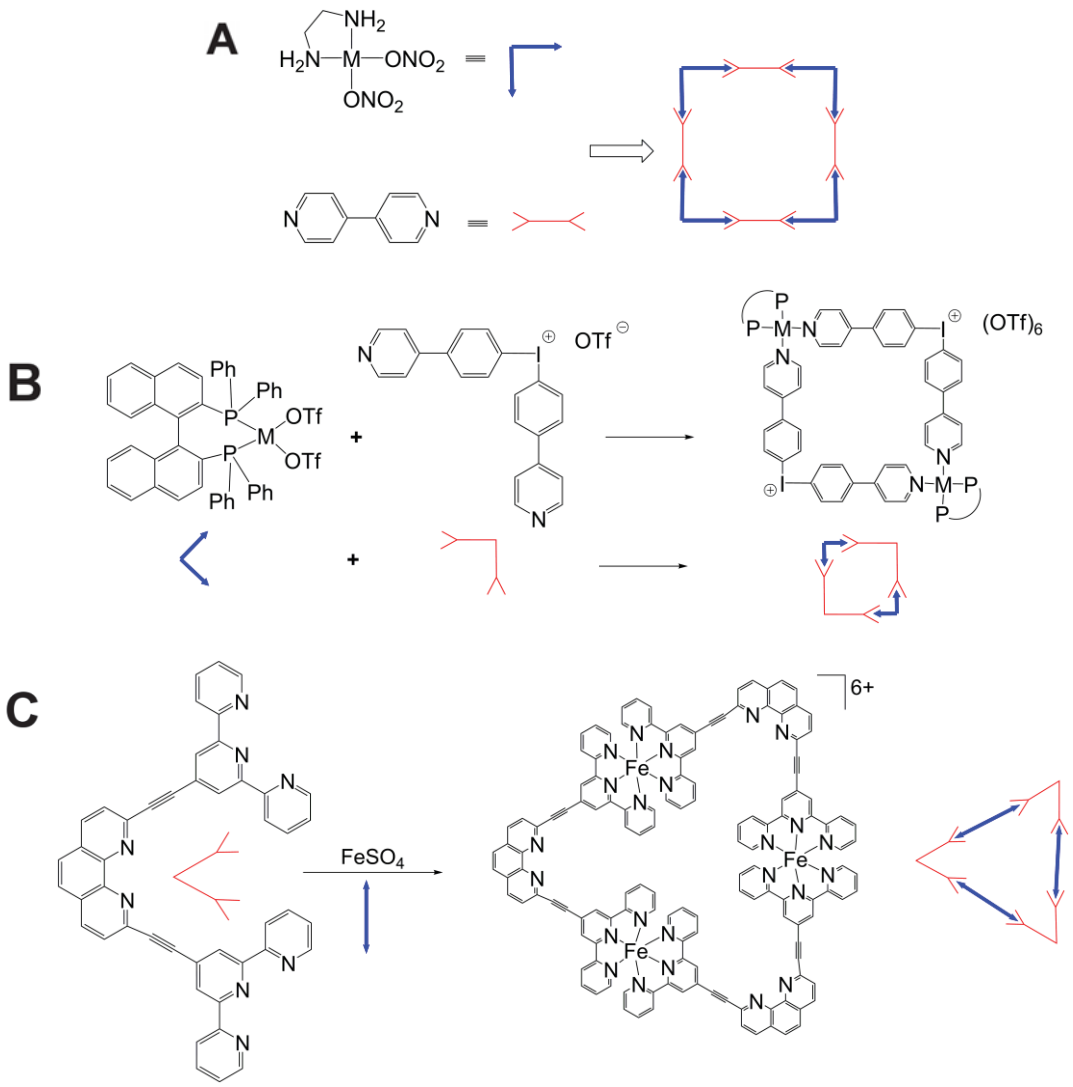


Figure 1.8. Supramolecular square formed from complementary linear and 90° components, A.[31] Supramolecular square formed from two 90° components, B.[32] Supramolecular triangle assembled from linear and 60° components, C.[33]

The concept of a polygonal library can be extended to describe DNA assemblies and is useful in explaining tertiary conformation in terms of the H-bonding arrangement of nucleobases. The DNA double helix has an approximately two-fold axis of symmetry (C2) which is a direct consequence of the approximately linear H-bonding arrangement of the nucleobases in base pairs. The three-fold symmetry of base triplets gives rise to DNA triplexes with ~C3 symmetry, and the four-fold symmetry of G-quartets results in a G-quadruplex with C4 symmetry (see Figure 1.9).

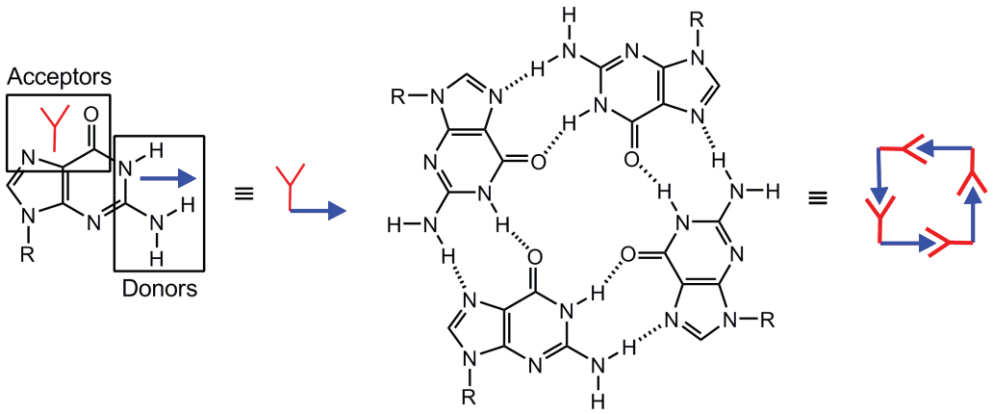


Figure 1.9. Representation of G-quartets as a supramolecular square. An individual guanine base possesses hydrogen bond donors and acceptors that orient at approximately right angles to each other, directly leading to square G-quartet assemblies.

1.2.3 Molecular Recognition

The definition of supramolecular chemistry as the study of intermolecular bonds necessitates the interaction between two or more components; this may include the unstructured aggregation of molecules by the hydrophobic effect or forces such as Van der Waals interactions, but the non-directional nature of these forces means that, in practise, they are of little use in the design of self-assembling components. The directed assembly of subunits into complex structures requires the individual components to have some manner of discriminating not just the identity, but also the orientation of other components. This property is known as molecular recognition and is mediated primarily by the directional non-covalent forces such as H-bonding, dipole-dipole interactions, π-π interactions and coordinate bonding. In supramolecular complexes composed of DNA, the primary interactions are H-bonds between nucleobases where the specific position and directionality of the hydrogen bond donors and acceptors provides the basis for molecular recognition. In the case of duplex DNA, each nucleobase is capable of discriminating with high selectivity between several potential partners in several different orientations. It should be mentioned that non-directional forces, specifically the hydrophobic effect and π-π interactions provide much of the stability of DNA complexes and, in fact, due to the partial ordering of nucleobases, even in single stranded DNA, these forces provide partial preorganisation of DNA – presenting nucleobases to the solution in such a way as to maximise the likelihood of favourable molecular recognition between bases.

1.2.4 Metal-Ligand Bonds

Although they are of little relevance to standard DNA assemblies, metal-ligand interactions are of particular interest to this project. Most non-covalent interactions provide very little additional stability; instead, it is the accumulation of large numbers of small contributions that determine the overall stability of an assembly. While these forces have obvious benefits

in terms of reversibility and self-assembly, as described above, these interactions are difficult to control dynamically. The key variables are pH, solvent polarity and temperature. The strength of metal-ligand coordinate bonds can rival that of covalent bonds. This project was intended to make use of the strength of this kind of bonding to spontaneously assemble borderline-stable complexes upon the addition of a suitable metal-ion. As mentioned above, supramolecular systems making use of metal-ligand bonding depend on high lability metal ions, that is, metal ions that have a high rate of ligand exchange. High lability allows the reversibility needed for a system to converge on a properly assembled structure. In contrast, a system with low lability is more likely to be 'locked' into a misfolded structure, requiring more time to converge on the lowest energy structure.

When designing a supramolecular assembly, it is important to consider that although metal ions typically have a preferred coordination geometry, they can often adopt a number of different geometries, depending on their situation. This means that metal ions can play multiple roles in a polygonal library, silver(I) for example, can commonly adopt tetrahedral or linear conformations. Metal ions also have a preferred type of ligand, described by the Hard-Soft Acid-Base concept (HSAB). Small species with high oxidation states or high electronegativity are classified as 'hard'. Large, low, or zero-charged species are classified as 'soft'. In general terms, hard metal ions prefer hard ligands, and soft metals prefer soft ligands, although there are borderline instances where a species falls between the two extremes. The HSAB principle can be used in the design of supramolecular systems to select an appropriate partner for a given metal ion or ligand.

1.3 Hierarchies in DNA Assemblies

1.3.1 Problems with Existing Definitions

Nucleic acids are capable of self-assembling into a variety of secondary and tertiary structures, although there is some confusion in the literature concerning the precise definition of the terms secondary- and tertiary structure and subsequently which nucleic acid assemblies are assigned to each category.[34-36] The problems seem to arise from a comparison to protein structure; to better explain the problems in nucleic acid structure, the definitions of peptide structure are briefly detailed. A peptide's amino acid sequence is considered primary structure. Secondary structure relates to the general arrangement of peptides without considering three dimensional positions; examples include α-helices, β-sheets and random coils. Tertiary structure is defined as the three dimensional positioning of peptides, while quaternary structure is the arrangement of multiple protein subunits into a larger structure where each subunit is considered to have tertiary structure. A hierarchy emerges where quaternary structure results from the assembly of tertiary structure, which in turn is

composed of the 3D arrangement of sections of secondary structure, which is itself composed of assembled peptides with defined primary structure.[35]

In direct analogy with protein primary sequence, DNA primary structure is defined by the exact sequence of nucleobases in an oligonucleotide.[36] This definition poses no problems, however, difficulties arise in the definition of higher-order assemblies. Secondary structure is variously defined as base-pairing (in this case excluding most non-duplex structures), or it is classified into categories such as duplexes or stem/loops (comparable to the classification of peptide secondary structures). Nucleic acid tertiary structures are defined by the three dimensional conformation of an assembly, usually classified into duplexes, triplexes, quadruplexes etc.[34] In this case there is some overlap between secondary and tertiary structures; duplexes feature in both categories. The tertiary, three dimensional coordinates of a duplex are of limited use when describing, for example, an extended, 150 nucleotide duplex floating free in solution. Further problems are encountered by the definition of quaternary structure: described as the interaction between separate oligonucleotides, the interactions between nucleic acids and proteins; or as the chromatin complex.

According to these definitions, a DNA duplex composed of two separate oligonucleotides can be described as secondary, tertiary, or quaternary, conversely, if both these oligonucleotides were combined into one large sequence that formed a hairpin duplex then the assembly would be considered secondary structure. A unimolecular G-quadruplex with four G-quartets is considered tertiary, but a comparable bimolecular or tetramolecular quadruplex would be quaternary. These definitions are not useful in describing higher-order structures such as in the dimerisation of G-quadruplexes or the structures known as G-wires, which clearly represent higher levels of organisation. These definitions appear to have arisen by analogy to protein structure from the perspective of molecular biologists with a focus on the forms of DNA and RNA observed *in vivo*; they are not particularly useful for describing structures found in DNA nanotechnology or supramolecular chemistry.

It is hoped the definitions described in this paragraph will simplify discussions of nucleic acid structures and address many of the problems described above.

1.3.2 Primary Structure

The most elementary nucleotide or nucleobase unit present in a given system. This, most often, will refer to oligonucleotides, but may also apply to individual nucleosides, nucleotides or nucleobases as these small units are, themselves, capable of self-assembly.

1.3.3 Secondary Structure

A particular motif formed from hydrogen bonding between units of primary structure, *regardless of the connectivity between those units*. In this case, a Watson-Crick base pair formed from H-bonding between two nucleobases (each a primary structural element) is considered an element of secondary structure; as would base-triplets found in triplexes, cytosine-cytosine base pairs found in i-motifs, and G-quartets found in G-quadruplexes.

1.3.4 Tertiary Structure

Assemblies consisting of several associated elements of secondary structure, ignoring their connectivity. For example: a duplex consists of stacks of base-pairs and G-quadruplexes consist of stacks of G-quartets. This definition also adequately describes DNA triplexes and i-motifs, and avoids problems defining unimolecular, bimolecular, and tetramolecular G-quadruplexes.

1.3.5 Quaternary Structure

The controlled, three-dimensional arrangement of two or more subunits, each of which has discrete tertiary structure. This adequately describes the face to face dimerisation of G-quadruplexes such as $(TGGGG)_4$, protein/nucleic acid complexes and extended arrays such as G-wires and DNA origami (large planned 2- and 3D structures assembled from DNA duplexes[37]).

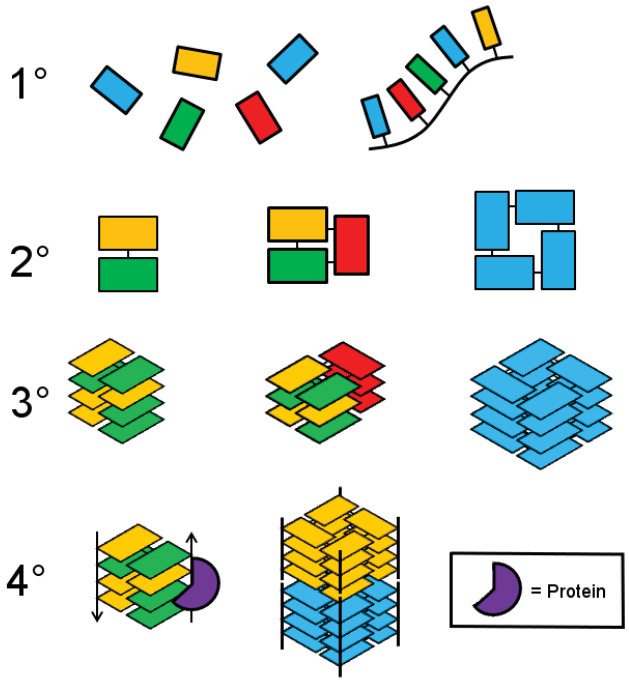


Figure 1.10. Schematic representation of hierarchies in DNA assembly. Each level is defined by interactions between units of the previous level.

1.4 Existing Metal-Coordinating Nucleotides

Previous research into metal-coordinating nucleotides has largely been limited to DNA duplexes: their application towards G-quadruplexes has been largely unexplored. The following pages contain a brief description of the most notable examples of metal-coordinating nucleotides.

1.4.1 Metal-Coordinating Ligands in Nucleic Acid Duplexes

The first example of a metal-mediated base-pair was demonstrated in 1999 by Tanaka *et al.* and consisted of a pair of nucleotides containing phenylenediamine coordinating a square planar Pd^{2+} cation, Figure 1.11.[38] The general design of metal binding nucleotides has remained relatively unchanged since; several examples are shown in Figure 1.12.

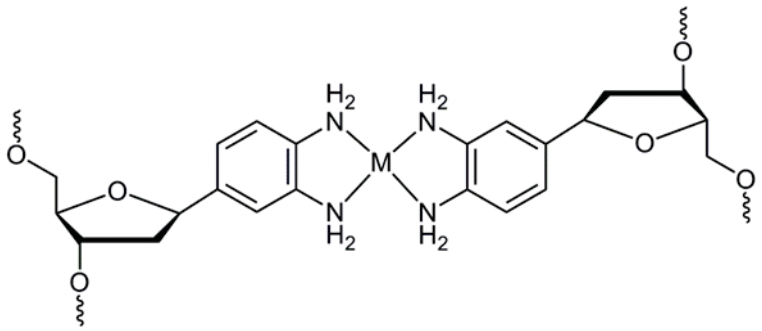


Figure 1.11. Tanaka's phenylenediamine base pair.

The majority of metal coordinating base pairs share several features; they are attached to DNA through the 1′ position of 2′-deoxyribose sugars, they bind metals through oxygen or nitrogen, and they typically form homodimers where both bases in a pair are identical. Exceptions to these general rules do exist, including the use of non-ribose based backbones, sulfur-metal binding, and the presence of hydrogen bonds to complement metal bonding.[39]

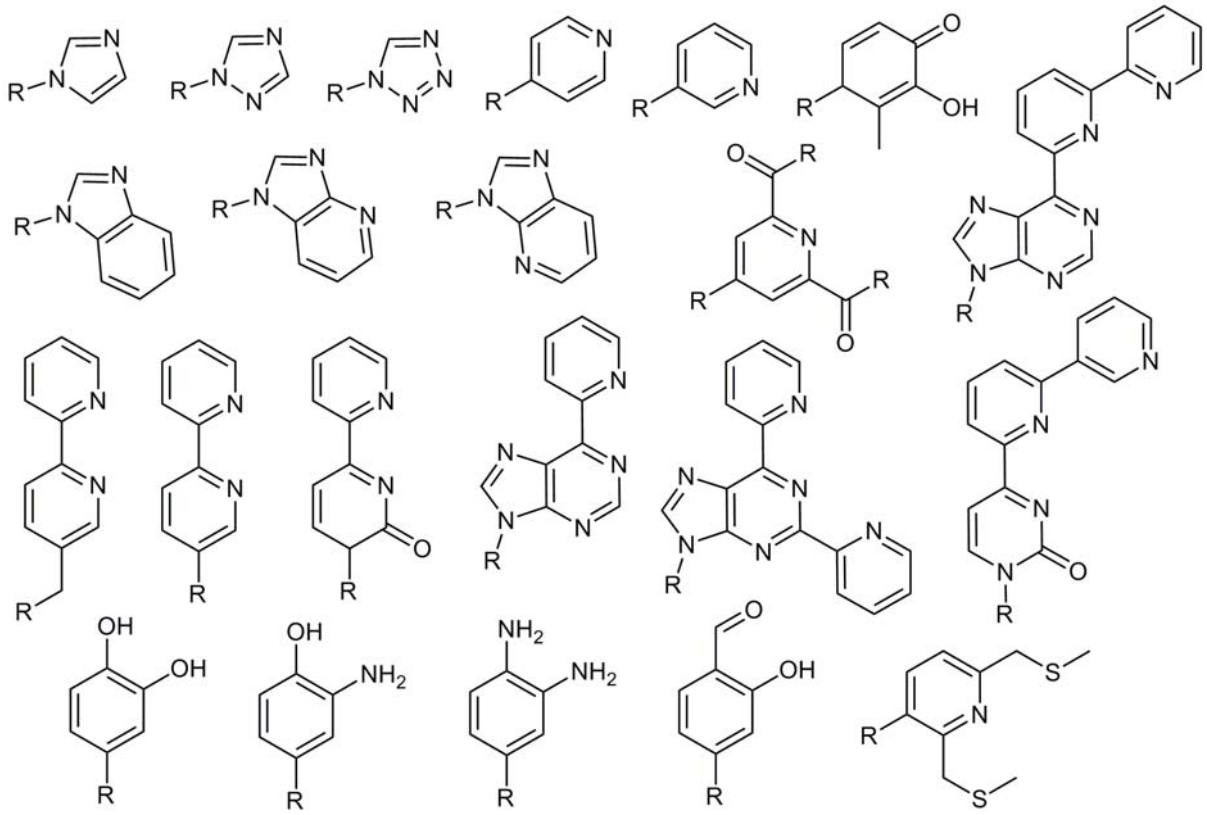


Figure 1.12. A selection of metal-coordinating base pairs found in the literature.

A 1,2,4-triazole based ligand, Figure 1.13, was used in the construction of an oligonucleotide that switches structure from a hairpin loop to a duplex upon the addition of Ag^+ ions.[40] The sequence used, $A_7\mathbf{T^r}_3T_7$ (where $\mathbf{T^r}$ is the triazole nucleotide), forms a hairpin loop under normal conditions with the modified nucleotides located in the loop regions. Upon addition of three equivalents of Ag^+ the oligonucleotide reorganises itself into a duplex containing three contiguous Ag^+ mediated base pairs in the middle of the sequence. This behaviour can be explained by the inability of the triazole ligands to hydrogen bond with each other – the absence of hydrogen bonding in the region where the three bases are located is energetically unfavourable and in these circumstances the hairpin is the most stable structure. The formation of metal-mediated base pairs in this region compensates for the absence of hydrogen bonding and so the duplex is the most stable conformation when Ag^+ is present. NMR studies of a similar duplex with three contiguous imidazole base pairs, rather than 1,2,4-triazole nucleotides, demonstrated a solution structure that is remarkably similar to that of a standard DNA duplex – even with three metal mediated base pairs the oligonucleotide was found to adopt a B-DNA conformation.[41] A difference in the size of the major and minor grooves of the artificial base pair when compared to their natural counterparts was found to cause a small change in the helical twist (the angle between adjacent base pairs) of the duplex and was observed as a slight unwinding of the duplex over the region containing the imidazole ligands. The space between adjacent imidazole base

pairs was slightly larger than that between natural base pairs, however, electrostatic repulsion between Ag^+ ions did not appear to be responsible for this increase as an even larger spacing was observed between imidazole and natural base pairs. Overall, the changes to B-DNA structure were surprisingly small, considering the significant differences between the metal-coordinating base pair and natural Watson-Crick base pairs.

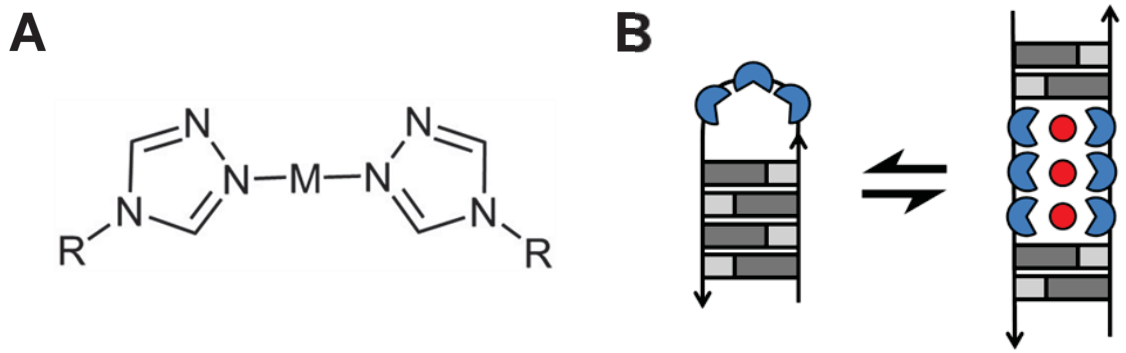


Figure 1.13. A 1,2,4-Triazole metal-mediated base pair, A. Addition of Ag^+ ions causes a change from hairpin loop to duplex, B.

Schlegel et al. developed a GNA containing several hydroxypyridone nucleotides which are capable of binding metal ions (Figure 1.14).[42] GNA is a simplified analogue of DNA where the 2′-deoxyribose sugars have been replaced by glycol units. GNA oligonucleotides typically have higher melting temperatures than an equivalent DNA sequence. Duplexes formed from the self complementary sequence CGHATHCG (where H is a hydroxypyridone nucleotide and all nucleotides are GNA) were found to have a T_M of 78 °C when coordinated by Cu^{2+}. In contrast, the related sequence CGAATTCG had T_M values of 26 °C (DNA) and 40 °C (GNA). X-ray structures revealed that GNA has a structure quite different to that of the familiar B-DNA double helix, with a greater distance between adjacent base pairs and an increased pitch.[43] The glycol units of the backbone appear to stack onto the bases of the previous nucleotide of the same strand, this stacking may be responsible for the increased stability of GNA – increasing the hydrophobic effects of the core of the helix.

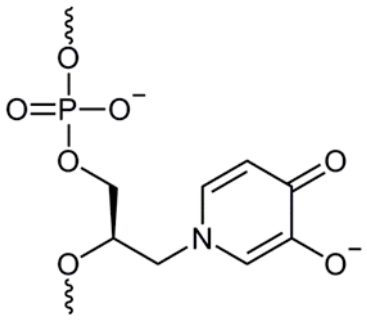


Figure 1.14. The structure of a single hydroxypyridone nucleotide of GNA.

It should be noted that naturally occurring nucleotides are also capable of forming metal mediated base pairs as in the case of the formation of thymidine or uridine base pairs with Hg^{2+}, and cytosine base pairs with Ag^{+}.[44] Several examples of duplexes containing T-Hg-T and C-Ag-C base pairs exist in the literature.[45-49]

1.4.2 Metal-Coordinating Ligands in G-Quadruplexes

To date we have found two examples of a G-quadruplex containing metal-binding ligands. Miyoshi's group developed a modified oligonucleotide, Figure 1.15, that can be switched between a dimeric quadruplex and an extended linear structure known as a guanine nanowire (Figure 1.16).[51,52] The quadruplex-forming sequence $G_4T_4G_4$, previously demonstrated to form G-wires in the presence of Mg^{2+} and spermidine,[53] was modified by replacing the four central thymines with 2,2′-bipyridine (commonly known as bipy) attached to DNA through polyether linkers.

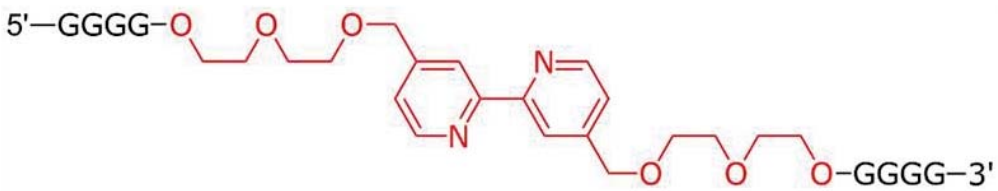


Figure 1.15. A guanine-rich oligonucleotide modified with a bipyridyl group.

This oligonucleotide forms a dimeric antiparallel G-quadruplex in the presence of 100 mM NaCl, with the bipy ligands and ethylene glycol linkers forming diagonal loops at either end of the quadruplex. Upon addition of Zn^{2+}, Co^{2+} or Ni^{2+} the oligonucleotide self-assembles into an extended parallel quadruplex bound by metal-bipy complexes. The two structures can be interchanged by addition of EDTA or another equivalent of divalent metal ions – the average efficiency for each cycle was found to be 95 %.

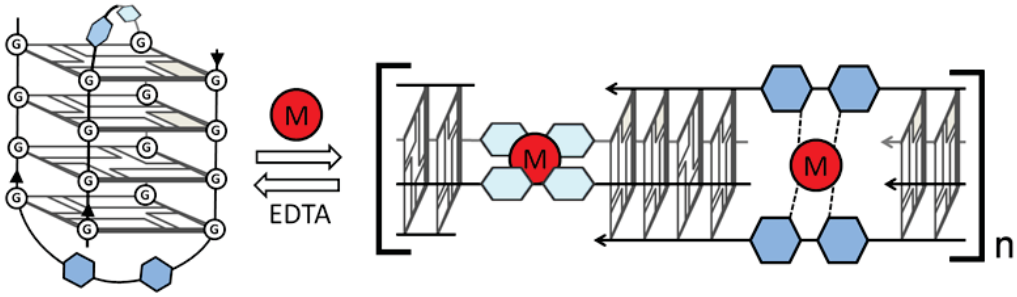


Figure 1.16. Miyoshi's oligonucleotide can switch conformations between a dimer and an extended quadruplex.[54]

During the course of our studies an example of a G-quadruplex making use of a T-Hg-T base pair was published by Smith *et* al.[50] The study detailed a series of G-quadruplexes based on

the same chair-type, three G-quartet core (Figure 1.17). The formation of a T-Hg-T base pair in certain positions between the two lateral loops was found to increase the thermal stability of the assembly by up to 10.5 °C, and also had the effect of reducing the number of potential conformations these oligonucleotides could assume. This system differs from our own, described in Chapter 2, in their use of a unimolecular G-quadruplex, and in the formation of the metal base pair in the loops of the structure, rather than directly on the surface of the terminal G-quartets, as in our case.

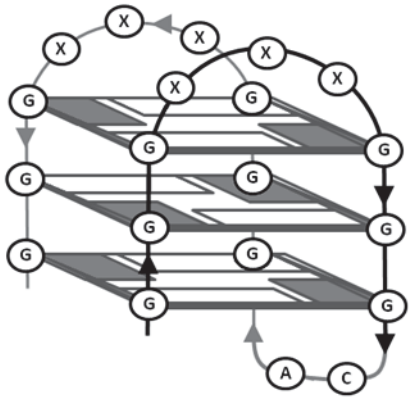


Figure 1.17. General chair-type G-quadruplex developed by Smith *et* al. The sequence of nucleotides in the two lateral loops (each labelled 'X') were altered to investigate a variety of potential T-Hg-T base pair conformations.

1.5 Oligonucleotide Nomenclature Used in this Thesis

This thesis presents and discusses a large number of oligonucleotides, many of which share similar sequences. To aid the reader, the last page of this document contains a fold-out list of each oligonucleotide and their sequence, listed by chapter. Because discussions of oligonucleotides are limited to comparisons *within* each chapter, a numbering system has been used that I believe simplifies these comparisons. Thus, each chapter has its own oligonucleotide series beginning with **ON1** and increasing from this point. In any situation where there may be ambiguity, I have included reference to the appropriate chapter.

1.6 Thesis Objectives

The goal of this thesis is to apply the principles of supramolecular assembly, with particular focus on metal-ligand coordination, to the design and assembly of G-quadruplex structures. Each of the chapters 2 – 5 investigate the effects that metal-coordinating nucleotides have on the assembly of different G-quadruplex systems.

Chapter 2 investigates the direct replacement of a Watson-Crick base pair in a 3+1 G-quadruplex formed from the human telomeric sequence. Chapter 3 details our attempt to use

the formation of a metal-mediated base pair to influence the position of equilibrium between two related G-quadruplex forming sequences and their ability to promote the formation of quaternary G-quadruplex assemblies. Chapter 4 focuses on tetramolecular G-quadruplexes and the use of metal-coordinating nucleotides to form metal-mediated quartets. Chapter 5 focuses on the use of metal coordination at cytosine-cytosine mismatches to promote duplex assembly and subsequent G-quadruplex assembly.

1.7 Methods

The following pages briefly cover the most commonly used techniques in DNA chemistry, including oligonucleotide synthesis and characterisation methods.

1.7.1 DNA Synthesis

Chemical methods for nucleic acid synthesis involve a cycle whereby a sequence is extended one nucleotide at a time, similar to peptide synthesis. For small-scale applications, synthesis takes place on a controlled pore glass (CPG) support. High-surface-area porous silica beads that are functionalised with a reactive, cleavable moiety from which oligonucleotides are extended. The first nucleobase of a desired sequence is typically pre-attached to the support. Synthesis on a solid support simplifies purification, as excess reagents and waste can simply be washed off at the end of each step.

The fact that DNA synthesis involves a large number of reactions in succession requires each step in the DNA synthesis cycle to have an extremely high yield, >99%, in order to produce usable quantities of material; even with 95% yield at each step, a short oligonucleotide with six bases would have a final yield of only 40%. The established methods make use of extremely reactive reagents and large excesses of reagents to ensure each step goes to completion. The standard method makes use of nucleotide precursors that contain 5′-O-dimethoxytrityl (DMT) protecting groups and 3′-β-cyanodiisopropyl phosphoramidites. Exocyclic amines on cytosine, guanine and adenine are protected to prevent their reaction with phosphoramidites and a β-cyanoethyl amidite group prevents two 3′-hydroxyl groups reacting with the same phosphorus centre.

Extension of an oligonucleotide sequence occurs via the following cycle:

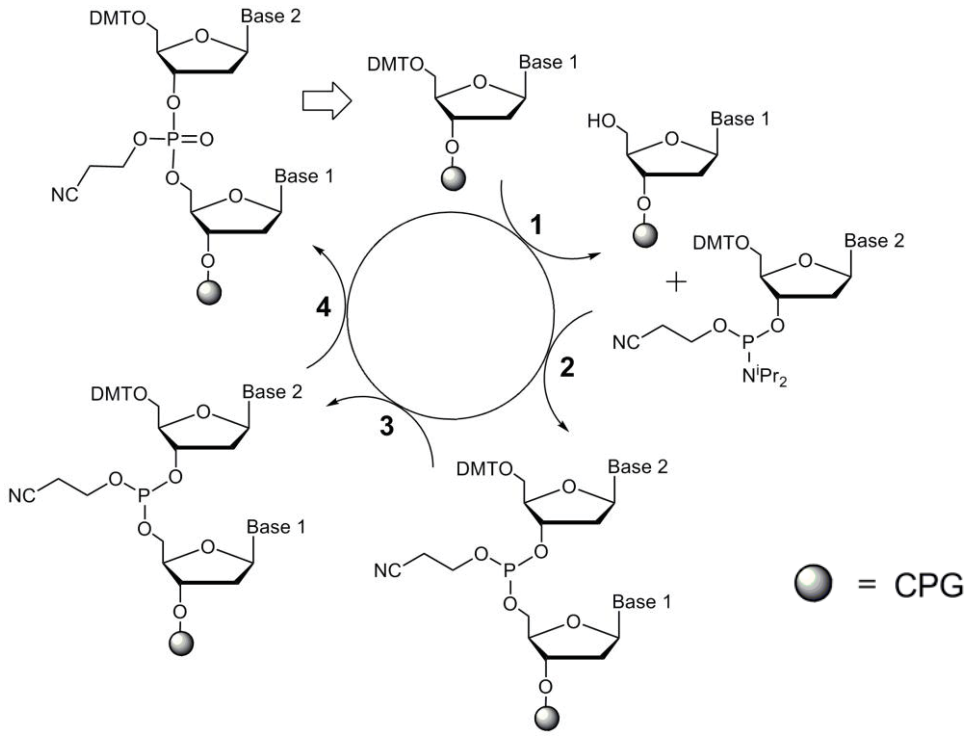


Figure 1.18. DNA synthesis cycle.

1. Deprotection: The acid-sensitive DMT protecting group is cleaved by a solution of 3 % trichloroacetic acid in DCM. Deprotection exposes a 5′-hydroxyl group.

2. Condensation: The next nucleoside, with a 3′-β-cyanodiisopropyl phosphoramidite group, is added in solution with an activator (such as 1*H*-tetrazole or 4,5-dicyanoimidazole) which is necessary for the reaction of the 3′-phosphine with the newly exposed 5′-hydroxyl of the previous base.

3. Capping: A solution containing acetic anhydride is added; any 3′-hydroxyl groups that did not react in the previous step are blocked from further reactions by an acetyl group.

4. Oxidation: A solution of I_2 in pyridine/H_2O is added to oxidise the newly formed phosphine to a more stable phosphate.

The cycle repeats as many times as is necessary, until all bases have been added. After the oligonucleotide is complete, cleavage from the CPG support and cleavage of the β-cyanoethyl protecting groups is carried out with 32% aqueous ammonia solution at r.t. for 30 minutes followed by raising the temperature to 55 °C for several hours to cleave the protecting groups on the nucleobases (cleavage of the phosphodiester backbone may occur if the temperature is raised before β-cyanoethyl groups are removed).

1.7.2 UV/Vis Spectroscopy

UV/Vis spectroscopy is used to determine the thermal stability and association kinetics of DNA assemblies, and to determine the concentration of solutions of DNA.

The concentration of a DNA sample is determined by the straightforward application of Beer's law. The extinction coefficient (ε) of an oligonucleotide depends on its nucleobase sequence. Several methods are available to predict the ε value of a given sequence. The nearest-neighbour method is the most common and is accurate to within approximately 4 % of the true figure.[55] Values of ε provided by this method are given for pairs of nucleotides – the total value is calculated by simply adding the appropriate number for a nucleotide, considering the identity of the nucleotide that preceded it. As an example, the ε value for the oligonucleotide TGCTA is calculated from the individual values for TG, GC, CT, and TA. This takes into account the effect of interactions between nucleobases on absorptivity. Variations on this method can be used to calculate DNA and RNA in single or double stranded forms. Online calculators are typically available on the websites of custom oligonucleotide suppliers.

The thermal stability of oligonucleotide assemblies is determined by monitoring changes in absorption at a particular wavelength while submitting the sample to a temperature ramp. The cooperative assembly of DNA generally leads a change in absorption that follows a sigmoidal curve. The point of inflection (the maximum in the first derivative) of these curves represents the temperature at which 50 % of an oligonucleotide is in single stranded form. This temperature is referred to as the T_M, or melting temperature and is used as a measure of the relative stability of a particular assembly. The dissociation of a DNA duplex is monitored by the increase in absorbance (hyperchromicity) observed at 260 nm, but changes in other structures may result in either hypo- or hyperchromicity, depending on the wavelength selected. G-quadruplexes, for example, are often monitored by hyperchromicity at 260 and 295 nm, and hypochromicity at 275 nm. In order for the T_M to be considered a true representation of thermodynamic stability, the dissociation and association of a structure must be reversible, that is, both the heating and cooling profiles must match.[56] Hysteresis is undesirable as the apparent T_M becomes a function of the rate of heating/cooling, although this may be reduced by lowering the temperature ramp. While not usually a problem for DNA duplexes, hysteresis is often unavoidable when studying bimolecular and tetramolecular G-quadruplexes as these complexes form so slowly that a sufficiently low ramp is impractical – in such cases the observed point of inflection for the heating curve is quoted as a $T_{1/2}$ value and the heating rate is specified.

In cases where hysteresis prevents calculation of a true T_M, UV/vis spectroscopy can be used to carry out isothermal association experiments. These experiments can provide data in the

form an association constant, k_{on}, to complement dissociation data represented by a $T_{1/2}$ value. Isothermal association experiments involve a using a concentrated DNA sample that is heated at a high temperature for sufficiently long to completely dissociate any high order structures. The sample is then immediately diluted into a larger volume and then monitored for changes in absorption. Isothermal association rates are often used to compare the kinetics of tetramolecular G-quadruplexes. The equation used to calculate association rates is:[57]

$$\alpha = [1 + C_o^{n-1} \cdot (n - 1) \cdot k_{on} \cdot t]^{1/(1-n)} \qquad \text{Equation 1.1}$$

Where α is the proportion of unfolded oligonucleotide, C_o is the initial oligonucleotide concentration, n is the order of the reaction (approximately 4 in the case of tetramolecular G-quadruplexes), k_{on} is the association rate, and t is time. α can be related to the absorption of a sample by the equation:

$$Abs = C_o[\alpha \cdot \varepsilon_{ss} + (1 - \alpha) \cdot \varepsilon_q] \qquad \text{Equation 1.2}$$

Where ε_{ss} and ε_q are the molar extinction coefficients for the single-stranded oligonucleotide and G-quadruplex forms, respectively.

Careful choice of oligonucleotide concentration is needed for these experiments – too high and the quadruplex will begin to associate before the experiment can be started, too low and the experiment will take too long. These difficulties highlight the major problems that must be addressed in order to extract useful data out of these experiments. Because of the nature of the equation used to model tetramolecular assembly, the very first reading recorded by the spectrophotometer is assumed to represent 100% unfolded DNA. Any delay when starting an experiment may have a significant effect on the resulting k_{on}. Such sensitivity makes reproducing data very difficult. Mixing the sample too much or too little can cause errors that only become apparent several hours later.

1.7.3 Circular Dichroism

Circular dichroism (CD) spectroscopy functions on the basis that chiral entities interact differently with left circularly polarised light (LCP) and right circularly polarised light (RCP). A CD experiment can be compared to measuring a spectrum by UV/Vis spectroscopy: a range of wavelengths are examined and the absorbance at each point is measured, the difference being that where UV/Vis data for each wavelength is displayed as simple absorbance, CD data is presented as the difference in absorption between LCP and RCP. The observed difference in absorption is reported in degrees of ellipticity or as molar ellipticity, the latter of which has been corrected for concentration. Circular dichroism spectroscopy is particularly useful in investigating the large-scale conformations of nucleic acids. Each of the

major tertiary conformations has a characteristic CD spectrum which can be described by noting the wavelength and polarity of each minimum and maximum, see Table 1.1. For example, parallel G-quadruplexes can be identified by a positive maximum at 260 nm and a negative minimum at 240 nm whereas antiparallel G-quadruplexes have a positive maximum at 295 nm and a negative minimum at 265 nm. For nucleic acid assemblies, the largest contributor to CD spectra is the syn/anti conformation of nucleobases, Figure 1.5.

Table 1.1. Characteristic peaks for various DNA species observed in CD spectra. Values are approximate and may be influenced by other factors such as external loops or ligand binding.[58,59]

Structure	Peak positions (nm)
Parallel G-quadruplex	+265, -240
Antiparallel G-quadruplex	+295, -265
DNA duplex	Variable*
ssDNA	Variable*

* CD spectra of duplex and ssDNA vary according to their sequence, particularly for short oligonucleotides.

CD spectra are less useful when it comes to differentiating between two structures of the same type – it would not be possible to reliably discern between two antiparallel triplexes, for instance, because of the broad nature of the signals obtained and because a variety of factors influence the final spectrum obtained, including the chemical environment and temperature. Because of the broad, qualitative nature of circular dichroism, data are often reported in degrees of ellipticity rather than the concentration-independent molar ellipticity as both units provides the same information regarding the position of maxima and minima. In this thesis many of the CD spectra have been normalised for intensity differences. This involves defining the point of maximum ellipticity as 1.0 and another point, typically at 350 nm, as 0.0 arbitrary units (a.u.). This method eliminates differences between spectra due to varying concentration of samples and allows a more direct comparison of the general shape and position of peaks.

1.7.4 Mass Spectrometry

Matrix-assisted laser desorption ionisation (MALDI) and electrospray ionisation (ESI) are the two most useful mass spectrometry methods for nucleic acid research. MALDI provides a method of confirming DNA synthesis efficiency and ESI provides information regarding the tertiary and quaternary conformation of nucleic acid assemblies. MALDI involves using a laser to desorb a sample that has been chemically ionised by an acidic or basic matrix. The matrix also serves to absorb the laser radiation, improving desorption of the sample and increasing the amount of analyte that subsequently enters the detector. The use of a laser limits MALDI's utility to easily ionised short oligonucleotides while the need for dry samples

co-crystallised with a matrix reduces a great deal of control needed to ensure the proper assembly of DNA complexes. These factors make it nearly impossible to observe nucleic acid assemblies. MALDI is, however, suitable for measuring the mass of individual oligonucleotides and is used to confirm the products of DNA synthesis. ESI has the advantage that samples may be introduced in solution – samples are aspirated and passed through a fine, charged nozzle producing charged droplets which undergo a number of desolvating processes that ultimately produce individual ions. ESI is a 'soft' ionisation technique which can produce multiply charged ions and can analyse nucleic acid assemblies. Data obtained from this technique are accurate enough, for example, to determine the number of oligonucleotides in a G-quadruplex and the identities of any counter-ions included in the central channel. However, ESI typically requires the use of volatile salts, necessitating the use of NH_4^+ in G-quadruplex assemblies and thereby reducing their stability.

1.7.5 Gel Electrophoresis

In gel electrophoresis, charged molecules are drawn through a cross-linked polymer gel by an electric field and are separated according to their size, charge, and molecular mass. Smaller analytes with higher charge/mass ratio migrate most quickly. Polyacrylamide gel electrophoresis (PAGE) is commonly used to assess the purity of relatively short nucleic acids up to ~200 nucleotides in length, and to investigate the assembly of supramolecular structures. The density of the polyacrylamide gel can be altered to provide varying levels of resolution; PAGE is capable of separating oligonucleotides differing by one nucleotide.[60] Nucleic acids have an approximately constant mass to charge ratio, this means that oligonucleotides are separated according to their size only – this enables easy comparison between similar structures, for example: single stranded DNA is separated based on the number of nucleotides in each sequence; similarly, duplexes are separated based on the number of base-pairs they contain. Care must be taken when comparing different assemblies since apparent size is not linearly related to nucleotide content. For example, a tetramolecular G-quadruplex composed of four six-base oligonucleotides does not migrate at the same rate as a twenty-four nucleotide single-stranded sequence, nor does a duplex consisting of twelve base-pairs. The size ladders included in PAGE figures in later chapters are single-stranded poly-thymidine sequences, thus T_{20} is an oligonucleotide of twenty thymidine nucleotides. When relating size of items on a gel, bands are described, for example, as migrating with T_{15}: this is not to say that the band in question is composed of fifteen nucleotides, rather, that it migrates at the same rate as a single-stranded oligonucleotide of that size. Two variants of PAGE include denaturing and native experiments. In denaturing PAGE, samples are incubated at high temperature in the

presence of a chaotropic agent, in many cases urea, to disrupt high order assemblies found in the sample. Chaotropic agents are also incorporated into the gel itself to prevent assembly during the experiment. Sodium dodecyl sulfate (SDS) or urea are frequently used for this purpose. In this way, bands migrate as unstructured, single-stranded nucleic acids which is useful for assessing the purity of the products of DNA synthesis. The standard conditions employed in denaturing PAGE are not always sufficient to ensure that nucleic acid assemblies are eliminated. G-quadruplexes, particularly those containing more than four G-quartets, are often stable at temperatures greater than 90 °C and, once assembled, are largely unaffected by urea, owing to the extremely low rate of G-quadruplex dissociation – chaotropic agents are never given the opportunity to disrupt H-bonding, even at high temperatures. The aim of native PAGE is to ensure that samples are capable of migrating on the gel as fully assembled structures. In this way, a sample's migration rate can provide information on its size and molecularity. If one sample produces several bands, native PAGE can also reveal the position of equilibrium that exists between an assembly and its components. To ensure that PAGE provides an accurate picture of nucleic acid assembly, mild conditions are used. Denaturing agents are obviously avoided and lower voltages are also used, increasing experiment time but reducing gel temperature and subsequent denaturing of samples during the experiment. A problem encountered when examining G-quadruplexes by electrophoresis is the fact that the stabilising cations found in the central channel of a quadruplex and the oligonucleotides themselves are drawn in opposite directions, effectively tearing a quadruplex apart and leading to an apparent stability lower than what would otherwise be observed. This problem can be alleviated to some degree by including cations in the gel and buffer – providing replacements for any ions drawn out of the central channel. This solution is somewhat limited and brings about further problems. For bimolecular and tetramolecular G-quadruplexes association is often so slow that providing replacement cations is of little use as the experiment will be over by the time a quadruplex is given the opportunity to reassemble. An increased ionic strength will also lead to further heating and subsequent destabilisation of assemblies – a lower voltage may address this problem but results in a longer experiment. After electrophoresis, oligonucleotide samples are typically visualised by immersing the gel in a solution of dye. The dyes used in this thesis are Stains-all or Gel-star. Stains-all is capable of intercalation and non-specific binding, allowing the staining of DNA, RNA and proteins. Gels are exposed to Stains-all for approximately five minutes, followed by destaining for approximately thirty minutes to remove dye bound to the polyacrylamide gel itself. DNA samples are stained purple and can be visualised using a standard camera. Gel-star is a fluorescent dye with greater sensitivity that stains-all. Gels are immersed in a dye solution for thirty minutes. Destaining is not necessary, as Gel-star only fluoresces upon intercalation into DNA. Because

of the requirement for intercalation for imaging, Gel-star is not effective in staining single stranded nucleic acids.

1.7.6 Capillary Electrophoresis

Capillary electrophoresis (CE) is conceptually similar to gel electrophoresis in that a sample is introduced at one end of the apparatus and its components are separated by a voltage applied across the system. The two methods differ in that samples are drawn through a capillary in CE, rather than a gel, as in PAGE. CE capillaries typically have a narrow diameter (40 μm) and are made of fused silica. These two factors lead to electroosmotic flow, a phenomenon that causes all solutes to migrate in the same direction. At pH above three the silanol groups lining the capillary are negatively charged, leading to the formation of a *diffuse double layer* consisting of firmly bound cations in a *fixed layer* and a second, loosely bound layer of cations in a *mobile layer*. The mobile layer is attracted to the negative electrode. Because of the extreme narrowness of the capillary, the solvated cations covering the internal surface represent a significant fraction of the contents of the capillary. As the solvated cations of the mobile layer move towards the cathode the electroosmotic flow generated is sufficient to ensure that all other contents of the capillary are dragged in the same direction. Solutes are separated according to their electroosmotic mobilities, a factor determined by size, charge, solvent viscosity and the applied voltage. A brief outline of CE procedure follows: an analyte is introduced into the system by immersing one end of a buffer-filled capillary into the sample and applying pressure so that a small volume (as little as several nanolitres) enters the column. Both ends of the capillary are immersed in a buffer solution and a voltage is applied. Solutes migrate through the capillary and pass through a UV/vis detector before eluting directly into the buffer solution at the cathode. As applied to nucleic acid assembly, CE's principal advantages over PAGE are its speed, low sample usage and easily processed output. A typical analysis takes approximately fifteen minutes, uses less than a tenth of the sample required by PAGE, and records output as an electropherogram that can be more easily quantified than the comparable array of bands on a stained gel. The narrowness of the capillaries also means that any heat generated during the process is quickly dissipated – reducing the thermal denaturation of samples. The most significant drawback of CE is its inability to separate all but the smallest oligonucleotides: although oligonucleotides have an approximately equal charge to mass ratio, their hydrodynamic radii (R_H, a measure of the effective size of a particle in solution, taking into account its shape and flexibility) differs to a degree that is most significant for small oligonucleotides. The difference in R_H between oligonucleotides of four and five bases is much more significant than the difference between oligonucleotides of eight and nine, or even twelve, bases. When viewed from the perspective of G-quadruplex study, this limitation means that with an

uncoated capillary CE can easily separate a single-stranded oligonucleotide with six bases from its G-quadruplex, for example, but is unable to resolve a single-stranded oligonucleotide with eight bases and its quadruplex. In these cases functionalised or matrix-filled capillaries may be used to improve separation.

1.7.7 NMR Spectroscopy

Nuclear magnetic resonance spectroscopy, NMR, can be both helped and hindered by the high guanine content of G-quadruplexes. Taking a tetramolecular G-quadruplex assembled from the sequence TGGGGT as an example, it can be seen that there are sixteen separate guanine nucleobases. In this tetramolecular example, the number of guanines observed may be reduced by the symmetry of the assembly, but each of the four guanines in a particular strand experiences a slightly different environment and will provide a slightly different set of signals in NMR. In the absence of four-fold symmetry, NMR spectra can feature a mass of highly overlapping, but slightly different, guanine signals that make it very difficult to distinguish individual bases when solving the three dimensional structure of a G-quadruplex. Conversely, the large amount of overlap can simplify the interpretation of one dimensional NMR spectra. For NMR, oligonucleotide samples are typically dissolved in an aqueous buffer containing 10 % D_2O. Guanine imino (H1) protons exposed to the solvent exchange with deuterium at a rate that usually causes the complete disappearance of H1 protons from NMR spectra. Because of this, in situations where it is suspected that an oligonucleotide may assemble into a G-quadruplex and where it is not necessary to completely solve the structure, ^{1}H NMR can reveal whether H1 of guanine is involved in H-bonding. The presence of peaks between 9 – 12 ppm is a strong indication of G-quadruplex assembly. A notable drawback of NMR is the requirement for relatively high oligonucleotide concentrations and sample volumes of ~200 μL. Samples for structure determination are often in the millimolar range, although simple ^{1}H NMR spectra may require as little as 200 μM. For comparison, PAGE requires oligonucleotide concentrations of 100 μM or less, whereas CD and UV/vis spectroscopy use as little as 2 μM.

1.8 References

(1) Doluca, O.; Withers, J. M.; Filichev, V. V.: Molecular Engineering of Guanine-Rich Sequences: Z-DNA, DNA Triplexes, and G-Quadruplexes. *Chem. Rev.* **2013**, *113*, 3044-3083.

(2) Watson, J. D.; Crick, F. H. C.: Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **1953**, *171*, 737.

(3) Nielsen, P.; Egholm, M.; Berg, R.; Buchardt, O.: Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science* **1991**, *254*, 1497-1500.

(4) Obika, S.; Nanbu, D.; Hari, Y.; Morio, K.-i.; In, Y.; Ishida, T.; Imanishi, T.: Synthesis of 2′-O,4′-C-methyleneuridine and -cytidine. Novel bicyclic nucleosides having a fixed C3, -endo sugar puckering. *Tetrahedron Lett.* **1997**, *38*, 8735-8738.

(5) Ueda, N.; Kawabata, T.; Takemoto, K.: Synthesis of N-(2,3-dihydroxypropyl) derivatives of nucleic bases. *J. Heterocycl. Chem.* **1971**, *8*, 827-829.

(6) Urata, H.; Shinohara, K.; Ogura, E.; Ueda, Y.; Akagi, M.: Mirror-Image DNA. *J. Am. Chem. Soc* **1991**, 8174-8175.

(7) Murayama, K.; Tanaka, Y.; Toda, T.; Kashida, H.; Asanuma, H.: Highly Stable Duplex Formation by Artificial Nucleic Acids Acyclic Threoninol Nucleic Acid (aTNA) and Serinol Nucleic Acid (SNA) with Acyclic Scaffolds. *Chem. Eur. J.* **2013**.

(8) Urata, H.; Shinohara, K.; Ogura, E.; Ueda, Y.; Akagi, M.: Mirror-image DNA. *J. Am. Chem. Soc.* **1991**, *113*, 8174-8175.

(9) Petersen, M.; Wengel, J.: LNA: a versatile tool for therapeutics and genomics. *Trends Biotechnol.* **2003**, *21*, 74-81.

(10) Zhang, L.; Peritz, A.; Meggers, E.: A Simple Glycol Nucleic Acid. *J. Am. Chem. Soc.* **2005**, *127*, 4174-4175.

(11) Busseron, E.; Ruff, Y.; Moulin, E.; Giuseppone, N.: Supramolecular self-assemblies as functional nanomaterials. *Nanoscale* **2013**, *5*, 7098-7140.

(12) Braun, E.; Keren, K.: From DNA to transistors. *Adv. Phys.* **2004**, *53*, 441-496.

(13) 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.* **2006**, *34*, 564-574.

(14) Guckian, K. M.; Schweitzer, B. A.; Ren, R. X. F.; Sheils, C. J.; Paris, P. L.; Tahmassebi, D. C.; Kool, E. T.: Experimental Measurement of Aromatic Stacking Affinities in the Context of Duplex DNA. *J. Am. Chem. Soc.* **1996**, *118*, 8182-8183.

(15) Basham, B.; Schroth, G. P.; Ho, P. S.: An A-DNA triplet code: thermodynamic rules for predicting A- and B-DNA. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 6464-6468.

(16) Choi, J.; Majima, T.: Conformational changes of non-B DNA. *Chem. Soc. Rev.* **2011**, *40*, 5893-5909.

(17) Herbert, A.; Rich, A.: Left-handed Z-DNA: structure and function. *Genetica* **1999**, *106*, 37-47.

(18) Rich, A.; Zhang, S.: Z-DNA: the long road to biological function. *Nat. Rev. Genet.* **2003**, *4*, 566-572.

(19) Felsenfeld, G.; Davies, D. R.; Rich, A.: Formation of a Three-Stranded Polynucleotide Molecule. *J. Am. Chem. Soc.* **1957**, *79*, 2023-2024.

(20) Frank-Kamenetskii, M. D.; Mirkin, S. M.: Triplex DNA Structures. *Annu. Rev. Biochem* **1995**, *64*, 65-95.

(21) Beal, P.; Dervan, P.: Second structural motif for recognition of DNA by oligonucleotide-directed triple-helix formation. *Science* **1991**, *251*, 1360-1363.

(22) Thuong, N. T.; Hélène, C.: Sequence-Specific Recognition and Modification of Double-Helical DNA by Oligonucleotides. *Angew. Chem. Int. Ed.* **1993**, *32*, 666-690.

(23) Doluca, O.; Boutorine, A. S.; Filichev, V. V.: Triplex-Forming Twisted Intercalating Nucleic Acids (TINAs): Design Rules, Stabilization of Antiparallel DNA Triplexes and Inhibition of G-Quartet-Dependent Self-Association. *ChemBioChem* **2011**, *12*, 2365-2374.

(24) Patel, P. K.; Hosur, R. V.: NMR observation of T-tetrads in a parallel stranded DNA quadruplex formed by Saccharomyces cerevisiae telomere repeats. *Nucleic Acids Res.* **1999**, *27*, 2457-64.

(25) Hazel, P.; Huppert, J.; Balasubramanian, S.; Neidle, S.: Loop-Length-Dependent Folding of G-Quadruplexes. *J. Am. Chem. Soc.* **2004**, *126*, 16405-16415.

(26) Mukundan, V. T.; Do, N. Q.; Phan, A. T.: HIV-1 integrase inhibitor T30177 forms a stacked dimeric G-quadruplex structure containing bulges. *Nucleic Acids Res.* **2011**, *39*, 8984-8991.

(27) Zhang, A. Y. Q.; Bugaut, A.; Balasubramanian, S.: A Sequence-Independent Analysis of the Loop Length Dependence of Intramolecular RNA G-Quadruplex Stability and Topology. *Biochemistry* **2011**, *50*, 7251-7258.

(28) Collie, G. W.; Parkinson, G. N.; Neidle, S.; Rosu, F.; De Pauw, E.; Gabelica, V.: Electrospray Mass Spectrometry of Telomeric RNA (TERRA) Reveals the Formation of Stable Multimeric G-Quadruplex Structures. *J. Am. Chem. Soc.* **2010**, *132*, 9328-9334.

(29) Lehn, J.-M.: Supramolecular Chemistry. Scope and Perspectives: Molecules, Supermolecules, Molecular Devices. In *Nobel Lecture*, 1987.

(30) Leininger, S.; Olenyuk, B.; Stang, P. J.: Self-Assembly of Discrete Cyclic Nanostructures Mediated by Transition Metals. *Chem. Rev.* **2000**, *100*, 853-908.

(31) Fujita, M.; Sasaki, O.; Mitsuhashi, T.; Fujita, T.; Yazaki, J.; Yamaguchi, K.; Ogura, K.: On the structure of transition-metal-linked molecular squares. *Chem. Commun.* **1996**, 1535-1536.

(32) Olenyuk, B.; Whiteford, J. A.; Stang, P. J.: Design and Study of Synthetic Chiral Nanoscopic Assemblies. Preparation and Characterization of Optically Active Hybrid, Iodonium–Transition-Metal and All-Transition-Metal Macrocyclic Molecular Squares. *J. Am. Chem. Soc.* **1996**, *118*, 8221-8230.

(33) Romero, F. M.; Ziessel, R.; Dupont-Gervais, A.; Van Dorsselaer, A.: Monitoring the iron(II)-induced self-assembly of preorganized tritopic ligands by electrospray mass spectrometry: unique formation of metallomacrocycles. *Chem. Commun.* **1996**, 551-553.

(34) Mathews, C. K.; van Holde, K. E.; Appling, D. R.; Anthony-Cahill, S. J.: *Biochemistry*; 4th ed.; Prentice Hall, 2012.

(35) *IUPAC Compendium of Chemical Terminology*; Blackwell Science, 1997.

(36) *Nucleic Acids in Chemistry and Biology*: 3rd ed.; Blackburn, G. M.; Gait, M. J.; Loakes, D.; Williams, D. M., Eds.; RSC Publishing, 2006.

(37) Rothemund, P. W. K.: Folding DNA to create nanoscale shapes and patterns. *Nature* **2006**, *440*, 297-302.

(38) Tanaka, K.; Shionoya, M.: Synthesis of a novel nucleoside for alternative DNA base pairing through metal complexation. *J. Org. Chem* **1999**, *64*, 5002-5003.

(39) Müller, J.: Metal-Ion-Mediated Base Pairs in Nucleic Acids. *Eur. J. Inorg. Chem.* **2008**, *2008*, 3749-3763.

(40) Bohme, D.; Dupre, N.; Megger, D. A.; Muller, J.: Conformational Change Induced by Metal-Ion-Binding to DNA Containing the Artificial 1,2,4-Triazole Nucleoside. *Inorg. Chem.* **2007**, *46*, 4300-4302.

(41) Johannsen, S.; Megger, N.; Bohme, D.; Sigel, R. K. O.; Muler, J.: Solution Structure of a DNA Double Helix with consecutive metal-mediated base pairs. *Nat. Chem.* **2010**, 1-6.

(42) Schlegel, M. K.; Zhang, L.; Pagano, N.; Meggers, E.: Metal-mediated base pairing within the simplified nucleic acid GNA. *Org. Biomol. Chem.* **2009**, *7*, 476-82.

(43) Schlegel, M. K.; Essen, L.-O.; Meggers, E.: Duplex structure of a minimal nucleic acid. *J. Am. Chem. Soc.* **2008**, *130*, 8158-9.

(44) Kosturko, L. D.; Folzer, C.; Stewart, R. F.: Crystal and molecular structure of a 2:1 complex of 1-methylthymine-mercury(II). *Biochemistry* **1974**, *13*, 3949-3952.

(45) Ono, T.; Yoshida, K.; Saotome, Y.; Sakabe, R.; Okamoto, I.; Ono, A.: Synthesis of covalently linked parallel and antiparallel DNA duplexes containing the metal-mediated base pairs T-Hg(II)-T and C-Ag(I)-C. *Chem. Commun.* **2010**, *47*, 1542-4.

(46) Torigoe, H.; Miyakawa, Y.; Kozasa, T.; Ono, A.: The specific interaction between two T:T mismatch base pairs and mercury (II) cation. *Nucleic Acids Symp. Ser.* **2007**, 185-6.

(47) Ono, A.; Cao, S.; Togashi, H.; Tashiro, M.; Fujimoto, T.; Machinami, T.; Oda, S.; Miyake, Y.; Okamoto, I.; Tanaka, Y.: Specific interactions between silver(I) ions and cytosine-cytosine pairs in DNA duplexes. *Chem. Commun.* **2008**, 4825-7.

(48) Petrovec, K.; Ravoo, B. J.; Muller, J.: Cooperative formation of silver(I)-mediated base pairs. *Chem. Commun.* **2012**, *48*, 11844-6.

(49) Okamoto, I.; Iwamoto, K.; Watanabe, Y.; Miyake, Y.; Ono, A.: Metal-ion selectivity of chemically modified uracil pairs in DNA duplexes. *Angew. Chem. Int. Ed.* **2009**, *48*, 1648-51.

(50) Smith, N. M.; Amrane, S.; Rosu, F.; Gabelica, V.; Mergny, J.-L.: Mercury-thymine interaction with a chair type G-quadruplex architecture. *Chem. Commun.* **2012**, *48*, 11464-6.

(51) Miyoshi, D.; Wang, Z.-M.; Karimata, H.; Sugimoto, N.: DNA nanowire sensitive to the surrounding conditions. *Nucleic Acids Symp. Ser.* **2005**, *49*, 43-44.

(52) Miyoshi, D.; Karimata, H.; Wang, Z. M.; Koumoto, K.; Sugimoto, N.: Artificial G-wire switch with 2,2 '-bipyridine units responsive to divalent metal ions. *J. Am. Chem. Soc.* **2007**, *129*, 5919-5925.

(53) Marsh, T. C.; Vesenka, J.; Henderson, E.: A new DNA nanostructure, the G-wire, imaged by scanning probe microscopy. *Nucleic Acids Res.* **1995**, *23*, 696-700.

(54) Miyoshi, D.; Karimata, H.; Wang, Z.-M.; Koumoto, K.; Sugimoto, N.: Artificial G-wire switch with 2,2'-bipyridine units responsive to divalent metal ions. *J. Am. Chem. Soc.* **2007**, *129*, 5919-25.

(55) Allawi, H. T.; SantaLucia, J.: Thermodynamics and NMR of Internal G·T Mismatches in DNA. *Biochemistry* **1997**, *36*, 10581-10594.

(56) Mergny, J.-L.; De Cian, A.; Ghelab, A.; Sacca , B.; Lacroix, L.; Cian, A. D.; Sacca, B.: Kinetics of tetramolecular quadruplexes. *Nucleic Acids Res.* **2005**, *33*, 81-94.

(57) Wyatt, J. R.; Davis, P. W.; Freier, S. M.: Kinetics of G-quartet-mediated tetramer formation. *Biochemistry* **1996**, *35*, 8002-8.

(58) Paramasivan, S.; Rujan, I.; Bolton, P. H. P. H.: Circular dichroism of quadruplex DNAs: applications to structure, cation effects and ligand binding. *Methods* **2007**, *43*, 324-31.

(59) Kypr, J.; Kejnovská, I.; Renčiuk, D.; Vorlíčková, M.: Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res.* **2009**, *37*, 1713-1725.

(60) Birrer, B. W.; Simon, M. I.; Lai, E.: The basis of high resolution separation of small DNAs by asymmetric-voltage field inversion electrophoresis and its application to DNA sequencing gels. *Nucleic Acids Res.* **1990**, *18*, 1481-1487.

2 1,2,4-Triazole Base Pairs in a 3+1 G-Quadruplex Assembly

2.1 Introduction

A 3+1 G-quadruplex is an antiparallel assembly in which three of the G-tracts are oriented in opposition to the fourth. Several examples of this conformation are known and it appears to be a common G-quadruplex motif.[1-4] The sequence d(G_3(TTAGGG)$_2$T), referred to as *htel3*, is particularly interesting as this example forms an *asymmetric* dimeric G-quadruplex where one oligonucleotide contributes more than the other. In typical dimeric quadruplexes each oligonucleotide will contribute two G-tracts; in the quadruplex formed from *htel3*, three G-tracts are provided by one oligonucleotide and the fourth by a second strand – the remaining G-tracts on this second strand remain unstructured in solution and do not take part in the assembly of other G-quadruplexes: a schematic diagram of this arrangement is illustrated in Figure 2.1. The basic structure of this *htel3* quadruplex has been modified[3] by removing the unstructured bases of the second strand, thereby creating a structure comprised of two different oligonucleotides with sequences d(GIGTTAGGGTTAGGGT) and d(TAGGGU), Figure 2.2. In the original paper, the inclusion of inosine ('I', Figure 2.1) at position 2 in place of guanine, was found to aid NMR assignment. The 3+1 G-quadruplex assembled from these two sequences had improved thermodynamic stability compared with the assembly formed from the *htel3* sequence.

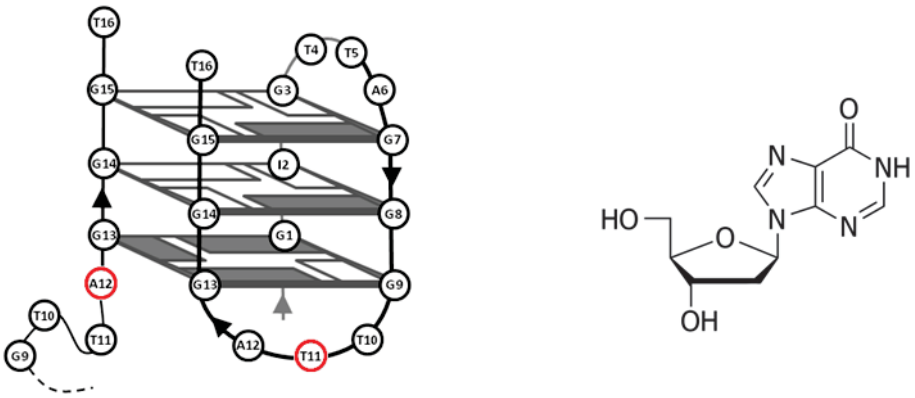


Figure 2.1. Schematic structure of the asymmetric dimeric 3+1 G-quadruplex formed from the *htel3* sequence, d(G_3(TTAGGG)$_2$T). A Watson-Crick base pair forms between the highlighted bases T11 and A12, left. Structure of 2′-deoxyinosine, right.

A notable feature of this family of structures is the presence of an A/T Watson-Crick base pair located on the surface of one of the terminal G-quartets; this base pair presumably stabilises the assembly by increasing H-bonding within the structure, and by shielding the G-

quartet face from the solvent. This chapter investigates the replacement of this A/T base pair with a metal-coordinating triazole-triazole base pair and the effect of this change on the stability of these 3+1 structures. Metal coordinating base pairs have been explored extensively for DNA duplexes, where reports have indicated that a single triazole-Ag-triazole base pair is capable of stabilising a duplex by up to 5 °C;[5] we anticipated that a 3+1 quadruplex structure incorporating metal-coordinating base pairs would be stabilised by transition metal ions capable of coordinating in a linear orientation, such as silver(I) or mercury(II) ions. A brief summary of notable metal-mediated base pairs that we considered in the design of our system can be found in the introduction to this thesis. The research of Smith *et al.* [6] bears some resemblance to this project (see Introduction, page 18). However, their research was published during the course of our work and was not used in the design of our assemblies.

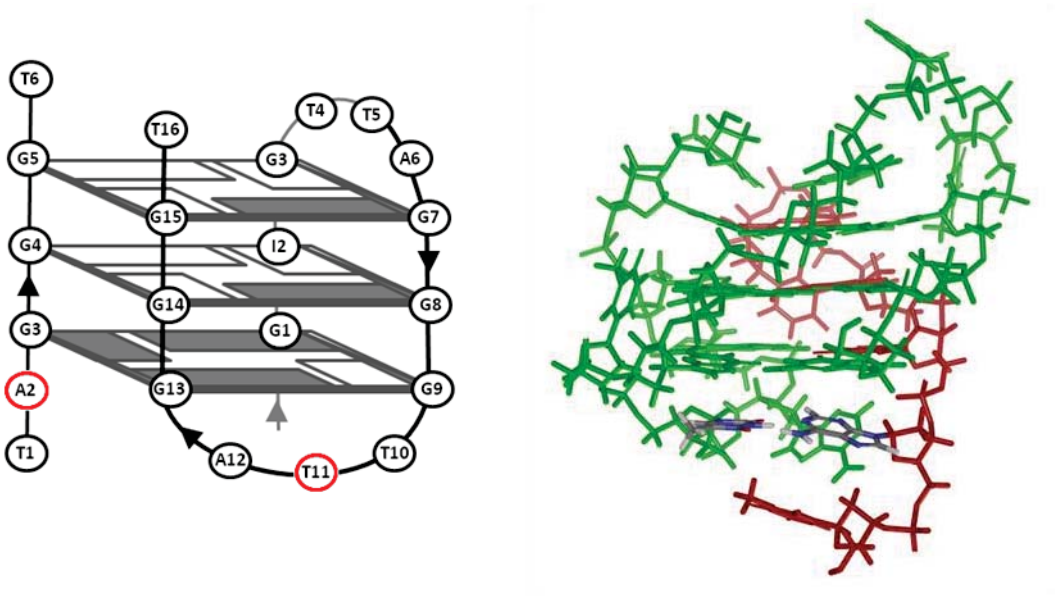


Figure 2.2. 3+1 htel G-quadruplex refined by Phan and coworkers.[3] Schematic topology, left, containing a Watson-Crick base pair between the highlighted bases T11 and A2. The structure as solved by NMR (PDB entry 2AQY), right. Sequences d(GIGTTAGGGTTAGGGT) and d(TAGGGU) have been coloured green and red, respectively.

2.2 Design of 3+1 G-quadruplex system

2.2.1 Selection of Metal-Coordinating Nucleotide

Our initial nucleotide targets were a series of novel pyridyl-based ligands, Figure 2.3A. We encountered difficulties synthesising the corresponding phosphoramidites. Molecular modelling suggested that the flexible nature of these nucleotides allowed the pyridyl nitrogen to assume a position adjacent to phosphorus, possibly sensitising the molecule to catalyse its own decomposition. To address this issue, we focused our efforts on investigating the metal-coordinating properties of 1,2,4-triazole nucleotides (Figure 2.3B) which feature rigid

furanose rings, thereby preventing triazole nitrogens from accessing the phosphorus. The triazole nucleotide, with several similar azole ligands, also had the benefit of a published synthetic route and the demonstrated ability to coordinate silver(I) and mercury(II) ions.[5,7,8]

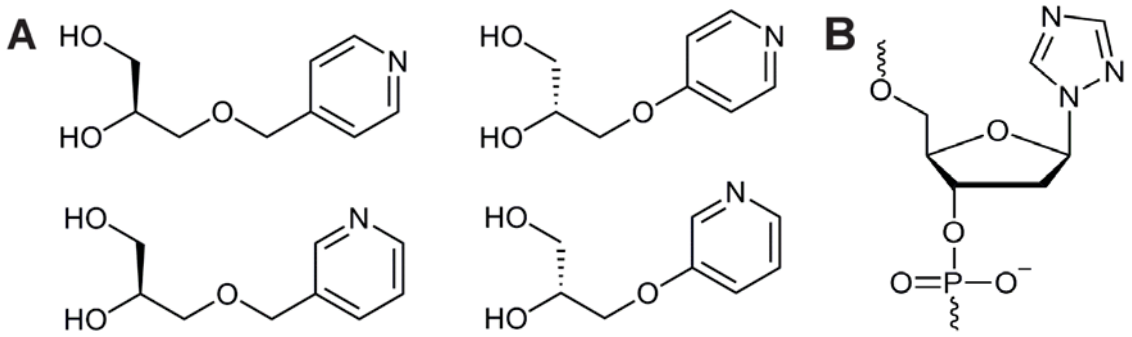


Figure 2.3. Our initial pyridyl-nucleotide targets, A. 1,2,4-Triazole nucleotide, B.

We opted to use the parent sequences d(GGGTTAGGGTTAGGGT) and d(TAGGGT), (the nucleobases participating in the Watson-Crick base pair are underlined) referred to as the control sequences **ON1** and **ON2**, respectively. Without the need to fully assign an NMR structure, the inosine nucleobase included in the original paper (Figure 2.1) increases the cost of oligonucleotide synthesis and may reduce the overall stability of the G-quadruplex assembly by reducing the number of hydrogen bonds present in the quartet. Figure 2.2 shows the relative positioning of the A/T base pair in the structure assumed by the control sequences. The triazole-nucleobase-containing oligonucleotides **ON3** and **ON4** have the sequences d(GGGTTAGGGT**T**^{r}AGGGT) and d(T**T**^{r}GGGT), respectively, differing from the control sequences only in the replacement of the T11 of **ON1** and A2 of **ON2** with triazole nucleotides. Molecular modelling, discussed in further detail below, suggested that a triazole-Ag-triazole base pair occupies approximately the same total space as a standard base pair and should not distort the assembly. Any transition-metal-dependent stabilisation, or destabilisation, would be apparent in a structure's T_M as determined by CD and UV/Vis spectroscopy.

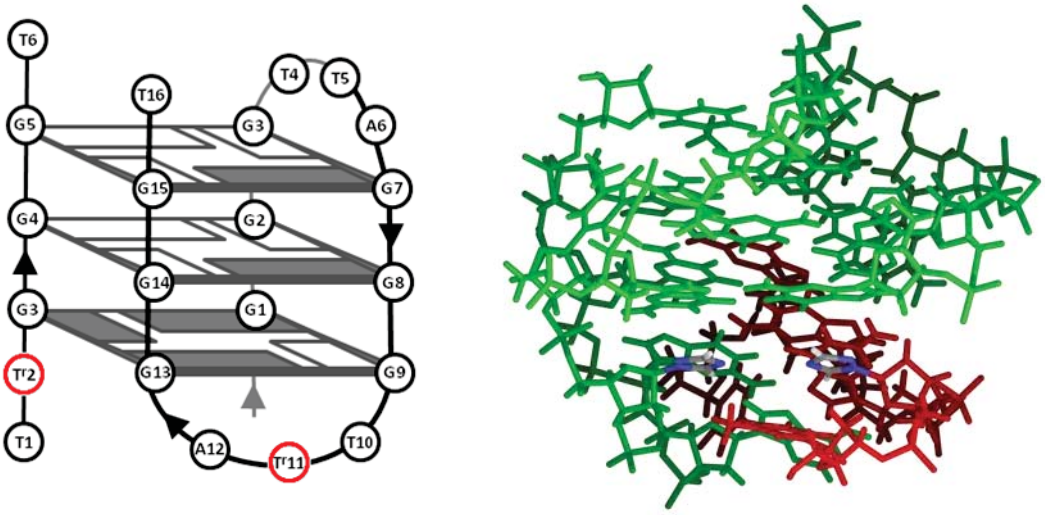


Figure 2.4. Schematic topology, left, containing a Triazole-triazole base pair between the highlighted bases **T**^{r}11 and **T**^{r}2. Calculated structure showing the position of the triazole base-pair, right. Sequences d(GGGTTAGGGT**T**^{r}AGGGT) and d(T**T**^{r}GGGT) have been coloured green and red, respectively.

2.2.2 Molecular Modelling of Metal-Coordinating 3+1 G-Quadruplex

We carried out molecular modelling based on the protein databank entry 2AQY, deposited by Phan and coworkers,[3] shown in Figure 2.2. The A/T Watson-Crick base pair between T11 of **ON1** and A2 of **ON2** was replaced by a triazole-triazole base pair. The modelling software used was not capable of processing transition metals; instead the base pair was constrained in a position consistent with an expected Ag-triazole length of 2.1 Å.[5] A series of molecular dynamic calculations were carried out, which consisted of random, small-scale alterations of residue positions in order to simulate motions of the assembly in aqueous solution. These models were carried out with the assumption that the triazole base-pair would remain intact. No attempt was made to determine whether molecular motions would be capable of disrupting this base pair. An example of the structures generated is included in Figure 2.5. The majority of these structures retained the conformation of the original assembly, with the largest variations occurring within the loops, having only a small influence on the overall stability of the structure.

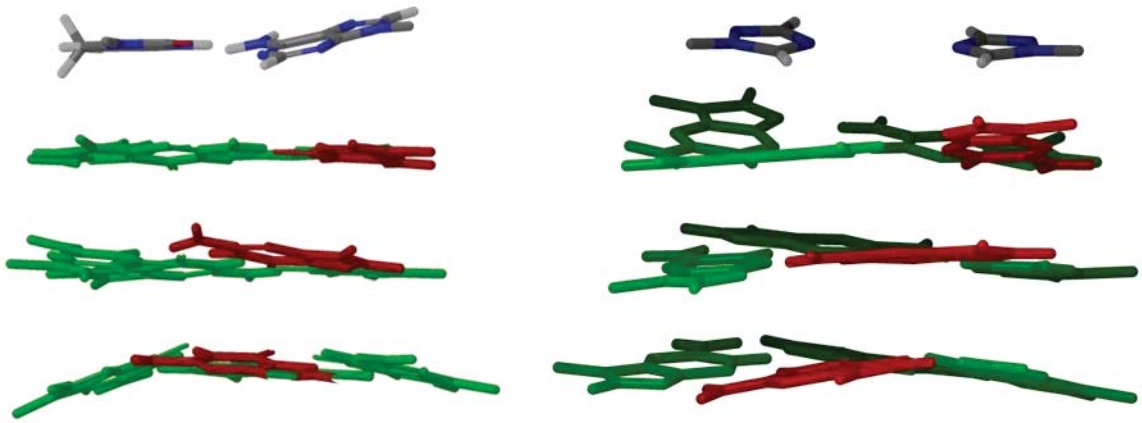


Figure 2.5. Relative positions of G-quartets and base pairs of 3+1 assemblies. NMR structure of control **ON1**/**ON2** assembly with A/T base pair, left. **ON1** and **ON2** are coloured green and red, respectively. **ON3**/**ON4** structure produced by molecular modelling, containing Triazole-Ag-Triazole base pair, right. **ON3** and **ON4** are coloured green and red, respectively. The DNA backbone and loops have been omitted for clarity.

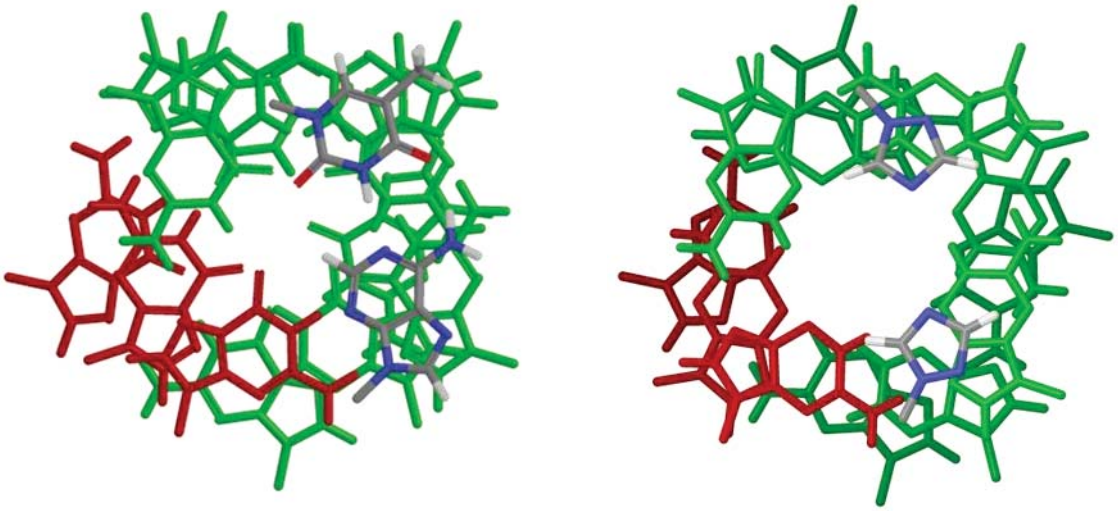


Figure 2.6. Relative position of base pairs. A/T base pair of the **ON1**/**ON2** complex, left. Triazole-Ag-Triazole of the **ON3**/**ON4** complex, right.

The 3+1 assembly formed from **ON1** and **ON2** forms an antiparallel G-quadruplex containing three G-quartets (the relative positioning of which are indicated in Figure 2.5) and two edgewise loops. The two loops each contain three nucleotides and are positioned on opposite sides of the assembly. The movement of one of the loops is somewhat restricted by the Watson-Crick A/T base pair; this restriction leads A12 to assume a position adjacent to the base pair, acting to shield the surface of the terminal G-quartet from the solvent – this position may be stabilised by a hydrogen bond between the NH_2 group of A12 and the furan

oxygen of the deoxyribose ring of G2 of **ON2**. T10 of the same loop assumes a position in one of the grooves. In the second loop, T4 assumes a position covering part of the terminal quartet; however, molecular modelling suggests that this nucleotide has some freedom of movement and may not act as an effective shield from the solvent. The NMR structure indicates that the remaining two nucleotides of this loop, T5 and A6, are largely unstructured: consistent with modelling, which reveals that a great deal of movement occurs in this loop.

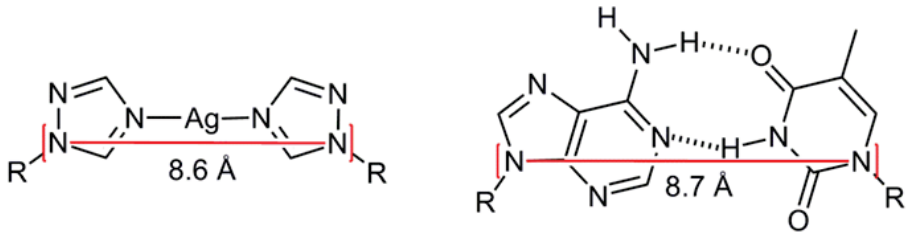


Figure 2.7. Triazole-Ag-Triazole base pair, left, Watson-Crick A/T base pair, right. Nitrogen-nitrogen distances are as indicated.

A comparison of the control sequence with the triazole-containing assembly reveals few large scale differences. The width of the triazole-Ag-triazole base pair, 8.6 Å, measured from N1 of each of the bases, is comparable to the size of the standard Watson-Crick base pair, 8.7 Å, Figure 2.7. This similarity suggests that any distortions of the structure would be minimal. The G-quartet immediately adjacent to the base pairs is most at risk of any deformation, but retains its conformation in the triazole assembly. The triazole rings of the base pair are coplanar, both with each other and with the G-quartets themselves. This arrangement serves to maximise Van der Waals contacts with the terminal G-quartet. In both triazole and Watson-Crick base pairs, A12 affords some degree of protection from the solvent, acting to shield the part of the terminal G-quartet not covered by the base pairs. The location of each of the base pairs is also conserved; in Figure 2.6 it can be seen that the triazole base pair assumes a similar position on one side of the quartet. The restrictions placed on the T10, T^{r}11, A12 loop by the triazole base pair match those of the parent structure: A12 and T10 occupy the same positions. The slight distortion in the G-quartet immediately adjacent to the triazole base pair is the only significant deviation from the parent structure: there are no other large-scale differences in the two structures. The similarity between these assemblies led us to the conclusion that the formation of a triazole-metal-triazole base pair at the positions mentioned would not result in serious disruption to the original assembly and that there are no structural features that would prevent the coordination of a transition metal to the triazole nucleotides.

2.3 Results and Discussion

2.3.1 Assessment of Buffer and Counter-Ions

The original experiments carried out by Phan *et al.* were performed in sodium chloride and sodium phosphate buffer.[3] The chloride anion is incompatible with our experiments due to the possible precipitation of AgCl from solution. Most of the common counter ions that do not precipitate with silver or mercury, such as nitrate or acetate, absorb strongly in the region of the UV spectrum used to monitor DNA assemblies (220 – 350 nm), for this reason we have used $NaClO_4$ and lithium cacodylate buffer which do not have significant absorbance in the wavelength range used in our CD and UV melting experiments. However, because our experiments used conditions different from that of the original paper it was necessary to ensure that that these conditions had no effect on the resultant G-quadruplex assembly.

The majority of previous studies on this 3+1 assembly have been carried out using NMR and have focused on structural details, as such, T_M studies appear absent in the literature and there exists only one example of a CD spectrum[9] – for the self association of the longer of the two strands – unfortunately recorded in conditions different to those used in our experiments. Due to the absence of CD data in the literature it was necessary to replicate the conditions of the original paper and perform our own measurements. Figure 2.8 compares the CD spectrum of the quadruplex formed from **ON1** and **ON2** using the conditions in Phan *et al.* with the spectrum using the conditions found in our experiments; the close match strongly suggests that the substitution of NaCl with $NaClO_4$ and Na_2HPO_4 buffer with lithium cacodylate has not led to the formation of a different structure. Both conditions gave rise to identical melting temperatures of 40 °C (determined by CD spectroscopy). The evidence strongly suggests that the G-quadruplex formed from **ON1** and **ON2** in our experiments is the same as that detailed in the original paper.

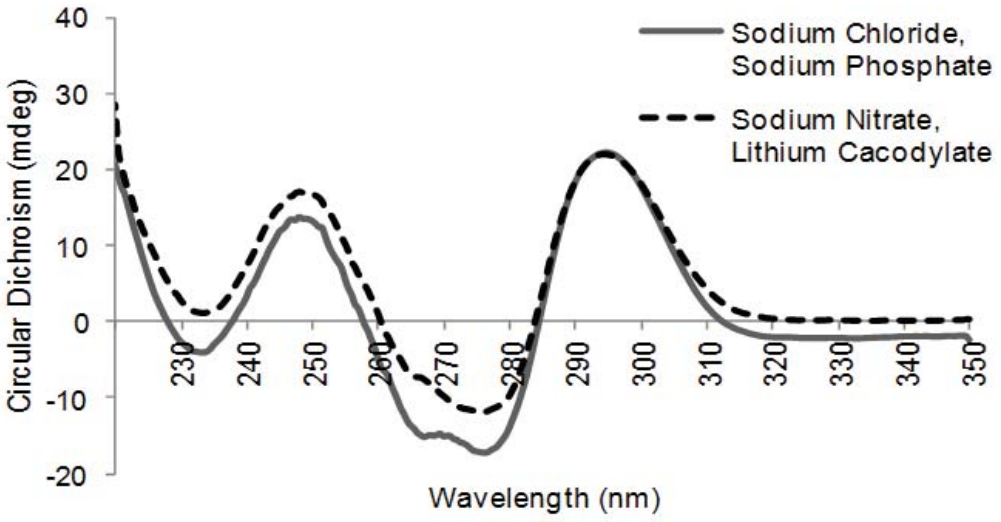


Figure 2.8. A comparison of the CD spectra of the 3+1 G-quadruplex formed from **ON1** & **ON2** under conditions used in the original paper (100 mM NaCl and 10 mM $NaPO_4$), solid line, and those used for melting experiments in this chapter (100 mM $NaClO_4$ and $Li(CH_3)_2AsO_2$), dashed line. Spectra were recorded at 20 °C with 10 μM DNA.

Due to the fact that **ON1** is capable of self associating into an **ON1**/**ON1** G-quadruplex, it was necessary to demonstrate that a mixture of **ON1** and **ON2** was in fact assembling into the expected **ON1**/**ON2** hybrid 3+1 G-quadruplex. CD spectra are not useful in distinguishing between the two possibilities as both assemblies are antiparallel and would provide similar spectra. We found that native PAGE experiments are capable of differentiating between the two possible assemblies and these experiments demonstrate support for the assembly of the **ON1**/**ON2** 3+1 Quadruplex: **ON1** self-assembles into a dimeric 3+1 G-quadruplex where two G-tract 'overhangs' remain unstructured – this assembly has a larger mass and is less compact, and thus would be expected to have a lower electrophoretic mobility than the **ON1**/**ON2** 3+1 assembly. PAGE experiments, shown in Figure 2.9, demonstrate that the addition of **ON2** to **ON1** (lane 3) results in a visible increase in electrophoretic mobility, greater than that of either **ON1** (lane 1) or **ON2** (lane 2) alone. It should also be noted that the combined **ON1**/**ON2** assembly is better defined than that formed from **ON1** alone, indicating that the product is more stable.

Interestingly, Figure 2.9 also indicates that there is no interaction between triazole-containing oligonucleotides **ON3** and **ON4** – the bands found in the lanes containing both (lanes 7 & 8) appear to be a simple combination of those bands of the individual oligonucleotides (lanes 5 & 6, respectively). No difference in mobility is observed, in contrast to the changes observed for **ON1** and **ON2**. The influence of silver and mercury ions on the assembly of **ON3** and **ON4** is discussed below.

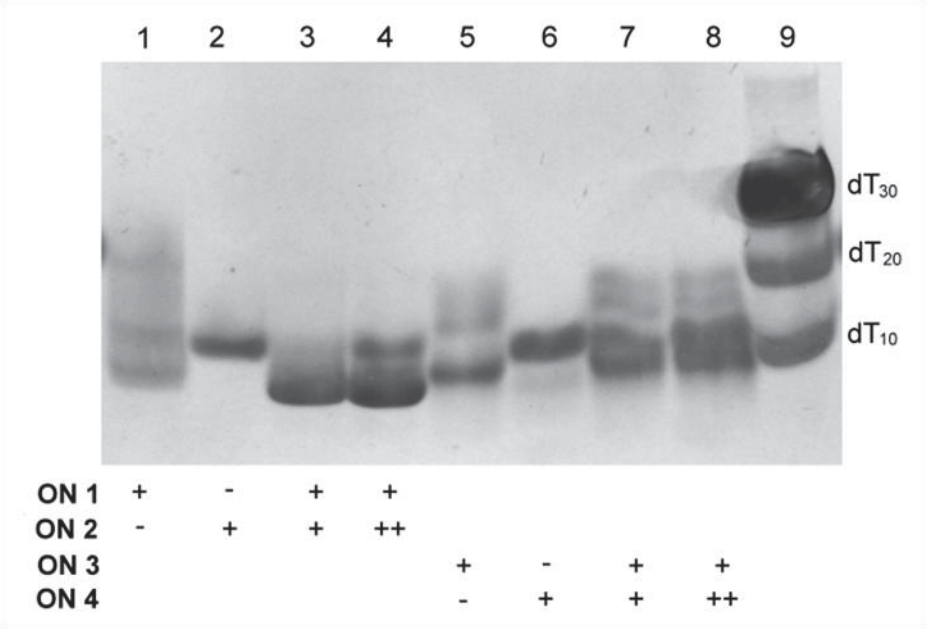


Figure 2.9. PAGE of 3+1 sequences. Native 12 % in TBE buffer (pH 8.4) supplemented with 100 mM $NaNO_3$. All samples were prepared in 100 mM $NaClO_4$, 10 mM lithium cacodylate buffer (pH 7.2). + is 1 eq.; ++ is 2 eq. DNA ladder (lane 9) is composed of oligothymidilates of lengths as indicated. Electrophoresis was carried out for 3 hours at +3 °C, 15 W.

Our preliminary experiments have demonstrated that the system we are investigating is comparable to that used in the literature, has indicated the formation of a hybrid **ON1/ ON2** 3+1 G-quadruplex, but has cast doubt on the formation of a similar assembly from **ON3** and **ON4**.

2.3.2 Thermal Stability Studies

Removing the Watson-Crick base-pair-forming nucleotides and replacing them with triazole nucleotides causes such a large reduction in stability that the hybrid 3+1 G-quadruplex no longer forms, eliminating the expected benefits derived from the preorganisation of the two triazole nucleotides. We investigated the ability of a linear-coordinating transition metal to improve the stability of the triazole-containing **ON3**/**ON4** assembly. Silver(I) and mercury(II) ions are both capable of coordinating in a linear manner, although they possess some flexibility in their preferred coordination geometry.[10] Both ions also have a suitably high lability to allow for reversible supramolecular assembly.[11] Considering the elimination of the Watson-Crick base pair had such a negative effect on the formation of a G-quadruplex, it was hypothesised that the addition of silver(I) or mercury(II) ions would promote the assembly of the **ON3**/**ON4** quadruplex through coordination to a triazole-triazole base pair on the surface of the G-quadruplex complex. Such an effect would be apparent in PAGE experiments and in the T_M of the resulting complex.

All samples were examined using both CD and UV/Visible spectroscopy, producing two sets of melting data from which T_M values could be extracted. T_M values have a degree of uncertainty introduced by the instrument used; for this reason values quoted in the literature are typically rounded to the nearest half degree. Data obtained by the UV/Vis spectrophotometer are likely to be more reliable owing to the improved sampling rate available on that instrument, however, both data sets were used to draw conclusions. Although the methods differ in their exact values, they typically reveal similar trends. Samples for CD and UV/Vis spectroscopy experiments were prepared in the same way; containing 10 mM lithium cacodylate buffer (pH7.2), 100 mM sodium perchlorate, 10 μM of each oligonucleotide strand and, when present, of either silver(I) nitrate or mercury(II) nitrate. Care was taken to ensure no compounds were present that would cause the precipitation of these metal ions. Thus, sodium perchlorate was used rather than sodium chloride to prevent the precipitation of AgCl. 10 μM of the metal ion corresponds to one equivalent, where one equivalent is the number of transition metal ions needed to form one triazole-triazole base pair per G-quadruplex.

Table 2.1. Melting temperatures of DNA assemblies.

	No metal	**Ag^+**	**Hg^{2+}**
ON1	37.0/38.0	NM	NM
ON1 + **ON2**	42.0/40.0	41.0/38.5	42.0/39.0
ON3	32.0/29.0	NM	NM
ON3 + **ON4**	32.0/32.0	27.0/31.5	33.0/32.0

NM – not measured. T_M values are within the uncertainty ±0.5 °C as determined by repetitive measurements, and reported in the format UV T_M/ CD T_M.

In the case of the control sequences, there is a large difference in T_M from 37.0 °C for samples containing **ON1** alone, to at least 40.0 °C for the samples containing both **ON1** and **ON2** – this suggests the formation of different species and is consistent with literature data suggesting that the hybrid **ON1/ON2** structure is thermodynamically more stable than its self-associated **ON1/ON1** counterpart. The triazole-containing sequences are much less stable than the control oligonucleotides: the T_M of **ON3** is reduced in comparison with **ON1** by as much as 9 °C (38.0 vs 29.0 °C). This is unsurprising, considering that PAGE results indicated that the elimination of the Watson-Crick base pair present in the **ON1**/**ON1** quadruplex was strongly destabilising.

Comparing the triazole-containing samples, there appears to be no difference in T_M between samples containing **ON3** and those containing both **ON3** and **ON4**, this supports the conclusion discussed above, and illustrated by Figure 2.9, that there is no interaction

between **ON3** and **ON4**. It is worth considering that even in the absence of base pairs (Watson-Crick or triazole-metal-triazole) the hybrid **ON3**/**ON4** quadruplex would be expected to have a higher thermal stability than the **ON3**/**ON3** assembly, thus the absence of the former also suggests the absence of the latter. By analogy with the control G-quadruplex, the core G-quadruplex for both **ON1**/**ON1** and **ON1**/**ON2** are identical. The differences in their stability can be explained by supramolecular principles: enthalpy favours the formation of a hybrid structure where a greater number of nucleobases are engaged in H-bonding as this satisfies the principle of maximal site occupancy. Entropy, through the hydrophobic effect, also favours the hybrid as a more compact structure reduces exposure of aromatic nucleobases surfaces to solvent. The triazole quadruplex will ultimately be less stable due to the absence of the base pair, but the same principles encourage the formation of the **ON3**/**ON4** quadruplex over **ON3**/**ON3**. As the evidence provided in Figure 2.9 and Table 2.1 indicates that an **ON3**/**ON4** quadruplex *does not form* it may be assumed that the less stable **ON3**/**ON3** quadruplex would also be absent. If this is the case, an explanation must be found for presence of a melting transition as some folded species must be present. The most likely possibility is that a G-triplex may be forming, Figure 2.10. Examples of G-triplexes have been reported in the literature.[12,13] It should be noted that an intramolecular triplex will be favoured over a quadruplex both kinetically and entropically.

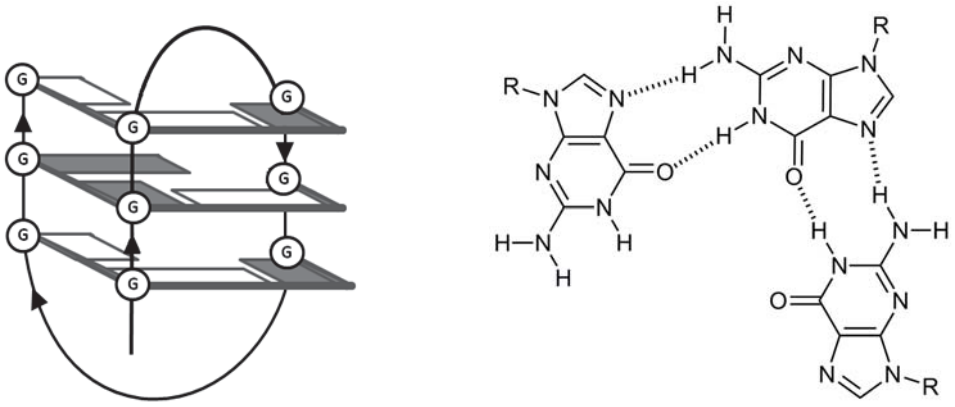


Figure 2.10. A schematic representation of the G-triplex formed from the sequence TTA(GGGTTA)$_3$, left.[12] Hydrogen bonding found in G-triplets, right.[13]

Also apparent from Table 2.1 is that the addition of silver or mercury ions does not increase the T_M of these assemblies and in some cases reduces their stability – this result suggests that the transition metal ions do not coordinate to the available triazole bases, or at least that their coordination does not result in an increase in thermal stability.

In summary, the T_M data supports the formation of an **ON1/ON2** structure but provides no evidence to suggest that silver or mercury promote the assembly of an **ON3**/**ON4** G-

quadruplex. In fact, T_M data provides evidence *against* the formation of either an **ON3**/**ON4** or an **ON3**/**ON3** G-quadruplex.

2.3.3 Examination of Metal Ion Coordination by PAGE

PAGE experiments cast further doubt on the ability of the triazole nucleotides to coordinate transition metals. Figure 2.11 demonstrates a complete lack of response of either **ON3** or **ON4** to silver or mercury ions, even in a large excess, and demonstrates that **ON3** and **ON4** do not interact – the bands observed in lanes containing both oligonucleotides (lanes 5-8) are simply the addition of the bands found in those lanes where **ON3** and **ON4** are, individually (lanes 1 & 2, and 3 & 4, respectively).

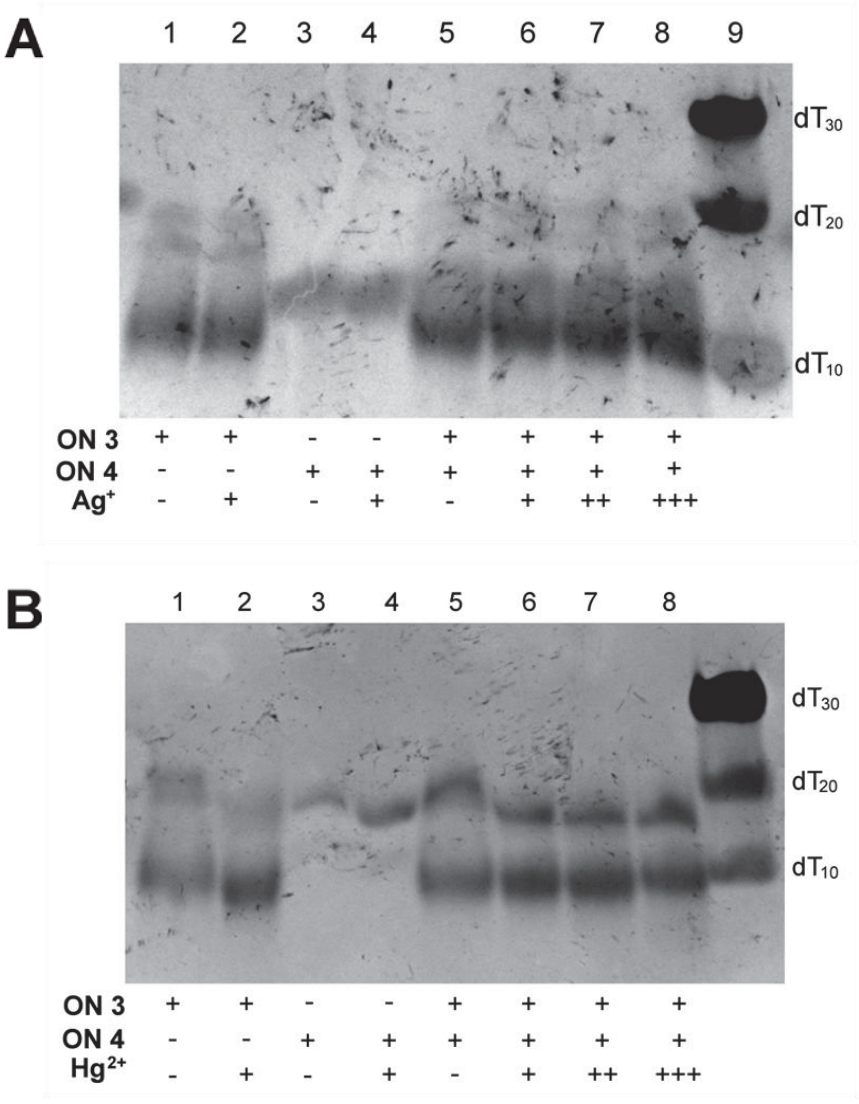


Figure 2.11. Effect of Ag^+ ions, A, and Hg^{2+} ions, B, on the assembly of **ON3** and **ON4** oligonucleotides. Native 12 % PAGE of 3+1 control sequences in TB buffer (pH 8.4) supplemented with 100 mM $NaNO_3$. All samples were prepared in 100 mM $NaClO_4$, 10 mM lithium cacodylate buffer (pH 7.2). + is 1 eq.; ++ is 2 eq.; +++ is 10 eq.. DNA ladder (lane 9) is composed of oligothymidilates of different length as indicated. Electrophoresis was carried out for 3 hours at +3 °C, 15 W.

2.3.4 Circular Dichroism Experiments

A comparison of the changes in CD spectra observed upon the addition of **ON2** to **ON1** and the addition of **ON4** to **ON3**, Figure 2.12, reveals small changes in intensity of maxima and minima and slight changes to the position of the maxima at ~250 nm. The data are consistent with the replacement of an **ON1**/**ON1** G-quadruplex with an **ON1**/**ON2** assembly – a situation where minimal changes to CD spectra are to be expected. It is, however, difficult to draw conclusions from the triazole containing system: T_M and PAGE experiments have indicated that the **ON3**/**ON4** quadruplex has not formed, but we are unable to distinguish between an **ON3**/**ON3** G-quadruplex and an **ON3** G-triplex as both of these species are expected to produce similar CD spectra.

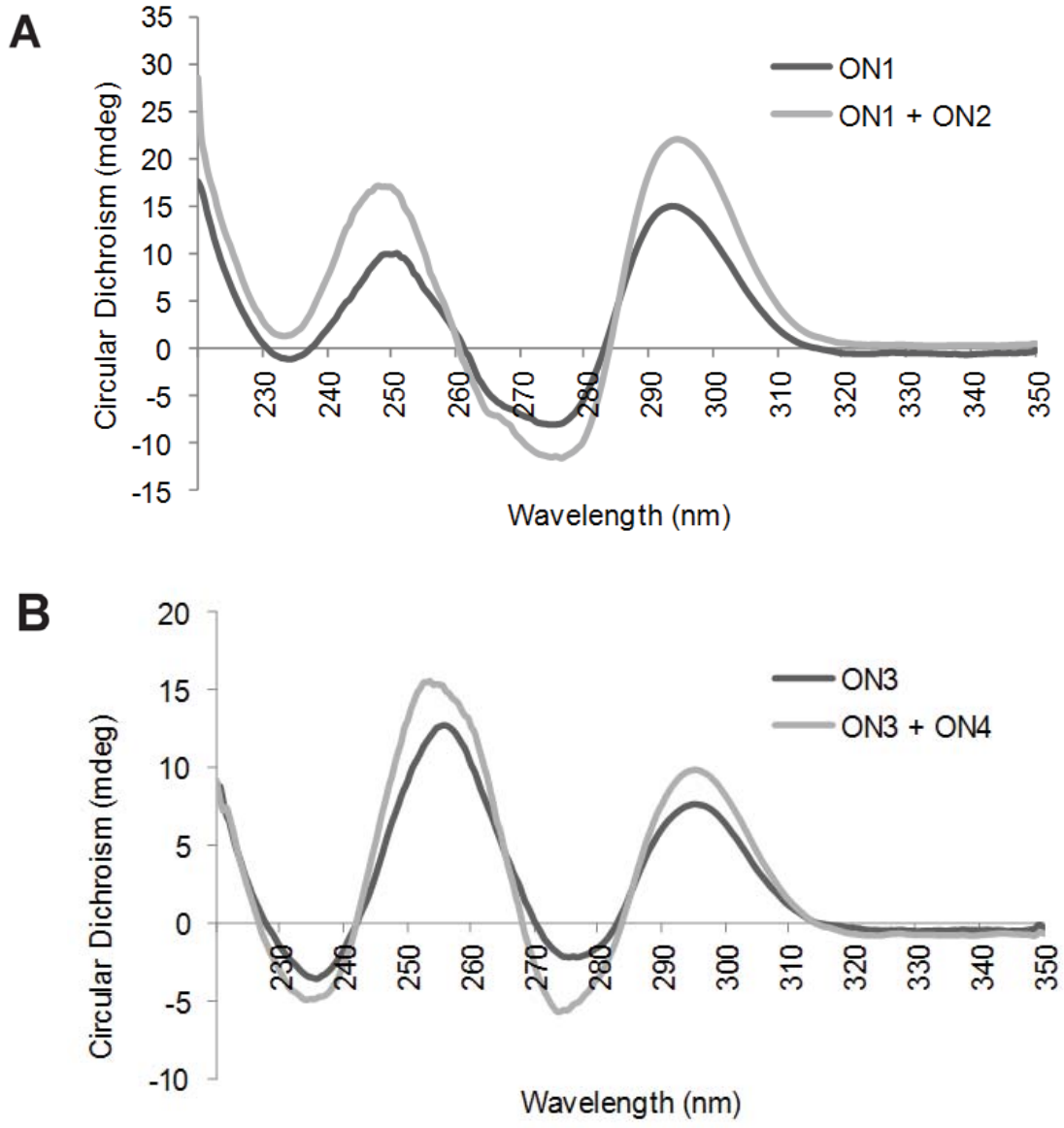


Figure 2.12. CD spectra of samples containing **ON1** alone and **ON1** + **ON2**, A, and **ON3** alone and **ON3** + **ON4**, B. Spectra were recorded at 20 °C in 100 mM $NaClO_4$, $Li(CH_3)_2AsO_2$) (pH 7.2) and 10 μM DNA.

Literature data of the three *htel* repeat sequence **ON1** was recorded in 100 mM KCl solution[9] and so is not directly comparably to our own data. Our CD spectra of the self-association of **ON1** and **ON3**, Figure 2.13, revealed positive maxima at 295 and negative minima at 275 nm indicating that both oligonucleotides assume an antiparallel conformation as expected. Although these CD spectra are similar in that each forms an antiparallel assembly, but they differ in the relative intensities of their maxima and minima. Due to the nature of CD spectroscopy, it is difficult to assign specific structural details to these differences although in this context several possible explanations for these differences may be considered: firstly, as these two sequences differ by only one nucleotide, it may be that the differences in spectra are simply due to the absence of the A/T base pair in the triazole-containing structure. In this scenario the rest of the structure remains intact. One of the problems facing this explanation has been described in the previous pages, in that the absence of an **ON3**/**ON4** assembly also suggest the absence of an **ON3**/**ON3** assembly; even further, the magnitude of differences between the CD spectra is larger than would be expected for a change in a single base pair, particularly as the CD spectrum is dominated by contributions from the 12 guanine nucleobases and that even in the absence of a triazole-metal-triazole base pair the triazole bases would be expected to occupy similar positions on the surface of the terminal quartet due to π-π interactions and the hydrophobic effect, meaning that their contribution to the CD spectra would be similar regardless of whether they coordinate a transition metal or not. A second possibility, as previously suggested, may be the presence of a G-triplex. The evidence supporting this possibility lays chiefly in the fact that in the absence of a stable G-quadruplex, an intramolecular triplex will be favoured kinetically and entropically. A third possibility is a mixture of G-quadruplex and G-triplex in solution.

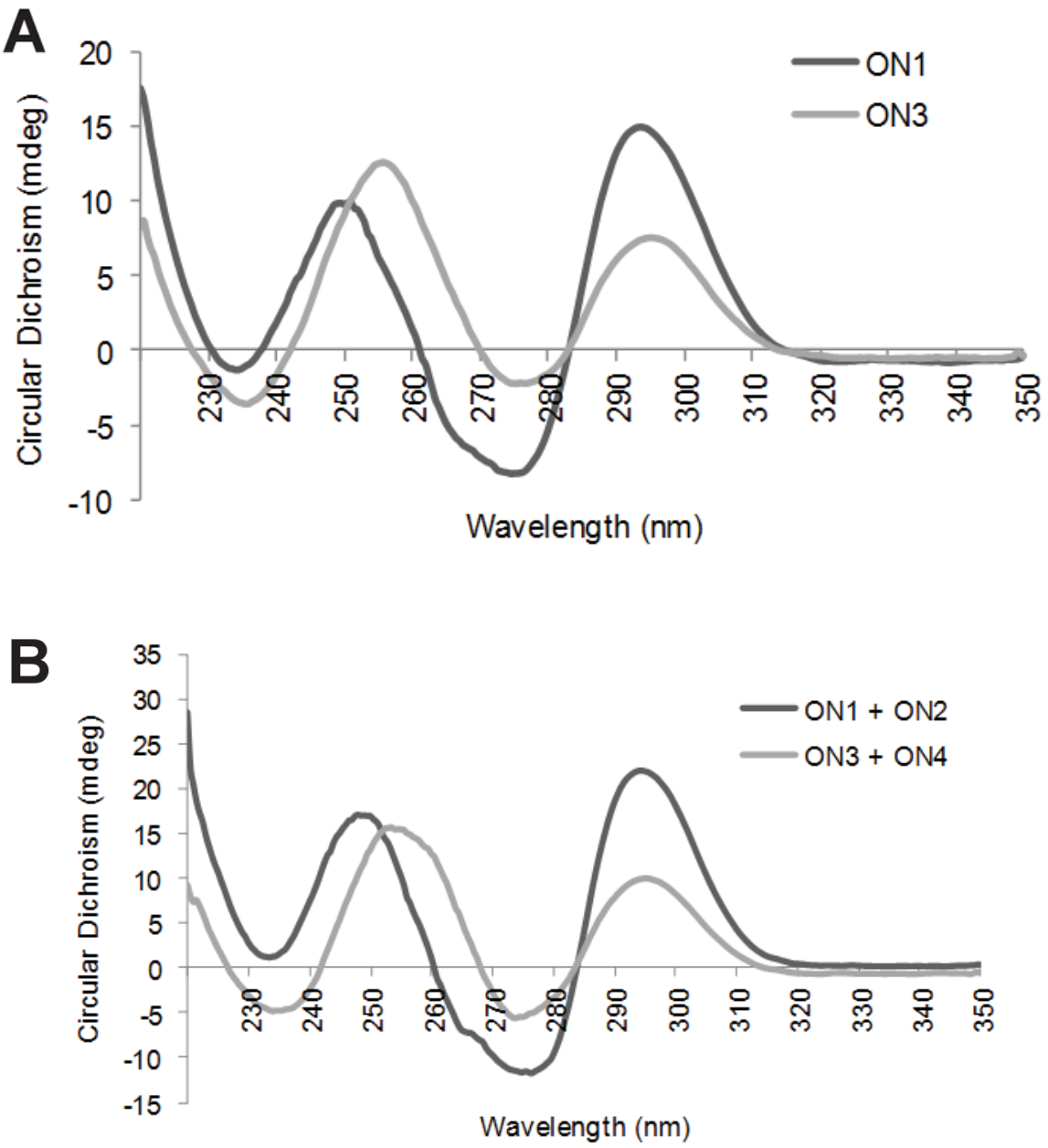


Figure 2.13. CD spectra comparing the self-assembly of **ON1** with **ON3**, A, and samples containing **ON1** & **ON2** with samples containing **ON3** & **ON4**, B. Spectra were recorded at 25 °C, with 10 mM lithium cacodylate buffer (pH 7.2), 100 mM $NaClO_4$ and with no transition metals present.

Reports in the literature suggest that the short, single repeat telomeric sequences, d(TAGGGT) and d(T**T**^{r}GGGT), are capable of forming ordered structures of limited stability.[3] Our results, Figure 2.14, indicate that both sequences produce CD spectra that largely resemble unstructured DNA with a positive maximum elipticity at 255 nm, in contrast to 260-265 nm observed for parallel G-quadruplexes.

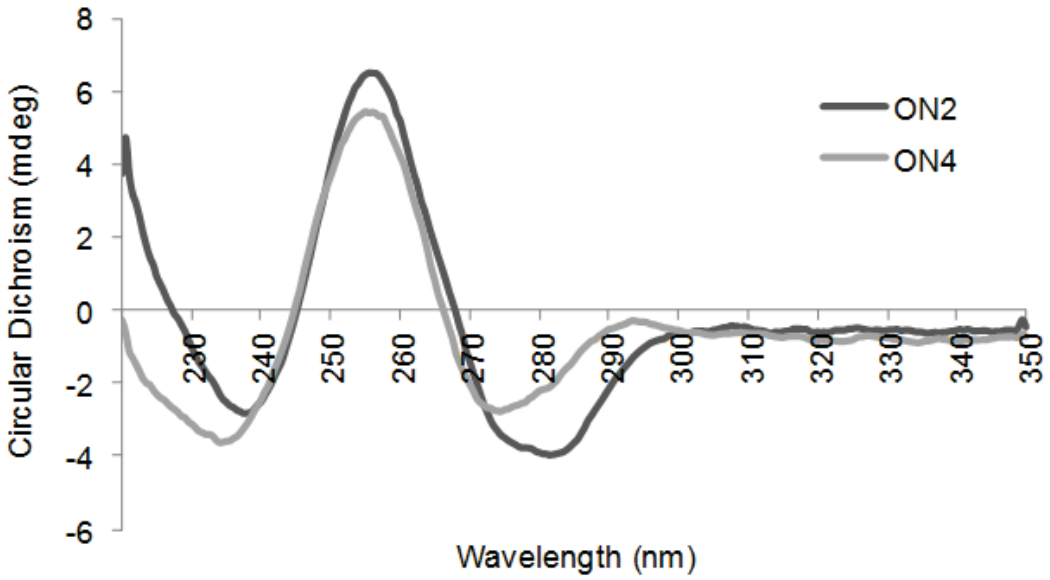


Figure 2.14. CD spectra of sequences **ON2** d(TAGGGT), solid line, and **ON4** d(TT^{r}GGGT), dashed line. Spectra were recorded at 25 °C, with 10 mM lithium cacodylate buffer (pH 7.2), 100 mM $NaClO_4$ and with no transition metals present.

2.3.5 Detection of 1,2,4-Triazole – Ag^+ interactions by ^{1}H NMR

Finally, we attempted to detect the coordination of silver by NMR spectroscopy. Experimental evidence for the coordination of silver(I) and mercury(II) ions to the toluoyl-protected 1,2,4-triazole nucleosides, Figure 2.15, has been demonstrated using ^{1}H NMR spectroscopy by Müller et al. who monitored the chemical shift of the two triazole protons.[5] We carried out similar experiments using the full **ON3**/**ON4** complex: titrating increasing quantities of $AgNO_3$ into a solution containing both oligonucleotides and examining the effects on the chemical shifts of peaks found in the aromatic region of ^{1}H NMR spectra.

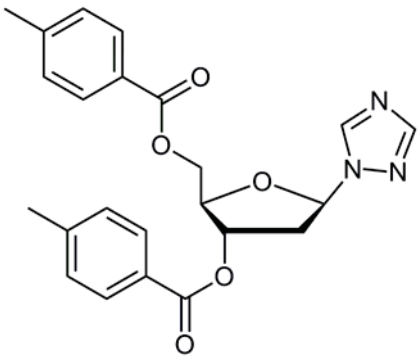


Figure 2.15. Toluoyl-protected 1,2,4-triazole nucleoside.

The chemical shift of the two protons can be found at 8.4 ppm in the free 1,2,4-triazole ring, and approximately 8.1 and 8.65 ppm when incorporated into a nucleoside; this region of the ^{1}H NMR spectra for an oligonucleotide is heavily populated – a consequence of the large

number of aromatic protons found in nucleotides. Making an exact assignment of the triazole protons is difficult without an in-depth structural study of these oligonucleotide complexes. The chemical shift of the free nucleosides is also expected to differ from that of the nucleotides within an assembly. However, full silver-ion occupancy of triazole base-pairs is expected to result in a change in chemical shift of ~0.15 ppm: triazole protons should be revealed through their gradual migration upon titration of increasing quantities of silver nitrate.

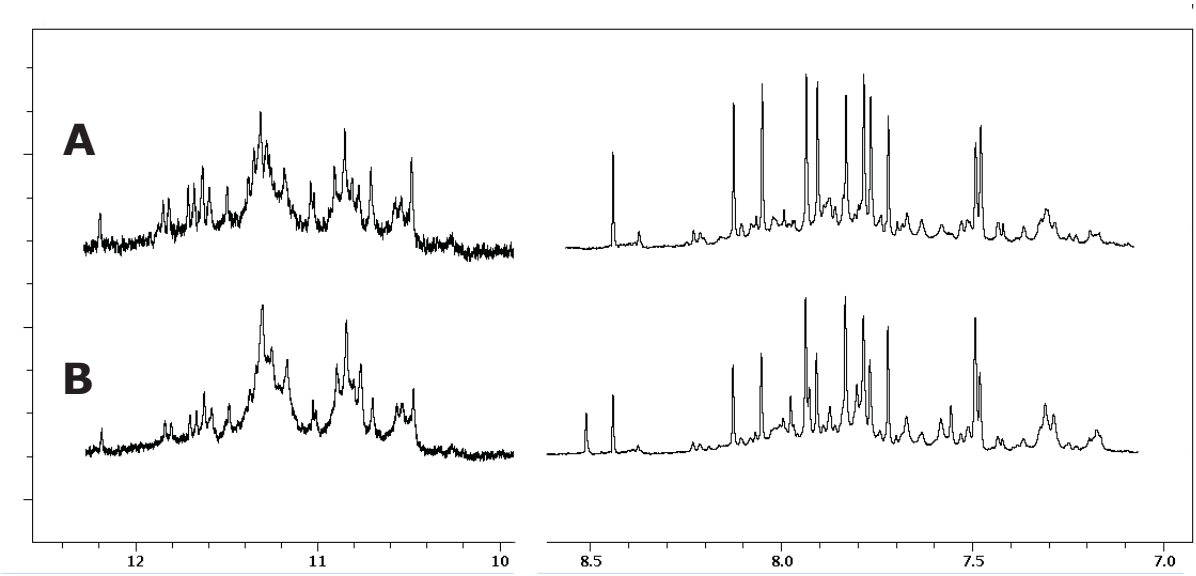


Figure 2.16. ^{1}H NMR spectra of **ON3** in 10 mM sodium phosphate buffer (pH 7.0), A, and **ON3**/**ON4** in 10 mM sodium phosphate buffer (pH 7.0) with 100 mM $NaClO_4$, B. The imino region from 10 - 12.5 ppm has been magnified relative to the aromatic region from 7 - 8.5 ppm. Samples contain 10 % D_2O and 5 μM trimethylsilyl propanoic acid (TSP) as an internal standard.

We first recorded spectra of **ON3** in 10 mM sodium phosphate buffer (pH 7.0) with 10 % D_2O, Figure 2.16A. Several peaks were present from 10-12 ppm, corresponding to the imino protons of guanine. Imino protons exchange rapidly with deuterium causing a broadening, and possibly the complete disappearance of peaks for protons exposed to the bulk solvent. The presence of peaks indicates that the oligonucleotide has assumed an ordered structure, shielding the imino protons. ^{1}H NMR spectra of a solution containing both **ON3** and **ON4** in 10 mM sodium phosphate buffer (pH 7.0), 100 mM $NaClO_4$ with 10 % D_2O, Figure 2.16B, shows no significant differences from that of **ON3** alone: peaks in the imino region show identical chemical shift, albeit with a loss in resolution as a consequence of the higher ionic strength of the sample. The aromatic region of the spectrum appears to be the sum of the spectra for the individual oligonucleotides. These spectra provide no evidence of interaction between **ON3** and **ON4** – consistent with the results of PAGE, CD and UV experiments.

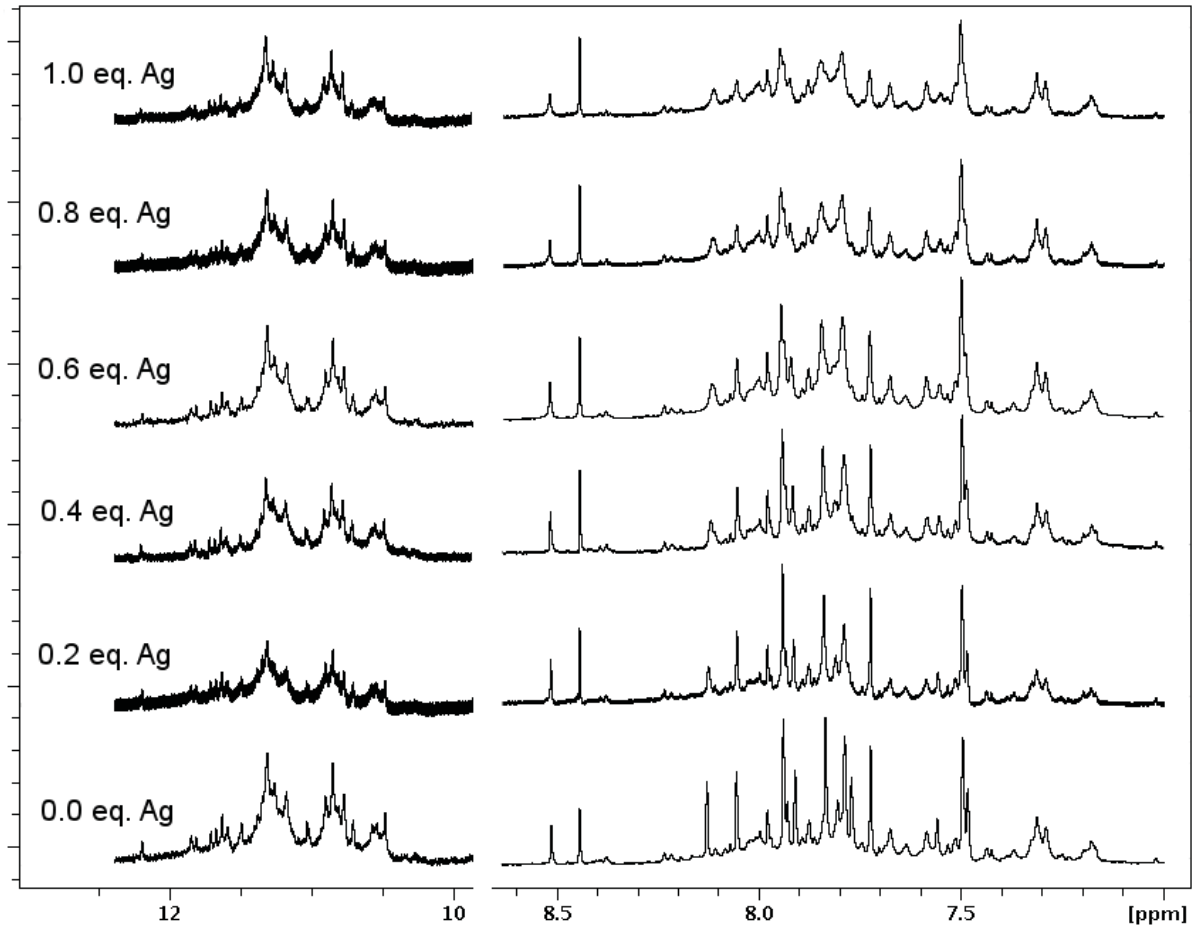


Figure 2.17. ^{1}H NMR spectra of the titration of **ON3** and **ON4** with $AgNO_3$. Samples contain 10 mM sodium phosphate buffer (pH 7.0), 100 mM $NaClO_4$, 10 % D_2O and 5 μM TSP as an internal standard. The imino region from 10 - 12.5 ppm has been magnified relative to the aromatic region from 7 - 8.5 ppm.

A solution of **ON3** and **ON4** in 10 mM sodium phosphate buffer, 100 mM $NaClO_4$ with 10 % D_2O was titrated with increments of 0.2 eq. of silver(I) nitrate. Resulting spectra are detailed in Figure 2.17. Some line broadening is apparent, particularly for spectra containing 0.4 – 1.0 eq. Ag^+, due to an earthquake-related reduction in shimming quality. The peaks in the aromatic region of these spectra show no changes in chemical shift with increasing silver concentration, suggesting that there is no specific triazole interaction with silver. The imino regions are also similar, indicating that there are no significant changes in the conformation of the assembly, providing even further evidence against the ability of silver to promote the formation of an **ON3**/**ON4** G-quadruplex.

2.4 Conclusion

Our experiments with the parent sequences are consistent with literature observations that an asymmetric 3+1 G-quadruplex is formed between **ON1** and **ON2**, this conclusion is supported by PAGE experiments and UV & CD spectroscopy. We had expected that the core, three G-quartet quadruplex would maintain a large degree of stability in the absence of the

A/T Watson-Crick base pair, thereby providing a high degree of preorganisation for the formation of a 1,2,4-triazole-metal base pair. However, the data indicate that the replacement of the Watson-Crick base-pair-forming nucleotides T11 of **ON1** and A2 of **ON2** with nucleotides containing 1,2,4-triazole (forming **ON3** and **ON4**, respectively) results in the complete destabilisation of hybrid assembly, eliminating much of the basis for our design. This destabilisation of the hybrid assembly removes any preorganisation benefit that could result from bringing the two triazole bases into close proximity. It is not surprising, then, that the stability is not recovered by the addition of silver(I) or mercury(II) ions – there is no evidence for the formation of a triazole-metal-triazole base pair nor evidence to suggest that either of these transition metals coordinate to triazole nucleotides in non-base-pair scenarios.

The destabilisation of the 3+1 structure by the removal of the Watson-Crick A/T base pair highlights its importance in the assembly of this G-quadruplex. This observation has not been noted in the literature. Our evidence suggests that the A/T base pair is also vital for the assembly of the **ON1**/**ON1** G-quadruplex, and that in its absence a G-triplex structure is probably the preferred conformation. Further structural characterisation is needed to confirm this.

The 3+1 G-quadruplex described in this chapter provided an interesting complex to investigate metal-mediated base pair formation and provided a conservative system with which to test base pair formation. These triazole nucleotides had a demonstrated ability to coordinate transition metals and molecular modelling had suggested that the position of the base pair within the quadruplex assembly would approximate the conformation within a duplex. Although we were unable to observe metal-coordination under the conditions tested, the behaviour of this 3+1 G-quadruplex system makes it an attractive model for further study of metal-coordinating nucleotides as any metal-dependent stabilisation is likely to have an obvious impact on the assembly of this structure.

2.5 References

(1) Wang, Y.; Patel, D. J.: Solution structure of the Tetrahymena telomeric repeat d(T2G4)4 G-tetraplex. *Structure* **1994**, *2*, 1141-1156.

(2) Luu, K. N.; Tu, A.; Kuryavyi, V.; Lacroix, L.; Patel, D. J.: Structure of the Human Telomere in K + Solution : An Intramolecular (3 + 1) G-Quadruplex Scaffold Favoring a Major G-Quadruplex Conformation for Struc. *J. Am. Chem. Soc.* **2006**, 9963-9970.

(3) Zhang, N.; Phan, A. T.; Patel, D. J.: (3 + 1) Assembly of Three Human Telomeric Repeats into an Asymmetric Dimeric G-Quadruplex. *J. Am. Chem. Soc* **2005**, *2005*, 17277-17285.

(4) Dai, J.; Dexheimer, T. S.; Chen, D.; Carver, M.; Ambrus, A.; Jones, R. A.; Yang, D.: An Intramolecular G-Quadruplex Structure with Mixed Parallel/Antiparallel G-Strands Formed in the Human BCL-2 Promoter Region in Solution. *J. Am. Chem. Soc.* **2006**, *128*, 1096-1098.

(5) Muller, J.; Bohme, D.; Lax, P.; Morell Cerda , M.; Roitzsch, M.: Metal ion coordination to azole nucleosides. *Chem. Eur. J.* **2005**, *11*, 6246-53.

(6) Smith, N. M.; Amrane, S.; Rosu, F.; Gabelica, V.; Mergny, J.-L.: Mercury-thymine interaction with a chair type G-quadruplex architecture. *Chem. Commun.* **2012**, *48*, 11464-6.

(7) Johannsen, S.; Megger, N.; Bohme, D.; Sigel, R. K. O.; Muler, J.: Solution Structure of a DNA Double Helix with consecutive metal-mediated base pairs. *Nat. Chem.* **2010**, 1-6.

(8) Müller, J.; Böhme, D.; Düpre, N.; Mehring, M.; Polonius, F.-A.: Differential reactivity of α and β 2′-deoxyribonucleosides towards protonation and metalation. *J. Inorg. Biochem.* **2007**, *101*, 470-476.

(9) Paul, A.; Sengupta, P.; Krishnan, Y.; Ladame, S.: Combining G-Quadruplex Targeting Motifs on a Single Peptide Nucleic Acid Scaffold: A Hybrid (3+1) PNA–DNA Bimolecular Quadruplex. *Chem. Eur. J.* **2008**, *14*, 8682-8689.

(10) Klausmeyer, K. K.; Feazell, R. P.; Reibenspies, J. H.: Two-, Three-, and Four-Coordinate Ag(I) Coordination Polymers Formed by the Novel Phosphinite PPh2(3-OCH2C5H4N). *Inorg. Chem.* **2004**, *43*, 1130-1136.

(11) Richens, D. T.: Ligand Substitution Reactions at Inorganic Centers†. *Chem. Rev.* **2005**, *105*, 1961-2002.

(12) Koirala, D.; Mashimo, T.; Sannohe, Y.; Yu, Z.; Mao, H.; Sugiyama, H.: Intramolecular folding in three tandem guanine repeats of human telomeric DNA. *Chem. Commun.* **2012**, *48*, 2006-2008.

(13) Limongelli, V.; De Tito, S.; Cerofolini, L.; Fragai, M.; Pagano, B.; Trotta, R.; Cosconati, S.; Marinelli, L.; Novellino, E.; Bertini, I.; Randazzo, A.; Luchinat, C.; Parrinello, M.: The G-Triplex DNA. *Angew. Chem. Int. Ed.* **2013**, *52*, 2269-2273.

3 Two- and Four-Repeat *htel* Sequences Modified with 1,2,4-Triazole Nucleotides

3.1 Introduction

The human telomeric sequence, *htel*, is capable of assuming a large number of different conformations depending on DNA concentration, solvent, cations present etc. Telomeric DNA is present in cells as long strands of approximately 200 nucleotides at the ends of chromosomes.[1,2] Such large DNA strands are difficult to work with experimentally and so research typically focuses on smaller, representative oligonucleotides usually no longer than four repeats of the *htel* sequence. Shorter oligonucleotides containing from one to three *htel* repeats form a diverse assortment of multimeric structures that are sometimes difficult to relate back to potential *in vivo* structures. Oligonucleotides containing >4 *htel* repeats are likely to provide structures more representative of native telomeric G-quadruplexes.

With any telomeric sequence care must be taken as the exact sequence used can play a dramatic role in determining the final structure. Because of the repetitive nature of the telomeric sequence, the basic unit is variously described as $(TTAGGG)_n$, $(GGGTTA)_n$, $(TAGGGT)_n$ etc. The three htel repeat, $d(G_3(TTAGGG)_2T)$, described by Anh Tuân Phan and coworkers has been discussed in detail in Chapter 2.[3] This chapter focuses on two different assemblies containing two and four htel repeats, each of which possess an interesting equilibrium between two different conformations. In particular, we sought to examine the influence of triazole nucleotides on the equilibria present in these systems, and on the formation of higher-order quaternary assemblies.

3.1.1 G-Quadruplexes Containing Four Repeats of the *htel* Sequence

The NMR structure of the four *htel* repeat sequence AGGGTTAGGGTTAGGGTTAGGG was determined twenty years ago and found to interchange between parallel and antiparallel conformations, Figure 3.1A and C.[4] The more stable parallel conformation dominates in solution where K^+ ions are present while the antiparallel conformation is favoured in Na^+-containing solution.

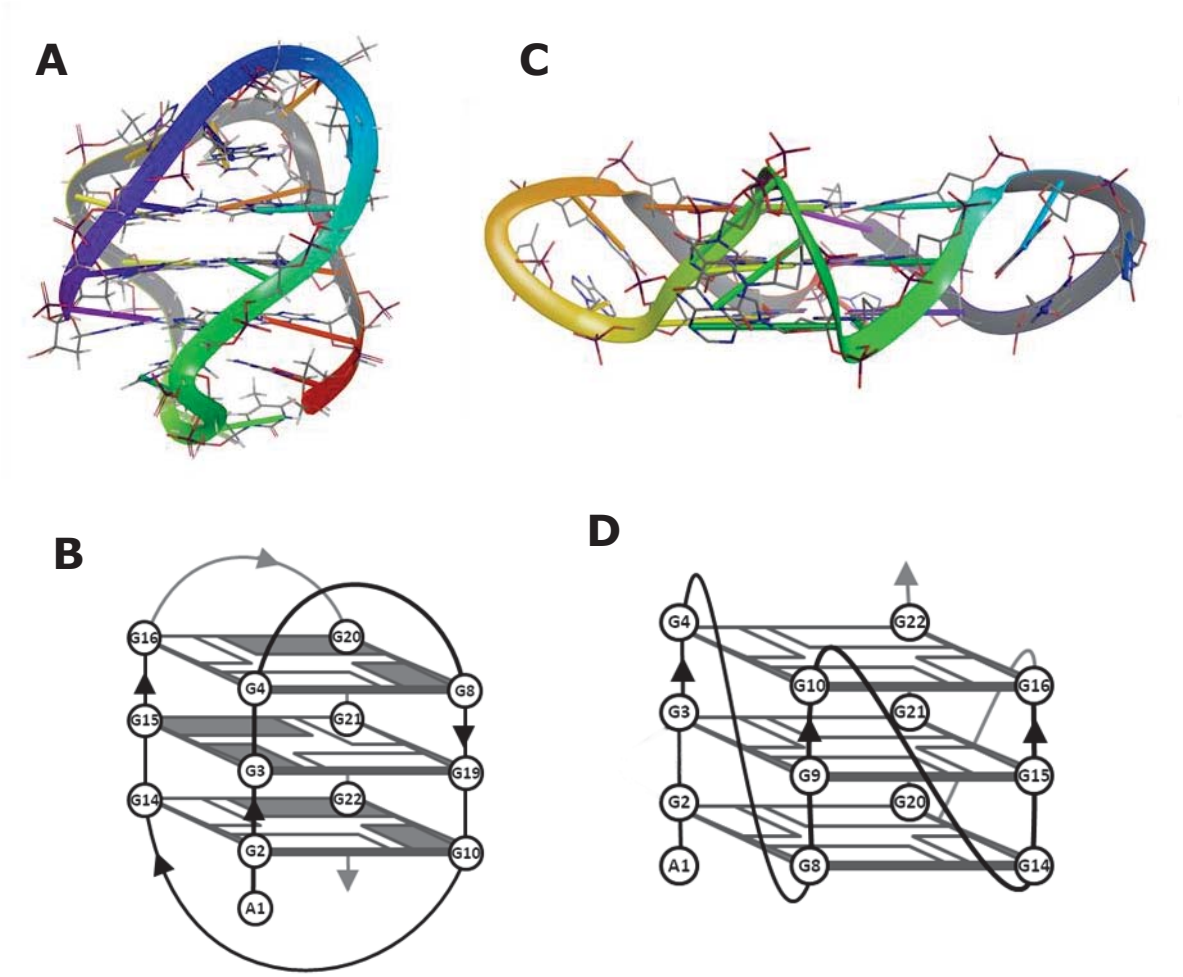


Figure 3.1. G-quadruplexes formed from the 4*htel* sequence, $AG_3T_2AG_3T_2AG_3T_2AG_3$. Antiparallel structure formed in Na^+ solution, as determined by NMR, A,[4] and shown schematically, B. Parallel structure formed in K^+ solution, as determined by X-ray crystallography, C,[5] and shown schematically, D.

A notable feature of these two conformations is that the 5′ and 3′ ends of the oligonucleotide change position relative to each other. In the parallel structure the termini are found on opposite sides of the G-quadruplex, Figure 3.1C. In the antiparallel arrangement they occupy opposite corners on the same quadruplex face, Figure 3.1B. This presents an interesting system to investigate 5′ and 3′ 1,2,4-triazole nucleotides: the different positioning of these nucleotides has the potential to give rise to high-order assembly, discussed below, and may influence the conformation assumed by these G-quadruplexes. The 4-*htel* oligonucleotides studied are detailed in Table 3.1.

Table 3.1. Oligonucleotides containing four *htel* repeats and terminal 1,2,4-triazole nucleotides.

	Sequence
ON1	**T^r**AGGGTTAGGGTTAGGGTTAGGG
ON2	**T^r**AGGGTTAGGGTTAGGGTTAGGG**T^r**

ON1 possesses a single 5′ triazole base whereas **ON2** is modified at both the 5′ and 3′ ends. Each triazole nucleotide is included as an addition to the native sequence, rather than replacing nucleotides, to minimise any structural changes that might be induced by their presence. These modifications coupled with the equilibrium between parallel and antiparallel structures give rise to four possible scenarios, each of which can give rise to unique metal-coordinating effects.

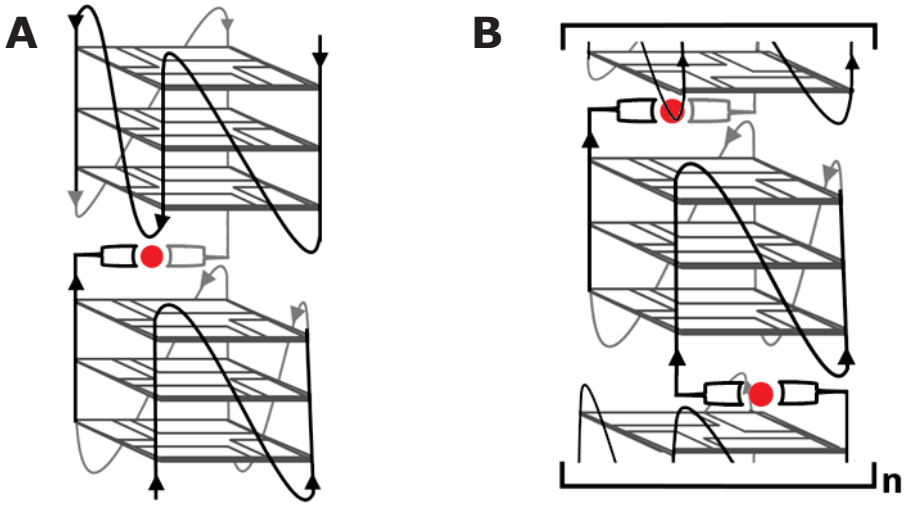


Figure 3.2. Potential quaternary G-quadruplex assemblies formed from **ON1** and **ON2**. The formation of a metal-ion triazole base pair between separate assemblies may promote the formation of G-quadruplex dimers from **ON1**, A, or polymers from **ON2**, B.

Considering the schematic diagrams and structural information depicted in Figure 3.1, the parallel conformation has a compact structure and exposed G-quartets that make the quadruplex susceptible to the formation of high-order quaternary assemblies. A single metal-coordinating nucleotide located at the 5′ end of the oligonucleotide sequence may serve to encourage the formation of a tail-to-tail dimer, Figure 3.2A. Similarly, modifications at both the 5′ and 3′ ends may allow the formation of polymers of G-quadruplexes, as in Figure 3.2B. Polymers of this type may be considered as model systems for telomeres *in vivo*. Currently it is not possible to create extended arrays of this type using conventional oligonucleotides such as $(TTAGGG)_n$ (where 'n' represents an arbitrarily large number) because there is no way to ensure the regular positioning of quadruplexes along the polymer.

The conformation assumed by the 4*htel* oligonucleotide in Na^+ solution features two lateral loops on one side, and a diagonal loop on the opposite face, these loops cover the surface of the terminal quartets, as can be seen in Figure 3.1B, and prevent any hydrophobic stacking between separate G-quadruplex assemblies, making the formation of high-order structures unlikely. The addition of both a 5′ and 3′ triazole nucleotide creates an interesting system: both nucleotides can be expected to be located on the same face of the quadruplex when assembled in Na^+ solution; however, the presence of the diagonal loop is likely to prevent any direct interactions between two triazole bases on the same G-quadruplex. One possibility is that the triazole bases may serve to stabilise the parallel conformation: examination of the parallel structure reveals that the 5′ and 3′ termini are located adjacent to each other on one corner of the quadruplex, more clearly seen in Figure 3.3. The formation of an intramolecular triazole base pair in this position could be expected to shift the equilibrium between the two conformations in favour of the parallel structure.

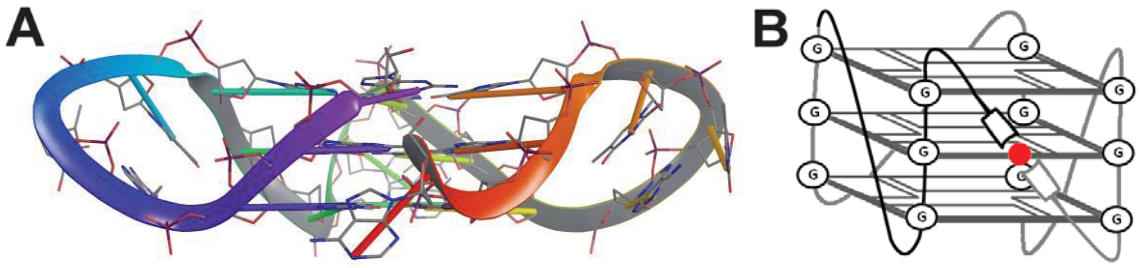


Figure 3.3. Parallel G-quadruplex formed from the 4*htel* sequence, A. The 5′ and 3′ termini (coloured purple and red, respectively) are located on the same 'corner' of the assembly, possibly allowing **ON2** to form a triazole base pair as a diagonal loop, drawn schematically, B.

3.1.2 G-Quadruplexes Containing Two Repeats of the *htel* Sequence

As a further example of the diversity of G-quadruplexes that can be formed from the human telomeric sequence, a sequence containing two *htel* repeats, TAGGGTTAGGGT, has been found to form two different bimolecular quadruplexes, shown in Figure 3.4.[6] Base substitution studies indicated that the antiparallel form predominates at low temperatures while the parallel form becomes the major component above 60 °C, and is also favoured in 90 mM aqueous KCl solution.

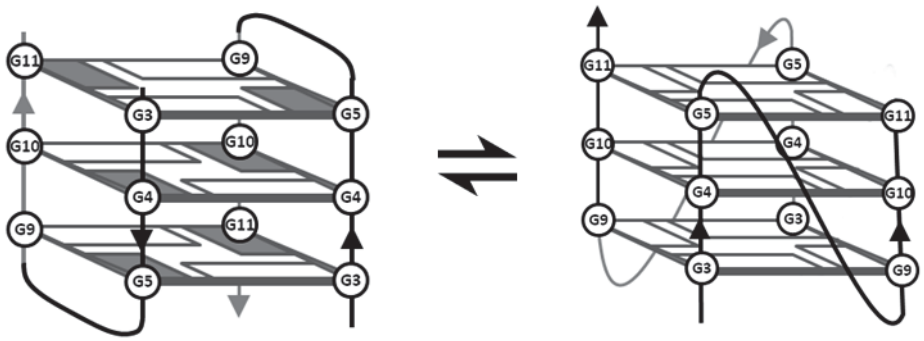


Figure 3.4. Bimolecular G-quadruplexes topologies formed from the the 2*htel* sequence $TAG_3T_2AG_3T$. Antiparallel, left, and parallel, right.

As with the 4*htel* DNA, we designed two sequences, listed in Table 3.2, to investigate the effect of including a triazole nucleotide at the 5′ and 3′ ends of the 2*htel* oligonucleotide.

Table 3.2. Oligonucleotides containing two *htel* repeats and terminal 1,2,4-triazole nucleotides.

	Sequence
ON3	**T**^{r}AGGGTTAGGGT
ON4	**T**^{r}AGGGTTAGGG**T**^{r}

The absence of the diagonal loop found in the 4*htel* structure makes the formation of a triazole-metal-triazole base pair across the surface of the terminal G-quartets more probable. We elected to replace the terminal T1 and T13 of the parent sequence in order to position the triazole bases closer to the surface of the G-quadruplex.

An antiparallel assembly formed from **ON3** would present one triazole base at each end of the G-quadruplex, a situation comparable to that of the parallel quadruplex formed from **ON2** and so there exists the possibility to form a similar polymer of G-quadruplexes, as in Figure 3.2B. A parallel quadruplex formed from **ON3** would have both triazole nucleotides diagonally opposite each other on the surface of the same quartet which could lead to the formation of a triazole base pair.

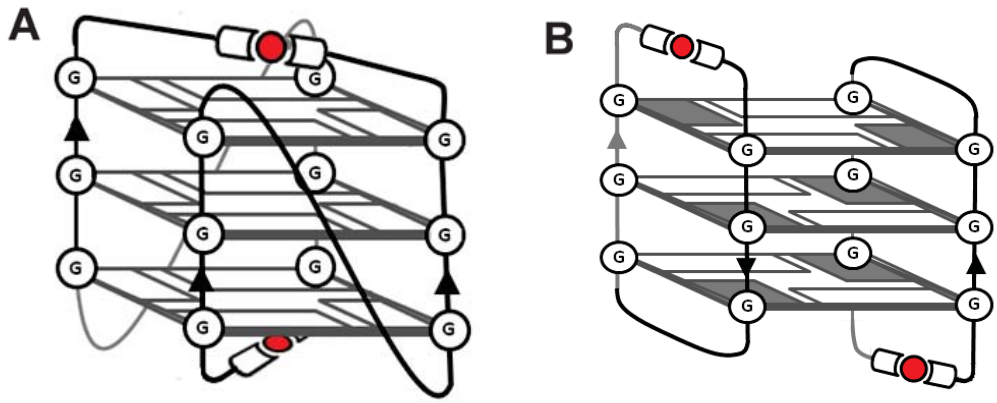


Figure 3.5. Potential metal-coordinating assemblies formed from **ON4**. Antiparallel conformation with diagonal triazole base pairs, A. Parallel conformation with edgewise triazole base pairs, B.

In contrast, an antiparallel quadruplex formed from **ON4** presents the triazole bases on the same edge of the quartet, Figure 3.5B. In this case, two base pairs would be present: one on each side of the assembly. A parallel arrangement formed from **ON4**, Figure 3.5A**,** allows the possibility of two diagonal base-pairs on the surface of each of the terminal G-quartets and, again, the ability to form a polymer of G-quadruplexes.

3.2 Results and discussion

3.2.1 Examination of Metal Ion Coordination by PAGE

The 1,2,4-triazole nucleotide was synthesised as detailed in the experimental section (Chapter 7). **ON1 – 4** were synthesised according to standard oligonucleotide solid phase synthesis methods and were characterised by MALDI-TOF mass spectrometry. Further experimental details can be found in Chapter 7.

Using PAGE, we screened **ON1 – 4** with a range of s- and d-block cations in solutions containing Na^+ and K^+ salts. Cations were present at 50 μM for **ON1** and **ON3** and at 100 μM for **ON2** and **ON4:** half the triazole nucleotide concentration, a level selected to allow the assembly of triazole-metal-triazole base pairs while minimising destabilisation of the G-quadruplex core via metal coordination to guanine. Several examples exist in the literature of silver coordination to 1,2,4-triazole nucleotides, and of less stable mercury coordination.[7] To the best of our knowledge, there have been no other reports of these nucleotides coordinating other metal ions. However, there are examples of metal-(1,2,4-triazole)$_2$ complexes being formed with Co(II), Ni(II), Cu(II), and Zn(II).[8]

For **ON1,** G-quadruplexes formed in Na^+ solution appeared to be marginally stable, Figure 3.6. Diffuse bands on the gel indicate the presence of a mixture of partially folded assemblies. No clear cation-specific response was observed under these conditions. G-quadruplexes formed in K^+ conditions formed bands with better definition indicating improved stability. A small gel-shift was observed with Ag^+ (lane 2, Figure 3.7). This shift in mobility is too small to result from G-quadruplex dimerisation and may be the result of either silver coordination to the triazole nucleotide, or a conformation change in the quadruplex itself.

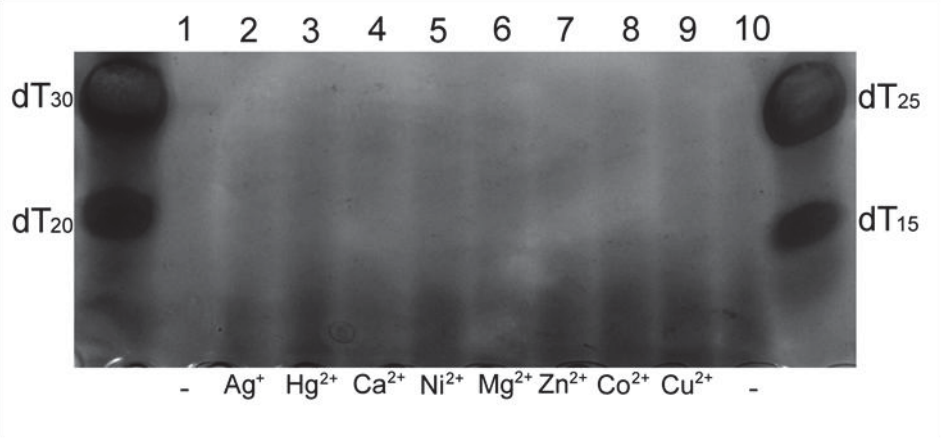


Figure 3.6. Native PAGE of **ON1** in Na^+ solution. Samples contain 100 μM DNA, 10 mM lithium cacodylate buffer (pH 7.2) and 100 mM $NaNO_3$. Metal ions included at 50 μM as their nitrate salts. Gel contains 12% acrylamide, TB buffer (pH 8.0) supplemented with 100 mM $NaNO_3$. Electrophoresis carried out at 18 W, r.t. for ~3 hours.

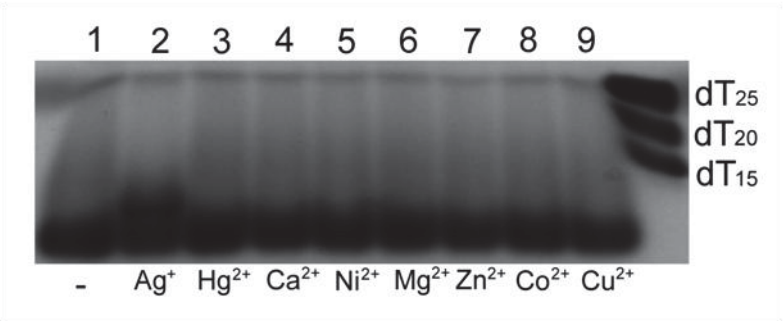


Figure 3.7. Native PAGE of **ON1** in K^+. Samples contain 100 μM DNA, 10 mM lithium cacodylate buffer (pH 7.2) and 100 mM KNO_3. Metal ions included at 50 μM as their nitrate salts. Gel contains 12% acrylamide, TB buffer (pH 8.0) supplemented with 100 mM KNO_3. Electrophoresis carried out at 18 W, r.t. for ~3 hours.

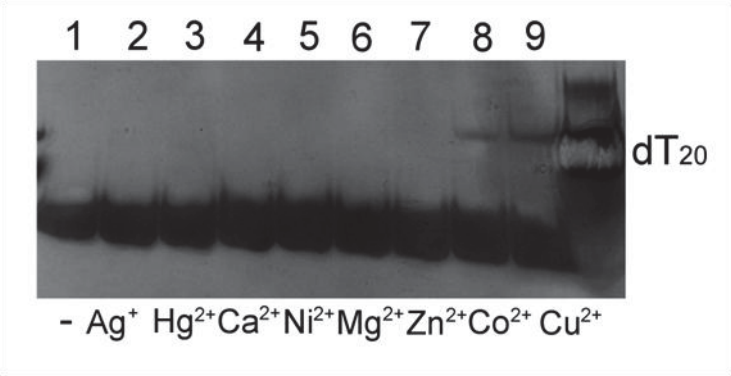


Figure 3.8. Native PAGE of **ON2** in Na^+. Samples contain 100 μM DNA, 10 mM lithium cacodylate buffer (pH 7.2) and 100 mM $NaNO_3$. Metal ions included at 50 μM as their nitrate salts. Gel contains 12% acrylamide, TB buffer (pH 8.0) supplemented with 100 mM $NaNO_3$. Electrophoresis carried out at 18 W, r.t. for ~3 hours. The low mobility bands observed in lanes 8 and 9 are caused by residual loading dye.

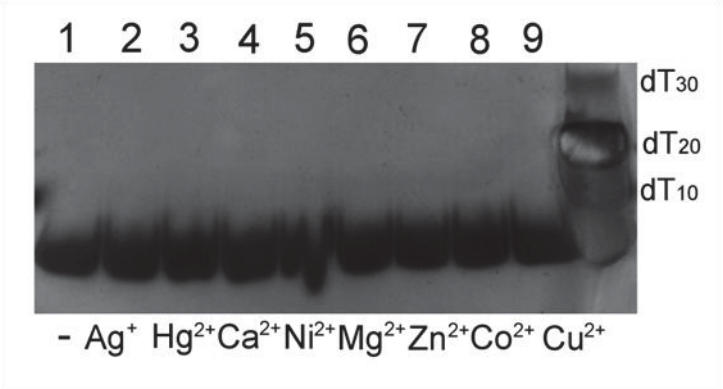


Figure 3.9. Native PAGE of **ON2** in K^+. Samples contain 100 μM DNA, 10 mM lithium cacodylate buffer (pH 7.2) and 100 mM KNO_3. Metal ions included at 50 μM as their nitrate salts. Gel contains 12% acrylamide, TB buffer (pH 8.0) supplemented with 100 mM KNO_3. Electrophoresis carried out at 18 W, r.t. for ~3 hours.

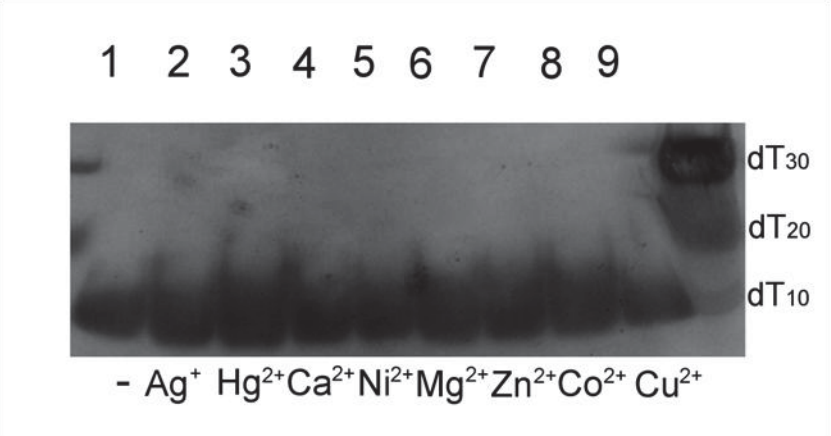


Figure 3.10. Native PAGE of **ON3** in Na^+. Samples contain 100 μM DNA, 10 mM lithium cacodylate buffer (pH 7.2) and 100 mM $NaNO_3$. Metal ions included at 50 μM as their nitrate salts. Gel contains 12% acrylamide, TB buffer (pH 8.0) supplemented with 100 mM $NaNO_3$. Electrophoresis carried out at 18 W, r.t. for ~3 hours.

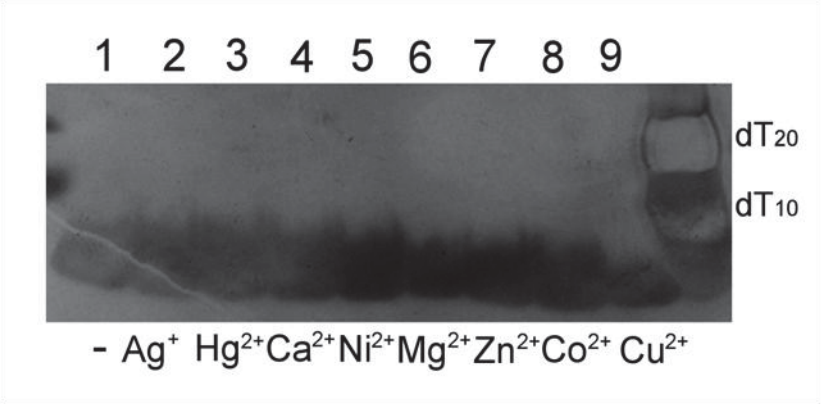


Figure 3.11. Native PAGE of **ON3** in K^+. Samples contain 100 μM DNA, 10 mM lithium cacodylate buffer (pH 7.2) and 100 mM KNO_3. Metal ions included at 50 μM as their nitrate salts. Gel contains 12% acrylamide, TB buffer (pH 8.0) supplemented with 100 mM KNO_3. Electrophoresis carried out at 18 W, r.t. for ~3 hours.

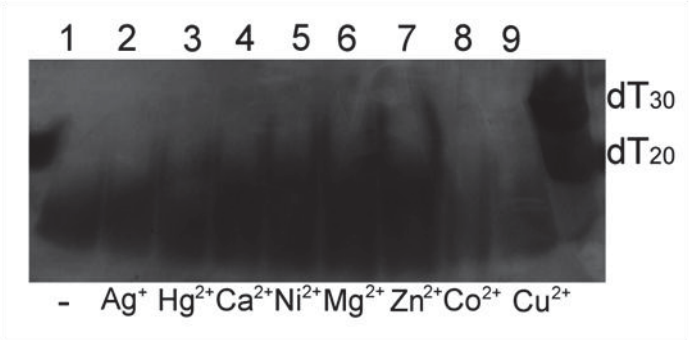


Figure 3.12. Native PAGE of **ON4** in Na^+. Samples contain 100 μM DNA, 10 mM lithium cacodylate buffer (pH 7.2) and 100 mM $NaNO_3$. Metal ions included at 50 μM as their nitrate salts. Gel contains 12% acrylamide, TB buffer (pH 8.0) supplemented with 100 mM $NaNO_3$. Electrophoresis carried out at 18 W, r.t. for ~3 hours. The increased size of bands near the centre of the image appears to have been caused by inconsistencies within the gel itself.

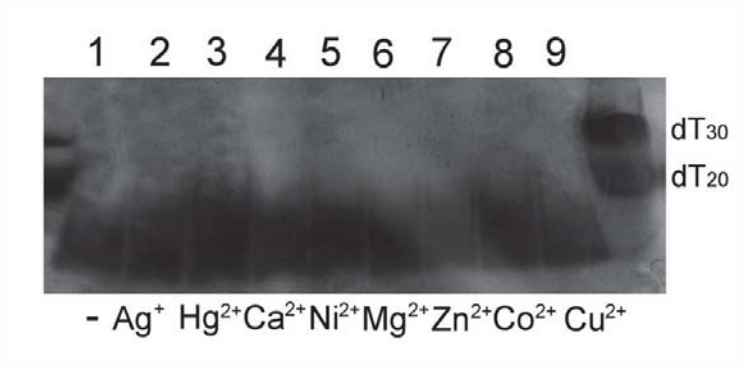


Figure 3.13. Native PAGE of **ON4** in K^+. Samples contain 100 μM DNA, 10 mM lithium cacodylate buffer (pH 7.2) and 100 mM KNO_3. Metal ions included at 50 μM as their nitrate salts. Gel contains 12% acrylamide, TB buffer (pH 8.0) supplemented with 100 mM KNO_3. Electrophoresis carried out at 18 W, r.t. for ~3 hours.

Inconsistencies in preparation and visualisation can cause variance in the appearance of some bands in the gel. With consideration of these artefacts, we determined that there were no significant cation-dependent changes for **ON2 – 4** in either Na^+ or K^+ conditions, as seen in Figure 3.8 – Figure 3.11. It is difficult to rationalise why **ON1** would show a response to silver ions, but not **ON2,** which presumably shares a similar conformation. The only expected difference between these two oligonucleotides is the presence of an extra 3′ triazole nucleotide in **ON2**, being remote from each other, the 5′ and 3′ triazole nucleotides should not interact: if the 5′ triazole of **ON1** can coordinate silver, so should that of **ON2**. One explanation may be that the introduction of the 3′ triazole nucleotide of **ON2** introduces a structural change that reduces its ability to coordinate metal ions.

3.2.2 Determination of Conformation by CD Spectroscopy

The oligonucleotides studied in this chapter are capable of forming complex systems containing a number of different conformations. We used circular dichroism spectroscopy to identify the most favoured conformation in Na^+ and K^+ solutions for each oligonucleotide. Considering the Ag^+-dependent gel shift observed for **ON1**, we also examined the influence of silver ions on each of these assemblies. It should be noted that the formation of the higher-order structures, suggested above, is not likely to be measureable using CD spectroscopy as the basic G-quadruplex 'subunit' structure is likely to remain unchanged.

3.2.2.1 CD Spectra of 4htel Oligonucleotides ON1 and ON2

CD spectra for **ON1** and **ON2** in Na^+ solution show strong similarities to each other indicating that both oligonucleotides assume the same conformation. These spectra have positive maxima at ~250 and ~295 nm with a negative minimum at ~265 nm, consistent with an antiparallel G-quadruplex, and with spectra published for the *4htel* parent sequence.[9] It can be assumed that the structure of **ON1** and **ON2** is similar to that shown in

Figure 3.1A, and schematically in Figure 3.1B. The addition of Ag^+ ions to these oligonucleotides does not result in any significant changes in spectra.

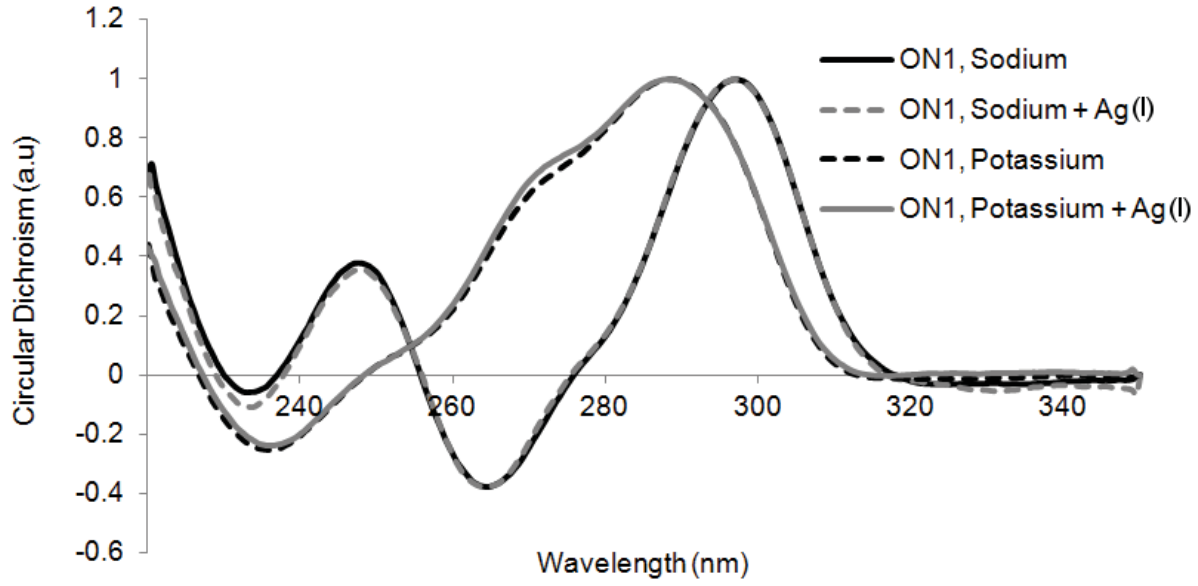


Figure 3.14. Normalised CD spectra of **ON1** in Na^+ and K^+ in the presence or absence of Ag^+. Samples contain 20 μM DNA, 10 mM lithium cacodylate buffer (pH 7.2), 100 mM of either sodium or potassium cacodylate, and 10 μM $AgNO_3$, when present.

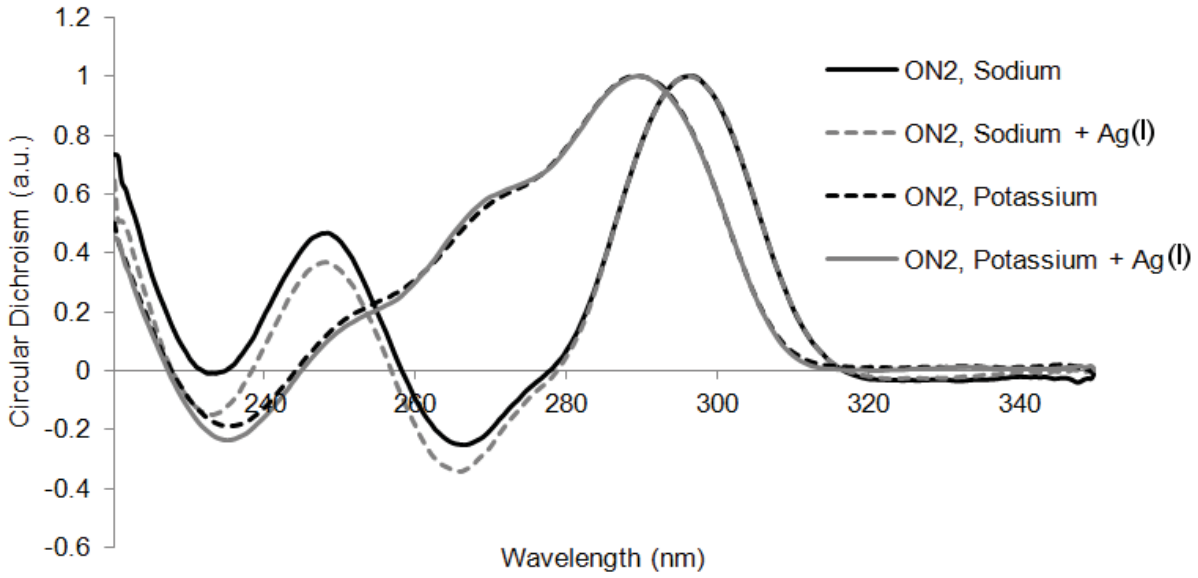


Figure 3.15. Normalised CD spectra of **ON2** in Na^+ and K^+ in the presence or absence of Ag^+. Samples contain 20 μM DNA, 10 mM lithium cacodylate buffer (pH 7.2), 100 mM of either sodium or potassium cacodylate, and 10 μM $AgNO_3$, when present.

In K^+ solution, both **ON1** and **ON2** show similar spectra that are consistent with spectra published for the parent *4htel* sequence in similar conditions,[9] where it has been suggested that these oligonucleotides form parallel G-quadruplexes with propeller loops, similar to those shown in Figure 3.1C, and schematically in Figure 3.1D. It should be noted that no

published NMR structure exists for the *4htel* in K^+ containing solution. X-ray structures have been published in these conditions,[5] but may not represent the structures assumed in the conditions used for Figure 3.14 and Figure 3.15. A positive maximum at ~290 nm is unusual for parallel quadruplexes and may indicate the presence of both parallel and antiparallel G-quadruplexes in 100 mM K^+ solution.

It can be seen that the addition of Ag^+ makes very little difference to **ON1** and **ON2** CD spectra recorded in either Na^+ or K^+ conditions. This indicates that there is no structural change in these assemblies, and in the case of **ON2** indicates that there is no change in the position of the equilibrium between parallel and antiparallel G-quadruplexes. The similarity in spectra suggests that the silver(I)-dependent gel shifts observed for **ON1** may be caused by silver-triazole coordination and not a conformation change. However, the similarity in spectra between the two oligonucleotides is problematic, as it is evidence against the explanation that differences in conformation account for the different coordination capabilities of **ON1** and **ON2**. However, as CD spectroscopy provides relatively broad information, it may still be the case that small-scale differences in conformation are responsible for the differences.

3.2.2.2 CD Spectra of 2htel Oligonucleotides ON3 and ON4

CD spectra for **ON3** and **ON4** recorded in Na^+ solution, Figure 3.16 and Figure 3.17, respectively, show spectra consistent with an antiparallel conformation with positive maxima at ~250 and ~295 nm with a negative minimum at ~265 nm. These spectra also match those recorded for the parent sequence,[9] indicating that these oligonucleotides form bimolecular antiparallel G-quadruplexes of the type shown in Figure 3.4. The addition of Ag^+ has some impact on **ON3** and **ON4**, although the structure retains an antiparallel conformation. These changes may be the result of destabilisation by coordination of Ag^+ to guanine residues. In contrast, Ag^+ has no significant effect on these assemblies in K^+ solution.

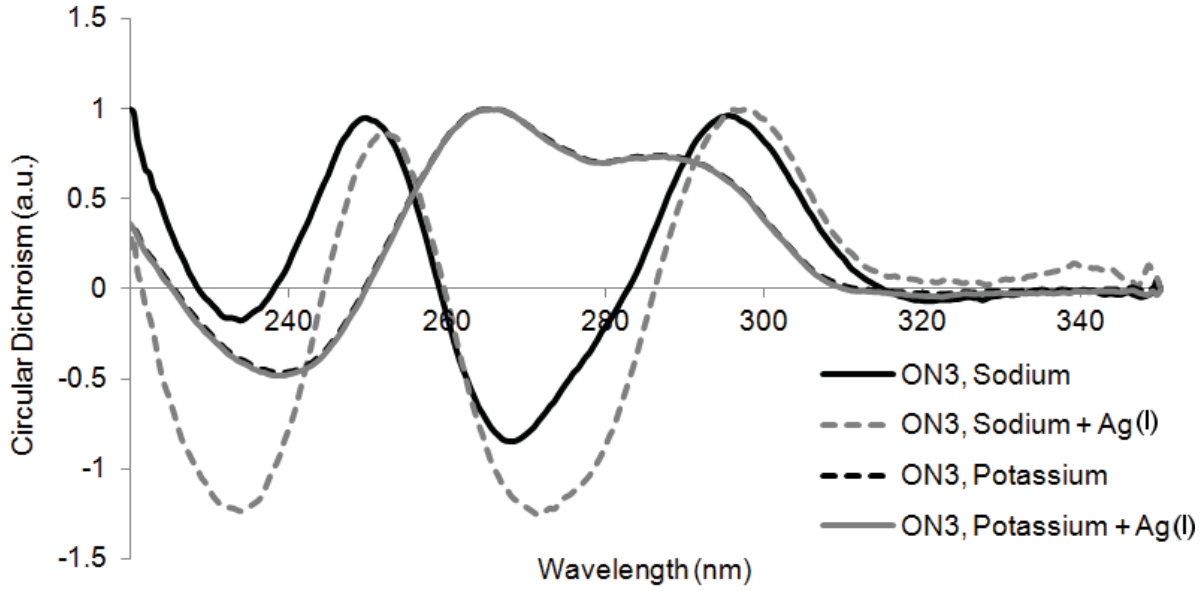


Figure 3.16. Normalised CD spectra of **ON3** in Na^+ and K^+ in the presence or absence of Ag^+. Samples contain 20 μM DNA, 10 mM lithium cacodylate buffer (pH 7.2), 100 mM of either sodium or potassium cacodylate, and 10 μM $AgNO_3$, when present.

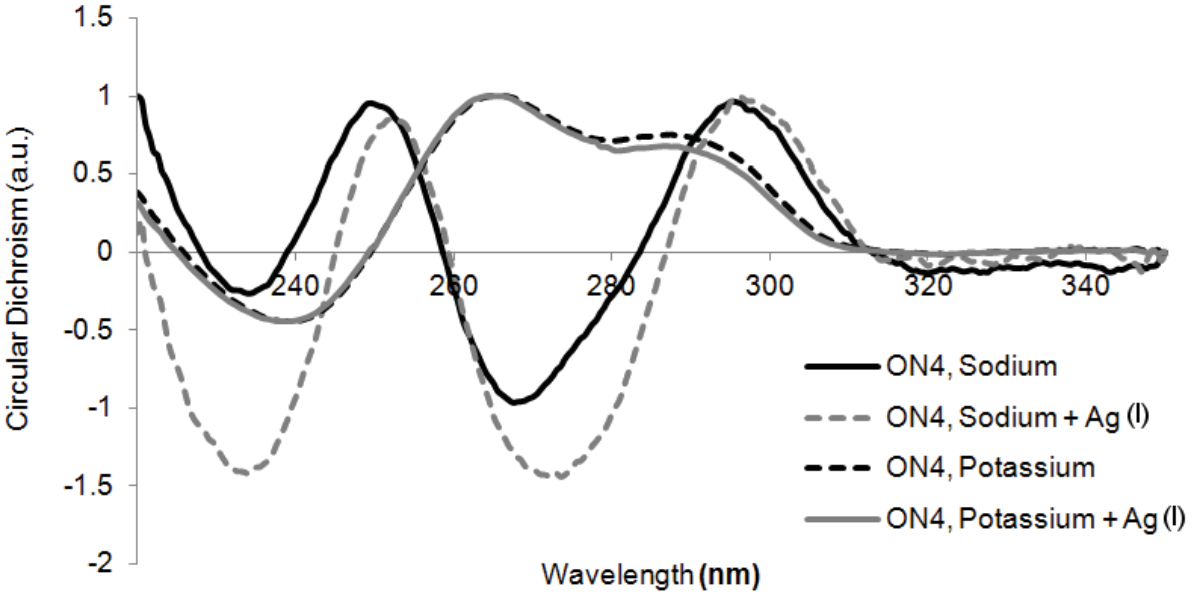


Figure 3.17. Normalised CD spectra of **ON4** in Na^+ and K^+ in the presence or absence of Ag^+. Samples contain 20 μM DNA, 10 mM lithium cacodylate buffer (pH 7.2), 100 mM of either sodium or potassium cacodylate, and 10 μM $AgNO_3$, when present.

In K^+ solution **ON3** and **ON4** show complex spectra with a negative minimum at ~240 nm and positive maxima at ~265 and ~290 nm, sharing features of both parallel and antiparallel G-quadruplexes. As mentioned above, published studies of the 2*htel* parent sequence revealed an equilibrium between the parallel and antiparallel conformations shown in Figure 3.4.[6] Literature CD experiments carried out in 100 mM KCl solution suggested that these species were present in an approximately 1:1 ratio.[9] The addition of silver(I) has no

noticeable effect on the spectra produced by **ON3** and **ON4** indicating that silver has no influence on the position of the equilibrium between the parallel and antiparallel conformations.

In summary, CD spectra confirmed that **ON1** and **ON2** form antiparallel G-quadruplexes in Na^+ solution, and parallel assemblies in K^+. **ON3** and **ON4** form antiparallel quadruplexes in Na^+, and an approximately 1:1 mixture of parallel and antiparallel assemblies in K^+.

3.2.3 Effect of Increasing Ag^+ Concentration on ON1 Assembly

Having observed a response of **ON1** to silver ions indicating that silver-triazole coordination was occurring, we investigated whether dimerisation would be encouraged by higher Ag^+ concentration. These gels were run at a lowered temperature (~3 °C) to improve the resolution of DNA separation.

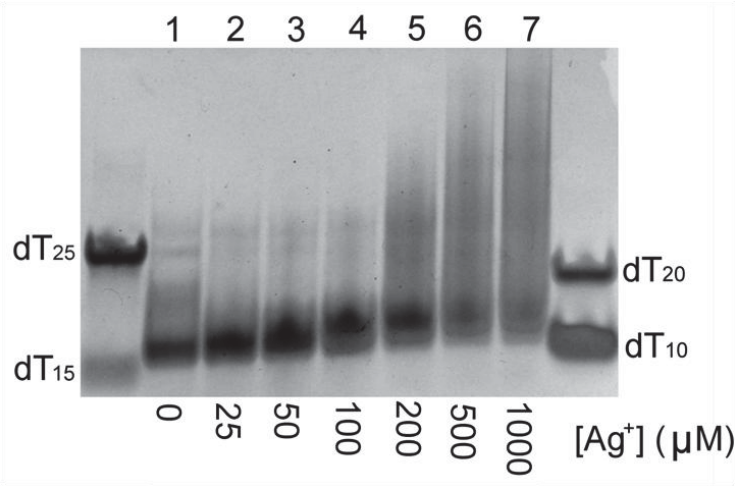


Figure 3.18. Native PAGE of **ON1** in Na^+ with increasing Ag^+ concentrations. Samples contain 100 μM DNA prepared in 10 mM lithium cacodylate buffer (pH 7.2), 100 mM $NaNO_3$, and $AgNO_3$ at the indicated concentration. Gel contains 20 % acrylamide, TB buffer (pH 8.0) supplemented with 100 mM $NaNO_3$. Electrophoresis carried out at 3°C, 15 W.

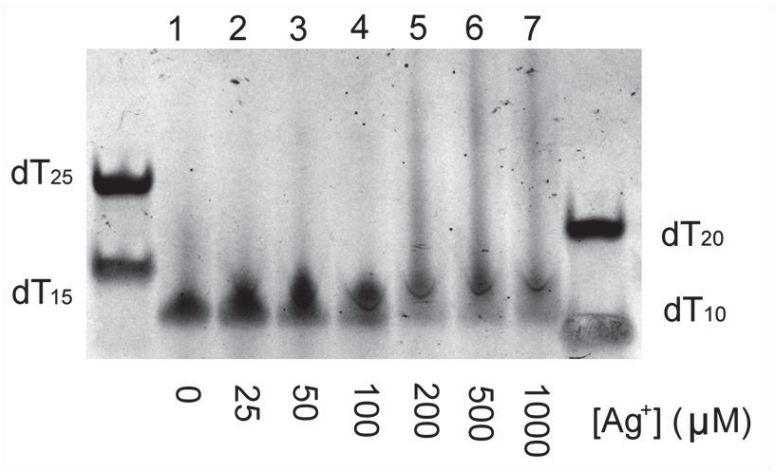


Figure 3.19. Native PAGE of **ON1** in K^+ with increasing Ag^+ concentrations. Samples contain 100 μM DNA prepared in 10 mM lithium cacodylate buffer (pH 7.2), 100 mM KNO_3, and $AgNO_3$ at the indicated concentration. Gel contains 20 % acrylamide, TB buffer (pH 8.0) supplemented with 100 mM KNO_3. Electrophoresis carried out at 3°C, 15 W.

Under these conditions a response to silver is observed for **ON1** under both Na^+, Figure 3.18, and K^+, Figure 3.19, conditions. Similar behaviour is observed in both gels – a slight gel shift is observed even at 25 μM Ag^+ which intensifies until 100 μM. At concentrations above 100 μM a smear is observed, indicating denaturation of the complex. This disruption is likely caused by coordination of silver ions to the amine groups of guanine, preventing their participation in G-quartet assembly.[10] It is interesting to note that denaturation becomes most apparent at ratios greater than 1:1 triazole nucleotide:silver ions which may indicate that the observed response is the result of coordination of a single silver ion. A specific silver-triazole interaction may prevent Ag^+ ions interfering with G-quadruplex assembly until after all triazole positions have been occupied.

3.2.4 Detection of 1,2,4-Triazole – Ag^+ Interactions by ^{1}H NMR

Using ^{1}H NMR spectroscopy we examined the specific interaction of silver ions with **ON1**. We carried out similar experiments as described in Chapter 2, by titrating increasing quantities of $AgNO_3$ into a solution containing **ON1** and examining the effects on the chemical shifts of peaks found in the aromatic region of ^{1}H NMR spectra. Triazole protons should be revealed through their gradual migration upon titration of increasing quantities of silver nitrate.

The following spectra were recorded in 100 mM Na^+ solution. We attempted to carry out similar titration experiments in K^+ solution, however the apparent NMR spectra were of poor quality. In K^+ containing conditions these oligonucleotides form parallel G-quadruplexes with exposed terminal G-quartets. It may be the case that the high DNA concentrations necessary for NMR experiments resulted in the aggregation of higher-order quadruplex structures, resulting in a decreased experimental sensitivity.

A solution of each oligonucleotide in 10 mM Na_2HPO_4 buffer, 100 mM $NaClO_4$ with 10 % D_2O, was titrated with increments of 0.2 eq. of silver(I) nitrate. The resulting spectra are detailed below. Some line broadening is apparent, particularly for spectra containing 0.4 – 1.0 eq. Ag^+, due to an earthquake-related reduction in shimming quality.

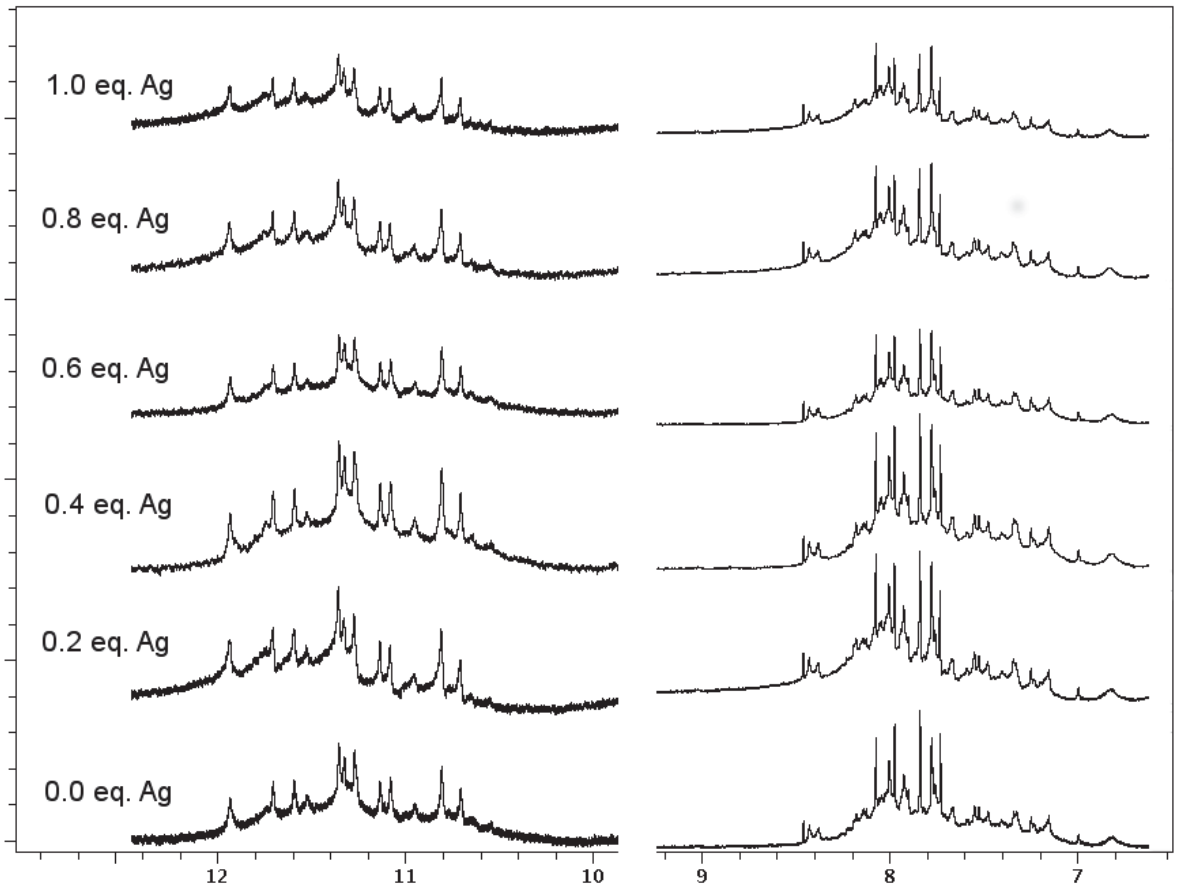


Figure 3.20. ^{1}H NMR spectra of the titration of **ON1** with $AgNO_3$. Samples contain 10 mM sodium phosphate buffer (pH 7.0), 100 mM $NaClO_4$, 10 % D_2O and 5 μM TSP as an internal standard. The imino region from 10 - 12.5 ppm has been magnified relative to the aromatic region from 7 - 8.5 ppm.

Interestingly, after titration of **ON1** with Ag^+ ions we observed no notable changes in NMR spectra. PAGE data using variable Ag^+ concentration had suggested a specific silver-oligonucleotide interaction, which should have an impact on the aromatic region of the NMR spectrum, but no changes in peak position were observed. This might indicate that the **ON1**/silver interaction is not the result of a specific interaction with the 1,2,4-triazole nucleotide. It may be the case, then, that there is a small-scale difference in conformation between **ON1** and **ON2** that explains this difference in behaviour. This may be supported by an examination of the ^{1}H NMR spectrum for **ON2**, Figure 3.21, which is similar to **ON1** in the imino region of the spectrum, indicating the central G-quadruplex core is conserved, and thus the larger structure is similar. However, there are a greater number of differences in the aromatic region, indicative of smaller-scale disparity in the assemblies.

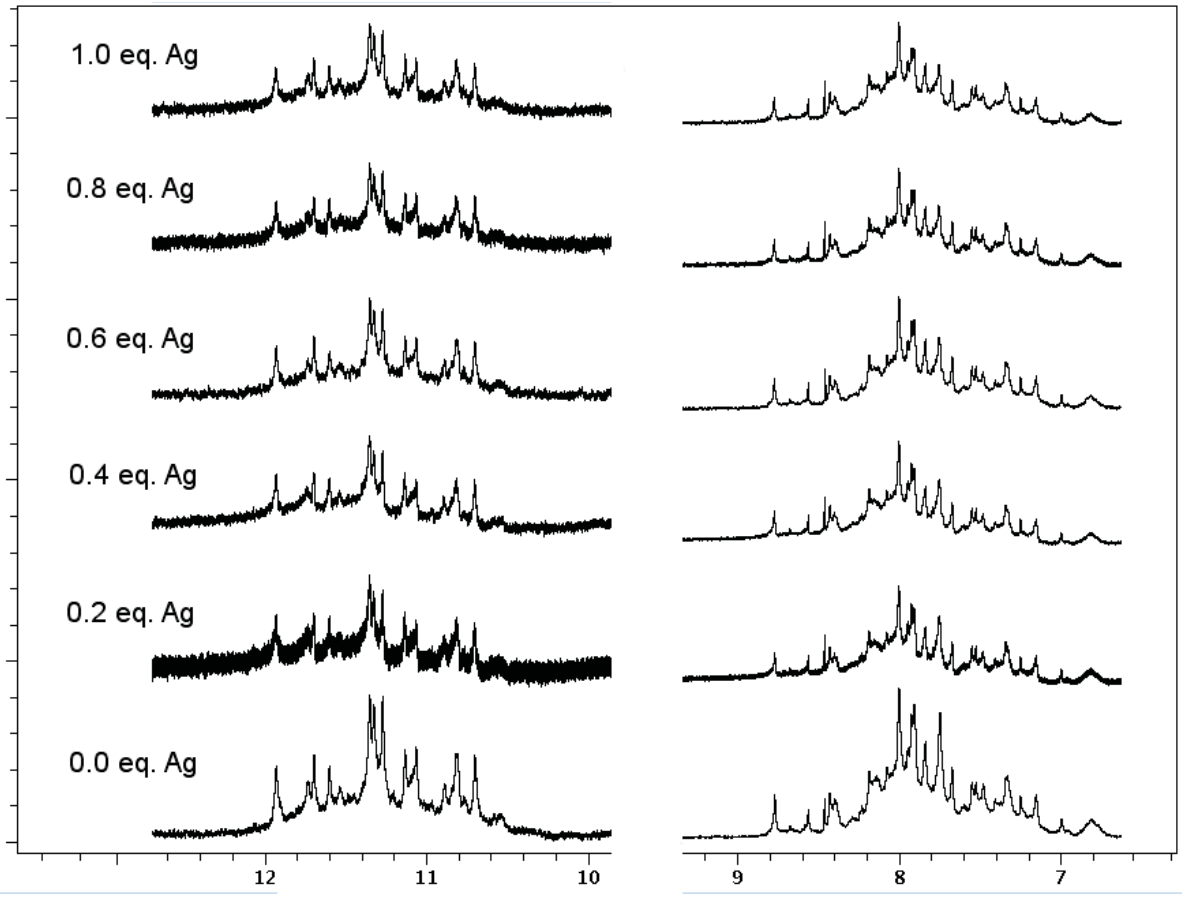


Figure 3.21. ^{1}H NMR spectra of the titration of **ON2** with $AgNO_3$. Samples contain 10 mM sodium phosphate buffer (pH 7.0), 100 mM $NaClO_4$, 10 % D_2O and 5 μM TSP as an internal standard. The imino region from 10 - 12.5 ppm has been magnified relative to the aromatic region from 7 - 8.5 ppm.

^{1}H NMR spectra for **ON3** and **ON4** do not show any changes with increasing silver(I) concentration, consistent with PAGE and CD data which also showed that these oligonucleotides showed no changes with the addition of silver. Imino protons exchange rapidly with deuterium in solution causing a broadening, and possibly the complete disappearance, of peaks for protons exposed to the bulk solvent. The small, broad imino signals for these oligonucleotides indicate that the structures formed from **ON3** and **ON4** are of relatively low stability.

The spectrum for **ON4** suffers from a low DNA concentration, resulting in particularly broad signals, although their chemical shift is unchanged. The low resolution of these peaks makes confident comparisons between the spectra more difficult. However, considered with PAGE and CD data, there is no evidence to indicate any silver interactions occurs with this oligonucleotide.

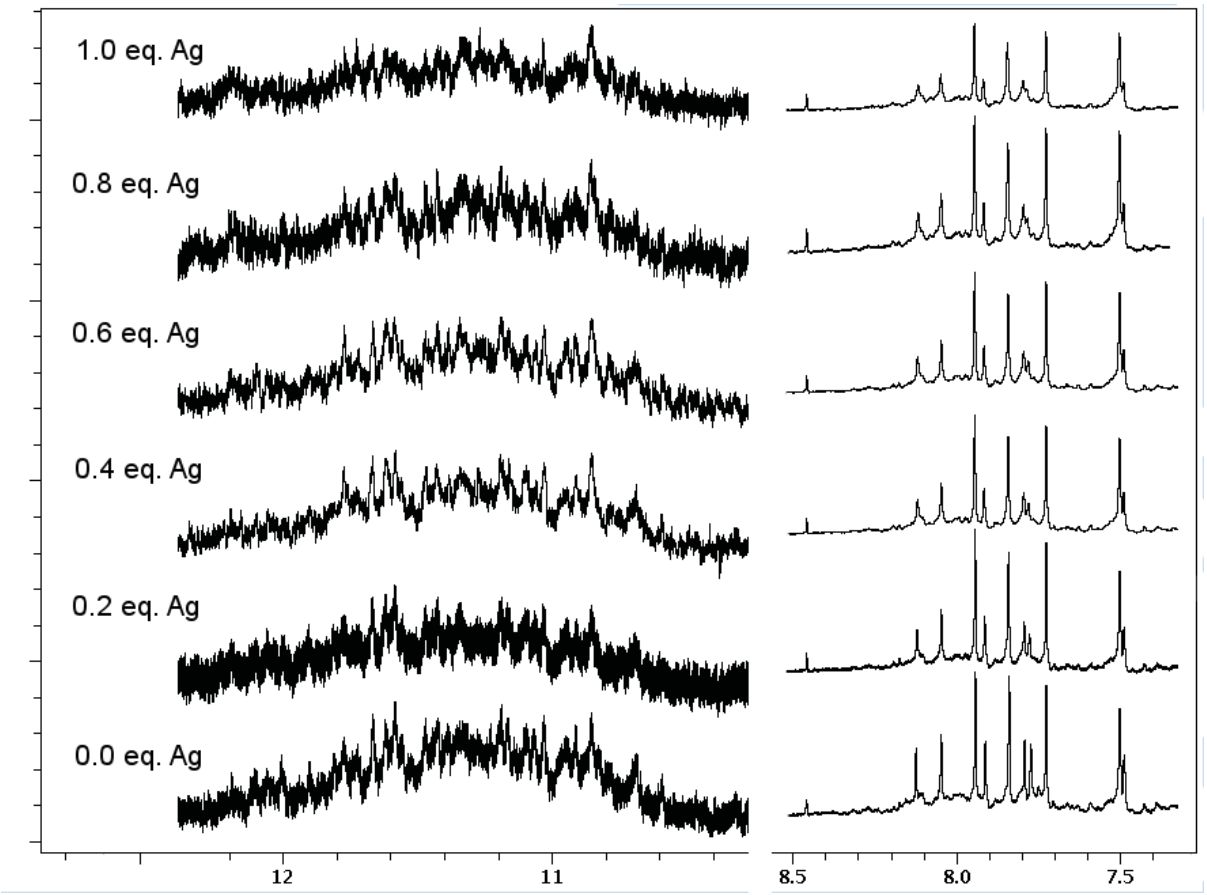


Figure 3.22. ^{1}H NMR spectra of the titration of **ON3** with $AgNO_3$. Samples contain 10 mM sodium phosphate buffer (pH 7.0), 100 mM $NaClO_4$, 10 % D_2O and 5 μM TSP as an internal standard. The imino region from 10 - 12.5 ppm has been magnified relative to the aromatic region from 7 - 8.5 ppm.

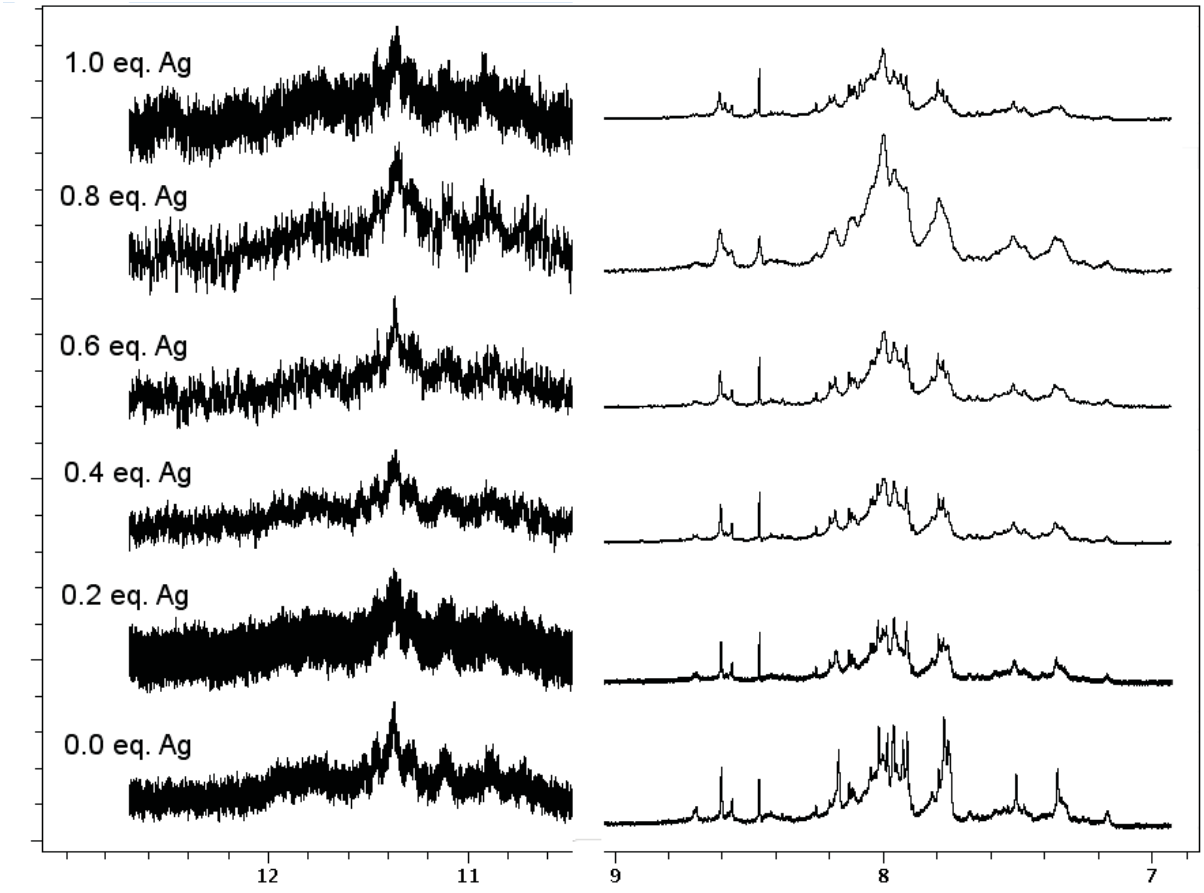


Figure 3.23. ^{1}H NMR spectra of the titration of **ON4** with $AgNO_3$. Samples contain 10 mM sodium phosphate buffer (pH 7.0), 100 mM $NaClO_4$, 10 % D_2O and 5 μM TSP as an internal standard. The imino region from 10 - 12.5 ppm has been magnified relative to the aromatic region from 7 - 8.5 ppm.

3.3 Conclusion

ON1 and **ON2**, based on four repeats of the human telomeric sequence (4*htel*) modified with 5′ and 3′ 1,2,4-triazole nucleotides, assume conformations consistent with structures published for their parent structures. These sequences form antiparallel G-quadruplexes in Na^+ solution and appear to form an equilibrium between parallel and antiparallel quadruplexes in K^+ solution. We assumed that the coordination of a transition metal to the triazole nucleotides would promote the quaternary assembly of these oligonucleotides into dimers or larger structures, however, our experiments showed that this was not the case. PAGE results hinted that **ON1** was capable of coordinating silver(I) ions, but not other metals. This coordination did not result in the formation of high-order quadruplex structures and appeared to involve the coordination of individual silver ions to the complex. CD spectroscopy revealed no large-scale changes in conformation in response to silver ions. ^{1}H NMR suggested that the interaction with silver was not occurring at the triazole nucleotide. Strangely, no response to metal ions was observed under any conditions for **ON2**, even though CD spectroscopy revealed strong similarities in their structures. This may indicate the presence of a small-scale difference in structures brought about by the inclusion of the 3′

triazole nucleotide of **ON2** that serve to prevent silver coordination to this structure. This proposal is speculative, and requires further investigation.

ON3 and **ON4**, based on a 2*htel* parent sequence and modified with 5′ and 3′ 1,2,4-triazole nucleotides, form a dimeric antiparallel G-quadruplex in Na^+ solution and mixture of parallel and antiparallel quadruplexes in K^+. These observations are consistent with literature studies.[9] Neither of these oligonucleotides showed any response to metal ions under any of the conditions tested.

3.4 References

(1) Luu, K. N.; Tu, A.; Kuryavyi, V.; Lacroix, L.; Patel, D. J.: Structure of the Human Telomere in K + Solution : An Intramolecular (3 + 1) G-Quadruplex Scaffold Favoring a Major G-Quadruplex Conformation for Struc. *J. Am. Chem. Soc.* **2006**, 9963-9970.

(2) Lipps, H. J.; Rhodes, D.: G-quadruplex structures: in vivo evidence and function. *Trends Cell Biol.* **2009**, *19*, 414-422.

(3) Zhang, N.; Phan, A. T.; Patel, D. J.: (3 + 1) Assembly of Three Human Telomeric Repeats into an Asymmetric Dimeric G-Quadruplex. *J. Am. Chem. Soc* **2005**, *2005*, 17277-17285.

(4) Wang, Y.; Patel, D. J.: Solution structure of the human telomeric repeat d[AG3(T2AG3)3] G-tetraplex. *Structure* **1993**, *1*, 263-282.

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

(6) Phan, A. T.; Patel, D. J.: Two-Repeat Human Telomeric d(TAGGGTTAGGGT) Sequence Forms Interconverting Parallel and Antiparallel G-Quadruplexes in Solution: Distinct Topologies, Thermodynamic Properties, and Folding/Unfolding Kinetics. *J. Am. Chem. Soc.* **2003**, *125*, 15021-15027.

(7) Muller, J.; Bohme, D.; Lax, P.; Morell Cerda , M.; Roitzsch, M.: Metal ion coordination to azole nucleosides. *Chem. Eur. J.* **2005**, *11*, 6246-53.

(8) Seel, F.; Rodrian, J.: “Azolato”-metallate des Kobalts, Nickels, Kupfers und Zinks. *Liebigs Ann. Chem.* **1974**, *1974*, 1784-1788.

(9) Paramasivan, S.; Rujan, I.; Bolton, P. H. P. H.: Circular dichroism of quadruplex DNAs: applications to structure, cation effects and ligand binding. *Methods* **2007**, *43*, 324-31.

(10) Blume, S. W.; Guarcello, V.; Zacharias, W.; Miller, D. M.: Divalent transition metal cations counteract potassium-induced quadruplex assembly of oligo(dG) sequences. *Nucleic Acids Res.* **1997**, *25*, 617-25.

4 1,2,4-Triazole Quartets in Tetramolecular G-quadruplexes

4.1 Introduction

4.1.1 Concept

The concept of metal-coordinating base-pairs in DNA duplexes has been well explored. It occurs to us that a natural progression in this research is the development of metal-coordinating triplets and quartets. Of these two possibilities, we have focused on G-quadruplexes for the following reasons: first, they possess a greater degree of symmetry than triplexes – their C4 axis of rotation lends itself to the coordination of square planar, tetragonal and octahedral transition metals. Second, the hollow, central core of a G-quadruplex provides an obvious means of access of transition metals to any internal metal-coordinating ligands, as it does for monovalent s-block cations such as sodium and potassium. Third, G-quadruplex assembly is often more straightforward, in that one type of short, self-assembling oligonucleotide may often suffice, whereas triplexes typically require three different, longer oligonucleotides.

The inclusion of a metal-coordinating quartet into a G-quadruplex structure creates a novel method of controlling oligonucleotide assembly. A system can be imagined where the inclusion of metal-coordinating nucleotides into a DNA sequence encourages self-assembly upon addition of a suitable metal-ion. The resulting metal quartet may also act to stabilise the entire structure. An example of an application where these features may be useful is in the G-quadruplex-forming oligonucleotide AS1411, an anti-cancer agent currently undergoing phase-II clinical trials. This oligonucleotide, with the sequence GGTGGTGGTGGTTGTGGTGGTGGTGG, forms a mixture of assemblies as a result of the sequence's poorly defined G-tracts, assuming a number of different conformations with similar stability. The biologically active component is currently unknown.[1,2] This example highlights a limitation when designing G-quadruplex assemblies: there is no alternative to the G-tetrad. Even the availability of just two different base pairs, A/T(U) or G/C, allows nucleic acid duplexes to store a vast amount of information, and in the case of RNA, allows for complicated three-dimensional structures such as tRNA and RNAzymes. A metal-coordinating quartet could function as a compatible, but orthogonal, partner to the standard G-quartet. In the case of AS1411, a metal quartet may allow the ability to 'lock in' the biologically active conformation, with the effect of improving the potency of the anti-cancer activity without increasing the dosage.

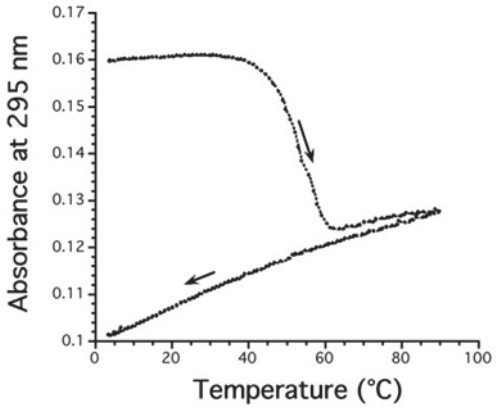


Figure 4.1. An example of the slow assembly rate of tetramolecular G-quadruplexes.[3] A metal coordinating quartet may increase the rate of G-quadruplex assembly and reduce hysteresis in melting

We propose two mechanisms by which a transition-metal-coordinating quartet can contribute to G-quadruplex assembly:

1) The coordinate bond formed between a metal-coordinating nucleotide and a transition metal ion may act to increase the rate of quadruplex assembly. One of the factors responsible for the slow association rates of tetramolecular quadruplexes, Figure 4.1, is their negative activation energy, which is a consequence of the need for several guanine bases come into contact in a very particular orientation for successful assembly.[3] The formation of a bond between triazole and a metal ion does not have the same stringent orientation requirements as guanine-guanine interactions and so it may be the case that a metal-coordinating quartet assembles quickly and serves to bring the rest of the structure into more favourable contact.

2) A metal-coordinating quartet may improve the melting temperature of a G-quadruplex assembly by acting as a more stable analogue of a standard G-quartet. Even if metal coordination occurs more slowly than G-quadruplex assembly it may be the case that an assembled quadruplex serves to preorganise the triazole bases into a position that is favourable for metal coordination, i.e. the quadruplex may be stabilised by a transition metal, but only after assembly has occurred.

As mentioned above, the concept of a transition-metal-coordinating quartet is, to the best of our knowledge, completely unique, there exist no examples in the literature from which to form a basis for design. A summary of research we considered in the development of our assemblies has been included in the introduction to this thesis and will not be repeated here.

4.1.2 Design of Model G-Quadruplex

Unimolecular and bimolecular G-quadruplexes can often form complicated systems with unexpected structures, see AS1411, above. For this reason we focused our investigations on tetramolecular complexes, which almost invariably form parallel G-quadruplexes – this

predictability makes them an ideal model system. Tetramolecular assemblies also have the advantage that they can be formed using shorter oligonucleotides, simplifying DNA synthesis. A consequence of having narrowed our focus to simple tetramolecular G-quadruplexes and metal-coordinating quartets is that we are limited to monodentate ligands. Multi-dentate ligands such as that used by Miyoshi,[4] described on page 17, are likely to be too large and may prompt the formation of undesired high-order structures. Figure 4.2 shows an idealised example of a G-quadruplex containing four monodentate nucleotides assembled into a metal-coordinating quartet.

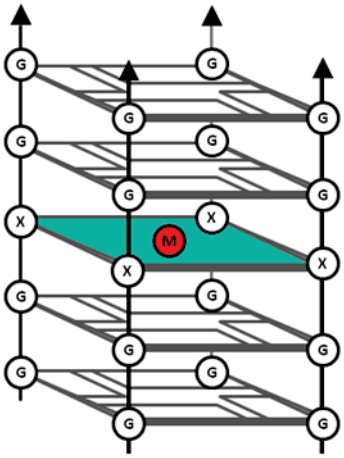


Figure 4.2. An idealised G-quadruplex incorporating metal-coordinating nucleotides, labelled as 'X', forming a metal-coordinating quartet, highlighted blue.

Many of the same principles that govern metal-mediated base-pair formation apply to metal-coordinating quartets. For our intended use in a tetramolecular G-quadruplex, a nucleotide need only have the ability to coordinate metal ions and be of a suitable size, such that four nucleotides cover an area of similar size to a G-quartet. As described previously, our initial targets were pyridyl-based ligands. Difficulties in synthesis led us to focus our efforts on 1,2,4-triazole nucleotides which, with several similar azole ligands, had the benefit of a published synthetic route and the demonstrated ability to coordinate silver(I) and mercury(II) ions.[5-7]

4.2 Results and Discussion

4.2.1 Molecular Modelling of 1,2,4-Triazole Quartets in a G-Quadruplex

We carried out molecular modelling based on a modified version of protein databank entry 2O4F, deposited by Creze et al.[8] This structure consists of a parallel tetramolecular G-quadruplex formed from the sequence TG_4T. We used this X-ray structure as the basis for an investigation into the stability of a 1,2,4-triazole quartet located in the centre of a G-quadruplex assembly, similar to the assembly shown in Figure 4.2. Our model was a tetramolecular parallel G-quadruplex formed from the self-assembly of four oligonucleotides with the sequence TGG**T**$^{\mathbf{r}}$GGT, where **T**$^{\mathbf{r}}$ refers to a 1,2,4-triazole nucleotide. The molecular modelling software we used was not capable of processing transition metal ions. Instead we constrained the distances between each of the triazole rings into a position consistent with their coordination to a square-planar Cu^{2+} ion, i.e. a Cu-N distance of 2.0 Å,[9-11] Figure 4.3. The starting position for this quartet was determined by a low-level minimisation that resulted in the triazole rings twisting slightly out of plane in a propeller conformation. The full (TGG**T**$^{\mathbf{r}}$GGT)$_4$ quadruplex was subjected to a series of molecular dynamic calculations which consisted of random, small-scale alterations of residue positions in order to simulate motions of the assembly in aqueous solution; these models were carried out with the assumption that the triazole quartet would remain intact, no attempt was made to determine whether molecular motions would be capable of disrupting this complex.

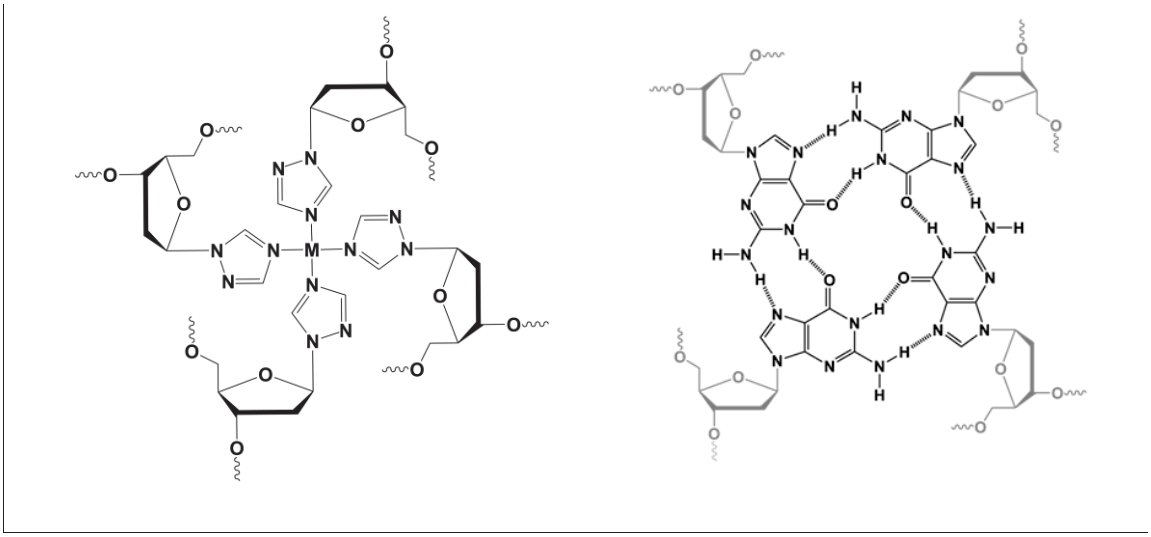


Figure 4.3. A metal-coordinating quartet, left, and a standard G-quartet, right.

Although the metal-coordinating quartet covers a similar area as a G-quartet, the increased height caused by the triazole propeller twisting resulted in a large distortion of the surrounding G-quartets, Figure 4.4. Interestingly, the most stable arrangement seems to sacrifice one side of the assembly. The most significant deformations were localised in one half of the structure where one pair of G-quartets were greatly distorted while the rest of the

structure remains largely intact. As a comparison, a model where the triazole quartet assumes a flat arrangement has been included.

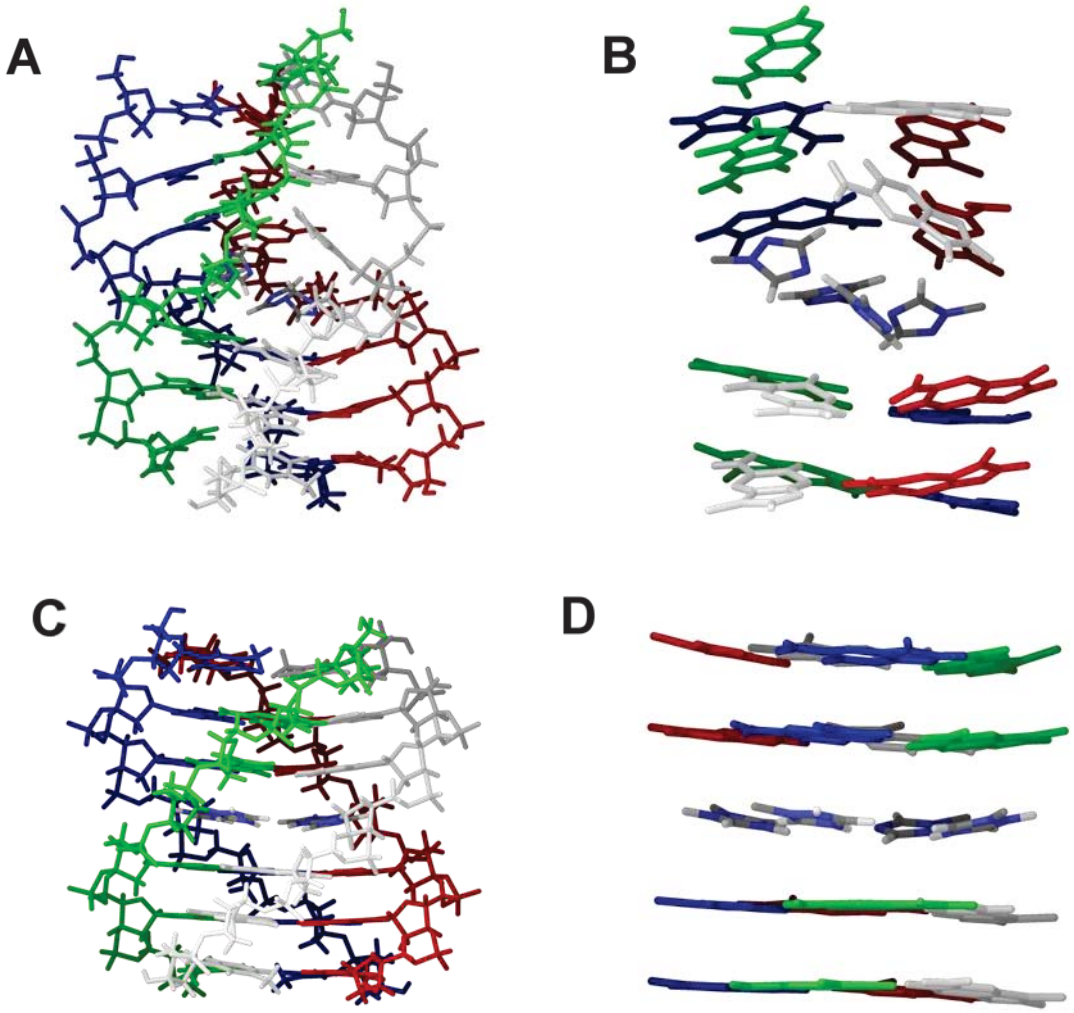


Figure 4.4. Calculated conformation of G-quadruplex assembled from (TGG**T**^{r}GGT)$_4$, A; distortions of G-quartets caused by the presence 1,2,4-triazole quartet, B. DNA backbone removed for clarity. For comparison: G-quadruplex conformation where the 1,2,4-triazole quartet remains planar, C; G-quartets surrounding planar triazole quartet, D.

These calculations suggested that a triazole quartet might be best tolerated on the ends of a G-quadruplex and that inclusion within a G-quadruplex was likely to prevent a stable quadruplex from forming. We carried out molecular modelling on a G-quadruplex formed from a related sequence: TGG_**T**^{r}_GGT, where '_' represents an abasic spacer: a nucleotide without an attached nucleobase, Figure 4.5; the rationale for this being that the abasic spacers would provide the triazole quartet with sufficient room to prevent the distortion of the neighbouring G-quartets. Performing the same molecular dynamics calculations on this revised assembly revealed that the flexibility introduced by the abasic spacers is such that we are not able to present a single model that was representative the assembly, i.e. dynamics

calculations could not converge on a single most stable structure in our experiments. From the range of models that were produced it could be determined that the abasic spacers behaved as intended, providing enough space around the central metal-coordinating quartet that the surrounding G-quartets remained largely intact – reducing, although not eliminating, distortion.

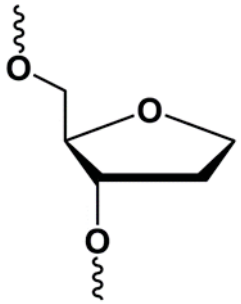


Figure 4.5. An abasic nucleoside spacer.

With the information provided by these models, we designed a series of oligonucleotides that explore the various positions a metal-coordinating quartet could assume within a G-quadruplex, Table 4.1. Oligonucleotides **ON1**, **ON2** and **ON3** contain triazole nucleotides in positions to direct the formation of metal-quartets on the surface of their quadruplexes and were intended to investigate whether there are differing metal coordination effects depending on positioning at the 5′ or 3′ ends of the quadruplex. **ON4** and **ON5** are the direct result of our modelling experiments, described above. We synthesised the 1,2,4-troazole nucleotides as detailed in the experimental section (Chapter 7). The oligonucleotides themselves were synthesised using standard phosphoramidite chemistry as described in the introduction. Characterisation of the triazole nucleotide and oligonucleotides is described in the experimental section.

Table 4.1. Oligonucleotides examined in this chapter.

	Sequence
ON1	TGGGG**T**^{r}T
ON2	T**T**^{r}GGGGT
ON3	T**T**^{r}GGGG**T**^{r}T
ON4	TGG_**T**^{r}_GGT
ON5	TGGG_**T**^{r}_GGGT

'_' Represents an abasic spacer, Figure 4.5.

As noted above, one of the primary mechanisms by which a metal ion may contribute to G-quadruplex stability is by coordination after G-quadruplex assembly has preorganised the

1,2,4-triazole quartet; this only becomes a factor if the G-quadruplex is capable of assembling first. Tetramolecular G-quadruplexes containing four contiguous G-quartets, as in the cases of quadruplexes formed from **ON1**, **ON2** and **ON3,** are expected to have a $T_{1/2}$ value of approximately 55 °C in Na^+ solution.[12,13] **ON5** potentially contains six G-quartets: two sets of three contiguous G-quartets, but an interruption in a G-tract is strongly destabilising, as are abasic spacers: **ON4** is likely to be only marginally stable for these reasons.[14] The added flexibility introduced by the abasic spacers also opens up the possibility that **ON4** and **ON5** may form a bimolecular structure of the type indicated in Figure 4.6. We used circular dichroism spectroscopy to determine whether these oligonucleotides formed G-quadruplexes in the absence of a transition metal, and to identify the conformations they assumed.

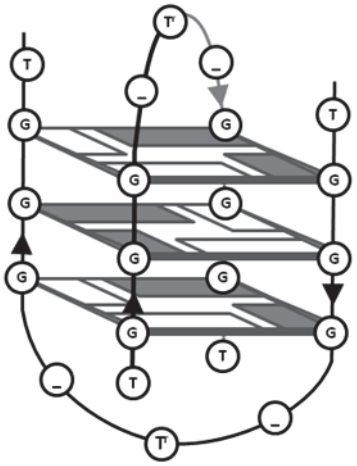


Figure 4.6. An antiparallel bimolecular G-quadruplex is one potential conformation of **ON5**.

4.2.2 Determination of G-Quadruplex Conformation by CD Spectroscopy

CD experiments were carried out in 10 mM lithium cacodylate buffer (pH 7.2) and 100 mM $NaClO_4$ at 20 °C (Figure 4.7). CD spectra for **ON1**, **ON2**, **ON3** and **ON5** are very similar: positive maxima at ~265 nm and negative minima at ~235 nm, this indicates that these oligonucleotides form parallel G-quadruplexes, as expected. It may be inferred that the four triazole bases are found at the same position within each G-quadruplex allowing for some degree of preorganisation for metal coordination. We found that **ON4** provides very different spectra and appears to be unstable under these conditions. It should be noted that the CD spectrum for **ON4** was much less intense than those for **ON1**, **ON2**, **ON3** and **ON5**, but appear matched because of the normalisation carried out.

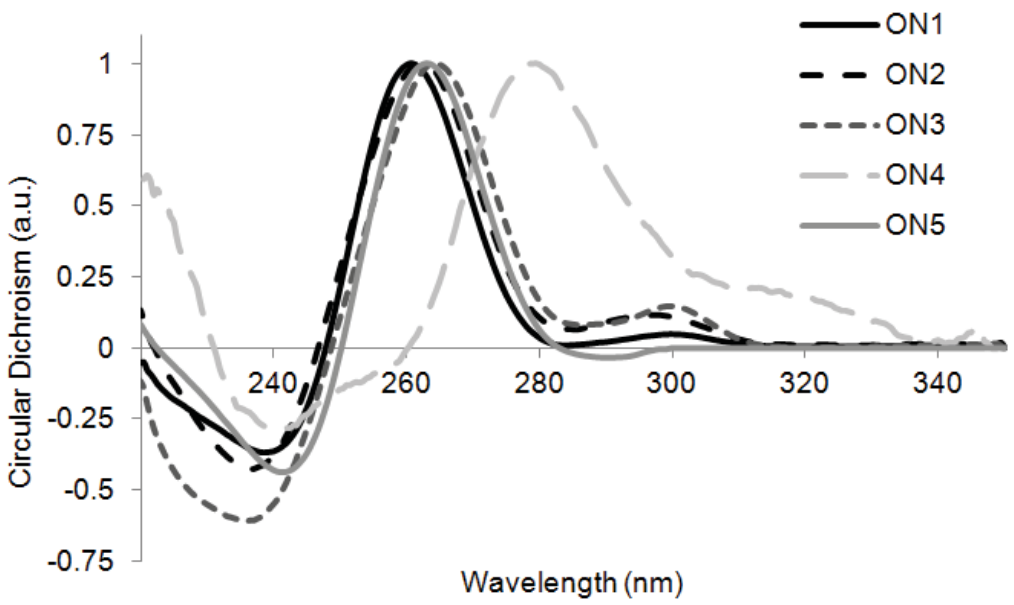


Figure 4.7. Normalised CD spectra of **ON1 – 5**. Recorded at 20 °C with 10 mM lithium cacodylate buffer (pH 7.2), 100 mM $NaClO_4$.

4.2.3 Examination of Metal Ion Coordination by PAGE

Using PAGE, we screened **ON1 – 5** with a range of s- and d-block cations. Cations were present at 25 μM**:** one quarter of the triazole nucleotide concentration, a level selected to allow the assembly of triazole-metal quartets while minimising destabilisation of the G-quadruplex core via metal coordination to guanine. As noted previously, several examples exist in the literature of silver and mercury coordination to 1,2,4-triazole nucleotides.[6] There also exist several examples of metal coordination to free 1,2,4-triazole, forming ML_2 complexes with Co(II), Ni(II), Cu(II), and Zn(II).[9]

The ladder used in these gels shows poor resolution, however, the assemblies formed by **ON1**, **ON2**, **ON3** and **ON5** migrate approximately between dT_{10} and dT_{20}, consistent with literature observations of tetramolecular G-quadruplexes formed from $(TG_4T)_4$,[15] indicating these oligonucleotides form tetramolecular G-quadruplexes. This is consistent with CD spectra (Figure 4.7), and with the appearance of signals in the imino region of ^1H NMR spectra, discussed later.

Figure 4.8 - Figure 4.11 reveal no metal-dependent changes under these conditions for any of the oligonucleotides tested. It should be noted that the apparent gel shift observed for **ON5** (Figure 4.11) in lane 3 caused by Hg^{2+} is an artefact caused by an air bubble within the gel.

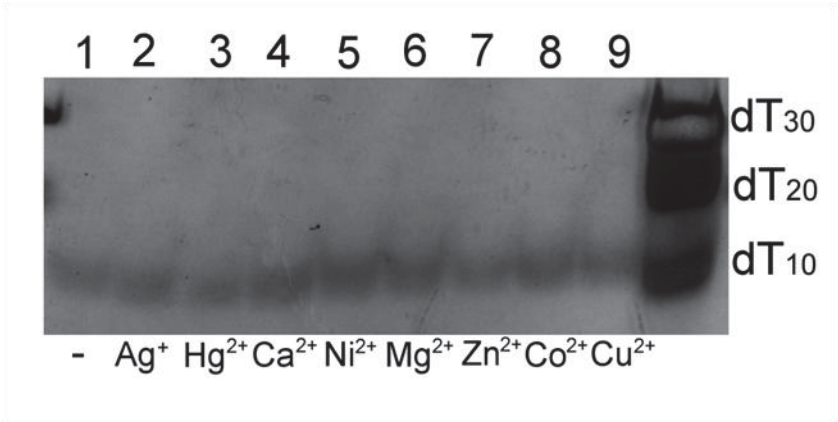


Figure 4.8. Native PAGE of **ON1**. Samples contain 100 μM DNA, 10 mM lithium cacodylate buffer (pH 7.2) and 100 mM $NaClO_4$. Metal ions included at 25 μM as their nitrate salts. Gel contains 12% acrylamide, TB buffer (pH 8.0) supplemented with 100 mM $NaNO_3$. Electrophoresis carried out at 15 W, 3 °C, 3 hours.

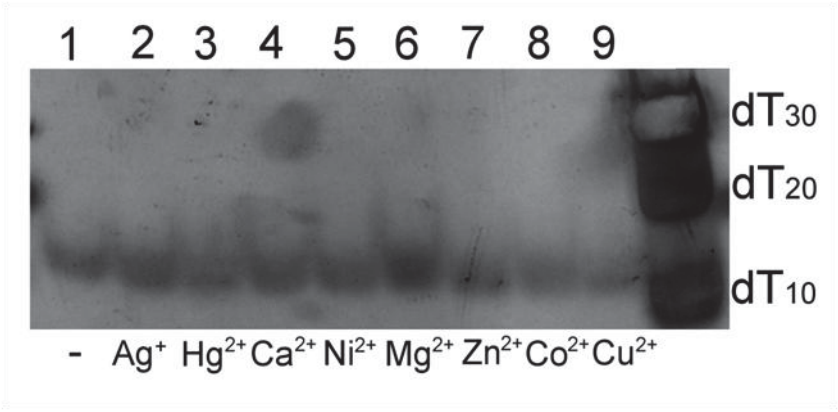


Figure 4.9. Native PAGE of **ON2**. Samples contain 100 μM DNA, 10 mM lithium cacodylate buffer (pH 7.2) and 100 mM $NaClO_4$. Metal ions included at 25 μM as their nitrate salts. Gel contains 12% acrylamide, TB buffer (pH 8.0) supplemented with 100 mM $NaNO_3$. Electrophoresis carried out at 15 W, 3 °C, 3 hours.

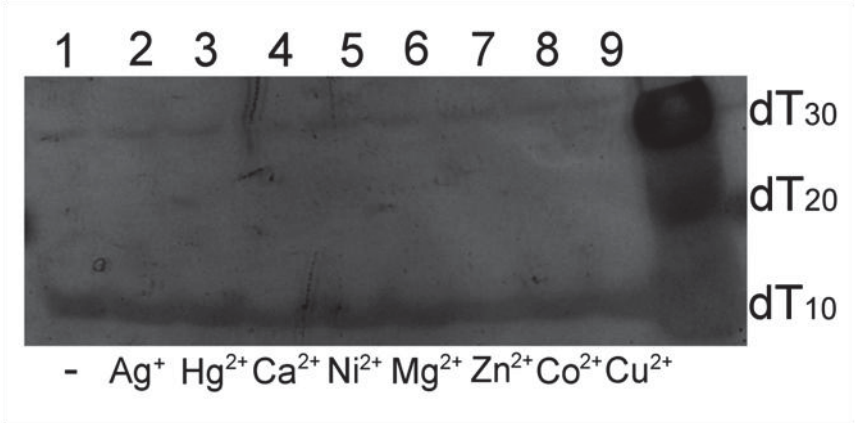


Figure 4.10. Native PAGE of **ON3**. Samples contain 100 μM DNA, 10 mM lithium cacodylate buffer (pH 7.2) and 100 mM $NaClO_4$. Metal ions included at 25 μM as their nitrate salts. Gel contains 12% acrylamide, TB buffer (pH 8.0) supplemented with 100 mM $NaNO_3$. Electrophoresis carried out at 15 W, 3 °C, 3 hours.

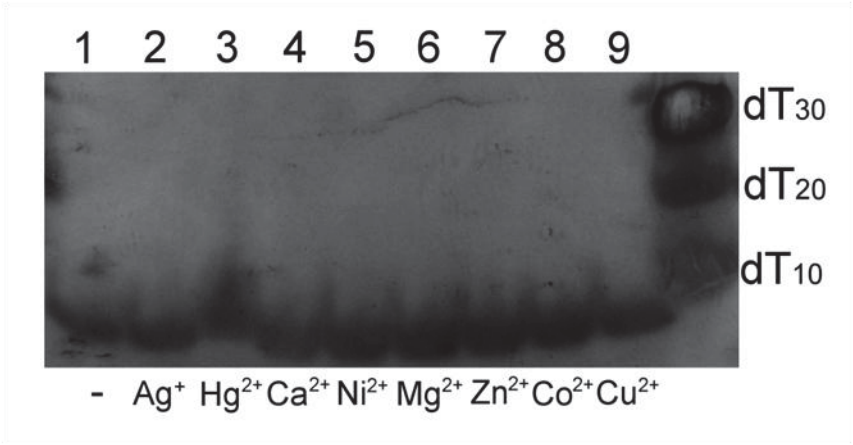


Figure 4.11. Native PAGE of **ON5**. Samples contain 100 μM DNA, 10 mM lithium cacodylate buffer (pH 7.2) and 100 mM $NaClO_4$. Metal ions included at 25 μM as their nitrate salts. Gel contains 12% acrylamide, TB buffer (pH 8.0) supplemented with 100 mM $NaNO_3$. Electrophoresis carried out at 15 W, 3 °C, 3 hours. The apparent gel shift observed in lane 3 (Hg^{2+}) is an artefact caused by an air bubble within the gel.

In the absence of data to suggest a more appropriate alternative, we selected silver(I) for further investigation because of its proven ability to coordinate to 1,2,4-triazole nucleotides, albeit in a different environment. Silver(I) most commonly assumes a tetrahedral geometry. Molecular modelling suggested that this geometry can be tolerated within the G-quadruplex structure, albeit with slightly more distortion than a square planar complex.

4.2.4 Influence of Ag^+ on Association Rates of ON5

To test our hypothesis that fast assembly of a metal-coordinating quartet serves to promote G-quadruplex formation, we investigated whether the addition of Ag^+ ions influenced the rate of G-quadruplex formation by measuring the association rate of **ON5** (Figure 4.12). Association rates were modelled according to the equation:

$$\alpha = [1 + C_o^{n-1} \cdot (n - 1) \cdot k_{on} \cdot t]^{1/(1-n)} \qquad \text{Equation 4.1}$$

Where α is the proportion of unfolded oligonucleotide, C_o is the initial oligonucleotide concentration, n is the order of the reaction (approximately 4 in the case of tetramolecular G-quadruplexes), k_{on} is the association rate, and *t* is time.[3,16] See introduction 1.4.1 for further detail.

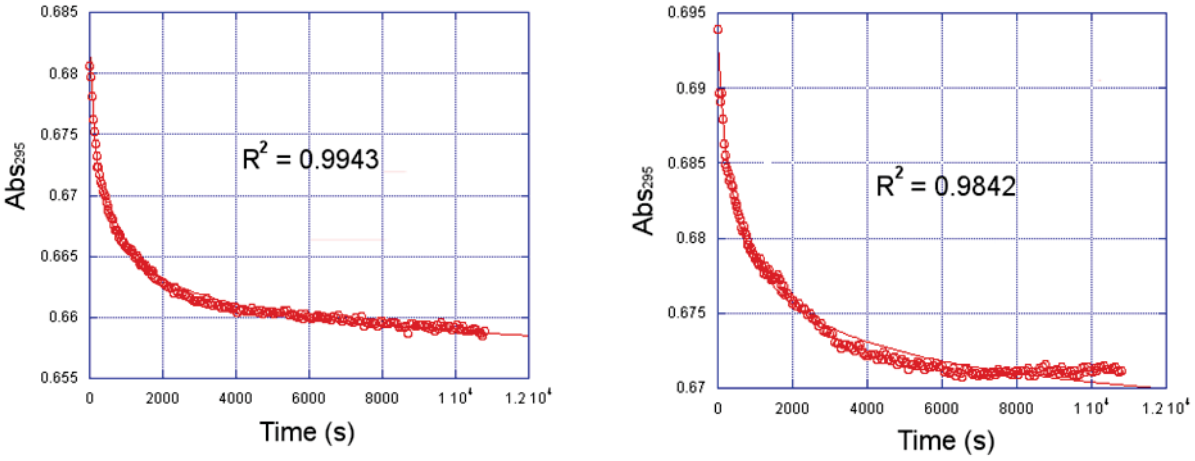


Figure 4.12. Association rates of **ON5** in the absence of transition metal ions, left, and with 1 eq. Ag^+, right. Recorded at 25 °C, 295 nm. Samples contain 10 μM DNA, 10 mM lithium cacodylate buffer (pH 7.2), 100 mM potassium cacodylate and, when present, 2.5 μM $AgNO_3$.

For **ON5** in the absence of transition metals k_{on} was found to be $3.49(0.14)\times10^{12}$ $M^{-3}s^{-1}$. The addition of silver ions had a negative effect on association rates, lowering k_{on} to $1.98(0.13)\times10^{12}$ $M^{-3}s^{-1}$. This decrease in association rate is likely brought about by the coordination of silver to guanine nitrogens which interferes with G-quartet formation, this mechanism is consistent with the silver-mediated destabilisation of assemblies observed by CD spectroscopy.[17] It should be noted that the observed decrease in association rate is strong evidence against the first of our proposed mechanisms by which the fast assembly of a metal-coordinating quartet serves to promote G-quadruplex assembly by bringing the rest of the structure into more favourable contact. It appears that for silver(I), and likely for other transition metals, the negative impact of metal coordination to guanine outweighs any positive influence of triazole coordination, if triazole-metal coordination is, in fact, occurring.

4.2.5 Thermal Stability Studies

Association data indicated that silver(I) ions have a negative impact on G-quadruplex formation rates; $T_{1/2}$ values are able to reveal whether silver(I) ions are capable of stabilising G-quadruplexes once they have assembled, i.e. whether the second of our proposed mechanisms for G-quadruplex stabilisation is valid.

$T_{1/2}$ data were collected using a CD spectrophotometer. Spectra were monitored at 260 nm and the midpoint of the transition between G-quadruplex and ssDNA was determined. Experiments were carried out in 10 mM lithium cacodylate buffer (pH 7.2), 100 mM $NaClO_4$ and, when present, 1 eq. $AgNO_3$. It should be noted that the dissociation of tetramolecular G-quadruplexes is independent of oligonucleotide concentration, being a unimolecular process

where the reverse direction (quadruplex reassociation) occurs slowly enough to be ignored on the time scale of the experiment.[3]

Table 4.2. Melting temperatures ($T_{1/2}$, ° C), determined by CD, for oligonucleotides in the presence or absence of 1 eq. $AgNO_3$. Samples contain 10 mM lithium cacodylate buffer (pH 7.2), 100 mM $NaClO_4$.

	No metal	**Ag^+**
ON1	44.0	43.5
ON2	46.0	45.5
ON3	48.5	47.5
ON4	-	-
ON5	49.5	47.0

Observed melting temperatures reveal that these G-quadruplexes show similar melting temperatures, however, all are lower than that for $(TG_4T)_4$ (~55 °C).[15,16] The difference in melting temperatures between **ON1** and **ON2** is interesting, as the 3′ triazole of **ON1** appears to destabilise the G-quadruplex to a greater extent than the 5′ modified **ON2**. Whereas **ON3**, with both 3′ and 5′ modifications has an increased melting temperature, possibly due to increased shielding the terminal G-quartets from the solvent. **ON5** has the highest stability measured, a result of its two additional G-quartets, although it should be noted that **ON5** (TGGG_**T**$^{\mathbf{r}}$_GGGT) is significantly less stable than, for comparison, $(TG_6T)_4$ which is unmeltable in similar conditions,[18] whereas a TG_3T quadruplex melts at 16 °C.[19] Most significantly, it can be seen that silver has a negative impact on stability and reduces the $T_{1/2}$ for all recorded assemblies, demonstrating that these assemblies are not stabilised by addition of silver(I) ions.

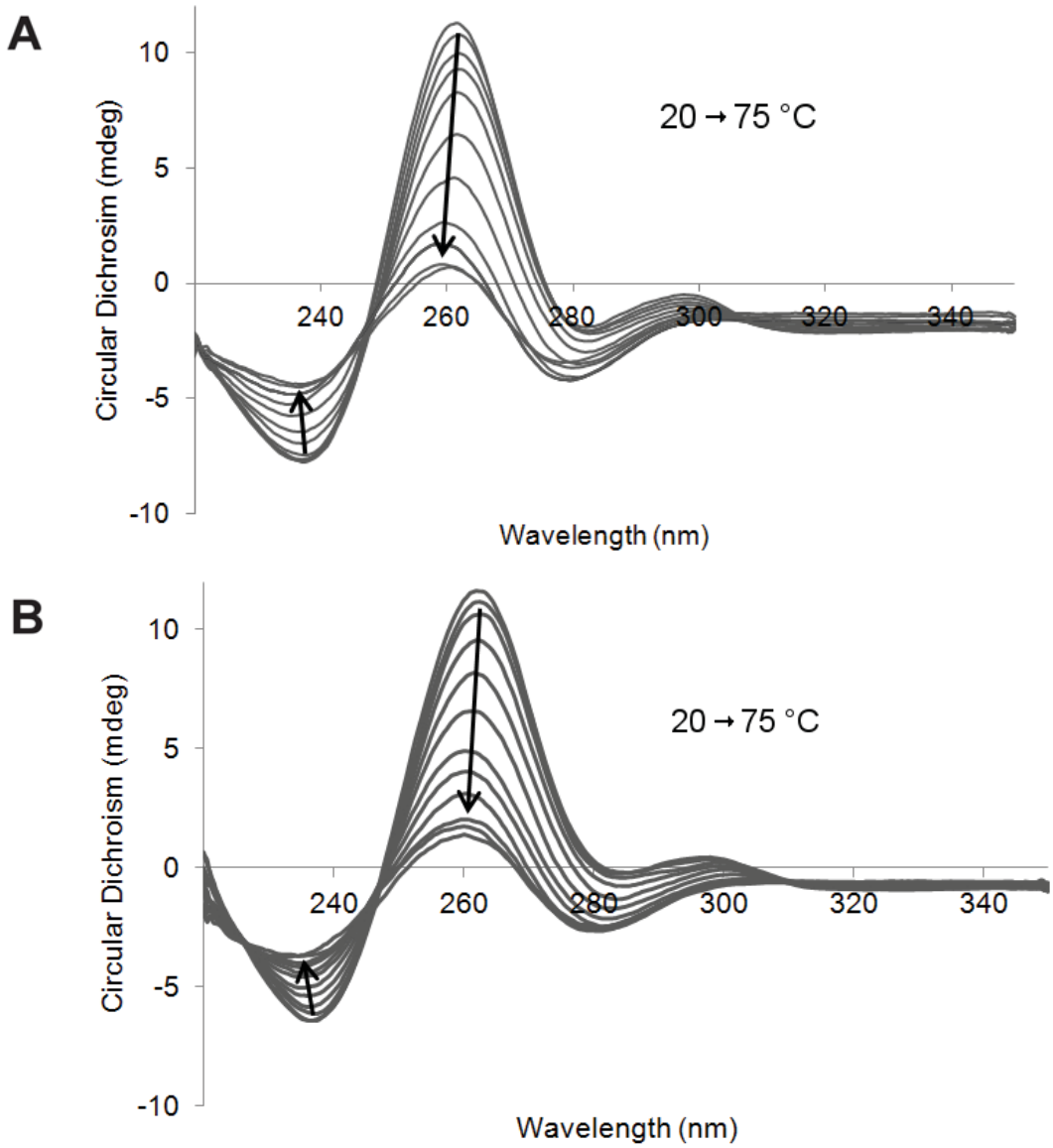


Figure 4.13. Representative CD spectra of changes observed during melting experiments from 20 – 75 °C. Spectra displayed are for **ON2** without Ag^+, A, and with 1 eq. Ag^+, B. Samples contain 10 mM lithium cacodylate buffer (pH 7.2), 100 mM $NaClO_4$. Spectra were recorded at 5 °C increments with a temperature ramp of 1 °C/min and 5 min for thermal equilibration at each increment.

Changes in CD spectra that occur during melting demonstrate a clear transition from parallel G-quadruplex to unstructured DNA. An example of changes observed in CD spectra during the experiment is included in Figure 4.13. The addition of silver does not have significant impact on the conformation of these structures, its effect seems limited to a slight destabilisation, as observed in $T_{1/2}$ values.

4.2.6 Examination of Ag^+ Coordination by NMR

Association data and melting temperatures indicated that the addition of silver(I) ions to these oligonucleotides does not stabilise the resultant G-quadruplexes by either of our prosed mechanisms. The experimental data described above may be explained by two

models: 1) No 1,2,4-triazole/silver quartet is formed, or 2) The formation of a 1,2,4-triazole/silver quartet does occur, but the stabilising influence of this quartet is offset by the negative impact of silver coordination to guanine bases. To distinguish between these possibilities we attempted to detect the coordination of silver by NMR spectroscopy. We carried out titration experiments using increasing quantities of $AgNO_3$ and examined the effects on the chemical shifts of peaks found in the aromatic region of the ^1H NMR spectra.

The chemical shift of the two 1,2,4-triazole protons can be found at 8.4 ppm in the free ring, and approximately 8.1 and 8.65 ppm when incorporated into a nucleoside.[6] Full silver-ion occupancy of triazole base-pairs is expected to result in a change in chemical shift of ~0.15 ppm: a change in position of this magnitude will be noticeable – the two triazole protons can thus be identified through their gradual migration upon titration of increasing quantities of silver nitrate. A solution of each oligonucleotide in 10 mM Na_2HPO_4 buffer (pH 7.0), 100 mM $NaClO_4$ with 10 % D_2O was titrated with increments of 0.2 eq. of silver(I) nitrate, the resulting spectra are detailed in Figure 4.14-Figure 4.17.

Examining the effect of silver(I) on each oligonucleotide, it can be seen that the peaks in the aromatic region of these spectra show no changes in chemical shift with increasing silver concentration, indicating that there is no specific triazole interaction with silver and providing strong evidence against the ability of silver to coordinate to any of these tetramolecular G-quadruplexes under these conditions. The imino regions of each oligonucleotide remain largely unchanged, indicating that there are no significant changes in the conformation of each assembly; a result that is consistent with CD spectroscopy. Some line broadening can be observed in the spectra for **ON5**, Figure 4.17, due to a reduction in shimming quality and small mis-configurations in the NMR instrument as a result of an earthquake. Although signal resolution is decreased for this sample, the position of each peak is unaltered: no changes were observed that would indicate silver(I) coordination.

It is interesting to note the differences in NMR spectra between oligonucleotides **ON1**, **ON2** and **ON3**. The central G-quadruplex cores for these assemblies are expected to be conserved, as CD spectra indicated these oligonucleotides assumed similar conformations, Figure 4.7. Small differences might be caused by the presence of impurities; however these are likely to be limited to the aromatic region of the displayed spectra – it is unlikely that impurities would contribute to the imino region as these signals are the result of non-exchangeable protons. Small oligonucleotides of this type are not likely to form side-products, and CD spectra indicate they assumed the expected parallel conformation. It may be the case that the observed imino signals for **ON1**, **ON2** and **ON3** would otherwise appear similar, but the influence of the triazole nucleotides causes guanine bases to have differing rates of exchange

with the solvent. In the case of **ON1**, the triazole base on the 3′ end of the quadruplex core may lead to increased solvent exchange of the 3′ guanines and thereby suppressing 3′ guanine imino signals. **ON2** would show similar behaviour at the 5′ end of the quadruplex. **ON3**, with both 5′ and 3′ triazole nucleotides, would show a combination of effects.

NMR experiments of oligonucleotide samples containing four equivalents of Ag^+ showed spectra with extremely broad signals consistent with the observation that silver ions contribute to G-quadruplex instability.

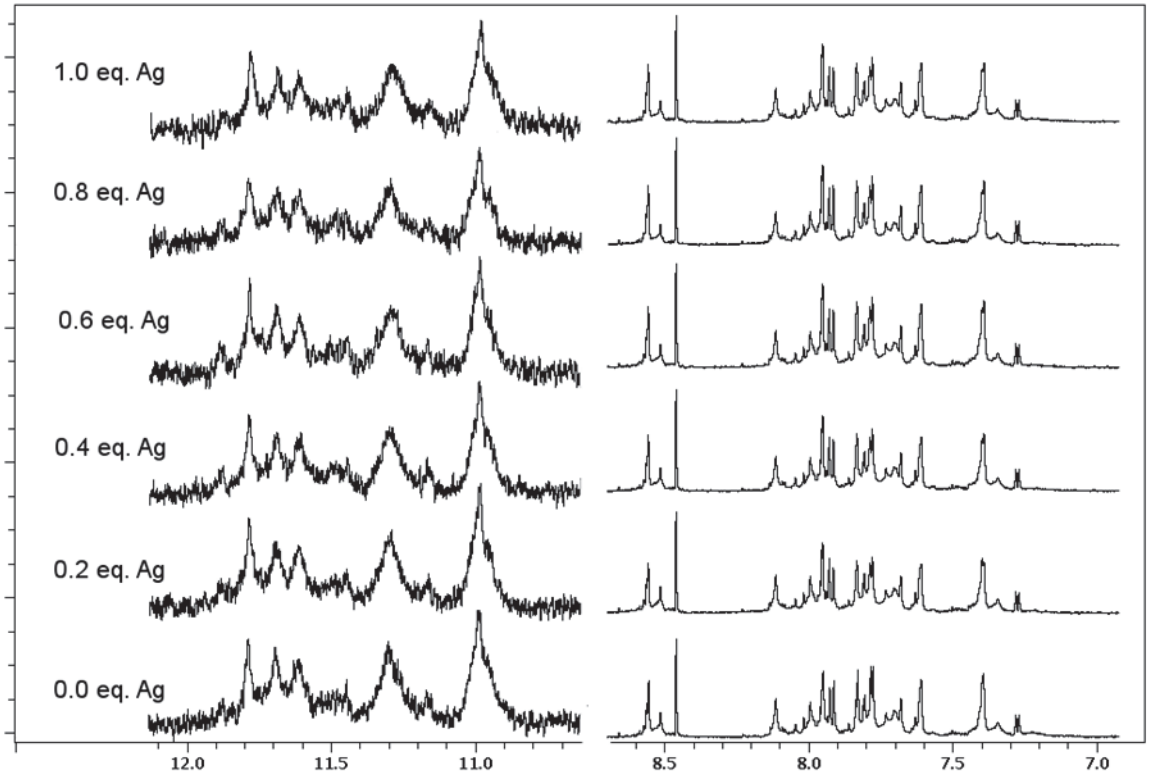


Figure 4.14. ^{1}H NMR spectra of the titration of **ON1** with $AgNO_3$. Samples contain 220 μM DNA, 10 mM sodium phosphate buffer (pH 7.0), 100 mM $NaClO_4$ and 10 % D_2O. The imino region from 10 - 12 ppm has been magnified relative to the aromatic region from 7 - 8.5 ppm.

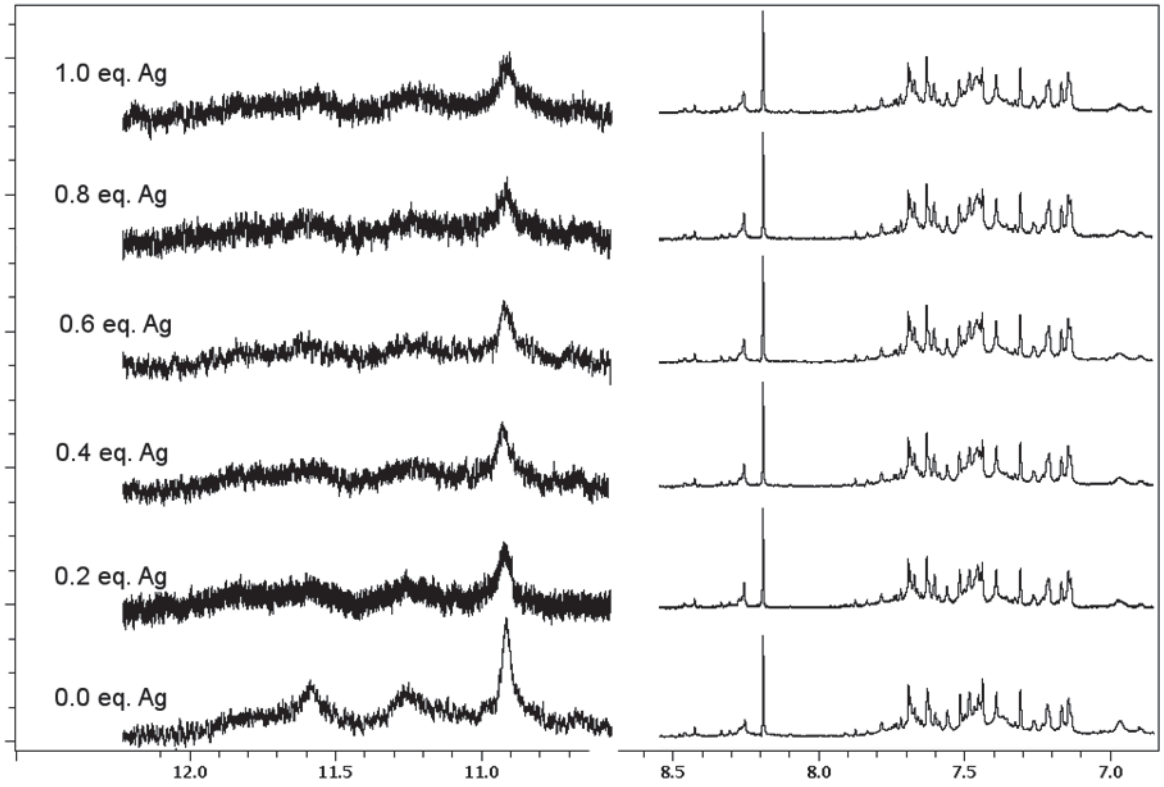


Figure 4.15. ^{1}H NMR spectra of the titration of **ON2** with $AgNO_3$. Samples contain 260 μM DNA, 10 mM sodium phosphate buffer (pH 7.0), 100 mM $NaClO_4$ and 10 % D_2O. The imino region from 10 - 12.5 ppm has been magnified relative to the aromatic region from 7 - 8.5 ppm.

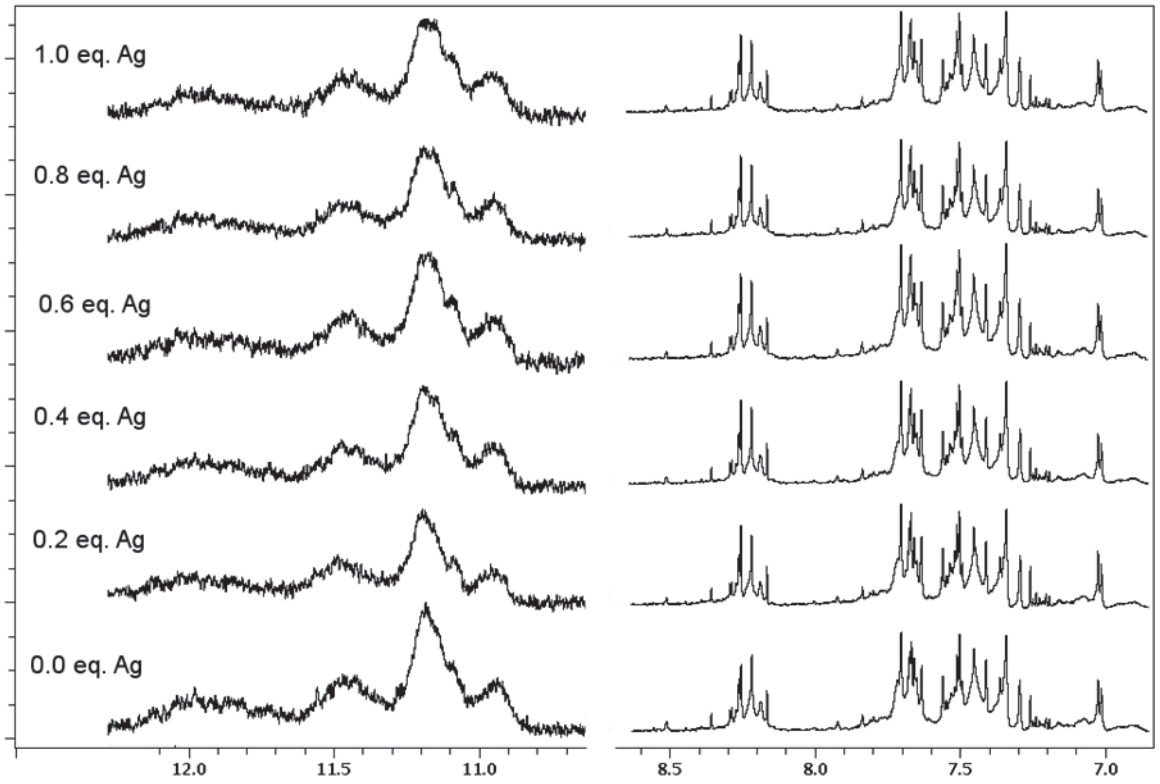


Figure 4.16. ^{1}H NMR spectra of the titration of **ON3** with $AgNO_3$. Samples contain 280 μM DNA, 10 mM sodium phosphate buffer (pH 7.0), 100 mM $NaClO_4$ and 10 % D_2O. The imino region from 10 - 12.5 ppm has been magnified relative to the aromatic region from 7 - 8.5 ppm.

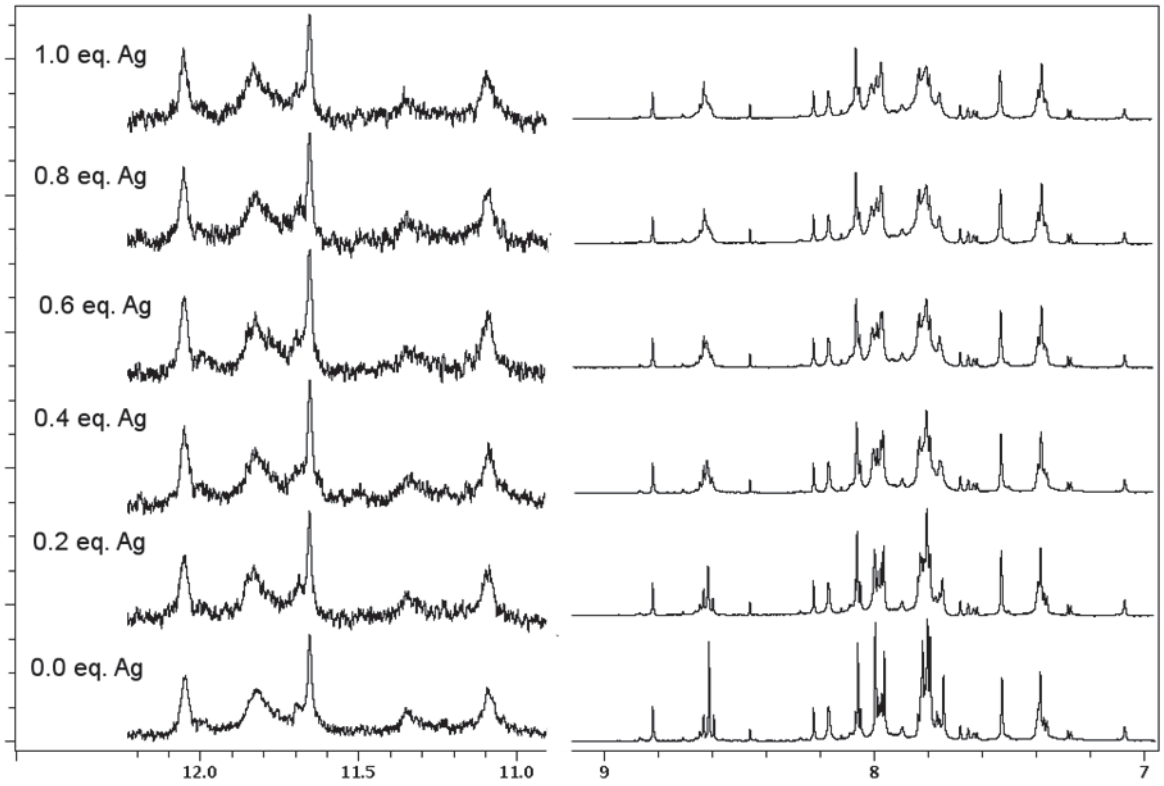


Figure 4.17. ^{1}H NMR spectra of the titration of **ON5** with $AgNO_3$. Samples contain 350 μM DNA, 10 mM sodium phosphate buffer (pH 7.0), 100 mM $NaClO_4$ and 10 % D_2O. The imino region from 11 - 12.5 ppm has been magnified relative to the aromatic region from 7 - 9 ppm.

4.3 Conclusion

Using information collected by molecular modelling, a series of oligonucleotides, **ON1 – 5**, was developed to promote the formation of a metal-coordinating 1,2,4-triazole quartet in a quadruplex assembly. Circular dichroism spectroscopy showed that the DNA sequences did form G-quadruplexes, but ^{1}H NMR indicated that silver ions did not coordinate to the triazole nucleotides. Melting and association experiments revealed that the presence of silver ions reduced both the association rate and thermodynamic stability of G-quadruplexes, demonstrating that the inclusion of a metal-coordinating quartet has a negative impact on quadruplex assembly. Our results indicate that the negative impact of silver(I) ions on G-quadruplex assembly and stability become significant before any potential benefits introduced silver-triazole coordination can be observed. This problem is likely to be present in any system G-quadruplex system containing 1,2,4-triazole nucleotides. The most straightforward method of addressing this problem is likely to involve the use of different metal-coordinating nucleotides that are able to more effectively compete with other binding sites in oligonucleotide assemblies.

Note: Between the initial submission of this thesis (11 October 2013) and the completion of final emendations (13 March 2014), a paper has been published (Engelhard, D. M., Pievo, R.,

Clever, G. H., Reversible Stabilization of Transition-Metal-Binding DNA G-Quadruplexes. Angew. Chem. Int. Ed., 2013, 52, 12843-12847.) that describes a metal coordinating G-quadruplex system similar to the concept presented in this chapter. This paper outlines a tetramolecular G-quadruplex that is stabilised by up to 20 °C by the coordination of Cu(II) ions to terminal pyridine-based nucleotides. This work serves to demonstrate that the concept behind our assemblies was sound, and also highlights areas where our own design may be improved in future.

4.4 References

(1) Reyes-Reyes, E. M.; Teng, Y.; Bates, P. J.: A New Paradigm for Aptamer Therapeutic AS1411 Action: Uptake by Macropinocytosis and Its Stimulation by a Nucleolin-Dependent Mechanism. *Cancer Res.* **2010**, *70*, 8617-8629.

(2) Wu, J.; Song, C.; Jiang, C.; Shen, X.; Qiao, Q.; Hu, Y.: Nucleolin Targeting AS1411 Modified Protein Nanoparticle for Antitumor Drugs Delivery. *Mol. Pharm.* **2013**.

(3) Mergny, J.-L.; De Cian, A.; Ghelab, A.; Sacca , B.; Lacroix, L.; Cian, A. D.; Sacca, B.: Kinetics of tetramolecular quadruplexes. *Nucleic Acids Res.* **2005**, *33*, 81-94.

(4) Miyoshi, D.; Karimata, H.; Wang, Z.-M.; Koumoto, K.; Sugimoto, N.: Artificial G-wire switch with 2,2'-bipyridine units responsive to divalent metal ions. *J. Am. Chem. Soc.* **2007**, *129*, 5919-25.

(5) Johannsen, S.; Megger, N.; Böhme, D.; SigelRoland, K. O.; Müller, J.: Solution structure of a DNA double helix with consecutive metal-mediated base pairs. *Nat. Chem.* **2010**, *2*, 229-234.

(6) Muller, J.; Bohme, D.; Lax, P.; Morell Cerda , M.; Roitzsch, M.: Metal ion coordination to azole nucleosides. *Chem. Eur. J.* **2005**, *11*, 6246-53.

(7) Müller, J.; Böhme, D.; Düpre, N.; Mehring, M.; Polonius, F.-A.: Differential reactivity of α and β 2′-deoxyribonucleosides towards protonation and metalation. *J. Inorg. Biochem.* **2007**, *101*, 470-476.

(8) Creze, C.; Rinaldi, B.; Haser, R.; Bouvet, P.; Gouet, P.: Structure of a d(TGGGGT) quadruplex crystallized in the presence of Li+ ions. *Acta Crystallogr. Sect. D* **2007**, *63*, 682-8.

(9) Seel, F.; Rodrian, J.: “Azolato”-metallate des Kobalts, Nickels, Kupfers und Zinks. *Liebigs Ann. Chem.* **1974**, *1974*, 1784-1788.

(10) Yao, J.-C.; Yao, F.-J.; Li, Y.-C.; Wang, J.-G.: Refinement of the Crystal Structure of tetra(imidazole)copper(II) diperchlorate. *Z. Kristallogr. - New Cryst. Struct.* **2007**, *222*, 211-212.

(11) Sletten, E.: Crystallographic studies of metal-nucleotide base complexes. I. Triclinic bis-(6-aminopurine)copper(II) tetrahydrate. *Acta Crystallogr. Sect. B* **1969**, 1480-1491.

(12) Petraccone, L.; Erra, E.; Esposito, V.; Randazzo, A.; Mayol, L.; Nasti, L.; Barone, G.; Giancola, C.: Stability and structure of telomeric DNA sequences forming quadruplexes containing four G-tetrads with different topological arrangements. *Biochemistry* **2004**, *43*, 4877-84.

(13) Dapic, V.: Biophysical and biological properties of quadruplex oligodeoxyribonucleotides. *Nucleic Acids Res.* **2003**, *31*, 2097-2107.

(14) Esposito, V.; Martino, L.; Citarella, G.; Virgilio, A.; Mayol, L.; Giancola, C.; Galeone, A.: Effects of abasic sites on structural, thermodynamic and kinetic properties of quadruplex structures. *Nucleic Acids Res.*, *38*, 2069-80.

(15) Pedersen, E. B.; Nielsen, J. T.; Nielsen, C.; Filichev, V. V.: Enhanced anti-HIV-1 activity of G-quadruplexes comprising locked nucleic acids and intercalating nucleic acids. *Nucleic Acids Res.* **2011**, *39*, 2470-2481.

(16) Wyatt, J. R.; Davis, P. W.; Freier, S. M.: Kinetics of G-quartet-mediated tetramer formation. *Biochemistry* **1996**, *35*, 8002-8.

(17) Blume, S. W.; Guarcello, V.; Zacharias, W.; Miller, D. M.: Divalent transition metal cations counteract potassium-induced quadruplex assembly of oligo(dG) sequences. *Nucleic Acids Res.* **1997**, *25*, 617-25.

(18) *Quadruplex Nucleic Acids*: 1st ed.; Niedle, S.; Balasubramanian, S., Eds.; RSC Publishing, 2006.

(19) Jin, R.; Gaffney, B. L.; Wang, C.; Jones, R. A.; Breslauer, K. J.: Thermodynamics and structure of a DNA tetraplex: a spectroscopic and calorimetric study of the tetramolecular complexes of d(TG3T) and d(TG3T2G3T). *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 8832-8836.

5 Silver Coordination to Duplex-Mediated G-Quadruplex Assemblies

5.1 Introduction

Although they possess extremely high thermal stabilities, the association rates of multi-molecular G-quadruplexes are typically very slow and strongly influenced by oligonucleotide concentration.[1] Tetramolecular quadruplex association follows forth-order association kinetics.[2] Assembly is very slow, even at high concentrations. The association rate of a six-quartet G-quadruplex, Figure 5.1A, although fast by G-quadruplex standards, still takes more than three hours. It can be seen that while beginning quickly, the rate of assembly slows to an almost negligible rate once the concentration of single-stranded oligonucleotide drops. Another feature typical of all tetramolecular quadruplexes is that assembly will not occur at concentrations below a certain threshold. Figure 5.1B demonstrates the melting and subsequent inability of TG_4T to re-associate following melting. In contrast, the folding of intramolecular G-quadruplexes, being a unimolecular reaction, is independent of concentration. Several groups have developed methods for improving the rates of assembly of G-quadruplexes by joining the ends of oligonucleotides with linkers of various designs, Figure 5.2. These structures combine the structural features of tetramolecular quadruplexes with the association kinetics of unimolecular quadruplexes.[3] The term Template-Assisted Synthetic G-quadruplex (TASQ) has been proposed for these assemblies.

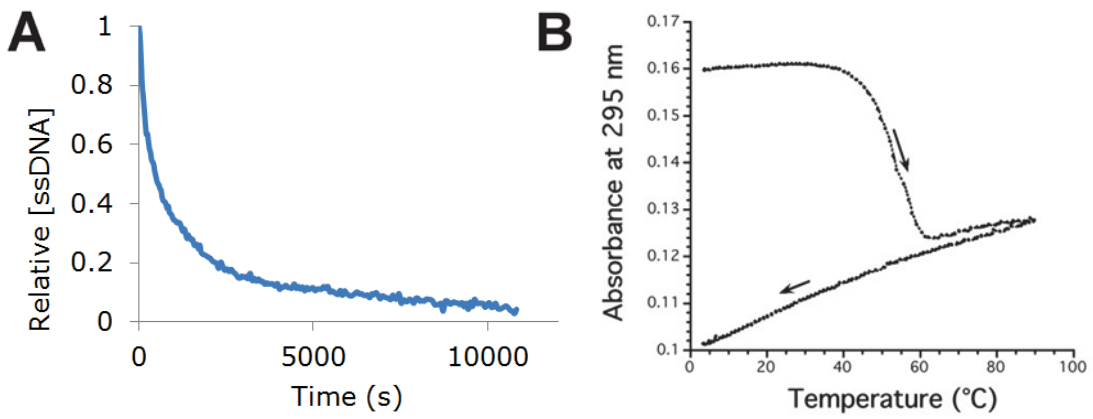


Figure 5.1. Rate of formation of a tetramolecular G-quadruplex containing six G-quartets (10 μM DNA, 20 °C), A. Melting and subsequent failure to reassociate of $(TG_4T)_4$, (20 μM, 110 mM Na^+, 0.18 °C/min), B.[2]

5.1.1 Template-Assisted Synthetic G-Quadruplexes

Several TASQs have been developed based on flexible, branched linkers (Figure 5.2A) but there exist other designs, including examples based on cyclic peptide rings (Figure 5.2B).[4,5] In general, studies have found that parallel G-quadruplexes formed by these assemblies possess a higher thermal stability than comparable un-linked tetramolecular G-quadruplexes and show concentration-independent association rates.

Oliviero et al. studied the 'tetra-end-linkers' of type L and S (Figure 5.2A) in G-quadruplexes based on the sequence TG_4T, and investigated both parallel and anti-parallel strand orientations.[6] Unexpectedly, it was found that tetra-end-linked-oligonucleotides using long linkers of type L invariably formed parallel-stranded G-quadruplexes, including those TASQs specifically designed to promote the formation of antiparallel G-quadruplexes, as in the case of constructs 3 and 4, Figure 5.2C. This behaviour was attributed to the length and flexibility of the linker allowing the formation of structures I-III and V, an assumption that was supported by molecular modelling. G-quadruplexes using these longer linkers were found to have improved stability over comparable tetramolecular complexes. Shorter linkers of type S were subsequently developed to restrict the number of potential topologies available and encourage the formation of structures IV and VI.[7] G-quadruplexes making use of these shorter linkers also formed parallel-stranded structures and were found, in general, to be less thermally stable than those making use of the longer, type L linkers. Molecular modelling suggested the formation of quadruplex species where the central G-quartet core was maintained, albeit distorted, and where the four thymine bases directly attached to the branched linker acted as an extension to the linker. To demonstrate real world applications of these structures, tetra-end-linked-ON analogues of Hotoda's sequence (TGGGAG) were tested, an oligonucleotide with anti-HIV-1 activity.[8] All tetra-end-linked variations of Hotoda's sequence showed higher antiviral activity (EC_{50} between 0.082 and 1.40 μM) when compared with their parent structures (EC_{50} >30 μM); those variants containing type L linkers consistently had higher activity than their type S counterparts.

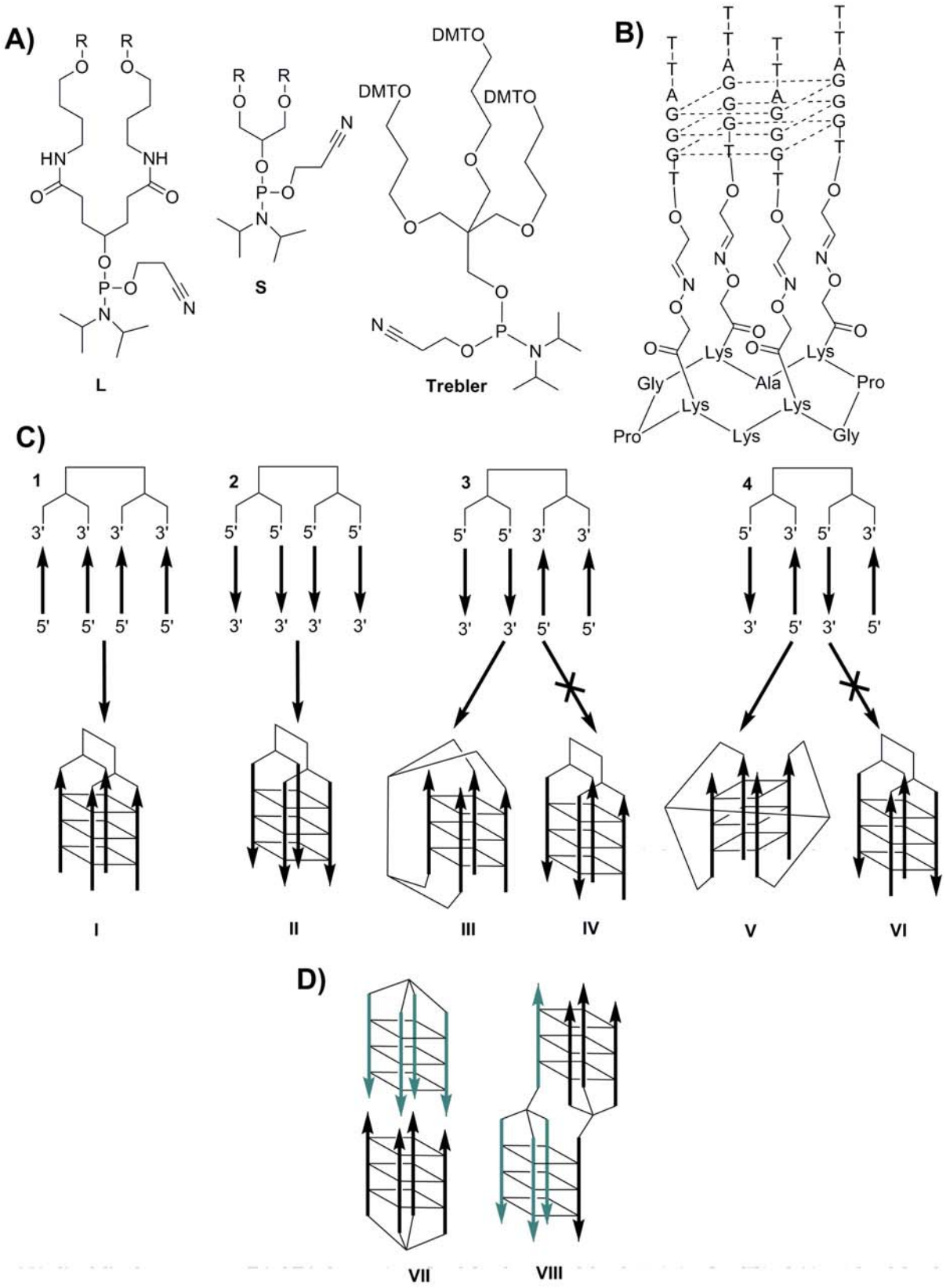


Figure 5.2. Various Template-assisted Synthetic G-quadruplexes (TASQs). The long, L, and short, S, linkers used by Oliviero et al, A.[7-9] TASQ based on cyclic peptide rings, B.[4,5] G-quadruplex topologies investigated by Oliviero et al., C. Structures proposed for dimeric TASQs, D.

Studies were made of tetra-end-linked G-quadruplex structures using the commercially available 'trebler' phosphoramidite shown in Figure 5.2A, which allows the synthesis of branched oligonucleotides. One of the four oligonucleotide strands of each quadruplex was modified by reversing its polarity.[10] These G-quadruplexes were also based on the sequence TG_4T. Parallel G-quadruplexes where all four oligonucleotide strands were linked by either the 3′ or 5′ end were found to have a T_M > 80 °C in $NH4^+$, Na^+ and in K^+. Structures containing one strand with an orientation in opposition to the remaining three were found to be less stable than their parallel counterparts. These antiparallel, tetra-end-linked oligonucleotides formed (3+1) antiparallel structures in NH_4^+. However, when Na^+ or K^+ were present CD spectra indicated the formation of parallel quadruplex structures analogous to those shown in Figure 5.2C. Interestingly, both parallel and anti-parallel structures were observed to form mixtures of both G-quadruplex monomers and dimers, where monomers were the major component. The authors proposed two possible dimeric structures, both shown in Figure 5.2D; 1) Face-to-face stacking of terminal T-tetrads (structure VII), supported by ESI-MS data and previously reported in the literature[11] and 2) For 3+1 G-quadruplexes, a structure where a pair of quadruplexes each exchange their fourth, antiparallel strand with their partner, thus forming two linked, parallel G-quadruplexes (structure VIII).

5.1.2 Concept

The TASQ's described above make use of covalent 'linker' moieties to promote G-quadruplex assembly. In contrast, we sought to use the association properties of DNA duplexes. This chapter details our attempts to combine the association kinetics of DNA duplexes with the thermal stability of G-quadruplexes. DNA duplexes follow extremely fast association kinetics when compared with multimeric G-quadruplexes, and are capable of assembly at much lower concentrations, although the exact association rates depend on temperature and oligonucleotide length.[12] We reasoned that the formation of short DNA duplexes could be used to promote G-quadruplex assembly by preorganising four G-tracts.

5.1.2.1 Features of Target Assemblies

Our basic design involves the self assembly of two oligonucleotides. Each oligonucleotide contains a central duplex-forming section and two G-tracts located at the 5′ and 3′ ends of the sequence. The duplex section serves to promote the association of two single stranded oligonucleotides, bringing together four G-tracts that can then rapidly associate into one of several antiparallel G-quadruplex topologies. A schematic diagram of our target structures is shown in Figure 5.3. The polythymidine linkers provide the flexibility needed for the G-tracts to assemble into a G-quadruplex. The use of a self-complementary palindromic duplex-forming sequence reduces the complexity of the system by eliminating the need to provide a

second, complementary oligonucleotide to complete the system. The intention of this design is to produce an assembly that resembles a tetramolecular G-quadruplex but which follows the association kinetics of a DNA duplex.

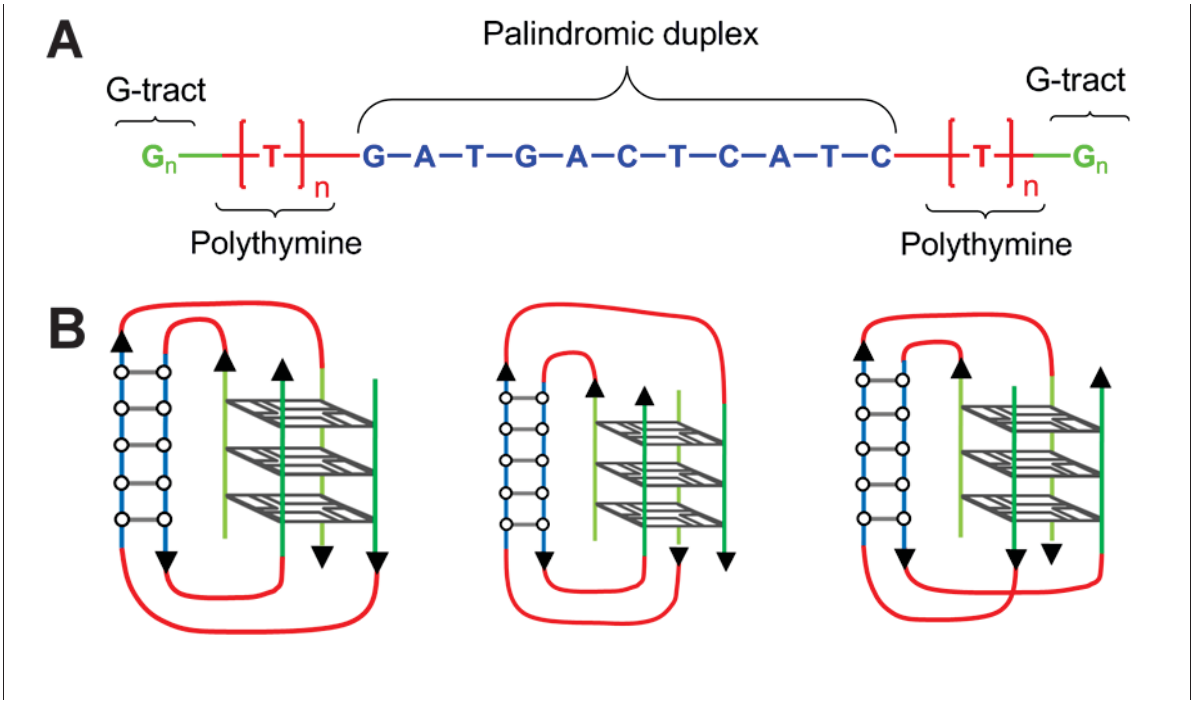


Figure 5.3. Schematic illustration of the target structures. An oligonucleotide containing a central duplex-forming section with two G-quadruplex-forming G-tracts attached via two polythymidine linkers, A. Potential conformations of the target duplex-mediated G-quadruplex assemblies, B.

5.1.2.2 Expected Hysteresis Profile

Ideally, an assembly as described above would have a melting profile that features the high melting temperature of its G-quadruplex component but which can only assemble at temperatures below the T_M of the duplex section. The duplex would promote fast association at low concentrations, but only at low temperatures. An example of such a melting profile is shown in Figure 5.4. The upper and lower temperature ranges of this system could be tuned by altering the length of the duplex and G-quadruplex forming sections.

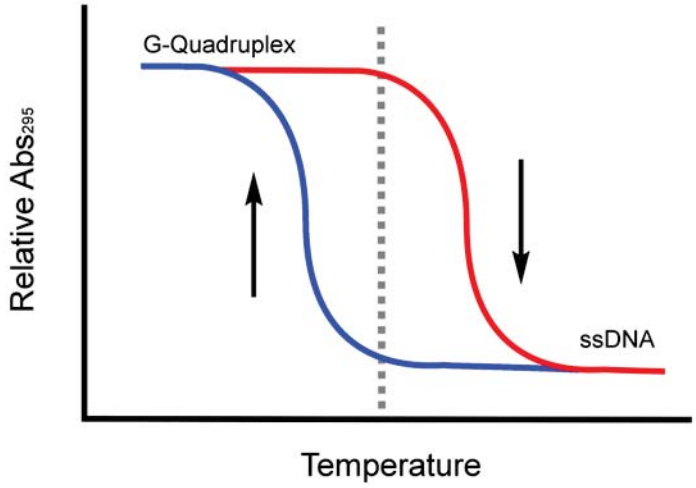


Figure 5.4. A duplex-mediated G-quadruplex assembly would, ideally, feature the thermal stability of a G-quadruplex (red curve) with the assembly properties of a duplex (blue curve). Compare with the assembly of a G-quadruplex alone, Figure 5.1.

5.1.2.3 Silver-Ion Promoted Assembly

One limitation of the system described above is that it can only be controlled by temperature, but no other means. This chapter explores the use of a metal coordinating base pair to prompt duplex formation. Rather than use the 1,2,4-triazole nucleotides detailed elsewhere in this thesis, for this application we have opted to use the ability of a cytosine-cytosine mismatch (Figure 5.5) to coordinate silver ions, a property that has been described extensively in the literature.[13-15] Coordination of a silver ion into one of these mismatched pairs has been reported to increase the T_M of a duplex by up to 5 °C.[16] In this chapter, we have made use of this property by incorporating a cytosine-cytosine mismatch in the centre of the palindromic duplex-forming section. If a suitably sized duplex section is selected, the incorporation of this mismatch will render the resulting duplex unstable when silver ions are absent: the intended result is to promote the spontaneous assembly of the entire G-quadruplex complex upon the addition of silver ions to the sample.

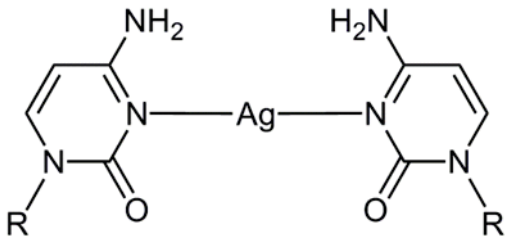


Figure 5.5. A cytosine-silver(I)-cytosine base pair.

5.1.3 Literature Precedents

When we began this project there were, to the best of our knowledge, no examples of duplex-mediated G-quadruplex formation. However, during the course of our research a similar system was developed by Zhou *et al.*, who published a communication on the assembly of a 'tri G-quadruplex'.[17] This system, shown in Figure 5.6, makes use of two short duplexes formed from three oligonucleotides to preorganise four guanine tracts. These oligonucleotides form a duplex in lithium buffer, and assemble into a G-quadruplex when Na^+ ions are added. The assemblies were characterised by PAGE and ESI mass spectrometry, and were found to form discrete structures to the exclusion of other larger aggregates. This assembly is notable as being the first published example of a G-quadruplex consisting of three separate oligonucleotide sequences.

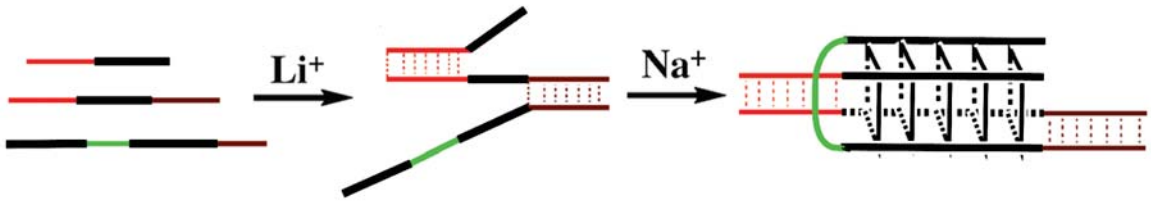


Figure 5.6. The tri-G-quadruplex developed by Zhou et al.[17] This assembly preorganises G-quadruplex assembly by making use of two duplex sections across three oligonucleotides. Complementary duplex sections are coloured red and brown, the flexible linker section is coloured green.

5.2 Results and discussion

5.2.1 Design of Duplex Core

Preliminary experiments were aimed at designing a suitable duplex-forming core for the duplex-mediated G-quadruplex assemblies detailed in Figure 5.3.

Four self-complementary sequences ranging in length from twelve to eight nucleotides were examined, each incorporating a central C/C mismatch. These oligonucleotides were selected based on melting temperatures predicted using the nearest-neighbour method.[18] The use of self-complementary sequences carries the risk that the formation of hairpins competes with the formation of the target bimolecular duplexes. Hairpins require loops with at least three nucleotides and were not expected to be formed from the short oligonucleotides in Table 5.1. We confirmed using PAGE studies (not shown) that hairpins were not formed.

Table 5.1. T_M values (° C) for duplex-forming oligonucleotides determined by CD. Each sequence is palindromic; the central cytosine, underlined, forms a cytosine-cytosine mismatch that is capable of coordinating silver(I) ions. Date were recorded in 10 mM lithium cacodylate buffer (pH 7.2), 100 mM potassium cacodylate, 2 μM DNA.

	Sequence	T_m (°C)		
		Predicted[18]	Actual	+ 1 eq. Ag^+
ON1	AGATGAC̲TCATCT	26.4	41.0	41.0
ON2	GATGAC̲TCATC	15.7	26.0	32.0
ON3	ATGAC̲TCAT	-5.8	-	-
ON4	TGAC̲TCA	-31.6	-	-

A summary of results for UV melting experiments is included in Table 5.1. A 6 °C stabilisation of **ON2** is observed upon the addition of silver. Interestingly, the longer **ON1** shows no similar stabilisation, presumably because the formation of a cytosine-silver-cytosine base pair contributes comparatively less towards the stability of this already stable duplex.

It should be noted that our intention was to select a duplex-forming sequence with marginal stability, allowing stabilisation by the addition of silver(I) to prompt spontaneous assembly. The T_M values in Table 5.1 demonstrate that **ON1** - **5** do not allow sufficient granularity in stability to achieve this goal. This has implications for the assembly of the target complex: the duplex core is likely to be stable under our experimental conditions, even in the absence of silver. This is likely to minimise the influence of silver(I) ions on the assembly.

Using the data in Table 5.1, **ON2**, with a duplex core containing 11 nucleotides was selected as the basis for further sequences as this was the only oligonucleotide that showed a measurable response to Ag^+. The melting temperature for **ON2** was expected to allow T_M experiments of the assembly to distinguish between the melting of the duplex section (26.0 °C) from that of the G-quadruplex section (~55.0 °C in these conditions).[2]

5.2.2 Design of Polythymidine Linker

Using the duplex-forming core **ON2**, we designed a series of oligonucleotides that varied in the length of the polythymidylate linker (Table 5.2). A suitable linker is required to provide enough flexibility to allow the guanine tracts to assemble into a G-quadruplex. A G-tract containing six guanines (expected to have a T_M >95 °C in these conditions)[2] was used in these oligonucleotides to ensure that a G-quadruplex would form if given the opportunity.

Table 5.2. Oligonucleotides with varying lengths of polythymidylate linkers, based on the duplex core of **ON2**.

	Sequence	Linker length
ON5	$G_6T_6GATGACTCATCT_6G_6$	6
ON6	$G_6T_7GATGACTCATCT_7G_6$	7
ON7	$G_6T_8GATGACTCATCT_8G_6$	8
ON8	$G_6T_9GATGACTCATCT_9G_6$	9

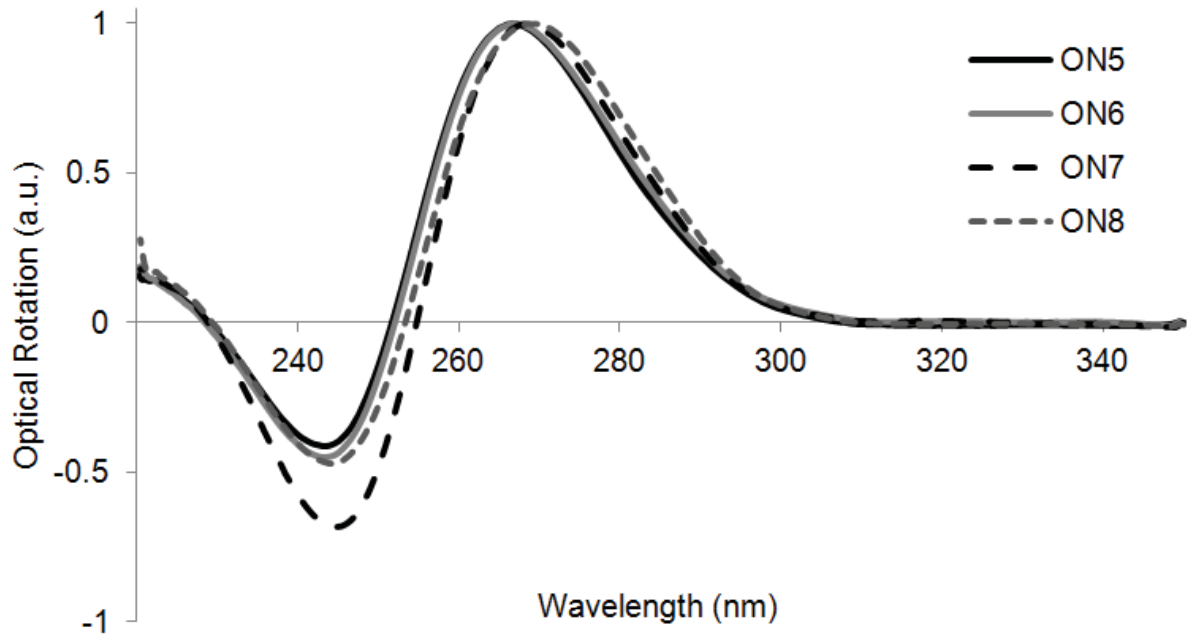


Figure 5.7. CD spectra of **ON5 – 8**. CD spectra were recorded at 20 °C with 2 μM oligonucleotide, 10 mM lithium cacodylate buffer (pH7.2) and 100 mM sodium cacodylate.

CD spectra, Figure 5.7, revealed that all oligonucleotides in this series assembled into parallel G-quadruplexes. The formation of a parallel assembly in preference to the expected antiparallel quadruplex was unexpected, but mirrors the results found by Oliviero et al. as in Figure 5.2.[7] These results indicate that even the shortest six-thymidine linker provides enough flexibility to permit quadruplex assembly. We examined **ON5 – 8** using gel electrophoresis to determine whether these oligonucleotides were, in fact, forming discrete assemblies. The three potential G-quadruplex conformations detailed in Figure 5.3 are expected to have very similar migration rates. The rate of assembly for these structures is expected to be high and so the presence of ssDNA or partially folded structures was considered unlikely. Thus, if **ON5 – 8** form discrete assemblies a single band would be the expected result.

It should be noted that PAGE experiments are, by necessity, carried out at oligonucleotide concentrations much higher than that for CD and UV spectroscopy; as a consequence, care must be taken in comparing this data to that collected by spectroscopy as higher concentrations favour intermolecular species and can lead to G-wires. Gel-Star® fluorescent dye was used to allow a reduction in DNA concentrations as much as possible, allowing visualisation of DNA duplexes and G-quadruplexes at concentrations as low as ~10 μM, which can be more confidently compared with CD data.

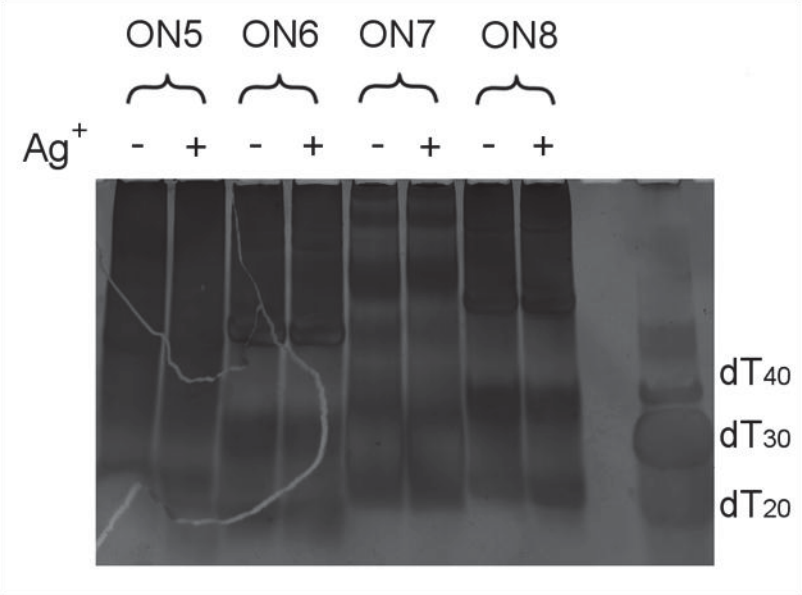


Figure 5.8. Native Gel electrophoresis of **ON5 – 8**. Samples were prepared with 50 μM DNA, 10 mM lithium cacodylate buffer (pH7.2) and 100 mM sodium cacodylate. Silver, when present is included as $AgNO_3$ at 1 equivalent (25 μM). Gel contains 12 % polyacrylamide in TB buffer (pH 7.4) and supplemented with 100 mM $NaNO_3$. Electrophoresis was carried out at 3 °C, 15 W for ~3 hours.

PAGE experiments of **ON5 – 8**, Figure 5.8, revealed that these oligonucleotides formed a large number of side products, visible as smears on the gel. Individual bands represent assemblies with discrete structures; the visible bands on the gel may be accounted for by

some, or all, of the following possibilities: single-stranded DNA, duplexes, hairpins, the target duplex-mediated G-quadruplex or several G-quadruplex side-products. The smears can be attributed to higher-order G-quadruplex products such as G-quadruplex dimers and polymer structures such as G-wires. As noted above, the higher DNA concentrations used for PAGE experiments prevent a direct comparison with being made with CD experiments; however, the extremely long six-guanine G-tracts used in these oligonucleotides are capable of promoting the formation of multimeric G-quadruplex species, even at low concentration.[19] The evidence provided by PAGE gives enough reason to suspect that high-order G-quadruplex species are present in the samples used for CD spectroscopy.

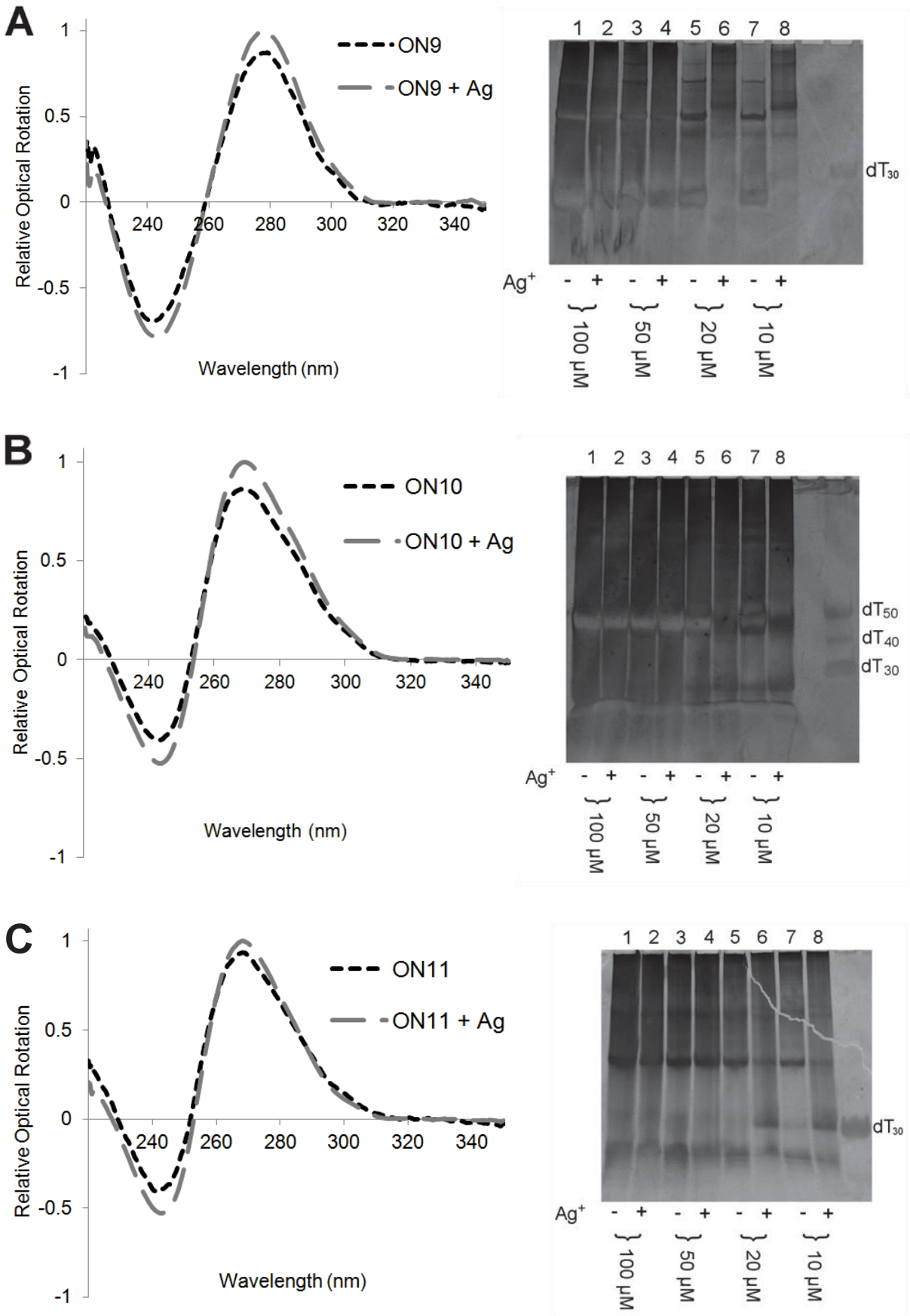
5.2.3 Refinement of G-Tract

With evidence from **ON5 – 8** indicating that oligonucleotides containing six-guanine G-tracts form large, multimeric assemblies, we investigated G-tracts of varying lengths both with and without thymidine caps in an attempt to eliminate the formation of side products. Oligonucleotides containing shorter G-tracts of four- and five-guanines are less likely to form high-order structures. Thymidine caps – thymine nucleotides included at the 5′ or 3′ terminus of a G-tract – block the surface of terminal G-quartets and serve to reduce the likelihood of higher-order G-quadruplex species forming. The oligonucleotides examined are detailed in Table 5.3. We investigated the assembly properties of these oligonucleotides using CD spectroscopy and PAGE.

Table 5.3. Oligonucleotides varying in the length of their G-tract, and the presence or absence of a thymidine cap.

	Sequence	G-tract
ON9	$G_4T_6GATGACTCATCT_6G_4$	G_4
ON10	$G_5T_6GATGACTCATCT_6G_5$	G_5
ON11	$G_6T_6GATGACTCATCT_6G_6$	G_6
ON12	$TG_4T_6GATGACTCATCT_6G_4T$	G_4T
ON13	$TG_5T_6GATGACTCATCT_6G_5T$	G_5T
ON14	$TG_6T_6GATGACTCATCT_6G_6T$	G_6T

The gels in Figure 5.9 show concentrations of 100, 50, 20 and 10 μM for each oligonucleotide in the presence and absence of 1 eq. $AgNO_3$. Samples for these gels were prepared at the stated concentrations in 10 mM lithium cacodylate buffer (pH 7.2) with 100 mM sodium cacodylate and were allowed to anneal for a minimum of 12 hours prior to loading on the gel.


A
Relative Optical Rotation
Wavelength (nm)
ON9
ON9 + Ag
1 2 3 4 5 6 7 8
dT30
Ag+
100 µM
50 µM
20 µM
10 µM
B
ON10
ON10 + Ag
dT50
dT40
dT30
C
ON11
ON11 + Ag

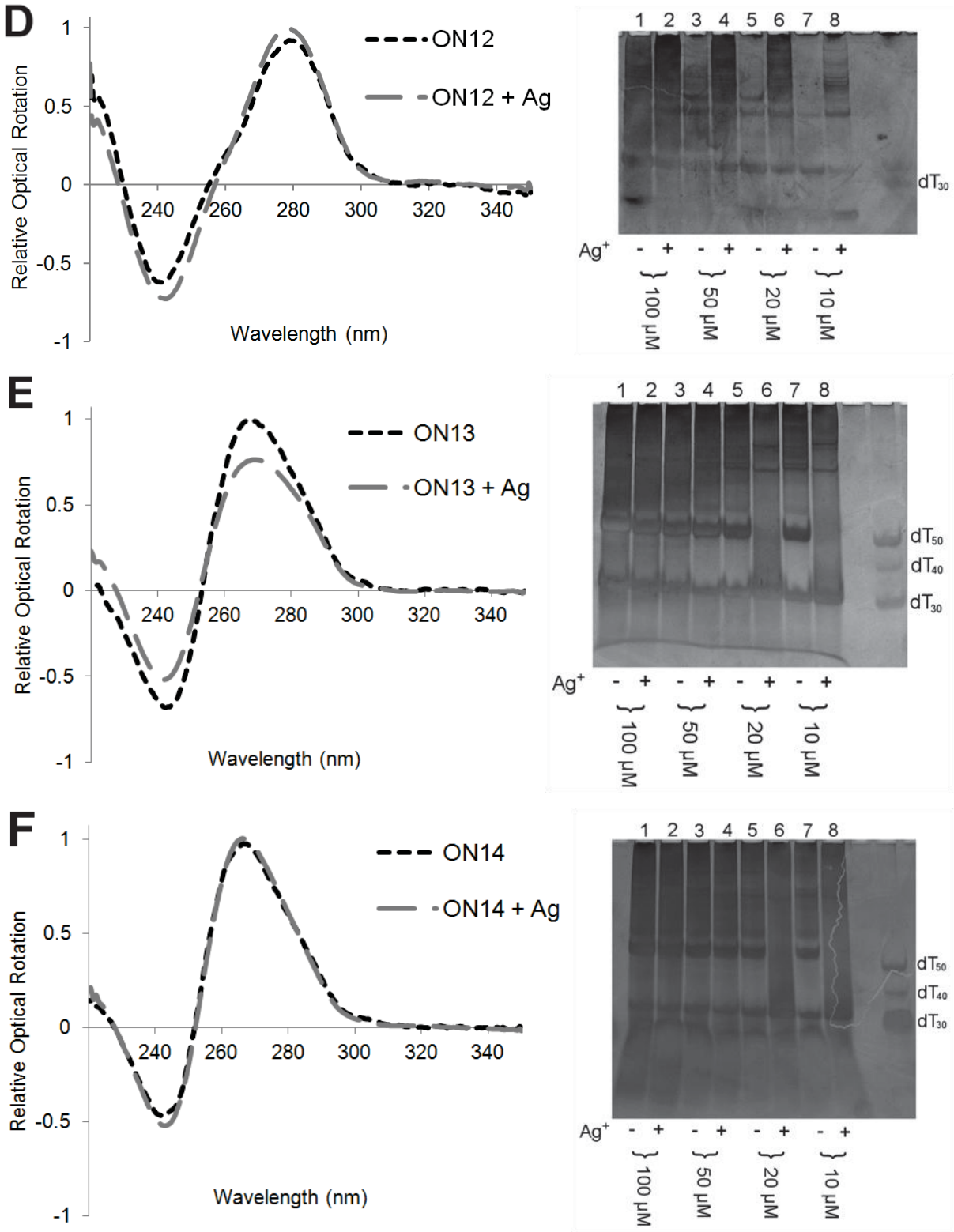


Figure 5.9. CD spectra and native gel electrophoresis of **ON9** – **14**. CD spectra were recorded at 20 °C with 2 µM oligonucleotide (to produce 1 µM duplex), 10 mM lithium cacodylate buffer (pH 7.2) and 100 mM sodium cacodylate. Silver, when present, is included as $AgNO_3$ at 1 equivalent (1 µM). PAGE samples were prepared at the indicated oligonucleotide concentrations with 10 mM lithium cacodylate buffer (pH7.2) and 100 mM sodium cacodylate. Silver, when present is included as $AgNO_3$ at 1 equivalent (50 % of the indicated oligonucleotide concentration). Native gels are 12 % polyacrylamide in TB buffer (pH 7.4) and supplemented with 100 mM $NaNO_3$. Electrophoresis was carried out at 3 °C, 15 W for ~3 hours.

An examination of the CD spectra in Figure 5.9 reveals that the addition of silver(I) ions has no effect on the final structure that is assembled from any of the oligonucleotides: although the maxima differ in intensity, there is no change in their position. It is also possible to determine that **ON9** and **ON12**, containing G-tracts of G_4 and G_4T respectively, form DNA duplexes (maxima at ~280 nm) but do not form G-quadruplexes. Conversely, **ON10** (G_5), **ON11** (G_6), **ON13** (G_5T), and **ON14** (G_6T) have maxima at ~265 nm – characteristic of parallel G-quadruplexes. **ON9** and **ON12** demonstrate that even when brought into close proximity by the duplex, the G-tracts are not in a favourable position to assemble a G-quadruplex; this leaves two possibilities: first, it may be that the polythymidine linkers are too short, distorting the quadruplex core and disrupting the shorter, less stable G-quadruplexes. A counterpoint may be found in comparing, again, the structures of Oliviero et al. where a four-quartet TASQ with very short linkers tolerated a significant degree of structural deformation, including the distortion of a terminal quartet, to achieve a parallel conformation. Second, the linkers may be too long. In this scenario the linkers are so long that the G-tracts have a low probability of meeting and thus behave independently. Longer linkers also introduce a greater number of negatively charged phosphates and so unnecessarily long polythymidine sections may destabilise quadruplex assembly via increased electrostatic repulsion.

PAGE results for **ON9 – 14** reveal smears and a large number of bands, indicating that a large number of side products are present, despite our efforts to reduce their formation. The presence of such a large number of assemblies in these PAGE experiments at DNA concentrations as low as 10 μM suggests that many of these assemblies are also be present in the CD samples (recorded at 2 μM). This indicates that the apparent CD spectra for each oligonucleotide is the sum of many individual components, i.e. the observed spectra show that parallel G-quadruplex species predominate in solution, not that a single parallel G-quadruplex species is present nor, most importantly, that other forms are absent. The presence of a mixture of different assemblies at 2 μM is supported by UV melting experiments which provide complex melting curves that defy T_M analysis, suggesting a complex mixture in solution.

The data provided by **ON9 – 14** can be considered in two contexts: the first in terms of G-tract length, the second by the presence or absence of the terminal thymidine cap. In general terms, a longer G-tract results in a more stable G-quadruplex and makes an oligonucleotide more likely to form higher-ordered structures. A thymine cap reduces the likelihood of high-ordered structures at the expense of slower G-quadruplex association.[20] It is interesting then, that the comparisons of **ON9** vs. **ON12**, **ON10** vs. **ON13** and **ON11** vs. **ON14** suggests that the presence of thymine caps has a negligible influence on the formation of higher-ordered

structures: both oligonucleotides in each pair display smears to a similar extent; it appears that at these concentrations the tendency for G-quadruplexes to aggregate overcomes the ability of a thymine cap to resist this assembly.

For most of these sequences a response to the presence of silver(I) ions is only noticeable at concentrations below 20 μM (lanes 5-8 in each of the gels). This is an unusual result considering that the proportion of silver remains the same in all situations. **ON12** is the exception, and shows influence at all concentrations. The most significant effect, visible in for all oligonucleotides, is an increase in the length of the smear in each of the silver-containing lanes; it is likely that this is due to the disruption of G-quadruplexes by the coordination of silver to the amine groups of guanine.[21]

As described above, **ON9** and **ON12** appear to form duplexes, but not G-quadruplexes under the conditions used for CD experiments. This suggests that discrete, target assemblies formed from **ON9** and **ON12** are unlikely to be present in PAGE experiments. The bands observed in the gels for **ON9** and **ON12** must then be dominated by higher-order G-quadruplexes, particularly as higher DNA concentrations favour intermolecular assembly. It was our expectation that the rapid assembly of the central DNA duplex would promote the formation of the discrete target structures before the formation of higher-order structures could take place – if the duplex *were* contributing to the rapid assembly of the bimolecular quadruplex we would expect to see this assembly dominate at the expense of higher-order structures, particularly at lower DNA concentrations. This is not observed. High-order intermolecular assemblies are clearly a major product indicating that the assembly rates of the target bimolecular quadruplexes are on a similar time scale as competing higher-order complexes. Duplex formation does not appear to be a large contributor to assembly of these G-quadruplexes, even when a duplex has preorganised their assembly.

In order to slow the rate of G-quadruplex assembly, relative to duplex formation, we prepared samples of **ON9 – 14** in lithium buffer (conditions not favourable for G-quadruplex formation) in the presence or absence of silver(I) ions. Na^+ ions were added after a delay of twelve hours to prompt quadruplex assembly. This procedure replicates the techniques used by Zhou *et al.* for their tri-G-quadruplex assemblies, Figure 5.6, which resulted in the formation of discrete assemblies. The results of our experiment are detailed in Figure 5.10. It can be seen that Ag^+ has no significant impact on the results. It also appears that even with delayed G-quadruplex formation a large number of side products are formed. This result indicates that the assembly of high-order structures is the result of aggregation *after* duplex formation. This fact, together with the evidence provided by **ON9** and **ON12**, strongly suggests that although duplex assembly does occur rapidly and at low concentration

it plays no role in the formation of G-quadruplex assemblies. This conclusion also explains the minimal impact of Ag^+ on these assemblies, i.e. the stabilisation of the duplex-forming section by Ag^+ will have no observed impact if those duplexes do not effect G-quadruplex formation.

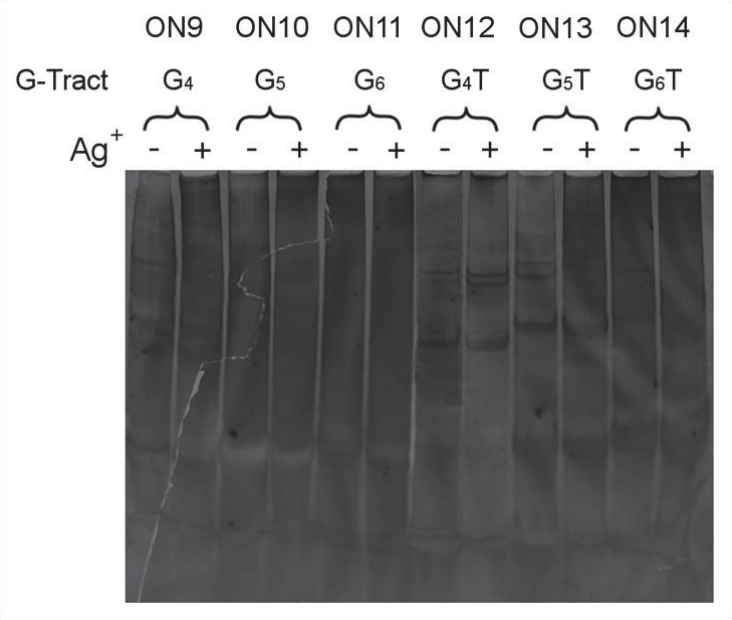


Figure 5.10. Delayed G-quadruplex assembly. 20 μM DNA samples were prepared at in 10 mM lithium cacodylate buffer (pH7.2) in the presence or absence of $AgNO_3$ at 1 equivalent (10 μM). 100 mM sodium cacodylate was added to samples after 12 hours, which were then allowed to assemble for an additional 2 hours before electrophoresis. Gels are 12 % polyacrylamide in TB buffer (pH 7.4) and supplemented with 100 mM $NaNO_3$. Electrophoresis was carried out at 3 °C, 15 W for ~3 hours.

5.2.4 G-quadruplex Assembly in Non-Duplex-Forming Oligonucleotides

With evidence indicating that duplex formation does not contribute to G-quadruplex assembly, we performed experiments on the oligonucleotides detailed in Table 5.4, which have had their duplex forming component replaced by a polythymidine tract. **ON15** and **ON16** have G-tracts comparable to **ON9** and **ON10**, respectively.

Table 5.4. Oligonucleotides featuring polythymidine in place of duplex-forming sequences.

	Sequence
ON15	$G_4T_{23}G_4$
ON16	$G_5T_{23}G_5$

PAGE experiments carried out on **ON15** and **ON16**, Figure 5.11, demonstrate the formation of a large number of high-order DNA structures, similar to those observed for **ON5 – 14**. Interestingly, it appears that the duplex-forming section of DNA may actually cause the formation of a greater number of side products. This is most evident in a comparison of

ON15 with **ON9**, which reveals that the duplex-forming **ON9** produces a greater number of high-order assemblies. In summary, the similarity of **ON15** and **ON16** to **ON9** and **ON10**, as demonstrated by PAGE experiments, suggests that duplex formation does not have a significant role in the assembly of these oligonucleotides.

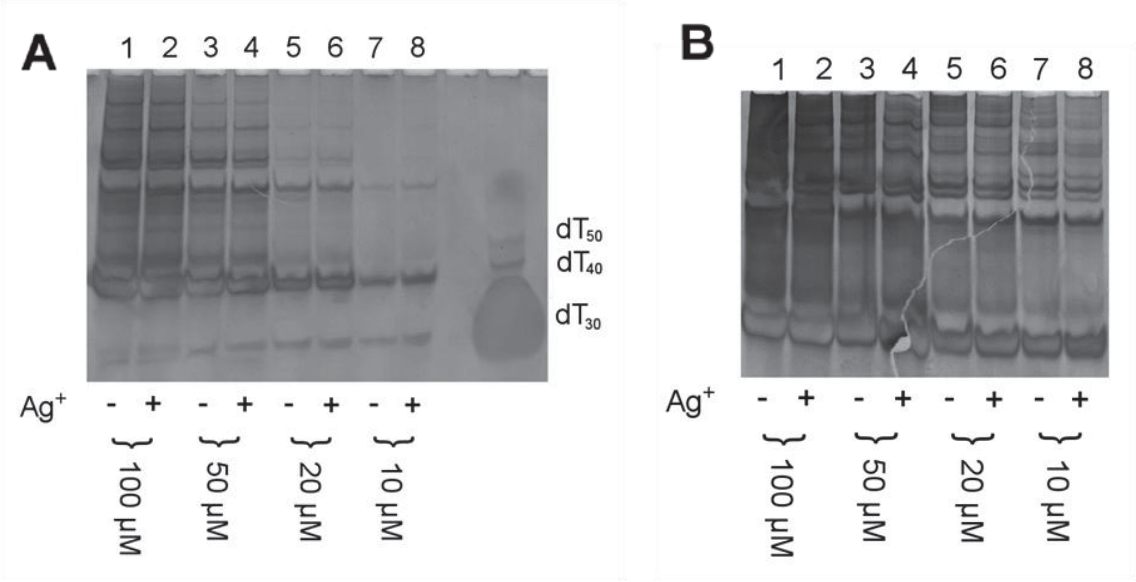


Figure 5.11. Gel electrophoresis of **ON15**, A, and **ON16**, B. Samples were prepared at the indicated oligonucleotide concentrations with 10 mM lithium cacodylate buffer (pH 7.2) and 100 mM sodium cacodylate. Silver, when present is included as $AgNO_3$ at 1 equivalent (50 % of the indicated oligonucleotide concentration). Gels are 12 % polyacrylamide in TB buffer (pH 7.4) and supplemented with 100 mM $NaNO_3$. Electrophoresis was carried out at 3 °C, 15 W for ~3 hours.

5.3 Conclusion

The inclusion of a duplex-forming DNA sequence into the centre of a pair of guanine tracts, as outlined in Figure 5.3, does not promote the formation of a discrete bimolecular G-quadruplex. Instead, a comparison of **ON15** with **ON9** suggests that the duplex-forming section may actually encourage the assembly of higher-order G-quadruplexes. The coordination of silver(I) ions to the cytosine-cytosine mismatch does stabilise the isolated duplex core, as in the case of **ON2**; however, the full oligonucleotides with G-tracts and polythymidine linkers (**ON5 – 14**) are capable of forming a duplex even in the absence of silver ions. The only effect the presence of silver ions appears to have on **ON5 – 14** is the partial disruption of G-quadruplex assemblies.

5.4 References

(1) Wyatt, J. R.; Davis, P. W.; Freier, S. M.: Kinetics of G-quartet-mediated tetramer formation. *Biochemistry* **1996**, *35*, 8002-8.

(2) Mergny, J.-L.; De Cian, A.; Ghelab, A.; Sacca , B.; Lacroix, L.; Cian, A. D.; Sacca, B.: Kinetics of tetramolecular quadruplexes. *Nucleic Acids Res.* **2005**, *33*, 81-94.

(3) Nikan, M.; Sherman, J. C.: Template-assembled synthetic G-quartets (TASQs). *Angew. Chem. Int. Ed.* **2008**, *47*, 4900-2.

(4) Murat, P.; Gennaro, B.; Garcia, J.; Spinelli, N.; Dumy, P.; Defrancq, E.: The Use of a Peptidic Scaffold for the Formation of Stable Guanine Tetrads: Control of a H-bonded Pattern in Water. *Chem. Eur. J.* **2011**, *17*, 5791-5795.

(5) Murat, P.; Cressend, D.; Spinelli, N.; Van der Heyden, A.; Labbé, P.; Dumy, P.; Defrancq, E.: A novel conformationally constrained parallel G-quadruplex. *ChemBioChem* **2008**, *9*, 2588-91.

(6) Oliviero, G.; Amato, J.; Borbone, N.; Galeone, A.; Petraccone, L.; Varra, M.; Piccialli, G.; Mayol, L.: Synthesis and characterization of monomolecular DNA G-quadruplexes formed by tetra-end-linked oligonucleotides. *Bioconjugate Chem.* **2006**, *17*, 889-98.

(7) Oliviero, G.; Borbone, N.; Amato, J.; D'Errico, S.; Galeone, A.; Piccialli, G.; Varra, M.; Mayol, L.: Synthesis of quadruplex-forming tetra-end-linked oligonucleotides: effects of the linker size on quadruplex topology and stability. *Biopolymers* **2009**, *91*, 466-77.

(8) Oliviero, G.; Amato, J.; Borbone, N.; D'Errico, S.; Galeone, A.; Mayol, L.; Haider, S.; Olubiyi, O.; Hoorelbeke, B.; Balzarini, J.; Piccialli, G.: Tetra-end-linked oligonucleotides forming DNA G-quadruplexes: a new class of aptamers showing anti-HIV activity. *Chem. Commun.* **2010**, *46*, 8971-3.

(9) Oliviero, G.; Borbone, N.; Mayol, L.: Synthesis and Characterization of DNA Quadruplexes Containing T-Tetrads Formed by. *Biopolymers* **2006**, *81*, 194-201.

(10) Ferreira, R.; Alvira, M.; Aviñó, A.; Gómez-Pinto, I.; González, C.; Gabelica, V.; Eritja, R.: Synthesis and Structural Characterization of Stable Branched DNA G-Quadruplexes Using the Trebler Phosphoramidite. *ChemistryOpen* **2012**, *1*, 106-114.

(11) Cáceres, C.; Wright, G.; Gouyette, C.; Parkinson, G.; Subirana, J. A.: A thymine tetrad in d(TGGGGT) quadruplexes stabilized with Tl+/Na+ ions. *Nucleic Acids Res.* **2004**, *32*, 1097-1102.

(12) Ouldridge, T. E.; Šulc, P.; Romano, F.; Doye, J. P. K.; Louis, A. A.: DNA hybridization kinetics: zippering, internal displacement and sequence dependence. *Nucleic Acids Res.* **2013**.

(13) Scharf, P.; Müller, J.: Nucleic Acids With Metal-Mediated Base Pairs and Their Applications. *ChemPlusChem* **2013**, *78*, 20-34.

(14) Okamoto, I.; Iwamoto, K.; Watanabe, Y.; Miyake, Y.; Ono, A.: Metal-ion selectivity of chemically modified uracil pairs in DNA duplexes. *Angew. Chem. Int. Ed.* **2009**, *48*, 1648-51.

(15) Knobloch, B.; Linert, W.; Sigel, H.; Gray, H. B.; Linertt, W.: Metal ion-binding properties of uridine , related in aqueous nucleosides solution. *Proc. Natl. Acad. Sci. USA*, *102*, 7459-7464.

(16) Muller, J.; Bohme, D.; Lax, P.; Morell Cerda , M.; Roitzsch, M.: Metal ion coordination to azole nucleosides. *Chem. Eur. J.* **2005**, *11*, 6246-53.

(17) Zhou, J.; Bourdoncle, A.; Rosu, F.; Gabelica, V.; Mergny, J.-L.: Tri-G-Quadruplex: Controlled Assembly of a G-Quadruplex Structure from Three G-Rich Strands. *Angew. Chem.* **2012**, *124*, 11164-11167.

(18) Allawi, H. T.; SantaLucia, J.: Thermodynamics and NMR of Internal G·T Mismatches in DNA. *Biochemistry* **1997**, *36*, 10581-10594.

(19) Smargiasso, N.; Rosu, F.; Hsia, W.; Colson, P.; Baker, E. S.; Bowers, M. T.; De Pauw, E.; Gabelica, V.: G-Quadruplex DNA Assemblies: Loop Length, Cation Identity, and Multimer Formation. *J. Am. Chem. Soc.* **2008**, *130*, 10208-10216.

(20) *Quadruplex Nucleic Acids*: 1st ed.; Niedle, S.; Balasubramanian, S., Eds.; RSC Publishing, 2006.

(21) Blume, S. W.; Guarcello, V.; Zacharias, W.; Miller, D. M.: Divalent transition metal cations counteract potassium-induced quadruplex assembly of oligo(dG) sequences. *Nucleic Acids Res.* **1997**, *25*, 617-25.

6 Future Directions

Our attempts to use metal coordination to direct the assembly of DNA structures has been largely unsuccessful. The reasons for this lack of success vary somewhat between each of the projects described in this thesis, but common problems can be observed that must be addressed for the future development of similar systems.

The methods used in this thesis provide only an averaged picture of all components in a system. It has become clear that if metal-coordination to 1,2,4-triazole nucleotides does in fact occur, then these complexes represent only a small fraction of the assemblies in the system. ESI mass spectrometry may offer a means to identify low levels of metal-coordination to G-quadruplex assemblies. ESI-MS has the advantage characterising individual assemblies in a mixture of species at a resolution sufficient to identify the coordination of single metal-ions. A drawback of ESI is the requirement for volatile components. For G-quadruplexes this typically requires the use of ammonium salts in place of sodium or potassium, which can change the stability of quadruplexes and can also influence their topology. We have previously attempted to carry out ESI mass spectrometry on several of the oligonucleotide assemblies described in this thesis, although we were able to observe the parallel $(TG_4T)_4$ G-quadruplex we were only able to observe ssDNA for the other oligonucleotides tested. Nevertheless, ESI-MS remains an attractive method for identifying low levels of metal-coordination.

For the assemblies described in this thesis, the destabilising effect of transition metals on G-quartets and quadruplexes are apparent before any coordination to triazole nucleotides is observed. Guanine bases are more numerous and each provides several possible coordination sites, the majority of which are necessary for G-quartet assembly. A metal-coordinating nucleotide must be able to compete successfully with guanine. We developed the G-quadruplexes in this thesis to provide a range of coordination environments with differing levels of preorganisation. It has become clear, however, that 1,2,4-triazole nucleotides are not capable of effectively competing for metal coordination in any of the situations tested. Future experiment on these G-quadruplex systems should focus on nucleotides which are capable of forming more stable metal-ligand bonds. The use of chelating moieties is appealing, but places limitations on the type of assemblies that could be investigated. It has been noted that the affinity of a ligand towards a metal ion generally increases with higher pK_a values.[1] Imidazole nucleotides, which have a higher pK_a than 1,2,4-triazole nucleotides (6.42 vs. 1.53)[2,3] may be able to more effectively compete. However, we have not been able to observe metal coordination to imidazole nucleotides in our own, limited, studies. It should also be noted that higher pK_a values also result in increased

competition between metal coordination and protonation. pK_a values are not the sole determining factor in the coordination of metals to these nucleotides.

Although we were not able to demonstrate metal-coordination to triazole nucleotides using the 3+1 G-quadruplex described in Chapter 2, it remains an interesting system for further study of metal-coordinating nucleotides. We had expected that the core, three G-quartet quadruplex would maintain a degree of stability in the absence of the A/T Watson-Crick base pair, and would thereby provide a degree of preorganisation for the formation of a metal base pair. The removal of the A/T base pair destabilised the quadruplex, eliminated any preorganisation and created a barrier to metal coordination that 1,2,4-triazole nucleotides were not able to overcome. However, it remains to be seen whether assembly of the 3+1 G-quadruplex could be encouraged by an artificial nucleobase capable of forming a more stable metal-mediated base pair.

The *htel* G-quadruplexes described in Chapter 3 take part in complex equilibria between different conformations. These equilibria are themselves the focus of study for several research groups. The use of metal-coordinating nucleotides to influence these systems is worthy of further study. The apparent differences in silver(I) coordination ability between the two 4*htel* sequences **ON1** and **ON2** is, at present, puzzling. Further work is needed to ensure that our results are reproducible. However, the information available to us suggests that both oligonucleotides share similar conformations and that silver(I) coordination does not occur at the triazole nucleotides. The difference in behaviour may be caused by small-scale differences in structure which require more detailed structural characterisation of each assembly.

We were unable to observe the formation of metal-ion coordinating quartets in the tetramolecular G-quadruplexes detailed in Chapter 4, however molecular modelling suggests that the concept itself remains sound. As with the other systems described in this thesis, these tetramolecular assemblies are likely to benefit from the use of a metal-coordinating nucleotide that can more effectively compete with guanine bases. Metal-coordinating quartets are likely to be sensitive to changes in the shape of nucleotides. Comparatively small 1,2,4-triazole ligands still occupy a large space due to the need to assume a propeller conformation in a quartet. The formation of a metal-coordinating quartet that remains completely coplanar with surrounding G-quartets is an interesting target for further research.

Chapter 5 contains a clear example of metal-coordination leading to improved thermal stability, if only for silver coordination to a short duplex containing a cytosine-cytosine mismatch, see page 103. However, the complete oligonucleotides (containing

polythymidylate linkers and guanine tracts) showed the same destabilising effect of transition metal ions observed for other assemblies in this thesis. This example demonstrates the challenges facing the targeted coordination of metal ions when a large number of guanine bases are present.

The use of metal-coordinating nucleotides remains a promising field of study. Recent discoveries of the biological relevance of G-quadruplexes and rising interest in their use as pharmaceuticals are likely to increase the importance of developing methods for controlling oligonucleotide assembly. Metal-coordinating nucleotides have the potential to provide a predictable, bioorthogonal method of improving the thermodynamic stability and biological activity of G-quadruplex assemblies.

6.1 References

(1) Müller, J.: Metal-Ion-Mediated Base Pairs in Nucleic Acids. *Eur. J. Inorg. Chem.* **2008**, *2008*, 3749-3763.

(2) Muller, J.; Bohme, D.; Lax, P.; Morell Cerda , M.; Roitzsch, M.: Metal ion coordination to azole nucleosides. *Chem. Eur. J.* **2005**, *11*, 6246-53.

(3) Müller, J.; Böhme, D.; Düpre, N.; Mehring, M.; Polonius, F.-A.: Differential reactivity of α and β 2′-deoxyribonucleosides towards protonation and metalation. *J. Inorg. Biochem.* **2007**, *101*, 470-476.

7 Experimental

7.1 Oligonucleotide Synthesis

Oligonucleotides were synthesised at 1 μmol scale using a MerMade 4 DNA synthesiser from Bioautomation Corporation. Standard DNA phosphoramidites were purchased from Innovassynth Technologies, India. Triazole phosphoramidites were synthesised as described below, and were coupled using extended coupling times (5 minutes) to ensure high yields. The DMT profiles, provided by the DNA synthesiser, showed synthetic efficiency >99 % for each of the oligonucleotides described in this thesis. The quantity of each oligonucleotide obtained can be found in the table on the following page. Oligonucleotides were cleaved from CPG solid support using 32 % aqueous ammonia for approximately 1 hour at room temperature then 12 hours at 65 °C. This ammonia solution was directly loaded onto, and purified by HPLC using a reverse-phase Phenomenex Clarity 5 μm Oligo-RP column (100×10 mm) and eluted using a gradient of acetonitrile in 0.05 mM triethylammonium acetate buffer. Oligonucleotide masses were confirmed using a MassLynx MALDI-TOF mass-spectrometer. DNA samples were loaded onto a plate with anthranilic acid or 6-aza-2-thiothymine as matrix and dibasic ammonium citrate as co-matrix.

Chapter 2: Expected mass **ON1:** 1776.2 $gmol^{-1}$, observed: 1779.8 $gmol^{-1}$.

Chapter 4: Expected mass **ON1:** 2104.3 $gmol^{-1}$, observed: 2110.8 $gmol^{-1}$; expected mass **ON2**: 2104.3 $gmol^{-1}$, observed: 2109.4 $gmol^{-1}$; expected mass **ON3**: 2350.5 $gmol^{-1}$, observed: 2356.9 $gmol^{-1}$; expected mass **ON4**: 2462.5 $gmol^{-1}$, observed: 2471.2 $gmol^{-1}$expected mass **ON5**: 3118.9 $gmol^{-1}$, observed: 3130.4 $gmol^{-1}$.

Oligonucleotides described in Chapter 5 were purchased from Integrated DNA Technologies. Denaturing PAGE was used to check the purity of these products.

It should be noted that **ON3** (Chapter 2) and **ON1** and **ON2** (Chapter 3) are too large to be observed using MALDI-TOF equipment and awaiting mass confirmation by ESI. DMT profiles, PAGE and HPLC studies have indicated that synthesis of these oligonucleotides was successful.

Table 7.1. Available quantities of oligonucleotides after synthesis and purification. Note: oligonucleotides that do not contain 1,2,4-triazole nucleotides were purchased commercially, quantities for these oligonucleotides are presented as received from the supplier. Quantity was determined by measuring UV-absorbance at 260 nm.

		Sequence	Mass obtained (µg)
Chapter 2	**ON1**	$G_3T_2AG_3TTAG_3T$	1400
	ON2	TAG_3T	188
	ON3	$G_3T_2AG_3T\mathbf{T^r}AG_3T$	122
	ON4	$T\mathbf{T^r}G_3T$	175
Chapter 3	**ON1**	$\mathbf{T^r}AG_3T_2AG_3T_TAG_3T_TAG_3$	751
	ON2	$\mathbf{T^r}AG_3T_2AG_3T_TAG_3T_TAG_3\mathbf{T^r}$	324
	ON3	$\mathbf{T^r}AG_3T_2AG_3T$	166
	ON4	$\mathbf{T^r}AG_3T_2AG_3\mathbf{T^r}$	87
Chapter 4	**ON1**	$TG_4\mathbf{T^r}T$	332
	ON2	$T\mathbf{T^r}G_4T$	275
	ON3	$T\mathbf{T^r}G_4\mathbf{T^r}T$	177
	ON4	$TG_2\_\mathbf{T^r}\_G_2T$	523
	ON5	$TG_3\_\mathbf{T^r}\_G_3T$	809
Chapter 5	**ON1**	AGATGA<u>C</u>TCATCT	222
	ON2	GATGA<u>C</u>TCATC	251
	ON3	ATGA<u>C</u>TCAT	275
	ON4	TGA<u>C</u>TCA	239
	ON5	G_6T_6GATGA<u>C</u>TCATCT_6G_6	462
	ON6	G_6T_7GATGA<u>C</u>TCATCT_7G_6	417
	ON7	G_6T_8GATGA<u>C</u>TCATCT_8G_6	287
	ON8	G_6T_9GATGA<u>C</u>TCATCT_9G_6	490
	ON9	G_4T_6GATGACTCATCT_6G_4	534
	ON10	G_5T_6GATGACTCATCT_6G_5	332
	ON11	G_6T_6GATGACTCATCT_6G_6	462
	ON12	TG_4T_6GATGACTCATCT_6G_4T	581
	ON13	TG_5T_6GATGACTCATCT_6G_5T	558
	ON14	TG_6T_6GATGACTCATCT_6G_6T	728
	ON15	$G_4T_{23}G_4$	277
	ON16	$G_5T_{23}G_5$	242

7.2 Buffer Preparation

Metal-ion impurities were removed from water and buffers used in the preparation of oligonucleotide samples using Chelex-100® resin from Bio-Rad Laboratories, Inc. This resin contains iminodiacetic acid groups and serves to chelate many divalent metal ions, including Pd^{2+}, Cu^{2+}, Fe^{2+}, Ni^{2+}, Pb^{2+}, Mn^{2+}, Ca^{2+}, Mg^{2+}, Co^{2+}, Cd^{2+}, Ba^{2+} and Zn^{2+}. 5 g of resin per 100 mL of solution was used. Solutions were kept over resin for ~12 hours and the resin was filtered off.

7.3 UV/Vis Spectroscopy

UV/Vis data were recorded using Varian Cary 100 and Cary 1 spectrophotometers. Samples were denatured by heating to 95 °C for five minutes and annealed by slow cooling to room temperature. Samples were left at r.t. for a minimum of 8 hours prior to data collection. A ramp of 0.5 °C/min per minute was used for melting experiments with data being recorded at wavelengths of 260, 275 and 295 nm at 1 °C increments. Where possible, $T_{1/2}$'s were calculated as the first derivative of the melting curves using software provided with the spectrophotometer. Alternatively, $T_{1/2}$'s were calculated by defining the maximum absorbance of a G-quadruplex sample as 0 % folded, defining the minimum absorbance as 100% folded: $T_{1/2}$ values are determined from the point of the resulting melting curve where 50 % of the DNA is folded.

Association data were recorded using Varian Cary 100 and Cary 1 spectrophotometers with a 1 cm pathlength. Samples contained 1 mL of a solution of 10 μM DNA, 10 mM lithium cacodylate buffer (pH 7.2), 100 mM potassium cacodylate. Samples were heated to 90 °C for ~20 minutes to ensure complete denaturation and were then cooled to 25 °C within 30 seconds with the aid of an ice/salt mixture. Data was recorded at 25 °C at wavelengths of 260, 275 and 295 nm at 30 second increments for the first 30 minutes and at 1 minutes increments thereafter. k_{on} values were determined as detailed in Chapter 1, and were modelled using Kaleidograph® from Synergy Software.

7.4 Circular Dichroism Spectroscopy

CD data were recorded using an Applied Photophysics Chirascan instrument. Samples were denatured by heating to 95 °C for five minutes and annealed by slow cooling to room temperature. Samples were left at r.t. for a minimum of 8 hours prior to data collection. CD spectra were recorded in increments of 5 °C using a temperature ramp of 1 °C/min. Once the temperature had reached each increment samples were allowed five minutes settling time to ensure thermal equilibration before data acquisition. CD spectra recorded in this way, for both melting and annealing processes, were used to determine melting temperatures. $T_{1/2}$ values were calculated by defining the maximum elipticity of a G-quadruplex sample at 260

nm as 100 % folded and defining the minimum elipticity as 0 % folded: $T_{1/2}$ is determined at the point of a melting curve where 50 % of the DNA is folded.

7.5 Polyacrylamide Gel Electrophoresis

Unless otherwise noted, PAGE experiments were carried out using 12 % acrylamide gels with TB buffer (pH 8.0). Gels were run at 3 °C at 15 W for ~4 hours. Oligonucleotide samples were prepared as described in figures. Samples were denatured by heating to 95 °C for five minutes and annealed by slow cooling to room temperature. Samples were left at r.t. for a minimum of 8 hours prior to loading on the gel. Gels and running buffer were supplemented with either 100 mM $NaNO_3$ or KNO_3, as described in each figure. Stains-all dye was used to visualise all gels in Chapters 2, 3 and 4. Gel-star was used to develop the gels described in Chapter 5.

7.6 NMR Spectroscopy

NMR spectra for all oligonucleotide samples were recorded using a Bruker 700 MHz NMR spectrometer. Samples were prepared in 10 mM Na_2HPO_4 buffer (pH 7.0), 10 % D_2O. Oligonucleotide concentrations of 87-330 μM were used. Deuterated trimethylsilyl propanoic acid (3-(trimethylsilyl)-2,2',3,3'-tetradeuteropropionic acid, TSP) was included at 5 μM as an internal standard. All spectra presented in this thesis were carried out in the presence of 100 mM $NaClO_4$. We attempted to record spectra of **ON1** – **4**, Chapter 3, in 100 mM potassium cacodylate, however sample aggregation in these conditions was found to result in poor spectra. For silver titrations, $AgNO_3$ was introduced as an aqueous solution using small volumes to reduce sample dilution. The $AgNO_3$ concentration was dependent on concentration of oligonucleotide in the sample and was prepared so that 2 μL contained 0.2 eq. Ag^+. Samples were mixed and left at r.t. for ~12 hours prior to data collection.

7.7 Molecular Modelling

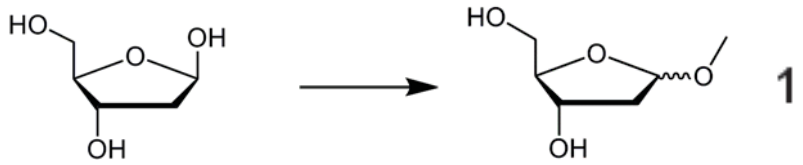
Molecular Modelling for the 3+1 assembly described in Chapter 2 was based on the PDB file 2AQY. This structure contains a bimolecular quadruplex where three G-stretches are provided by one oligonucleotide (sequence GGGTTAGGGTTAGGGT) and the remaining G-stretch is provided by another short oligonucleotide (sequence TAGGGT). An A-T Watson-Crick base-pair exists in this structure on one surface of the G-quadruplex stack (bases underlined in the sequences above). For our calculations the A-T base pair was replaced with two 1,2,4-triazole nucleotides. The adenine and thymine bases were removed leaving the 2-deoxyribose sugars in their original conformations. Triazole rings were added to these deoxyribose sugars, attached through N1. An amber* minimisation was carried out on the structure with water as solvent, constant dielectric electrostatic treatment with a dielectric constant of 1.0, extended cutoff and charges from force field. Because Amber* cannot handle

Ag atoms, the distance between N4 of both triazole rings was constrained to be 4.286 ± 0.05 Å (the distance of two nitrogen silver bonds) with a penalty of 1000 kJ/Å^2 for falling outside these boundaries. Editing was carried out in Maestro®, calculations were performed using Macromodel®; both applications are available from Schrodinger Inc. Methods used were based on those developed by Stephenson *et al.*[1]

Modeling for the assemblies described in Chapter 4 (TGG**T^r**GGT and TGG_**T^r**_GGT (**ON5**)) was based on the PDB file 2o4F. This structure contains a tetramolecular G-quadruplex formed from the self-association of the sequence TGGGGT. Our 1,2,4-triazole quartet models were created by constraining N4 of each triazole ring into a position consistent with their coordination with a Cu^{2+} ion.[2,3] For **ON5**, the 2O4F structure was separated at the centre of the four G-quartets and a triazole quartet inserted in the centre. An amber* minimisation was carried out on the structure with water as solvent, constant dielectric electrostatic treatment with a dielectric constant of 1.0, extended cutoff and charges from force field. Because amber* cannot process transition metal atoms, the distance between N4 of both triazole rings was constrained to be 4.02 ± 0.04 Å (the distance of two nitrogen copper bonds) with a penalty of 1000 kJ/Å^2 for falling outside these boundaries. Editing was carried out in Maestro®, calculations were performed using Macromodel®; both applications are available from Schrodinger Inc. Methods used were based on those developed by Stephenson *et al.*[1]

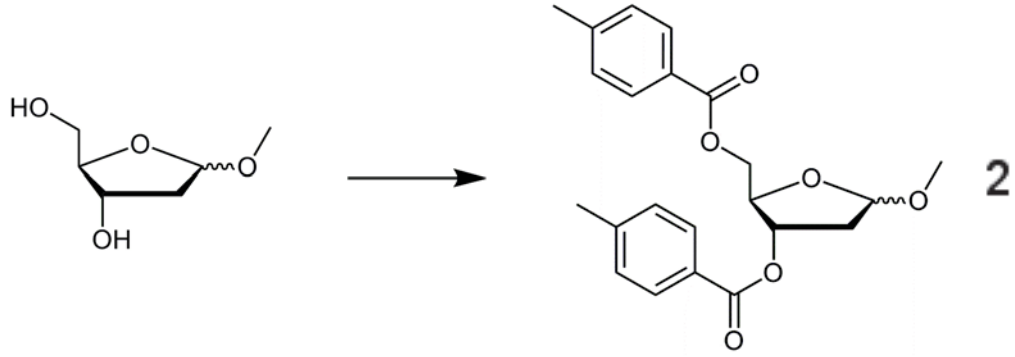
7.8 Synthesis of 1,2,4-Triazole Nucleotide

7.8.1 (*2R,3S*)-2-(Hydroxymethyl)-5-methoxytetrahydrofuran-3-ol (1)



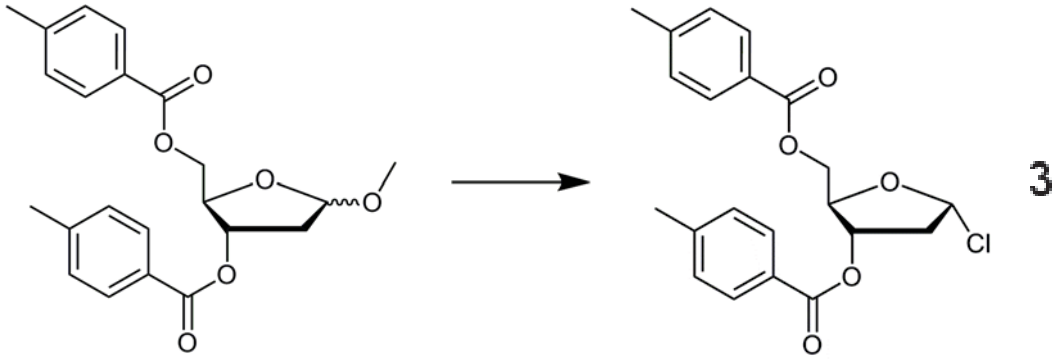
2′-Deoxy-D-ribose (11 g, 81 mmol) was dissolved in dry methanol (200 mL), acetyl chloride (140 µL, to make a 7 % solution) was added and stirred at room temperature. TLC (5 % MeOH in DCM) indicated that the starting material had been consumed after 45 minutes. The reaction was quenched with IRA-400 resin (400 g) and a small amount of pyridine. The resin was filtered off and the solution reduced under vacuum leaving the product **1** as a colourless oil in quantitative yield. Product was used without further purification. Method based on a modified procedure found in the literature.[4-6]

7.8.2 (*2R,3S*)-5-Methoxy-2-[(4-methylbenzoyloxy)methyl] tetrahydrofuran-3-yl 4-methylbenzoate (2)



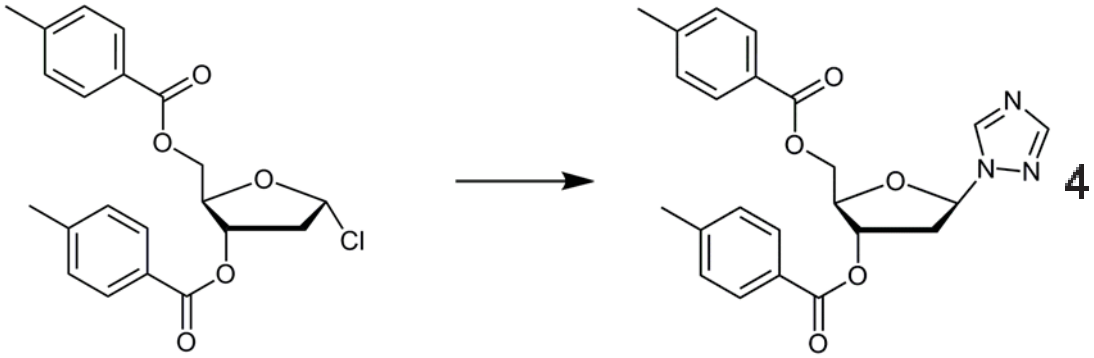
Compound **1** (10.86 g, 73 mmol) was dissolved in dry pyridine (200 mL) and cooled to 0 °C. p-Toluoyl chloride (30 mL, 3 eq.) was added dropwise over 5 minutes and stirred for 10 hours; the reaction mixture was allowed to return to room temperature. The reaction was quenched with a sat. $NaHCO_3$ (100 mL) and stirred for 2 hours. The reaction mixture was washed with water (3×30 mL), aqueous layers were back-extracted with DCM and combined organics were reduced under vacuum. The crude product was purified by silica gel column chromatography (10×25 cm) and eluted with a solution of 5 % methanol in DCM giving product **2** (24 g, 85 %) as a pale brown oil. Method based on a modified procedure found in the literature.[4,7]

7.8.3 (*2R,3S,5R*)-5-Chloro-2-[(4-methylbenzoyloxy)methyl] tetrahydrofuran-3-yl 4-methylbenzoate (3)



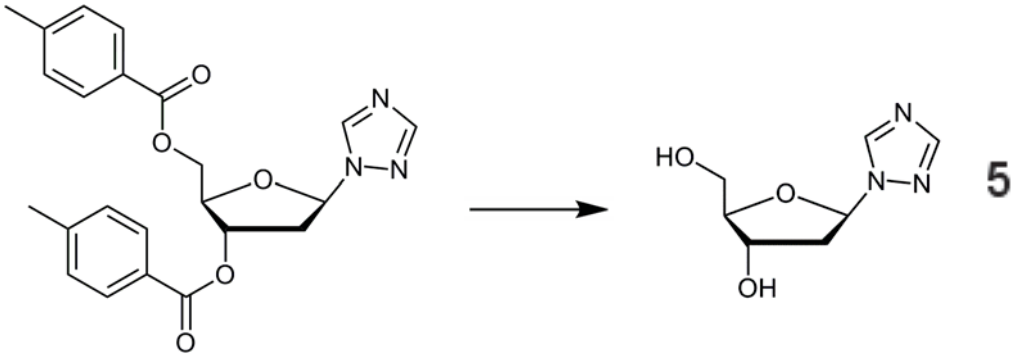
Compound **2** (15.2 g, 40 mmol) was dissolved in a saturated solution of HCl in acetic acid (75 mL) and stirred at 10 °C. After 20 minutes the product **3** had precipitated from solution to such a degree that further stirring became impossible. The reaction was left for a further 40 minutes. Diethyl ether (40 mL) was added and the precipitate was filtered and washed with ether. The product **3** was dried under vacuum, leaving a white powder (11.5 g, 75 %) that was used without further purification within 12 hours; observations reported in the literature suggest that storage for extended periods results in decomposition. [7,8]

7.8.4 [(*2R,3S,5R*)-3-(4-Methylbenzoyloxy)-5-(1*H*-1,2,4-triazol-1-yl)tetrahydrofuran-2-yl]methyl 4-methylbenzoate (4)



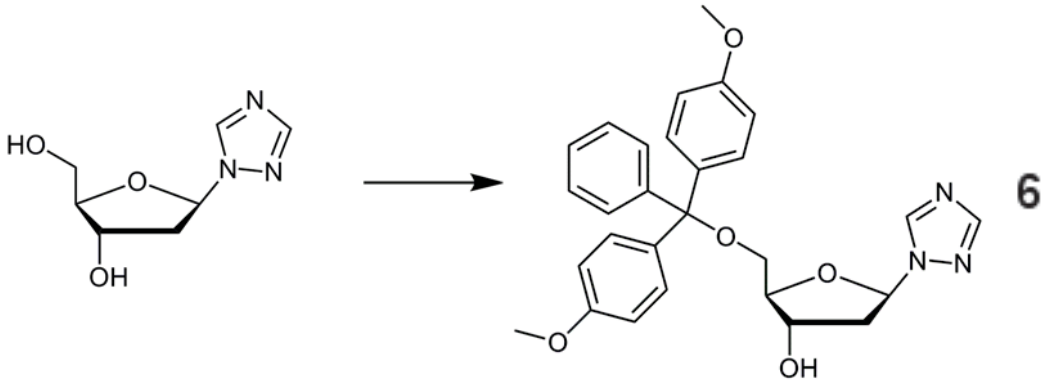
Sodium hydride (0.75 g of 60 % suspension in oil, 18 mmol) was suspended in dry acetonitrile (200 mL), 1,2,4-triazole (1.52 g, 22 mmol) was added and the reaction was stirred for 30 minutes. Compound **3** (6.2 g, 16 mmol) was added in small portions over 30 minutes and then stirred at room temperature for 12 hours. A colour change from a white suspension to a pale brown solution was observed. The reaction mixture was poured into water and extracted with DCM; combined organic layers were reduced *in vacuo* and then purified by silica gel column chromatography (5×20 cm) using 5 % methanol solution in DCM as eluent. Compound **4** (3.0 g, 44 %) was obtained as a brown oil. ^{1}H NMR were consistent with those found in the literature.[9]

7.8.5 (*2R,3S,5R*)-2-(Hydroxymethyl)-5-(1*H*-1,2,4-triazol-1-yl) tetrahydrofuran-3-ol (5)



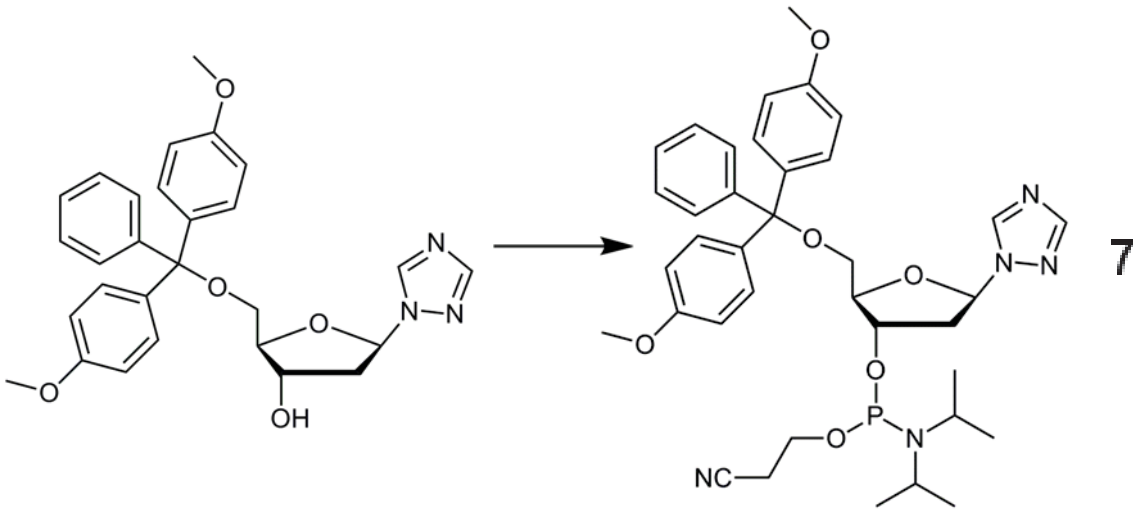
Compound **4** (2.5 g, 6 mmol) was dissolved in methanol (125 mL), 25 % aqueous ammonia (124 mL) was added and the solution stirred at room temperature until TLC indicated that the starting material had been consumed. The reaction mixture was freeze-dried and purified by silica gel column chromatography (1.5×15 cm) using 5 % methanol solution in DCM as eluent. Compound **5** (0.93 g, 84 %) was obtained as a brown oil. ^{1}H NMR were consistent with those found in the literature.[9]

7.8.6 (*2R,3S,5R*)-2-{[bis(4-Methoxyphenyl)(phenyl)methoxy] methyl}-5-(1*H*-1,2,4-triazol-1-yl)tetrahydrofuran-3-ol (6)



Compound **5** (1.73 g, 9.3 mmol) and dimethoxytrityl chloride (4.11 g, 1.3 eq.) were dissolved in dry pyridine (40 mL) and stirred under argon for 5 hours at room temperature. The reaction was quenched with a dilute $NaHCO_3$ solution (20 mL), the product was extracted with DCM and combined organic layers were reduced *in vacuo*. The crude product was purified by silica gel column chromatography (5×20 cm) using DCM with 1 % triethylamine and a gradient of methanol (0-10 %), affording compound **6** (4.0 g, 90 %) as a brown oil. ^{1}H NMR (500 MHz, CDCl3) δ 2.41 (s, 3H, CH3), 3.33 (m, 2H, H5′), 3.71 (m, 2H, H2′), 4.53 (m, 1H, H3′), 4.56 (m, 1H, H4′), 6.31 (t, 1H, H1′), 7.29 (d, 4H, Tol), 7.78 (d, 4H, Tol), 8.04 (s, 1H, Triazole-H), 8.69 (s, 1H, Triazole-H) ppm. NMR spectra were consistent with the literature.[9]

7.8.7 (*2R,3S,5R*)-2-{[bis(4-Methoxyphenyl)(phenyl) methoxy]methyl}-5-(1*H*-1,2,4-triazol-1-yl)tetrahydrofuran-3-yl 2-cyanoethyl diisopropylphosphoramidite (7)



Diisopropylethyl amine (180 μL) was added to compound **6** (0.2 g, 0.4 mmol) under argon followed by DCM (7 mL). Chloro-*N*,*N*-diisopropylamine-β-cyanoethyl phosphine (180 μL)

was added dropwise. TLC indicated the reaction was complete within 30 minutes. The solution was poured into ethyl acetate (30 mL) and washed with water; organic layers were dried with $MgSO_4$ and reduced under vacuum. The crude product was purified by silica gel column chromatography (2×25cm) using DCM:EtOAc:Net_3 (73:25:2) as eluent, affording pure **7** (70 mg, 25 %) as a colourless oil. ^{1}H NMR (500 MHz, $CDCl_3$) δ 2.61 (m, 2H, H2′), 3.22 (m, 2H, H5′), 3.73 (s, 6H, OCH_3), 4.18 (m, 1H, H4′), 4.54 (m, 1H, H3′), 6.22 (t, 1H, H1′), 6.74 (dd, 4H, DMT), 7.11-7.21 (m, 3H, DMT), 7.24 (dd, 4H, DMT), 7.34 (~d, 2H, DMT), 7.83 (s, 1H, Triazole-H), 8.43 (s, 1H, Triazole-H) ppm. ^{31}P NMR (200 MHz, $CDCl_3$) δ 148.7, 148.9 ppm (1:1). NMR spectra were consistent with the literature.[9]

7.9 References

(1) Stephenson, A. W. I.; Partridge, A. C.; Filichev, V. V.: Synthesis of β-Pyrrolic-Modified Porphyrins and Their Incorporation into DNA. *Chem. Eur. J.* **2011**, *17*, 6227-6238.

(2) Yao, J.-C.; Yao, F.-J.; Li, Y.-C.; Wang, J.-G.: Refinement of the Crystal Structure of tetra(imidazole)copper(II) diperchlorate. *Z. Kristallogr. - New Cryst. Struct.* **2007**, *222*, 211-212.

(3) Sletten, E.: Crystallographic studies of metal-nucleotide base complexes. I. Triclinic bis-(6-aminopurine)copper(II) tetrahydrate. *Acta Crystallogr. Sect. B* **1969**, *25*, 1480-1491.

(4) Ogawa, A. K.; Abou-zied, O. K.; Tsui, V.; Jimenez, R.; Case, D. A.; Romesberg, F. E.: A Phototautomerizable Model DNA Base Pair. *J. Am. Chem. Soc.* **2000**, *122*, 9917-9920.

(5) Zinner, H.; Wulf, G.; Heinatz, R.: Derivate der Zucker-mercaptale, XXXV. Darstellung und Mercaptalbildung der 2-Desoxy-D-xylose. *Chem. Ber.* **1964**, *97*, 3536-3540.

(6) Deriaz, R. E.; Overend, W. G.; Stacey, M.; Wiggins, L. F.: Deoxy-sugars. Part VI. The constitution of β-methyl-2-deoxy-L-ribopyranoside and of αβ-methyl-2-deoxy-L-ribofuranoside. *J. Chem. Soc.* **1949**, 2836-2841.

(7) Dhimitruka, I.; SantaLucia Jr, J.: Efficient Preparation of 2-Deoxy-3,5-di-O-p-toluoyl-α-D-ribofuranosyl Chloride. *Synlett* **2004**, *2004*, 0335-0337.

(8) Chin, T.-M.; Huang, L.-K.; Kan, L.-S.: Improved Synthesis of Halofuranose Derivatives with the desired alpha configuration. *J. Chin. Chem. Soc.* **1997**, *44*, 413-413.

(9) Durland, R. H.; Rao, T. S.; Bodepudi, V.; Seth, D. M.; Jayaraman, K.; Revankar, G. R.: Azole substituted oligonucleotides promote antiparallel triplex formation at non-homopurine duplex targets. *Nucleic Acids Res.* **1995**, *23*, 647-53.

7.10 List of Oligonucleotides

Chapter 2:

ON1	$G_3T_2AG_3T\underline{T}AG_3T$
ON2	$T\underline{A}G_3T$
ON3	$G_3T_2AG_3T\underline{\mathbf{T}^r}AG_3T$
ON4	$T\underline{\mathbf{T}^r}G_3T$

Chapter 3:

ON1	$\mathbf{T}^rAG_3T_2AG_3T_TAG_3T_TAG_3$
ON2	$\mathbf{T}^rAG_3T_2AG_3T_TAG_3T_TAG_3\mathbf{T}^r$
ON3	$\mathbf{T}^rAG_3T_2AG_3T$
ON4	$\mathbf{T}^rAG_3T_2AG_3\mathbf{T}^r$

Chapter 4:

ON1	$TG_4\mathbf{T}^rT$
ON2	$T\mathbf{T}^rG_4T$
ON3	$T\mathbf{T}^rG_4\mathbf{T}^rT$
ON4	$TG_2\_\mathbf{T}^r\_G_2T$
ON5	$TG_3\_\mathbf{T}^r\_G_3T$

Chapter 5:

ON1	AGATGA$\underline{C}$TCATCT
ON2	GATGA$\underline{C}$TCATC
ON3	ATGA$\underline{C}$TCAT
ON4	TGA$\underline{C}$TCA
ON5	G_6T_6GATGA$\underline{C}$TCATCT_6G_6
ON6	G_6T_7GATGA$\underline{C}$TCATCT_7G_6
ON7	G_6T_8GATGA$\underline{C}$TCATCT_8G_6
ON8	G_6T_9GATGA$\underline{C}$TCATCT_9G_6
ON9	G_4T_6GATGACTCATCT_6G_4
ON10	G_5T_6GATGACTCATCT_6G_5
ON11	G_6T_6GATGACTCATCT_6G_6
ON12	TG_4T_6GATGACTCATCT_6G_4T
ON13	TG_5T_6GATGACTCATCT_6G_5T
ON14	TG_6T_6GATGACTCATCT_6G_6T
ON15	$G_4T_{23}G_4$
ON16	$G_5T_{23}G_5$