# Investigating the Effect of Small Molecule Ligands and Cations on i-Motif DNA

**Henry Albert Day**

A thesis submitted for the degree of Doctor of Philosophy

University of East Anglia

School of Pharmacy

January 2015

## Declaration

This thesis is submitted to the University of East Anglia for the Degree of Doctor of Philosophy and has not been previously submitted at this or any university assessment or for any other degree. Except where stated, and reference and acknowledgment is given, this work is original and has been carried out by the author alone.

Henry Albert Day

# Acknowledgements

Carrying out my PhD has been one of the biggest challenges I have faced and there are of course many people whom I need to thank for helping me throughout this process. Firstly I would like to thank my supervisors Dr Zoë Waller and Professor Mark Searcey for allowing me to work in their laboratories and for their continued encouragement and guidance. I must also thank Novartis and the University for funding my work.

There are parts of this work which have been carried out in collaboration so I would like to thank Camille Huguin (for the FRET ligand screen and silver FRET titration) and Colin MacDonald (for NMR on the i-motif). I would also like to thank Professor Rob Field for donating the compound library and Dr Andy Gates for donating the cations to be screened in chapter 4. Throughout this work I have been fortunate enough to learn a variety of different techniques and work in a number of different groups so I would like to thank Maria O'Connell, Myles Cheesman, Andrew Mayes, Clare Stephenson and Richard Bowater for allowing me access to your laboratories, to use your equipment and in some cases for training me to use them. The technician team, particularly Tim Lane, James Goillau and Rebecca Baldwin have also been of great help to keep the lab running smoothly and safely when equipment has gone wrong.

Teaching and demonstrating has been an unexpected source of enjoyment throughout my PhD and I have been fortunate enough to be able to contribute to the chemical education of a large number of students both in the teaching labs and in the research labs. Your questioning has kept me on my toes and helped me to improve my own communication skills and my own understanding of the subject. Never stop asking questions.

I would like to thank Lesley and Onur for their general support and Professor Ganesan for sharing his laboratory with us. I have been fortunate enough to make many new friends since moving to UEA, particularly Angus, Niall, Hannae, Teresa, Louis and Qiran for making the lab a friendly place to work. To the Searcey, Hamilton and Mathews' groups thank you for your generousity with advice, chemicals and equipment.

Thank you to my friends who have stood by me throughout this process, I am sorry for the times I have neglected you and I thank you for your forgivness and support. To Jess, you have made this last and most stressful year also one of the most enjoyable and I wish you the very best with your own PhD. To John you have continually helped build my self confidence at the times when it has been at its lowest. Finally to my parents, thank you for your continued love and support. Your positive attitudes despite everything life throws at you, are and always will be an inspiration.

# Abstract

The i-motif is an alternative DNA secondary structure motif formed in sequences rich in cytosine, consisting of four strands, stabilised by hemi-protonated cytosine-cytosine$^+$ base pairs. The motif forms in sequences complimentary to the G-quadruplex however, far less is known about the i-motif and most research to date has focused on its application in nanotechnology. Despite this, recent progress in the field has indicated that the i-motif may be a possible therapeutic target in certain cases. In order to study this structure in more detail a chemical tool box of ligands and conditions is needed, which can be used to probe its potential biological function.

Herein the effect of small molecule ligands and cations has been investigated. A previously identified i-motif binding compound **BisA** has been characterised in detail with a range of biophysical experiments, showing that it does bind to the i-motif but causes the DNA to condense. A high throughput screen has been carried out finding a number of potential new i-motif binding ligands and, through a range of experiments, two lead compounds **mitoxantrone** and **tilorone** have been identified with micromolar affinities from which novel i-motif binding analogues and structure activity relationships can be developed. Finally, the effect of cations on the i-motif has been studied, including a wider selection from across the periodic table than has previously been investigated. This has shown that at neutral pH, silver (I) ions have the ability to induce i-motif formation and that this is reversible in the presence of cysteine. While at acidic pH, copper (II) ions have the ability to induce hairpin formation reversibly in the presence of EDTA. This could enable the formation of multiple structures from the same oligonucleotide sequence in response to different conditions. The combination of these results should provide useful tools to further study the i-motif structure.

# Table of Contents

# List of Figures

## Chapter 1: Introduction

**Chapter 2: Characterisation of the i-Motif Binding Compound BisA**

**Chapter 5: Conclusions and Future Work**

# List of Tables

# List of Schemes

## Chapter 2: Characterisation of the i-Motif Binding Compound BisA

# List of Abbreviations

| | |
|---|---|
| Å | Ångstrom |
| °C | Degrees celsius |
| APCI-MS | Atmospheric pressure chemical ionisation mass spectrometry |
| aq | Aqueous |
| Ar | Aromatic |
| Bcl-2 | B-cell lymphoma-2 gene |
| BHQ | Black hole quencher |
| **BisA** | **Bisacridine** |
| br | Broad |
| CD | Circular dichroism |
| c-jun | A proto-oncogene promoter sequence containing an i-motif |
| c-ki-ras | An oncogene promoter sequence containing an i-motif |
| c-kit | An oncogene promoter sequence containing an i-motif |
| c-myb | An oncogene promoter sequence containing an i-motif |
| c-Myc | An oncogene promoter sequence containing an i-motif |
| **CV** | **Crystal violet** |
| d | Doublet |
| DCM | Dichloromethane |
| **DETA** | **Diethylenetriamine** |
| DLS | Dynamic light scattering |
| DMF | Dimethylformamide |
| DNA | Deoxyribose nucleic acid |
| EDTA | Ethylenediamine tetraacetic acid |
| Eq | Equivalents |
| FAM | 6-Carboxyfluorescein |
| FRET | Förster resonance energy transfer |
| GQD | Graphen quantum dot |
| HEG | Hexaethylene glycol |
| HeLa | Henrietta Lacks cancer cell line |
| HIF-1α | Hypoxia-inducible factor 1 alpha gene |
| hnRNP | Heterogenous ribonuclear protein |
| HRMS | High resolution mass spectrometry |

| | |
|---|---|
| hTelo | Human telomeric region |
| hTERT | Human telomerase reverse transcriptase gene |
| Hz | Hertz |
| ICD | Induced circular dichroism |
| ILPR | Insulin linked polymorphic region |
| m | Multiplet |
| **MonoA** | **Monoacridine** |
| **MWNT** | **Multi-walled carbon nanotube** |
| NMR | Nuclear Magnetic Resonance |
| n-Myc | An oncogene promoter sequence containing an i-motif |
| PAGE | Polyacrylamide gel electrophoresis |
| PDB | Protein data bank |
| PDGF | Platelet-derived growth factor gene |
| pH | Power of hydrogen |
| q | Quartet |
| Rb | Retinoblastoma |
| RET | Rearranged during transfection gene |
| RNA | Ribose Nucleic Acid |
| s | Singlet |
| SPR | Surface plasmon resonance |
| **SWNT** | **Single-walled carbon nanotube** |
| t | Triplet |
| TAMRA | 6-Carboxytetramethyl rhodamine |
| TBE | Tris/Borate/EDTA |
| TDS | Thermal Difference Spectrum |
| Th | Tyrosine hydroxylase gene |
| THF | Tetrahydrofuran |
| TLC | Thin layer chromatography |
| **TMPyP4** | **5,10,15,20-tetra-(*N*-methyl-4-pyridyl)porphine** |
| UV | Ultraviolet |
| VEGF | Vascular endothelial growth factor gene |

# List of work within this thesis that has been published or is in preparation for publication

1. Henry A. Day, Camille Huguin, Zoë A. E. Waller; Silver Cations Fold i-Motif at Neutral pH, *Chem. Commun.*, 2013, **49**, 7696-7698.

2. Henry A. Day, Pavlos Pavlou, Zoë A. E. Waller; i-Motif DNA: Structure, Stability and Targeting with Ligands, *Bioorganic and Medicinal Chemistry*, 2014, **22**, 4407-4418.

   Special edition symposium in print: Unlocking Nature through Chemistry, edited by Professor Shankar Balasubramanian.

# Chapter 1: Introduction

## 1.1 DNA Secondary Structure

The iconic double helix structure of DNA was first suggested by James Watson and Francis Crick in 1953 (figure 1.1a).[1] Since then, research efforts have focused on whether DNA can form other structures and what role they may play in biology. This lead to the identification of now well-known structures including B-, A- and Z-form duplexes, hairpins, triplexes, 3-way and 4-way junctions and quadruplexes, illustrating that DNA may exist in many different structural forms.[2] The G-quadruplex structure in particular (figure 1.1b), has seen some of the most extensive investigation.

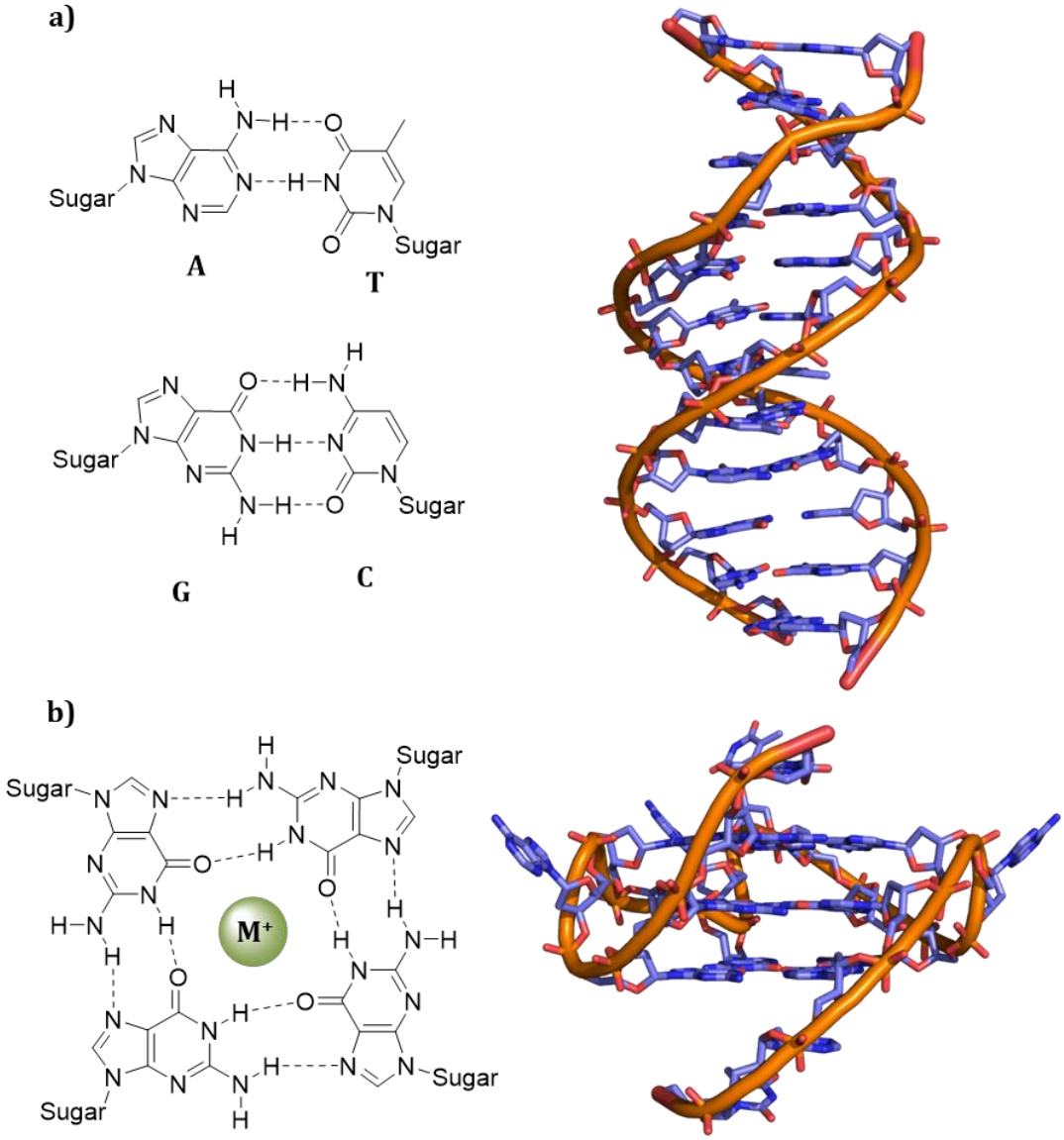



*Figure 1.1: a) Watson and Crick base pairing and the structure of B-form DNA. PDB ID: 1BNA,[3] b) Hoogsteen base pairing in a G-tetrad and the structure of a G-quadruplex PDB ID: 2LBY.[4]*

These are four stranded structures formed from sequences with a repeating series of guanine bases that form a G-tetrad using Hoogsteen hydrogen bonding. These tetrads are able to stack on top of one another forming the G-quadruplex stabilised by monovalent cations.[5] Sequences capable of forming these structures have been found throughout telomeric and promoter regions of the genome, while Balasubramanian and co-workers have been able to identify their existence in living human cells.[6–9] Some G-quadruplex binding ligands have been shown to be effective inhibitors of telomerase and have been used to both up regulate and down regulate oncogene expression.[10,11] Proteins have also been discovered which bind to G-quadruplex structures indicating a natural biological role for the structure.[12–14]

## 1.2 i-Motif DNA

First proposed in 1993 by Gehring, Leroy and Guéron, the i-motif is another four stranded structure formed from sequences rich in cytosine, which are complimentary to G-quadruplex forming sequences.[15] The structure forms readily under acidic conditions and is stabilised by hemi-protonated cytosine-cytosine$^+$ base pairs, forming

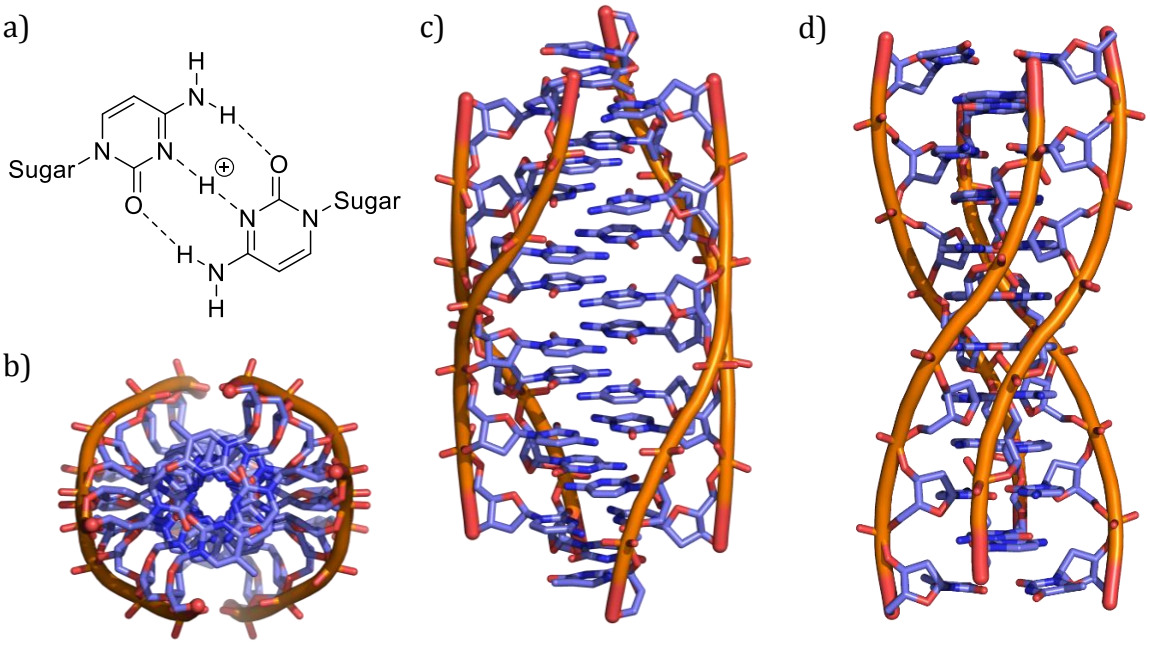



*Figure 1.2: The structure of tetrameric i-motif DNA. a) A hemi-protonated cytosine-cytosine$^+$ base pair, b) the top view of d(TC$_5$) c) side view of d(TC$_5$) showing the major groove, d) side view of d(TC$_5$) showing the minor groove. PDB ID: 225D.[15] Figure adapted from Henry Day et. al. (see appendix).[16]*

two parallel duplexes that intercalate with one another in an antiparallel orientation (figure 1.2).[15] Due to a reliance on acidic pH for the stability of the structure, there is still some skepticism in the field whether the i-motif can form under physiological conditions. Therefore, the majority of i-motif research so far has focused on its potential uses in nanotechnology. The ability of the i-motif to rapidly fold and unfold in response to changes in pH has enabled it to be used in pH sensing,[17,18] logic operations for computation,[19–21] nanomachines[22,23] and nano-scale assembly.[24,25] However, much less is known about the natural occurrence of these structures and whether they have a role to play in biology.

### 1.2.1 Structure and Topology of i-Motif DNA

The first i-motifs identified were intermolecular structures consisting of four separate strands of cytosine.[15] Hemi-protonated cytosine base pairs were identified as early as 1962[26] and evidence of hairpin secondary structures in cytosine rich sequences was observed in 1989.[27] Despite this, the identification of an intercalated tetrameric structure was surprising and the intercalated (i) motif is still one of the only structures of this type to have been identified. Sequences with two stretches of cytosines can also form dimeric i-motif structures where two hairpins intercalate.[28] More interestingly, from a biological perspective, natural sequences with four stretches of cytosines separated by other bases can fold into an intramolecular i-motif structure (figure 1.3). Since the original i-motif discovery, a number of crystal structures of intermolecular i-motifs have been reported,[29–33] however, there as yet no published examples of intramolecular i-motif crystal structures.

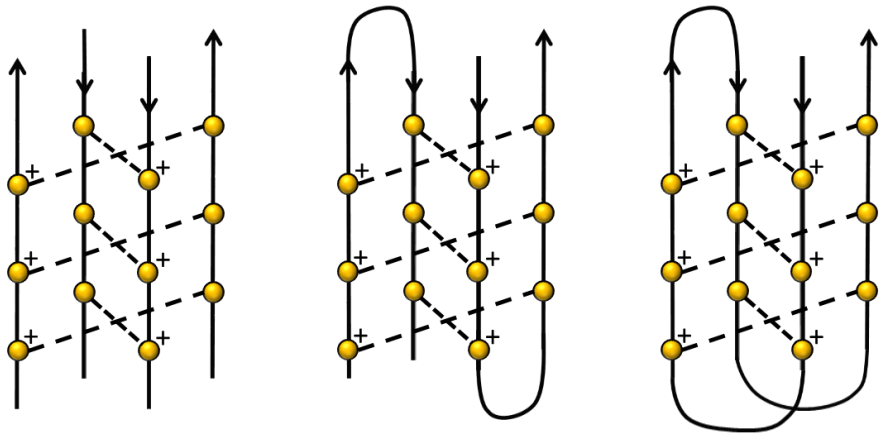


*Figure 1.3: Tetrameric, dimeric and intramolecular i-motif structures.*

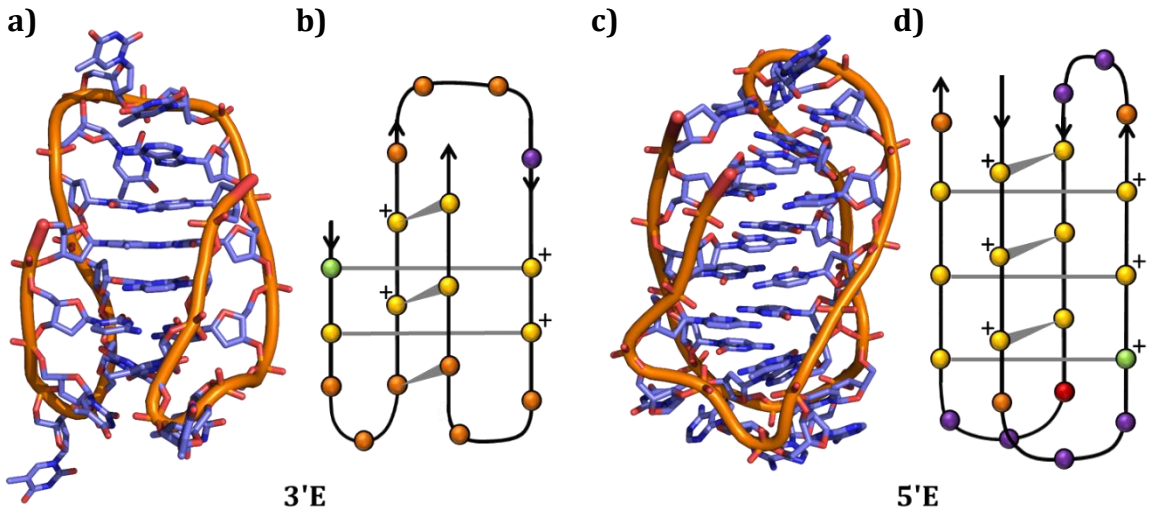



*Figure 1.4: A comparison of 3'E & 5'E i-motif structures. a) an NMR structure of d(5mCCT$_3$CCT$_3$ACCT$_3$CC) PDB ID: 1A83,[34] b) a schematic diagram of this 3'E i-motif, c) an NMR structure of the human telomeric i-motif d(CCCTA$_2$5mCCCTA$_2$CCCUA$_2$CCCT) PDB ID: 1EL2,[35] d)a schematic diagram of this 5'E i-motif. Figure adapted from Henry Day et. al. (see appendix).*[16]

The intercalation of the i-motif structure can occur in different ways to form two different topologies. One in which the outermost cytosine pair are from the 3' end of the sequence and another where the outermost pair are from the 5' end. First labelled R-form and S-form,[36] they were later re-named 3'E and 5'E respectively.[37,38] A third topology denoted T-form is also possible in which the outermost C-C$^+$ pair are not intercalated but this is less stable.[39] As will be discussed, there are several competing factors affecting i-motif stability so whichever of these topologies is favoured for any given i-motif will be a compromise of these multiple effects, depending on sequence (figure 1.4).

As well as the 3'E and 5'E classifications, intramolecular i-motifs have also been classified by Laurence Hurley and co-workers based upon the length of the loop regions.[40] i-Motifs with short loops have been described as class I while those with long loops have been denoted class II (figure 1.5). Of the i-motifs they classified as class I or class II, it was shown that the class II i-motifs were more stable, which they hypothesised was due to interactions between bases in the loop regions providing a stabilising effect. However, John Brazier's group have since identified a class I i-motif in the HIF-1α gene promoter which has unexpectedly high stability, with a transitional

pH (the pH at which half the DNA is folded and half unfolded) close to neutral ($pH_T$ = 7.2).[41]

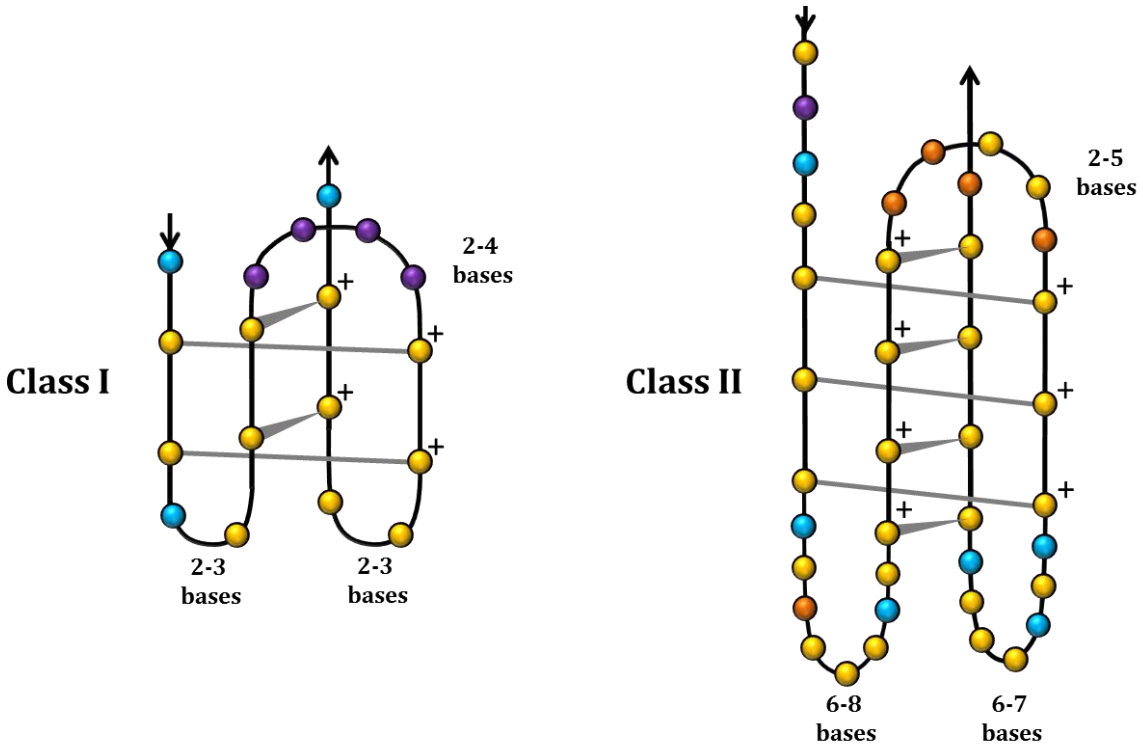



*Figure 1.5: A schematic diagram illustrating the difference between short looped Class I and longer looped Class II i-motif structures. In this example, the class I i-motif is from the promoter region of the Retinoblastoma oncogene (Rb) and the Class II i-motif is from the P1 promoter region of B-cell lymphoma 2 (Bcl-2).*[40]

### 1.2.2 Stability of i-Motif DNA

Most investigations into i-motif stability are driven by the quest to find a natural sequence that is able to form an i-motif structure under physiological conditions of pH and temperature. It has been shown that the i-motif can form at neutral pH at low temperature ($T$ = 4 °C) but this in itself isn't particularly physiologically relevant.[42] Molecular crowding is another important consideration. The cellular and nuclear environment is more akin to a gel than simple aqueous conditions and simulation of this crowded environment can be achieved using high molecular weight polyethylene glycol polymers.[43] Under these conditions it has been shown that G-quadruplex and i-motif structures are both favoured and duplex and single stranded DNA are actively disfavoured.[44,45] This is due to an increase in p$K_a$ of the N3 position of cytosine which

needs to be protonated. This makes protonation more favourable at neutral pH and hence the corresponding i-motif is easier to form.[46,47]

Another key factor affecting the favourability of formation of both i-motif and G-quadruplex structures, is their respective destabilising effects on the remaining double helical strand. It has been suggested that a minimum spacing of five base pairs of unwound DNA is needed between the i-motif structure and the start of the double helix in order to maintain the stability of each structure.[48] What may facilitate this is the negative superhelical stress that is induced in the DNA single strand when it is unwound for transcription. Formation of an alternative secondary structure such as the i-motif, can relieve this superhelical stress and hence stabilise the DNA structure as a whole.[49] Laurence Hurley's group were able to show this effect by placing the i-motif and G-quadruplex forming sequences from the c-Myc oncogene promoter into a supercoiled plasmid which simulated the negative superhelicity generated during transcription.[49] Using enzymatic and chemical footprinting methods they showed the formation of both G-quadruplex and i-motif structures at neutral pH. This indicates a plausible mechanism for the formation of i-motif structures under physiological conditions during transcription.

Under physiological conditions there are a range of different cations present at varying concentrations. Therefore, the effect of cation concentration on i-motif stability is an important factor to assess. The effects of $Na^+$, $K^+$ and $Mg^{2+}$ on the formation of the i-motif in the c-jun proto-oncogene have been studied.[50] Saxena and co-workers showed that initially in solutions of 100 mM $Na^+$ at pH 5.7, only duplex DNA was present. However, in solutions of 100 mM $K^+$ at the same pH, i-motif and G-quadruplex structures were favoured. When 10 mM $Mg^{2+}$ and 100 mM $Na^+$ were assessed in combination, all 3 structures - duplex, G-quadruplex and i-motif were found. Later, Mergny and co-workers showed that at pH 4.8, (close to the $pK_a$ of cytosine) increasing the concentration of sodium chloride to 100 mM destabilised i-motif structure, but that a further increase in concentration to 300 mM did not destabilise the structure any further.[51] This effect they showed, was due to the lowering of the $pK_a$ of the N3 of cytosine in low salt conditions, making i-motif formation more favourable at that pH in low salt rather than high salt concentration. They also showed, that in the presence of 100 mM NaCl at pH 6.4, there was no effect on i-motif stability with the addition of 5

mM $Mg^{2+}$, $Ca^{2+}$, $Zn^{2+}$, $Li^{+}$ or $K^{+}$ ions. Other recent studies have continued to show a decrease in the stability of the i-motif in the presence of a high concentration of cations. In particular, Sung Eun Kim and co-workers have shown that $Li^{+}$ ions promote unfolding of the i-motif without hindering the folding of the i-motif.[52] They suggest this is due to the small ionic radius of $Li^{+}$ (0.7 Å) which is small enough to fit between the cytosine-cytosine$^{+}$ base pairs (1.8 Å,[35]) but does not have the ability to facilitate cytosine base pairing itself and hence disrupt the structure. Aside from these few experiments, little has been done to look at the effect of cations on i-motif DNA. Therefore, this topic has been studied as part of this thesis and will be discussed in detail in chapter 4.

## 1.3 Biological Significance

Recently, there has been increasing interest in the potential biological significance of the i-motif. Due to the complimentarity of i-motif forming sequences with G-quadruplex forming sequences, initial research has so far looked at the telomeric and oncogene promoter regions where G-quadruplexes have already been well characterised. The results of some of these studies are described in this section and show that the i-motif may have potential as a therapeutic target.

### 1.3.1 Telomeric Regions of Eukaryotic DNA

The telomeric regions of eukaryotic DNA are the repeating sequences of linear double and single stranded DNA at the end of chromosomes and they are rich in guanine bases with the repeat sequence 5'-d$(TTAGGG)_n$-3'. They include a single stranded overhang of the same sequence at the 3' end which is known to form G-quadruplexes.[53] Normally the telomeric DNA sequence is in a complex with several proteins and forms a cap to the chromosome which protects the DNA from degradation during cell division. This is known as the telomere. In fact, the telomeres act like a "replication clock" which defines how many times the cell can divide.[54] During each replication, a portion of the telomere is lost, until eventually there is no telomere left and replication would damage the coding DNA region at which point senescence is triggered. The enzyme telomerase is able to regenerate the telomeric repeat region. In normal healthy cells, this enzyme is switched off, conversely, in over 85% of cancers, telomerase is over expressed.[54,55] This

is one mechanism by which cancer cells can become immortal, enabling them to divide uncontrollably.

It has already been demonstrated that stabilisation of G-quadruplex structures in the human telomere can inhibit telomerase.[56] In fact, there are now a large number of G-quadruplex binding ligands which are also potent telomerase inhibitors.[57] The complimentary strand of the human telomeric DNA sequence is rich in cytosine with the ability to form an intramolecular i-motif structure and was one of the first intramolecular i-motifs to be discovered.[58] It has the sequence 5'-d(CCCTAA)$_n$-3' and forms a class I type of structure with short loop sequences consisting of the 3 TAA bases that sit between the repeating cytosine tracts, this is defined as a 3:3:3 loop topology.[35] Loop topology is the number of bases forming the loops between cytosine tracts for example in figure 1.5, Rb has a loop topology of 2:4:2 while Bcl-2 has a loop topology of 8:5:7 VEGF and RET both have 2:3:2 loop topologies but different sequences.[40]

Targeting of the human telomeric i-motif has also been shown to inhibit telomerase (figure 1.6).[59] Using carboxyl-modified single walled carbon nanotubes which bind specifically to i-motif structures, Xiaogang Qu and co-workers showed they were able to inhibit telomerase and induce telomere uncapping, resulting in a DNA damage response and apoptosis (see section 1.3.2).[59] There are several proteins which have been shown to bind to the C-rich sequence of telomeric DNA however, it is not certain whether these proteins are able to bind to the i-motif structure or whether they only bind to single stranded DNA.[60,61]

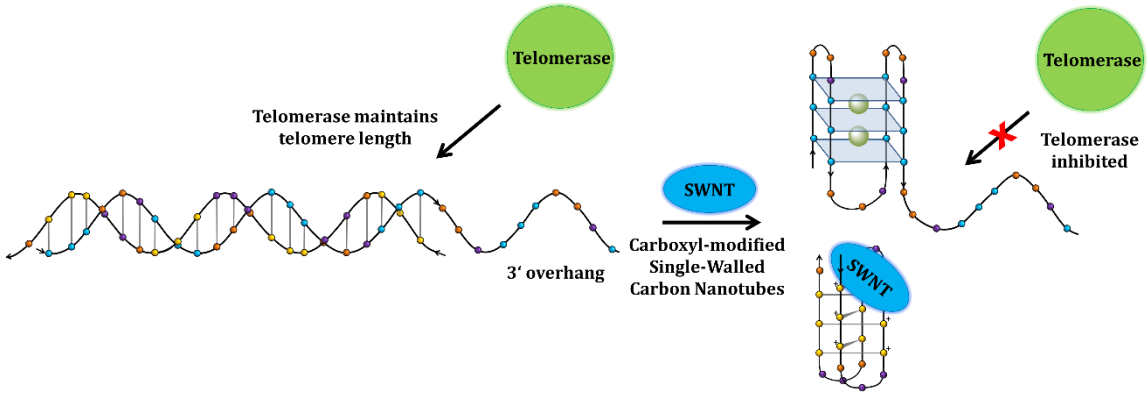



*Figure 1.6: A schematic illustration of the binding of carboxyl-modified single-walled carbon nanotubes (**SWNTs**) to the human telomeric i-motif and the resultant inhibition of telomerase which may be as a result of corresponding G-quadruplex formation.*

### 1.3.2 Gene Promoter Regions

One of the most well studied i-motifs is that of the sequence within the promoter region of the c-Myc oncogene.[62] This gene codes for a transcription factor with a variety of different functions in the control of cell cycle progression and cell growth.[63] Its dysregulation is one of several characteristic hallmarks of cancer. The c-Myc promoter region consists of seven nuclease hypersensitive elements or NHEs. The NHE $III_1$ is of particular interest as this is responsible for approximately 90% of c-Myc transcription.[64] The NHE $III_1$ is rich in cytosine bases and is therefore able to form an i-motif structure, it is also in this C-rich region that the transcription factor binding sites for Sp1 and heterogeneous ribonucleoprotein K (hnRNP K) are found.[63] The c-Myc promoter i-motif has a relatively high transitional pH of 6.6.[40] *In vitro* the sequence forms an i-motif with a 5:5:5 loop topology, neither class I or class II, utilising eight cytosine base pairs.[62] However, Laurence Hurley's group have shown that under conditions of negative superhelical stress, such as those generated during transcription, the c-Myc i-motif can form at neutral pH with a class II 6:2:6 loop topology and only using 6 cytosine base pairs.[49] The reason for this is, as yet, unknown but the presence of longer loops may provide additional stability or may be part of any possible biological role of the i-motif.

The P1 promoter of the B-cell lymphoma 2 (Bcl-2) oncogene has been shown to form a stable i-motif ($pH_T$ = 6.6) with a large 8:5:7 loop topology (class II).[40,65] Bcl-2 is over expressed in many cancers and works to protect the cell against apoptosis.[66,67] Interestingly, for the same reason Bcl-2 is under expressed in neurodegenerative diseases such as Alzheimer's and Parkinson's resulting in an increase in cell death.[68] Laurence Hurley and co-workers have studied the Bcl-2 i-motif in detail and recently showed that the i-motif structure is in fact in dynamic equilibrium with a hairpin structure that is more flexible than the i-motif (figure 1.7).[69] Hurley's group have also identified ligands which bind to the Bcl-2 i-motif and flexible hairpin structures using a high throughput screen. They went on to show that stabilisation of the i-motif with the ligand **IMC-48** (see section 1.3.3) resulted in an increase in gene transcription, conversely the hairpin ligand **IMC-76** (see section 1.3.3) resulted in a decrease in gene transcription. The effects on gene expression were seen at both the transcription level and the translation level and is the first evidence for a possible biological function for

the i-motif.[69] This result is particularly interesting as it was previously hypothesised that formation of the i-motif is likely to down regulate gene expression in a similar way to G-quadruplexes, by preventing the binding of transcription factors to the single strand. In fact, Laurence Hurley's group also identified a natural transcription factor (hnRNP LL) which binds to the two lower loops of the i-motif and then unfolds the structure to activate transcription (see section 1.2.3).[70] The hnRNP LL protein is analogous to the hnRNP K protein which binds to the c-Myc promoter and it has now been hypothesised that unlike the G-quadruplex, the i-motif may be an "on switch" for gene transcription, acting as a scaffold to enable more favourable binding of transcription factors to the c-rich sequence.

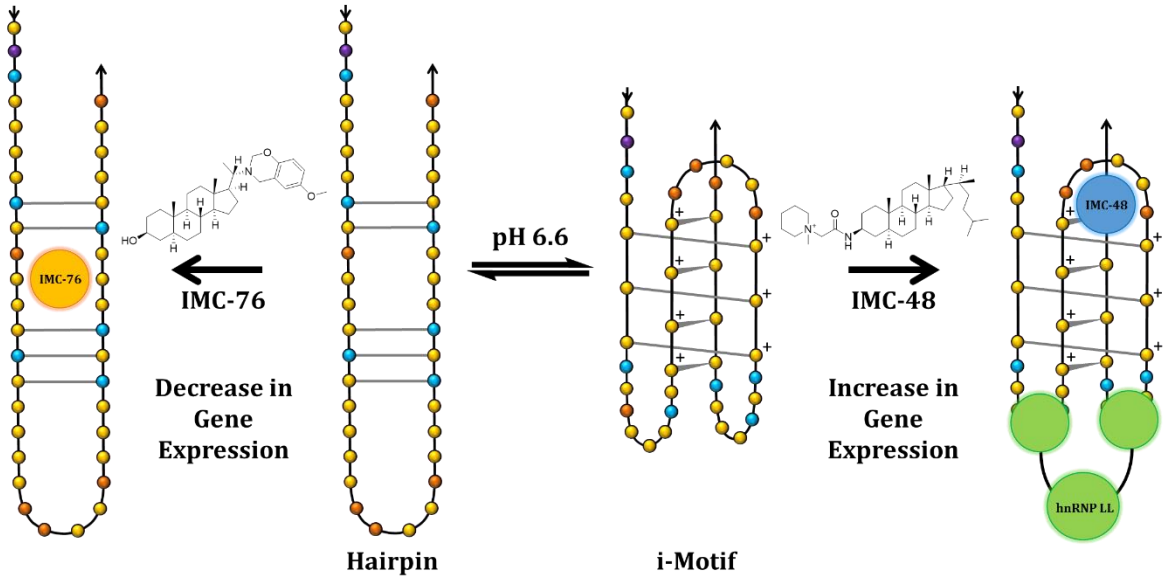



*Figure 1.7: A schematic illustration to show the dynamic equilibrium between the i-motif structure and a flexible hairpin structure in the Bcl-2 promoter region and the effect of compounds **IMC-48** and **IMC-76** on gene expression.*

Largely thanks to the work of Laurence Hurley and co-workers, c-Myc and Bcl-2 are the two most well studied promoter i-motif sequences. Despite this, there is still a lot of work needed to cement the i-motif as a possible therapeutic target. However, an increasing number of i-motif forming gene promoter sequences have been identified and more are currently under investigation. These are most often in promoter regions where the complimentary G-quadruplex structures have already been characterised.

A range of different i-motif forming sequences from promoter regions have been listed in Table 1.1. Many such as retinoblastoma (Rb),[71,72] vascular endothelial growth factor (VEGF),[73] c-ki-ras,[74] c-kit,[75,76] n-Myc[77] and RET[78] are part of promoters for oncogenes which are over expressed in a variety of different cancers. The proximity of the potential i-motif structures to important transcription start sites offers circumstantial evidence for an involvement in gene expression and, even if they do not form naturally, could be investigated as potential targets for ligands. Like G-quadruplexes however, the i-motif should not be considered as a solely cancer related target. The identification of an i-motif forming sequence in the insulin-linked polymorphic region (ILPR),[28,79,80] and a possible i-motif structure in the fragile X chromosome repeat sequence,[81] show that the i-motif could be investigated in the context of other diseases such as diabetes or fragile X syndrome.

Due to the complimentary nature of G-quadruplex and i-motif sequences, a logical starting point to look for new i-motifs is in those promoter regions where G-quadruplexes have already been characterised. John Brazier and co-workers have started to do just this. They examined the cytosine rich strands of c-KIT, c-myb, hTERT, PDGF-A and HIF-1α, all of which already contain known G-quadruplexes in their G-rich strands.[41] They showed that all of these sequences were able to form i-motifs at pH 5.0. More interestingly, they showed that PDGF-A and HIF-1α were able to form unusually stable i-motifs even at pH 7.0 with $T_m$'s of 32 and 27°C respectively. This is surprising as both these sequences form what would be described as class I i-motifs with short loop sequences under Hurley's earlier classification system.[40] This goes against previous evidence which suggests that class II i-motifs with longer loops are more stable.[40] However, these particular sequences do have much longer cytosine stretches able to form up to 10 (PDGF-A) and 7 (HIF-1α) cytosine base pairs which may be responsible for this extra stability. It is also not yet clear what the precise folding topology of these i-motifs are and depending on which cytosines are used for the core i-motif structure, they may in fact have longer loops than expected, with the loops containing cytosines or they may have long flanking regions at the 3' and 5' ends which could also have a stabilising effect.

It has been shown that there are over 376,000 putative G-quadruplex forming sequences in the human genome,[82,83] with 43% of all gene and 69% of oncogene promoters containing a possible G-quadruplex.[84] It is conceivable therefore that there

could be at least as many putative i-motif forming sequences. However, the stability of the i-motif is much more variable with length and composition of the sequence and no precise folding rules have yet been formulated. The example of Bcl-2 and of other sequences that are able to form multiple hairpin or i-motif structures adds another layer of complexity to the possible biological role of the i-motif. While G-quadruplexes are generally favoured within shorter sequences forming compact structures, the stability of longer looped i-motifs and their possible dynamic nature indicates that i-motifs may require longer sequences to form and it may be the case that there are i-motif forming sequences in regions where G-quadruplexes are less stable or their formation is not possible at all.

| Name | Sequence 5'→3' |
|---|---|
| hTeloC | d(CCCTAACCCTAACCCTAACCCT)$_n$ |
| c-Myc | d(CCCCACCTTCCCCACCCTCCCCACCCTCCCC) |
| Bcl-2 | d(CAGCCCCGCTCCCGCCCCCTTCCTCCCGCGCCCGCCCCT) |
| Rb | d(GCCGCCCAAAACCCCCCG) |
| RET | d(CCGCCCCCGCCCCGCCCCGCCCCTA) |
| VEGF-A | d(GACCCCGCCCCCGGCCCGCCCCGG) |
| ILPR | d(TGTCCCCACACCCCTGTCCCCACACCCCTGT)$_n$ |
| c-ki-ras | d(GCTCCCTCCCTCCCTCCTTCCCTCCCTCCC) |
| c-kit | d(CCCTCCTCCCAGCGCCCACCCT) |
| n-Myc | d(ACCCCCTGCATCTGCATGCCCCCTCCCACCCCCT) |
| c-myb | d(TCCTCCTCCTCCTTCTCCTCCTCCTCCGTGTCCTCCTCCTCC) |
| hTERT | d(CCCCGCCCCGTCCCGACCCCTCCCGGGTCCCCGGCCCAGCCCCCACCGGGCCCTCCCAGCCCCTCCCC) |
| PDGF-A | d(CCGCGCCCCTCCCCCGCCCCCGCCCCCGCCCCCCCCCCCCC) |
| HIF-1α | d(CGCGCTCCCGCCCCCTCTCCCCTCCCCGCGCGCCCGAGCGCGCCTCCGCCCTTGCCCGCCCCCTG) |
| c-jun | d(TAACCCCCTCCCCCTCCCCCCTTTAAT) |
| Th (rat) | d(TCTCGTCGCCCTCGCTCCATGCCCACCCCCGCCTCCCTCAGGCAC) |

*Table 1.1: Example sequences of different i-motif structures forming within the human genome, which have been investigated in the literature. hTeloC and ILPR are both tandem repeat sequences.*

### 1.3.3 i-Motif Binding Proteins

There are many proteins which are known to bind to C-rich DNA but it is often assumed that this is either to the double strand or single strand. A protein has been identified in HeLa cell extracts, which is specific for telomeric repeat sequences and sequences with

at least 4 cytosine tracts over other cytosine rich sequences in promoter regions and d($C_{22}$).[60] This is circumstantial evidence that the i-motif may affect its binding, either to promote or inhibit, however the binding of the protein has only been tested at pH 7.4 where the i-motif is not expected to form. Guéron and co-workers have investigated the binding of yeast proteins to the yeast tetrahymena sequence and vertebrate telomeric sequences showing binding from pH 6 to 8 using an electrophoretic mobility shift assay (EMSA),[85] unfortunately binding across this pH range means it is not clear whether the proteins bind to the i-motif, the single strand or both.

The first protein to be identified and characterised as binding to the i-motif structure itself, is the transcription factor hnRNP LL.[70] In their investigation of the Bcl-2 i-motif, Hurley and co-workers used a pull down assay to identify proteins which bind to the Bcl-2 i-motif. They identified 35 proteins, including 9 with a role in transcription. hnRNP LL was of particular interest as it is a paralog of the RNA binding protein hnRNP L, having four DNA binding domains which recognise the sequences CCCGC and CGCCC which are found in loops 1 and 3 of the Bcl-2 i-motif. The protein was shown to be specific for i-motif over a mutant single strand by EMSA at pH 6.8. Surface plasmon resonance was then used to measure the binding affinity of the protein for both the Bcl-2 i-motif and the Bcl-2 single strand, showing that hnRNP LL had a much stronger affinity ($K_d$ = 19.4 pM) for the i-motif than for the single strand ($K_d$ = 69.8 pM). The spacing of the CCCGC sequences that are recognised by the protein provides further evidence that the i-motif is the natural target of hnRNP LL. Hurley's group used EMSA to assess the binding of hnRNP LL to sequences with between 2 and 17 thymine bases between the two recognition sequences. These experiments showed that a spacing of 13 bases resulted in the strongest binding, this is the same number of bases separating the recognition sequences in the natural Bcl-2 sequence. Using circular dichroism and bromine footprinting, it has been shown that hnRNP LL works by binding to the two lateral loops of the folded i-motif structure, it then unfolds the i-motif within the protein active site to enable transcription. The difference in footprinting patterns showed that this unfolded form was not the same as the alternative hairpin structure identified previously.[70] Molecular population dynamics experiments were used to show that on short time scales (60 s) the protein favoured binding to the i-motif structure, whilst on longer time scales (180 s), binding to the single strand was favoured.[86] This supports the hypothesis that the function of the i-motif structure is to

act as a scaffold to present the recognition sequences of the protein in the most kinetically favourable orientation for the protein to bind, after which the protein unfolds the structure to the most thermodynamically stable single-strand.

The identification of hnRNP LL is an important discovery in deciphering the possible biological function of the i-motif. Is it possible that there are hnRNP type proteins or similar which bind to all i-motifs activating transcription? hnRNP K is a transcription factor which binds to the C-rich i-motif forming sequence of c-Myc.[63] hnRNP K is very similar to hnRNP LL but recognises CCCT and TCCC sequences with the same 13 base separation which are found in loops 1 and 3 of the c-Myc i-motif.[49,87] Hurley and colleagues have hinted in their work on Bcl-2, that they have seen similar results for the hnRNP K/c-Myc system, but this has not yet been published. However, hnRNP K has also been shown to bind to telomeric i-motif forming sequences which also contain repeats of CCCT but in this case these recognition sites are within the cytosine core rather than the loop regions.[61] Lacroix *et al*. carried out experiments in the pH range from 6 to 9.2 showing binding of hnRNP K across this range. It was not ascertainable from their results whether hnRNP K bound to the single strand and hence unfolded i-motif at the lower pH's or, whether it preferred to bind to the i-motif and hence stabilised the i-motif at higher pH's. The example of hnRNP LL suggests that hnRNP K probably binds to the i-motif structure in c-Myc and unfolds it to activate transcription, whilst in the telomeric region hnRNP K most likely only binds to the single strand. Further investigation of this case could serve to strengthen the hypothesis that the i-motif in promoter regions can serve to activate transcription, whilst in telomeric regions can serve to inhibit protein recognition and binding.

Although the discovery of hnRNP LL and other similar proteins suggests that the i-motif may have a natural biological function, this is still a subject that is dividing opinion in the field and is a long way from being understood. Most recently, Kasturi Banerjee and co-workers have identified both G-quadruplex and i-motif forming sequences in the promoter region of the tyrosine hydroxylase (Th) gene that co-insides with the binding site of hnRNP K and they show that this protein activates transcription of the gene.[88] However, rather than agree with Hurley's suggestion that hnRNP K activates transcription through binding of the i-motif, they suggest that formation of the i-motif blocks the binding of hnRNP K and decreases transcription. They support

this hypothesis by showing that the compound **TMPyP4** which binds to both G-quadruplex and i-motif, has an inhibitory effect on transcription.[88] It is possible though, that since the i-motif is less stable than the G-quadruplex under physiological conditions, that **TMPyP4** is only binding to the G-quadruplex in this case and that the decrease in gene expression is soley through stabilisation of this structure. To fully understand the role of the i-motif within the Th promoter, a more specific ligand will be needed to deconvolute the possibly contrasting effects of the two structures.

Even if the i-motif turns out not to have a natural role in biology, the fact that this structure can form under certain conditions, should enable the i-motif to be targeted with small molecules to investigate its biological effect. At the time of writing, there are a very small number of i-motif ligands in the literature, this is in stark comparison to the G-quadruplex field where many ligands have been identified.[89] Despite this, it has already been shown that an i-motif specific ligand can inhibit telomerase,[59] whilst ligands specific for Bcl-2 can control gene expression.[69] This suggests that the i-motif is a worthwhile target to investigate but a wider range of ligands is required. This is one of the areas that will be investigated in this thesis.

## 1.4 Techniques for Investigating i-Motif – Ligand Interactions

In order to discover new i-motif binding ligands, a range of biophysical techniques are required to identify and investigate any i-motif – ligand interaction. There are already a wide variety of experimental methods that have been developed to study ligand interactions with duplex and G-quadruplex DNA.[90] However, the application of these experiments to study the i-motif, poses its own difficulties. *In vitro* most i-motif structures are only stable under acidic conditions, therefore to be certain that the ligand is binding to the i-motif structure, assays must be conducted under acidic conditions. This can have a quenching effect on the fluorescence of ligands or labels. It can also have an effect on the solubility of the ligand and the charge of the ligand may affect its ability to bind to the structure. Despite these potential problems, a number of existing methods have been successfully applied to the study of i-motif ligand interactions and these techniques will be utilised through out this research.

### 1.4.1 UV Absorbance Spectroscopy

When irradiated, DNA absorbs light at 260 nm. Due to the difference in π-π stacking interactions, folded DNA absorbs differently to unfolded DNA. It is therefore possible to use UV absorbance to monitor folding or unfolding of the structure. Typically, the i-motif shows an increase (hyperchromism) in absorbance at 260 nm and a decrease (hypochromism) in absorbance at 295 nm when it unfolds.[91]

UV absorbance can also be used to measure binding affinity for compounds interacting with the i-motif. Many DNA binding compounds are UV active due to their planar, aromatic structure. If these compounds absorb in a different region of the spectrum to the DNA, then by titrating DNA into the sample and monitoring the change in absorbance for the compound, a binding affinity can be measured. This was put to good effect by Laurence Hurley and co-workers when they identified the first i-motif binding compound, the porphyrin TMPyP4.[92] They used the absorbance of TMPyP4 at 425 nm which showed significant hypochromism (33%) upon addition of the sequence $d(AACCC)_4$ as well as a bathochromic shift indicating a binding event. From this data they were able to measure a binding affinity of 45 μM.[92] The disadvantage to this method is that ligands which absorb in the same region as DNA cannot be assessed as it is difficult to deconvolute the change in absorbance that is due to the ligand from that which is due to the DNA.

### 1.4.2 UV Melting and UV Difference Spectroscopy

The difference in absorbance between folded and unfolded DNA can be used in conjunction with a thermal peltier system to carry out DNA melting experiments. By monitoring the change in absorbance at a particular wavelength (typically 295 nm for the i-motif) with increasing temperature, the melting temperature of the DNA structure can be measured which is used to describe the stability of the structure.[91] By comparing melting temperatures of different i-motif sequences, their relative stabilities can be compared if performed under the same conditions. This was used recently by John Brazier and colleagues in identifying stable i-motifs in sequences complimentary to known G-quadruplexes.[41] As well as measuring thermal stability they also measured the change in UV absorbance with changing pH in order to measure

the transitional pH ($pH_T$) at which half the DNA is folded and half unfolded, identifying two i-motifs (PDGF-A and HIF-1α) with unusually high $pH_T$ as described in section 1.3.2.[41]

As well as just simply monitoring the change in UV absorbance at a specific wavelength, Mergny and co-workers have shown that by subtracting the complete UV spectrum of the folded (low temperature) sample from the spectrum of the unfolded (high temperature) sample, a new spectrum is generated which they term the thermal difference spectrum (TDS).[93] Mergny *et al*. carried out this experiment for a wide range of different structures including AT duplex, GC duplex, Z-DNA, parallel stranded DNA, GA DNA duplex, Hoogsteen base paired duplex, i-motif, G-quadruplex and pyrimidine triplex DNA. The data was taken between 220 and 330 nm, was normalised so that the maximum ΔAbs. = 1 and corrected to ΔAbs. = 0 at 320 nm as DNA should not give any absorbance at this wavelength. By testing multiple different sequences which form the same type of structure, they were able to show a great deal of similarity between TDS spectra of the same structure but clear differences between spectra of different structures. They have therefore identified the thermal difference spectrum as a useful

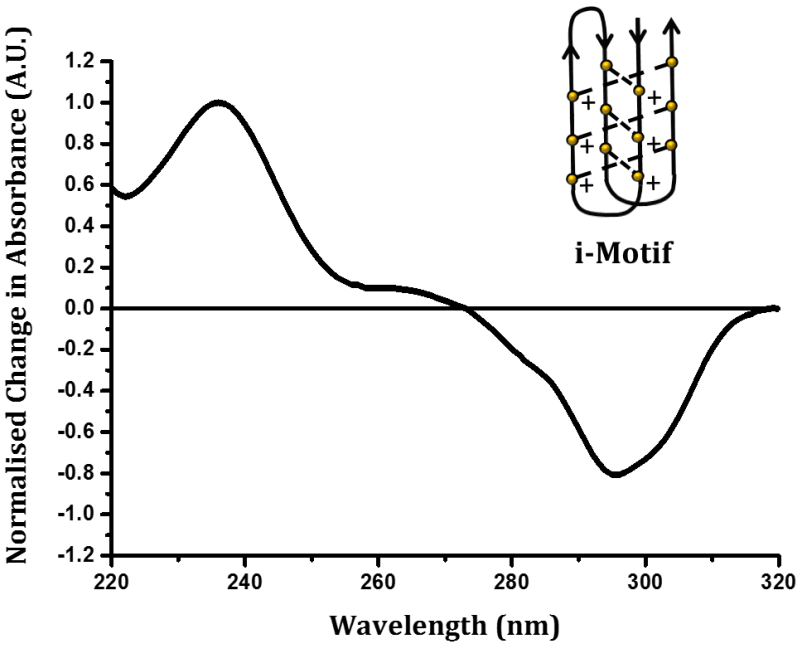



*Figure 1.8: UV thermal difference spectrum of human telomeric i-motif at pH 5.5. High (unfolded) T = 95°C, low (folded) T = 20°C. [DNA] = 2.5 μM, buffer = 10 mM sodium cacodylate and 5 mM sodium chloride.*

diagnostic tool to identify the type of DNA structure present in solution.[93] The thermal difference spectrum of i-motif DNA (figure 1.8) gives a characteristic positive peak at 240 nm and a negative peak at 295 nm. Thermal difference spectra were used by Jun Zhou and co-workers to demonstrate i-motif formation at neutral pH.[42] By taking the difference between the UV spectrum at 4°C and at 25°C they generated a TDS spectrum characteristic of the i-motif. This, in combination with other experiments, showed that at low temperature, i-motif formation could occur at neutral pH.[42]

### 1.4.3 Fluorescence Emission Spectroscopy

Analogous to UV absorbance titrations, if the potential ligand has fluorescent properties, then a binding affinity can be measured by fluorescence emission spectroscopy. Fluorescence occurs when a molecule in a singlet excited electronic state, transitions back to the ground state, releasing energy in the form of light.[94] In fluorescence titrations, the ligand is excited at a specific excitation wavelength that has been pre-determined for the molecule. The corresponding emission is then monitored at a separate emission wavelength as DNA is titrated into the sample. When the ligand binds to the DNA a change in fluorescence occurs because the binding effects the electronic distribution of the ligand. This change in fluorescence can either be an increase or a decrease depending on the fluorescence properties of the compound and can depend on the conditions of the experiment such as buffer and pH. If all other conditions are kept constant, the change in fluorescence can be assumed to be dependent on the binding of the ligand to the DNA and can therefore be used to determine binding affinities. The advantage of using this method over UV is that ligands which absorb in the same UV region as DNA can still be tested. However, the affinity of ligands which are not fluorescent or do not become fluorescent upon binding cannot be measured. The application of fluorescence spectroscopy to measure binding to the i-motif poses its own problems due the effect of pH on fluorescence. The i-motif generally requires acidic conditions in order to form *in vitro*, typical experiments are carried out at pH 5.5. However, many ligands are no longer fluorescent at this pH due to protonation of the compound, causing a change in the energy levels of the chromophore.

### 1.4.4 Förster Resonance Energy Transfer

Another method to use fluorescence spectroscopy is to attach fluorophores to the DNA. This can enable the monitoring of ligands which are not fluorescent themselves. If two different fluorophores are attached to seperate parts of the DNA sequence that move into close proximity upon folding, then the effect known as Förster resonance energy transfer (FRET) can be exploited. FRET is a physical phenomenon where by the fluorescence of one fluorophore (the donor) is transferred to a different fluorophore (the acceptor).[95] The effect relies on there being a sufficient energy match between the excited states of the two fluorophores, this is called resonance. The energy transfer occurs by a non-radiative mechanism that relies on a dipole-dipole interaction between the two fluorophores and is distance dependent, typically working in the range of 10 - 80 Å.[95] The efficiency of energy transfer ($E_T$) can be calculated from the change in fluorescence of the donor (equation 1.1) or from the distance between the two fluorophores (equation 1.2) and hence the change in fluorescence of the donor can be directly related to the distance between the two fluorophores.[95] This has already been put to use in the study of G-quadruplex ligands for a number of years.[96,97]

*Equation 1.1:* $$E_T = 1 - \frac{FI_d}{FI_d^0}$$

*Equation 1.2:* $$E_T = \frac{R_0^6}{(R^6 + R_0^6)}$$

Where $E_T$ is the FRET efficiency, $FI_d$ is the fluorescence intensity of the donor, $FI_d^0$ is the fluorescence of the donor in the absence of the acceptor, $R$ is the distance between the two fluorophores and $R_0$ is the Förster distance ($E_T = 0.5$).[95]

In 1999 Jean-Louis Mergny investigated the use of this technique to monitor the folding of an intramolecular i-motif.[98] Using a dual labelled oligonucleotide with the donor fluorophore FAM (6-carboxyfluorescein) attached to the 5' end and the acceptor fluorophore TAMRA (6-carboxytetramethylrhodamine) attached to the 3' end (figure 1.9). When the DNA is unfolded, the two fluorophores are sufficiently far apart for no FRET to occur (> 10 nm) but when the DNA folds into the i-motif structure, the two fluorophores are brought into close proximity, enabling energy transfer from FAM to TAMRA, resulting in a reduction in the fluorescence signal of FAM. This means that the

folding or unfolding of the DNA can be monitored by measuring the change in fluorescence of FAM at 515 nm (figure 1.10). This has been used recently by Laurence Hurley and co-workers who screened a library of compounds and identified i-motif stabilising compounds by a decrease in fluorescence and destabilising compounds by an increase in fluorescence though he used a different acceptor fluorophore called black hole quencher (BHQ).[69]

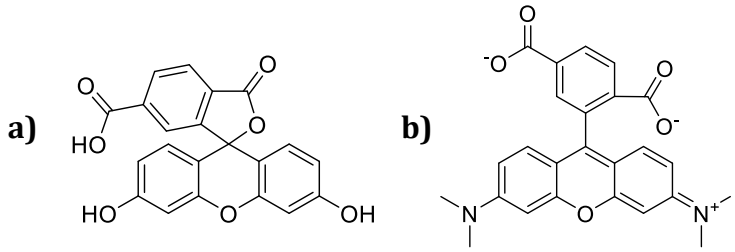



*Figure 1.9: Structures of a) donor fluorophore 6-carboxyfluorescein (**FAM**) and b) acceptor fluorophore 6-carboxytetramethylrhodamine (**TAMRA**).*

As well as looking at folding as a result of ligands, the FRET effect can also be used to measure a melting temperature of the i-motif. Mergny showed that monitoring the fluorescence of FAM with increasing temperature results in a sigmoidal curve from which the melting temperature can be determined in an analogous way to UV melting curves.[91,98] Using this stratergy, Khondaker Rahman and co-workers were able to screen over 2000 compounds for interaction with the human telomeric G-quadruplex, based on their stabilisation potential ($\Delta T_m$).[99] This method can also be applied to discover i-motif ligands. This was used for the first time by Patrizia Alberti and co-workers to identify the i-motif binding compound **BisA** showing that **BisA** stabilised the human telomeric i-motif by +33°C at pH 6.8.[100] This technique holds advantages over UV melting in that, due to the high fluorescence intensity of FAM, experiments can be carried out at much lower DNA and ligand concentrations and ligands which absorb in the same UV region as DNA can also be assessed. The disadvantage to this technique is that the FRET effect alone does not prove that i-motif formation is occurring, only that some form of folding is bringing the two fluorophores close together and i-motif formation specifically needs to be shown with a different technique such as UV difference or circular dichroism.

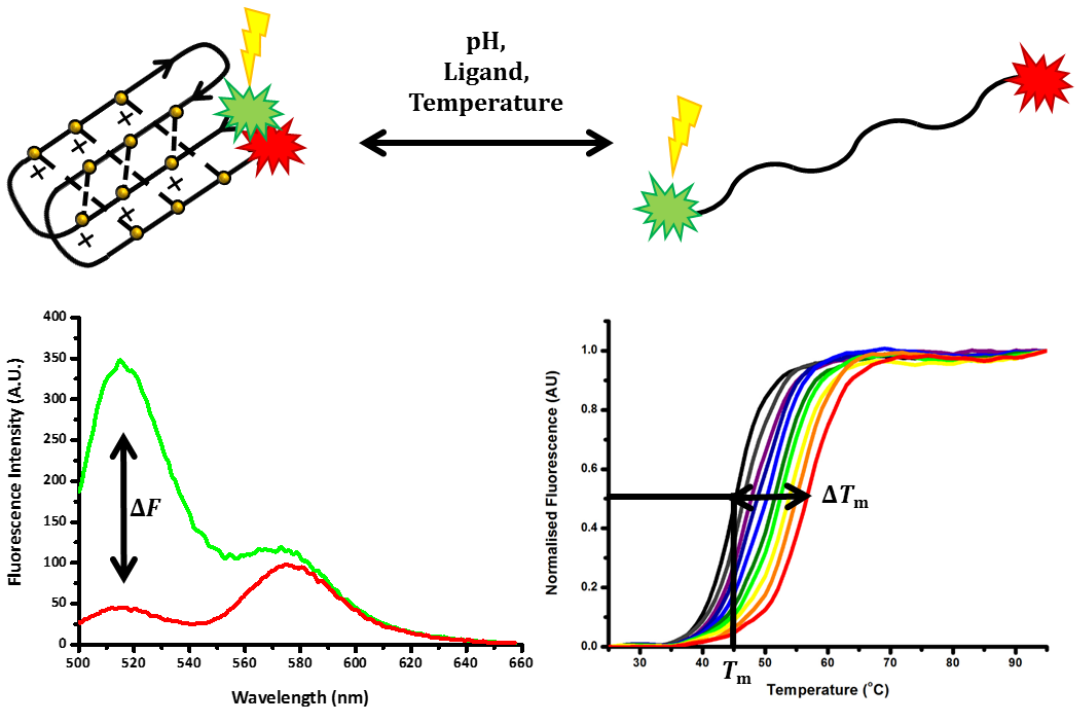



*Figure 1.10: An illustration of Förster resonance energy transfer used in FRET melting and FRET titration experiments.*

### 1.4.5 Circular Dichroism

Circular dichroism (CD) is the physical property of an optically active molecule where it is able to absorb left and right circularly polarised light differentially. A circularly polarised wave is the result of combining two waves which are perpendicular to each other and are out of phase by 90°. The summation of left and right circularly polarised light results in a plane-polarised light wave.[101] By the nature of their different glycosidic bond angles, different DNA secondary structures have different optical activities and hence exhibit different circular dichroism. CD spectroscopy exploits this property by measuring the CD signal of a sample (often quoted in molar ellipticity) with changing wavelength. This gives a specific spectrum depending on the type of DNA structure and hence can be used as a diagnostic tool to indicate which structure or structures may be present in the sample.[102] i-Motif DNA structure gives a characteristic CD spectrum with a positive peak at 288 nm and a negative peak at 255 nm (figure 1.11) whereas single stranded DNA gives a spectrum with a positive peak at 275 nm and a negative peak at 250 nm and is less optically active, giving a weaker signal. CD spectroscopy can therefore be used to monitor folding of single stranded DNA to i-

motif and vice versa as a result of changing pH, temperature or ligands. Ligands interacting with the i-motif do not need to be optically active themselves, as a ligand binds to the DNA, as well as possibly affecting the conformation of the DNA structure, the ligand itself can become optically active through the process of chirality transfer. This results in a new induced CD signal (ICD). In much the same way as with UV and fluorescence changes, by measuring the change in CD signal with changing ligand concentration, the data can be fitted to obtain a binding constant.[11]

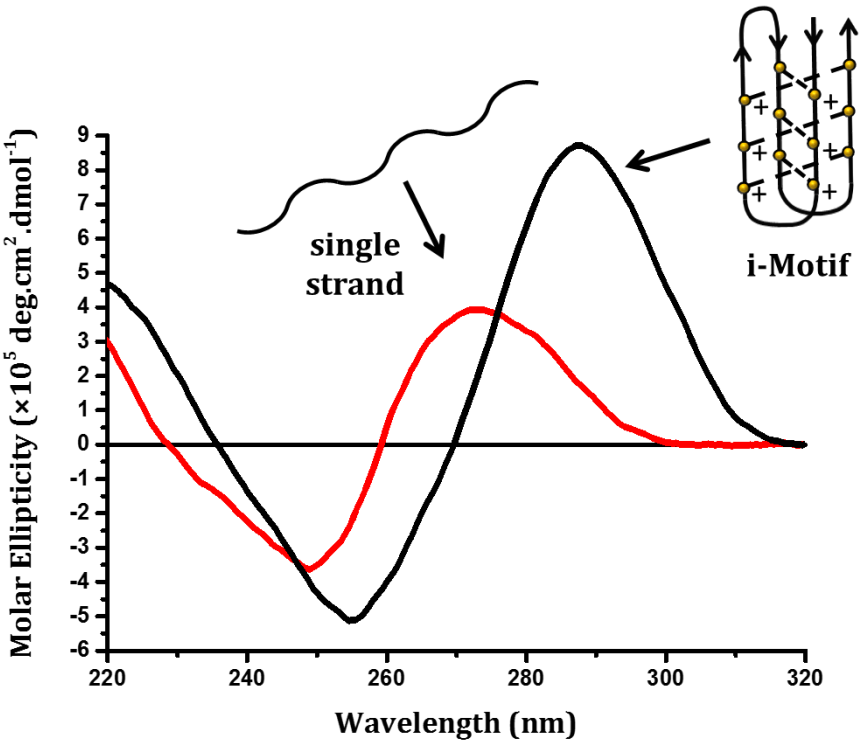



*Figure 1.11: Example circular dichroism spectra of single stranded and i-motif DNA. [DNA] = 10 μM, buffer = 50 mM sodium cacodylate, pH 5.5 (i-motif) or pH 7.4 (ssDNA).*

### 1.4.6 Surface Plasmon Resonance

Surface Plasmon Resonance (SPR) is a very sensitive technique for measuring binding affinities by monitoring binding to a target structure attached to a gold film. Using oligonucleotides labelled with Biotin (figure 1.12) at the 5' end of the sequence, the DNA structure of interest (i.e. the i-motif) can be attached to a streptavidin coated gold chip due to the very high affinity of streptavidin for biotin ($K_d$ ~0.01 pM).[103] Laser light is passed through a prism onto the underside of the gold at the correct angle to make plasmons in the gold surface resonate, creating an interference band in the reflected laser light which is detected. Running buffer continuously flows over the surface of the

chip and different concentrations of compound are injected into the system. When a compound binds to the DNA, it creates a change in the refractive index in the solution immediately above the gold surface. This affects the resonance of the gold plasmons and in turn moves the position of the interference band within the reflected laser light. This is what is measured as the SPR response (figure 1.13). This is a powerful technique which can measure very high binding affinities down to picomolar sensitivity. Binding affinities can be obtained in two different ways. By testing a range of different ligand concentrations in a titration experiment, and monitoring the response at equilibrium with ligand concentration, a binding curve can be produced and fitted with standard equilibrium binding models.[104] Alternatively the SPR sensorgram itself can be fitted using a kinetics based model to obtain the association

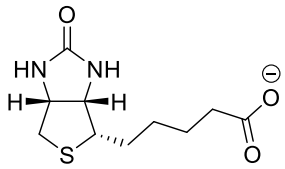



*Figure 1.12: The structure of Biotin.*

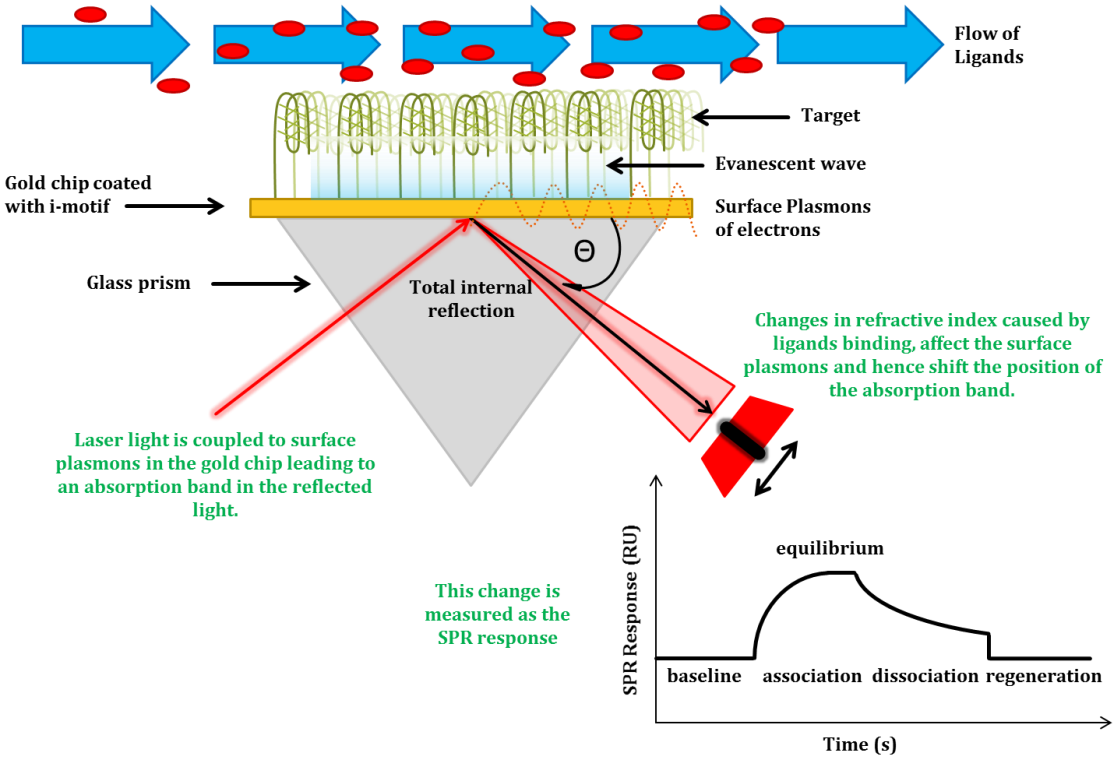



*Figure 1.13: Schematic diagram of a surface plasmon resonance experiment.*

and dissociation rates, which can be used to obtain the equilibrium binding affinity. This technique was used by Laurence Hurley and co-workers in their recent study of the Bcl-2 i-motif. They used SPR to determine the strong binding affinity of the transcription factor hnRNP LL measuring a $K_d$ of 19.4 pM at pH 6.5.[70]

Aside from measuring binding affinities, SPR has more frequently been used to monitor folding and unfolding of the i-motif under different conditions. Zhi-Xiong Zeng and colleagues used SPR to indicate i-motif formation in the human telomeric sequence as it inhibited the binding of a single strand binding protein from *E. coli*.[105] Hongwei Xia and co-workers showed that the density at which the DNA sequence is grafted to the chip is important, as higher densities prevent i-motif formation due to crowding effects.[106] Finally Xiaogang Qu's group have used SPR to monitor i-motif folding in a DNA nanomachine driven by single-walled carbon nanotubes which selectively induce i-motif formation.[107,108] Despite this though, SPR has not yet been used to measure the affinity of a small molecule ligand for the i-motif.

### 1.4.7 Gel Electrophoresis

Gel electrophoresis can be used to distinguish between different DNA structures and whether a small molecule, peptide or oligonucleotide sequence is bound to the target sequence based on their different mobilities through the gel.[90,109] This is often termed an electrophoretic mobility shift assay (EMSA) and is typically used to determine DNA-protein interactions.[110] Small single stranded sequences are able to move through the gel more easily than folded structures. However, small more compact structures are able to move through the gel more easily than bigger structures. Finally charged sequences such as DNA will move through the gel more rapidly than neutral sequences due to the charge repulsion or attraction from the electrodes. As multiple DNA strands come together to form a structure, or a protein, peptide or small molecule binds to the DNA, a complex is formed with a higher molecular weight which therefore moves more slowly through the gel. There can also be some charge neutralisation occurring as, for example, a positively charged ligand binds to the negatively charged oligonucleotide. This again has the effect of slowing down the movement of the sample through the gel. In comparing gel results, it is important that the same control sample is present in one of the lanes in each gel which will move consistently under the same conditions. By

comparing the relative position of the bands to the control sample it can then be determined whether a change has occurred. Gels can be imaged by fluorescence, using fluorescently labelled oligonucleotide sequences or dyes, by UV using a dye such as ethidium bromide to stain the gel or by radioactivity using $^{32}$P-labelled DNA.[110] This technique has been used in a number of studies to show i-motif formation including at neutral pH[42] and formation of multiple different i-motifs in the c-Myc promoter sequence.[62] The EMSA has also been used to show binding of proteins to C-rich single stranded DNA[60,61] and the i-motif.[70]

## 1.5 Existing i-Motif Ligands

### 1.5.1 TMPyP4 and BisA

The first compound to be investigated as a potential i-motif ligand was the porphyrin 5,10,15,20-tetra-(*N*-methyl-4-pyridyl)-porphine **TMPyP4** (figure 1.14).[92] This compound is a known G-quadruplex ligand which has already been shown to decrease c-Myc expression,[111] inhibit telomerase and tumour growth.[112] Hurley and co-workers investigated **TMPyP4**'s interaction with telomeric repeat sequences forming an intermolecular i-motif d(CCCAAT)$_4$ and d(AATCCC)$_4$.[92] They showed, using an electrophoretic shift assay, that not only did **TMPyP4** bind to the i-motif but that it promoted its formation at pH 4.5. Despite this, **TMPyP4** had little effect on the melting temperature of the i-motif ($\Delta T_m$ <20°C) and they measured a relatively weak dissociation constant of 45 μM using UV titrations. **TMPyP4** has a stronger affinity for G-quadruplex ($K_d$ = 0.5 μM) and duplex DNA ($K_d$ = 1.2 μM) in comparison.[113] Finally Hurley's group suggested a cooperativity in the binding of **TMPyP4** and using NMR modelling experiments, indicated that **TMPyP4** bound in a 2:1 stoichiometry by stacking on the top and bottom of the four stranded structure. The affinity of **TMPyP4** for i-motif may vary however, as its interaction with the intramolecular human telomeric i-motif was assessed, giving a $K_d$ of 1 μM, but this result was calculated from UV melting data so is not directly comparible.[114] A similar sequence where adenines in the loops were substituted with thymines was also studied, showing a difference in stoichiometry of binding between the mutant (1:2 and 1:1) and natural (just 1:1) sequences. This suggests that the adenines in the loop region or the difference between the topology of the two i-motif structures can alter the binding of **TMPyP4**. In their

previous work on **TMPyP4** with G-quadruplexes Laurence Hurley and co-workers showed that **TMPyP4** prefers to bind to intermolecular structures over intramolecular structures promoting the formation of anaphase bridges, down regulating gene expression and inhibiting telomerase.[111,112] This may also be the case for i-motif structures, but is yet to be explored.

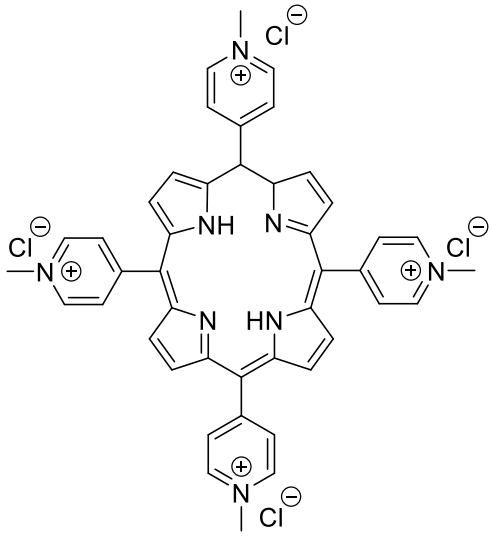



*Figure 1.14: The structure of porphyrin 5,10,15,20-tetra-(N-methyl-4-pyridyl)porphine,* ***TMPyP4****.*

The macrocyclic acridine compound **BisA** (figure 1.15) was the first i-motif interacting compound to be identified using Förester resonance energy transfer (FRET) DNA melting in 2001 as described in section 1.4.4.[100] In this example, they showed that **BisA** had a significant stabilising effect on both the human telomeric i-motif at pH 6.8 ($\Delta T_m$ = +33°C) and the human telomeric G-quadruplex at pH 7.2 ($\Delta T_m$ = +15°C). **BisA** was also shown to inhibit telomerase with an $IC_{50}$ of 0.75 μM but further explorations of this ligand were focused on its interaction with other DNA structures. A more detailed investigation of **BisA**'s interaction with i-motif DNA is discussed in chapter 2.

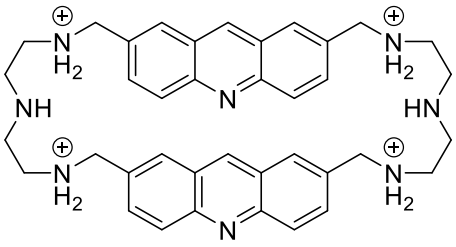



*Figure 1.15: The structure of the macrocyclic acridine* ***BisA****.*

### 1.5.2 Carboxyl-Modified Single-Walled Carbon Nanotubes

One of the most well-studied and unusual i-motif ligands is carboxyl-modified single-walled carbon nanotubes (**SWNTs** figure 1.16a). Investigations by Xiaogang Qu and co-workers in 2006 used a wide range of experiments including FRET, CD, UV melting and NMR to show that **SWNTs** could inhibit duplex formation and promote human telomeric i-motif formation in conditions up to pH 8.0 with a significant stabilisation temperature of +22°C at pH 5.5.[107] The carboxyl-modified nanotubes were prepared by sonicating the SWNTs in a mixture of concentrated sulfuric acid and nitric acid (3:1 vol/vol) so it is not clear how many acid groups are present on any single carbon nanotube.[115] This is a particular concern with respect to the pH sensitivity of the i-motif structure but Qu *et al*. state that the pH did not change in their experiments.[107] Due to their size (diameter = 1.1 nm[107]), Qu and co-workers hypothesised that the **SWNTs** bind in the 5' major groove of the i-motif, enabling the carboxyl groups to reach into the core of the structure and pick up favourable interactions with the hemi-protonated cytosine base pairs. Larger carboxyl-modified multi-walled carbon nanotubes prepared in the same way, which are too big to fit into the major groove, were also tested and shown not to stabilise the i-motif providing further evidence that it is more than just the presence of acidic groups having an effect.[59] Further studies by the group showed that **SWNTs** accelerated the S1 nuclease cleavage rate and that they are able to promote i-motif formation at neutral pH under molecular crowding conditions suggesting they may be able to have an effect *in vivo*.[116,117] Most recently, Xiaogang Qu's group carried out an extensive study into the biological effects of **SWNTs** binding to telomeric i-motif in cancer cells.[59] They showed that **SWNTs**, which are specific for i-motif over other DNA structures, are able to inhibit the activity of telomerase (figure 1.6), resulting in DNA damage response, displacement of telomere binding proteins, telomere uncapping, cell cycle arrest and finally apoptosis. They showed this both *in vitro* with an $IC_{50}$ of 0.27 ± 0.02 μg.mL$^{-1}$ and in living K562 and HeLa cancer cells with an $IC_{50}$ of 10.2 ± 0.8 and 7.5 ± 0.45 μg.mL$^{-1}$ respectively. They were able to show that the **SWNTs** accumulated in the telomeric regions of nuclear DNA using fluorescently labelled analogues, which co-localised with a marker for interphase telomeres TRF1.[59] What is not yet clear is whether the resultant telomerase inhibition is due to formation of the i-motif directly, or due to the fact that formation of the i-motif and destabilisation of the duplex will result in concomitant G-quadruplex formation, which is already

shown to inhibit telomerase.[118] Qu and co-workers suggest that this may well be the case.

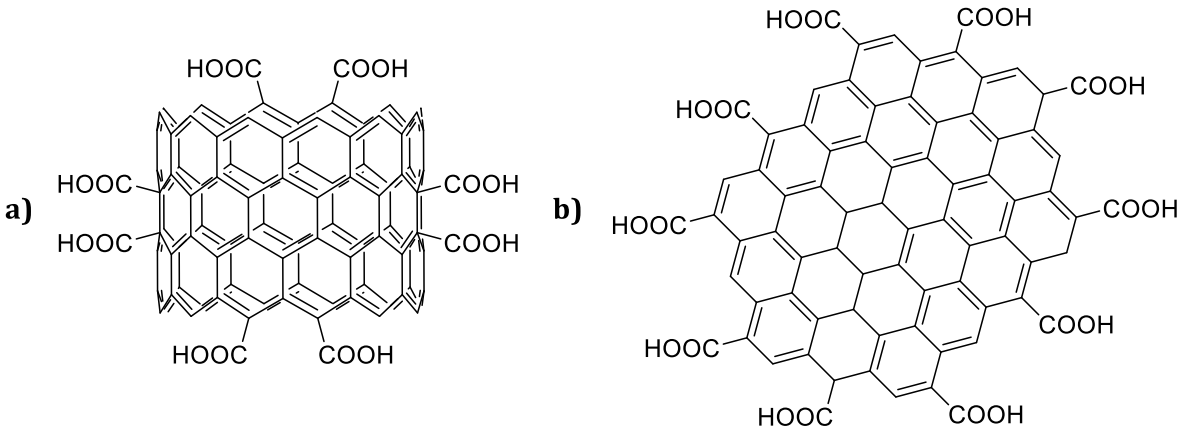



*Figure 1.16: Illustrative structures of a) carboxyl-modified single-walled carbon nanotubes, **SWNTs**, b) carboxyl-modified graphene quantum dots, **GQDs**.*

Carboxyl-modified graphene quantum dots (**GQDs** figure 1.16 b) have also been investigated as a logical extension of work on **SWNTs**.[119] These were also shown to both stabilise and promote i-motif formation at up to pH 8.0 by end stacking interactions with the loops. This was shown for both telomeric and c-Myc i-motif sequences with stabilisation of +20 °C and +16 °C respectively, dependent on concentration, but again the precise number of carboxyl groups on each **GQD** is not certain so maintaining pH is an important consideration.[120]

### 1.5.3 Bcl-2 Specific Ligands

Two of the most promising ligands for deciphering a potential biological function for the i-motif are the compounds **IMC-48** and **IMC-76** (figure 1.17) which were identified by Laurence Hurley and co-workers in 2014.[69] Having already shown that the Bcl-2 i-motif forming sequence is in dynamic equilibrium between the i-motif structure and a flexible hairpin structure, they used a FRET based high throughput screen to screen the National Cancer Institute (NCI) Diversity set of 1990 compounds for interaction with Bcl-2 i-motif based on changes in fluorescence at pH 5.8. **IMC-48** showed a decrease in fluorescence and was hence identified as an i-motif stabilising compound, whereas **IMC-76** showed an increase in fluorescence and hence was identified as an i-motif destabilising compound. In FRET studies, the two compounds were shown to be

very specific for the Bcl-2 i-motif sequence in comparison to a Bcl-2 mutant sequence, the c-Myc i-motif, the VEGF i-motif, the Bcl-2 G-quadruplex and the BcL-2 duplex sequence. Some changes in fluorescence were seen for the VEGF i-motif but both compounds were found to have stronger affinities for Bcl-2 (**IMC-48** $K_d$ = 0.49 μM, **IMC-76** $K_d$ = 1.01 μM) than VEGF (**IMC-48**: $K_d$ = 4.7 μM, **IMC-76**: $K_d$ = 5.73 μM). NMR studies showed that **IMC-48** favoured the formation of i-motif structure while **IMC-76** favoured the formation of the hairpin structure. This was confirmed by circular dichroism and bromine footprinting experiments. The NMR also showed that the two compounds were antagonistic towards one another.

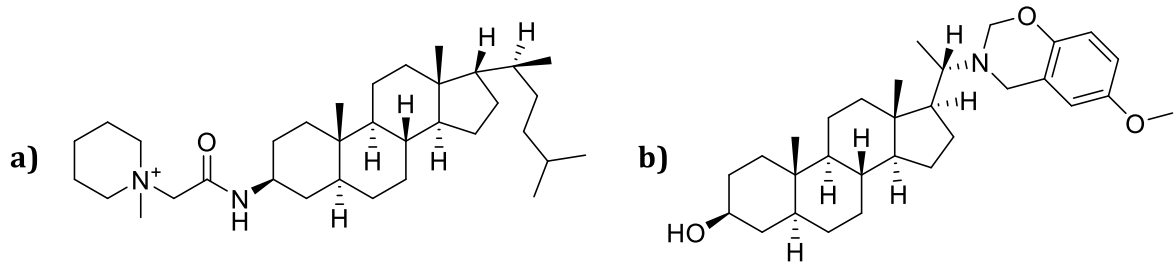



*Figure 1.17: The structures of the Bcl-2 specific ligands a)* ***IMC-48*** *and b)* ***IMC-76****.*

**IMC-48** did not bind to the c-Myc i-motif, however, when the c-Myc sequence was mutated so that the central loop had the same sequence as Bcl-2, binding did occur. This demonstrated that the central loop is the binding site for **IMC-48**. Mutations to the Bcl-2 lateral loops also did not affect **IMC-48** binding. As well as the observed structural antagonistic effects, **IMC-48** and **IMC-76** also had opposite effects on Bcl-2 gene expression. It has previously been hypothesised that, like the G-quadruplex, stabilisation of the i-motif structure would act as a block to the "transcription machinery" and hence result in a decrease in gene expression. In contrast, Hurley and co-workers showed that treatment with **IMC-48** actually increased Bcl-2 gene expression whereas **IMC-76**, favouring the hairpin, decreased gene expression. This effect was seen both at the level of transcription and translation, with the antagonistic relationship between **IMC-48** and **IMC-76** also apparent. As discussed earlier, these effects on gene expression were shown to be due to the presence of a transcription factor, hnRNP LL, which binds specifically to the i-motif and unfolds the structure in order to activate transcription.[70] Importantly, by binding to the central loop of the i-motif, **IMC-48** does not interfere with the binding mode of hnRNP LL and increases gene expression by increasing the amount of i-motif available for hnRNP LL to bind to.

Conversely, **IMC-76** decreases gene expression by reducing the amount of i-motif structure available for hnRNP LL to bind. Molecular population dynamics experiments showed that **IMC-48** favoured the i-motif structure over short time periods (60 s) but over longer periods (180 s) favoured unfolding of the structure to the single strand.[86] This is similar to the effect of hnRNP LL (see section 1.3.3) and indicates that **IMC-48** may have an activating effect on transcription of its own through unfolding of the i-motif after binding. Hurley's group also tested four smaller compounds (figure 1.18) where the positively charged piperidine unit was kept constant and the amide linked group was varied.[69] Each of these compounds were shown to stabilise the i-motif by FRET whilst **IMC-42** was also tested for its effect on Bcl-2 expression and shown to increase expression to a level comparable with **IMC-48** in MCF-7 cells. This suggests it is the positively charged piperidine that is important in stabilising the i-motif, most likely through interactions with the negative phosphate backbone, whilst the rest of the molecule binds to the central loop through Van der Waals and stacking interactions.

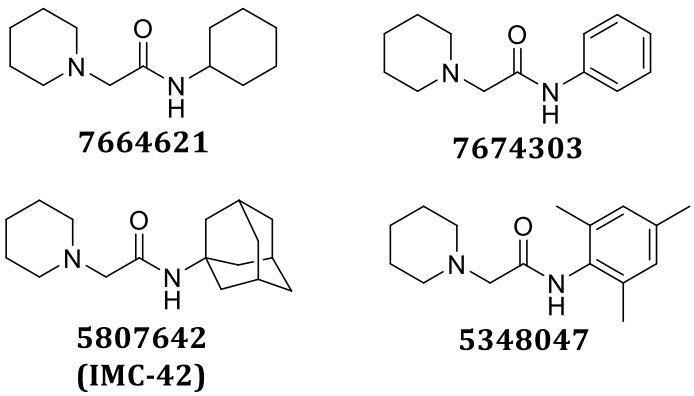



*Figure 1.18: The structures of Bcl-2 binding piperidine derivatives including **IMC-42**.*

Finally, Hurley and co-workers demonstrated, that in cells resistant to the chemotherapy drugs etoposide (B95.8 cells) and cyclophosphamide (GRANTA-519 cells), co-dosing with **IMC-76** caused chemosensitisation to the drugs as a result of reduced Bcl-2 expression. Caspase-3 activity was monitored as a marker for apoptosis resulting in a 2.5-fold and a 1.9-fold increase in each case.[121] The antagonistic effect of **IMC-48** was also apparent as when **IMC-76** and **IMC-48** were used in a 1:1 ratio with cyclophosphamide the chemosensitisation did not occur. Chemosensitisation to etoposide was also observed *in vivo* in a resistant lymphoma severe combined immunodeficient (SCID) mouse model with GRANTA-519 lymphoma xenografts. Treatment with 6 – 8 mg/kg of **IMC-76** resulted in a 20% decrease in Bcl-2 mRNA level

and a greater 48% decrease at a dose of 10 mg/kg. This was also observed at the protein level. Co-treatment of the mice with 10 mg/kg of **IMC-76** and etoposide resulted in a 65% inhibition of tumour growth.

### 1.5.4 Other i-Motif Ligands

There are very few other i-motif ligands of note and most have not been investigated to the same extent as those discussed previously. In most cases they show weak affinities for the i-motif and poor selectivity over other DNA structure types. For example the phenanthroline derivatives (figure 1.19) have each been shown to be mildly stabilising of the i-motif at acidic pH ($\Delta T_m$ 7 – 10°C with 10 eq. of ligand) with binding affinities of 4 – 8 μM.[122] However they have stronger affinity for G-quadruplexes ($K_d$ = 1.6 – 2.5 μM) and slightly weaker affinity for double stranded DNA ($K_d$ = 6.7 – 12 μM).

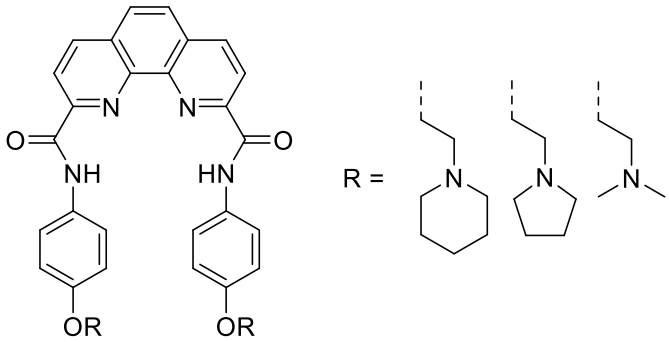



*Figure 1.19: The structures of the phenanthroline derivatives.*

The **neomycin-perylene** conjugate (figure 1.20) showed some degree of affinity for the i-motif in competition dialysis experiments, measuring a bound concentration of ~1.2 μM which was stronger than other G-quadruplexes and duplex tested (~0.2 μM bound concentration) but weaker than for telomeric G-quadruplex (~4 μM bound concentration).[123] However these experiments were all carried out at pH 7 where the i-motif is normally unfolded so it may be that the observed affinity is for the single strand rather than the i-motif structure itself.

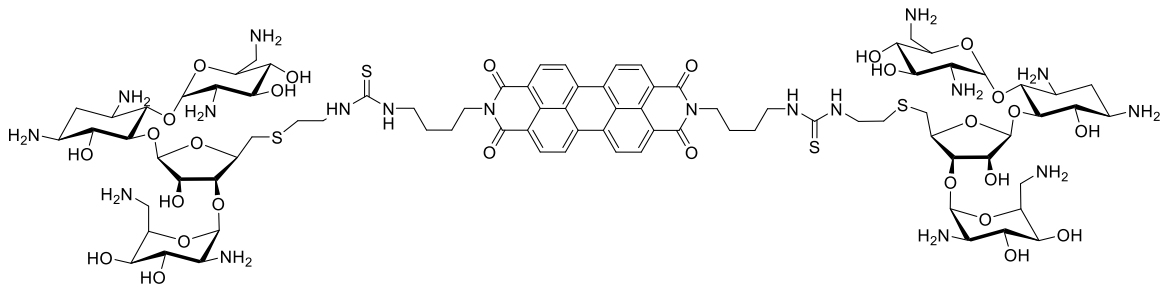


*Figure 1.20: The structure of the **neomycin-perylene** conjugate.*

**Crystal violet** (figure 1.21) has been investigated as an i-motif ligand by several groups due to its fluorescent properties upon binding to DNA secondary structures. While **crystal violet** had previously been shown to bind to G-quadruplex DNA and to favour G-quadruplex over double stranded and single stranded DNA,[124] it had not been studied as a potential i-motif ligand until 2011 when Chung-Lang Leung and co-workers developed a DNA-based OR logic gate.[125] Using UV titrations they were able to measure relatively strong binding affinities for **crystal violet** with a model 29 base sequence d$(C_5T_3)_3C_5$ ($K_d$ = 0.83 μM), a 21 base sequence d$(C_3T_3)_3C_3$ ($K_d$ = 1.4 μM) and a 17 base sequence d$(C_2T_3)_3C_2$ ($K_d$ = 1.3 μM). This shows a trend in affinity consistent with the trend in i-motif stability of the sequences however, all of these affinities are significantly stronger than the affinity of **crystal violet** for the comparable 28 base and 21 base G-quadruplex sequences with $K_d$'s of 53 μM and 28 μM respectively. Using molecular docking experiments they suggest that crystal violet binds to the simplified intermolecular i-motif $(ACCCT)_4$ *via* end stacking interactions with a binding energy of -38.44 kcal.mol$^{-1}$.[125] More recently, the binding of **crystal violet** to i-motif DNA and its subsequent increase in fluorescence has been utilised by Zhou *et al*. as a probe for a NOTIF logic gate, consisting of a G-quadruplex and i-motif within the same strand,[126] and by Zhang *et al*. as a probe for a surface based electrochemical pH-switch.[127]

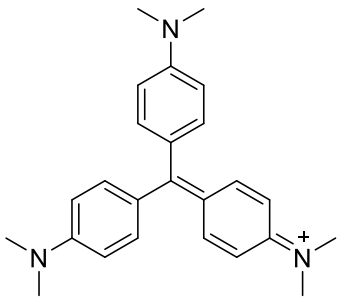
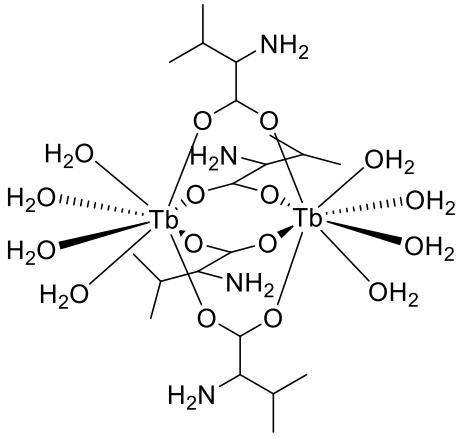
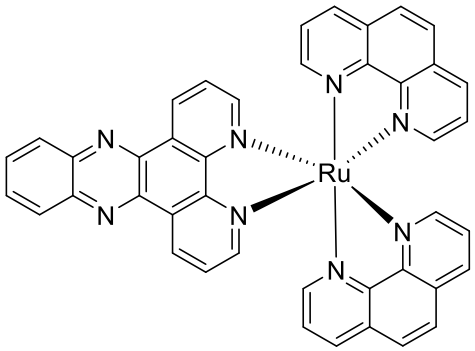


*Figure 1.21: The structure of crystal violet, **CV**.*

Complexes of terbium and ruthenium (figure 1.22), which are known G-quadruplex ligands, have been investigated as potential i-motif ligands. The amino acid complexes **$[Tb_2(_{DL}\text{-Cys})_4(H_2O)_8]Cl_2$** and **$[Tb_2(_{DL}\text{-HVal})_4(H_2O)_8]Cl_6$** were shown to bind to telomeric i-motif at pH 5.5 with comparable affinity ($K_d$ = 22 μM and 30 μM respectively) to the telomeric G-quadruplex at pH 7.1 ($K_d$ = 26 μM and 23 μM respectively).[128] However, neither complex was able to stabilise the structures, showing a small decrease in melting temperature for both the i-motif ($\Delta T_m$ = -0.5°C and -4.0°C respectively) and the G-quadruplex ($\Delta T_m$ = -4.0°C and -3.0°C respectively). The ruthenium complexes **$[Ru(bpy)_2(dppz)]^{2+}$** and **$[Ru(phen)_2(dppz)]^{2+}$** are slightly more promising as they have stronger affinities than the terbium complexes ($K_d$ = 5.6 μM and 1.5 μM respectively) and do not destabilise the i-motif structure.[129,130] However they are still not specific. **$[Ru(bpy)_2(dppz)]^{2+}$** has a stronger affinity for the G-quadruplex ($K_d$ = 0.7 μM vs. 5.6 μM) while **$[Ru(phen)_2(dppz)]^{2+}$** has a comparable affinity for the G-quadruplex ($K_d$ = 2.1 μM vs. 1.5 μM) all at pH 5.5. However their increase in fluorescence signal upon binding enable their potential use as a light switch for G-quadruplex and i-motif structures.[130]




**$[Tb_2(_{DL}\text{-HVal})_4(H_2O)_8]Cl_6.2H_2O$**




**$[Ru(phen)_2dppz]^{2+}$**

*Figure 1.22: The structures of terbium and ruthenium complexes.*

Pyrene or pyrene derivatives may also have potential as i-motif ligands. So far they have not been investigated as free ligands but they have been examined as part of covalent modifications to the i-motif structure. Modification of a terminal adenine base with pyrene (figure 1.23a) for use as a switch off fluorescent probe for i-motif formation resulted in a stabilisation of the structure ($\Delta T_m$ = +11.4°C).[131] Further modification of a terminal adenine and a central loop adenine in a dimeric sequence,

resulting in incorporation of 4 pyrene units into the i-motif structure, gave an even greater stabilisation of +23.5°C at pH 5.0.[132] However, incorporation of the pyrene modified adenine base, into the central loop of an intramolecular i-motif resulted in a destabilising effect ($\Delta T_m$ = -8.9°C) indicating that stacking on the terminal end of the i-motif is preferable to incorporation inside the small central loop of this sequence.[133] These results have been supported by investigations of other groups also showing stabilising effects of terminal modifications and destabilising effects of central loop modifications. For example, incorporation of the pyrene intercalator **TINA** (figure 1.23b) at the 3 different positions of the central loop of the telomeric i-motif resulted in destabilisation of the structure at pH 5.2 ($\Delta T_m$ = -4.5°C to -8°C) but very little change at pH 6.2 ($\Delta T_m$ = 0°C to -1°C).[134] Finally, a dual pyrene labelled c-Myc sequence (figure 1.23 c) was both stabilised and destabilised depending on pH and whether an extra base was incorporated.[135] For the sequence with an additional adenine base, the terminal pyrenes had a destabilising effect at lower pH ($\Delta T_m$ = -6.5°C at pH 5.0) but a stabilising effect at higher pH ($\Delta T_m$ = +25.8°C at pH 7.0). Whereas, without an additional base, the sequence was stabilised at both pHs ($\Delta T_m$ = +1.4°C at pH 5.0 and +7.1°C at pH 7.0).[135]

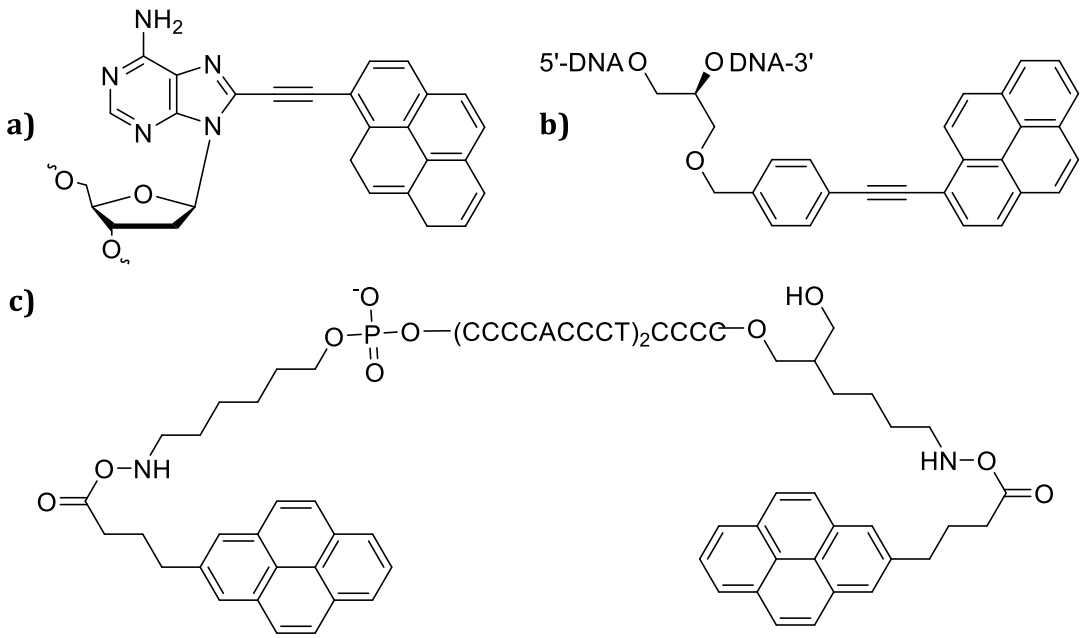



*Figure 1.23: The structures of a) pyrene modified adenosine, b) the pyrene intercalator **TINA**, c) the pyrene labelled c-Myc sequence.*

## 1.6 Challenges in i-Motif Research

In summary, the i-motif is an alternative DNA secondary structure which can form *in vitro* in sequences complimentary to the G-quadruplex, in telomeric regions and close to transcription start sites with a variety of different topologies depending on the length and composition of the sequence.[16] However, due to the instability of the structure and its reliance on acidic pH to form *in vitro*, there is still a lot of doubt as to whether the structure is physiologically relevant and therefore the potential biological application of the i-motif has had very little investigation. Work by the groups of Xiaogang Qu[59] and Laurence Hurley[69] have begun to relate some biological effects of telomerase inhibition and control of gene expression to i-motif binding compounds, suggesting that, at least in these special cases, the i-motif is a potential drug target and it is worthy of further investigation. As described in section 1.5, there are relatively few i-motif ligands in the literature and most are not specific, unlike the G-quadruplex for which there have been many ligands identified.[89] So, in order to study the i-motif in more detail, a better chemical toolbox of ligands is needed.

As Laurence Hurley and co-workers have shown for Bcl-2,[69] the i-motif may turn out to be quite a dynamic structure, in equilibrium with other structures and factors that may affect i-motif stability and this equilibrium, such as the presence of cations, also need to be investigated. As described in section 1.4 there are a variety of techniques available which have been used extensively in the G-quadruplex field and are now being used to study the i-motif. However, applying these techniques to the i-motif introduces its own challenges with factors such as fluorescence and solubility that are strongly affected by pH.

## 1.7 Aims and Objectives

The aim of this project is to investigate the effect of both small molecules and cations on the structure and stability of i-motif DNA. This is in order to identify new i-motif binding ligands which can stabilise the structure and to look at how different cationic conditions may affect i-motif formation. The overall objective is to discover a range of tools that could be used to investigate the chemical biology of i-motif structures.

This work will be carried out in 3 stages described in this thesis. First in chapter 2, the existing i-motif stabilising ligand, **BisA**,[100] is examined in detail using a variety of biophysical techniques that have not previously been used to look at this compound's interaction with the i-motif. This is in order for it to be used as a positive control when looking for new i-motif ligands. In chapter 3, a library of 960 compounds are screened using a FRET melting assay, for interaction with the i-motif based on their ability to stabilise or destabilise the DNA. The purpose of this screen is to discover new i-motif binding compounds. Because there are so few i-motif ligands in the literature, a very diverse library is used to include a wide variety of structure types and chemical properties. Finally in chapter 4, the same screening methodology is applied to look at the effect of different cations on i-motif forming sequences both at acidic and neutral pH to see what effect they have on the stability of the structure. Some small studies have been carried out previously[50–52] but this is expanded in this thesis to include other physiologically relevant cations not previously investigated.

# Chapter 2: Characterisation of the i-Motif Binding Compound BisA

# 2.1 Introduction

### 2.1.1 BisA is a known DNA binding compound

The bisacridine macrocycle **BisA** (figure 1.15) was first synthesised by Marie-Paule Teulade-Fichou in Jean-Marie Lehn's group in 1995.[136] The group were particularly interested in **BisA** for its fluorescence properties and developed the compound as a tool for the selective molecular recognition of nucleosides and nucleotides. The stacked planar structure of the molecule gave sufficient space between the two acridine units (7 Å,[137]) to allow it to intercalate with a nucleoside or nucleotide base resulting in a change in fluorescence and effectively forming a receptor for these molecules.

Using fluorescence titrations, they showed that the fluorescence of **BisA** was significantly enhanced by 70 – 130% in the presence of pyrimidine bases, while in the presence of purine bases the fluorescence of **BisA** was quenched by 70 – 80%. This change in fluorescence was consistent with both adenine and guanine vs cytosine and uracil. They also tested the nucleoside as well as the nucleotide mono-, di- and tri-phosphates showing the same enhancement or quenching of the fluorescence in all cases.

After this initial work, the group moved on to investigate the interaction of **BisA** with hairpin DNA structures.[138,139] They showed that **BisA** had a high binding affinity for the hairpin sequence d(GCGAAACGC) and selectively stabilised the hairpin structure over the alternative bulged duplex structure. On comparing the fluorescence spectra of **BisA** bound to the hairpin with the spectra of **BisA** in the presence of single stranded adenine sequences ($dA_3$ and $dA_{20}$) as well as adenine monophosphate (AMP) they observed the same $\lambda_{max}$ (441 ±2 nm). They therefore concluded that the binding site of **BisA** was the central adenine loop region with the most central adenine likely to be intercalating between the two acridine units of the structure.[138] Using polyacrylamide gel electrophoresis where the duplex and hairpin structures are clearly visible as separate bands, they were able to show that increasing amounts of **BisA** actively shifted the equilibrium from duplex to hairpin structures.[139] However, they also showed that higher concentrations of **BisA** (>25 μM) did cause aggregation of the DNA.

This was seen in gel experiments where, at lower **BisA** concentrations, they observed a 1:1 DNA:**BisA** complex which moved freely through the gel, whereas at higher **BisA** concentrations, aggregation was indicated by bands not migrating from the well and a smearing effect of the bands that did migrate.[139]

In 1998 Jean-Marie Lehn and co-workers, exploited the preference of **BisA** for single stranded DNA over double stranded DNA to show that **BisA** could be used for selective photocleavage of circular single stranded DNA in the presence of supercoiled circular double stranded DNA.[140] Meanwhile, control experiments with **MonoA** (figure 2.1) showed equal cleavage of both single and double stranded DNA sequences. Soon after, the group showed that while **BisA** preferentially cleaved single stranded DNA over normal double stranded DNA, it could in fact be used to cleave duplex DNA containing abasic sites.[141] They showed that **BisA** formed a 1:1 complex with an abasic site containing duplex and that **BisA** stabilised this sequence ($\Delta T_m$ = +14°C). The binding site was confirmed by NMR studies showing that one acridine unit intercalated between the adjacent base pair, while the other acridine unit resided in the pocket created by the abasic site.[142] **BisA** is also able to destabilise duplex DNA, promoting the disassociation of the DNA into separate strands. The strong binding affinity of **BisA** for single stranded DNA then enables **BisA** to effectively compete with single strand binding (SSB) proteins.[143]

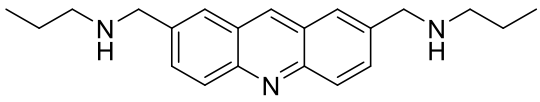



*Figure 2.1: Structure of the monoacridine **MonoA.***

Finally in their most recent work on **BisA**, Marie-Paule Teulade-Fichou and co-workers have investigated the interaction of **BisA** with mismatched base pairs in duplex DNA.[137,144,145] **BisA** was tested with 17 base pair duplexes containing a central TT, TC or TG mismatch and compared to the matched TA containing duplex. As expected, the introduction of a mismatched base pair destabilised the duplex and the effect was more pronounced with TT and TC mis-pairs, compared to a TG mis-pair. Addition of up to 2 equivalents of **BisA** stabilised the mismatched duplexes by up to 7.1°C, restoring the stability almost back to that of the matched sequence. Conversely **BisA** had no effect

on the stability of the matched duplex, indicating a selectivity for mismatch containing DNA. The mono-acridine control compound showed no such selectivity, imparting a slight stabilisation on both matched and mismatched duplexes ($\Delta T_m$ = 0.6 – 2.1°C).[144] By looking at the effect of **BisA** on permanganate oxidation of thymine in the mismatched duplexes, they also suggested that binding of **BisA** to the mismatched base pair results in base flipping of the thymine base out of the double strand into an extrahelical position, where it is more readily oxidised. This mimicked the effect of base modifying enzymes such as glycosylases and methyltransferases.[144] This was later confirmed by electrospray ionisation mass spectrometry studies (ESI-MS), which showed that **BisA** had a much higher affinity for mismatch sequences ($K_d$ = 0.5, 0.55 and 3.1 μM for TT, TC and TG respectively) over the matched sequence ($K_d$ = 10 μM).[137]

### 2.1.2 BisA stabilises i-motif and G-quadruplex structures

In the course of their investigations on the DNA binding properties of **BisA** Jean-Marie Lehn and co-workers examined the interaction of **BisA** with G-quadruplex and i-motif DNA structures, assessing its potential as an inhibitor of telomerase.[100] In collaboration with Jean-Louis Mergny, a pioneer in the use of FRET to study the i-motif,[98] they were the first group to identify an i-motif stabilising ligand using this technique. Initially they carried out FRET melting experiments with the human telomeric G-quadruplex sequence (F21GT: FAM-5'-d(GGGTTAGGGTTAGGGTTAGGG)-3'-TAMRA) and the complimentary human telomeric i-motif sequence (F21CT: FAM-d(CCCTAACCCTAACCCTAACCC)-TAMRA). Comparing **BisA** and **MonoA** they measured significant stabilisation of both G-quadruplex and i-motif ($\Delta T_m$ = +15°C and +33°C respectively) for **BisA** while **MonoA** showed no significant change in melting temperature for either structure.[100]

UV absorbance spectra for **BisA** showed a hypochromic affect at 360 nm upon addition of 21G (24%) or 21C (13%) which they took as indication of a binding event. The spectrum of **MonoA** also showed some hypochromism suggesting that although **MonoA** did not stabilise either G-quadruplex or i-motif structures, it may still weakly bind. This was confirmed by fluorescence titrations of **MonoA,** which indicated a weak binding affinity lower than $10^6$ $M^{-1}$ but was not determined exactly due to weak signal variations. The affinity of **BisA** for i-motif or G-quadruplex was not determined.

Equilibrium dialysis experiments were carried out to show the specificity of **BisA** for i-motif and G-quadruplex structures over other types of DNA and RNA structure. In this experiment, 19 different nucleic acid sequences/structures are placed in separate dialysis tubes inside a common ligand solution. This enables the ligand to move freely under the same buffer conditions and bind to multiple different structures. Once equilibrium has been reached after 24 hours, the concentration of bound ligand is measured by UV-Vis spectroscopy to give an idea of specificity.[146] Alberti and co-workers measured higher bound concentrations for **BisA** with 3 different G-quadruplex sequences and the i-motif forming sequence poly $d(C)_4$, in comparison to duplex DNA, Z-DNA, single stranded DNA, triplex DNA and triplex RNA. However, caution should be taken in interpreting these results as in order to have all structures present in the same conditions, the equilibrium dialysis experiments were carried out at pH 7.0. At this pH the i-motif is not expected to form so the measured bound concentration is almost certainly for the single strand rather than the i-motif structure.

The effect of **BisA** on duplex formation, as described in section 2.1.1, has previously been investigated but for a different sequence.[143] In this study the effect of **BisA** on the telomeric double stranded sequence was assessed. F21GT in its G-quadruplex structure was titrated with the complimentary sequence 21C and the increase in fluorescence of FAM upon G-quadruplex unfolding and strand hybridisation was monitored over time with and without **BisA**. These experiments showed that **BisA** is able to inhibit duplex formation by stabilising the G-quadruplex structure. Conversely, the control compound **MonoA** had no effect on duplex formation. Finally, the ability of **BisA** to inhibit telomerase was assessed in a modified TRAP assay.[147] The telomeric repeat amplification protocol (TRAP) assesses the ability of the telomerase enzyme to synthesise telomeric DNA by using the polymerase chain reaction (PCR) to amplify the telomeric product sequence.[148] Inhibition of telomerase can then be observed by a reduction in the PCR products however, the compound must also be assessed for inhibition of the Taq polymerase enzyme used in the PCR reaction as a control. Alberti and co-workers showed that **BisA** inhibited telomerase with an $IC_{50}$ of 0.75 μM, whilst **MonoA** did not ($IC_{50}$ >> 10 μM). **BisA** had a weaker inhibition of Taq polymerase ($IC_{50}$ = 2.6 μM) measured at four times the concentration of **BisA** needed for telomerase inhibition indicating that the observed result was not due to interaction with Taq polymerase.[100]

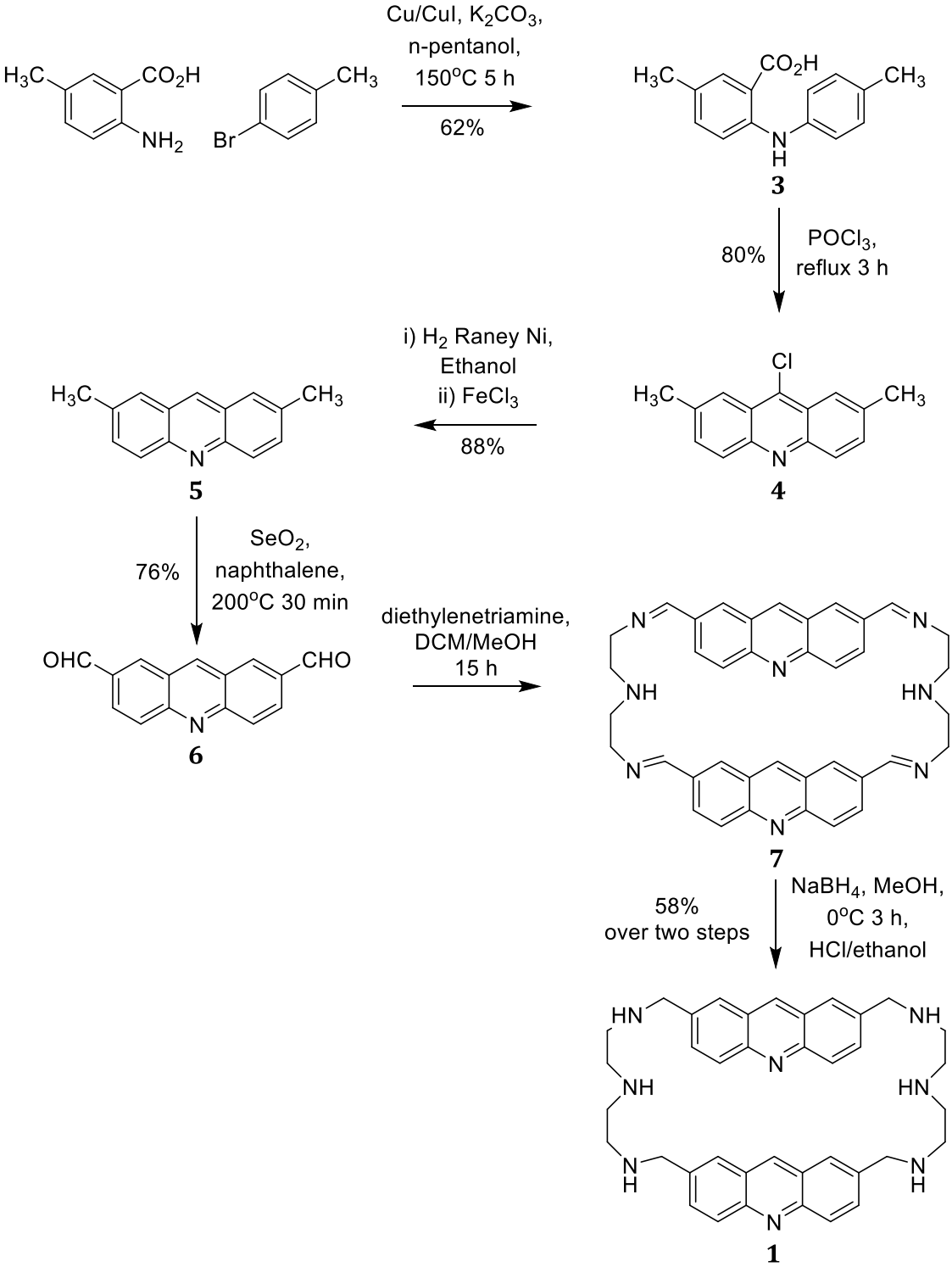
### 2.1.3 Aims and Objectives

Despite the i-motif field now rapidly growing, there are still very few i-motif ligands in the literature and there is no obvious trend as to what constitutes a good i-motif ligand. There have also been no other compounds assessed as i-motif binding ligands with a similar type of structure to **BisA**. In this chapter, **BisA** is synthesised with the aim of fully characterising in detail the effect of **BisA** on the i-motif structure.In previous work on **BisA**, other experiments such as fluorescence emission spectroscopy and polyacrylamide gel electrophoresis have been used to study the compound's interaction with hairpin DNA but not the i-motif, as they have been applied in this chapter. As a comparison the compound **MonoA**, which has been used as a control throughout the previous work on **BisA**, has been tested. This will indicate whether it is simply the acridine unit that has an effect or whether there is something in particular about the **BisA** structure that enables its interaction with i-motif DNA.

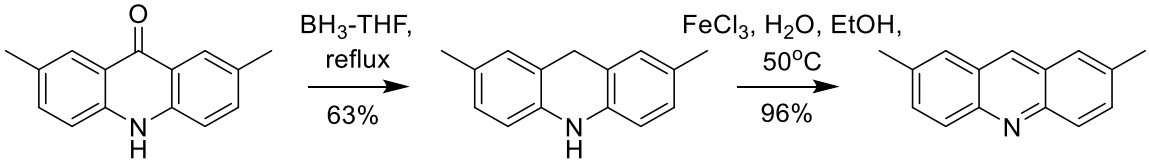
## 2.2 Synthesis of BisA and MonoA

As mentioned in section 2.1.1, **BisA** was first synthesised by Marie-Paule Teulade-Fichou in 1995 (Scheme 2.1).[136] This procedure was used as a starting point for the synthesis of **BisA**. 4,4'-Dimethyldiphenyl-2-carboxylic acid **3** was prepared from 2-amino-5-methylbenzoic acid and para-bromotoluene which are both commercially available. In Teulade-Fichou's procedure, this was then cyclised to 9-chloro-2,7-dimethylacridine **4** and dechlorinated by hydrogenation to give 2,7-dimethylacridine **5**. However, attempts to follow this route failed at this stage as hydrogenation of **4** did not produce any product. This is because the cyclisation reaction did not give the 9-chloro-2,7-dimethylacridine **4** but instead produced the 2,7-dimethylacridone **8**. It has previously been shown that acridones are effectively reduced using borane-THF complex to give the acridane and then re-oxidised to provide the desired acridine.[149] Therefore, the 4,4'-dimethyldiphenyl-2-carboxylic acid **3** was cyclised by heating to reflux in concentrated sulfuric acid and the resultant novel 2,7-dimethylacridone **8** was reduced to the novel 2,7-dimethylacridane **9** and re-oxidised to the desired 2,7-dimethylacridine **5** (scheme 2.2). The methyl groups were oxidised to aldehydes **6** and

the final product **1** was formed by reductive amination with the diethylenetriamine linker *via* the imine **7** following Teulade-Fichou's procedure (scheme 2.1).




*Scheme 2.1: The synthesis of **BisA** outlined by Marie-Paule Teulade-Fichou.*[136]

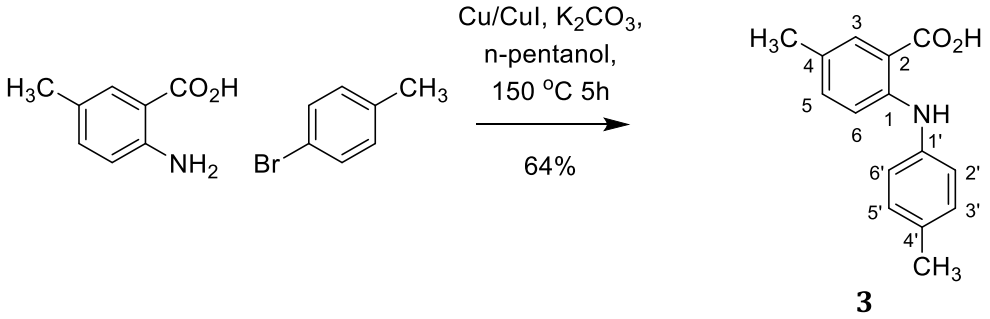



**8** **9** **5**

*Scheme 2.2: Formation of acridines from acridones using borane-THF complex.*[149]

### 2.2.1 Synthesis of 4,4'-dimethyldiphenyl-2-carboxylic acid (3)

The first step in the synthesis of **BisA** was carried out using a copper catalysed Ullmann condensation reaction to couple 2-amino-5-methylbenzoic acid to *p*-bromotoluene to produce 4,4'-dimethyldiphenyl-2-carboxylic acid **3**.[136] The starting materials were heated to reflux at 150°C in n-pentanol with 1 equivalent of potassium carbonate base and catalytic amounts of copper and copper iodide. The crude product was purified by crystallisation from ethanol giving 64% yield from 4 crops and characterised by $^1$H NMR.



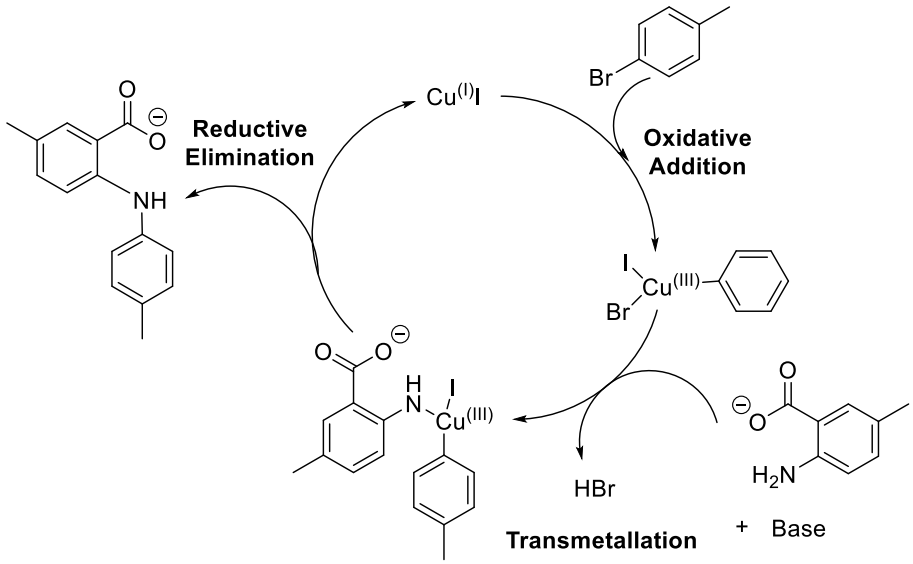

**3**

*Scheme 2.3: The synthesis of 4,4'-dimethyldiphenyl-2-carboxylic acid (**3**).*

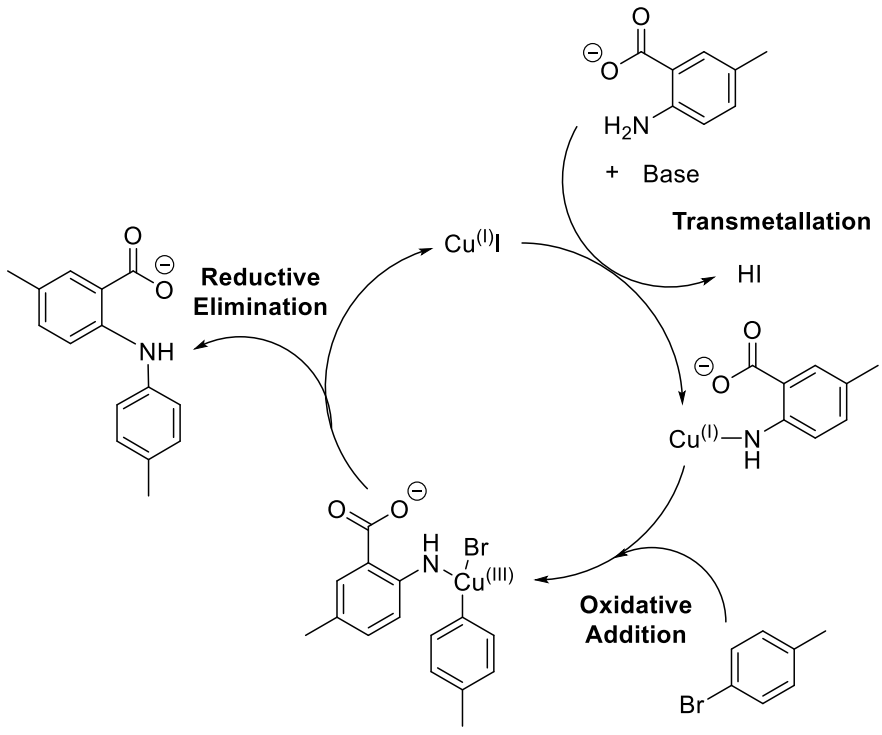
The $^1$H NMR was carried out in deuterated chloroform and showed a broad singlet at 9.05 ppm corresponding to the NH proton. The seven aromatic protons were shown in 2 multiplet peaks, the most downfield at 7.82 ppm representing one proton at position 3, with the multiplet at 7.12 ppm representing the other 6 protons. The two methyl groups are represented by separate singlet peaks at 2.35 and 2.26 ppm. This corresponded satisfactorily with the published NMR data.[136]

The exact mechanism of the Ullmann condensation reaction still has not been fully elucidated. In their review in 2010, Elena Sperotto and colleagues discuss the different proposals in the literature.[150] Generally, it is thought that the reaction proceeds *via* one of two routes which differ in the order of when oxidative addition and transmetallation occur (illustrated in schemes 2.4 and 2.5). In the first pathway (scheme 2.4), oxidative

addition of the copper catalyst to the C-Br bond of p-bromotoluene occurs first. The amine group, is then able to substitute for the bromine at the copper catalyst, releasing hydrogen bromide. Finally a reductive elimination completes the coupling of the nitrogen to the adjacent ring and the free acid is obtained in aqueous work up. In the second pathway (scheme 2.5), the copper catalyst first reacts with the nucleophilic amine before oxidative addition occurs with *p*-bromotoluene and finally reductive elimination regenerates the catalyst.



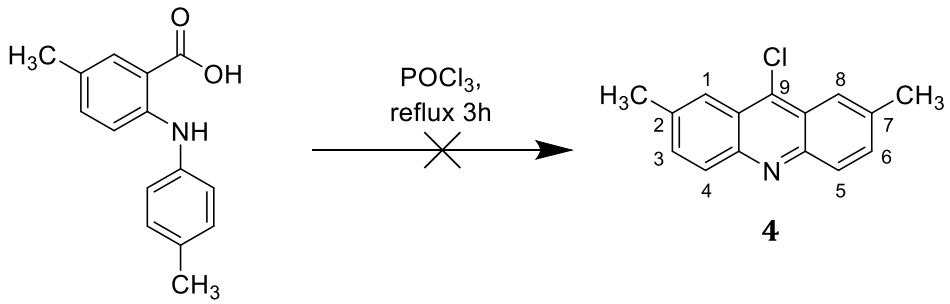

*Scheme 2.4: The possible mechanism of the Ullmann condensation reaction used in the first step of the synthesis of* ***BisA***.[150]




*Scheme 2.5: The alternative mechanism of the Ullmann condensation reaction for step 1.*

Other mechanisms have also been proposed using $Cu^{(0)}$ that is oxidised to $Cu^{(II)}$, as well as mechanisms using σ-bond metathesis or π-complexation where the oxidation state of copper does not change. A radical mechanism was previously disfavoured as early "radical-clock" experiments did not give the expected product. However, recent computational studies by Buchwald and Houk have suggested that a single electron transfer mechanism may still occur.[151]

### 2.2.2 Attempted synthesis of 9-chloro-2,7-dimethylacridine (4) and 2,7-dimethylacridine (5).



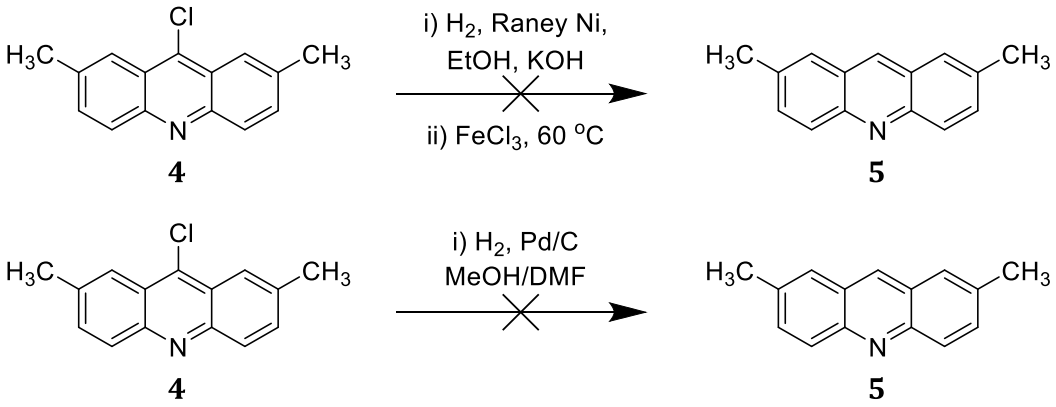

**3** **4**

*Scheme 2.6: The $POCl_3$ synthesis of 9-chloro-2,7-dimethylacridine (**4**).*

According to Teulade-Fichou's synthesis, 4,4'-dimethyldiphenyl-2-carboxylic acid **3** was cyclised using phosphorous oxychloride to give the 9-chloro-2,7-dimethylacridine **4** (Scheme 2.6). 4,4'-dimethyldiphenyl-2-carboxylic acid **3** was dissolved in neat phosphorous oxychloride and heated under reflux conditions for 3 hours before the reaction was concentrated *in vacuo*, washed with 15% ammonium hydroxide solution, extracted with DCM, washed with water and dried over $MgSO_4$. The product was recrystallised from ethanol and characterised by $^1H$ NMR in $d^6$-DMSO. The $^1H$ NMR spectrum showed peaks at 8.01 ppm, 7.55 ppm and 7.46 ppm representing the aromatic protons and the methyl protons were found at 2.41 ppm.

The shifts of these peaks did not match the published example (8.09, 8.07, 7.71 and 2.58 ppm), however there were the right number of peaks representing the correct number of protons with the same multiplicities. Therefore the product was taken forward to the next step of hydrogenation to give 2,7-dimethylacridine **5**. This was done by stirring **4** with potassium hydroxide and raney nickel in ethanol under an atmosphere of hydrogen. After this, the catalyst was removed, the solution was heated

to 60 °C and $FeCl_3$ was added before being allowed to cool. The crude mixture was washed with 15% ammonium hydroxide and the volume reduced to dryness *in vacuo*. $^1H$ NMR indicated that unfortunately no reaction had occurred as there was no new proton at position 9. In a second attempt to hydrogenate **4**, Palladium on carbon was added as an alternative catalyst and the mixture was stirred under hydrogen overnight but this also showed no reaction by $^1H$ NMR.




*Scheme 2.7: Attempted hydrogenations of 9-chloro-2,7-dimethylacridine (**4**).*

The mass spectrum of **4** after the attempted hydrogenation, gave a molecular ion peak at m/z = 224.0 which was indicative of the 2,7-dimethylacridone compound **8** rather than the desired product **5** (m/z = 208). The NMR of the synthesis of **4** using $POCl_3$ was re-examined in comparison to compound **8**, which showed that the 2,7-dimethylacridone **8** was the actual product of the reaction. This may be as a result of the purification procedure as, after a second attempt to produce **4** by this method, the crude NMR matched that in the literature. However, the NMR after purification by crystallisation from ethanol matched that of compound **8**.

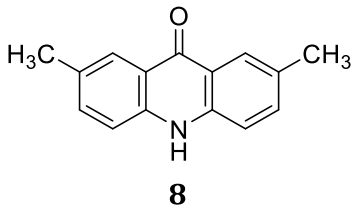
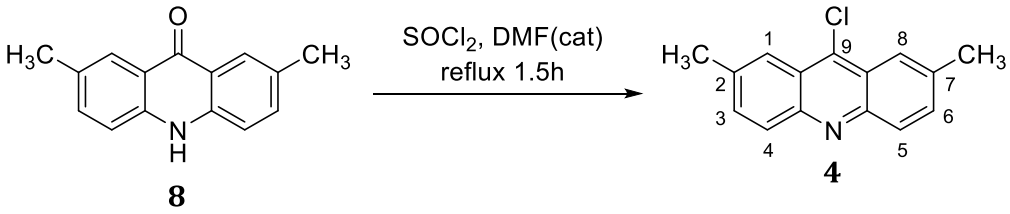



*Figure 2.2: The structure of 2,7-dimethylacridone (**8**).*



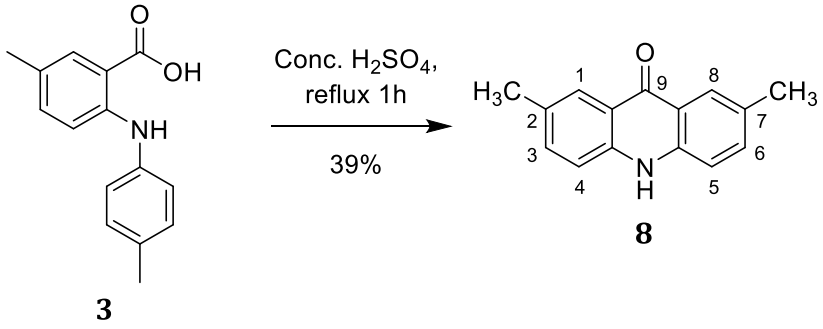

**8** **4**

*Scheme 2.8: The $SOCl_2$ synthesis of 9-chloro-2,7-dimethylacridine (**4**).*

The 2,7-dimethylacridone **8** was successfully converted into 9-chloro-2,7-dimethylacridine **4** using thionyl chloride as outlined by Atwell and co-workers (Scheme 2.8).[152] The acridone **8** was dissolved in thionyl chloride with a catalytic amount of DMF and heated to reflux at 90°C for 1.5 hours. The excess thionyl chloride was removed by trituration with benzene and the crude product was purified by flash coloumn chromatography in DCM before being characterised by mass spectrometry and $^{1}H$ NMR in deuterated chloroform. The mass spectrum gave a peak at m/z = 242.0734, corresponding to the M + $H^+$ ion for the $^{35}Cl$ isotope and a corresponding peak with one third the intensity was observed at m/z = 244.0704 corresponding to the $^{37}Cl$ isotope. The $^{1}H$ NMR gave a doublet at 8.02 ppm corresponding to the protons at position 4 and 5, a singlet at 7.97 ppm corresponding to the protons at position 1 and 8 and a doublet of doublets at 7.50 ppm corresponding to the protons at position 3 and 6. These protons showed a reciprocated 3 bond coupling to the protons at position 4 and 5 ($J$ = 8.8 Hz) and a long range coupling to the protons at position 1 and 8 ($J$ = 1.6 Hz).

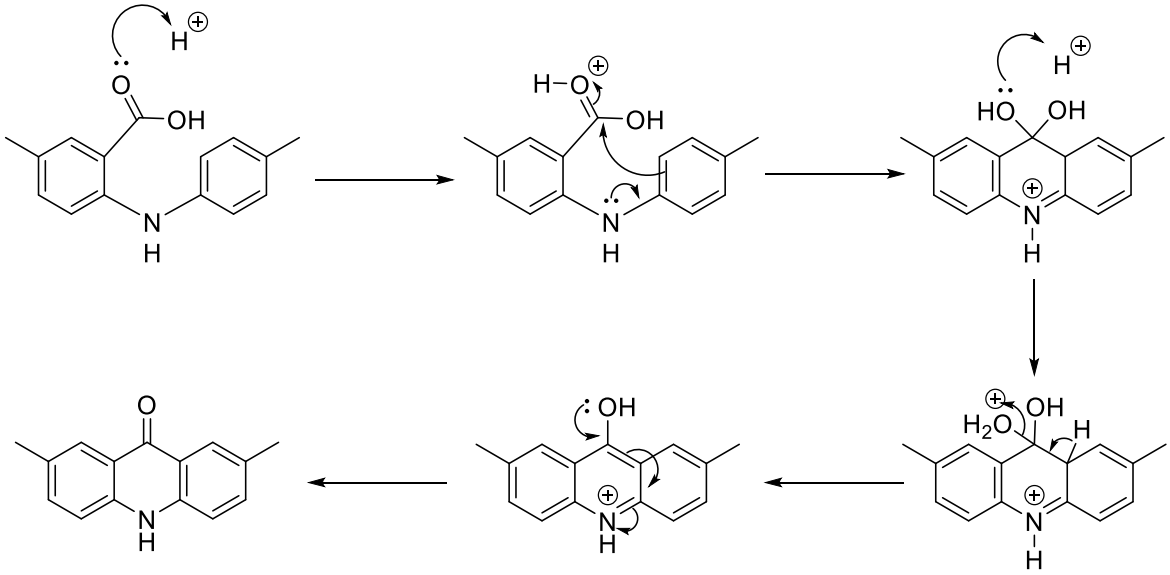
With 9-chloro-2,7-dimethylacridine to hand, it was subsequently submitted to hydrogenation conditions to convert it to 2,7-dimethylacridine **5**. Unfortunately neither raney nickel nor palladium on carbon conditions succeeded as could be seen by a $^{13}C$ NMR distortionless enhancement by polarisation transfer (DEPT) experiment, which reduces the $^{13}C$ NMR spectrum to just carbons that are attached to protons.[153] This did not show the required number of carbon peaks for the desired product **5**. The proton NMR did show some changes but did not show the presence of an additional proton for the 9 position. Further attempts to hydrogenate **4** were not made, as an alternative route to **5** was used that is described in sections 2.2.4 and 2.2.5.

### 2.2.3 Synthesis of 9-oxo-9,10-dihydro-2,7-dimethylacridone (8).




*Scheme 2.9: The synthesis of 9-oxo-9,10-dihydro-2,7-dimethylacridone (**8**).*

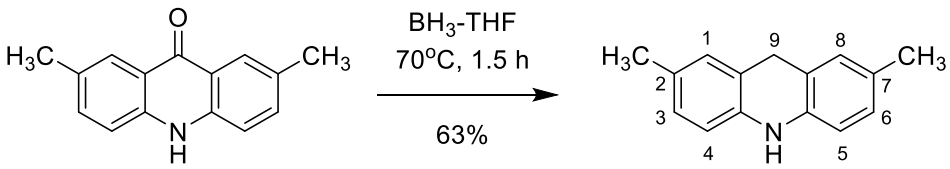
4,4'-dimethyldiphenyl-2-carboxylic acid **3** was cyclised to give **8** by heating at reflux in sulfuric acid for 1 hour (scheme 2.9).[154] The solid precipitate was filtered and recrystallised from DMF to give **8** as a yellow solid. The product is a novel compound and was fully characterised by melting point analysis, Infra-Red spectroscopy, mass spectrometry, $^{1}H$ and $^{13}C$ NMR as well as 2D correlated spectroscopy (COSY) NMR. The mechanism begins with protonation of the carbonyl oxygen (scheme 2.10), which has the effect of making the carbonyl carbon more electrophilic. The adjacent π-system is able to attack the carbonyl driven by the conjugation of the nitrogen lone pair into the ring in an enamine type reaction. The resultant unstable acetal, loses water to give the enol, which then tautomerises to give the desired acridone.




*Scheme 2.10: Mechanism of the acid catalysed cyclisation to form the acridone.*

The melting point of **8** was >355 °C similar to other acridones in the literature.[152] The IR spectrum showed a peak for the carbonyl bond at 1636 $cm^{-1}$ with peaks at 2784 and 1479 $cm^{-1}$ representing C-H and C=C respectively. NMR experiments were carried out in $d^6$-DMSO, the most downfield peak at 11.59 ppm is a singlet representing the N-H proton. The six aromatic protons are represented by 3 peaks, an apparent singlet at 8.02 ppm for the protons at position 1 and 8. A doublet of doublets at 7.55 ppm representing the protons at position 3 and 6. This shows a 3-bond coupling to the adjacent protons at position 4 and 5 of $J$ = 8.5 Hz and a long range 4-bond 'w-coupling' between position 3 and 6 and 1 and 8 of $J$ = 2.0 Hz. The 3rd aromatic peak at 7.44 ppm is a doublet for the protons at position 4 and 5, reciprocating the 3 bond coupling to those protons at 3 and 6 of $J$ = 8.5 Hz. Finally the methyl groups appear as a singlet at 2.42 ppm. The $^{13}C$ NMR consisted of 8 peaks, the most downfield at 176.31 ppm representing the carbonyl carbon of the acridone. Due to the symmetry of the molecule, the 12 aromatic carbons were represented by 6 peaks between 110 and 140 ppm while the methyl carbons were represented by the most up-field peak at 20.56 ppm. An accurate mass was obtained using atmospheric-pressure chemical ionisation (APCI) and was found at m/z = 224.1067 corresponding to the $M+H^+$ ion, calculated to be m/z = 224.1070.

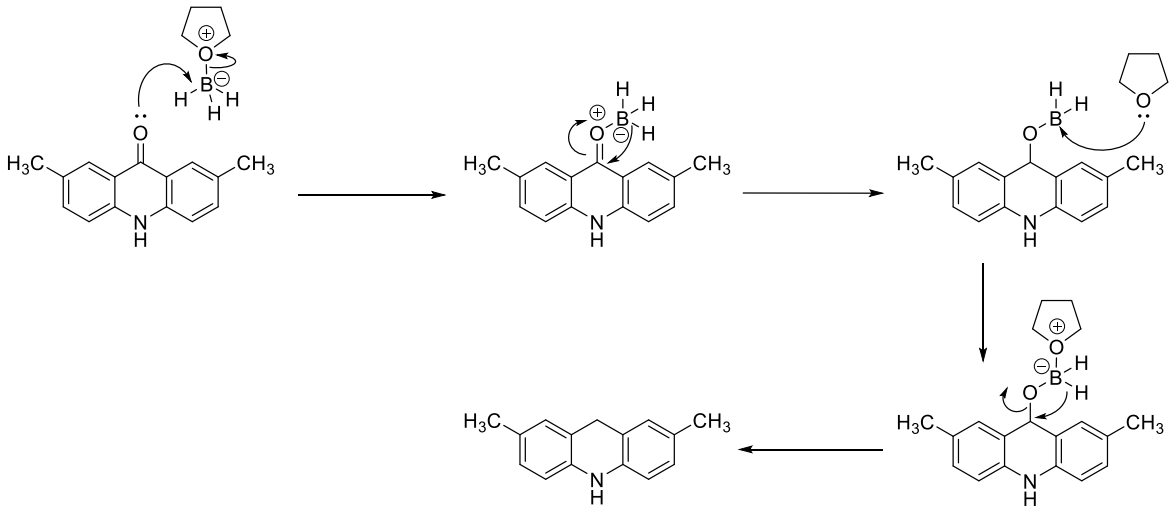
### 2.2.4 Synthesis of 9,10-dihydro-2,7-dimethylacridine (9)




**8** **9**

*Scheme 2.11: The synthesis of 9,10-dihydro-2,7-dimethylacridine (**9**).*

The acridone **8** was reduced to 9,10-dihydro-2,7-dimethylacridine **9** by heating to reflux at 70°C with a slight excess of 1 M borane-THF complex in THF for 1 hour 30 minutes, following a procedure outlined by Nicolas Desbois.[149] The mechanism of the reduction is shown in scheme 2.12. The unstable crude product was purified by short flash column chromatography (DCM) and characterised by melting point, IR, $^1H$, $^{13}C$ and 2D COSY NMR before being used in the next step.



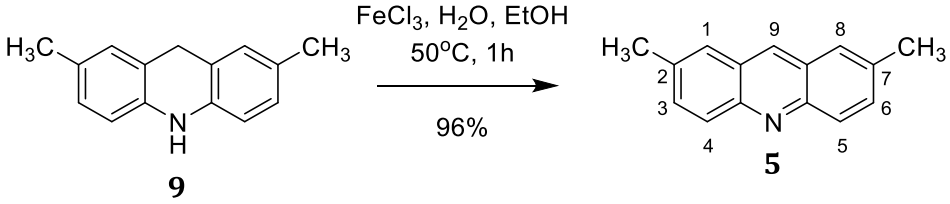
*Scheme 2.12: The mechanism of reduction of the acridone by borane-THF complex.*

The melting point was measured to be 218°C which is much lower than the acridone pre-cursor. This is consistent with the loss of electron density in the π-system from losing the carbonyl group which weakens the possible π-π stacking interactions. The IR spectrum showed peaks at 3400 $cm^{-1}$ characteristic of the N-H stretch, the rest of the spectrum consisted of peaks at 2920 $cm^{-1}$ and 1610 $cm^{-1}$ for C-H and C=C respectively and the absence of a C=O peak. The compound was more soluble than the acridone precursor as the reduction of the carbonyl to a $sp^3$ carbon diminishes the π-π stacking capability and the NMR spectra could be obtained in deuterated chloroform. The compound has a line of symmetry through the central carbon and nitrogen so the protons at positions 1, 3 and 4 are identical to the protons at positions 8, 6 and 5, as are the two methyl groups. The aromatic protons were shifted significantly up-field with the protons at position 1 and 8 appearing as a singlet at 6.91 ppm, the protons *meta* to the amine at position 3 and 6 appear as a doublet at 6.87 ppm and the protons *ortho* to the amine at position 4 and 5 appear as a doublet at 6.55 ppm. These two sets of doublets show coupling to one another ($J$ = 7.9 Hz) but it was not possible to see the expected long range coupling between position 3 and 6 and 1 and 8 under the experimental conditions. There is a broad singlet at 5.79 ppm representing the NH proton, the new $CH_2$ protons at position 9 appear as a singlet at 3.98 ppm and the two methyl groups appear as a singlet at 2.26 ppm.

Due to the symmetry of the compound, the $^{13}C$ NMR has 8 peaks, the 6 aromatic carbons appear between 110 and 140 ppm, the reduced carbon at position 9 appears at 31.38 ppm and the methyl groups appear at 20.59 ppm. The accurate mass was measured

using APCI and found an m/z = 210.1277 which corresponds to $M+H^+$. This corresponds to loss of oxygen and gain of 2 hydrogens consistent with the expected structure.

### 2.2.5 Synthesis of 2,7-dimethylacridine (5).

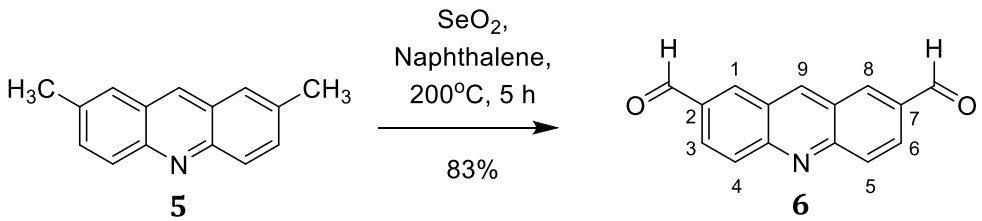



**9** **5**

*Scheme 2.13: The synthesis of 2,7-dimethylacridine (**5**).*

The acridane **9** synthesised in section 2.2.4 is unstable and readily loses a proton in the presence of a Lewis acid in order to form the more stable aromatic acridine. Therefore, following the procedure outlined by Desbois,[149] the acridane was stirred with an excess of iron trichloride in a mixture of water and ethanol at 50°C for 1 hour. This afforded the acridine **5** in a 96% yield and was of sufficient purity to use directly in the next step. The acridine was characterised using $^1$H NMR. This showed a singlet at 8.55 ppm for the proton at position 9, the two protons *ortho* to the nitrogen group at position 4 and 5 appear as a doublet at 8.11 ppm. The protons at position 1 and 8 appear as a singlet at 7.73 ppm and the protons at position 3 and 6, *meta* to the nitrogen, appear as a doublet of doublets at 7.59 ppm. Finally the methyl group protons appear as a singlet at 2.58 ppm. The protons at position 4 and 5 show a 3 bond coupling to those at position 3 and 6 ($J$ = 8.9 Hz) as well as a long range coupling to the protons at position 1 and 8 ($J$ = 1.9 Hz). This is in agreement with previously published spectra.[136]

### 2.2.6 Synthesis of 2,7-acridinedicarboxaldehyde (6).

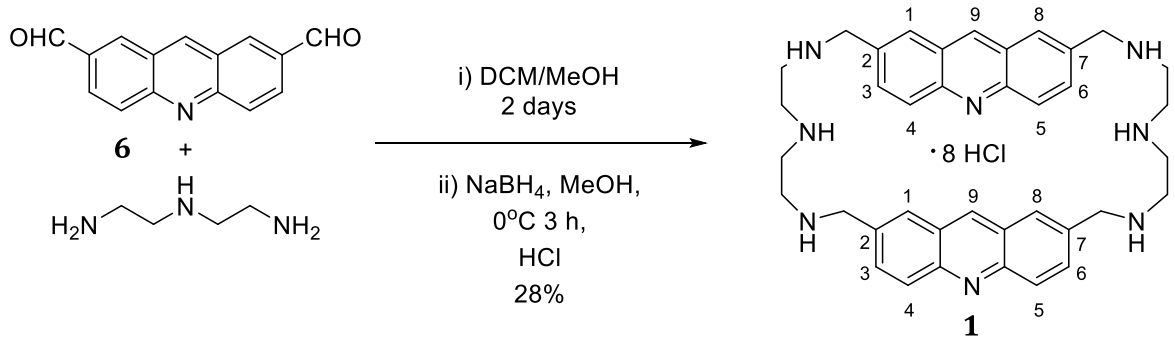



**5** **6**

*Scheme 2.14: Synthesis of 2,7-acridinedicarboxaldehyde (**6**).*

The 2,7-dimethylacridine **5** was converted to the dicarboxaldehyde **6** according to Teulade-Fichou's procedure.[136] The acridine was oxidised using selenium dioxide at 200°C with molten naphthalene as solvent. The crude product was purified by washing with hexane and the remaining naphthalene was removed by sublimation. The product was characterised by $^1$H NMR. This showed a downfield peak at 10.27 ppm for the aldehyde protons, singlets at 9.17 ppm for the proton at position 9 and at 8.60 ppm for the two protons at position 1 and 8 and a multiplet at 8.35 ppm for the remaining 4 aromatic protons. The presence of the aldehyde protons and the absence of a peak at 2.58 ppm for the methyl protons clearly demonstrates the formation of the product. This is in agreement with the spectrum in the literature.[136]

### 2.2.7 Synthesis of 2,5,8,21,24,27-Hexaaza[9,9](2,7)acridinophane (BisA)

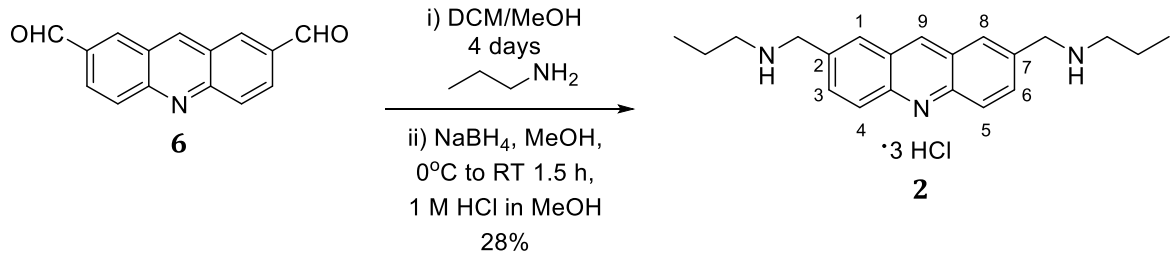



*Scheme 2.15: The synthesis of 2,5,8,21,24,27-Hexaaza[9,9](2,7)acridinophane (**BisA**).*

The final product was synthesised *via* reductive amination with diethylenetriamine according to Teulade-Fichou's procedure.[136] Diethylenetriamine was added slowly drop-wise to the acridinedicarboxaldehyde **6** in DCM and methanol and stirred at room temperature for 2 days. The resulting imine **7** (scheme 2.1) was reduced to the amine (**1**) by careful addition of sodium borohydride at 0°C. The crude product was purified by repeated washes with different organic solvents and then acidified with hydrochloric acid to give the 8 × HCl salt in 28% yield. The pure product was characterised by $^1$H NMR in deuterated water. The most downfield peak at 9.60 ppm represents the aromatic proton at position 9. The protons at positions 1 and 8 appear as a singlet at 8.39 ppm, those at positions 3, 4, 5 and 6 appear as two doublets at 8.21 and 8.09 ppm, showing a 3 bond coupling between them ($J$ = 9.1 Hz). The 8 protons between the acridine units and the amine linker appear as a singlet at 4.49 ppm. Finally the protons for the diethylenetriamine linker appear as two doublets at 3.40 and 3.31

ppm, showing a 3 bond coupling between them (*J* = 6.0 Hz). This is in agreement with the published spectra.[136]

### 2.2.8 Synthesis of 2,7-(Di-n-propylaminomethyl)acridine (MonoA)




*Scheme 2.16: The synthesis of 2,7-(Di-n-propylaminomethyl)acridine (**MonoA**).*

Similarly to **BisA**, the control compound **MonoA** was synthesised *via* reductive amination with n-propylamine according to Teulade-Fichou's procedure.[136] n-Propylamine was added drop-wise to the acridinedicarboxaldehyde (**6**) and stirred at room temperature for 2 days before being heated to reflux for a further 2 days to push the reaction towards completion. The imine product was reduced by stirring with sodium borohydride at 0°C for 30 minutes and then at room temperature for a further hour. The crude product was prepared in the salt form by washing with 1 M HCl in methanol and then purified by dissolving in water and washing repeatedly with DCM, ether, ethylacetate and hexane before being neutralised with saturated sodium hydrogen carbonate and re-extracted with DCM to give the pure product which was re-made into the salt with 1 M HCl in methanol in 28% yield. The pure product was characterised by $^1$H NMR in deuterated water. The aromatic protons were found between 9.6 and 8.1 ppm with the proton at the 9 position being most downfield at 9.59 ppm. The protons at position 1 and 8 appear as a singlet at 8.42 ppm, the protons at position 4 and 5 appear as a doublet at 8.28 ppm, while the protons at position 3 and 6 appear as a doublet of doublets at 8.12 ppm. These protons show a 3 bond coupling to the 4 and 5 position of *J* = 9.0 Hz and a long range coupling to the 1 and 8 position of *J* = 1.8 Hz. The methylene protons between the amine and acridine units appear as a singlet at 4.48 ppm. The n-propyl protons appear further up-field in order of their distance from the amine. $NH_2\mathbf{CH_2}CH_2CH_3$ appears as a triplet (*J* = 7.8 Hz) at 3.07 ppm. $NH_2CH_2\mathbf{CH_2}CH_3$ appears as a sextet (*J* = 7.7 Hz) at 1.69 ppm. $NH_2CH_2CH_2\mathbf{CH_3}$ appears

as a triplet ($J$ = 7.5 Hz) at 0.91 ppm. The characterisation data compares well with the published spectra.[136]

## 2.3 Biophysical Characterisation of BisA with i-Motif DNA

### 2.3.1 Buffer conditions and the effect of BisA on pH.

As described in the introduction to this thesis, the reliance on hemiprotonated cytosine-cytosine$^+$ base pairs to form the i-motif structure means that the stability of the structure is highly pH dependent. The most stable i-motifs are around the p$K_a$ of cytosine (4.6). Therefore, when assessing potential i-motif ligands it is important that the buffer is of a sufficient concentration to maintain the pH with addition of ligand. A decrease in pH could be misinterpreted as a ligand induced stabilisation while an increase in pH could be perceived as a ligand induced destabilisation. The biophysical characterisation of **BisA** was carried out in 3 different pHs: firstly, at pH 6.8, this is the pH at which Alberti and co-workers carried out their experiments but as this is only marginally acidic, the i-motif is not very stable. Secondly at pH 5.5 where the i-motif is more stable and finally at pH 7.4, physiological pH, where the telomeric i-motif is not normally able to form and exists as a single strand in a random coil structure. Sodium cacodylate is the preferred buffer as its p$K_a$ is independent of temperature within the temperature range of our experiments and has been used previously for i-motif melting studies.[93,98] The change in pH of each buffer with the addition of 10 μM and 100 μM **BisA** was measured (Table 2.1). As can be seen, the change in pH of each buffer is minimal and actually increases slightly in the case of 10 μM **BisA** at pH 7.4. Due to the potential for change in pH, the concentration of **BisA** did not exceed 100 μM in any experiment and was kept well below 100 μM for the majority of experiments.

| Buffer | ΔpH | |
|---|---|---|
| | 10 μM **BisA** | 100 μM **BisA** |
| pH 5.5 | -0.02 | 0.07 |
| pH 6.8 | 0.01 | 0.09 |
| pH 7.4 | -0.10 | 0.13 |

*Table 2.1: Change in pH upon addition of **BisA** to different pH buffers. pH 5.5 and pH 7.4 buffer = 10 mM sodium cacodylate 100 mM NaCl, pH 6.8 buffer = 10 mM sodium cacodylate. Error = ± 0.01.*

The buffers used throughout the experiments contain 10 mM sodium cacodylate and different amounts of salt, normally 100 mM sodium chloride (NaCl) is added to replicate a more physiologically relevant ion concentration. However it was not entirely clear whether Alberti and co-workers used 100 mM NaCl in their FRET melting experiments at pH 6.8 and the addition of salt can have a destabilising effect upon the i-motif structure.[51] As described in section 1.4.5 circular dichroism is an optical technique that gives a different spectrum dependent on the type of DNA structure present in solution.[102] i-Motif DNA gives a characteristic spectrum with a positive peak at 288 nm and a negative peak at 255 nm this can be seen at pH 5.5 in figure 2.3. Single stranded DNA gives a positive peak at 275 nm and a negative peak at 250 nm (figure 2.3 pH 7.4). As the i-motif is already less stable at pH 6.8, it can be seen in figure 2.3

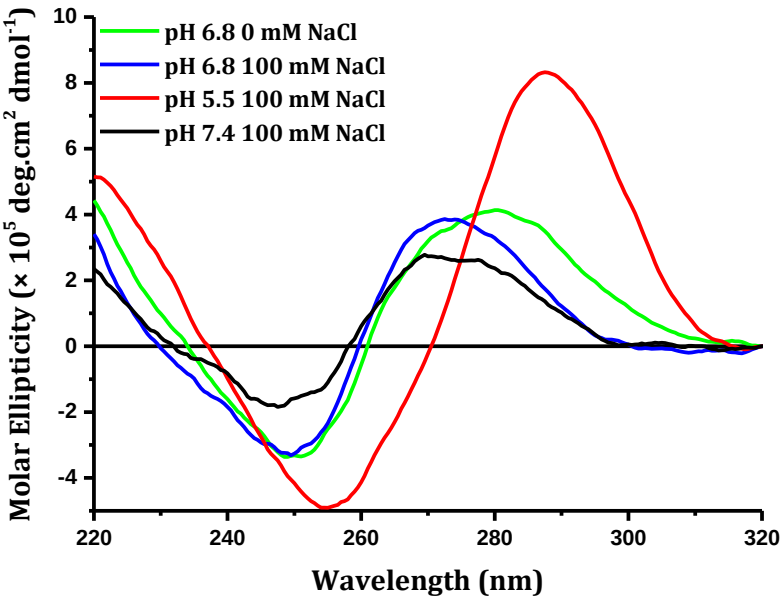



*Figure 2.3: CD spectra of the human telomeric i-motif sequence in different buffer conditions. Buffer = 10 mM sodium cacodylate with 0 or 100 mM NaCl.*

that at pH 6.8 with 100 mM NaCl the spectrum resembles that of ssDNA with peaks at 275 and 250 nm showing that no i-motif was detected by CD. Whereas at pH 6.8 with 0 mM NaCl there is a shift in the positive peak towards 288 nm, indicating a small amount of i-motif may be present. Therefore experiments at pH 6.8 were carried out in 10 mM sodium cacodylate with no additional salt whilst experiments at pH 5.5 and 7.4 still contained 100 mM NaCl.

Although **BisA** was synthesised as the 8 times hydrochloride salt, the p$K_a$'s of the nitrogens are different and therefore **BisA** will have a different charge depending on the buffer conditions. Marie-Paule Teulade-Fichou and colleagues estimated the p$K_a$ of the different nitrogens in **BisA** based on previous experiments on **MonoA** and a related naphthalene macrocycle.[136] Their experiments suggest that the central amine of the linker has a lower p$K_a$ of 2.5-3, whereas the amines at the ends of the linkers have higher p$K_a$'s of 7.4-10 (figure 2.4). The acridine nitrogen of **MonoA** has a p$K_a$ of 3.65 and the acridine nitrogens of **BisA** are likely to be slightly lower due to the polycationic environment of the cyclic compound. This indicates that **BisA** is likely to have a +4 charge in each of the 3 buffer conditions and **MonoA** will have a +2 charge. This is important because the positive charge can affect the compounds affinity for the i-motif through electrostatic interactions and many of the existing i-motif ligands discussed in section 1.5 carry a positive charge.

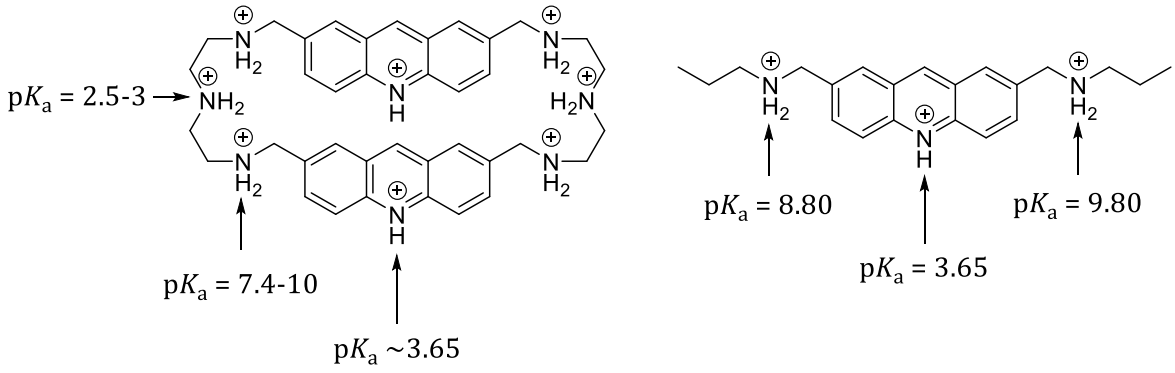



*Figure 2.4: Charges of **BisA** and **MonoA** in experimental buffer conditions. p$K_a$'s of **BisA** estimated from a naphthalene analogue and p$K_a$'s of **MonoA** calculated from pHmetric titration by Marie-Paule Teulade-Fichou and co-workers.*[136]

### 2.3.2 FRET melting studies

With **BisA** in hand, the FRET melting study of Teulade-Fichou and co-workers was first repeated, to see whether the conditions of the experiments gave similar results. **BisA** was tested at a concentration of 1 μM against a 200 nM solution of fluorescently labelled oligonucleotide from the human telomeric region which is able to form an i-motif. (hTeloC$_{FRET}$: 5'-FAM-d(TAACCCTAACCCTAACCCTAACCC)-TAMRA-3'). The samples were prepared in 10 mM sodium cacodylate buffer at pH 6.8 to be consistent with the previously published examples.[136] In their initial experiments, an increase in $T_m$ of +33°C was measured for **BisA** (1 μM) from normalised melting curves. The melting temperature is defined as the point at which half the sample is unfolded and half is folded.[91] In this case, normalisation assumes that at the lowest fluorescence intensity the DNA is completely unfolded and at the maximum fluorescence intensity the DNA is completely folded. Therefore, the data is converted to a graph of fraction folded between 0 and 1 versus temperature and the melting temperature is taken at fraction folded = 0.5. On repeating this procedure an even larger stabilisation was observed of +53°C (figure 2.5).

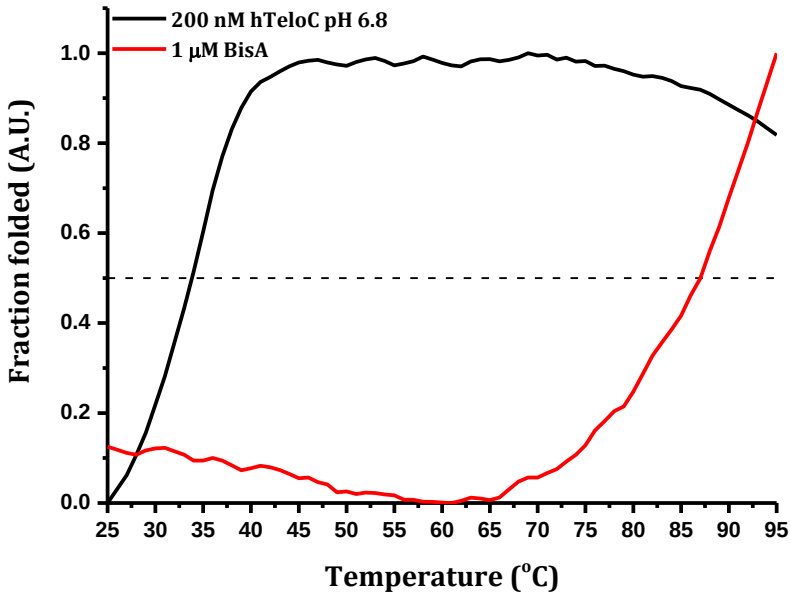



*Figure 2.5: Normalised FRET melting curves showing the fraction folded of human telomeric i-motif, with and without 1 μM **BisA** at pH 6.8. [DNA] = 200 nM, buffer = 10 mM sodium cacodylate.*

However, looking at the raw melting curves (figure 2.6) it can be seen that the addition of 1 μM **BisA** significantly quenches the fluorescence of FAM, only starting to increase towards 90°C and there is not a complete transition observed under the conditions of the experiment. This may be because **BisA** interacts and stabilises the structure to such an extent that the i-motif does not melt in this temperature range, alternatively this may be due to precipitation or aggregation of the DNA or quenching of the fluorescence due to an interaction between the ligand and the fluorophores. It can also be seen that the starting level of the fluorescence in the absence of **BisA** is quite high and increases immediately with temperature. This suggests that the DNA is not fully folded under these conditions, consistent with the low stability of the i-motif at pH 6.8.

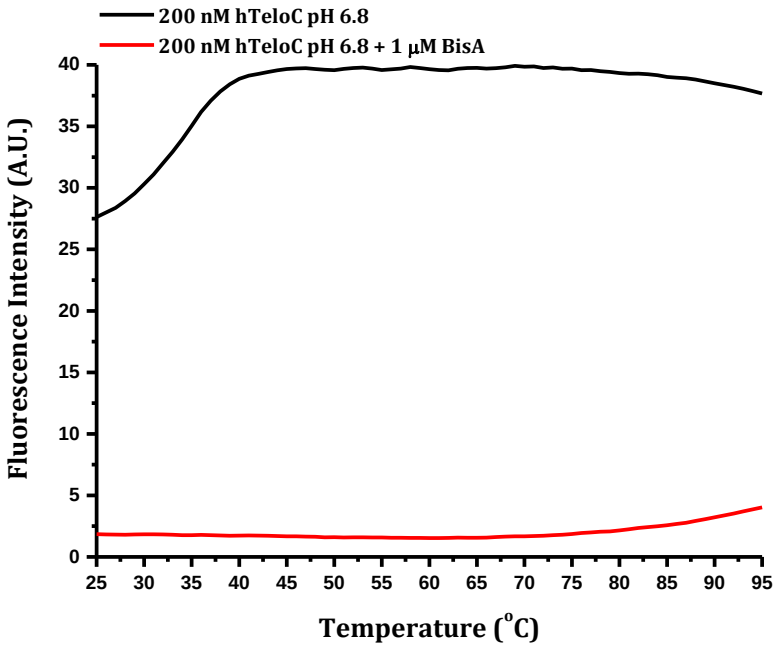



*Figure 2.6: Raw FRET melting curves for human telomeric i-motif, with and without 1 μM **BisA** at pH 6.8. [DNA] = 200 nM, buffer = 10 mM sodium cacodylate.*

To investigate this further, a range of concentrations between 0 and 1 μM were tested (figure 2.7). Here it can be seen that with increasing concentration, there is an overall decrease in the fluorescence signal across the temperature range. As an alternative to normalising the spectra, the first derivative may be taken as a measure of the melting temperature. Here the minimum point in the curve corresponds to the inflection point of the melting curve (figure 2.8).[91] This method can only be used for intramolecular systems and still may not equal the precise melting temperature. A third method of melting point determination may also be used called the "base line" method. Here lines

are drawn tangential to the low fluorescence baseline and the high fluorescence base line. A line is then draw bisecting the two and where this crosses the data is then taken as the melting temperature.[91] However, this method is subjective and it is not a good choice for melting curves which are not sigmoidal such as in figure 2.7 because one or more of the baselines cannot be accurately determined. Using the first derivative method, it can be seen that with increasing concentration of **BisA**, the minimum point is actually shifting towards lower temperatures, suggesting that **BisA** may actually be destabilising the i-motif at this pH (table 2.2).

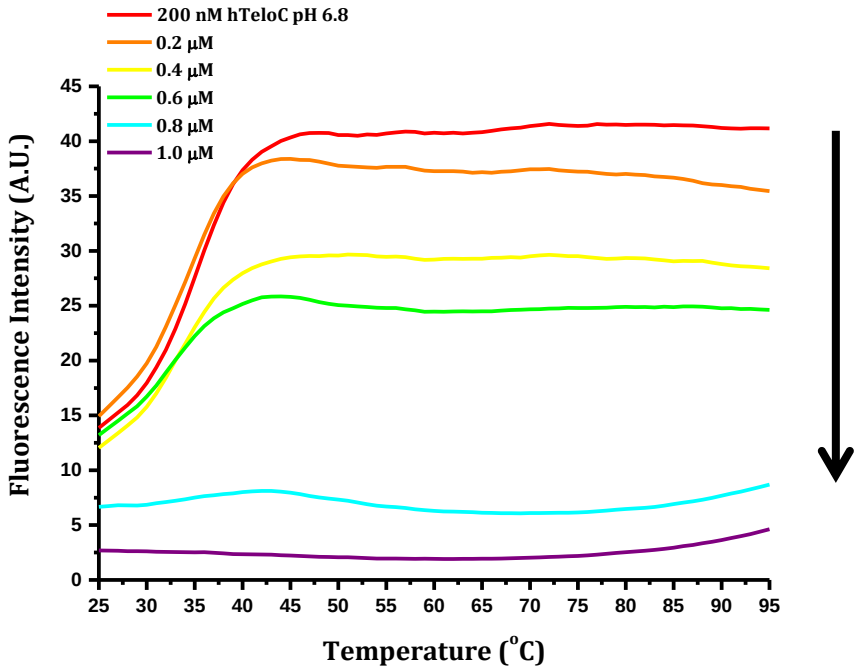



*Figure 2.7: Raw FRET melting curves for human telomeric i-motif with 0 – 1 µM of **BisA** at pH 6.8. [DNA] = 200 nM, buffer = 10 mM sodium cacodylate.*

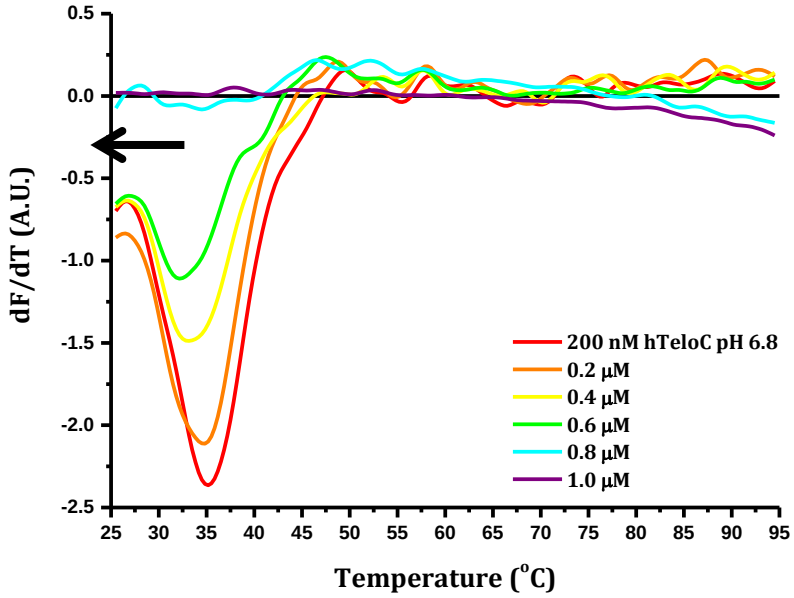



*Figure 2.8: First derivative of FRET melting curves for human telomeric i-motif with 0 – 1 µM of **BisA** at pH 6.8. [DNA] = 200 nM, buffer = 10 mM sodium cacodylate.*

| [**BisA**] (µM) | $\Delta T_m$ (°C) | |
|---|---|---|
| | pH 6.8 | pH 5.5 |
| 0.2 | -0.3 | -0.3 |
| 0.4 | -2 | -1 |
| 0.6 | -3 | -1.3 |
| 0.8 | nd* | -2.7 |
| 1.0 | nd* | -5.7 |

*Table 2.2: Change in melting temperature with increasing [**BisA**] at pH 6.8 and pH 5.5. * nd = not determined as there is no clear inflection point but could be >95°C. Error = ± 0.3°C.*

As discussed earlier, the human telomeric i-motif is not very stable at pH 6.8 ($T_m$ = 35°C) so **BisA** was further tested at pH 5.5 where the i-motif is much more stable ($T_m$ = 45°C). The same range of concentrations between 0 and 1 µM were tested against $hTeloC_{FRET}$ in pH 5.5 buffer (figure 2.9). This gave similar results with decreasing fluorescence signal across the pH range and apparent decrease in melting temperature (figure 2.10), though the effects were not as pronounced as at pH 6.8 (table 2.2).

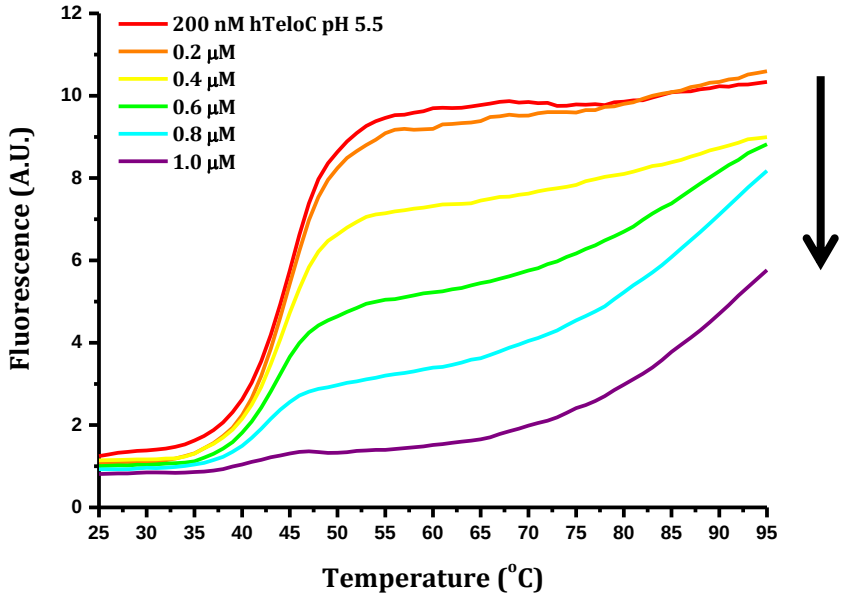



*Figure 2.9: Raw FRET melting curves for human telomeric i-motif with 0 – 1 µM of **BisA** at pH 5.5. [DNA] = 200 nM, buffer = 10 mM sodium cacodylate and 100 mM NaCl.*

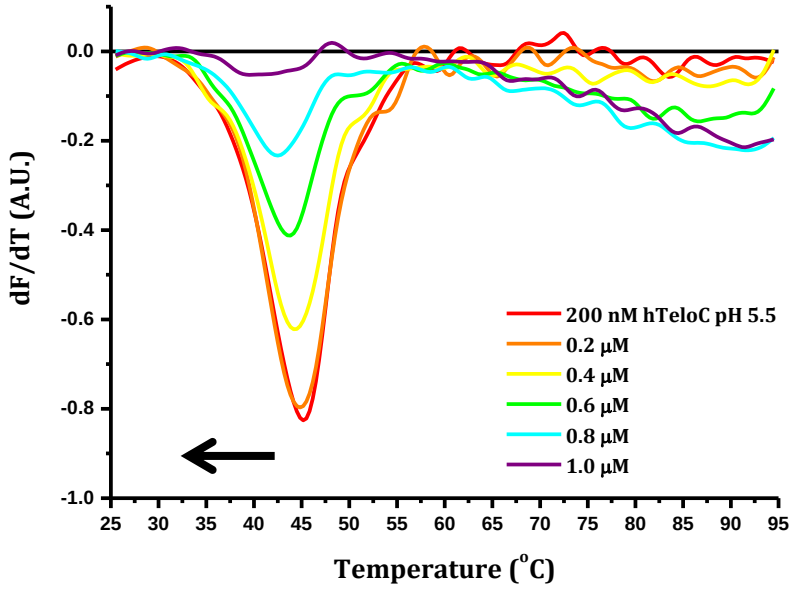



*Figure 2.10: First derivative of FRET melting curves for human telomeric i-motif 0 – 1 µM of **BisA** at pH 5.5. [DNA] = 200 nM, buffer = 10 mM sodium cacodylate and 100 mM NaCl.*

The FRET melting results of **BisA** at pH 7.4 are also interesting (figure 2.11). At this pH the hTeloC$_{FRET}$ sequence is unfolded in a single strand and gives a high florescence signal across the temperature range. With increasing **BisA** concentration, the starting fluorescence at 25°C decreases. This indicates folding of the DNA and at [**BisA**] = 0.8 –

1 μM a melt curve is observed suggesting that addition of **BisA** has caused the formation of a folded structure which melts with increasing temperature. However, it is not possible to determine from these FRET melting results alone, whether the DNA is folding into an i-motif structure or an alternative structure.

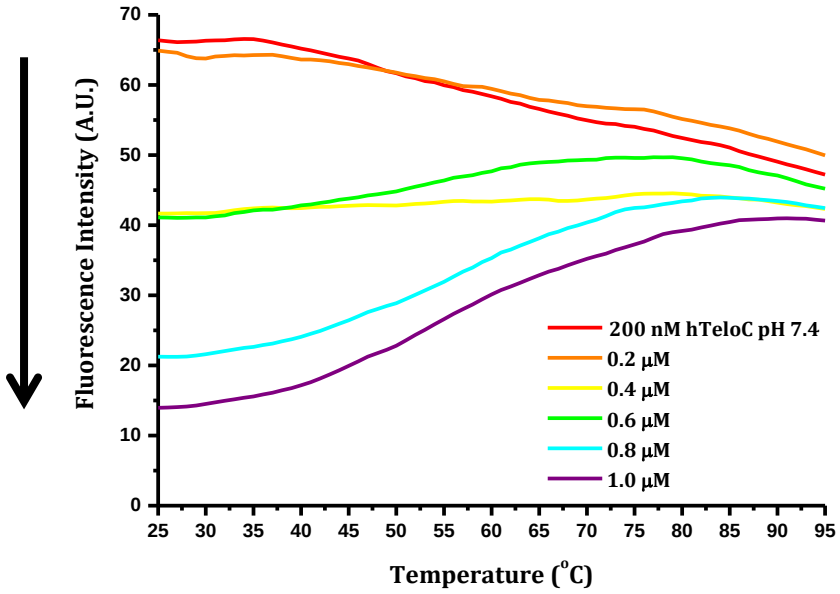



*Figure 2.11: Raw FRET melting curves for human telomeric i-motif with 0 – 1 μM of **BisA** at pH 7.4. [DNA] = 200 nM, buffer = 10 mM sodium cacodylate and 100 mM NaCl.*

Control FRET melting experiments were performed with the monoacridine (**MonoA**) in order to see whether the observed effects are as a result of the acridine unit itself or more specific to **BisA** and hence confirming the results of Alberti and colleagues.[100] At pH 6.8, **MonoA** showed little to no effect on the melting temperature or on the fluorescence up to a concentration of 1 μM (figure 2.12). At pH 5.5 where the i-motif is more stable, it can clearly be seen that **MonoA** had no effect on the fluorescence or melting temperature of the i-motif, even at the higher concentrations (figure 2.13). The diethylenetriamine linker was also tested at pH 5.5 and shown to have little or no effect, even at high concentrations (figure 2.14).

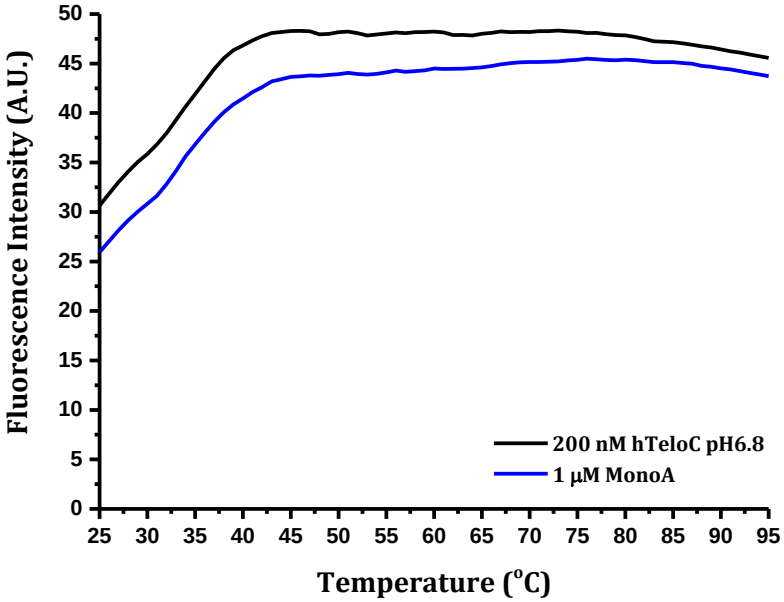



*Figure 2.12: Raw FRET melting curves for human telomeric i-motif with and without 1 µM of **MonoA** at pH 6.8. [DNA] = 200 nM, buffer = 10 mM sodium cacodylate.*

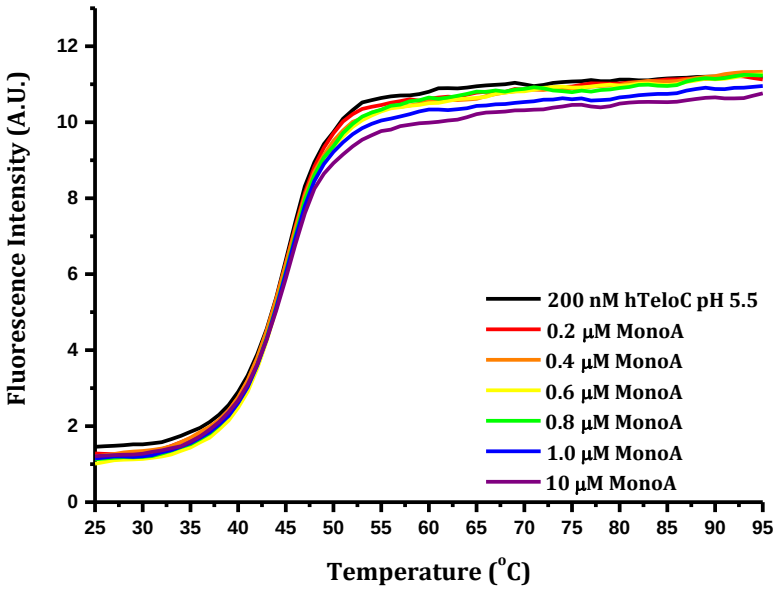



*Figure 2.13: Raw FRET melting curves for human telomeric i-motif with 0 – 10 µM of **MonoA** at pH 5.5. [DNA] = 200 nM, buffer = 10 mM sodium cacodylate and 100 mM NaCl.*

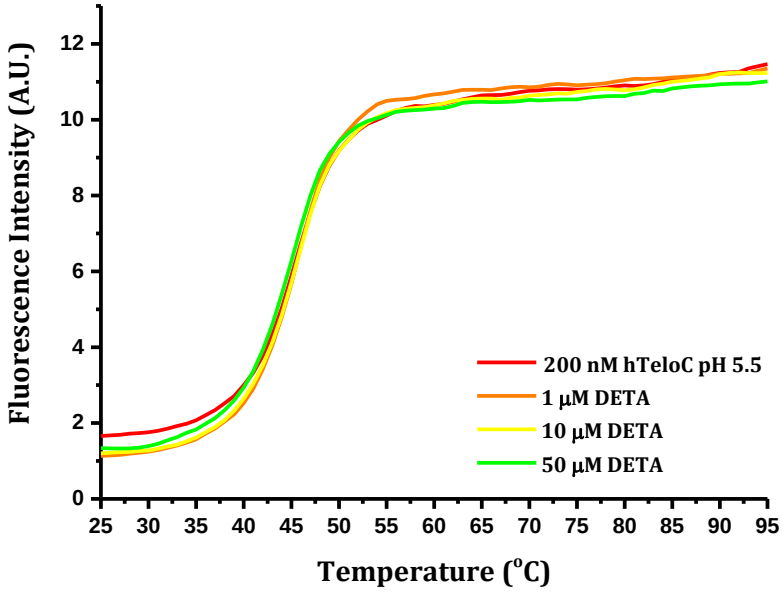



*Figure 2.14: Raw FRET melting curves for human telomeric i-motif with 0 – 100 µM of Diethylenetriamine (**DETA**) at pH 5.5. [DNA] = 200 nM, buffer = 10 mM sodium cacodylate and 100 mM NaCl.*

In summary, repetition of the previously published FRET melting experiments where the data was normalised (figure 2.5),[100] indicated that **BisA** significantly stabilised the i-motif structure ($\Delta T_m$ = 53°C). However, looking at the raw data in detail (figure 2.5 and 2.6) shows that the fluorescence decreases throughout the temperature range and does not show a melting transition within the temperature range of the experiment. This could be due to **BisA** stabilising the DNA so much that it doesn't melt, it could be as a result of **BisA** interacting with fluorophores and quenching fluorescence or it could be as a result of aggregation of the DNA. In any of these cases, the melting temperature determined from normalising the data is not accurate.

Using the first derivative of the data provides an alternative measure of the melting temperature. Here the data indicates that **BisA** actually destabilises the i-motif (figure 2.8) and this effect is also observed at pH 5.5 (figure 2.10) where the i-motif is more stable than at pH 6.8. Conversely the FRET melting data at pH 7.4, starting with the hTeloC sequence in an unfolded random coil, indicates folding of the DNA with increasing concentration of **BisA** (figure 2.11). Control experiments with the diethylenetriamine linker (**DETA**) and the monoacridine (**MonoA**) show that these compounds do not significantly affect the i-motif until much higher concentrations are

used. This is in agreement with Alberti and co-workers who showed that **MonoA** did not stabilise the i-motif.[100]

It is clear from these results that **BisA** is having an effect upon i-motif DNA and therefore it must be binding to the DNA in some way. However, FRET melting alone can only tell us whether the DNA is folding or unfolding, i.e. are the fluorophores moving closer together or further apart. This doesn't provide any information about how the structure of the DNA is changing and for this, further experiments are needed.

### 2.3.3 Circular Dichroism Experiments

In order to examine in more detail what effect **BisA** is having on the i-motif structure as observed in the FRET melting experiments, circular dichroism spectroscopy was employed. As discussed in section 1.4.5 and 2.3.1, circular dichroism is a very useful technique for studying DNA secondary structure and gives a characteristic spectrum depending on the type of structure present.[102] Folded i-motif DNA gives a positive peak at 288 nm and a negative peak at 255 nm. The CD spectrum of hTeloC oligonucleotide (5'-d(TAACCC)$_4$-3') was measured at a concentration of 10 μM at pH 5.5, 6.8 and 7.4 (figures 2.15, 2.16 and 2.17) in 10 mM sodium cacodylate buffer (with 100 mM sodium chloride at pH 5.5 and 7.4). **BisA** was titrated into the sample in 1 equivalent aliquots (10 μM) up to 10 equivalents (100 μM).

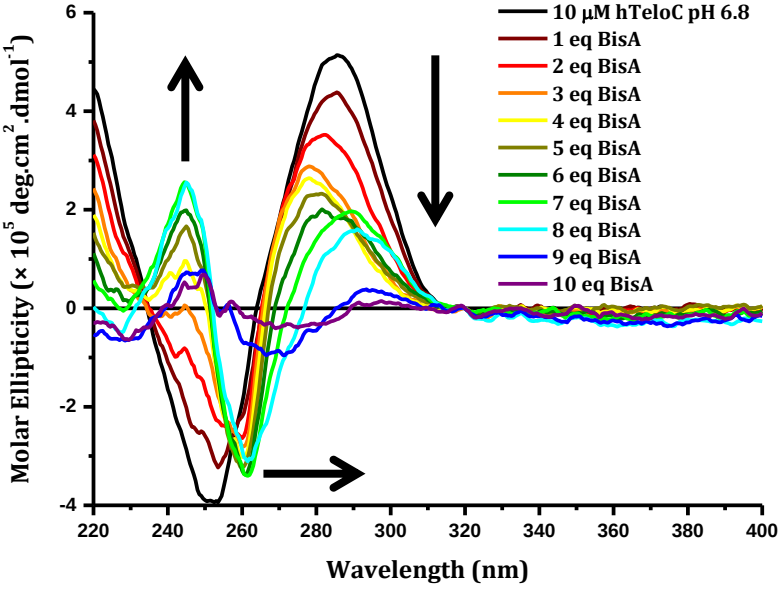



*Figure 2.15: CD titration of human telomeric i-motif with increasing concentration of **BisA** at pH 6.8. [DNA] = 10 μM, buffer = 10 mM sodium cacodylate.*

The CD spectrum at pH 6.8 (figure 2.15) in the absence of any ligand shows the presence of the i-motif structure with a negative peak at 253.5 nm and positive peaks at 285.5 and 220 nm. As **BisA** is titrated into the sample, the spectrum shows a decrease in molar ellipticity and a hypsochromic shift in the positive peak to 278 nm after addition of 4 equivalents of **BisA**. This could be indicative of unfolding of the DNA however, this change is not accompanied by a shift in the negative peak to 250 nm that is expected for single stranded DNA. In fact the negative peak shifts in the opposite direction to 260 nm indicating a different structural change. After addition of more than 4 equivalents of **BisA** the positive peak also begins to shift towards higher wavelengths indicating a further structural change. After addition of 9 and 10 equivalents of **BisA**, precipitation can be observed with very low molar ellipticity across the spectrum and with less smooth curves. Other changes are also observed at lower wavelengths where there is a decrease in molar ellipticity at 220 nm and formation of a new positive peak at 245 nm. This new positive peak is in the same spectral region as where **BisA** absorbs (248 nm) and could be an induced CD signal for the **BisA** molecule upon its binding to the DNA. These changes in the CD spectrum, clearly show that **BisA** is interacting with the i-motif and causing some form of structural change.

Whether or not the new peak at 245 nm is due to **BisA** or the DNA, the change in molar ellipticity shows that **BisA** binds to the DNA and can be fitted to obtain a measurement of the binding affinity. The change in molar ellipticity at 245 nm with increasing concentration of **BisA** was normalised to give the change in terms of fraction of **BisA** bound between 0 and 1 (figure 2.16). As can be seen, the data does not reach a plateau, this is because at concentrations above 80 μM, precipitation occurred, causing a large decrease in molar ellipticity. The data was fitted with a 2 to 1 binding model (equation 2.1) as this gave the best fit of the data.[155]

*Equation 2.1:* $$\theta = \frac{K_1[BisA]+2K_1K_2[BisA]^2}{1+K_1[BisA]+K_1K_2[BisA]^2}$$

Where $K_1$ and $K_2$ are the equilibrium association constants for the first and second binding sites. This gave rise to dissociation constants of $K_{D1}$ = 59.3 μM and $K_{D2}$ = 96.0

μM. This indicates that **BisA** has a moderate binding affinity for the i-motif with two possible binding sites.

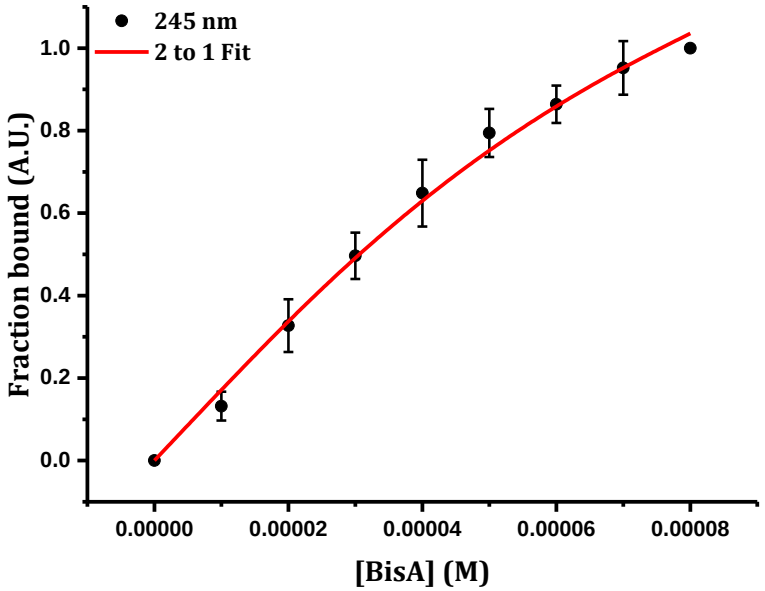



*Figure 2.16: Graph to show the fraction bound with increasing concentration of **BisA** from the change in molar ellipticity at 245 nm at pH 6.8. Error bars show the standard deviation across two repeats. [DNA] = 10 μM, buffer = 10 mM sodium cacodylate.*

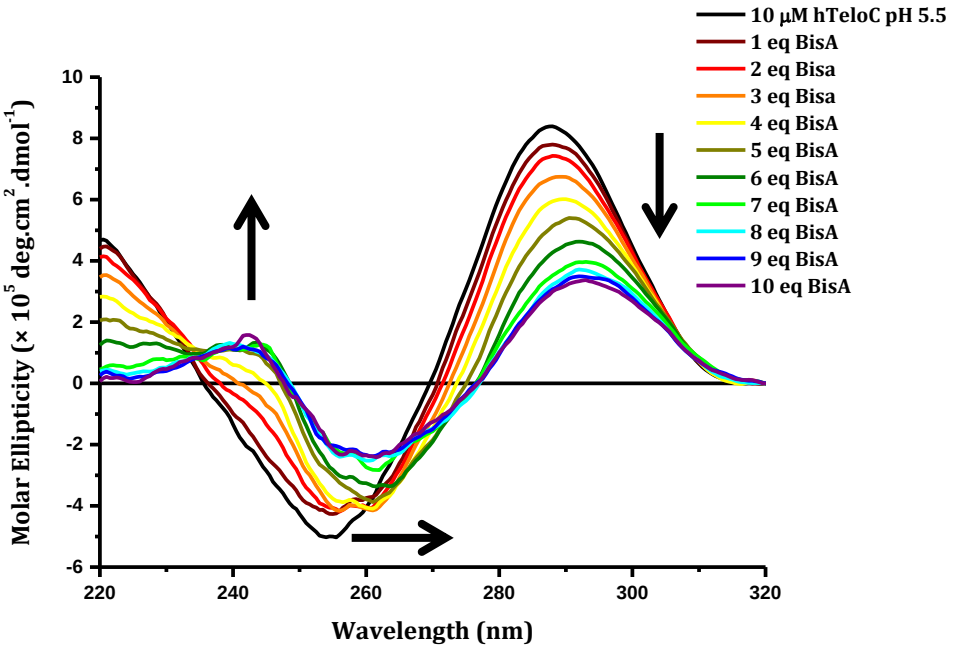



*Figure 2.17: CD titration of human telomeric i-motif with increasing concentration of **BisA** at pH 5.5. [DNA] = 10 μM, buffer = 10 mM sodium cacodylate and 100 mM NaCl.*

The CD titration at pH 5.5 (figure 2.17), gave similar final results to the titration at pH 6.8. A decrease in molar ellipticity, a shift in the positive signal from 288 nm towards

300 nm, a shift in the negative signal from 255 nm to 260 nm and the formation of an ICD peak, this time at 242 nm were all observed. However, unlike at pH 6.8, the positive peak moves smoothly towards the higher wavelength rather than first moving towards 278 nm and then back to 300 nm as was seen in figure 2.15. These observations suggest that in both cases the same or a very similar overall change in structure occurs however, at the two different pH's this occurs by different routes. This may be because of the lower stability of the i-motif at pH 6.8. Another difference between the two pH's is that under the more acidic conditions, precipitation did not occur which is illustrated by the much smoother spectrum and the fact that the molar ellipticity did not dramatically drop at high **BisA** concentrations such as at pH 6.8.

As before, the change in CD signal is indicative of **BisA** binding to the i-motif and by fitting the change in molar ellipticity a binding affinity can be measured. In this case the data did reach a plateau (figure 2.18) and was first fitted with the two binding site model in equation 2.1. This gave dissociation constants of $K_{D1}$ = 30.8 μM and $K_{D2}$ = 259 μM suggesting a moderate binding affinity for the first binding site with a much weaker second binding site. However, compared to pH 6.8 the data does show a greater degree of sigmoidal character, which means it may be better fitted by the Hill equation. This fitting not only gives a measure of the binding affinity but also of the cooperativity of binding meaning that the initial binding of **BisA** affects the binding of subsequent **BisA** molecules. The data was fitted with equation 2.2 where $K$ is the equilibrium dissociation constant, equal to the concentration at which half the binding sites are occupied.[156] $n$ is the Hill coefficient which gives a measure of cooperativity, if $n$ = 1 there is no cooperativity, $n$ > 1 indicates positive cooperativity and $n$ < 1 shows negative cooperativity.

*Equation 2.2:* $$\theta = \frac{[BisA]^n}{K^n + [BisA]^n}$$

Fitting with equation 2.2 gave a better fit of the data and gave rise to a dissociation constant of $K_D$ = 2.5 μM with a Hill coefficient of $n$ = 2.3. These results suggests that **BisA** may have a much higher affinity for the i-motif and that the binding of **BisA** may show positive cooperativity.

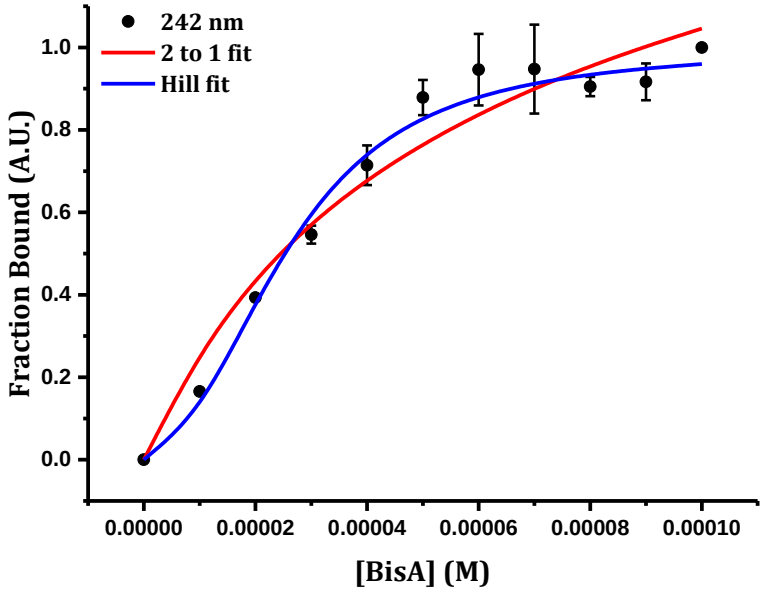



*Figure 2.18: Graph to show the fraction bound with increasing concentration of **BisA** from the change in molar ellipticity at 242 nm at pH 5.5. Error bars show the standard deviation across two repeats. [DNA] = 10 μM, buffer = 10 mM sodium cacodylate and 100 mM NaCl.*

At pH 7.4, the i-motif is not formed and the sequence takes a random coil structure. This gives a different CD spectrum to the i-motif with a positive peak at 275 nm and a negative peak at 250 nm. When **BisA** was titrated with hTeloC at pH 7.4, similar reductions in signal could be seen at both 275 nm and 250 nm with the formation of a smaller ICD at 248 nm also occurring (figure 2.19). This data provides supporting evidence that **BisA** is not simply destabilising and unfolding the i-motif structure because if that was the case, very little change would be expected when **BisA** is added to already unfolded DNA. Here it is apparent that some condensation of the DNA is occurring, in fact at pH 7.4 and 6.8, visible precipitate was observed at the higher concentrations of **BisA** (70 and 80 μM respectively). The presence of an ICD signal gives further evidence that a binding event is occurring and the compound must bind to the DNA if it is causing aggregation.

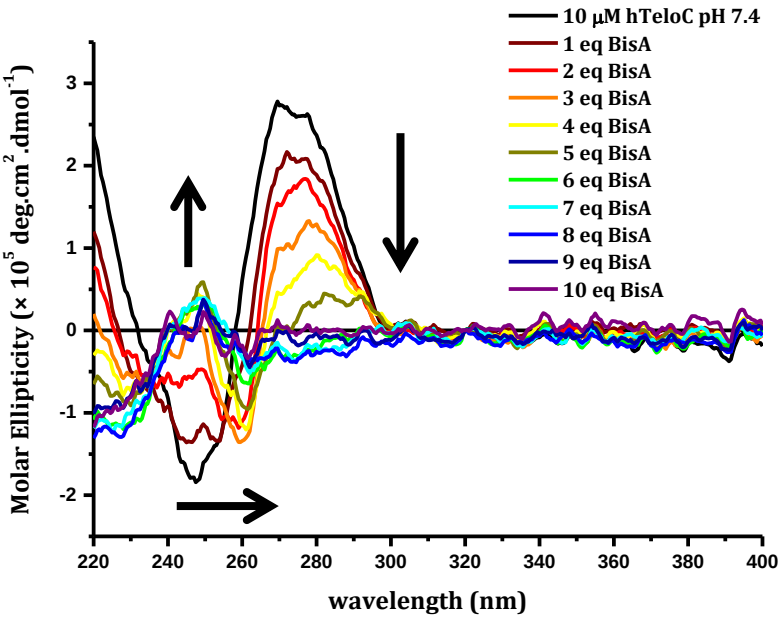



*Figure 2.19: CD titration of human telomeric i-motif with increasing concentration of* ***BisA*** *at pH 7.4. [DNA] = 10 μM, buffer = 10 mM sodium cacodylate and 100 mM NaCl.*

To give an indication of whether the signals seen at 242-248 nm were indeed ICDs rather than a signal originating from **BisA** in the absence of DNA, the CD spectrum of **BisA** was measured at the top concentration (100 μM) in each pH buffer (figure 2.20). **BisA** is not chiral so no CD signal is expected however, in all three cases a very weak negative signal is observed between 240 and 260 nm. The ICD signals in the presence of the DNA are positive and of a higher magnitude. The presence of a small signal in this region in the absence of DNA suggests that the ICD seen in figures 2.15, 2.17 and 2.19 could be from **BisA** rather than a conformational change in the DNA.

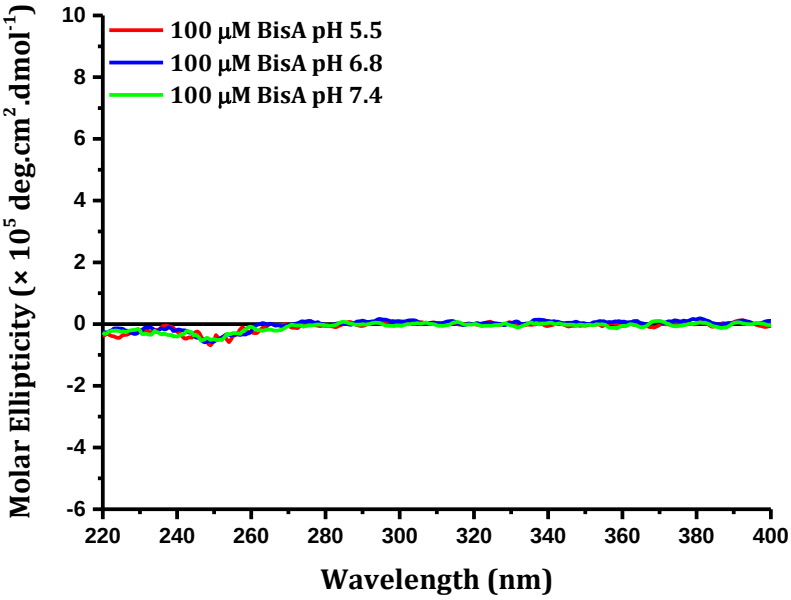



*Figure 2.20: CD spectrum of **BisA** at pH 5.5, 6.8 and 7.4. [**BisA**] = 100 µM, buffer = 10 mM sodium cacodylate and 100 mM NaCl pH 5.5 and 7.4, buffer = 10 mM sodium cacodylate pH 6.8.*

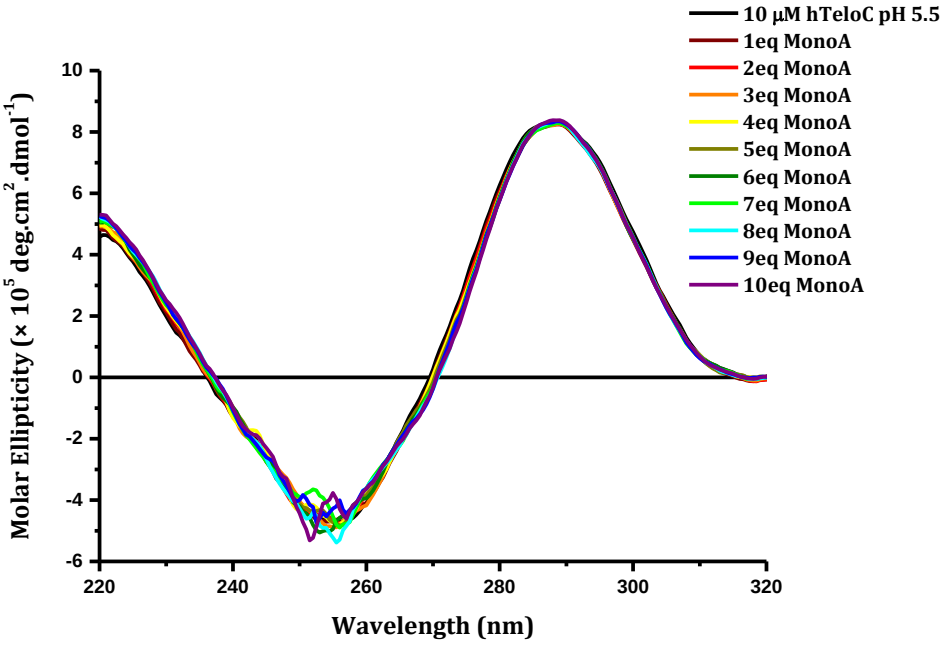



*Figure 2.21: CD titration of human telomeric i-motif with increasing concentration of **MonoA** at pH 5.5. [DNA] = 10 µM, buffer = 10 mM sodium cacodylate and 100 mM NaCl.*

Control experiments were carried out with **MonoA** which show no change in the CD spectrum of hTeloC at pH 5.5 (figure 2.21). Therefore, if **MonoA** is binding to the i-motif it is not having any effect on the structure at this pH. At pH 6.8 there is a decrease in signal and a shift in the positive peak from 288 nm to 275 nm suggesting that at this

pH, where the i-motif is less stable, **MonoA** causes some unfolding of the structure (figure 2.22). As **MonoA** appears to cause some unfolding of the DNA at pH 6.8 it may be binding to the DNA, however there were no observable induced CD signals.

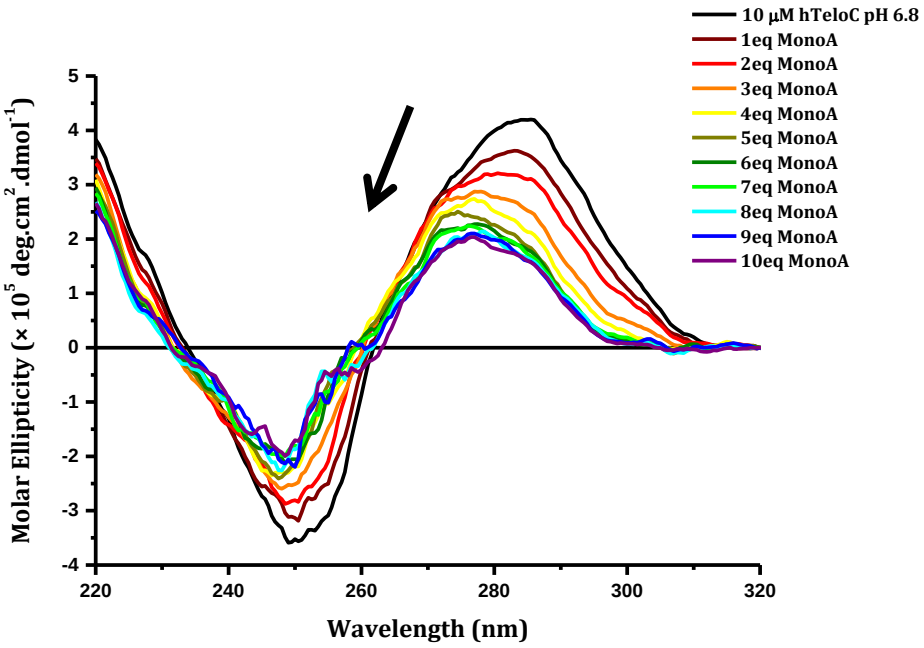



*Figure 2.22: CD titration of human telomeric i-motif with increasing concentration of* ***MonoA*** *at pH 6.8. [DNA] = 10 µM, buffer = 10 mM sodium cacodylate.*

Once again the change in molar ellipticity as a result of **MonoA** can be fitted to obtain a binding curve. Following the same method outlined above for **BisA**, the change in molar ellipticity at 285.5 nm was normalised to give a measure of the fraction of DNA that is unfolded (figure 2.23) and this was fitted with both equation 2.1 and 2.2. Using the 2 to 1 model gave equilibrium dissociation constants of $K_{D1}$ = 35.5 µM and $K_{D2}$ = 194 µM suggesting a moderate affinity for one binding site and a weaker affinity for a second binding site. Using the Hill equation gave a better fit of the data at higher concentrations of **MonoA** and was within the error at the lower concentrations. This fit gave the **MonoA** concentration at which half the DNA is unfolded ([**MonoA**]$_{50}$) as 2.58 µM. This suggests a stronger affinity for the structure. The cooperativity was measured to be $n$ = 2.5 suggesting a positive cooperativity. Alternatively, as no change is seen at pH 5.5, **MonoA** may not be binding to the i-motif at all and may just be increasing the ionic strength of the sample due to its positive charge and hence this may be the cause of unfolding at pH 6.8.

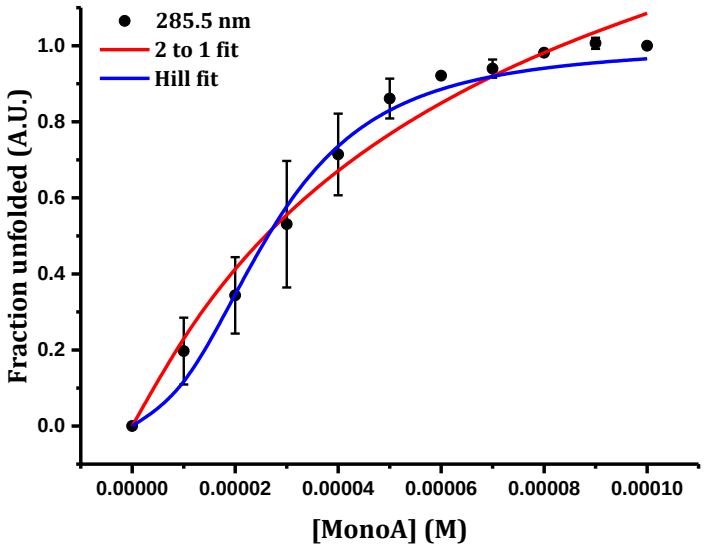



*Figure 2.23: Average change in molar ellipticity with increasing concentration of* ***MonoA*** *at pH 6.8. Error bars show the standard deviation across two repeats. [DNA] = 10 μM, buffer = 10 mM sodium cacodylate.*

The combined circular dichroism results show that both **BisA** and **MonoA** interact with the human telomeric i-motif but they have differing effects which are also dependent on pH. At pH 6.8 **BisA** causes a change in structure of the DNA which is indicated by the change in CD spectrum as the positive peak at 285 nm first undergoes a hypsochromic shift which is followed by a bathochromic shift as more equivalents of **BisA** are added. Finally resulting in precipitation after 9 to 10 equivalents. At pH 5.5, the positive peak of the spectrum undergoes a red shift but there is no precipitation. In both cases the negative peak shifts towards 260 nm and there is the formation of a new peak at around 242 to 245 nm. This suggests that at both pHs, the same overall structural change is occurring but at pH 6.8 there is a different change, perhaps unfolding, that occurs as part of the process. This may be because of the lower stability of the i-motif at this pH, making it much more susceptible to unfolding. The peaks at 245 and 242 nm may be due to the conformational change of the DNA but, as **BisA** absorbs in a similar UV region (248 nm) there is a good chance that this peak is an induced CD signal arising from **BisA** as it binds to the DNA. **BisA** itself only shows a very weak negative signal in this region in the absence of DNA. Unlike **BisA**, the CD titration of the i-motif with **MonoA** showed no change in the spectrum at pH 5.5 and at pH 6.8 the changes in the CD indicate unfolding. This suggests that **MonoA** may be binding to the i-motif and has a destabilising effect on the already unstable structure but it may also be the case that

**MonoA** is simply increasing the ionic strength of the sample due to its positive charge and that this is enough to cause unfolding at pH 6.8.

### 2.3.4 Fluorescence Titration Experiments

Fluorescence emission spectroscopy provides an alternative means to measure a binding affinity of **BisA** for the i-motif. In circular dichroism, the concentration of DNA is kept constant and the amount of **BisA** is varied but in the fluorescence titration, it is the amount of DNA that is varied with **BisA** constant. Depending on the properties of the ligand, fluorescence spectroscopy can be carried out at lower concentrations and can enable the assessment of higher DNA:ligand ratios. 1 μM samples of **BisA** were excited at 255 nm and the emission was observed between 290 and 490 nm with peaks at 326 and 426 nm (figure 2.24). The change in fluorescence was measured as increasing amounts of telomeric i-motif (hTeloC) were titrated into the sample. Upon addition of the DNA, the fluorescence of **BisA** can be seen to decrease, reaching a plateau at around 15 equivalents of hTeloC to **BisA**. This decrease in fluorescence is consistent with previous observations of **BisA** fluorescence as described earlier (section 2.1.1) showing that **BisA** fluorescence is decreased between 70-80% in the presence of purine bases.[136] The average decrease in fluorescence shown in these experiments is 76%, consistent with Teulade-Fichou and co-worker's findings.[136] While the sequence of hTeloC is predominantly rich in pyrimidine bases cytosine and thymine, each loop of the i-motif contains two adenine bases. It has previously been shown that **BisA** is able to bind to the adenine loops of hairpin DNA.[138,139] This decrease in fluorescence may therefore be the result of **BisA** binding to the adenine bases in the loops of the i-motif. This decreasing fluorescence signal could also be the result of condensation or precipitation of the DNA and **BisA.** In the previous FRET and circular dichroism experiments which indicate possible precipitation there is an excess of **BisA** compared to DNA however, in this experiment there is an excess of DNA and this means that there is less likely to be a precipitation effect.

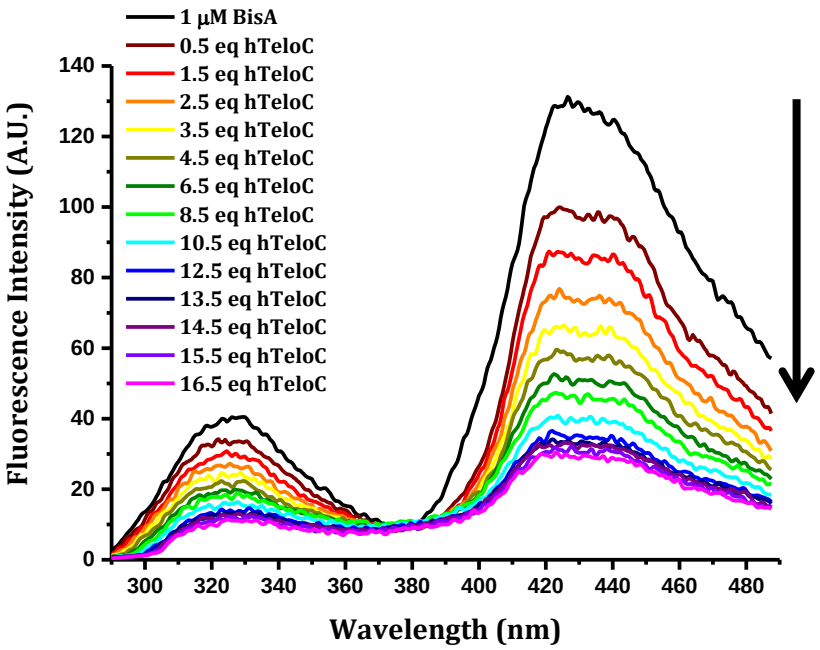



*Figure 2.24: Fluorescence titration of **BisA** with increasing concentration of human telomeric i-motif at pH 5.5. [**BisA**] = 1 µM, buffer = 10 mM sodium cacodylate and 100 mM NaCl.*

The change in fluorescence at 426 nm was normalised and fitted with a 2 to 1 binding model (equation 2.1, figure 2.25) where $K_1$ and $K_2$ are the equilibrium association constants for the first and second binding sites. This gave an average $K_{D1}$ = 2.1 ± 0.1 µM and $K_{D2}$ = 146 ± 10.2 µM. This indicates a relatively strong binding affinity for the 1st binding event with a much weaker 2nd binding interaction and compares well with the $K_D$ measured from the Hill fit of the CD titration (figure 2.18, $K_D$ = 2.5 µM).

*Equation 2.1:* $$\theta = \frac{K_1[DNA]+2K_1K_2[DNA]^2}{1+K_1[DNA]+K_1K_2[DNA]^2}$$

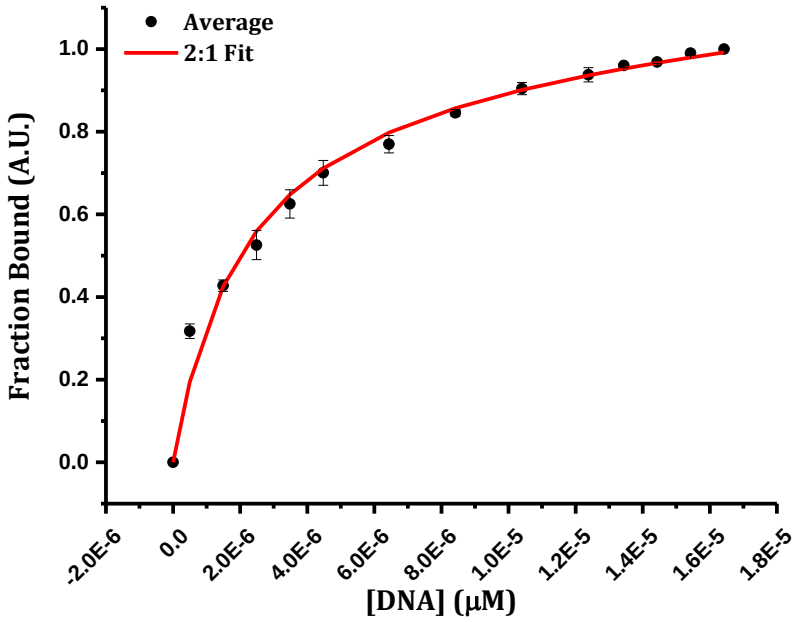



*Figure 2.25: 2 to 1 Binding curve for the normalised change in fluorescence of **BisA** with increasing concentration of human telomeric i-motif at pH 5.5. Error bars show the standard deviation across three repeats. [**BisA**] = 1 μM, buffer = 10 mM sodium cacodylate and 100 mM NaCl.*

These results provide further evidence that **BisA** binds to the i-motif. The decrease in **BisA** florescence is consistent with previous examples of **BisA** DNA binding[136] and may indicate that **BisA** binds to the adenine loop region of the i-motif.[138,139] These results may also be showing condensation of the DNA but as there is an excess of DNA it is more likely that the decrease in fluorescence is a binding effect. The change in fluorescence was fitted with a 2 to 1 binding model giving equilibrium binding constants of $K_{D1}$ = 2.1 μM and $K_{D2}$ = 146 μM. This is in contrast to the binding constant determined by circular dichroism and suggests a stronger affinity for the first binding site. However, these different results are not directly comparable as circular dichroism is carried out at higher **BisA** concentrations and the two techniques monitor two different properties; change in fluorescence of the ligand versus change in optical activity of the DNA and ligand. The fluorescence titration is likely to give the more reliable result because the experiment should not be affected by precipitation at these ligand:DNA ratios.

### 2.3.5 FRET Titration Experiments

In order to investigate further the structural changes observed in CD experiments, a FRET titration was carried out. The CD titration at pH 6.8 was unusual in that it showed an initial shift in the positive peak to the left and then back to the right, this might indicate an unfolding event followed by a different change in structure. As described in section 1.4, by exciting the FRET labelled DNA at 490 nm and taking the emission spectrum from 500 to 650 nm, the fluorescence of both FAM and TAMRA can be observed at 515 and 580 nm respectively. When the DNA is folded the emission of FAM is expected to be low, if it unfolds the emission should increase at 515 nm. At pH 6.8 (figure 2.26) the fluorescence of FAM in the absence of **BisA** is already at an intermediate level (120 A.U.) compared to the fully folded DNA at pH 5.5 (figure 2.27) where the fluorescence of FAM is very low (5 A.U.) This suggests that the i-motif is not fully and compactly folded at this pH, as was seen by the incomplete melting curves in figure 2.7. If folding occurs, the fluorescence of FAM is expected to decrease but the fluorescence of TAMRA should not change. If the fluorescence of both fluorophores decreases then this is indicative of precipitation.

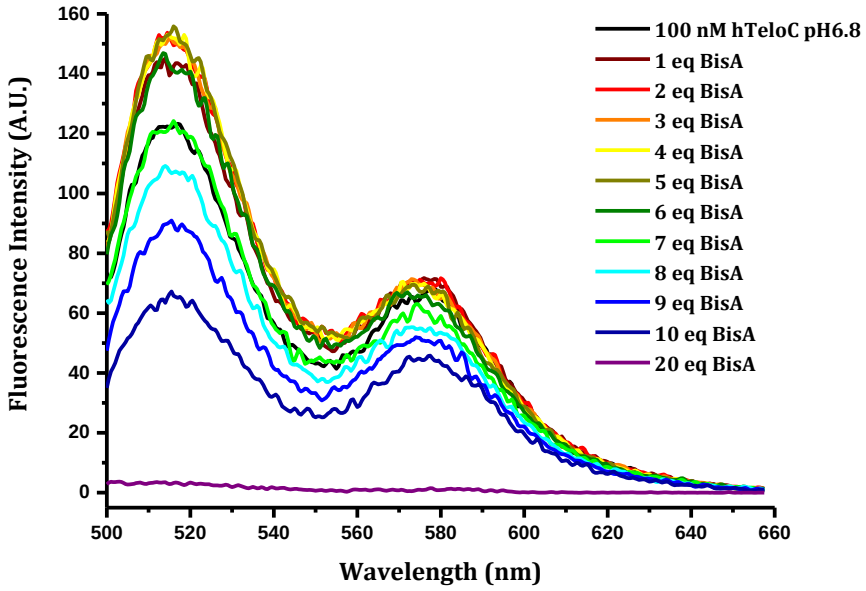



*Figure 2.26: Example FRET titration of hTeloC with increasing equivalents of **BisA** at pH 6.8. [DNA] = 100 nM, buffer = 10 mM sodium cacodylate.*

At pH 6.8 (figure 2.26), it can be seen that upon addition of **BisA**, the fluorescence of FAM at 515 nm increases with a negligible change in the fluorescence of TAMRA at 580

nm. This is indicative of the DNA unfolding. However, after addition of 5 equivalents of **BisA**, further additions of the compound cause the fluorescence of both fluorophores to decrease, eventually reaching a base level after addition of 20 equivalents. This is consistent with a condensation or precipitation of the DNA. However, the decreasing ratio between FAM and TAMRA (2:1 at 8 eq **BisA**, 1.5:1 at 10 eq **BisA**) does suggest that folding may be occurring or that the fluorophores are coming into close proximity during condensation.

At pH 5.5 (figure 2.27), it can be seen that the fluorescence of FAM is already very low showing that the DNA is completely folded at this pH. As **BisA** is titrated into the sample, there is no increase in the fluorescence of FAM showing that at this pH, the addition of **BisA** does not cause unfolding of the i-motif. This is consistent with the CD spectrum in figure 2.17 which, unlike at pH 6.8 (figure 2.15), does not show a shift in the positive peak towards 275 nm. In the FRET titration experiment, as the concentration of **BisA** is increased, there is a decrease in the fluorescence of both FAM and TAMRA indicating condensation of the DNA.

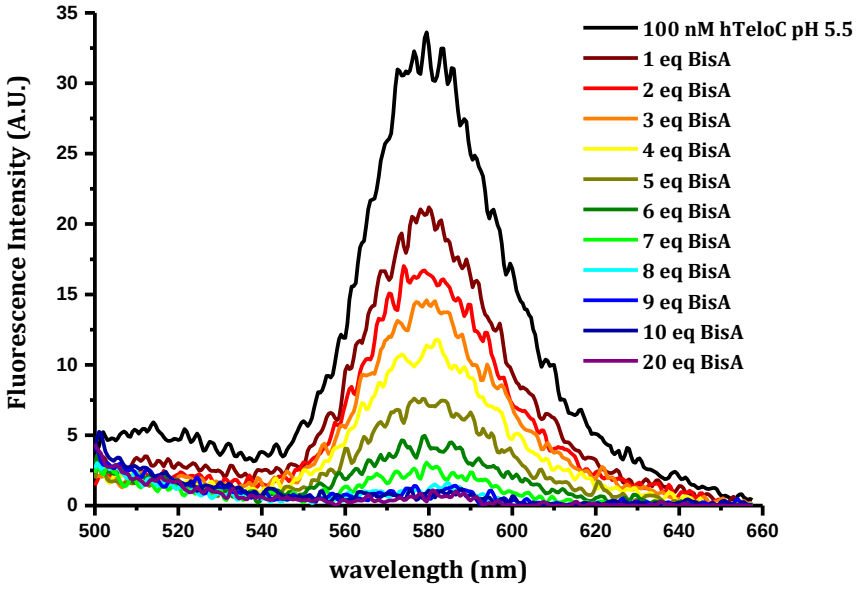



*Figure 2.27: FRET titration of hTeloC with increasing equivalents of **BisA** at pH 5.5. [DNA] = 100 nM, buffer = 10 mM sodium cacodylate and 100 mM NaCl.*

These results show that **BisA** causes condensation or precipitation of i-motif DNA at both pH 6.8 and pH 5.5. The titration at pH 6.8 is consistent with the CD observations showing that the DNA begins to unfold before condensation occurs whereas at pH 5.5,

condensation occurs immediately. With similar shaped CD spectra at the end of both titrrations it was already hypothesized that the same type of structure is forming at both pHs. As both FRET titrations seem to show condensation or precipitation of the DNA it would indicate that this is also what is being observed in the CD titrations.

### 2.3.6 UV Titration Experiments

To further show the interaction between **BisA** and the i-motif and to see if there were other signs of precipitation or condensation, the compound was tested with the DNA in a UV titration experiment. The absorbance of hTeloC i-motif was measured at a concentration of 2.5 µM at 260 nm (figure 2.28). **BisA** was titrated into the DNA sample in 1 equivalent aliquots and the sample was mixed. In the first titrations carried out, the absorbance could be seen to increase with a shift in wavelength towards 248 nm, consistent with an increasing concentration of **BisA** (figure 2.28). However, there was also a small but consistent increase in absorbance in the base line between 500 and 800 nm, indicative of a scattering effect.

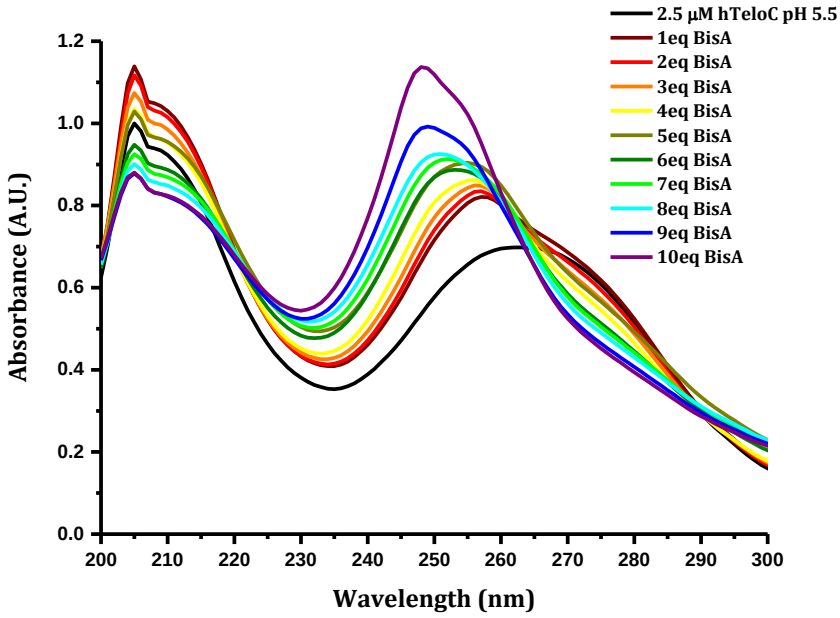



*Figure 2.28: Example UV titration of hTeloC with increasing concentration of **BisA** at pH 5.5. [DNA] = 2.5 µM, buffer = 10 mM sodium cacodylate and 100 mM NaCl.*

To ensure that there was no dust particles in the sample causing scattering, the titration was repeated but the sample was mixed and centrifuged at 6000 rpm before measuring each new concentration (figure 2.29).

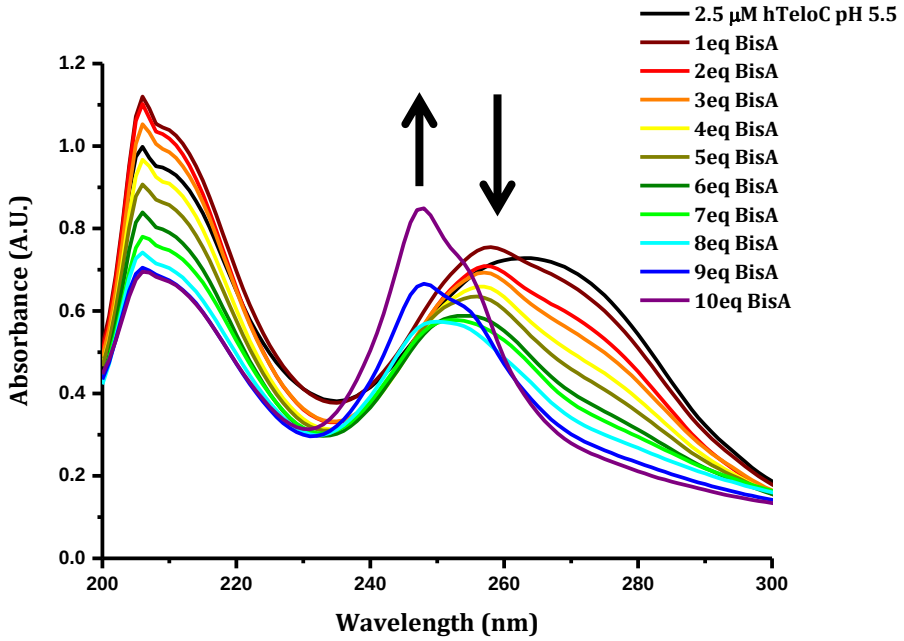



*Figure 2.29: UV titration of human telomeric i-motif with increasing concentration of **BisA** at pH 5.5. [DNA] = 2.5 μM, buffer = 10 mM sodium cacodylate and 100 mM NaCl.*

At pH 5.5 (figure 2.29) there is a significant decrease in the absorbance of DNA at 260 nm (figure 2.30). This would indicate the precipitation of DNA which is then separated from the sample during centrifugation. After addition of 7 equivalents of **BisA** a peak begins to appear at 248 nm which corresponds to the absorbance of **BisA** itself and suggests that the majority of the available **BisA** binding sites had been filled.

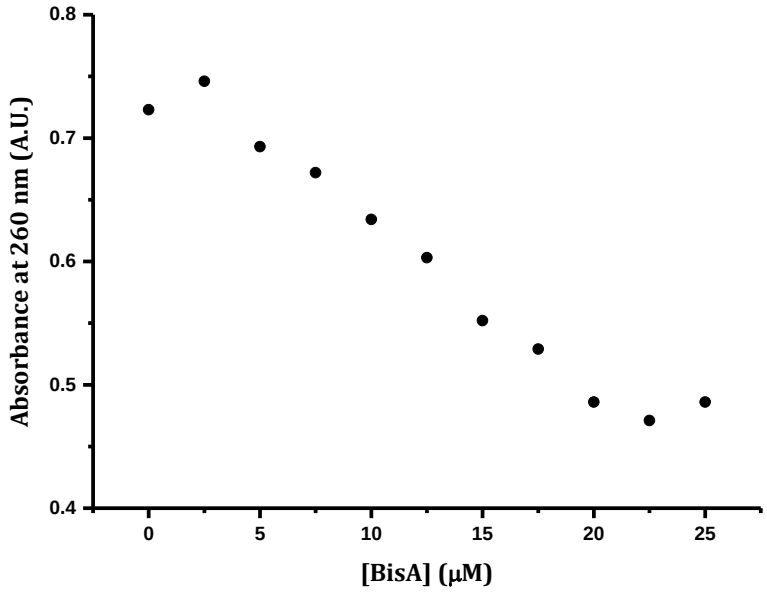



*Figure 2.30: Change in UV absorbance at 260 nm with increasing concentration of **BisA** at pH 5.5. [DNA] = 2.5 μM, buffer = 10 mM sodium cacodylate and 100 mM NaCl.*

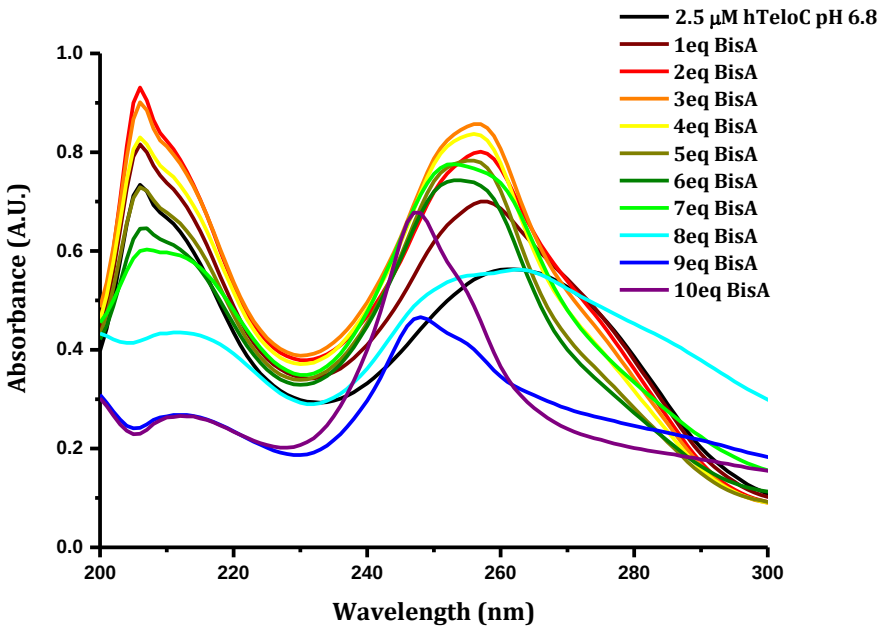



*Figure 2.31: UV titration of human telomeric i-motif with increasing concentration of **BisA** at pH 6.8. [DNA] = 2.5 µM, buffer = 10 mM sodium cacodylate.*

At pH 6.8 (figure 2.31), there was an initial increase in absorbance at 260 nm which may be indicative of the i-motif unfolding or **BisA** binding, but is quickly followed by a similar decrease in absorbance as the concentration of **BisA** is increased, indicating the same condensation of the DNA (figure 2.32). This is consistent with the earlier CD and FRET titration experiments which also indicate initial unfolding of the DNA at this pH.

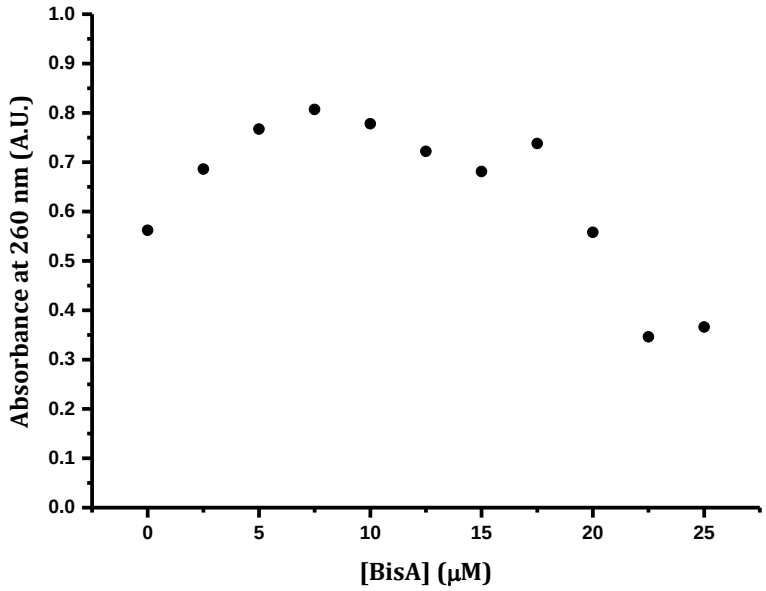



*Figure 2.32: Change in UV absorbance at 260 nm with increasing concentration of **BisA** at pH 6.8. [DNA] = 2.5 µM, buffer = 10 mM sodium cacodylate.*

To see whether **BisA** was simply aggregating with itself an analogous titration was carried out with just **BisA** in buffer at both pH 5.5 and 6.8 to check that the absorbance remained linear according to the Beer-Lambert law (equation 2.3) (figure 2.33 & 2.35). Where $A$ is the absorbance, $\varepsilon$ is the molar extinction coefficient, $c$ is the concentration and $l$ is the path length of the sample which in this case was 1 cm. This equation shows that the absorbance of **BisA** should remain linear with increasing concentration.

*Equation 2.3:* $A = \varepsilon cl$

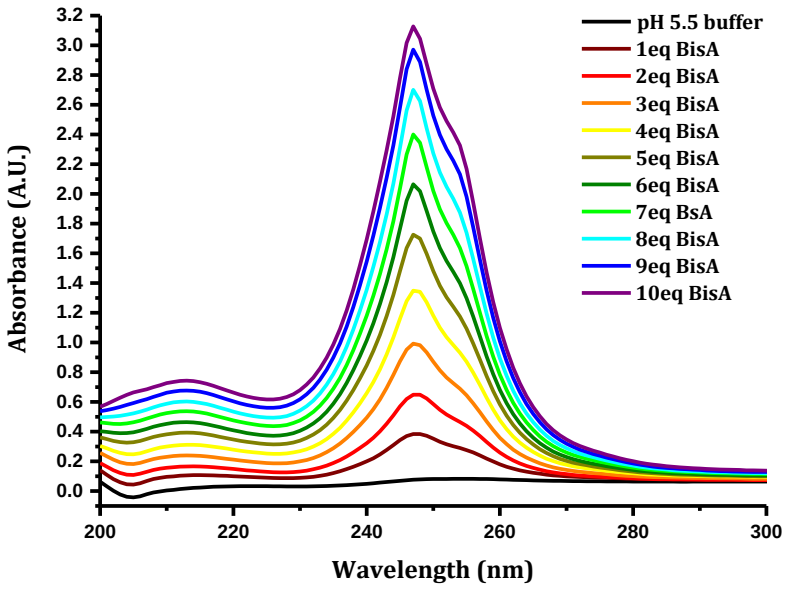



*Figure 2.33: UV titration of **BisA** in the absence of any DNA at pH 5.5. Buffer = 10 mM sodium cacodylate and 100 mM NaCl.*

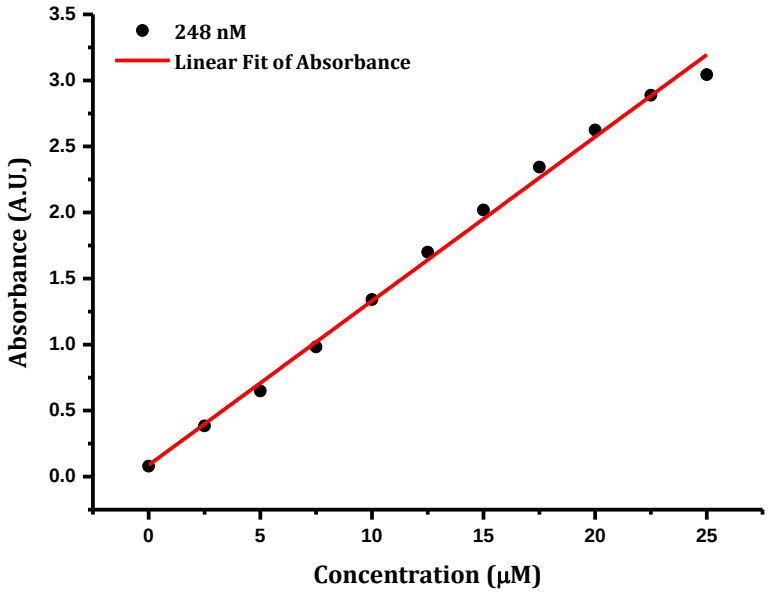



*Figure 2.34: Change in absorbance of **BisA** at 248 nm in pH 5.5 buffer showing it is linear, obeying the Beer-Lambert law. Buffer = 10 mM sodium cacodylate and 100 mM NaCl.*

These results show that as **BisA** is titrated into the buffer, the absorbance at 248 nm increases as expected (figure 2.33). It also shows that the absorbance at 260 nm increases, therefore, it would be expected that addition of **BisA** to DNA would result in an increase in absorbance at 260 nm, contradictory to what was observed. By plotting the absorbance at 248 nm with concentration of **BisA** (figure 2.34) it can be seen that the data is linear and therefore obeys the Beer-Lambert law showing, that it does not aggregate with itself over this concentration range.

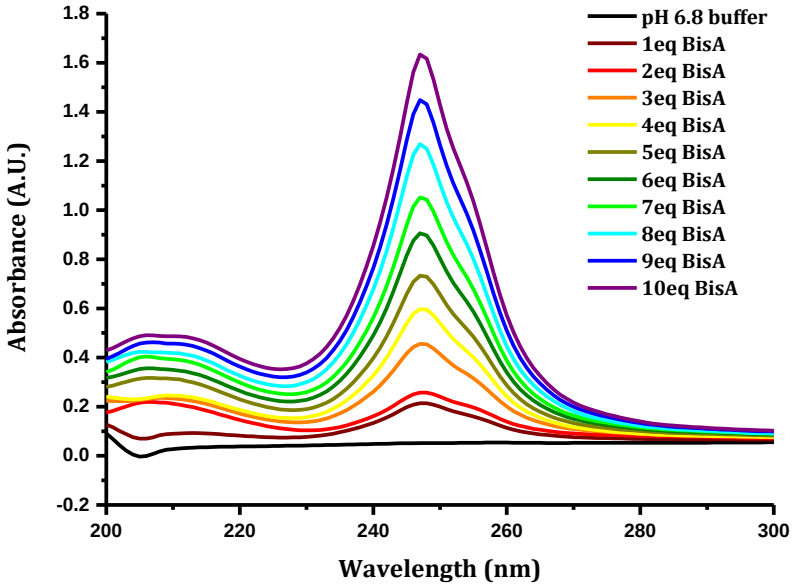



*Figure 2.35: UV titration of **BisA** in the absence of DNA at pH 6.8. Buffer = 10 mM sodium cacodylate.*

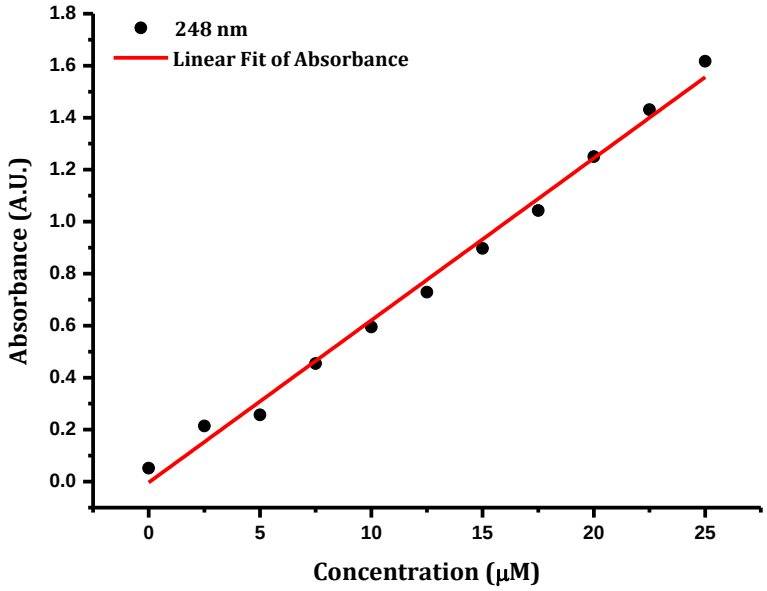



*Figure 2.36: Change in absorbance of **BisA** at 248 nm in pH 6.8 buffer showing it obeys the Beer-Lambert law. Buffer = 10 mM sodium cacodylate.*

The titration of **BisA** alone at pH 6.8 is very similar to that at pH 5.5 (figure 2.35). Again showing that the absorbance at 260 nm with DNA would be expected to increase and that the absorbance at 248 nm is linear showing that **BisA** does not aggregate with itself at this pH.

The combined UV titration results are consistent with the observations made in CD and FRET titrations at these pH values, even though each experiment was carried out at different DNA concentrations. The UV results indicate that addition of **BisA** at pH 6.8 initially causes the DNA to unfold, after which precipitation or condensation occurs. At pH 5.5, there is no unfolding observed and the addition of **BisA** appears to cause immediate condensation. The UV titrations of **BisA** alone in buffer show that these observations are not as a result of **BisA** aggregating with itself. In combination with the experiments carried out in the previous sections, these results provide strong evidence that **BisA** causes the condensation of the DNA. To further support this working hypothesis, further experiments such as polyacrylamide gel electrophoresis and dynamic light scattering were used to look for evidence of a condensation effect.

### 2.3.7 Polyacrylamide Gel Electrophoresis Experiments

As described in section 1.4.7, gel electrophoresis experiments are a useful tool to show ligands interacting with DNA structures. As ligands bind to the DNA this increases the molecular weight of the resultant complex and if binding occurs through an electrostatic interaction, some charge neutralisation can occur. Both of these changes can result in the sample moving more slowly through the gel. Aggregation of the DNA would also be indicated in gel experiments as small single DNA structures are able to move through the gel much more easily than aggregates. 15% Polyacrylamide was used rather than agarose due to the small size of oligonucleotides used (18 – 24 bases), for which polyacrylamide can provide a better resolution. Native polyacrylamide gel electrophoresis experiments were performed with the labelled oligonucleotide hTeloC$_{FRET}$ so that the gels could be visualised by fluorescence imaging and an 18-base sequence of single stranded DNA was used as a marker (FAM-5'-GTAAAACGACGGCCAGTG-3'). The gels were run in pH 5.5 10 mM sodium cacodylate, 100 mM sodium chloride running buffer at 100 volts for 1 hour 15 minutes. The gels were carried out under these native conditions to reduce the possibility of denaturing the DNA and of breaking up any complexes formed. Denaturing gels typically include a component such as urea to disrupt the DNA secondary structure and are run at higher temperatures and voltages.[157]

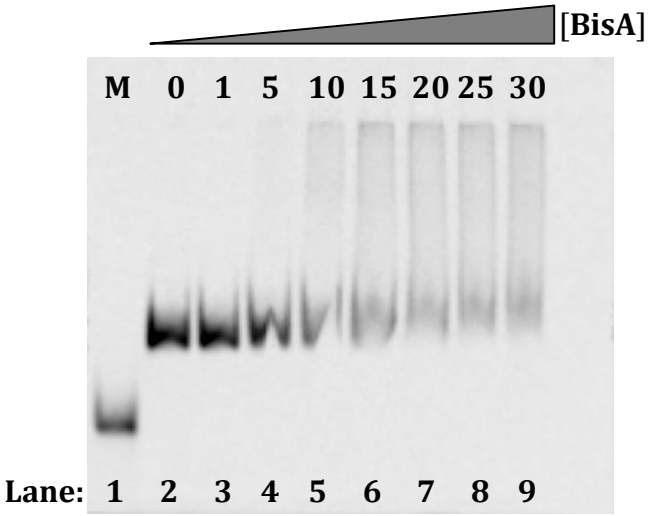



*Figure 2.37: **BisA** titration gel. Lane: 1 = 18 base marker, 2 = 1 μM hTeloC, lanes 3 – 9 = 1 μM hTeloC + 1, 5, 10, 15, 20, 25, 30 equivalents of **BisA**. [DNA] = 1 μM, buffer = 10 mM sodium cacodylate and 100 mM NaCl pH 5.5, Voltage = 100 V, Time = 1 hour 15 minutes.*

The effect of **BisA** on the mobility of the i-motif through the gel was tested by combining 1 μM hTeloC$_{FRET}$ with 0, 1, 5, 10, 15, 20, 25 and 30 equivalents of **BisA** in lanes 2 to 9. The 18 base marker was placed in lane 1 (figure 2.37). It can clearly be seen, that with increasing **BisA** concentration from left to right, the progress of the band through the gel is retarded and at higher **BisA** concentrations there is a smearing effect. This demonstrates that **BisA** is binding to the DNA and suggests that there is condensation of the DNA occurring as the smearing may be caused by different sized pieces of DNA moving through the gel at different rates. These results agree with earlier gel experiments carried out on **BisA** by Slama-Schwok and co-workers.[139] As described in section 2.1.1, their gels also showed a retardation and a smearing of the bands at **BisA** concentrations above 25 μM. In their experiments, aggregation happened slowly over the course of a day. However, in these experiments, condensation was observed more quickly in the time it took to add **BisA**, mix the sample and run the experiment which only lasted for 1 hour 15 minutes. The condensation effect may have been accelerated by centrifugation as the samples were centrifuged at a slow speed (3000 rpm) for 30 seconds to enable easier loading onto the gel.

Control gels were carried out with **MonoA** and the diethylenetriamine linker (**DETA**) (figure 2.38). 1 μM hTeloC$_{FRET}$ was combined with 0, 2, 10, 20, 30, 40, 50 and 60

equivalents of either **MonoA** or **DETA** in lanes 2 to 9 and the 18 base marker was again placed in lane 1. Double the concentrations were used for the experiments with **MonoA** and **DETA** in order to match the equivalents of each compound present in **BisA**. Here it can clearly be seen that the bands are not retarded and there is no difference in the bands compared to hTeloC alone. This suggests that **MonoA** and **DETA** are not interacting with the DNA or if they are, they are not affecting the progress of the DNA through the gel.

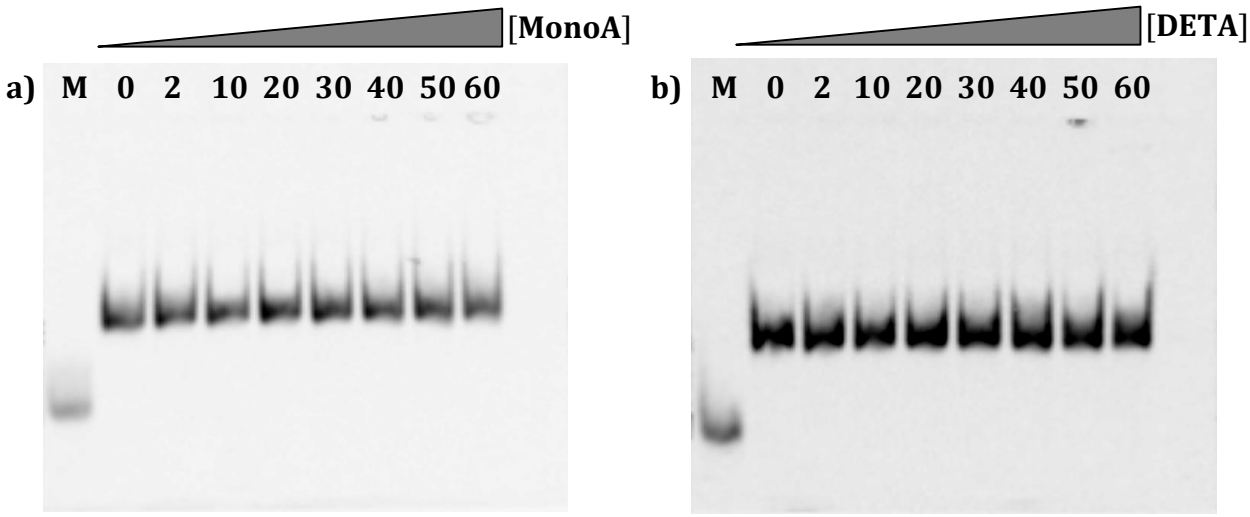



*Figure 2.38: a) **MonoA** titration gel b) **DETA** titration gel. Lane: 1 = 18 base marker, 2 = 1 μM hTeloC, lanes 3 – 9 = 1 μM hTeloC + 2, 10, 20, 30, 40, 50, 60 equivalents of **MonoA** or **DETA** respectively. [DNA] = 1 μM, buffer = 10 mM sodium cacodylate and 100 mM NaCl pH 5.5, Voltage = 100 V, Time = 1 hour 15 minutes.*

These results provide further evidence that **BisA** is binding to i-motif DNA but appears to indicate that higher concentrations of **BisA** cause condensation or aggregation of the DNA. The results of the gels in figure 2.38 show that neither **MonoA** nor **DETA** have this same condensation effect at equivalent concentrations to **BisA**. This implies that the observed effects are a particular characteristic of the **BisA** structure, rather than simply the presence of two acridine units or two **DETA** units.

### 2.3.8 Dynamic Light Scattering experiments

In order to characterise the possible condensation effect of **BisA**, dynamic light scattering was used to try and observe the presence of condensed particles. This

technique monitors the reflection of laser light from particles forming in the solution and uses this to measure the size of particles formed. This can indicate whether there is general aggregation occurring, forming a polydisperse mixture of different sized particles, or if there is more uniform condensation to a particular particle size.[158] By measuring the difference in scattering intensity over time, the instrument generates a correlation function (figure 2.39). The change in scattering over time is related to the size of the particles formed *via* the diffusion coefficient. The larger the particles, the slower they diffuse, hence the change in scattering is slower. By fitting the correlation curves with an exponential function an intensity distribution is calculated which is centered on the average particle size (figure 2.40). Looking at the correlation graph itself can also give useful qualitative information; if no particles form, the data given has a very low y-intercept, whereas particle formation is shown by a higher y-intercept closer to 1 and a sigmoidal curve. The time at which the inflection point occurs is indicative of the size of the particles formed, bigger particles at longer times. Finally the gradient of the sigmoidal curve is indicative of how polydisperse the sample is, i.e. the spread in the distribution of particle sizes. How flat the baseline is of the sigmoidal curve is indicative of the quality of the data as larger dust particles create an uneven baseline.

As outlined above for the UV absorbance experiments, the sample was centrifuged at a relatively slow speed (6000 rpm) before being tested and after each addition of **BisA**. For DLS experiments this is important as this process helps remove any much larger dust particles that may contaminate the sample and give an erroneous result. Dust contamination is also minimised by passing the sample through a 0.2 μm filter and by rinsing tubes and pipette tips with filtered buffer. A concentration of 2.5 μM hTeloC was tested at the same three pH's (pH 5.5, 6.8 and 7.4) that have been used throughout the previous experiments. In all three cases, the DNA alone, both in i-motif form at acidic pH and as a single strand at neutral pH, showed no particle formation (figure 2.39). However, addition of just 1 equivalent of **BisA** resulted in the formation of a polydisperse mixture of particles. This is indicated by the increase in the correlation function (figure 2.39) giving a sigmoidal curve. Interestingly at pH 5.5, the inflection point of the correlation curves occur at longer times (figure 2.39) indicating that larger particles are formed. This can be seen in the intensity plot (figure 2.40), where the distribution of particles is centered at different sizes depending on the pH. At lower pH

the size distribution of particles formed by DNA-**BisA** aggregation is centered on 600 nm in diameter with a second smaller distribution of particles centered around 100 nm. Whereas, at the higher pH's the population distribution is centered on the smaller 100 nm particles with fewer larger particles. This is most likely due to the extra charge neutralisation of the phosphate backbone that occurs at acidic pH. This enables the strands to come into close proximity with minimum repulsion whereas, at higher pH's, the phosphate backbone is still negatively charged and will cause electrostatic repulsion between the strands.

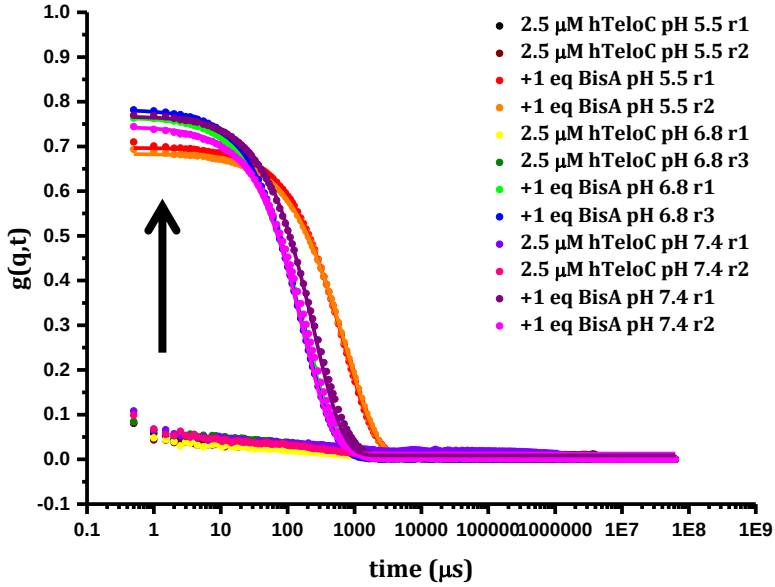



*Figure 2.39: Intensity correlation function graph of hTeloC in the absence and presence of 1 equivalent of **BisA** at pH 5.5, 6.8 and 7.4. [DNA] = 2.5 μM, buffer = 10 mM sodium cacodylate and 100 mM NaCl at pH 5.5 and 7.4, buffer = 10 mM sodium cacodylate at pH 6.8.*

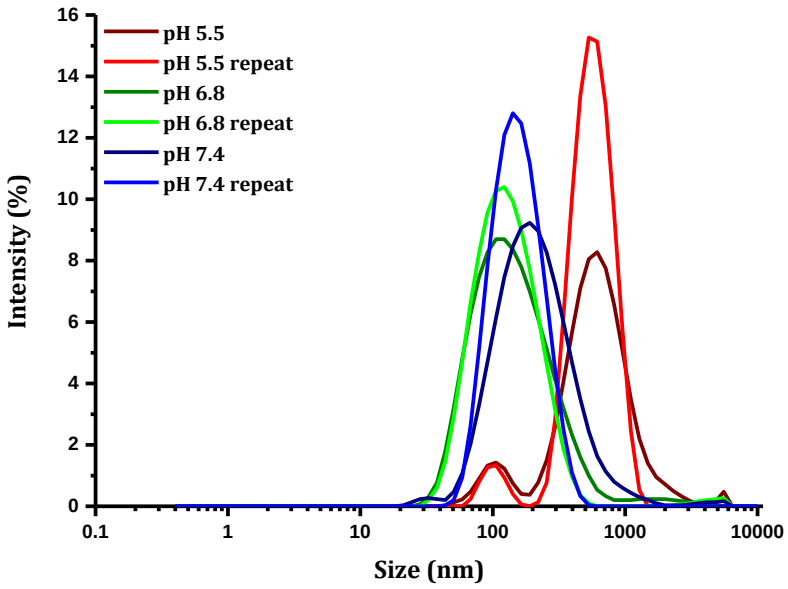



*Figure 2.40: Intensity distribution graph showing the size range of particles formed upon addition of 1 equivalent of **BisA** to human telomeric i-motif. [DNA] = 2.5 µM, buffer = 10 mM sodium cacodylate and 100 mM NaCl at pH 5.5 and 7.4, buffer = 10 mM sodium cacodylate at pH 6.8.*

In a titration experiment (figure 2.41) it can be seen that the size of particles formed increases when the concentration of **BisA** is increased up to 10 equivalents (25 µM) as is illustrated by the increase in time for the inflection point of the correlation curve.

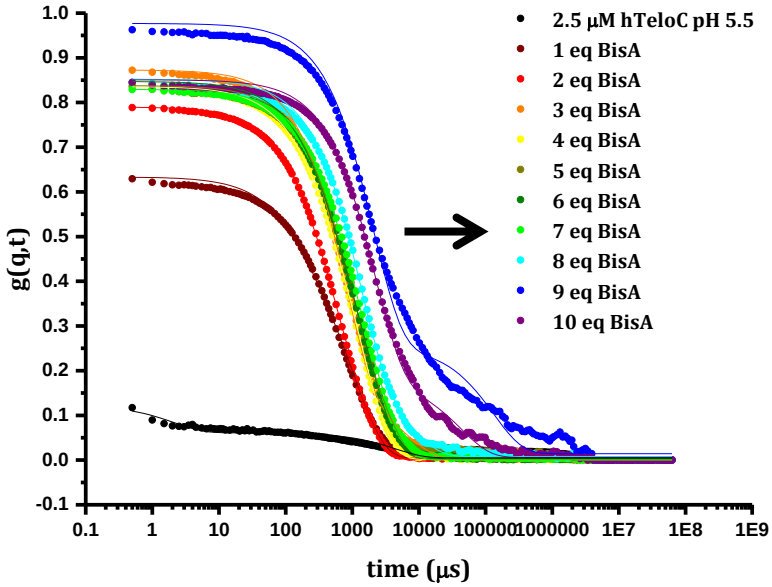



*Figure 2.41: Intensity correlation graph of hTeloC showing formation of increasing particle size with increasing concentration of **BisA**. [DNA] = 2.5 µM, buffer = 10 mM sodium cacodylate and 100 mM NaCl at pH 5.5.*

In analogous control experiments, addition of 1 equivalent of **MonoA** showed no particle formation (figure 2.42). This indicates that **MonoA** does not cause the DNA to condense into particles under the conditions of this experiment.

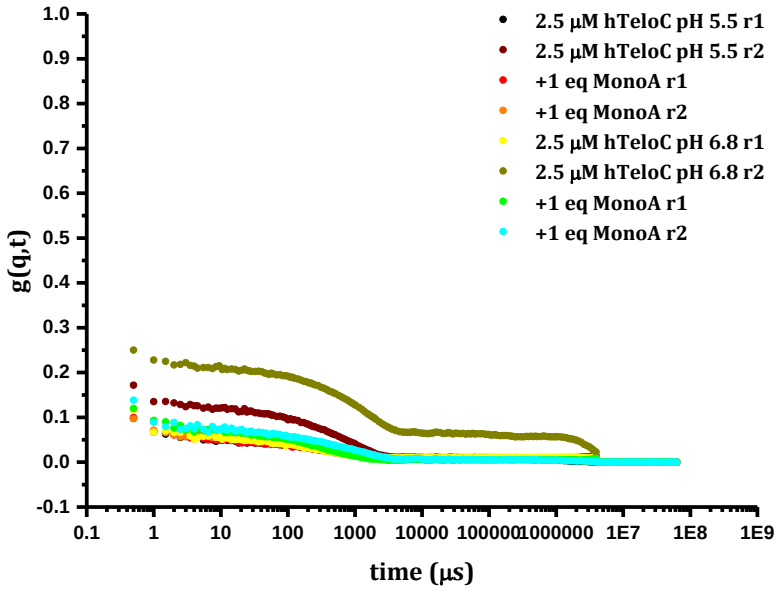



*Figure 2.42: Intensity correlation graph of hTeloC showing that 1 equivalent of **MonoA** does not enable particle formation. [DNA] = 2.5 µM, buffer = 10 mM sodium cacodylate and 100 mM NaCl at pH 5.5, buffer = 10 mM sodium cacodylate at pH 6.8.*

The results of the DLS experiments provide the strongest evidence of this investigation that **BisA** is causing condensation of the DNA. With these experiments it was possible to actually observe particle formation and to measure the size distribution of the particles. These results also demonstrate that on average, larger particles are formed at pH 5.5 compared to pH 6.8 and 7.4. This is likely to be due to the protonation of the phosphate backbone which will reduce electrostatic repulsion between strands. The fact that only 1 equivalent of **BisA** was required to cause a condensation effect suggests that **BisA** may not be a good ligand to study the potential chemical biology of the i-motif as it is able to cause condensation at low concentrations making it difficult for biological observations to be made.

## 2.4 Is this effect specific for i-motif?

Having seen that **BisA** may bind to the human telomeric i-motif but also appears to cause the DNA to condense, we next wanted to investigate whether this effect was specific for the i-motif or whether **BisA** would also cause condensation of telomeric G-quadruplex or double stranded DNA.

### 2.4.1 FRET Melting Studies

Analogous FRET melting experiments to those shown in section 2.3.2 for the i-motif, were performed for a human telomeric G-quadruplex and a double stranded sequence (hTeloG$_{FRET}$ and DS$_{FRET}$) at pH 5.5, 6.8 and 7.4. The sequences were tested with 0, 0.2, 0.4, 0.6, 0.8 and 1.0 μM concentrations of **BisA**. In each case, it can be seen that the fluorescence does decrease with increasing **BisA** concentration but not to the same extreme as with the i-motif sequence shown in section 2.3.2.

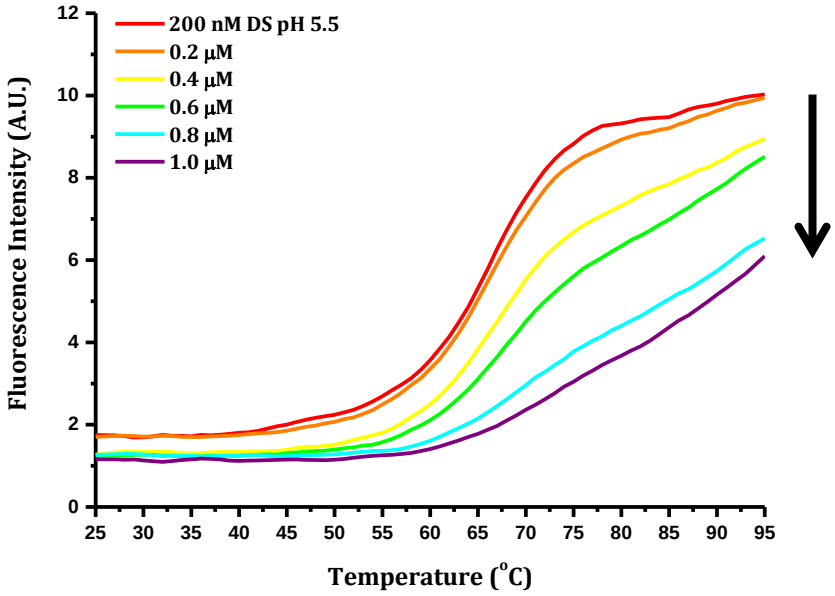



*Figure 2.43: FRET melting curves for duplex DNA with increasing concentration of **BisA** at pH 5.5. [DNA] = 200 nM, buffer = 10 mM sodium cacodylate and 100 mM NaCl.*

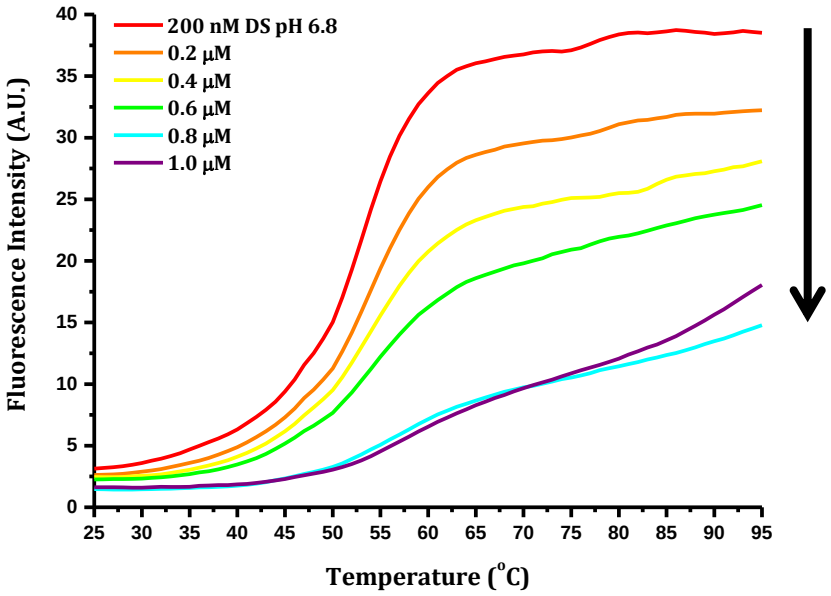



*Figure 2.44: FRET melting curves for duplex DNA with increasing concentration of **BisA** at pH 6.8. [DNA] = 200 nM, buffer = 10 mM sodium cacodylate.*

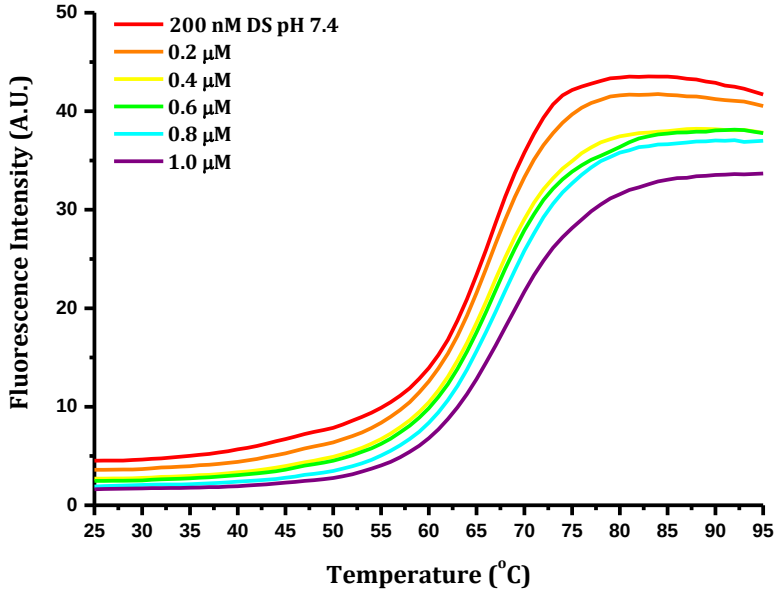



*Figure 2.45: FRET melting curves for duplex DNA with increasing concentration of **BisA** at pH 7.4. [DNA] = 200 nM, buffer = 10 mM sodium cacodylate and 100 mM NaCl.*

These results show that at pH 7.4 (figure 2.45), **BisA** has very little effect on the duplex stability ($\Delta T_m$ = 1.7°C at 1 μM). However at pH 5.5 and 6.8 (figure 2.43 and 2.44), while there is little change in the temperature of the melting transition there is a decrease in fluorescence suggesting that the DNA does not fully melt within the temperature range

of the experiment. This may be due to stabilisation of a different or partially folded structure, but taking into account the previous experiments on **BisA**, it is more likely to be a condensation effect.

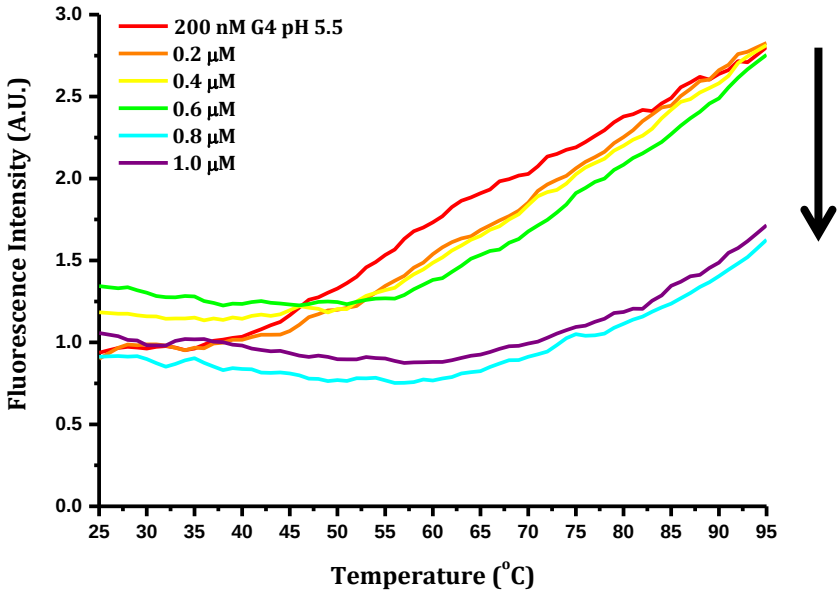



*Figure 2.46: FRET melting curves for human telomeric G-quadruplex with increasing concentration of **BisA** at pH 5.5. [DNA] = 200 nM, buffer = 10 mM sodium cacodylate and 100 mM NaCl.*

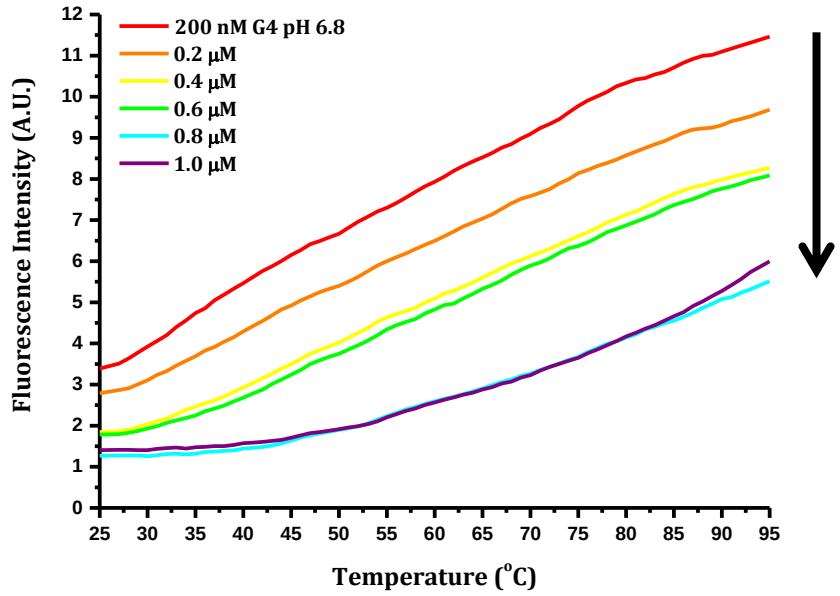



*Figure 2.47: FRET melting curves for human telomeric G-quadruplex with increasing concentration of **BisA** at pH 6.8. [DNA] = 200 nM, buffer = 10 mM sodium cacodylate.*

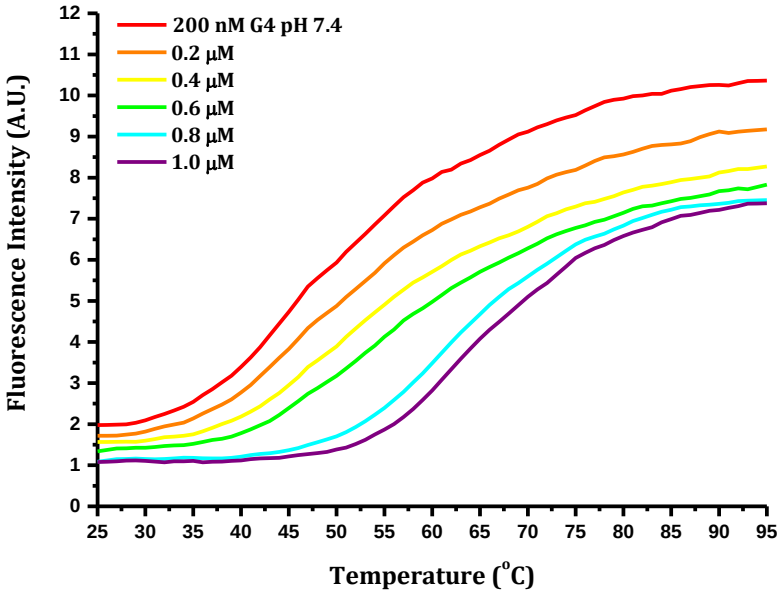



*Figure 2.48: FRET melting curves for human telomeric G-quadruplex with increasing concentration of **BisA** at pH 7.4. [DNA] = 200 nM, buffer = 10 mM sodium cacodylate and 100 mM NaCl.*

A similar pattern is observed for the G-quadruplex, at pH 5.5 and pH 6.8 (figure 2.46 and 2.47) there is a decrease in fluorescence at higher temperatures. However, at pH 7.4 (figure 2.48), sigmoidal melt curves are observed showing the expected increase in fluorescence upon melting. **BisA** clearly causes an increase in melting temperature ($\Delta T_m$ = +16°C at 1 μM), which is consistent with the results observed by Alberti and co-workers ($\Delta T_m$ = +15°C).[100] These results suggest that **BisA** may have a similar condensation effect on the G-quadruplex and duplex sequences used in this study but only under acidic conditions, as at neutral pH, normal melt curves were observed. However, the G-quadruplex does not have good sigmoidal melting curves at acidic pH even in the absence of **BisA** so the structure maybe less stable or only partially formed.

### 2.4.2 Polyacrylamide gel electrophoresis experiments

To further understand the interaction between **BisA**, duplex and G-quadruplex DNA, a native polyacrylamide gel was used (figure 2.49). This enables a direct visual comparison of the effect of **BisA** on i-motif, G-quadruplex and double stranded DNA at pH 5.5. Although double stranded and G-quadruplex structures are normally examined under neutral pH conditions, this would result in the i-motif sequence unfolding so the gel needed to be carried out at acidic pH. This is also the pH at which evidence of

condensation has already been observed (figure 2.37). An 18 base marker was placed in lane 1, as before. Lanes 2-4 contained hTeloC i-motif, lanes 5-7 contained double stranded DNA and lanes 8-10 contained hTeloG G-quadruplex, each with 0, 1 and 30 equivalents of **BisA** respectively. It can be seen that 1 equivalent of **BisA** has little to no effect on the movement of the band for each type of DNA, but at higher **BisA** concentrations, a retardation and smearing of the band can be seen for each type of DNA. In particular the band in lane 10 is very faint as a result. This suggests that the condensation caused by **BisA** is not specific to the i-motif.

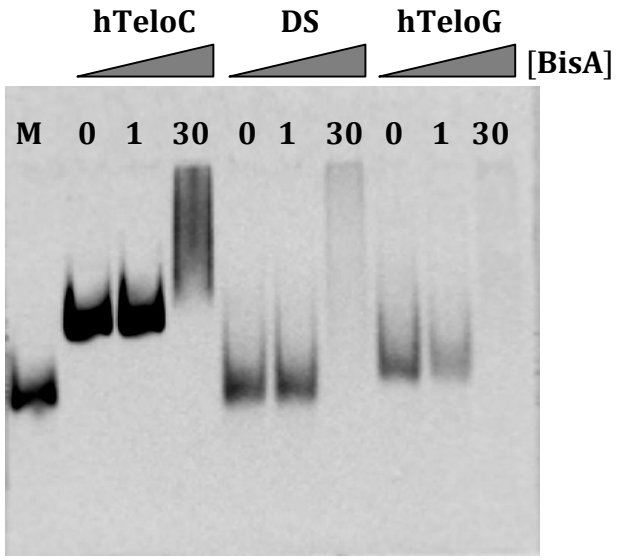



*Figure 2.49: Combined i-motif, G-quadruplex and DS gel at pH 5.5. Lane 1 = 18 base marker. Lane 2 - 4 = 1 μM hTeloC +0 eq, 1 eq & 30 eq **BisA**. Lane 5 – 7 = 1 μM duplex +0 eq, 1 eq & 30 eq **BisA**. Lane 8 – 10 = 1 μM hTeloG +0 eq, 1 eq & 30 eq **BisA**. [DNA] = 1 μM, buffer = 10 mM sodium cacodylate and 100 mM NaCl.*

The gel shown in figure 2.49 provides further evidence to compliment the FRET results that **BisA** binds and causes condensation of both G-quadruplex and duplex DNA at acidic pH. Like the i-motif, this is probably a result of the combination of a bis-intercalator as well as the charge neutralisation provided by the acidic conditions. The bis-intercalation properties of **BisA** enable it to interact with two separate strands simultaneously and bring them into close proximity. The acidic conditions protonate the phosphate back bone which would otherwise cause electrostatic repulsion when the strands are brought together.

## 2.5 Conclusions

In this chapter, the previously identified ligand **BisA** was synthesised and its interaction with the i-motif has been investigated in detail with a much wider range of experiments than previously published.[100] First the FRET melting experiments carried out by Alberti and co-workers were repeated and a high stabilisation potential ($\Delta T_m$ = +55°C) was observed after normalisation of the data, greater than the published result ($\Delta T_m$ = +33°C).[100] However, the raw data suggested that **BisA** may actually destabilise rather than stabilise the i-motif as indicated by the first derivative where the minimum which corresponds to the melting temperature can be seen to decrease with increasing concentration of **BisA**. Higher concentrations indicated possible aggregation and precipitation with a base level fluorescent signal throughout the temperature range.

Circular dichroism was used to look at the interaction in more detail. This gave further evidence of **BisA** binding to the DNA and showed that it caused a change in structure. At pH 6.8 this change in structure appeared to transition *via* a partial unfolding event whereas at pH 5.5 the change occurred without unfolding. In both cases a new peak appeared in the spectrum at 242 to 245 nm which could be due to the change in structure or as a result of an induced CD signal from **BisA** which absorbs in a similar region (248 nm). By fitting this change in CD binding constants could be obtained suggesting that **BisA** bound to the DNA with a moderate affinity in the low micromolar range. The observations made in the CD titrations were supported by FRET and UV titration experiments, both of which indicated that **BisA** caused condensation of the DNA but that at pH 6.8 the DNA began to unfold first. A fluorescence emission titration of **BisA** with i-motif DNA provided another measure of the binding affinity under conditions which were unlikely to favour condensation (excess DNA). This suggested that **BisA** bound to the i-motif in a 2 to 1 model with a stronger first binding site ($K_{D1}$ = 2.1 μM) and a weaker second binding site ($K_{D2}$ = 146 μM). Finally, further evidence that **BisA** caused condensation of the DNA was provided by polyacrylamide gel electrophoresis experiments and dynamic light scattering. This experiment enabled the particles formed to be measured, showing that larger particles (~600 nm) form at pH 5.5 while smaller particles (~100 nm) form at pH 6.8 and pH 7.4. Most likely due to

the protonation of the phosphate backbone reducing electrostatic repulsion as a barrier to condensation.

The DNA condensation effect of **BisA** was not restricted to just the telomeric i-motif sequence. FRET melting studies and gel electrophoresis showed similar effects for **BisA** with double stranded and G-quadruplex DNA but to a lesser extent. At different stages in the study, the effects of **BisA** were compared to the monoacridine compound **MonoA**. These experiments showed that it had little to no effect on the i-motif with no change in melting temperature, and little effect on the structure by CD other than to cause unfolding at pH 6.8 when the i-motif is already unstable. **MonoA** also showed no aggregation effect in gel or DLS experiments. The diethylenetriamine linker **DETA** also had no effects when tested on its own in FRET melting and gel experiments. Conversely other bisacridine analogues that are linear rather than cyclic, do appear to cause some condensation or precipitation effects in FRET melting and CD experiments. The combination of these results suggests that it is not simply the charge of **BisA** (represented by **DETA**) that causes condensation and not simply the acridine unit itself but the combination of both these entities in the **BisA** structure.

These initial investigations of the interaction of **BisA** with the i-motif highlight some of the potential pitfall of converting FRET melting data to fraction folded in order to measure the melting temperature as, unless the data gives a sigmoidal curve, this can cause effects such as condensation or precipitation to be misinterpreted as stabilisation of the structure. Whilst these results indicate that **BisA** is not a good choice of ligand for investigating the potential biological effects of i-motif DNA, the DNA condensation properties of this molecule could make it a very useful tool in order to try and obtain a crystal structure. Further work based on this chapter would be to try and use **BisA** to form a DNA-ligand crystal at acidic pH. This would show whether the DNA changed structure upon condensing or whether multiple i-motif structures condensed together. It would also show where the **BisA** molecule is bound to the DNA, enabling the identification of potential binding sites to aid the design of future i-motif ligands.

# Chapter 3: Identification of i-Motif Binding Compounds Using a FRET Melting Based Screen

# 3.1 Introduction

### 3.1.1 The need for new i-motif binding ligands.

Currently, there are very few i-motif binding ligands published in the literature. As discussed in section 1.5, the majority of i-motif ligands that have been identified often have stronger binding affinities for duplex or G-quadruplex DNA and impart little stabilisation on the i-motif structure. Despite this, there have been some good ligands which have enabled an insight into the potential biological functions of the i-motif. In particular, carboxyl-modified single-walled carbon nanotubes have shown that inhibition of telomerase is possible through stabilisation of the telomeric i-motif.[59] Whilst Laurence Hurley and co-workers have identified a ligand specific for the Bcl-2 i-motif which increases expression of the gene.[69] These results provide strong evidence that the i-motif is worthy of investigation as a potential therapeutic target. However, the i-motif is inherently unstable due to the need for hemiprotonation of cytosine and *in vitro*, does not usually form at neutral pH. In order to fully understand the i-motif structure, new high affinity ligands are needed which are able to stabilise the structure sufficiently under physiological conditions.

### 3.1.2 Aims and Objectives.

In this chapter, a 960 compound library was screened against the human telomeric i-motif *via* a medium throughput FRET melting assay with the aim of identifying new i-motif interacting ligands based on their ability to stabilise or destabilise the structure. The library was a very diverse group of known drugs and biologically active compounds, with a wide range of structure types, molecular weights and chemical properties. Such a diverse set of compounds was deliberately chosen because there are so few existing ligands and as yet there is no obvious i-motif binding chemotype. Therefore, it was important to cover a wide range of options as possible i-motif ligands.

## 3.2 The initial screen

### 3.2.1 Library and screening method

The screen was carried out in collaboration with Camille Huguin using the Gen-Plus library from Microsource Discovery Systems Inc. consisting of 960 drug standards with approval in Europe, Japan or the USA supplied as 10 mM solutions in DMSO which were diluted to 1 mM in 96 well plates. The compounds were screened against the human telomeric i-motif (200 nM) in pH 5.5 buffer with 10 mM sodium cacodylate and 5 mM sodium chloride. These conditions were chosen as they are the same as previous FRET melting studies on the i-motif which have been demonstrated in the literature.[98] Sodium cacodylate is an ideal choice of buffer for melting experiments because its p$K_a$ remains constant within the temperature range of the experiment.[93] pH 5.5 was chosen as it offers the best compromise to ensure formation of a stable i-motif structure, whilst not being so acidic as to completely quench the fluorescence of the fluorophore. A higher pH such as pH 7.4 was not chosen because at this pH the DNA is unfolded and although a decrease in fluorescence might be observed, which would indicate folding of the DNA, the FRET melting data itself does not give information about the type of structure that is formed. The primary aim of this study is to identify ligands which bind to the i-motif structure and therefore conditions were chosen in order to be certain that a stable i-motif is present in solution as a binding target.

The compounds were screened at a relatively high concentration of 100 μM. This is because there are so few existing i-motif ligands in the literature that it is still not clear what particular structure types and characteristics favour i-motif binding. As the library chosen was a diverse library it was important to choose a high initial concentration in order to increase the chance of observing an interaction and to maximise the number of possible hits. These initial hits can then be further refined to find the best compounds through further experiments. After dilution in buffer, the samples still contain 10% DMSO, which has a destabilising effect on the i-motif. Therefore, changes in melting temperature ($\Delta T_m$) were calculated in relation to a control sample containing i-motif in the presence of 10% DMSO.

The potential draw-backs of this screening method are that compounds which bind to the i-motif but do not alter its melting temperature may be over looked. This screen also relies on measuring changes in fluorescence of the labelled DNA, therefore compounds which interact with the fluorophores rather than the i-motif may be incorrectly identified as hits. The FRET melting screen does however, offer several advantages. By using a fluorescently labelled oligonucleotide the screen can be carried out at much lower DNA concentrations (200 nM) than for example a UV based assay(>1 μM). This is more cost effective and enables compounds to be tested at much higher ligand:DNA ratios easily. Using this assay, the samples could also be tested in a medium throughput manner, enabling the compounds to be screened in a short period of time.

### 3.2.2 Hit identification

Normalisation of data is a technique where the raw data is transformed so that the lowest point is equal to zero and the highest point is equal to 1, this simplifies the interpretation of the data and makes it easier to compare different data sets. By assuming that the lowest fluorescence corresponds to completely folded DNA and the highest corresponds to completely unfolded DNA, the normalisation of FRET melting data gives a graph representing the fraction of folded DNA. The melting temperature is defined as the point of mid transition and on this basis the melting temperature ($T_{1/2}$) is measured at the point where the fraction folded = 0.5, the point at which 50% of the DNA is folded and 50% is unfolded. As described in section 2.3.2, it is not always appropriate to assume that the DNA is completely folded or unfolded and using this method can lead to falsely identified changes in melting temperature if the data is not a sigmoidal curve. Particularly if the data gives a low fluorescence signal across the whole temperature range. Therefore, as well as measuring the $T_{1/2}$, the data from the screen was analysed by taking the first derivative. As discussed in section 2.3.2, the minimum in this graph represents a point of inflection in the melting curve which should represent the point of mid transition and hence the melting temperature ($T_m$).[91] This method is only valid for intramolecular systems and still may not represent the precise melting temperature. However, this method avoids the possible errors that normalisation of the data can introduce. The third method for determining the melting temperature, the “base line” method, is not suited to the analysis of melting curves where one or both baselines cannot be accurately determined.

The melting temperature derived from the first derivative method was assessed in combination with the raw melt curves to see that the melting point chosen made sense and did indeed reflect a transition in the melt curve. As the compounds were tested at quite a high concentration, this resulted in a number of them giving a base level fluorescence signal across the whole temperature range and not showing a melting transition at all. As has been illustrated by the previous chapter, this may be the result of condensation or precipitation of the DNA, however this may also be due to a highly stabilising effect of the ligand or an interaction with the fluorophore and cannot be verified without further experiments. Therefore for the purpose of hit identification, ligands which gave a base level fluorescence signal were assumed to have a $T_m$ > 95 °C and were included as a hit. Initially 34 hits were identified which gave any degree of change in melting temperature. The hit compounds gave a wide spread in $\Delta T_m$ values from -13.7 to +47 °C (table 3.1) and included a wide variety of different structures (figure 3.1). The only trends that were observed between their stabilisation potential ($\Delta T_m$) and any structural or chemical properties was that positive charge and a higher number of hydrogen bond donors tended to favour stabilisation. This is most likely due to the fact that the i-motif structure has two very narrow grooves where the phosphate backbones come into close proximity, the sugar phosphate backbone has a large number of oxygens which can act as hydrogen bond acceptors so molecules with a larger number of hydrogen bond donors will have more opportunity to form favourable binding interactions with the structure.

| | Compound | $\Delta T_{1/2}$ (°C)[a] | $\Delta T_m$ (°C)[b] |
|---|---|---|---|
| P2C5 | **Cetylpyridinium chloride** | 49 | 47 |
| P9C11 | **Alexidine hydrochloride** | 35 | 40 |
| P1A3 | **Chlorhexidine** | 31 | 38 |
| P9G8 | **Methylbenzethonium Chloride** | 26 | 25.7 |
| P3A4 | **Tyrothricin** | 24 | 24 |
| P9F11 | **Pararosaniline Pamoate** | 19 | 24.7 |
| P12B3 | **Amiodarone hydrochloride** | 19 | 7.7 |
| P12A6 | **Astemizole** | 18 | -7.6, 19[c] |
| P4B6 | **Tamoxifen Citrate** | 16 | 12.3 |
| P7D5 | **Bromocriptine mesylate** | 15 | 17.3 |
| P4E9 | **Hexetidine** | 14 | 11.3 |
| P11H11 | **Tilorone** | 13 | 12.3 |
| P10B5 | **Benzalkonium chloride** | 12 | 8 |
| P12C5 | **Flunarizine hydrochloride** | 11 | 13.3 |
| P9H8 | **Mitoxantrone** | 9 | 41.3 |
| P10C8 | **Clopidogrel sulfate** | 9 | 5.3 |
| P8F9 | **Fosfosal** | 7 | 5.3 |
| P2A6 | **Roxarsone** | 6 | 6 |
| P12F6 | **Pyrvinium** | 6 | 5.3 |
| P6H5 | **Halazone** | 5 | 3.7 |
| P8G9 | **Cefamandole sodium** | 5 | 4.7 |
| P1G6 | **Nicotine ditartrate** | 3 | 5 |
| P12G3 | **Geneticin** | 3 | 7 |
| P1C9 | **Sisomicin sulfate** | 1 | 4 |
| P8F8 | **Diacerin** | -3 | -7.3 |
| P11F5 | **Sertraline hydrochloride** | -6 | -9.3, 27.3[c] |
| P3A2 | **Neriifolin** | -6 | -2.3 |
| P5E4 | **Oxacillin sodium** | -7 | -10, 4.7[c] |
| P3A5 | **1,4-Naphthoquinone** | -7 | -6.7 |
| P3A3 | **Diazoxide** | -8 | -7 |
| P3A8 | **Rosolic acid** | -9 | -7 |
| P6H7 | **Hexachlorophen** | -9 | -9 |
| P6H6 | **Haloperidol** | -11 | -11.3 |
| P6G6 | **Gallamine** | -14 | -13.7 |

*Table 3.1: Melting temperatures of initial 34 hit compounds.* [a] *$T_{1/2}$ = melting temperature from the normalised data at fraction folded = 0.5, error = ± 1°C.* [b] *$T_m$ = melting temperature from the minimum of the first derivative, error = ± 0.3°C.* [c] *A bimodal melting curve was observed so both melting temperatures were recorded.*

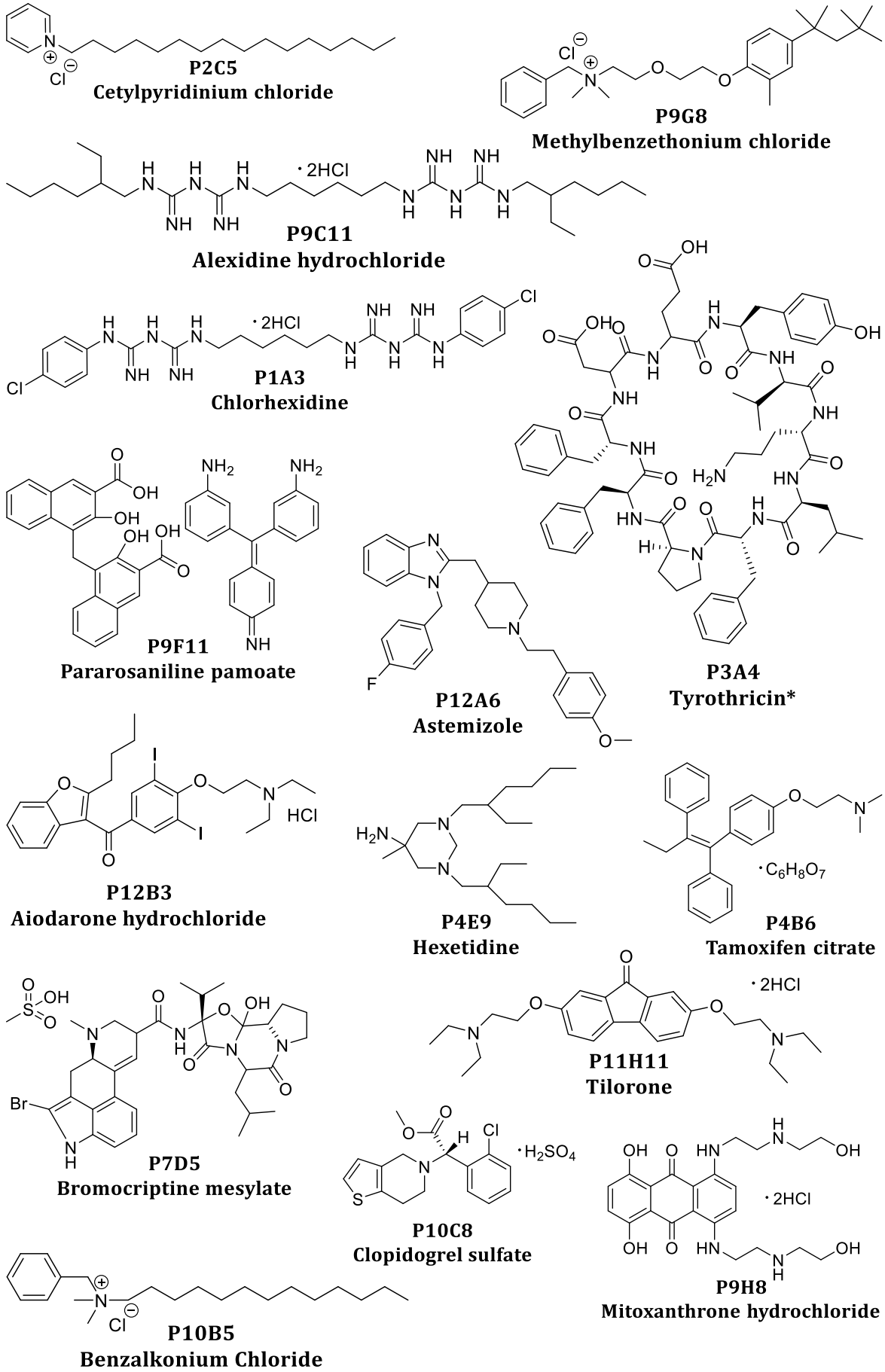



*Figure 3.1: Structures of the 34 hit compounds.* **Tyrothricin** is a peptide mixture.*

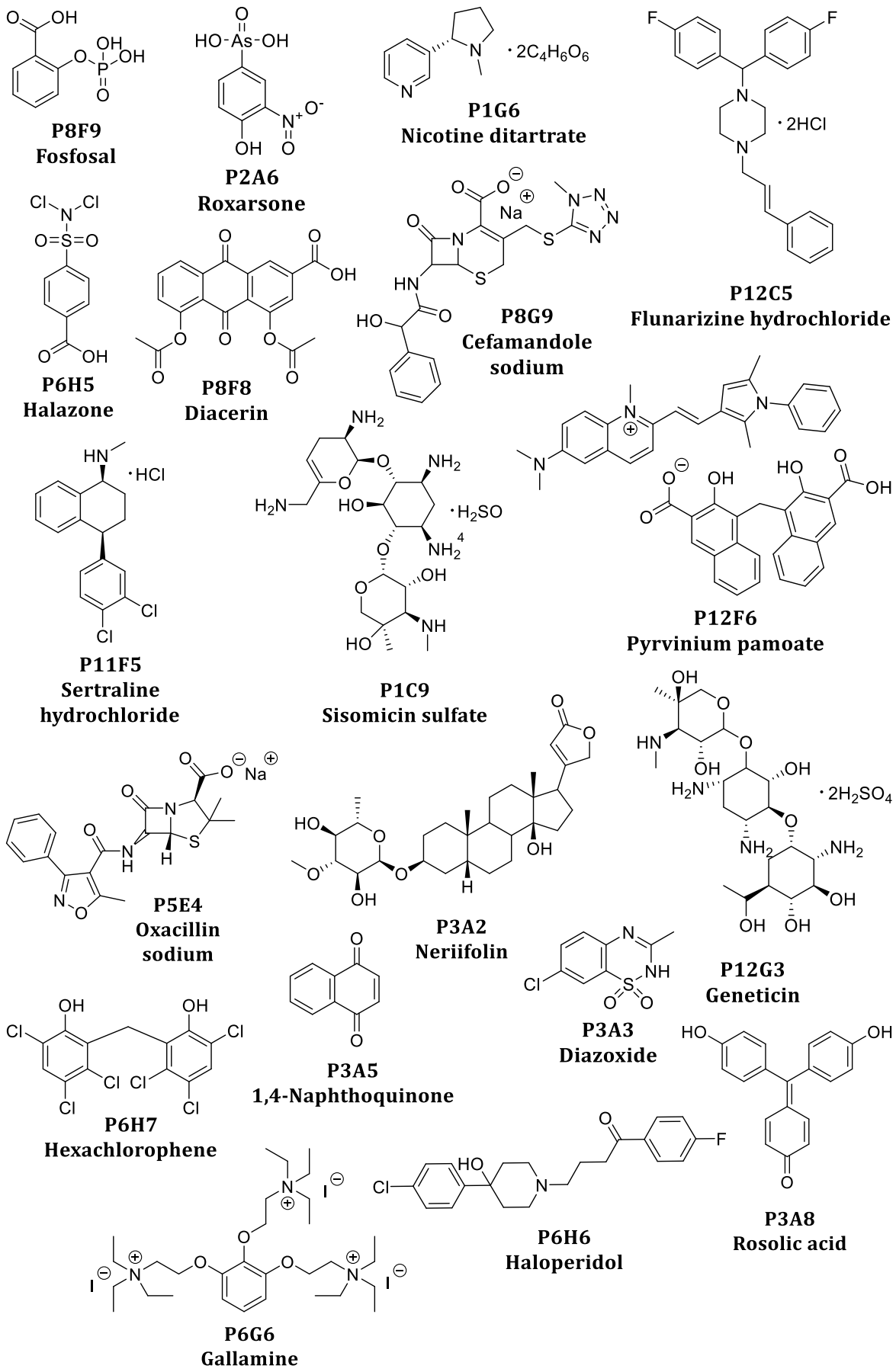
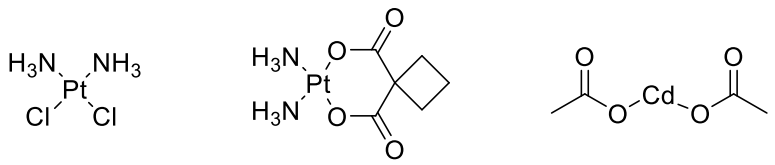



*Figure 3.1: Structures of the 34 hit compounds.* **Tyrothricin** is a peptide mixture.*

### 3.2.3 Potential False Positives and Negatives

As well as the 34 hits, a number of false positives were also identified (figure 3.2). Although these compounds gave base level fluorescence across the temperature range which could be due to a high degree of stabilisation, looking at the structures they are known DNA cross-linking agents and were therefore discounted as the high stabilisation is likely to be as a result of cross-linking[159]. Cadmium acetate also gave a base level fluorescence signal but is not a small molecule so was also discounted.




**P2A5**
**Cisplatin**

**P2D8**
**Carboplatin**

**P2H4**
**Cadmium acetate**

*Figure 3.2: Structures of false positives **Cisplatin** and **Carboplatin**, known DNA cross-linking agents, as well as **cadmium acetate**.*

There were also a range of compounds which gave a maximal fluorescence across the whole temperature range. Looking at their structures (figure 3.3) it is likely that these compounds are able to interact with the fluorophores of the labelled oligonucleotide or in the case of **P3A6** may fluoresce themselves under the same excitation conditions of this experiment due to its similar core structure to FAM. The increase in fluorescence observed is greater than that observed in other experiments for simple unfolding of the DNA, giving further support to a possible interaction with the fluorophores. This means it is not possible, under these experimental conditions, to tell whether they are interacting with the i-motif at all or what effect they are having and hence are potential false negatives. To investigate these compounds further, a non-fluorescence based assay would need to be used.

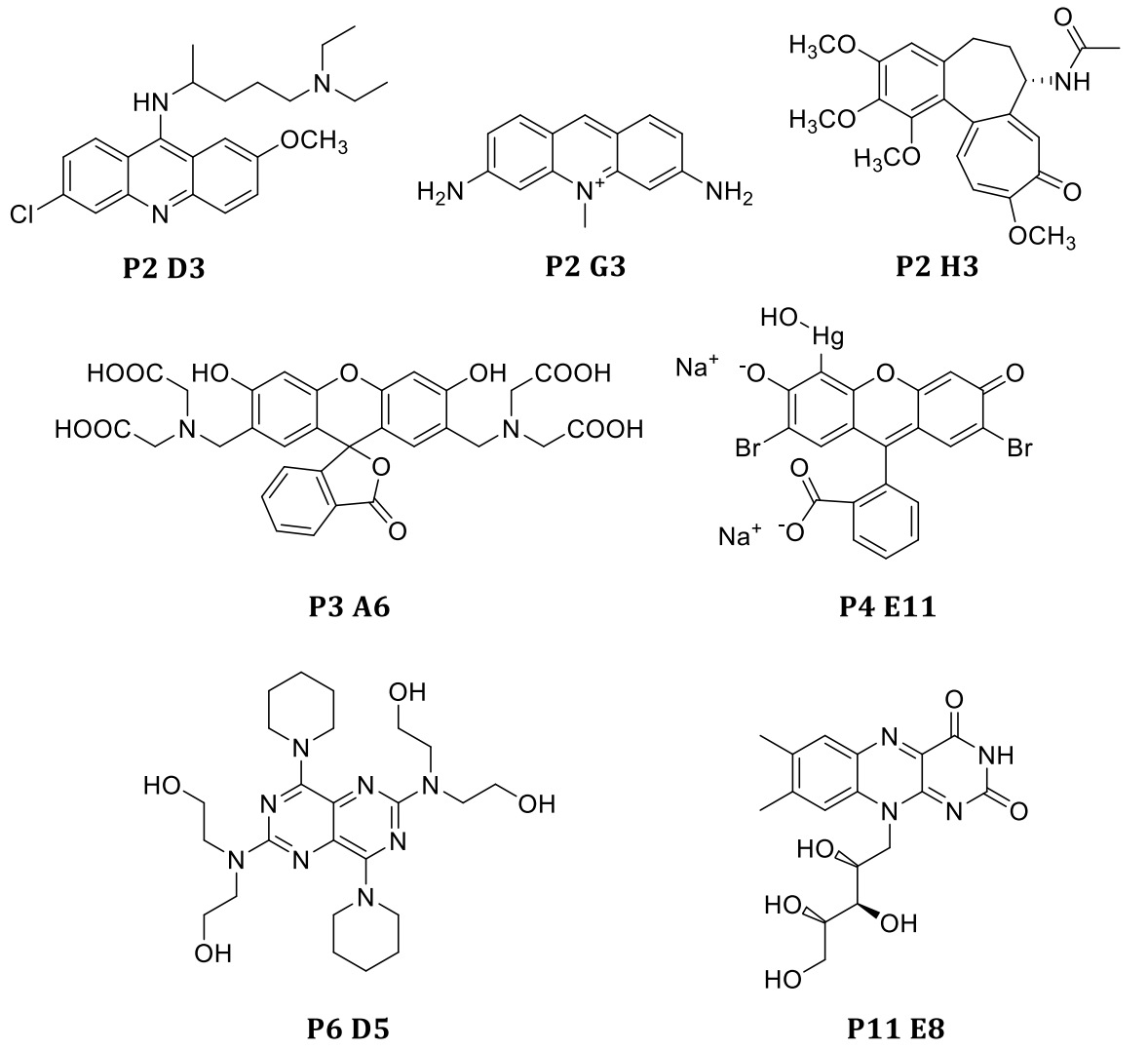



*Figure 3.3: Structures of potential false negative compounds which gave high fluorescence signals.*

### 3.2.4 Repeat Testing of Initial Hits

Next, the 34 hit compounds were tested again in 100 mM sodium chloride and 5 mM sodium chloride buffers to confirm the results of the initial screen. Both salt concentrations were tested because the i-motif is more stable with lower salt concentration but the higher salt concentration can reduce the effect of more non-specific electrostatic interactions. Many compounds showed a smaller effect on repeat screening (table 3.2). This may be as a result of regression toward the mean, a common occurrence in high throughput screens.[160] This effect occurs because initial screen data may contain more error as a result of testing a large number of compounds. The repeat results are likely to have less error as a smaller number of compounds have been tested.[161]

| | Compound | Initial screen $\Delta T_m$ (°C) | 100 mM $\Delta T_m$ (°C) | 5 mM repeat $\Delta T_m$ (°C) |
|---|---|---|---|---|
| P2C5 | **Cetylpyridinium chloride** | 47 | $T_m$ >95 | $T_m$ >95 |
| P9C11 | **Alexidine hydrochloride** | 40 | $T_m$ >95 | $T_m$ >95 |
| P1A3 | **Chlorhexidine** | 38 | 3, 25[a] | $T_m$ >95 |
| P9G8 | **Methylbenzethonium Chloride** | 25.7 | -1, 31.7[a] | 38 |
| P3A4 | **Tyrothricin** | 24 | 22.3 | 41 |
| P9F11 | **Pararosaniline Pamoate** | 24.7 | 0.7 | $T_m$ >95 |
| P12B3 | **Amiodarone hydrochloride** | 7.7 | 0.3, 10[a] | 7.3 |
| P12A6 | **Astemizole** | -7.6, 19[a] | -0.7 | 0.3 |
| P4B6 | **Tamoxifen Citrate** | 12.3 | 1 | 21.7 |
| P7D5 | **Bromocriptine mesylate** | 17.3 | -1 | $T_m$ >95 |
| P4E9 | **Hexetidine** | 11.3 | 1 | 14.3 |
| P11H11 | **Tilorone** | 12.3 | 0 | 50 |
| P10B5 | **Benzalkonium chloride** | 8 | -1.7, 28[a] | 38 |
| P12C5 | **Flunarizine hydrochloride** | 13.3 | 0 | 0.7 |
| P9H8 | **Mitoxantrone** | 41.3 | 33.3 | $T_m$ >95 |
| P10C8 | **Clopidogrel sulfate** | 5.3 | 2.7 | 1.7 |
| P8F9 | **Fosfosal** | 5.3 | 1.3 | 1 |
| P2A6 | **Roxarsone** | 6 | 1.3 | 5 |
| P12F6 | **Pyrvinium** | 5.3 | -0.3 | 1 |
| P6H5 | **Halazone** | 3.7 | 2 | 1 |
| P8G9 | **Cefamandole sodium** | 4.7 | 0 | 0.7 |
| P1G6 | **Nicotine ditartrate** | 5 | 2 | 1.3 |
| P12G3 | **Geneticin** | 7 | 0.3 | 6.3 |
| P1C9 | **Sisomicin sulfate** | 4 | 0 | 6.3 |
| P8F8 | **Diacerin** | -7.3 | 0.7 | 0.3 |
| P11F5 | **Sertraline hydrochloride** | -9.3, 27.3[a] | 0 | -0.3 |
| P3A2 | **Neriifolin** | -2.3 | 0.3 | 0.3 |
| P5E4 | **Oxacillin sodium** | -10, 4.7[a] | 0 | 0.3 |
| P3A5 | **1,4-Naphthoquinone** | -6.7 | -0.3 | 0.3 |
| P3A3 | **Diazoxide** | -7 | 0.3 | 0 |
| P3A8 | **Rosolic acid** | -7 | 0.3 | 0.3 |
| P6H7 | **Hexachlorophen** | -9 | -0.3 | 0.7 |
| P6H6 | **Haloperidol** | -11.3 | -1.7 | -0.3 |
| P6G6 | **Gallamine** | -13.7 | 0 | -3.7 |

*Table 3.2: Melting temperatures of 34 hit compounds when in 100 mM and 5 mM NaCl, 10 mM sodium cacodylate buffers at pH 5.5.* [a] *A bimodal melting curve was observed so both $T_m$'s were recorded. Error = ± 0.3°C.*

### 3.2.5 Results

As the compounds had initially been screened at a relatively high ligand concentration in order to give the best chance of seeing an effect, a minimum change in melting temperature criteria of ±5 °C was set in order to narrow down the hits to ones with a significant effect. This reduced the initial 34 hits to 13 lead compounds (figure 3.4). Here it can be seen that some of the compounds are only stabilising in 5 mM sodium chloride buffer, while others are highly stabilising in both salt concentrations. The structures of the hits are shown in figure 3.5, there are a variety of different structure types with a range of different sizes and properties.

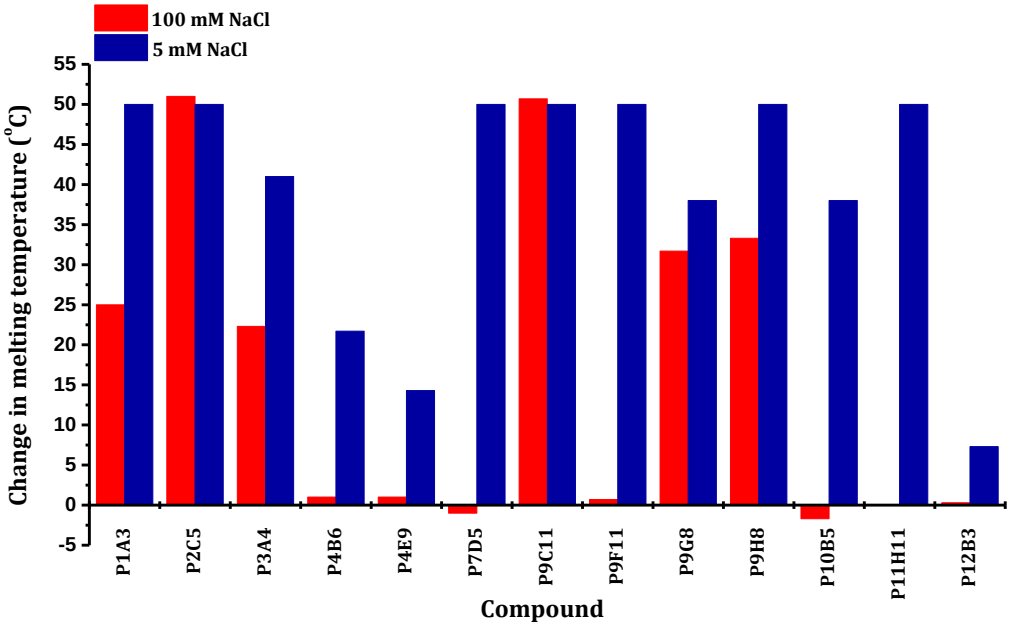



*Figure 3.4: Comparison of the $\Delta T_m$'s of the 13 hit compounds in both 100 mM and 5 mM sodium chloride buffer. DNA = 200 nM hTeloC, buffer = 10 mM sodium cacodylate and 100 mM or 5 mM NaCl at pH5.5.*

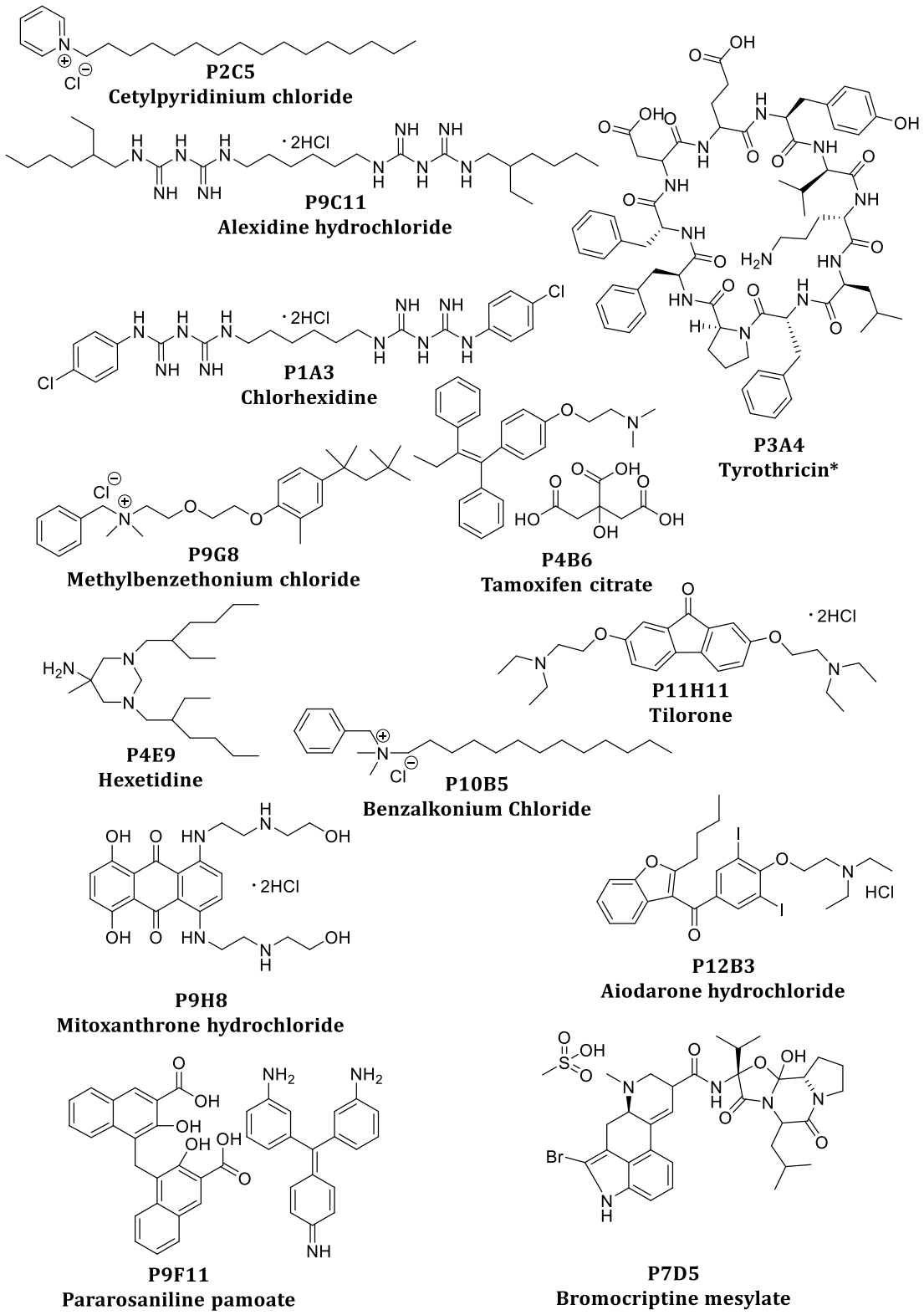



*Figure 3.5: Structures of the 13 lead hit compounds.* **Tyrothricin** is a peptide mixture.*

It should be noted that hit **tyrothricin**, is actually a mixture of two different classes of peptides called gramicidins and tyrocidines. The gramicidins are linear and make up approximately 37% of the mixture while the tyrocidines are cyclic and make up 62% of the mixture.[162] Gramicidin was tested separately as part of the screen and did not

come out as a hit so it is likely that the resultant stabilisation of the i-motif is due to the tyrocidine component but **tyrothricin** will continue to be tested as a mixture here for the purposes of hit validation.

## 3.3 Dose Response Studies

### 3.3.1 Dose response versus telomeric i-motif with 100 mM NaCl

To investigate the stabilisation potential of the hit compounds further, a range of concentrations between 0 and 100 μM were tested for each potential ligand (figure 3.7 – 3.12 and appendix section A.1). Care had to be taken in interpreting the raw data as the different concentrations are prepared in buffer *via* serial dilution from stock solutions in DMSO and hence each concentration has a different percentage DMSO content. Therefore analogous dose response studies of just DMSO in buffer were carried out and the change in melting temperature for each ligand at each concentration was calculated in reference to the respective DMSO control (table 3.3). An example DMSO dose response showing the destabilising effect of DMSO is shown in figure 3.6.

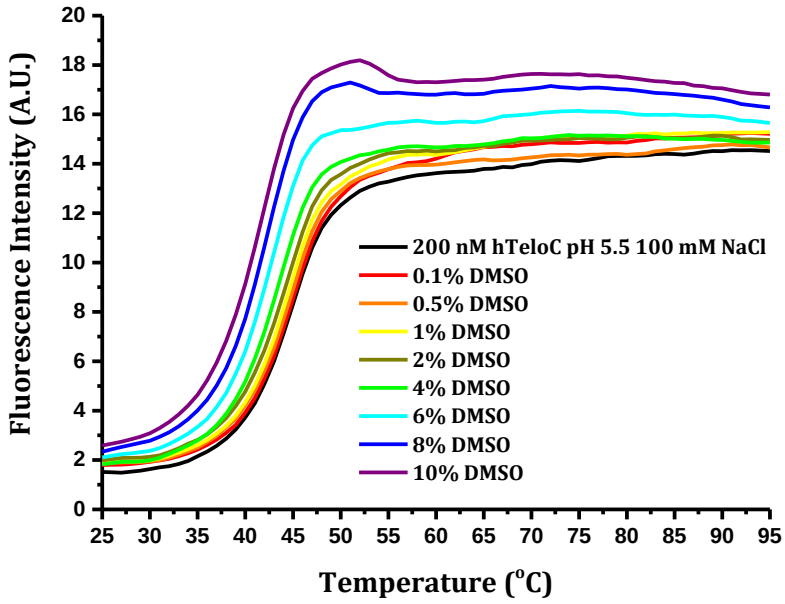



*Figure 3.6: Example FRET melting curves for human telomeric i-motif with increasing percentage of DMSO. [DNA] = 200 nM hTeloC, buffer = 10 mM sodium cacodylate and 100 mM NaCl at pH 5.5.*

| Hit | $\Delta T_m$ (°C) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 100 μM | 80 μM | 60 μM | 40 μM | 20 μM | 10 μM | 5 μM | 1 μM |
| **Chlorhexidine** | 3.2 | 5.2 | 4.0 | 2.5 | 0.5 | 0.5 | -0.2 | 0.2 |
| **Cetylpyridinium chloride** | $T_m$ >95 | $T_m$ >95 | $T_m$ >95 | $T_m$ >95 | $T_m$ >95 | 0.0 | -0.2 | 0.2 |
| **tyrothricin** | 0.0, 17.2[a] | -0.2 | 0.3 | 0.0 | -0.2 | 0.2 | 0.0 | 0.2 |
| **Tamoxifen citrate** | 3.0 | 3.3 | 2.3 | 0.8 | 0.0 | 0.0 | -0.2 | 0.2 |
| **Hexetidine** | 0.5 | 0.2 | 0.7 | 0.5 | -0.3 | 0.0 | -0.2 | 0.2 |
| **Bromocriptine mesylate** | 0.5 | -0.3 | 0.0 | 0.3 | -0.2 | 0.2 | 0.0 | -0.2 |
| **Alexidine hydrochloride** | 31.7 | 27.3 | 22.0 | 18.0 | 3.0 | 1.5 | -0.2 | -0.2 |
| **Pararosaniline pamoate** | 0.5 | 0.2 | 0.5 | 0.2 | -0.2 | -0.2 | 0.0 | -0.2 |
| **Methylbenzethonium chloride** | 35.8 | -1.7, 18.8[a] | -0.8, 17.3[a] | 0.2, 15.0[a] | 0.2, 14.0[a] | 0.2 | 0.0 | -0.2 |
| **Mitoxantrone hydrochloride** | 31.0 | 30.0 | 27.2 | 21.2 | 12.5 | 0.5 | -0.3 | -0.2 |
| **Benzalkonium chloride** | -1.0, 18.3[a] | -1.2, 17.3[a] | -0.3, 18.0[a] | 0.2, 15.2[a] | 0.3 | 0.3 | -0.2 | 0.0 |
| **Tilorone** | 1.0 | 0.7 | 1.0 | 0.5 | 0.5 | 0.3 | -0.2 | 0.0 |
| **Aiodarone hydrochloride** | 3.2 | 1.5 | 0.3 | 0.7 | -0.2 | 0.2 | 0.0 | -0.2 |

*Table 3.3: Average change in melting temperature with change in concentration of hit compounds.[a] A bimodal melting curve was observed so both melting temperatures were recorded. Error = ± 0.3°C.*

Example FRET melting curves for compounds with the most significant effects are shown in figure 3.7 to 3.12. Example FRET melting curves for the remaining compounds are shown in the appendix (section A.1). **Hexetidine** (P4E9) and **Aiodarone hydrochloride** (P12B3) showed no significant change in melting temperature at any concentration and had little effect at 5 mM salt concentration so they were no longer considered as hits and were discounted from further experiments. However, **chlorhexidine** (P1A3), **tamoxifen citrate** (P4B6), **bromocriptine mesylate** (P7D5), **pararosanilone pamoate** (P9F11) and **tilorone** (P11H11), which also did not show significant stabilisation, were taken forward due to their high stabilisation in 5 mM salt conditions.

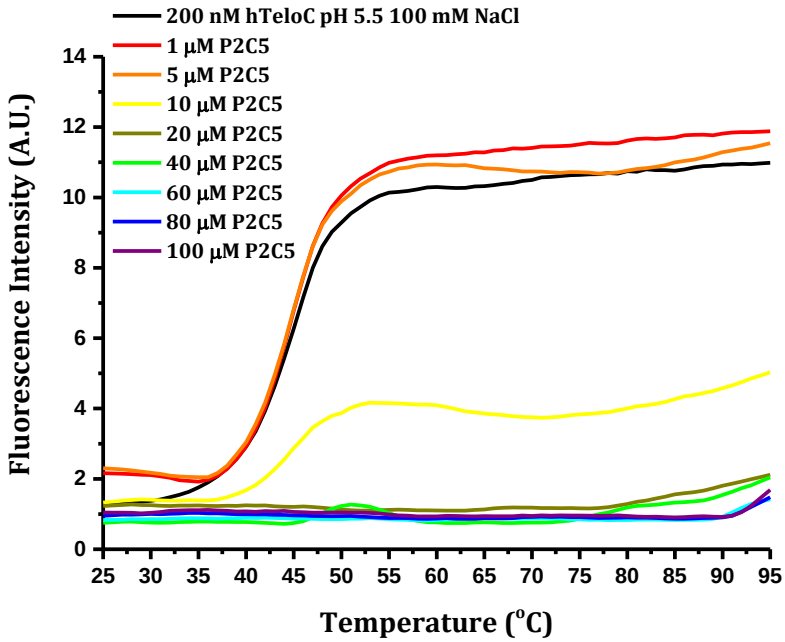



*Figure 3.7: Example FRET melting curves for human telomeric i-motif with increasing concentration of* ***cetylpyridinium chloride****. [DNA] = 200 nM hTeloC, buffer = 10 mM sodium cacodylate and 100 mM NaCl at pH 5.5.*

The results for **cetylpyridinium chloride** (figure 3.7) show that at concentrations above 10 µM, the DNA does not melt within the temperature range of the experiment as the fluorescence signal stays at a minimum from 25 to 90°C, only starting to increase slightly at 95°C. This may mean that the ligand has a highly stabilising affect on the i-motif and that if the experiment could continue to higher temperatures, a normal melting transition would be observed. Alternatively the very low fluorescence signal may be as a result of precipitation of the sample as discussed in chapter 2. This makes it very difficult to measure a melting temperature and it is not possible to determine which of these two effects are being observed without using an alternative experiment such as circular dichroism that can give information about the actual structure of the DNA. For the purposes of comparing this compound to the other hits therefore, the melting temperature is assumed to be ≥95°C.

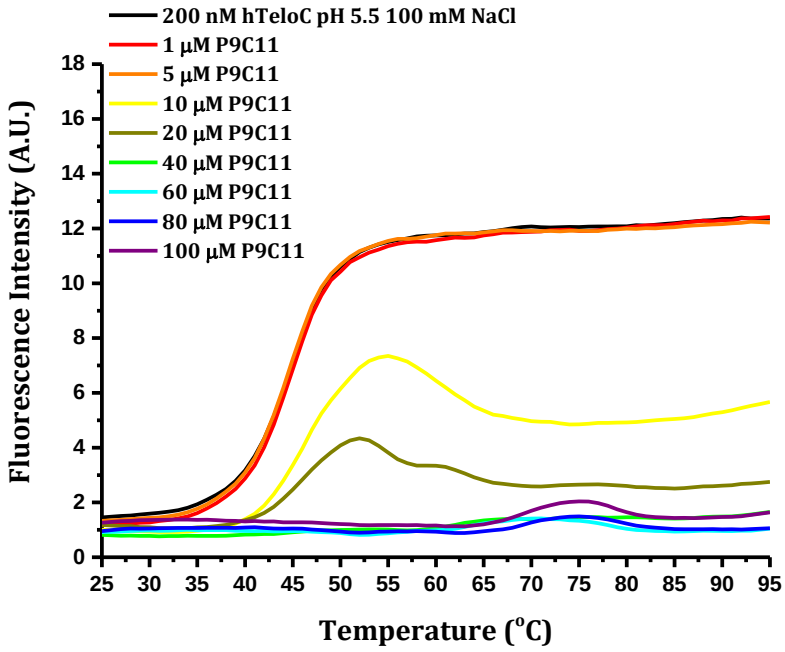



*Figure 3.8: Example FRET melting curves for human telomeric i-motif with increasing concentration of **alexidine hydrochloride**. [DNA] = 200 nM hTeloC, buffer = 10 mM sodium cacodylate and 100 mM NaCl at pH 5.5.*

Unlike the previous example, **alexidine hydrochloride** (figure 3.8) does show small increases in fluorescence and it is at this point where the melting point has been measured. However, after these small increases, the fluorescence begins to decrease again and remains at a low intensity showing that the fluorophores are still in close proximity to one another. This suggests that the DNA may be only partially unfolding and then re-forming into an alternative conformation or that some ligand induced precipitation occurs but not until the DNA has begun to unfold. This means that the melting temperature determined from this data may not be an accurate representation of the complete melting of the DNA but instead indicates at what temperature the DNA is beginning to unfold.

The data in figure 3.9 for **mitoxantrone** indicates that this is a very good example of a stabilising ligand which does not appear to cause any precipitation within the conditions of the experiment as the fluorescence still increases to high intensity levels close to that of the control sample in the absence of ligand (black line). The temperature of the transition point of the curves can clearly be seen to increase as the concentration of **mitoxantrone** increases and the gradient of the transition also appears to become

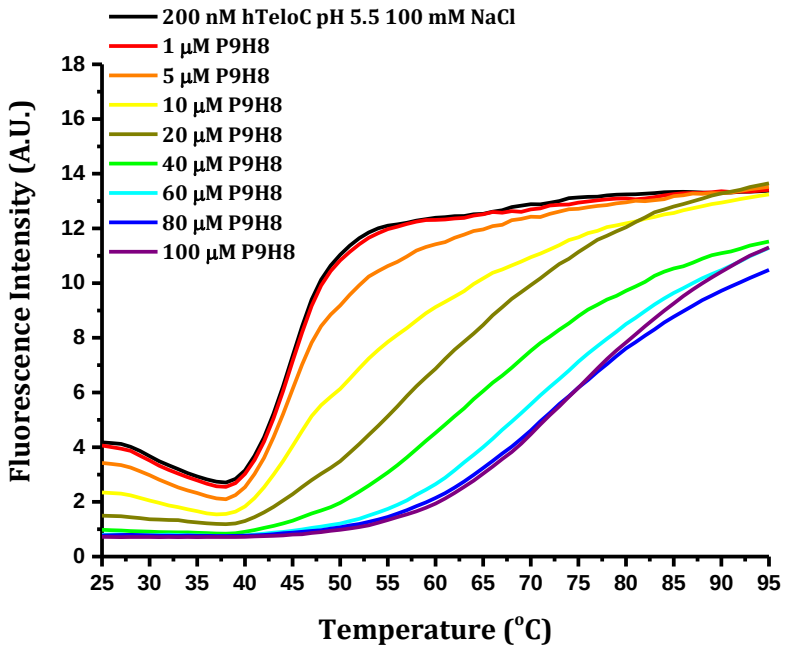



*Figure 3.9: Example FRET melting curves for human telomeric i-motif with increasing concentration of* ***mitoxantrone****. [DNA] = 200 nM hTeloC, buffer = 10 mM sodium cacodylate and 100 mM NaCl at pH 5.5.*

shallower. This indicates that the DNA also melts more slowly and over a wider temperature range than in the absence of any ligand. The curves only show one transition point suggesting that only one DNA conformation is present in the sample.

From these results it can be seen that **cetylpyridinium chloride**, **alexidine hydrochloride** and **mitoxantrone** are the most strongly stabilising hits as these cause the biggest increase in $T_m$ with concentration of ligand (e.g. $\Delta T_m$ = +50.7, +18.0 and +21.2°C respectively at 40 μM). However **mitoxantrone** is the most promising of the three as it gives the clearest melting transitions with no observable precipitation. The other two compounds may be highly stabilising but the data suggests that they may also cause some precipitation or change in conformation. These two compounds do not give a clear dose response and instead provide sudden changes in melting temperature. CD spectroscopy will be used in section 3.4 to examine the effect of these compounds on the i-motif structure in more detail.

The compounds that show the next highest changes in melting temperature are **tyrothricin** ($\Delta T_m$ = +17.2°C at 100 μM), **methylbenzethonium chloride** ($\Delta T_m$ = +15°C at 40 μM) and **benzalkonium chloride** ($\Delta T_m$ = +15.2°C at 40 μM). These compounds

are interesting as they show bimodal melting curves. Initially the fluorescence increases upon the first melting transition but then decreases again before the second melting transition. This suggests that in the presence of these compounds, there is more than one conformation of DNA in solution with different stabilities under the conditions of the experiment. The fact that the fluorescence increases and then decreases again before the second melt implies that one conformation of the DNA may be partially or completely unfolding before the sequence then refolds into a different structural form with a higher melting temperature stabilised by the ligands.

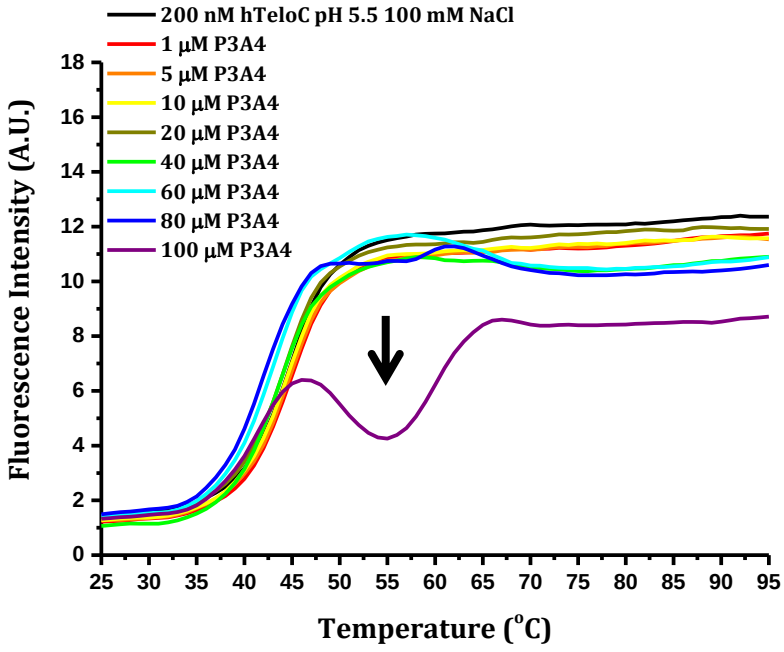



*Figure 3.10: Example FRET melting curves for human telomeric i-motif with increasing concentration of **tyrothricin**. [DNA] = 200 nM hTeloC, buffer = 10 mM sodium cacodylate and 100 mM NaCl at pH 5.5.*

**Tyrothricin** only shows this bimodal melting curve clearly (figure 3.10) at a concentration of 100 μM but there is also a small fluctuation in the curve at 80 μM. At lower ligand concentrations there is no change in the melting curve at all suggesting that **tyrothricin** is only interacting with the i-motif to a significant extent at high concentrations. However, since tyrothricin is a mixture of peptides, it is likely that only one or a small number of the constituent peptides interacts with the i-motif and these will be at a much lower concentration than the mixture as a whole. Alternatively, the constituent peptides may each have a number of weak interactions that accumulate to

give the observed effect at higher concentrations. The first melting transition is almost the same as the i-motif in the absence of the ligand, suggesting that this transition is for the i-motif structure, whereas the second transition is for an alternative structure with a higher melting temperature that is stabilised by or forms as a result of the peptide binding.

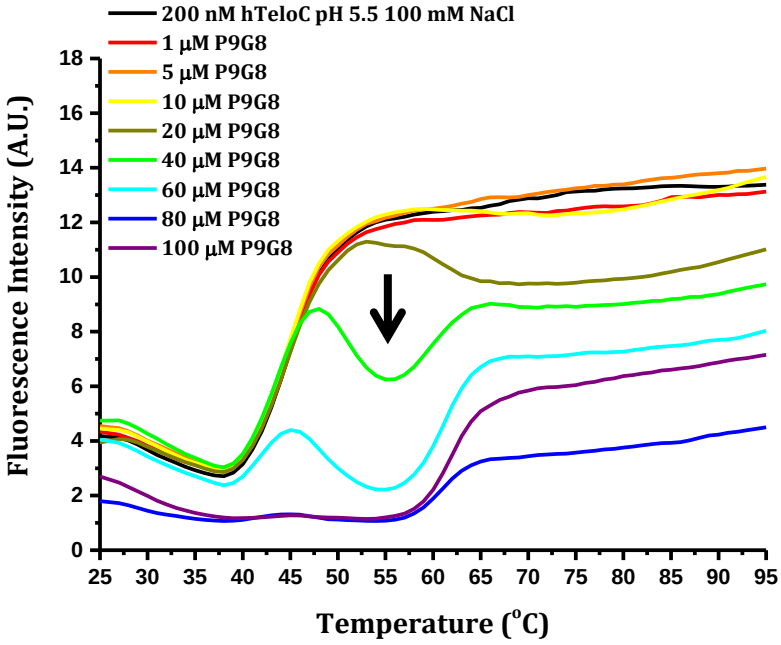



*Figure 3.11: Example FRET melting curves for human telomeric i-motif with increasing concentration of* ***methylbenzethonium chloride****. [DNA] = 200 nM hTeloC, buffer = 10 mM sodium cacodylate and 100 mM NaCl at pH 5.5.*

**Methylbenzethonium chloride** (figure 3.11) and **benzalkonium chloride** (figure 3.12) show very similar effects to **tyrothricin** (figure 3.10) but the changes in the melting curves begin at a lower ligand concentration (40 µM). Again the first melting transition occurs at almost the same temperature as the i-motif alone and implies that this melting event is that of the i-motif structure without any stabilising effect from the ligand and that it is the second melting point that represents an alternative conformation. Interestingly, in both examples, the change in fluorescence for the first transition is reduced at higher concentrations to the point that at 80 to 100 µM there is almost no observable melting transition at all. This implies that the two structures may be in equilibrium and that increasing concentrations of ligand shift the equilibrium towards that of the alternative structure. This structure could be a different i-motif conformation, a partially folded structure or a hairpin.

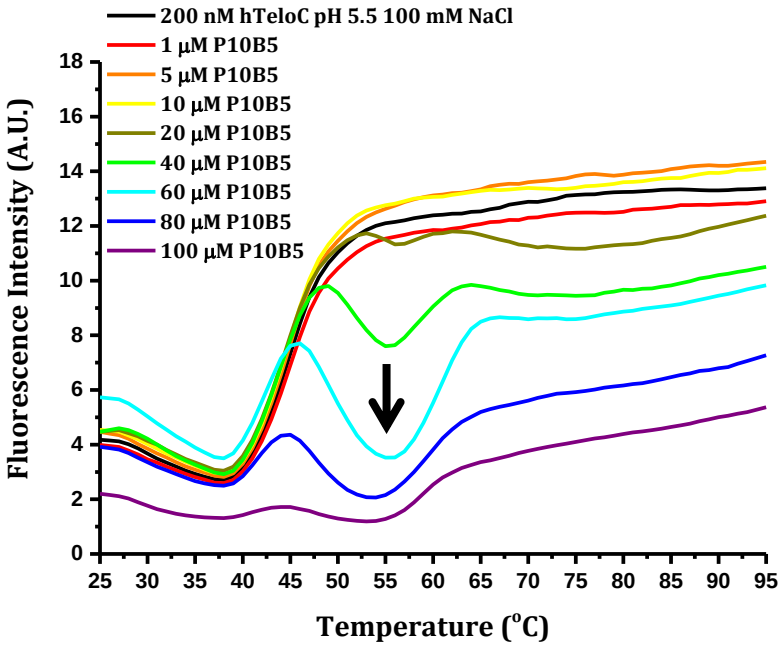



*Figure 3.12: Example FRET melting curves for human telomeric i-motif with increasing concentration of* ***benzalkonium chloride****. [DNA] = 200 nM hTeloC, buffer = 10 mM sodium cacodylate and 100 mM NaCl at pH 5.5.*

One factor that may have an effect is that **methylbenzethonium chloride**, **benzalkonium chloride** and **cetylpyridinium chloride** consist of quaternary ammonium cations attached to lipophilic hydrocarbon chains and hence have the ability to form micelles. This is where multiple compounds in a polar solution such as water, can group together to form a spherical structure where the lipophilic portion of the molecules align together in the centre of the sphere, leaving the hydrophilic part of the molecule on the surface where it interacts more favourably with the aqueous solution. These structures occur spontaneously at higher ligand concentrations as it is more favourable for the lipophilic groups to interact with one another, rather than the surrounding solution. The ligand concentration at which spontaneous micelle formation occurs is called the critical micelle concentration (cmc) and this can vary depending on the solution conditions such as buffer and salt concentration. The cmc for these compounds is unknown in these buffer conditions and therefore it is unclear whether these compounds will be forming micelles at the concentrations that are being tested.

### 3.3.2 Discounting the effect of citrate and pamoate counter ions.

As **tamoxifen** and **pararosaniline** have been tested as citrate and pamoate salts respectively, it is important to ascertain that the observed stabilisation is as a result of the ligand rather than of the pamoic or citric acid. To do this, **tamoxifen** free base and **Pararosaniline chloride** were tested. As pamoic acid is an aromatic compound with the ability to π-stack, it is likely that it is able to bind to DNA so this was also tested separately (figure 3.13). From these results it can be seen that pamoic acid has no effect on the melting temperature of the i-motif, so the stabilisation effects observed for **pararosaniline pamoate** (P9F11) are due to the **pararosaniline** moiety.

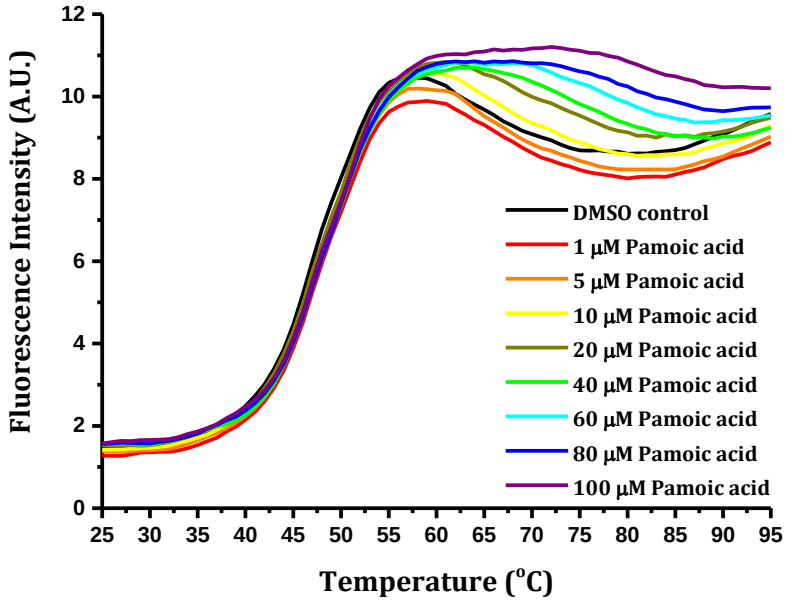



*Figure 3.13: Example FRET melting curves for human telomeric i-motif with increasing concentration of pamoic acid. [DNA] = 200 nM hTeloC, buffer = 10 mM sodium cacodylate and 5 mM NaCl at pH 5.5.*

For **tamoxifen**, in 100 mM salt conditions the free base (figure 3.14) was more stabilising of the i-motif than the citrate salt with a maximum average stabilisation potential ($\Delta T_m$) of 12.5°C compared to 3°C at a concentration of 100 µM. As **tamoxifen citrate** only contains one citrate molecule per tamoxifen, it can be assumed that **tamoxifen** is at the same concentration in both cases. This suggests that the citrate may have been having some effect to cause the lower magnitude of stabilisation. Therefore, from here on, the free base form of **tamoxifen** (P4B6fb) was tested rather than the citrate salt form.

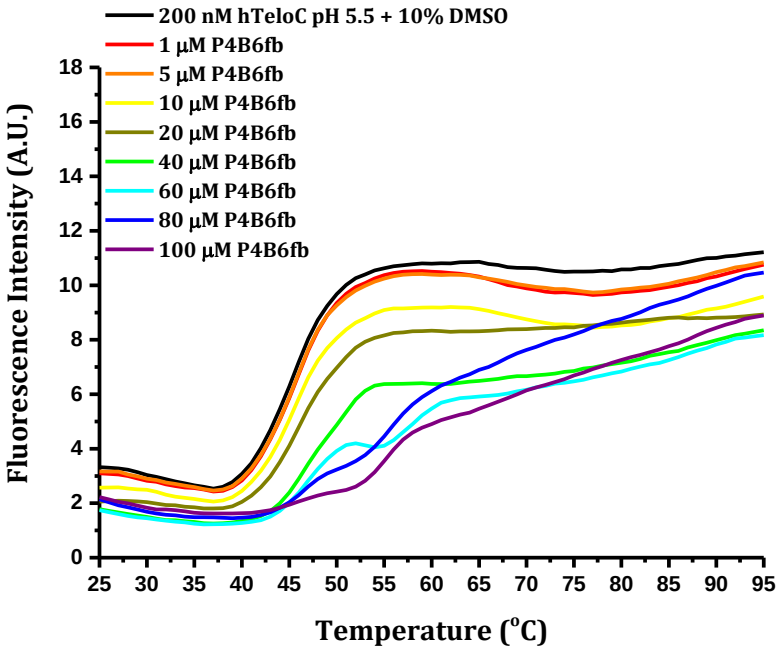



*Figure 3.14: Example FRET melting curves for human telomeric i-motif with increasing concentration of **tamoxifen** free base. [DNA] = 200 nM hTeloC, buffer = 10 mM sodium cacodylate and 100 mM NaCl at pH 5.5.*

### 3.3.3 Dose response versus telomeric i-motif with 5 mM NaCl

Having narrowed down the list of hits from 13 to 11 by exclusion of **hexetidine** and **aiodarone hydrochloride**, dose response studies were also carried out against the human telomeric i-motif in 5 mM salt conditions for comparison (table 3.4). Although 11 hits are under consideration, only 9 hits were tested as the structures of some of the compounds are so similar that it was assumed they would have similar effects. In Particular **chlorhexidine** and **alexidine hydrochloride** are both bisbiguanides. While **cetylpyridinium chloride** and **benzalkonium chloride** are both quaternary amine compounds with an aromatic head and a long aliphatic tail. Hence only **alexidine hydrochloride** and **cetylpyridinium chloride** were tested.

| Hit | $\Delta T_m$ (°C) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 100 μM | 80 μM | 60 μM | 40 μM | 20 μM | 10 μM | 5 μM | 1 μM |
| **Cetylpyridinium chloride** | 47.7 | 47.7 | 47.7 | 44 | 47.7 | 47.7 | -0.3 | 0.0 |
| **Tyrothricin** | $T_m$ >95 | $T_m$ >95 | $T_m$ >95 | 40.7 | 18.7 | 0.0 | 0.0 | 0.0 |
| **Tamoxifen free base** | 17.0 | 15.8 | 13.5 | 10.5 | 2.7 | 0.3 | 0.2 | 0.2 |
| **Bromocriptine mesylate** | 6.7 | 6.0 | 4.7 | 0.3 | 0.3 | 0.3 | 0.3 | 0.0 |
| **Alexidine hydrochloride** | $T_m$ >95 | $T_m$ >95 | $T_m$ >95 | $T_m$ >95 | $T_m$ >95 | $T_m$ >95 | 3.7 | 0.0 |
| **Pararosaniline chloride** | 4.7 | 4.0 | 4.0 | 4.7 | 0.3 | 0.0 | 0.3 | 0.3 |
| **Methylbenzethonium chloride** | $T_m$ >95 | $T_m$ >95 | $T_m$ >95 | $T_m$ >95 | 16.7 | -0.7<br>14.0[a] | -0.3 | -0.3 |
| **Mitoxantrone** | $T_m$ >95 | $T_m$ >95 | $T_m$ >95 | $T_m$ >95 | $T_m$ >95 | 45.3 | 44.3 | -1.3 |
| **Tilorone** | $T_m$ >95 | $T_m$ >95 | $T_m$ >95 | $T_m$ >95 | -1.7 | -0.7 | 0.3,<br>4.3[a] | 0.0 |

*Table 3.4: Change in melting temperature with change in concentration of hit compounds.[a] A bimodal melting curve was observed so both melting temperatures were recorded. Error = ± 0.3°C.*

As was seen in section 3.2.5, most of the compounds are more stabilising in the 5 mM NaCl buffer compared to the 100 mM NaCl buffer. **Pararosaniline** and **bromocriptine** only give a small stabilisation at higher concentrations however **tilorone**, which had no significant stabilising effect in 100 mM salt conditions, is significantly more stabilising. This suggests that **tilorone** binds through more electrostatic interactions as a high cation concentration would reduce this effect. **Methylbenzethonium chloride** still gave rise to a bimodal melting curve but at a lower concentration than previously (10 μM). While **tyrothricin** only showed a very small change in fluorescence before the melting transition at 20 μM. This suggests that the alternative DNA conformations that these compounds may be stabilising are even more favoured under these conditions.

## 3.4 Effect of Hits on i-Motif Structure: Circular Dichroism

In order to see what effect the hit compounds were having on the i-motif structure, circular dichroism was used. As discussed in chapters 1 and 2, circular dichroism is a very useful tool for looking at structural changes in DNA as different structures give characteristically different spectra.[102] The human telomeric i-motif was tested at a concentration of 10 μM with increasing number of equivalents of ligand, up to 20 equivalents (figure 3.15 – 3.23). However, if precipitation occurred before 20 equivalents, further aliquots of ligand were not added.

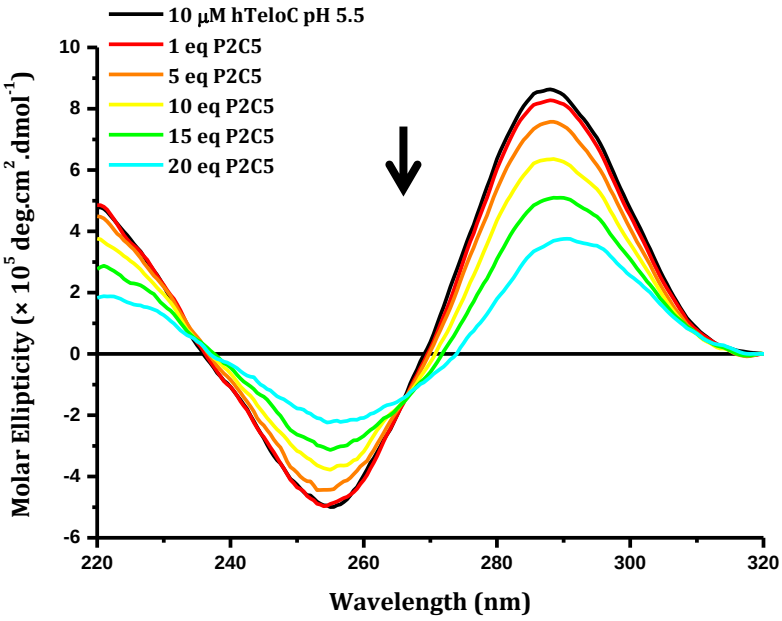



*Figure 3.15: Example CD titration of human telomeric i-motif with increasing concentration of* ***cetylpyridinium chloride****. [DNA] = 10 μM hTeloC, buffer = 10 mM sodium cacodylate and 100 mM NaCl at pH 5.5.*

The CD spectrum of **cetylpyridinium chloride** with the i-motif (figure 3.15) shows a decrease in signal at both 255 nm and 288 nm suggesting a decrease in the amount of i-motif structure present in the sample but no unfolding is observed as there is no shift in the peak towards 275 nm. There was no indication of precipitation from observation of the sample and there were no scattering effects which would be demonstrated by the appearance of signals above 320 nm. However, these scattering effects were not always observed with **BisA** in chapter 2 which did cause condensation and it is likely

that the general decrease in molar ellipticity across the spectrum is indicative of precipitation.

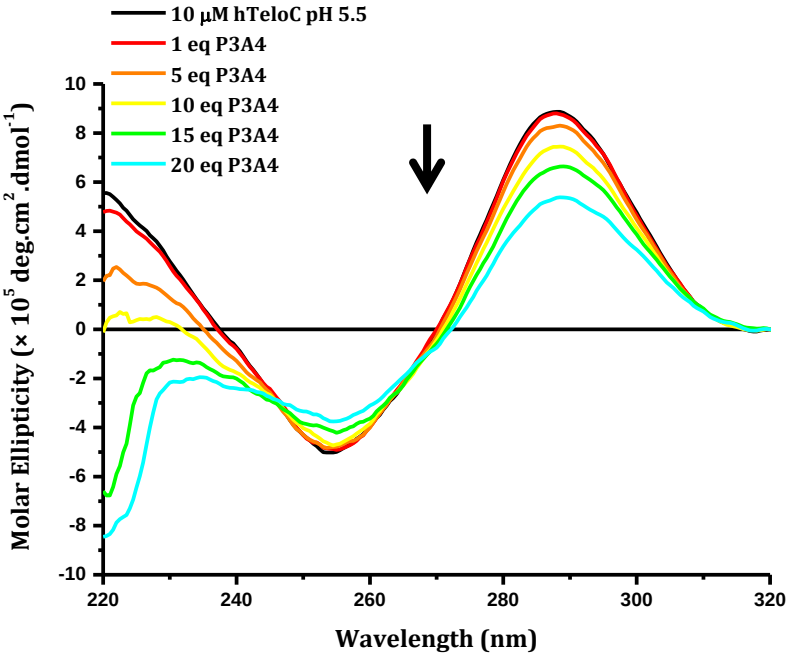



*Figure 3.16: Example CD titration of human telomeric i-motif with increasing concentration of **tyrothricin**. [DNA] = 10 µM hTeloC, buffer = 10 mM sodium cacodylate and 100 mM NaCl at pH 5.5.*

Although the CD spectrum of **tyrothricin** (figure 3.16) does not show such a large decrease in molar ellipticity as the previous example, precipitation was observed at a concentration of 100 µM (10 eq). The increasingly negative peak in the spectrum at 220 nm is likely to be due to the compound itself as the tyrocidines, a major constituent of **tyrothricin** have been shown to give a negative CD signal between 210 and 220 nm.[162]

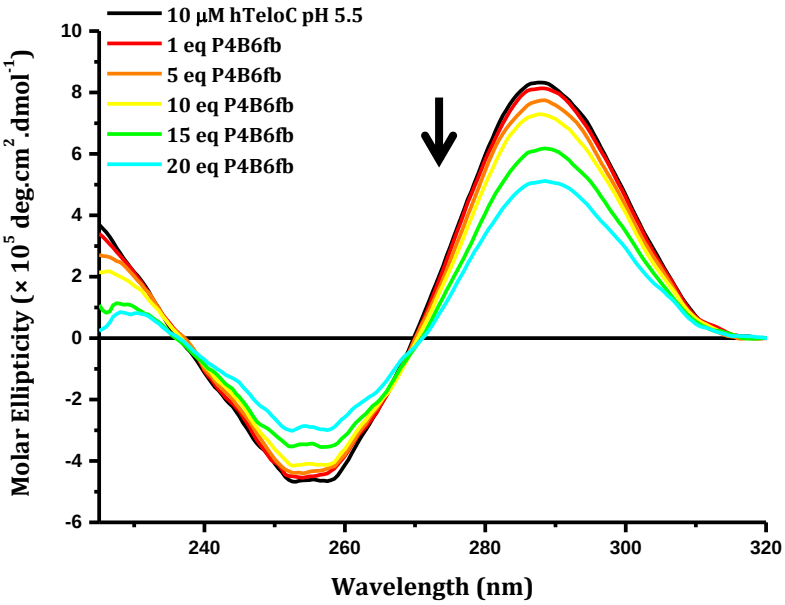



*Figure 3.17: Example CD titration of human telomeric i-motif with increasing concentration of **tamoxifen**. [DNA] = 10 µM hTeloC, buffer = 10 mM sodium cacodylate and 100 mM NaCl at pH 5.5.*

The CD spectrum of **tamoxifen** (figure 3.17) with the i-motif gives a very similar spectrum to **tyrothricin** with a decrease in molar ellipticity at 255 and 288 nm. In this case however, **tamoxifen** is not chiral and there is no peak forming at 220 nm. As with **tyrothricin**, precipitation could be observed by eye at higher concentrations and this may be represented in the CD spectrum by the small decrease in molar ellipticity. This suggests that **tamoxifen** does not have a significant effect on the i-motif structure up to 10 equivalents.

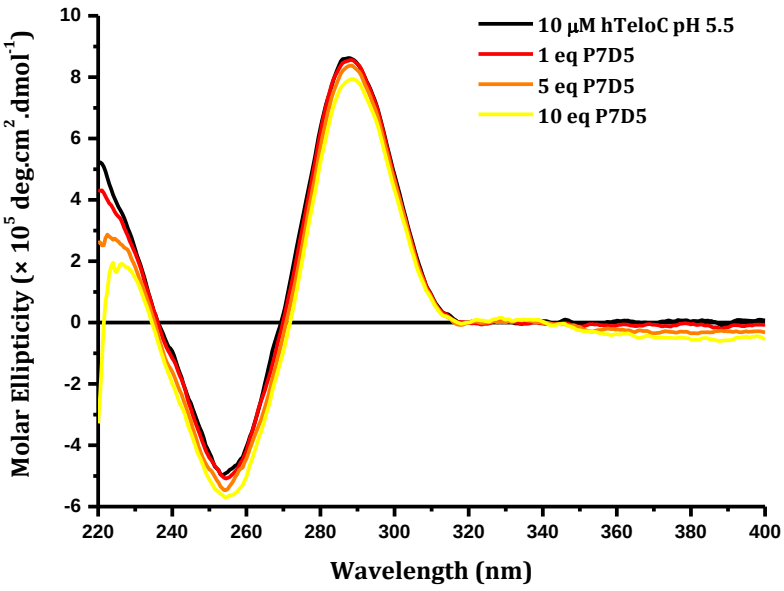



*Figure 3.18: Example CD titration of human telomeric i-motif with increasing concentration of* ***bromocriptine mesylate****. [DNA] = 10 μM hTeloC, buffer = 10 mM sodium cacodylate and 100 mM NaCl at pH 5.5.*

The CD titration of hTeloC with **bromocriptine mesylate** (figure 3.18) showed very little change up to 10 equivalents (100 μM) but began to show precipitation at this point which can be seen by the presence of a signal at 380 to 400 nm. A DNA sample should not give any signal in this region and this signal is a result of scattering effects. Higher concentrations of **bromocriptine** were not tested as precipitation had already occurred. The fact that the compound does not change the signal at 288 nm suggests that it may not be binding to the i-motif at all under these conditions or that if it does so, then there is no perturbation of the structure to alter its conformation. In earlier FRET melting experiments (Table 3.2 and 3.3) **bromocriptine** had little to no effect in buffer conditions containing 100 mM NaCl but did cause a base line fluorescence signal in 5 mM NaCl conditions. This suggests that the higher salt concentration may prevent any potential binding.

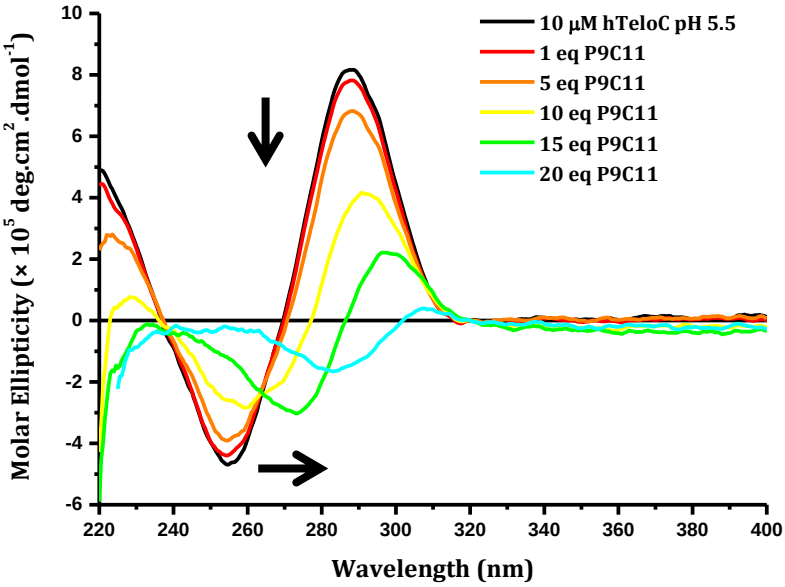



*Figure 3.19: Example CD titration of human telomeric i-motif with increasing concentration of **alexidine hydrochloride**. [DNA] = 10 μM hTeloC, buffer = 10 mM sodium cacodylate and 100 mM NaCl at pH 5.5.*

In figure 3.19, the CD spectrum shows significant changes upon addition of **alexidine hydrochloride** with not only a decrease in molar ellipticity but a shift in the peak towards 300 nm. This is similar to the CD spectra observed for the condensation effects of **BisA** in chapter 2 (figure 2.15) and was accompanied by visible precipitation. It may well be the case that for compounds such as **bromocriptine**, **tamoxifen** and **tyrothricin**, for which precipitation was observed but there is little change in the CD spectrum, the observed precipitation is merely the ligand itself precipitating from the buffer solution. Whereas, in the case of **alexidine hydrochloride** or other compounds where the precipitation is accompanied by dramatic changes in the CD spectrum, this suggests that the compound does alter the conformation of the DNA and the observed precipitate may be the complex of the DNA and ligand.

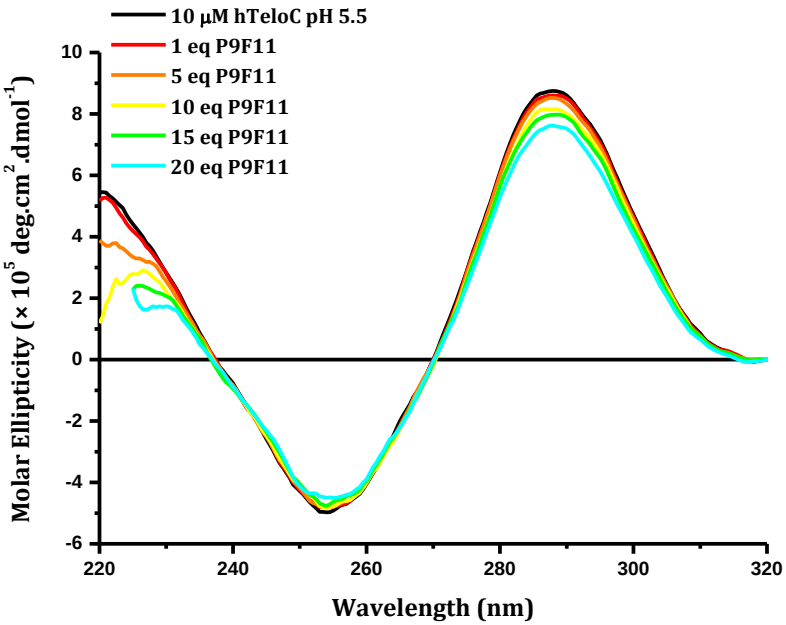



*Figure 3.20: Example CD titration of human telomeric i-motif with increasing concentration of* ***pararosaniline****. [DNA] = 10 µM hTeloC, buffer = 10 mM sodium cacodylate and 100 mM NaCl at pH 5.5.*

Figure 3.20 shows a slight decrease in signal at 288 nm that is most likely a result of dilution and is not significant. This indicates that if **pararosaniline** is binding to the i-motif, it does not have any effect on the structural conformation that is significant enough to be observed by CD. This implies that **pararosaniline** may not bind to the i-motif at all or if it does, that it is through an electrostatic or stacking interaction on the outside of the structure. Any binding within the loops, grooves or the core of the i-motif is likely to cause at least a small disturbance to the structure. Like **bromocriptine** in figure 3.18, **pararosaniline** only showed significant effects in FRET experiments in buffer containing 5 mM NaCl. Therefore the combined FRET and CD experiments suggest that the higher salt concentrations are inhibiting any possible binding interactions from taking place.

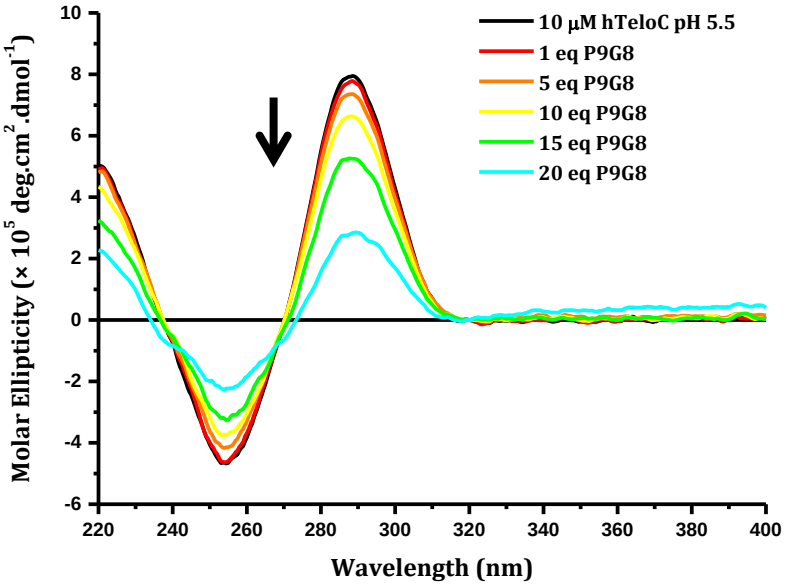



*Figure 3.21: Example CD titration of human telomeric i-motif with increasing concentration of* ***methylbenzethonium chloride****. [DNA] = 10 μM hTeloC, buffer = 10 mM sodium cacodylate and 100 mM NaCl at pH 5.5.*

The CD spectrum of **methylbenzethonium chloride** with the i-motif (figure 3.21) is very similar to that of **cetylpyridinium chloride** which is perhaps unsurprising as they share some similarities in their structures and both contain quaternary ammonium cation centres. This common functionality may be key to the compounds interaction with the i-motif and like **cetylpyridinium chloride** (figure 3.15) no precipitation was observed despite a decrease in molar ellipticity that suggests there is less i-motif structure present in solution. As both these compounds are able to form micelles, it may be the case that under the conditions of the experiment, micelles are being formed which interact with the i-motif instead of the individual compound.

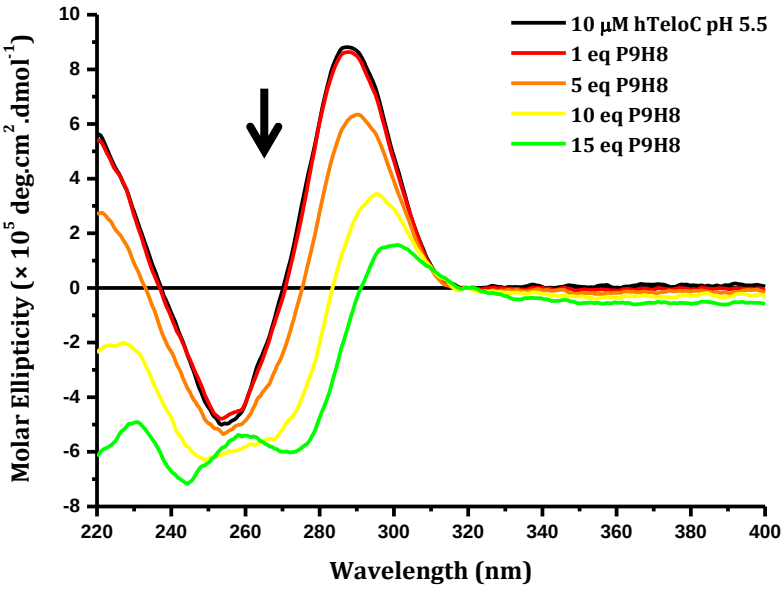



*Figure 3.22: Example CD titration of human telomeric i-motif with increasing concentration of* ***mitoxantrone****. [DNA] = 10 µM hTeloC, buffer = 10 mM sodium cacodylate and 100 mM NaCl at pH 5.5.*

Although **mitoxantrone** did not appear to cause any precipitation in FRET melting studies, at the higher concentrations used for CD spectroscopy, precipitation was observed at 10 and 15 equivalents (figure 3.22). This was observable by the naked eye and is apparent in the CD spectrum where the signal at 288 nm decreases and shifts towards 300 nm. Like figure 3.18 for **alexidine hydrochloride** this indicates that mitoxantrone is binding to the i-motif and forming a complex which may have a different structure upon condensation and precipitation. There are also significant changes in the spectrum between 220 and 280 nm which may be as a result of scattering or due to structural changes. **Mitoxantrone** is still a promising compound to investigate however as it does not cause precipitation at lower concentrations (50 µM).

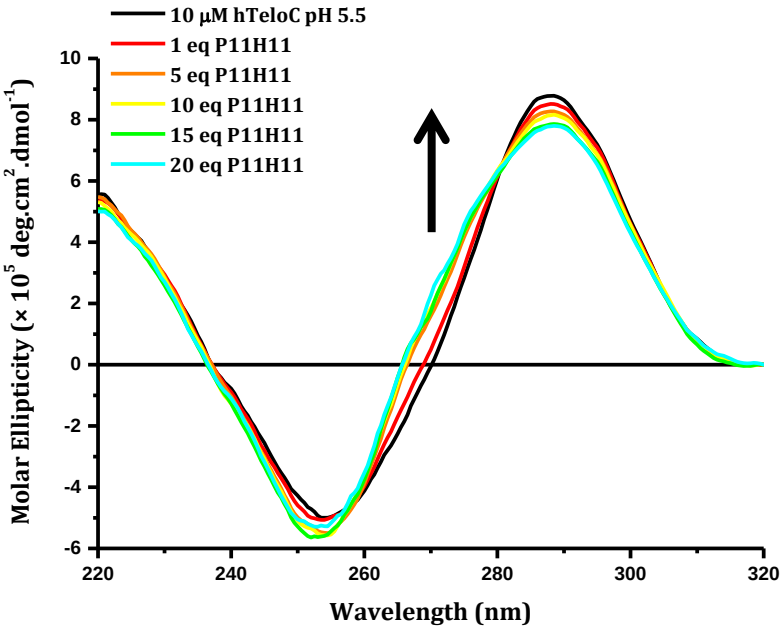



*Figure 3.23: Example CD titration of human telomeric i-motif with increasing concentration of **tilorone**. [DNA] = 10 μM hTeloC, buffer = 10 mM sodium cacodylate and 100 mM NaCl at pH 5.5.*

Finally the CD titration of **tilorone** with the i-motif (figure 3.23) does not show a significant decrease in the signal at 288 nm and no precipitation was observed. In fact there is an increase in signal at 270 nm showing the formation of an ICD with increasing concentration of **tilorone**. This indicates binding and can be used to measure a binding affinity.

By fitting the change in molar ellipticity with increasing concentration of ligand, a binding affinity can be measured. The change in molar ellipticity at 270 nm was normalised, which transforms the data so that the lowest point is equal to zero and the highest is equal to one, this gives a measure of the fraction of ligand bound to the DNA. The data was then fitted with both 1 to 1 and 2 to 1 binding models (equation 3.1 and 3.2). The two models gave very similar fits (figure 3.24) resulting in similar $K_d$'s for the first binding site (table 3.5) but the 2 to 1 model with a very weak affinity for the second binding site gave a slightly better fit.

$$\text{Equation 3.1: } \theta = \frac{K_a[Ligand]}{1+K_a[Ligand]} \qquad \text{1 to 1 binding model}$$

$$\textit{Equation 3.2: } \theta = \frac{K_1[Ligand]+2K_1K_2[Ligand]^2}{1+K_1[Ligand]+K_1K_2[Ligand]^2} \qquad \textit{2 to 1 binding model}$$

Where $K_a$ is the association constant, $K_1$ and $K_2$ are association constants for binding site 1 and 2 respectively. $K_d$ is the reciprocal of $K_a$.

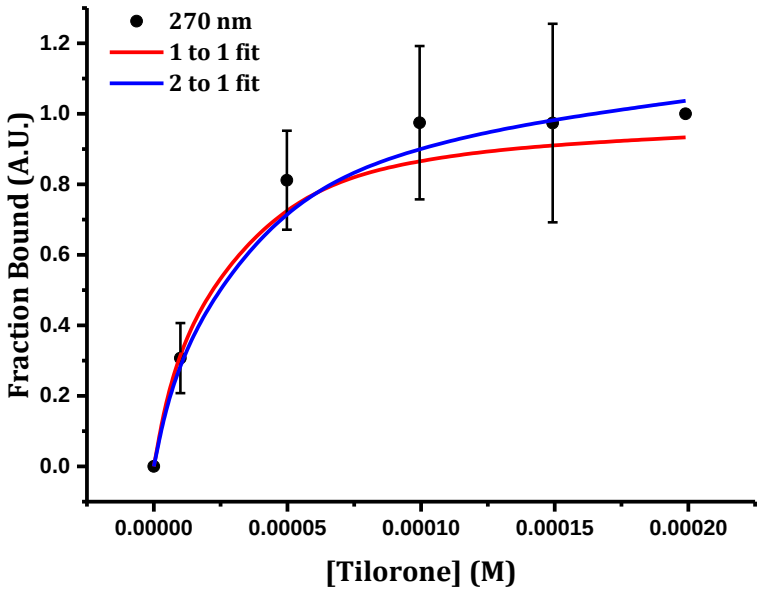



*Figure 3.24: Normalised average change in CD at 270 nm for **tilorone** with human telomeric i-motif, fitted with 1 to1 and 2 to 1 binding models. Error bars show the standard deviation across 2 repeats. [DNA] = 10 µM hTeloC, buffer = 10 mM sodium cacodylate and 100 mM NaCl at pH 5.5.*

| Binding affinity (µM) | | |
|---|---|---|
| 1 to 1 | $K_d$ | 14 ± 10 |
| 2 to 1 | $K_{d1}$ | 17 ± 10 |
| | $K_{d2}$ | 1478 ± 2077 |

*Table 3.5: Binding affinities of **tilorone** for telomeric i-motif from CD titrations. The error given is the maximum difference between the average and individual data.*

These results indicate that **tilorone** has a moderately strong affinity for the i-motif with dissociation constants ($K_d$) measured in the low micromolar range. Whilst the 2 to 1 binding model may give a marginally better fit, the very high dissociation constant for the second binding site and the similarity of the $K_d$ for the first binding site to the

one given by the 1 to 1 model suggests that there is only one significant binding site for **tilorone** with the human telomeric i-motif. This binding affinity is comparable to those identified for other existing i-motif ligands such as the terbium complexes discussed in section 1.5.4 ($K_d$ = 22 μM and 30 μM) but is not as strong as the affinity **tilorone** has for double stranded DNA (1.05 μM for poly(dA-dT)), though this was measured under different conditions.[163] As **tilorone** has not been shown to bind to the i-motif before this does still suggest that it is a good starting structure to develop analogues from.

## 3.5 Variation in stabilisation potential

After excluding **alexidine hydrochloride** for its precipitation effects as well as **cetylpyridinium chloride** and **methylbenzethonium chloride**, which may be forming micelles, the six remaining compounds were taken forward for further study. Dose response studies were carried out to assess whether there was any degree of selectivity in the hits stabilising effect compared to other DNA structures. The dose response experiments were repeated with 3 different structures. Firstly the c-Myc i-motif, as described in section 1.3.2, is an oncogene promoter i-motif sequence. Therefore this represents a different form of anticancer target to that of the telomeric i-motif. This i-motif has a different structure to hTeloC with larger loops and the sequence 5'-TCCCCACCTTCCCCACCCTCCCCACCCTCCCCA-3'. Secondly, the hits were tested against the human telomeric G-quadruplex sequence. This is the complementary strand to the i-motif forming the very different G-quadruplex structure. As discussed in section 1.1, telomeric G-quadruplexes have been shown to inhibit telomerase and being complimentary to the i-motif, it is not yet clear what the dynamics between the two structures are. The telomeric G-quadruplex sequence is 5'-GGGTTAGGGTTAGGGTTAGGG-3'. Finally, the hits were tested against a double stranded sequence 5'-TATAGCTATA-HEG(18)-TATAGCTATA-3'. This sequence was chosen because it has been used previously in the literature by a variety of groups in comparison to the G-quadruplex.[164,165] Ideally the hit compounds would have some degree of selectivity for the i-motif structure over other types of secondary structure and may also be selective between different types of i-motif.

In figure 3.25 to 3.30 the change in melting temperature with increasing concentration is plotted for each of the 6 potential ligands. This is to see whether there is any variation in the ability of the ligands to stabilise the different DNA structures at different concentrations. However, care should be taken when comparing the maximum change in melting temperatures that are reached at higher ligand concentrations as double stranded DNA has a higher melting point than either the i-motif or the G-quadruplex under these conditions. This means there is a narrower range of temperature between the melting point and the maximum scale of the experiment of 95°C so only a lower maximum $\Delta T_m$ is possible (~ 30°C).

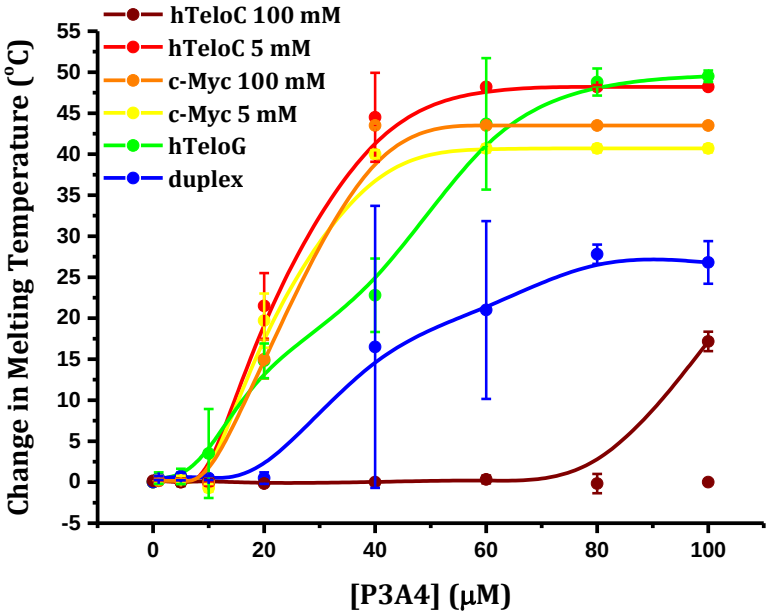



*Figure 3.25: Change in melting temperature with increasing concentration of* ***tyrothricin*** *with human telomeric i-motif (hTeloC), c-Myc i-motif, human telomeric G-quadruplex (hTeloG) and duplex DNA. Error bars show the standard deviation across 2 repeats. DNA = 200 nM, buffer = 10 mM sodium cacodylate and 100 mM or 5 mM NaCl. Two points are shown where biphasic melting curves were observed.*

The combined FRET melting results for **tyrothricin** (figure 3.25) show that in 100 mM salt conditions, this cyclic peptide is not strongly stabilising of the telomeric i-motif except at a concentration of 100 μM where a biphasic melting curve was observed, hence there are two data points on the graph. Conversely in 5 mM salt conditions, **tyrothricin** is highly stabilising, reaching maximum possible stabilisation potentials just above 40 μM. What is interesting is that for the c-Myc i-motif, there was very little

difference in stabilisation under the two salt conditions with the structure being highly stabilised under both conditions. This suggests that the interaction with the telomeric i-motif is perhaps more electrostatic because a higher salt concentration will disfavor this. Whereas the interaction with c-Myc may be less electrostatic with the larger loop sequences allowing easier access for π-π stacking between loop bases and aromatic residues in **tyrothricin** peptides. These results also indicate that **tyrothricin** is highly stabilising of the G-quadruplex and duplex DNA reaching maximum stabilisation potentials at higher concentrations (>80 μM). However, there is some degree of preference at concentrations of 20 μM, where **tyrothricin** is around 20 degrees more stabilising of i-motif and G-quadruplex sequences than of duplex. While at 40 μM, **tyrothricin** is a further 25 degrees more stabilising of the i-motif than of the G-quadruplex.

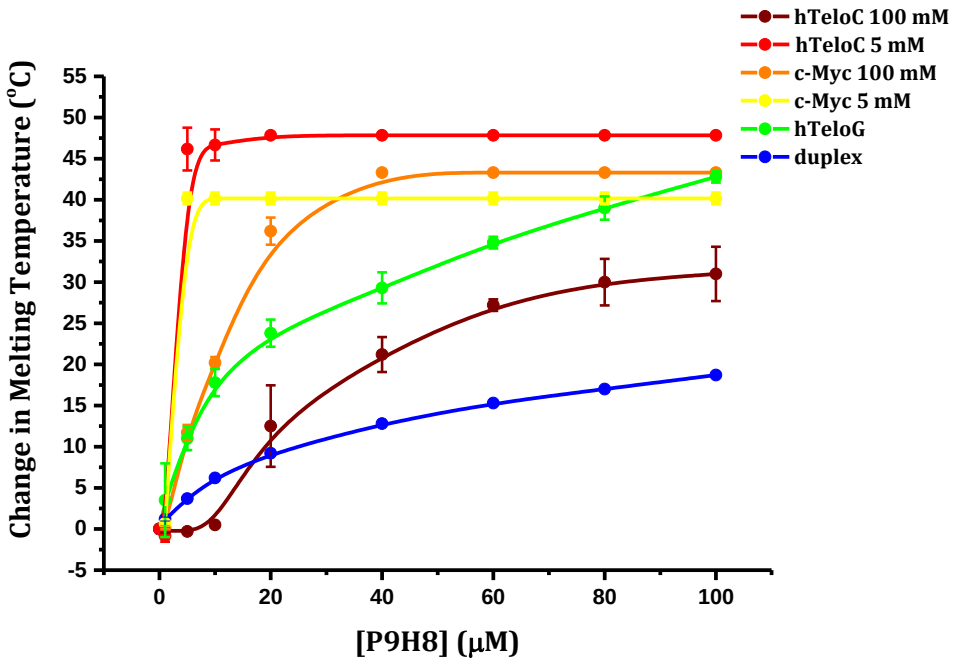



*Figure 3.26: Change in melting temperature with increasing concentration of* ***mitoxantrone*** *with human telomeric i-motif (hTeloC), c-Myc i-motif, human telomeric G-quadruplex (hTeloG) and duplex DNA. Error bars show the standard deviation across 2 repeats. DNA = 200 nM, buffer = 10 mM sodium cacodylate and 100 mM or 5 mM NaCl.*

The results for **mitoxantrone** (figure 3.26) again show that the compound is stabilising of all the DNA structures tested. This is not surprising however, as it is known to bind to duplex DNA as part of its inhibition of topoisomerase II[166] and its flat planar structure is similar to many other G-quadruplex ligands.[89] The data suggests

that even at higher concentrations, **mitoxantrone** is more stabilising of hTeloC i-motif in 5 mM salt conditions compared to 100 mM salt conditions but is highly stabilising of c-Myc i-motif under both conditions. Again this may be as a result of the larger loops enabling better access for the compound to interact with the bases in the structure. Some degree of preference can be observed at low concentrations up to 5 μM where mitoxantrone stabilises c-Myc and hTeloC i-motifs in 5 mM salt conditions far more than the other structures.

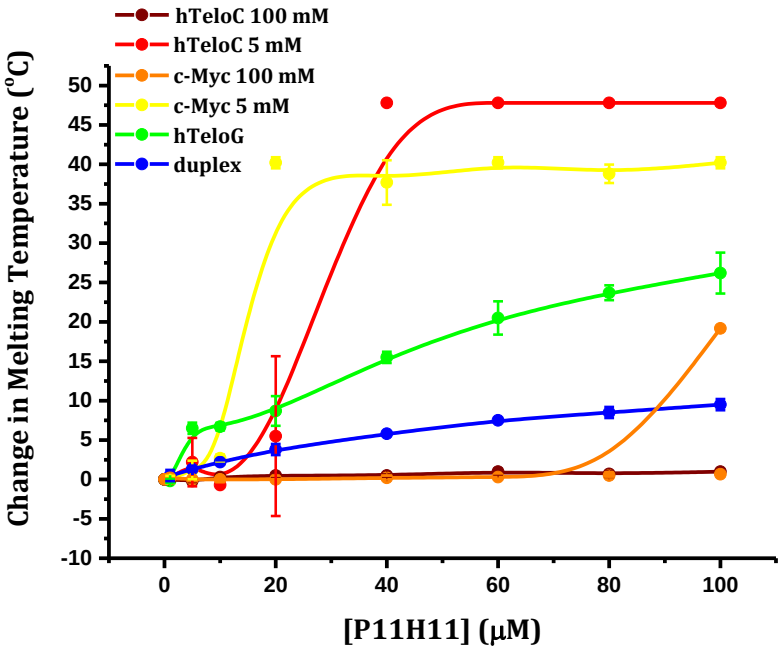



*Figure 3.27: Change in melting temperature with increasing concentration of **tilorone** with human telomeric i-motif (hTeloC), c-Myc i-motif, human telomeric G-quadruplex (hTeloG) and duplex DNA. Error bars show the standard deviation across 2 repeats. DNA = 200 nM, buffer = 10 mM sodium cacodylate and 100 mM or 5 mM NaCl. Two points are shown where biphasic melting curves were observed.*

Although **tilorone** is known to bind to duplex DNA, only a small stabilisation of duplex DNA was observed (figure 3.27). The sequence used in this study is A-T rich and **tilorone** has a preference for A-T rich DNA,[163] therefore, the compound is likely to still be binding to the DNA but without affecting the melting temperature significantly. In these results, it can be seen that **tilorone** has very little effect on either the telomeric or c-Myc i-motif structures in buffer containing 100 mM NaCl but is much more highly stabilising in buffer containing 5 mM NaCl. **Tilorone** does show some stabilisation of the c-Myc i-motif in 100 mM NaCl at 100 μM where biphasic melting curves were

observed but this is still not of the same magnitude as the stabilisation in 5 mM NaCl. Pleasingly, despite **tilorone's** planar structure, the compound does not stabilise the G-quadruplex to the same extent as the i-motif however this may again be due to the higher salt concentration. It is apparent from these results that **tilorone** is more stabilising of the i-motif than other structures at concentrations of 20 to 40 μM making it a worthwhile lead to follow, even if it is assumed the interaction is more electrostatic in nature.

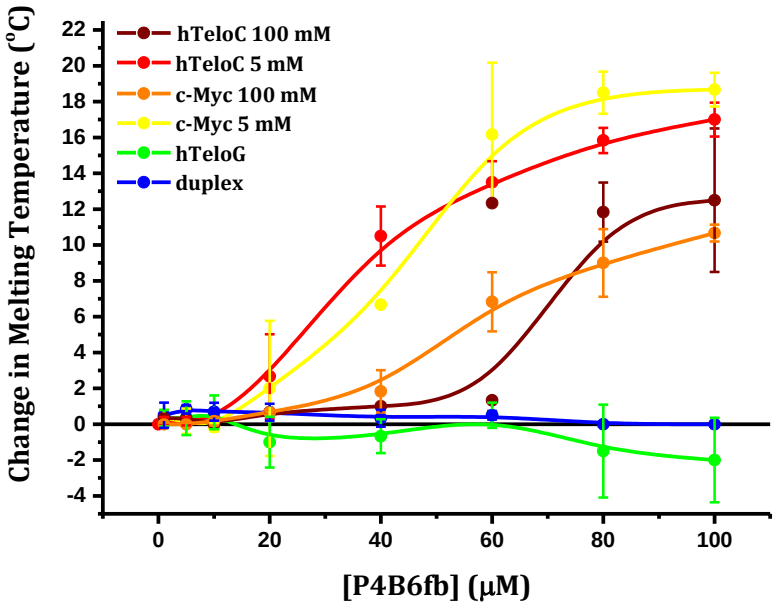



*Figure 3.28: Change in melting temperature with increasing concentration of* ***tamoxifen*** *with human telomeric i-motif (hTeloC), c-Myc i-motif, human telomeric G-quadruplex (hTeloG) and duplex DNA. Error bars show the standard deviation across 2 repeats. DNA = 200 nM, buffer = 10 mM sodium cacodylate and 100 mM or 5 mM NaCl.*

**Tamoxifen** does not stabilise the i-motif to the same extent as the previous three compounds discussed (figure 3.28), however what is pleasing is that it shows no stabilisation of duplex DNA and even shows a small destabilisation of the G-quadruplex though this is likely to be within error of the experiment. Also, the observed stabilisation of the i-motif is less affected by the difference in salt concentration compared to some of the other hits. The reason for this specificity is most likely due to the three dimensional structure of **tamoxifen**. Although the structure of **tamoxifen** consists of three phenyl rings around a carbon-carbon double bond, the structure as a whole is not planar with a large extended π-system.[167] This means it is less likely to stack on top of the G-quadruplex and is too large to be able to intercalate into duplex

DNA. However, the smaller phenyl rings are probably more able to slot into the loops of the i-motif to pick up π-stacking interactions with the loop bases.

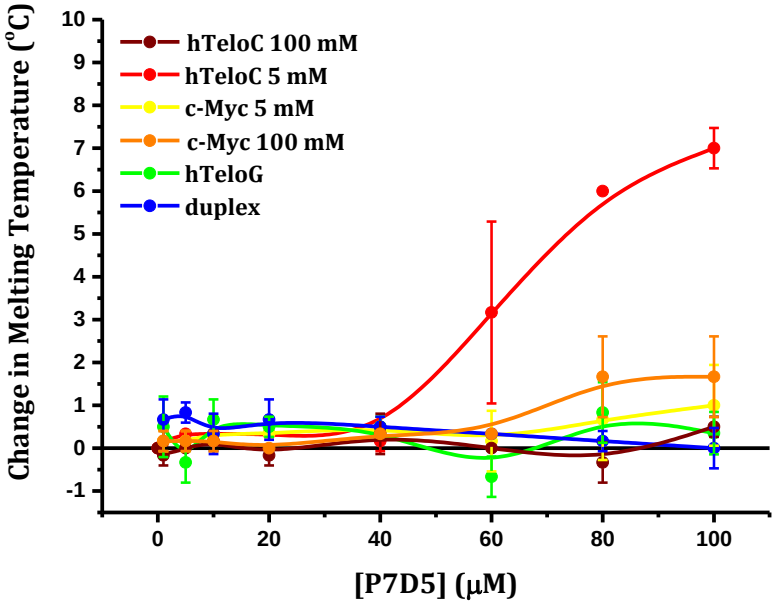



*Figure 3.29: Change in melting temperature with increasing concentration of* ***bromocriptine mesylate*** *with human telomeric i-motif (hTeloC), c-Myc i-motif, human telomeric G-quadruplex (hTeloG) and duplex DNA. Error bars show the standard deviation across 2 repeats. DNA = 200 nM, buffer = 10 mM sodium cacodylate and 100 mM or 5 mM NaCl.*

**Bromocriptine mesylate** shows only small stabilisation of the telomeric i-motif (6.7°C at 100 μM) in buffer containing 5 mM NaCl (figure 3.29). The compound shows almost no stabilisation of the other DNA structures with only a very weak stabilisation of 1.7°C for the c-Myc i-motif in 100 mM NaCl conditions. This suggests that **bromocriptine** does not interact strongly with the i-motif and most likely does so with a non-specific electrostatic interaction. This is why it shows much smaller effects in higher salt conditions as a higher concentration of sodium ions will disrupt the possible electrostatic interaction with the positive **bromocriptine** molecule and reduce its ability to bind to the DNA.

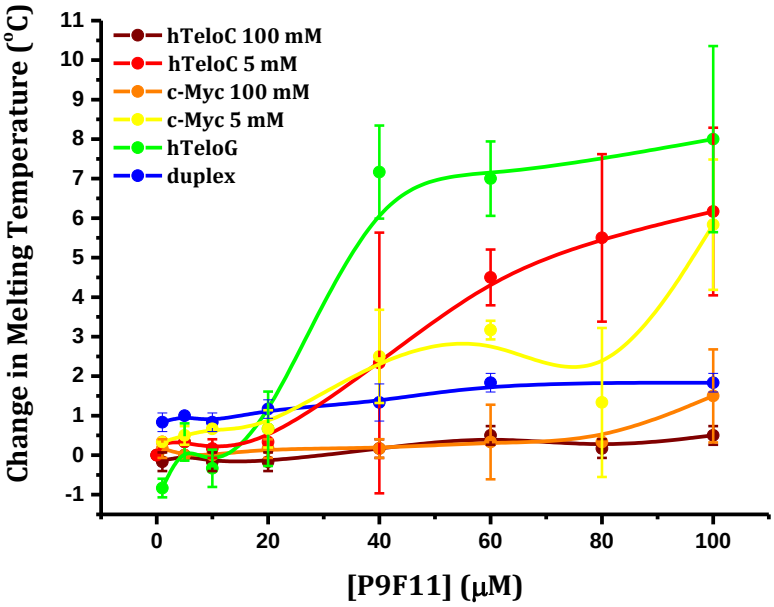



*Figure 3.30: Change in melting temperature with increasing concentration of* ***pararosaniline*** *with human telomeric i-motif (hTeloC), c-Myc i-motif, human telomeric G-quadruplex (hTeloG) and duplex DNA. Error bars show the standard deviation across 2 repeats. DNA = 200 nM, buffer = 10 mM sodium cacodylate and 100 mM or 5 mM NaCl.*

The results for **pararosaniline** (figure 3.30) show that it provides only a small stabilisation of the i-motif in low salt concentration buffer and gives a higher degree of stabilisation to the G-quadruplex. Although this is also still lower than other compounds such as **mitoxantrone**, **tilorone** and **tyrothricin**. In some ways this is surprising because **pararosaniline** has a very similar structure to **crystal violet** (figure 1.21) which is a known i-motif binding compound ($K_d$ = 0.8 – 1.4 μM).[125,127] However the stabilisation potential of **crystal violet** was not assessed so, like **pararosaniline**, it may not stabilise the i-motif to a large extent. Equally, **pararosaniline** may still have a good binding affinity for the i-motif despite not causing a large increase in stability. This highlights a drawback to screening for i-motif ligands based on their ability to stabilise the structure as some compounds which bind to the i-motif but do not affect its stability may be overlooked.

Unfortunately it is clear from these results that the most stabilising compounds **tyrothricin**, **mitoxantrone** and **tilorone** also seem to stabilise G-quadruplex and double stranded DNA. This isn't particularly surprising from looking at their structures and as discussed, some of these compounds are already known to bind to double

stranded DNA. However, each of these three ligands do appear to impart a greater stabilisation on the i-motif than other structures at lower concentrations. The other compounds **tamoxifen**, **bromocriptine** and **pararosaniline** do not stabilise G-quadruplex or double stranded DNA to the same extent but these compounds are also less strongly stabilising of the i-motif. From these results, **mitoxantrone** is the most promising ligand in terms of stabilisation potential as it imparts the greatest increase in melting temperature with concentration of ligand compared to the others. **Tilorone** and **tamoxifen** show the greatest degree of selectivity for the i-motif over G-quadruplex and duplex, particularly **tamoxifen**, which does not stabilise duplex or G-quadruplex at all. **Pararosaniline** and **bromocriptine** are the least stabilising of the 6 ligands. The fact that **bromocriptine** only increases the melting temperature of the i-motif in low salt conditions suggests that it interacts *via* a non-specific electrostatic interaction. However, the comparison of **pararosaniline** with **crystal violet** highlights that compounds may still bind to the i-motif even if they do not stabilise the structure and therefore these compounds will still be included in experiments to measure binding affinity.

## 3.6 Measuring the Binding Affinity

Having narrowed down the initial list of 34 compounds, which showed some effect on the melting temperature of the i-motif, to 6 lead compounds the next step in the investigation was to measure a binding affinity of the ligands for the i-motif. This can be done using a range of different methods as discussed in section 1.4, such as UV/Vis absorbance spectroscopy, fluorescence emission spectroscopy and surface plasmon resonance. Earlier in this chapter (section 3.4) the binding affinity of **tilorone** for the telomeric i-motif was measured using circular dichroism (figure 3.24). This gave a $K_d$ of 14 μM, assuming a 1:1 binding model. In both UV and fluorescence titrations, the absorbance or emission of the compound is observed as increasing concentrations of DNA are added and the change in the spectrum is monitored. UV titrations generally only work well if the compound itself absorbs in a different part of the spectrum to the DNA, which absorbs in a range with a maximum around 260 nm. Otherwise it is difficult to deconvolute which changes are a result of binding and which are simply an increase in concentration of the UV active components. Fluorescence titrations do not have this

drawback because if it is unlabelled, the DNA is not fluorescent. However this does rely on the ligand having some degree of fluorescence. Surface plasmon resonance does not have these constraints as this experiment measures the effect of a binding event on the plasmons within a gold surface and therefore does not depend on any spectroscopic properties of either the target structure or the binding molecule.

## 3.7 Surface Plasmon Resonance Studies on Tilorone and Mitoxantrone

Having been unable to measure a binding constant for the hits by fluorescence emission or UV absorption titrations, surface plasmon resonance (SPR) was used to measure their binding affinities. This is a highly sensitive technique as described in section 1.4.5 that measures binding interactions of compounds with a target that is immobilised on the surface of a gold chip. A binding event results in a change of refractive index above the chip surface which affects the resonance of surface plasmons in the gold chip and is therefore measured as a binding response. Importantly, this means that the measurement of a binding response does not rely on any spectroscopic property of the target or the compound being tested.

### 3.7.1 Preparation of the chip

The chip used was a streptavidin coated sensor chip purchased from GE Healthcare. This chip was chosen because biotinylated DNA can then be easily attached to the chip surface due to the extremely high affinity of biotin for streptavidin ($K_d$ ~0.01 pM).[168] The chip surface is divided into four separate flow cells which are in series with one another, this enables the simultaneous testing of ligands against up to 3 different DNA targets. The first cell was left blank as a control to account for any non-specific binding to the chip surface. In this experiment, the human telomeric i-motif sequence (5'-Biotin-d(TAACCCTAACCCTAACCCTAACCC)-3') was immobilised on flow cell 2 and the c-Myc oncogene promoter i-motif sequence (5'-Biotin-d(CCTTCCCCACCCTCCCCACCCTCCCCA)-3') on flow cell 3 in order to compare the affinity of the hits for these two different i-motif structures. Finally, to look for specificity for i-motif over double stranded DNA a 22 base pair duplex sequence (5'-

Biotin-d(GGCATAGTGCGTGGGCGTTAGC)-3' with its complement 5'-d(GCTAACGCCCACGCACTATGCC)-3') was immobilised on flow cell 4, leaving flow cell 1 blank as a control.

### 3.7.2 Initial binding response tests

Having prepared the chip surface, each compound was first tested at a concentration of 10 μM to see whether there was a binding response.

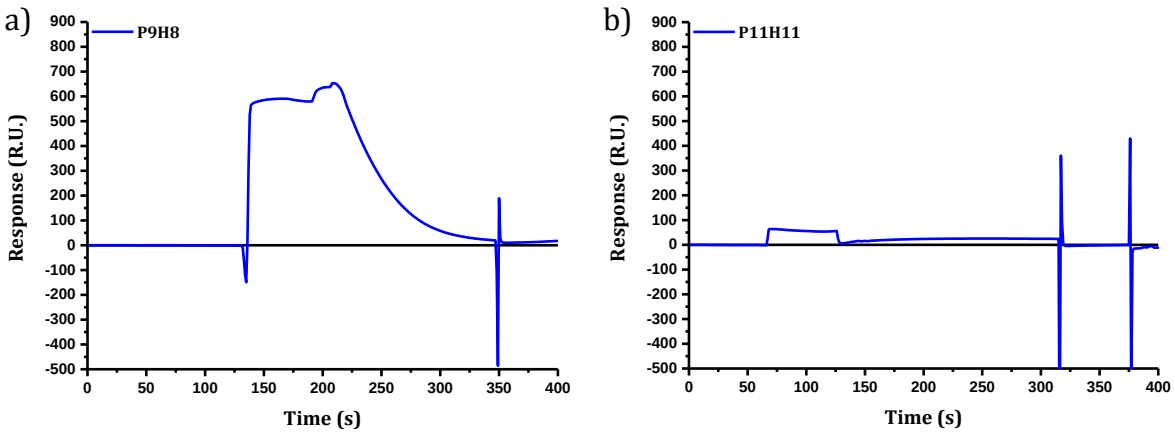



*Figure 3.31: Sensorgrams of initial binding response tests for a)* ***mitoxantrone*** *and b)* ***tilorone*** *with human telomeric i-motif. The sensorgrams are normalised with respect to a blank cell containing no DNA. Running buffer = 10 mM sodium cacodylate, 100 mM NaCl and 0.05% tween 20 at pH 5.5.*

**Mitoxantrone** (figure 3.31a) gave the clearest binding response. The response shows a sharp increase with a steep gradient after the ligand is injected which indicates a fast association rate ($k_a$). The response levels off showing that equilibrium is reached and, after continuous flow of running buffer, the response decreases as the compound dissociates. The shallower gradient indicates a slower dissociation rate ($k_d$) than the association. This is good because it indicates that the compound binds strongly as it does not become unbound as quickly. **Tilorone** (figure 3.31 b) gave a much lower magnitude binding response suggesting it has a weaker affinity for the i-motif. The response curve showed a sharp increase in the association phase and a sharp decrease in the dissociation phase indicating that the compound has fast association and dissociation rates.

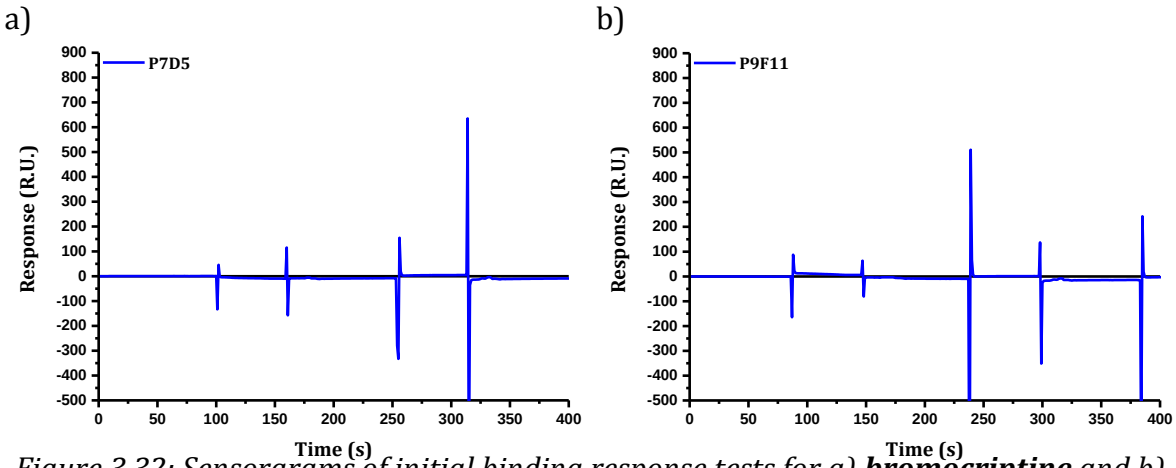



*Figure 3.32: Sensorgrams of initial binding response tests for a)* ***bromocriptine*** *and b)* ***pararosaniline*** *with human telomeric i-motif. The sensorgrams are normalised with respect to a blank cell containing no DNA. Running buffer = 10 mM sodium cacodylate, 100 mM NaCl and 0.05% tween 20 at pH 5.5.*

In contrast, the compounds **bromocriptine** and **pararosaniline** both showed almost no response at all (figure 3.32). This suggests that they do not bind to the i-motif with any significant degree of affinity. These compounds also showed very limited ability to stabilise the i-motif structure. The SPR studies were carried out in buffer containing 100 mM NaCl so it may be that the compounds would show a greater binding response in the low salt conditions but this is likely to be a more non-specific electrostatic interaction. Because of this, these two compounds were not taken forward to try and measure their low binding affinities.

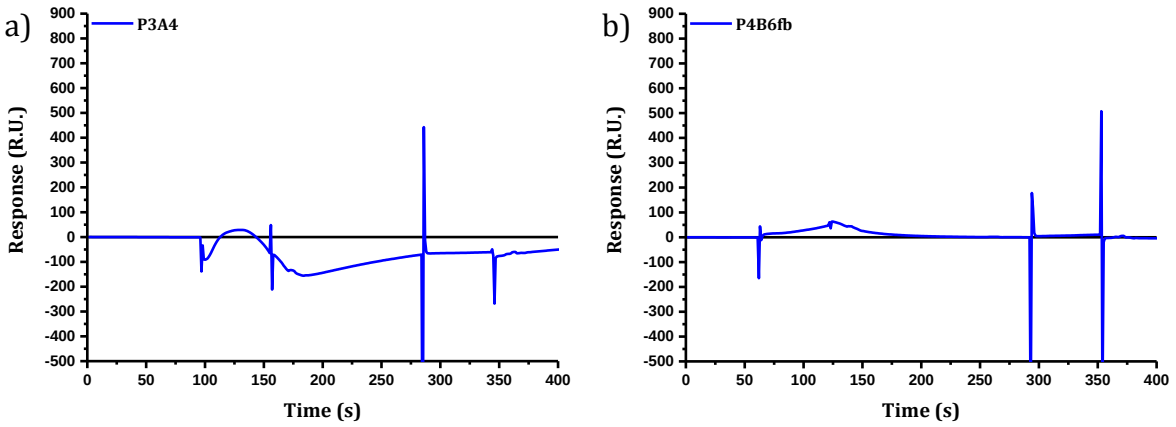



*Figure 3.33: Sensorgrams of initial binding response tests for a)* ***tyrothricin*** *and b)* ***tamoxifen*** *with human telomeric i-motif. The sensorgrams are normalised with respect to a blank cell containing no DNA. Running buffer = 10 mM sodium cacodylate, 100 mM NaCl and 0.05% tween 20 at pH 5.5.*

Finally the compounds **tyrothricin** and **tamoxifen** showed some unusual and anomalous sensorgrams (figure 3.33). These effects are due to the DMSO content of the sample because of the low solubility of the compounds. DMSO is a more dense solvent than water and introduction of DMSO into the system creates a large bulk refractive index change when it moves across the surface of the chip, this is not an effect of DMSO binding to the chip or the DNA, but a change in the composition of the buffer mixture which results in an anomalous response measurement and it is not possible to clearly measure the binding response of the compounds. This can be accounted for by using DMSO in the running buffer and performing a solvent correction to calibrate the instrument to the presence of DMSO, however, no more than 10% DMSO can be used in the running buffer due to incompatibility with the plastic seals and tubing within the instrument. Despite 10% DMSO being used in the initial FRET screen it was found that **tyrothricin** precipitated at a concentration of 100 μM in the SPR running buffer with this percentage of DMSO. Although in FRET melting studies, **tamoxifen** appeared to stabilise the i-motif more than other structures, showing a greater degree of selectivity than the other compounds, it still shows only a weak response in the SPR experiment. Therefore it was not taken forward for further analysis.

From these results, **mitoxantrone** and **tilorone** have been identified as the most promising i-motif binding compounds with the biggest binding responses. The fact that they are both water soluble also enables their binding affinity to be determined without the need to correct for DMSO. Therefore, these compounds were taken forward in order to try and measure binding affinities for the i-motif by SPR titrations.

### 3.7.3 Determining the binding affinity of mitoxantrone for i-motif by SPR.

After completing the initial binding response tests discussed in section 3.7.2, a titration was carried out over a wider range of concentrations in order to measure an accurate binding affinity. A large spread of concentrations between 0 and 100 μM was used in order to give the best chance of reaching a plateau and of having concentrations both above and below the $K_D$ as this gives the most reliable determination of affinity. The change in response over time as the compound was injected was plotted for each concentration to give a sensorgram (figure 3.34). The increase in response for each concentration plateaus out before finally decreasing again as the compound dissociates

and this point represents the equilibrium point. By measuring the change in response at the plateau with increasing hit concentration, the data can be fitted to obtain the equilibrium dissociation constant ($K_D$).[104] For each concentration, two curves are shown. These are repeat injections from the same sample and are used to show the reproducibility of the sensorgram.[169] The software uses both sets of data when fitting to find the $K_D$ with a 1 to 1 binding model (equation 3.3). An example binding curve for **mitoxantrone** with human telomeric i-motif is shown in figure 3.35. Further examples are given in the appendix (section A.2). The data was double referenced by subtracting both the response from the blank flow cell and by subtracting a series of blank injections that do not contain any compound. This accounts for non-specific binding to the chip's surface.

$$\textit{Equation 3.3:} \quad R_{eq} = \frac{K_A[\boldsymbol{Mitoxantrone}]R_{max}}{1+K_A[\boldsymbol{Mitoxantrone}]n},$$

Where $R_{eq}$ is the response at equilibrium, $R_{max}$ is the response at saturation, $K_A$ is the association constant and [**Mitoxantrone**] is the concentration of the compound. $n$ is a constant to factor for steric interference and is set to 1 assuming a 1 to 1 binding model. The equilibrium dissociation constant $K_D = \frac{1}{K_A}$.

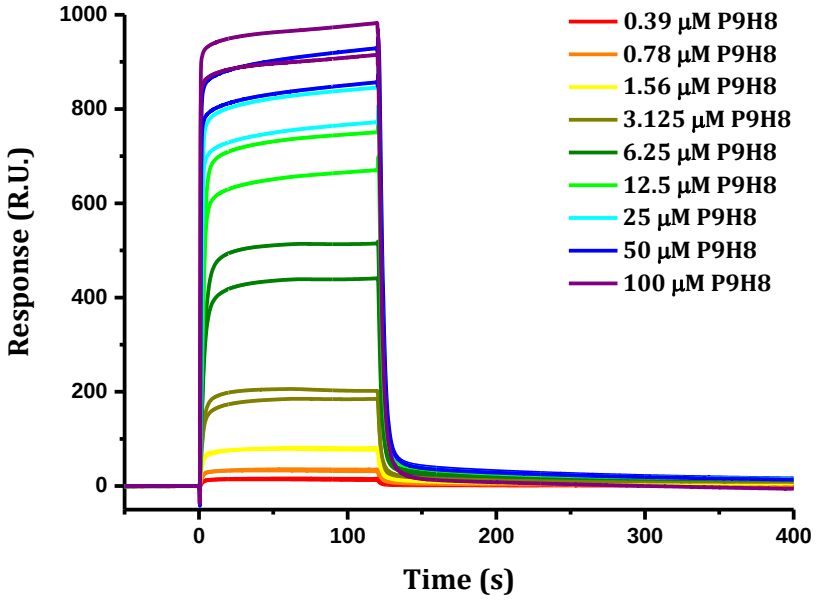



*Figure 3.34: Example sensorgram for **mitoxantrone** with human telomeric i-motif. The sensorgrams are double referenced by subtracting the response on the blank flow cell containing no DNA and by subtracting blank injections of just buffer. Running buffer = 10 mM sodium cacodylate, 100 mM NaCl and 0.05% tween 20 at pH 5.5.*

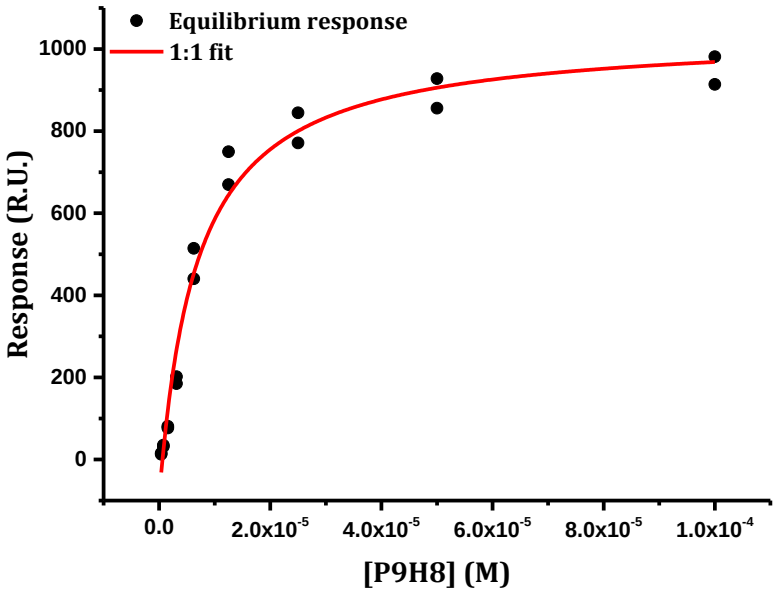



*Figure 3.35: Example SPR binding curve for **mitoxantrone** with human telomeric i-motif. Running buffer = 10 mM sodium cacodylate, 100 mM NaCl and 0.05% tween 20 at pH 5.5.*

| **Mitoxantrone** Binding Affinity $K_D$ (μM) | |
|---|---|
| hTeloC | 11.0 ± 4.3 |
| c-Myc | 15.0 ± 6.5 |
| Duplex | 19.8 ±3.0 |

*Table 3.6: Average binding affinities from 1:1 affinity fit. The error given is the maximum difference between the average and individual data.*

These results (table 3.6) show that **mitoxantrone** has a moderately strong affinity for the i-motif with dissociation constants in the low micromolar ($K_D$ = 11 and 15 μM for hTeloC and c-Myc) range and a similar though slightly better affinity for the i-motif structure over duplex DNA in this experiment. These affinities are of a similar magnitude to the affinities of existing i-motif ligands such as the phenanthrolines (4 to 8 μM)[122] and terbium complexes (22 to 30 μM).[128] This offers a good starting point for further development of analogues with better affinity and specificity.

### 3.7.4 Determining the binding affinity of tilorone by SPR

After measuring low micromolar $K_D$'s for **mitoxantrone**, the binding affinity of **tilorone** was measured using the same procedure. This compound however has a weaker affinity for the i-motif and at the top concentrations, the sensorgrams do not plateau sufficiently suggesting that they have not reached equilibrium before dissociation occurs. An example sensorgram for the human telomeric i-motif is shown in figure 3.36. The result of this is that when the data is fitted, the curve does not plateau (figure 3.37) and this creates a greater error in the determination of the binding affinity as illustrated by the greater variance between the two repeats.

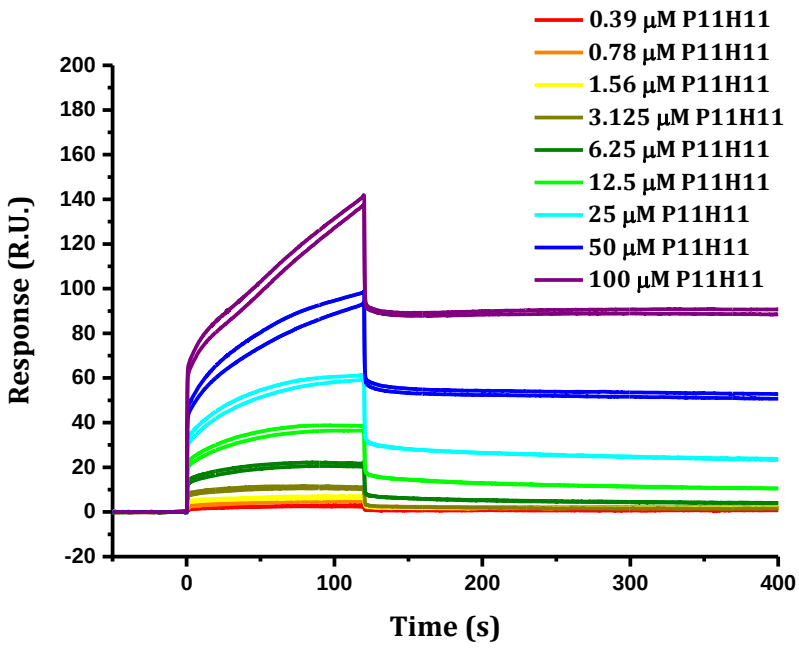



*Figure 3.36: Example sensorgram for **tilorone** with human telomeric i-motif. The sensorgrams are double referenced by subtracting the response on the blank flow cell containing no DNA and by subtracting blank injections of just buffer. Running buffer = 10 mM sodium cacodylate, 100 mM NaCl and 0.05% tween 20 at pH 5.5.*

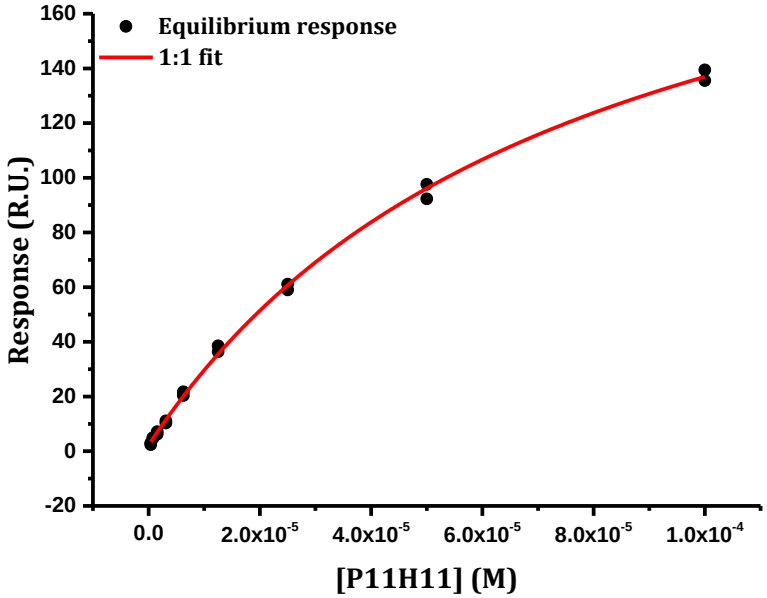



*Figure 3.37: Example binding curve for **tilorone** with human telomeric i-motif. Running buffer = 10 mM sodium cacodylate, 100 mM NaCl and 0.05% tween 20 at pH 5.5.*

| **Tilorone** Binding Affinity $K_D$ (μM) | |
|---|---|
| hTeloC | 157.9 ± 80.9 |
| c-Myc | 862.4 ± 684.7 |
| Duplex | 14.7 ± 4.3 |

*Table 3.7: Average binding affinities from 1:1 affinity fit. The error given is the maximum difference between the average and individual data.*

These results (table 3.7) show that **tilorone** has a much weaker affinity for i-motif than **mitoxantrone** ($K_D$ = 158 μM vs. 11 μM) and has a stronger preference for double stranded DNA ($K_D$ = 14.7 μM). This is unsurprising as **tilorone** is known to bind to double stranded DNA and the affinity that was measured in these experiments is actually slightly weaker than that measured previously in the literature by other methods ($K_b$ = 1.05 μM for poly(dA-dT) and $K_b$ = 7.30 μM for poly(dC-dG)).[163] Despite this, neither **tilorone** nor **mitoxantrone** have previously been shown to bind to the i-motif and they offer two new i-motif binding chemotypes which could be used as a starting point for development of analogues with improved binding to the i-motif and

could enable structure activity relationship studies to look at what features favour i-motif binding.

## 3.8 Discussion

In this chapter, a library of 960 known biologically active compounds has been screened against the human telomeric i-motif using a FRET melting based assay. The initial screen identified 34 compounds which caused some change in the melting temperature of the DNA. This list of compounds were reduced to six lead compounds through a series of FRET and circular dichroism studies. These six compounds were tested by SPR and finally, binding affinities in the low micromolar range were measured for two of the compounds **mitoxantrone** ($K_D$ = 11 μM) and **tilorone** ($K_D$ = 158 μM). **Mitoxantrone** is clearly the better of the two compounds in terms of affinity and is the primary hit from the screen. Although these two compounds are not novel structures themselves, they have never been identified as i-motif binding ligands before. Both **mitoxantrone** and **tilorone** are flat planar molecules with the ability to intercalate into DNA but it is this characteristic that enables them to stack on top of G-quadruplex or i-motif structures also. In further work on this project, it may be possible to increase the selectivity of these compounds by adding different side chains to the core structure which are able to pick up more specific interactions within the loop region of the i-motif. The identification of two new i-motif binding ligands is of importance to the field of i-motif research as they provide two new structure types which may be used as starting points for development of novel ligands with improved affinity and selectivity for the structure. In the future, analogues of these ligands may then be used as tools to probe the potential biological functions of i-motif structures.

These results have also highlighted that the FRET melting assay can be used to screen libraries of compounds for interaction with the i-motif. This method holds particular advantages for identifying compounds based on their ability to stabilise or destabilise the structure. The use of FRET also has advantages over UV melting for example, because lower concentrations of DNA and ligands are required and compounds which may absorb in the same UV region as DNA can still be examined. However, this study has highlighted several drawbacks. Some compounds may interact with the

fluorophores rather than the DNA or they may fluoresce themselves with the same excitation wavelength. For these compounds it is not possible to tell from the FRET assay whether they are interacting with the DNA or not. Another drawback is that for many compounds or for higher concentrations of most compounds, the results did not give sigmoidal melting curves making it difficult to accurately determine a melting temperature. Finally, by focusing on changes in melting temperature, this method may overlook compounds which bind to the i-motif but do not alter the melting temperature. More recently, an alternative FRET based screening method has been used by Laurence Hurley and co-workers to identify a compound interacting with the Bcl-2 i-motif.[69] Their method did not look at the melting of the DNA but looked at the change in fluorescence at room temperature as an indication of whether the i-motif is favoured (low fluorescence) or disfavoured (high fluorescence). This may be a better screening method as it holds many of the advantages of FRET melting but does not overlook compounds which do not change the melting temperature.

Although these compounds proved not to be selective for the i-motif, this is not a great surprise when screening a library of compounds with existing biological activity. For example **mitoxantrone** is a DNA intercalator used in chemotherapy for the treatment of a range of cancers including breast and lung cancer, melanoma, lymphoma, leukemia and prostate cancer. The drug's mechanism of action is to inhibit the enzyme topoisomerase II.[170] The function of this enzyme is to cleave the phosphate backbone and rejoin new DNA, relieving torsional stress in the process, **mitoxantrone** inhibits this process by intercalating at the cleavage site and trapping the complex.[171] **Mitoxantrone** has also been shown to inhibit telomerase with an $IC_{50}$ of 2 μM.[172] This is particularly interesting as inhibition of telomerase has also been demonstrated as a result of stabilising the telomeric i-motif.[59] These results suggest the possibility that **mitoxantrone** may inhibit telomerase by binding to the i-motif as well as through binding to the G-quadruplex. However, further experiments would be required to demonstrate this.

**Tilorone** is an antiviral agent that works by inducing the interferon response and has been shown to bind to AT rich double stranded DNA.[163,173,174] While more recently, analogues of **tilorone** have been investigated for their anticancer activity.[175,176] The interferon regulatory factor (IRF) genes code for transcription factors which activate

the transcription and production of interferon and their promoter regions are rich in cytosine.[177] For example, the Sp1 binding site -393 to -362 bases from the start site of the IRF-1 gene has the sequence 5'-GATATACCTCCGCCCCCGCCCCGCCAGGAGG-3', while the Sp1 binding site -383 to -350 bases from the start site of the IRF-2 gene has the sequence 5'-GCGGCCTGCGCCCCCCGCCCCGGACCCCGACTC-3'.[178] Although in these experiments, **tilorone** only showed weak binding affinity for the human telomeric and c-Myc i-motifs, the compound may well have a stronger affinity for different i-motif structures. As an example, this possibility has been demonstrated by Laurence Hurley's group when they identified a Bcl-2 i-motif binding compound that is very specific and does not interact strongly with other i-motifs such as c-Myc.[69] It has not yet been investigated whether there are any sequences capable of forming i-motifs within the IRF gene promoter regions but if there were, then the effect of **tilorone** on these structures could be investigated as a possible new mode of action.

Further work should also be carried out to investigate the interaction of **tyrothricin** with the i-motif structure. **Tyrothricin** is a natural product produced by the gram-positive bacteria *Bacillus Brevis* and is involved in the regulation of sporulation.[179] As mentioned earlier, **tyrothricin** is a mixture of peptides called gramicidins and tyrocidines and has antibacterial properties due to its ability to lyse the cell membrane.[162] Interestingly it has been shown that tyrocidine acts to reduce superhelical stress in plasmids and it was theorised that this is due to intercalation of the aromatic residues into the DNA.[179,180] However, as mentioned in chapter 1, it has also been shown that formation of the i-motif can reduce superhelical stress.[49] Although it was not possible to accurately measure a binding affinity of **tyrothricin** by SPR, it is clear that there is some response which is masked by the effect of DMSO. Likewise, the fact that the compound gave rise to bimodal FRET melting curves with a higher melting temperature clearly shows that there is some interaction between **tyrothricin** and the i-motif that is worthy of investigation. A key aspect of this would be to separate **tyrothricin** into its constituent peptides and test them separately as it may be that only one of these is responsible for the effect, or that they may have a higher magnitude affect individually than as a mixture. The early investigations into tyrocidine by Bohg et. al.[179,180] were published before the i-motif was identified so if a strong binding affinity for the i-motif was shown it would be worth investigating whether the effect of tyrocidine on superhelicity is linked to stabilisation of the i-motif.

# Chapter 4: The Effects of Cations on i-Motif DNA

# 4.1 Introduction

### 4.1.1 The effect of cations on i-motif forming DNA sequences

As discussed briefly in chapter 1, the effect of cations on i-motif DNA is an important consideration due to their abundance under physiological conditions. Despite this, there have been relatively few investigations in to the effects they may have on i-motif stability. Saxena and co-workers have investigated the effect of $Na^+$, $K^+$ and $Mg^{2+}$ ions on the c-jun i-motif forming sequence.[50] Their work showed no i-motif formation in the presence of 100 mM $Na^+$ but the i-motif did form in the presence of $K^+$ and $Mg^{2+}$. This work supports an earlier study by Jean-Louis Mergny showing that even at pH 4.8 (close to the p$K_a$ of cytosine, 4.6), the presence of 100 mM NaCl destabilised the i-motif.[51] However, they showed that increasing the sodium ion concentration further to 300 mM did not further destabilise the structure. At pH 6.4 in the presence of 100 mM $Na^+$, the addition of 5 mM $Mg^{2+}$, $Ca^{2+}$, $Zn^{2+}$, $Li^+$ or $K^+$ also caused no additional destabilisation of the structure. This may be due to the already highly unstable nature of the i-motif at this pH and in the presence of a higher concentration of $Na^+$ which could outweigh the effect of adding other cations.

More recently, the effect of $Li^+$ has been studied in more detail.[52] Sung Eun Kim and colleagues have shown that $Li^+$ ions do not prevent i-motif formation, but do promote unfolding of the structure at concentrations between 30 and 500 mM. In contradiction to the earlier studies, they showed using a single-molecule FRET experiment that analogous concentrations of $Na^+$ and $K^+$ ions did not destabilise the i-motif. However, their experiments were carried out between pH 6 and 6.4 where the i-motif is already less stable. They suggested that the reason $Li^+$ destabilises the i-motif is due to its smaller size (0.7 Å) which enables it to fit in the centre of cytosine-cytosine$^+$ base pairs that have an N3 to N3 distance of 1.8 Å,[35] yet it is unable to support cytosine-cytosine base pairing because it does not have the low energy d-orbitals needed to coordinate to both cytosine nitrogens and cannot sufficiently replicate the effects of a proton. No other extensive investigations have been carried out into the effect of cations on i-motif DNA and, with the exception of $Zn^{2+}$, these works have focused on group 1 and 2 cations and neglected transition metal and group 3 cations, several of which are of biological

relevance. Cations can interact with DNA in a variety of different ways, not just facilitating bonds between bases but also with the phosphate groups, with oxygens in the sugar moieties and between bases due to interactions with nitrogen lone pairs and the π-systems. All of these interactions can affect the stability of the DNA structure and could favour one structure over another. Therefore a wider screen of a variety of different cations would be an important step towards fully understanding the effect of cations on i-motif forming DNA sequences.

### 4.1.2 Metallated DNA

Metallated DNA is the incorporation of metal ions into DNA structure.[181,182] It is a topic that has seen a surge in recent research due to its potential applications in electronics and nanotechnology. A range of different cations have been studied including zinc, copper, nickel, mercury and silver, mostly with the aim of incorporating a number of metal-mediated base pairs into double stranded DNA.[183,184] This has enabled applications as molecular wires and sensors. Most of these applications have utilised non-natural "bases" to bind to the metal ion. For example the structure of a DNA duplex containing a sequence of consecutive imidazole-$Ag^{+}$-imidazole base pairs has been solved by NMR.[181] Non-natural substitutes for bases are also popular as they can include multi dentate ligands which bind much more strongly than the natural DNA bases.[185]

Akira Ono and co-workers showed in 2007 that mercury ions bound specifically to T-T mismatched bases forming a very stable T-$Hg^{(II)}$-T pair which stabilised double stranded DNA (figure 4.1).[186] They were able to observe $^{15}N$ – $^{15}N$ coupling across this bond using NMR and most recently obtained a crystal structure of duplex DNA containing two consecutive T-$Hg^{(II)}$-T base pairs.[182,187,188] Another group have also shown that mercury ions can facilitate T-T mispairing in a DNA polymerase reaction.[189]

As well as T-T base pairs stabilised by mercury it has also been demonstrated that silver cations can stabilise C-C,[190] C-A and C-T mismatches.[191,192] Akira Ono and colleagues showed that both silver and mercury cations were able to stabilise a C-T mismatch in duplex DNA by around 4 °C and 7 °C respectively (figure 4.1).[191] Later, it was shown that silver cations were able to stabilise a C-A mismatch, enabling the

formation of a full length replication product by a DNA polymerase using a template containing cytosine and dATP.[192] In the absence of silver, the polymerase reaction was stopped at the cytosine position and if the system was inverted so that the template contained an adenine base, no incorporation of cytosine occurred even in the presence of silver. The group hypothesised that the C-A mispairing occurs by using the Hoogsteen face of adenine and that when adenine is incorporated into the primer it is unable to rotate sufficiently to enable this interaction (figure 4.1).

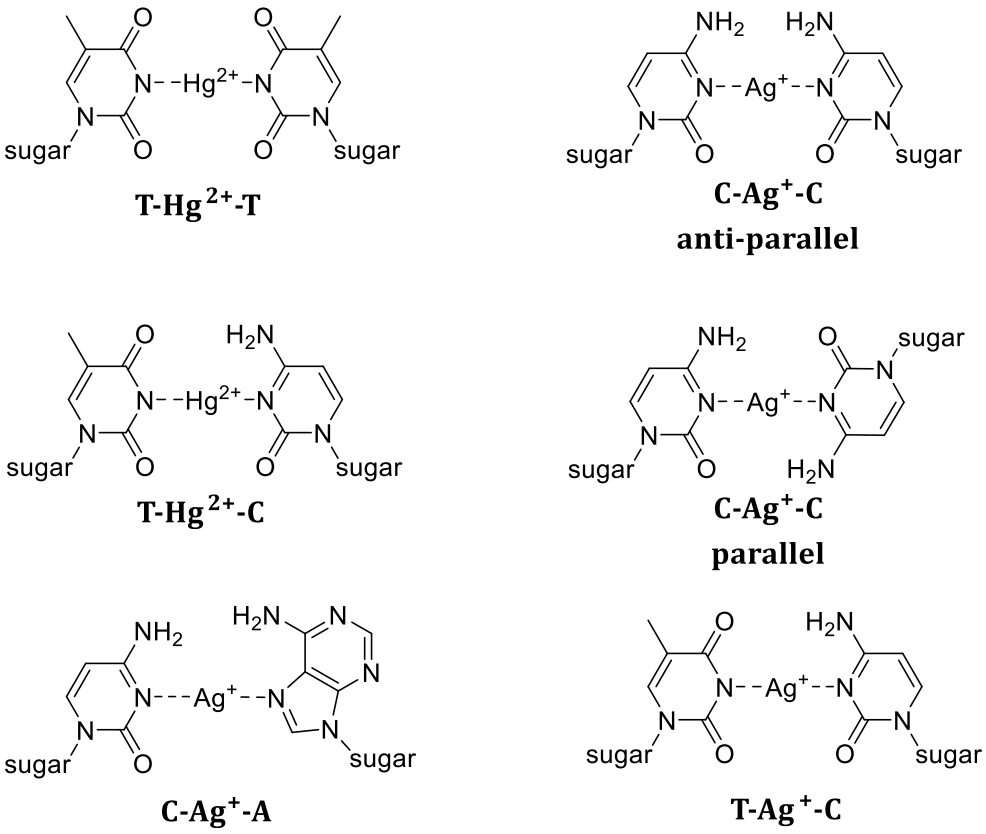



*Figure 4.1: Proposed structures of mismatched base pairs stabilised by Mercury (II) and Silver (I) cations.*

Of these structures, only the T-$Hg^{2+}$-T base pair has been structurally elucidated by NMR and X-ray crystallography.[182,188] As Akira Ono and co-workers have been looking at these base pairs in the context of double stranded DNA, the structures they have proposed have the bases orientated in the same way as natural base pairs in an anti-parallel duplex with both deoxyribose units on the same side. In the context of non-duplex DNA structures, the strands may interact in a parallel direction such as in the structure of the i-motif.[35] In this case the pairs form in the reverse orientation.[191]

### 4.1.3 Silver stabilises C-C base pairs

As a follow up to their previous work on mercury mediated thymine-thymine base pairs in DNA, more specific research has been carried out by Akira Ono and colleagues who showed that silver (I) cations are able to bind specifically to C-C base pairs and stabilise a C-C mismatch containing duplex (figure 4.1).[190] They then utilised this discovery to produce a DNA based sensor for silver cations where the presence of silver stabilises the formation of a DNA hairpin using C-C mismatched pairs. The hairpin sequence is labelled with two fluorophores which show FRET upon hairpin formation. This work has been applied by other groups to DNA based sensors for silver as well as the amino acid cysteine, a strong chelator of silver.[193] Rather than a DNA hairpin, Wan Yi Xie *et. al.* used silver to promote formation of duplex DNA between two separate strands. One of the strands was labelled with a FAM fluorophore next to the cytosine base and upon formation of the cytosine-$Ag^{+}$-cytosine base pair the fluorescence was quenched due to a photoinduced electron transfer (PET) effect. In the presence of cysteine, the silver is chelated and no longer facilitates the C-C mismatch base pair resulting in strand dissociation and an increase in FAM fluorescence.[193] The same group went on to utilise this effect in the design of a DNA based logic gate.[194]

Some investigations have already been published on the interaction between the i-motif and silver but these have only looked at the formation of silver nanoclusters rather than stabilisation of the i-motif itself. Small nanoclusters of silver have strong fluorescent properties and are synthesised using oligonucleotide templates which enable the formation of clusters with different emission wavelengths.[195] Bidisha Sengupta and colleagues used two different i-motif forming sequences $(dTA_2C_4)_4$ and $(dC_4A_2)_3C_4$ to template the formation of silver nanoclusters with red or green fluorescence depending on experimental conditions. They showed that formation of clusters with red fluorescence were favoured at acidic pH when the i-motif is stable and that clusters with green fluorescence were favoured at basic pH when the i-motif sequence is expected to be unfolded. Interestingly, they show using size exclusion chromatography that the DNA forms a folded i-motif conformation at acidic pH, but that at higher pHs the strand-cluster conjugate appears to take the same type of structure.[195] This is an early indication that silver cations may enable i-motif formation

and Sengupta et. al. postulate that an i-motif like structure may be forming as the DNA wraps around the silver clusters.

Wei Li and co-workers have also completed a series of studies on the i-motif as a template for silver nanocluster formation, in comparison to other DNA secondary structures, namely single stranded DNA, G-quadruplex and duplex.[196–198] Their results suggest that the i-motif is a better template than G-quadruplex and double stranded DNA, stabilising nanoclusters for a longer period of time.[198] They used UV melting to measure a decrease in melting temperature of the dimeric i-motif d[$C_4A_4C_4$] from $T_m$ = 43.5 °C to $T_m$ = 34.4 °C upon addition of $Ag^+$ which they suggest is due to the binding of $Ag^+$ to the N3 of cytosine and disrupting the C-$H^+$-C base pairing. At pH 7.0 they recorded CD spectra which suggests possible i-motif formation with a shift in the peak from 275 nm to 287 nm upon addition of $Ag^+$ or Ag, which they explain is due to formation of C-Ag-C base pairs.[196] With these hints and suggestions at i-motif formation in the presence of silver, a specific study into whether silver cations stabilise i-motif formation was necessary.

### 4.1.4 Aims and Objectives

As described in section 4.1.1, so far only a small number of investigations have been carried out to look at the effect of different cations on i-motif formation and these studies have only focused on a small selection of group 1 and group 2 cations. Akira Ono and co-workers have looked in more detail at metal-mediated base pairs showing that mercury (II) and silver (I) ions are able to stabilise T-T and C-C base pairs respectively. However, the focus of their work has been on duplex and hairpin DNA structures and no other secondary structures. The interaction between i-motif forming oligonucleotide sequences and silver (I) cations has been studied in the context of silver nanocluster formation (section 4.1.3), which has indicated the presence of folded structures at above neutral pH. Despite this, no detailed investigation has been carried out to look at i-motif formation in the presence of silver cations. This is the aim of the first part of this chapter. In the second part of the chapter, this work is extended to look at the effect of a wide range of other cations from across the periodic table, many of which are of biological relevance, rather than just group 1 and 2 cations which have

been studied previously. The objective of this research is to give a fuller understanding of the effect of different cationic conditions on i-motif formation.

## 4.2 The Effect of Silver (I) Ions on i-Motif Forming DNA.

### 4.2.1 FRET melting

Using the established FRET melting experiment used in chapters 2 and 3, the human telomeric i-motif forming sequence (hTeloC) was tested in 10 mM sodium cacodylate, 5 mM sodium chloride buffer to investigate the interaction between silver cations and i-motif forming DNA. A low salt concentration was used so that any effects of silver cations were not overwhelmed by the presence of a large amount of sodium ions. Previous FRET studies within the Waller group were carried out with up to 1 M $AgNO_3$, which showed that even at an $Ag^+$ concentration of 100 µM, the DNA did not melt. Therefore in these experiments, a range of concentrations of $AgNO_3$ were tested between 0 and 100 µM. $AgNO_3$ was used because this salt is very water soluble, whereas other

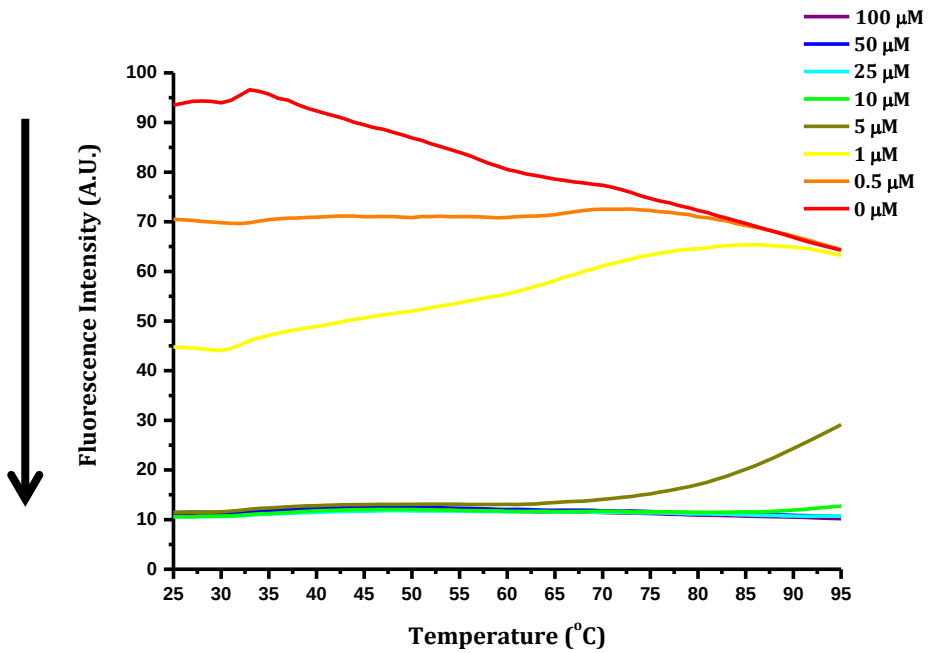



*Figure 4.2: FRET melting curves for human telomeric i-motif with increasing concentration of $AgNO_3$ at pH 7.4. [DNA] = 200 nM, buffer = 10 mM sodium cacodylate and 5 mM NaCl. Figure adapted from Henry Day et. al.*[199] *by permission of The Royal Society of Chemistry.*

salts such as the halides have a much lower solubility. At pH 5.5 the hTeloC sequence forms a stable i-motif structure, while at pH 7.4 the oligonucleotide exists as a random coil. Therefore silver (I) cations were tested at both pHs to show any effects on an existing i-motif structure and to see whether silver could induce DNA folding. This would be apparent by a decrease in fluorescence signal as shown in figure 4.2.

At pH 7.4, it can be seen that in the absence of any silver cations, there is a high fluorescence signal and no melting transition, consistent with the DNA taking a random coil conformation (Figure 4.2). During the experiment, as the concentration of silver ions is increased, there is a decrease in fluorescence signal, indicating that the fluorophores are coming into close proximity. This indicates that a folding event has occurred and at a $AgNO_3$ concentration of 1μM a melting transition can be observed with $T_m$ = 65 °C indicating that a stable folded structure has formed which then melts with increasing temperature. At 5 μM the folded DNA is strongly stabilised only beginning to melt at temperatures above 80 °C and at even higher concentrations, there is a low intensity fluorescence signal which does not show any sign of melting within

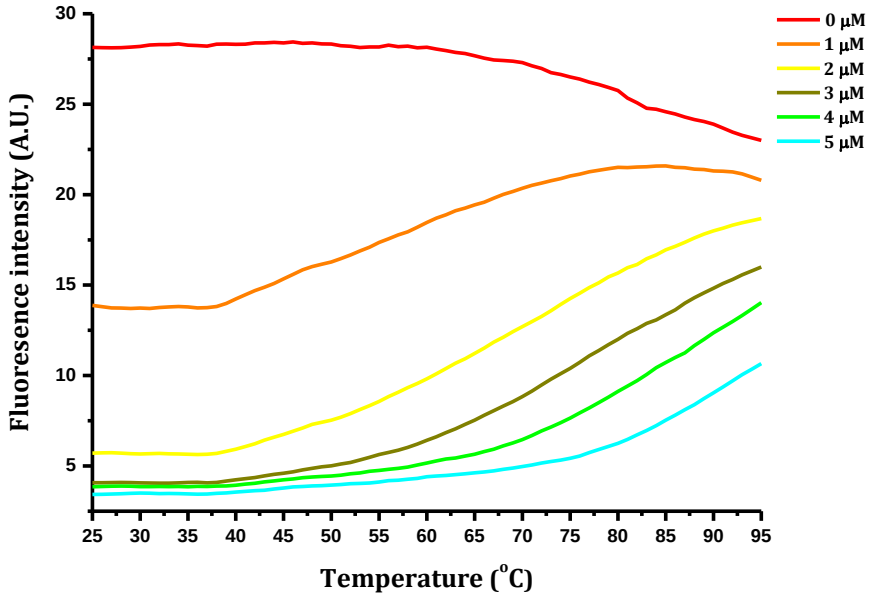



*Figure 4.3: Later work showing FRET melting curves for human telomeric i-motif with $AgNO_3$ between 0 and 5 μM carried out by Zoë Waller. [DNA] = 200 nM, buffer = 10 mM sodium cacodylate and 5 mM NaCl. Figure adapted from Henry Day et. al.*[199] *by permission of The Royal Society of Chemistry.*

the temperature range of the experiment ($T_m$ > 95°C). Later studies within the group examined the change in melting temperature at pH 7.4 between 0 and 5 μM silver nitrate (figure 4.3). These results showed a progressive increase in melting temperature with increasing silver nitrate concentration until at 4 and 5 μM, the folded DNA does not melt within the temperature range of the experiment.

These results suggest that the presence of silver cations can induce the DNA to fold and that increasing concentrations of $Ag^+$ increase the stability of the structure. It is not possible to tell from the FRET melting experiments alone, whether or not the folded structure is actually an i-motif. The low intensity fluorescence at 100 μM $Ag^+$ may also be the result of precipitation as has been shown in the previous chapters.

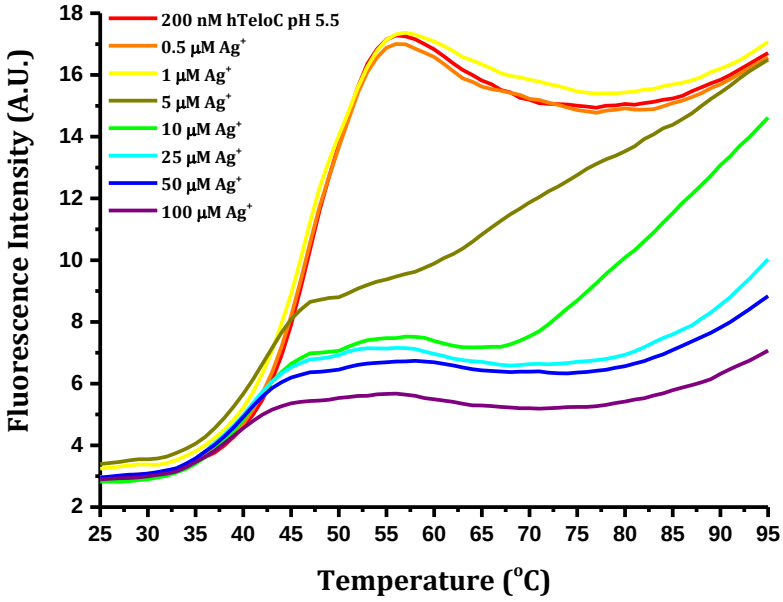



*Figure 4.4: FRET melting curves for human telomeric i-motif with increasing concentration of $AgNO_3$ at pH 5.5. [DNA] = 200 nM, buffer = 10 mM sodium cacodylate and 5 mM NaCl.*

At pH 5.5, in the absence of silver (I) ions, the i-motif shows a single melting transition at $T_m$ = 47 °C (figure 4.4). With the addition of silver nitrate, at low concentrations there was no change in melting temperature, but at concentrations of 5 μM and above, biphasic melting curves are observed with two melting temperatures, one lower than the i-motif alone ($T_m$ = 43 °C) and one at a higher temperature ($T_m$ = 65 °C) consistent with the decrease in melting temperature observed by Yan Fu and co-workers.[196]

Interestingly, the higher melting point is the same as that observed at neutral pH in figure 4.2, suggesting that the same structures or ones with similar stability are forming in both cases. The presence of biphasic melting curves indicates that two different conformations of DNA are present. The lower melting temperature could be an otherwise normal i-motif formed using cytosine-$H^+$-cytosine base pairs that has simply been destabilised by the presence of a higher salt concentration. Whilst the higher melting temperature may represent an alternative structure such as an i-motif or hairpin stabilised by cytosine-$Ag^+$-cytosine base pairs. Alternatively one or both melting temperatures may represent alternative structures stabilised by different mixtures of C-$H^+$-C and C-$Ag^+$-C base pairs.

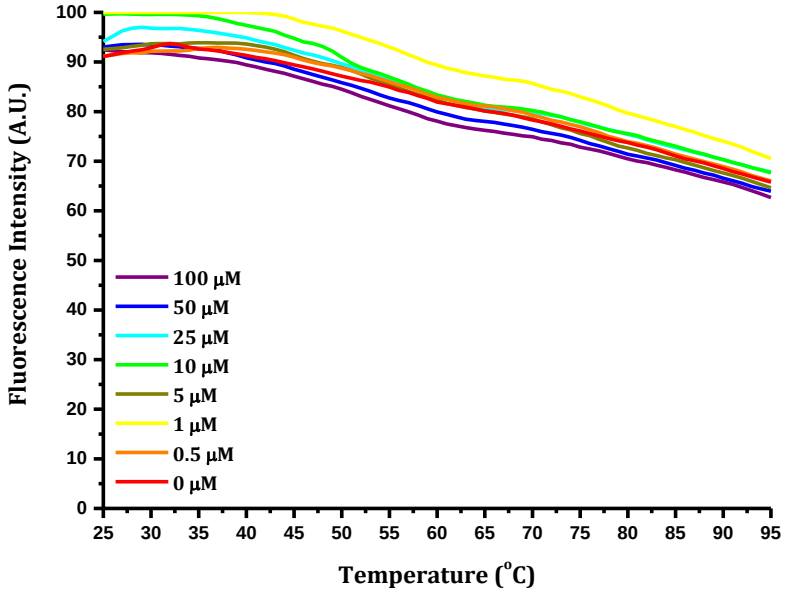



*Figure 4.5: FRET melting curves for human telomeric i-motif with increasing concentration of $NaNO_3$ at pH 7.4. [DNA] = 200 nM, buffer = 10 mM sodium cacodylate and 5 mM NaCl. Figure adapted from Henry Day et. al.*[199] *by permission of The Royal Society of Chemistry.*

To account for any effects from the nitrate anion, sodium nitrate was tested at the same concentrations in both pH 7.4 and pH 5.5 buffers as a control (figure 4.5 and 4.6). From these tests it can be seen that the nitrate anion did not induce any folding of the i-motif sequence at pH 7.4 and had no effect on the stability of the i-motif structure at pH 5.5. This shows that the observed stabilisation of the i-motif sequence in figures 4.2 and 4.4 is solely due to the presence of $Ag^+$ cations.

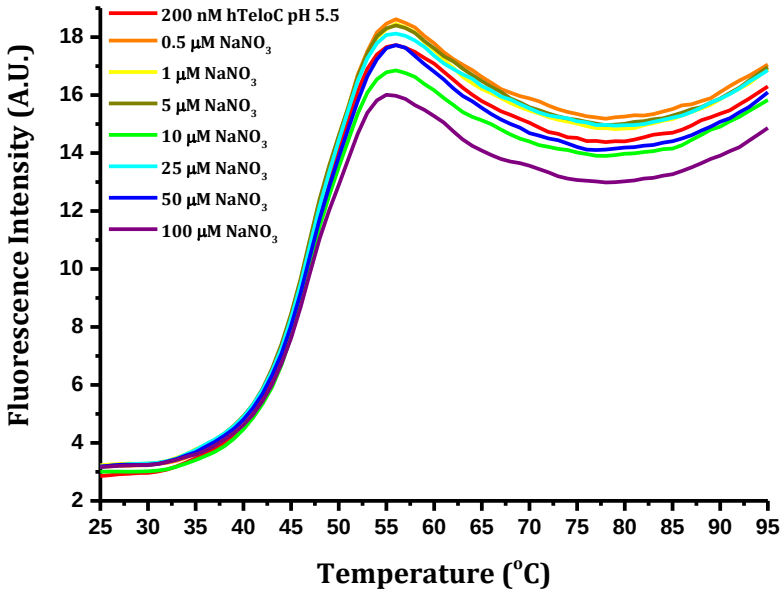



*Figure 4.6: FRET melting curves for human telomeric i-motif with increasing concentration of $NaNO_3$ at pH 5.5. [DNA] = 200 nM, buffer = 10 mM sodium cacodylate and 5 mM NaCl. Figure adapted from Henry Day et. al.*[199] *by permission of The Royal Society of Chemistry.*

In summary, the FRET melting data demonstrates that at pH 7.4, when the DNA starts in a random coil conformation, increasing the concentration of silver (I) ions causes the DNA to fold into an increasingly stable structure. At pH 5.5 where the i-motif structure is already stable, addition of increasing amounts of silver (I) ions leads to the appearance of biphasic melting curves, suggesting the presence of two different structures with different stabilities. As already described, these results are clearly due to the presence of the silver cations rather than the nitrate anion as control experiments with sodium nitrate showed it had no effect. These FRET melting results clearly show that silver (I) ions interact with the human telomeric i-motif forming sequence. To investigate this interaction in more detail, further experiments are required which are discussed in the next sections.

### 4.2.2 Circular Dichroism

As discussed in previous chapters, circular dichroism spectroscopy gives a specific signal dependent on the type of DNA secondary structure present in solution and therefore enables the monitoring of structural changes.[102] The CD spectrum of single

stranded DNA has a positive peak at 275 nm and a negative peak at 250 nm. The CD spectrum of the i-motif is significantly different with a positive peak at 288 nm and a negative peak at 255 nm. The DNA was tested at a concentration of 10 μM in 10 mM sodium cacodylate 5 mM NaCl buffer at pH 7.4 (figure 4.7). These conditions were chosen so that any observed changes in structure could be related to the apparent folding effect that was seen in the FRET melting experiment in figure 4.2.

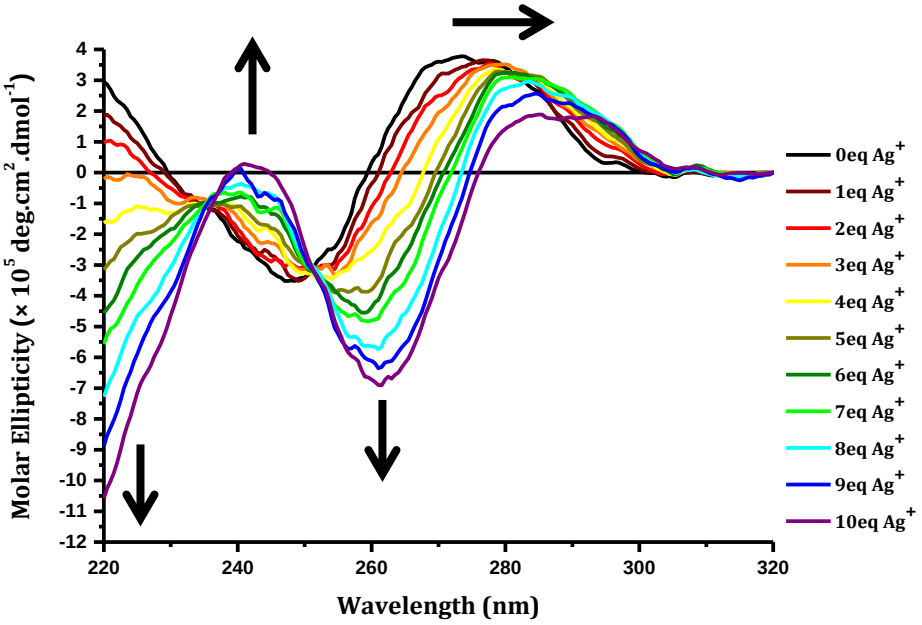



*Figure 4.7: CD spectrum of human telomeric i-motif sequence with increasing concentration of $AgNO_3$ at pH 7.4. [DNA] = 10 μM, buffer = 10 mM sodium cacodylate and 5 mM NaCl.*

In the absence of silver (I) ions, the DNA gave a spectrum with a positive peak at 275 nm and a negative peak at 250 nm indicative of a random coil structure (figure 4.7). Silver nitrate was titrated into the sample in 1 equivalent portions up to a concentration of 10 equivalents (100 μM). With increasing concentrations of $Ag^+$, the spectrum clearly shows a structural change with the positive peak undergoing a bathochromic shift from 275 nm to a broad peak at 288 nm while the negative peak also shifts from 250 nm to 260 nm (figure 4.7), this is indicative of formation of an i-motif-like structure. As well as these changes, the molar ellipticity at 220 nm switches from positive to negative and there is the appearance of a new positive peak at 241 nm. Whilst the other peaks are characteristic of i-motif formation, it is unlikely that a silver stabilised i-motif structure would give an identical CD spectrum to an acid stabilised i-motif due to the need to accommodate the $Ag^+$ cation (figure 4.8). The distance between

cytosine base pairs in the i-motif is 1.8 Å,[35] whereas the $Ag^+$-N distance in $Ag(NH_3)_2NO_3$ has been shown to be 2.07 Å.[200] This means the distance between cytosines in a C-$Ag^+$-C base pair could be more than double that of a normal i-motif, which would affect the glycosidic bond angles responsible for the optical activity observed in the CD experiment.

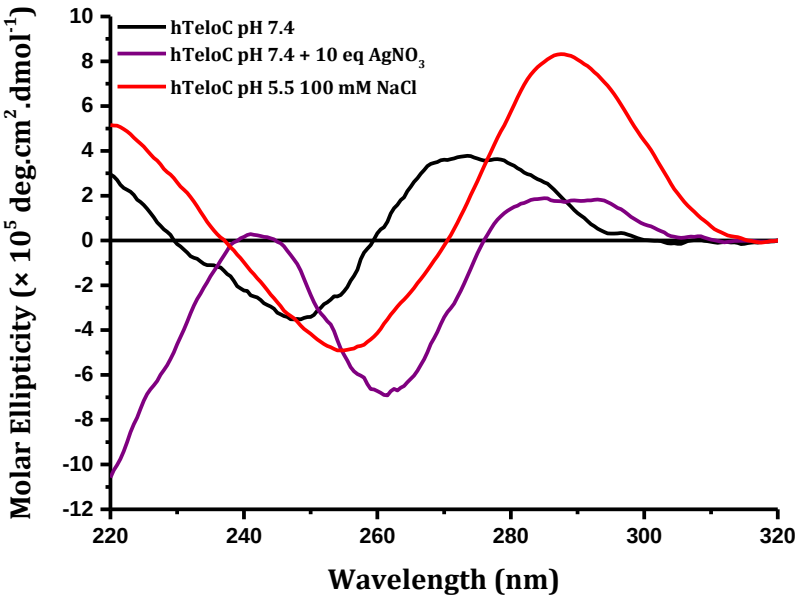



*Figure 4.8: Comparison of CD spectra for hTeloC in single stranded (pH 7.4) and i-motif (pH 5.5) conformations to the spectrum of the silver induced i-motif-like structure. [DNA] = 10µM, [$AgNO_3$] = 100 µM, buffer = 10 mM sodium cacodylate and 5 mM NaCl at pH 7.4, buffer = 10 mM sodium cacodylate and 100 mM NaCl at pH 5.5.*

The observed CD spectrum of hTeloC in the presence of $Ag^+$ at pH 7.4, is remarkably similar to that published by Jiangjiexing Wu and co-workers as part of a study on the growth of fluorescent silver clusters.[197] Their spectrum of the i-motif with increasing concentration of silver also showed the formation of a new peak at 240 nm. Another investigation published by Yan Fu and colleagues looked at G-/C-rich DNA as a template for silver nanocluster formation, their CD data for the sequence $C_4A_4C_4$ shows bimolecular i-motif formation at pH 5.0 with and without silver cations, while at pH 7.0 the sequence $G_4T_4G_4$-$C_4A_4C_4$ shows a shift to 287 nm in the presence of $Ag^+$ which could indicate i-motif formation in the C-rich strand.[196] They also measure a CD spectrum of just $C_4A_4C_4$ at pH 7.0 which shows a random coil conformation but the presence of $Ag^+$ results in peaks at 287 nm, 260 nm and 240 nm consistent with the spectrum in figure

4.7. These observations support the suggestion that i-motif formation is occurring in the presence of silver cations at neutral pH.

In order to measure the stoichiometry of the interaction with silver, the method of continuous variation binding analysis was used to create a Job plot (figure 4.9).[201] A series of samples were prepared where the ratio of $Ag^+$:DNA was varied between 0:1 and 10:1 but the overall concentration of DNA + $Ag^+$ was kept constant at 20 μM and the CD spectrum of each sample was measured. An analogous set of samples with the same DNA concentrations but in the absence of $Ag^+$ were also tested this data was subtracted from the first set to give the change in molar ellipticity as a result of the silver (I) ions. Plotting the change in molar ellipticity at 275 nm against the molar ratio of DNA gives a graph with a change in gradient which corresponds to the stoichiometry of the interaction, this is known as a Job plot.[201]

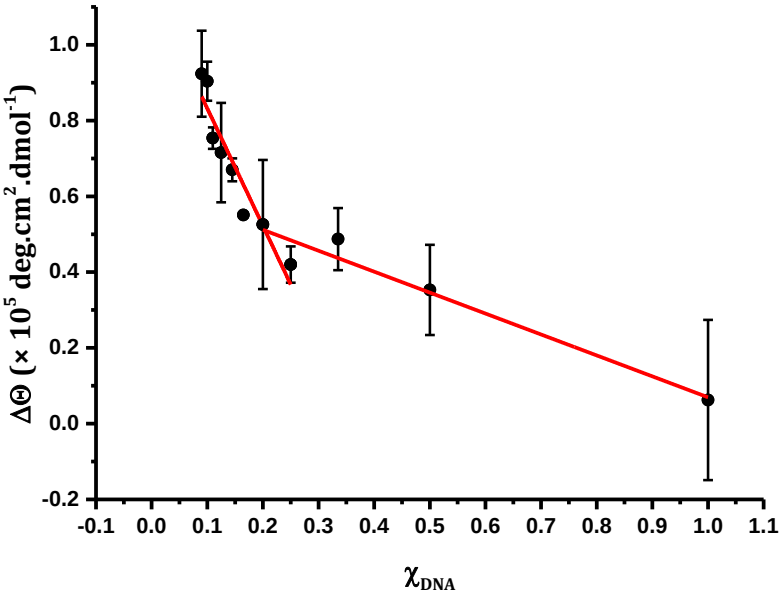



*Figure 4.9: Job plot showing the stoichiometry of the interaction between silver and human telomeric i-motif DNA. Error bars show the standard deviation across 2 repeats [DNA] = 0 – 20 μM, [$AgNO_3$] = 0 – 20 μM, buffer = 10 mM sodium cacodylate and 5 mM NaCl at pH 7.4.*

To calculate the stoichiometry the data was fitted with two linear lines of best fit which intersect. Solving the two linear equations simultaneously gives the molar ratio of DNA at the inflection point, which can then be converted into a stoichiometry. Analysis of these results led to an average stoichiometry of 4:1 $Ag^+$:DNA. This is surprising as there

are 6 possible cytosine-cytosine pairs in the i-motif so it is reasonable to expect that the stoichiometry would be 6. However, the average stacking interval between cytosine pairs in a normal i-motif structure is 3.1 Å,[35] but the distance between silver (I) ions in a duplex containing silver imidazole pairs was shown to be between 3.79 and 4.51 Å.[181] This suggests that incorporation of 6 cytosine-$Ag^+$-cytosine base pairs would cause significant perturbation of the structure and may be unfavourable. This result suggests that not all the possible cytosine-$Ag^+$-cytosine pairs need to form in order to stabilise an i-motif-like structure. That may be because only 4 cytosine-$Ag^+$-cytosine pairs, provide enough stability or it may suggest that the silver is stabilising the structure in a different way, perhaps by forming a nanocluster which the DNA then wraps around in an i-motif-like formation.

### 4.2.3 Reversibility

In previous published studies on the interaction of silver and cytosine it has been well documented that changes can be reversed by the use of cysteine to chelate the silver and remove it from the C-$Ag^+$-C pair.[194] For example, Wan Yi Xie and colleagues used a double stranded DNA probe as a sensor for silver based on the formation of C-$Ag^+$-C base pairs which caused the quenching of a fluorescent label upon formation of the double stranded DNA but in the presence of cysteine, this change was reversed, restoring the fluorescence.[193] They later applied this method to produce a DNA based logic gate for the detection of silver (I) ions and cysteine.[194] Around the same time, Chun-Xia Tang and co-workers also developed a different double stranded DNA based electrochemiluminescent sensor for the recognition of silver and cysteine.[202]

In order to examine whether L-cysteine could be used to reverse the apparent i-motif formation induced by silver ions, a CD titration was carried out on the human telomeric i-motif sequence in the presence of 10 equivalents of $Ag^+$ with increasing concentrations of L-cysteine in 1 equivalent portions. It can be seen from the resultant spectrum (figure 4.10) that L-cysteine did effectively reverse the changes caused by silver with a shift in the positive peak back to 275 nm, the negative peak back to 250 nm and a disappearance of the positive peak at 240 nm. After addition of 10 equivalents of L-cysteine shown in purple, the CD spectrum is almost identical to the starting spectrum of single stranded DNA, shown in grey.

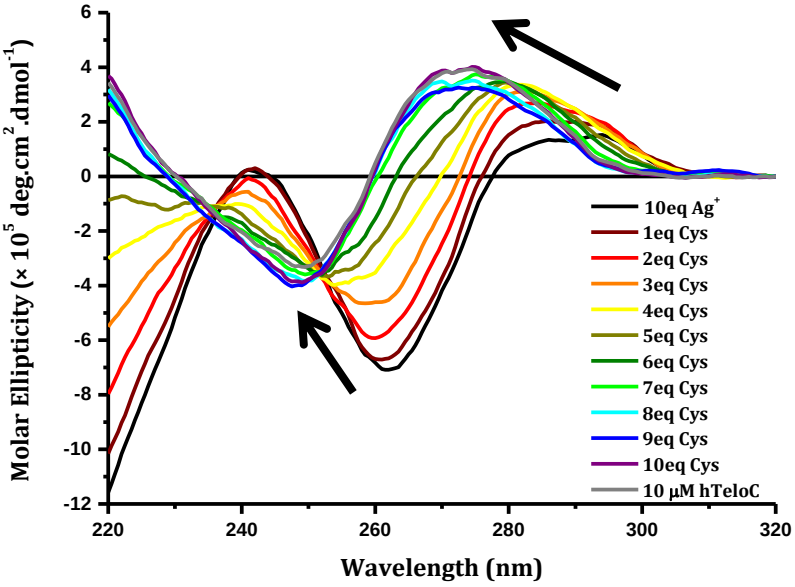



*Figure 4.10: CD spectrum of human telomeric i-motif sequence with 100 μM $AgNO_3$ and increasing concentration of L-Cysteine at pH 7.4. DNA: 10 μM, buffer = 10 mM sodium cacodylate and 5 mM NaCl. Figure adapted from Henry Day et. al.*[199] *by permission of The Royal Society of Chemistry.*

Having seen that equimolar amounts of cysteine can reverse the changes induced by silver and return the DNA to its single stranded form, further addition of silver was then able to return the DNA to the folded form. It was then shown that repeated addition of silver and cysteine could repeatedly fold and unfold the DNA. This could be seen in the CD spectrum itself and the molar ellipticity was measured at 250 nm and 261 nm after sequential addition of 10 eq. $Ag^+$ and 10 eq. cysteine showing repeated changes in the molar ellipticity (figure 4.11). This suggests the possibility that silver and cysteine could be used as an alternative switch mechanism for the controllable folding and unfolding of i-motif DNA rather than changing pH which is what has been used throughout other published examples.[203] The recognition of silver and cysteine by the telomeric i-motif sequence could also be developed as a sensor for silver or cysteine similar to other examples involving hairpin or duplex DNA.[190,193,194,202]

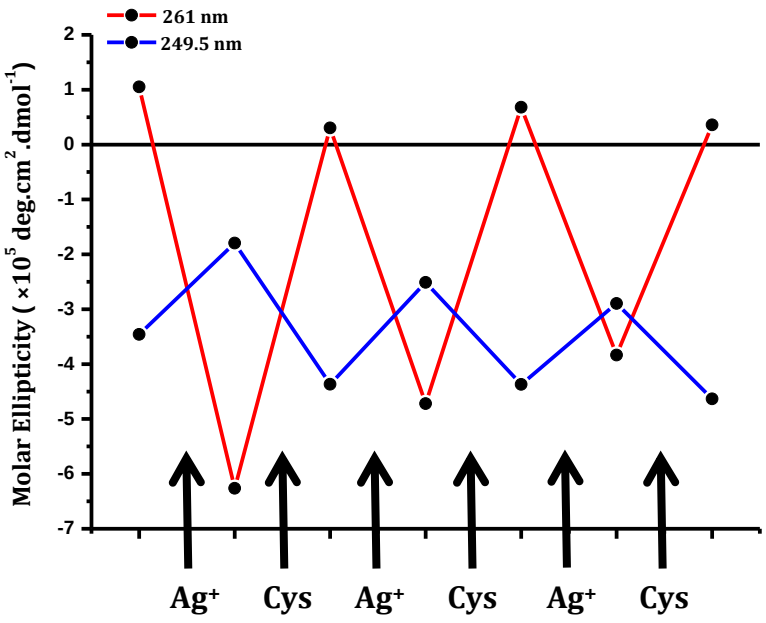



*Figure 4.11: Change in molar ellipticity of human telomeric i-motif sequence with repeat additions of 10 eq of $AgNO_3$ and 10 eq of cysteine at pH 7.4. [DNA] = 10 μM, buffer = 10 mM sodium cacodylate and 5 mM NaCl. Figure adapted from Henry Day et. al.*[199] *by permission of The Royal Society of Chemistry.*

These results show that the apparent folding and change in structure of the telomeric sequence in the presence of silver (I) cations is reversible by addition of equal amounts of L-cysteine. Not only did the addition of cysteine return the CD spectrum to that previously shown for single stranded DNA but it was possible to repeatedly fold and unfold the DNA by successive additions of silver (I) ions and cysteine. This is important as the process needs to be reversible for it to find use in a nanotechnology application such as the sensing of silver or cysteine.[193,194]

### 4.2.4 FRET Titration Experiments

As described in section 1.4.4, another way in which the folding of the DNA can be monitored is by using the fluorescently labelled DNA in a FRET titration experiment. By exciting the FRET labelled human telomeric i-motif sequence at 490 nm and measuring the fluorescence emission from 500 to 650 nm, the fluorescence signal from both the FAM and TAMRA fluorophores can be monitored simultaneously at 515 nm and 580 nm respectively (figure 4.12). Unlike the FRET melting experiments, this

allows the observation of folding in real time. Upon addition of silver nitrate, the spectrum shows a decrease in fluorescence of FAM while the fluorescence of TAMRA does not change (figure 4.12). This indicates a folding event, bringing the fluorophores into close proximity so that Förster resonance energy transfer can occur. The fact that the fluorescence of TAMRA does not change provides further support that this is a conformational change rather than a precipitation event as this would result in a decrease in fluorescence of both FAM and TAMRA.

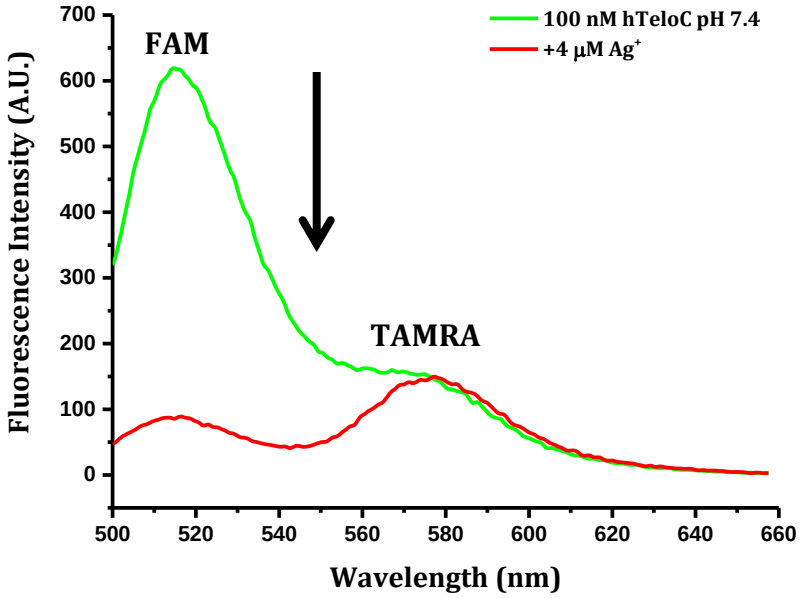



*Figure 4.12: Fluorescence spectrum of hTeloC labelled with FAM and TAMRA at pH 7.4 and in the presence of $AgNO_3$ showing folding. [DNA] = 100 nM, [$AgNO_3$] = 4 μM, buffer = 10 mM sodium cacodylate and 5 mM NaCl.*

In previous work carried out in the Waller group by Camille Huguin a FRET titration was carried out to show the folding of hTeloC with increasing concentrations of silver nitrate at pH 7.4. From this experiment a positive cooperativity ($n$ = 3.3) was determined and the concentration of silver (I) ions at which half the DNA was folded and half unfolded was also measured ($[Ag^+]_{50}$ = 1.9 μM).

The reversibility of folding using cysteine was also seen in FRET titrations. The telomeric sequence was folded at pH 7.4 by addition of 4 μM $AgNO_3$. The folded DNA was then titrated with increasing concentrations of L-cysteine. This caused an increase in the fluorescence of FAM at 515 nm indicating unfolding of the DNA (figure 4.13). As

can be seen, the fluorescence returned to starting levels after addition of 5 µM of L-cysteine suggesting that the DNA was completely unfolded.

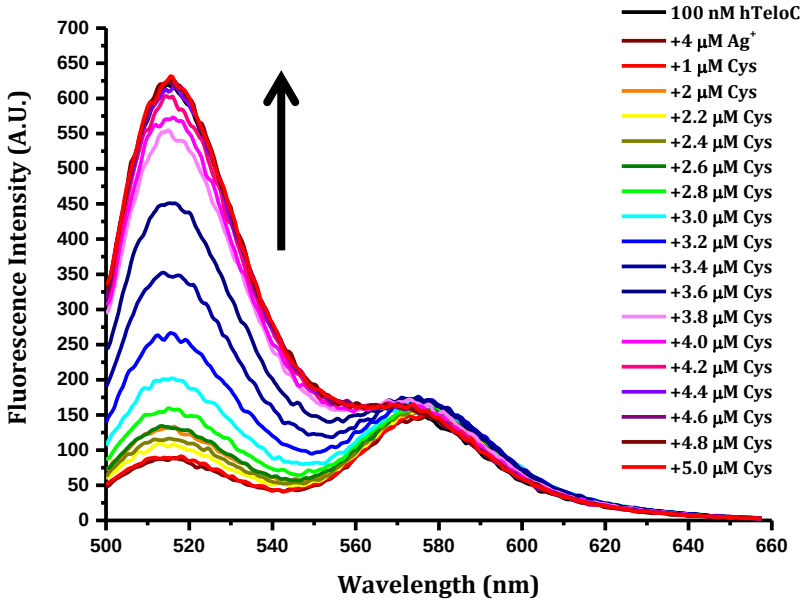



*Figure 4.13: FRET titration of hTeloC sequence folded with $AgNO_3$, with increasing concentration of L-cysteine. [DNA] = 100 nM, [$AgNO_3$] = 4 µM, buffer = 10 mM sodium cacodylate and 5 mM NaCl.*

The change in fluorescence can be used to give the FRET efficiency using equation 4.1. Where $FI_d$ is the fluorescence intensity of the donor fluorophore FAM and $FI_d^0$ is the fluorescence intensity of FAM in the absence of the acceptor fluorophore TAMRA which is taken as the fluorescence of FAM when the DNA is fully unfolded before any addition of $Ag^+$. This gives a decreasing curve, indicating unfolding (figure 4.14). The FRET efficiency was fitted with equation 4.3 to obtain the $[Cysteine]_{50}$ and the cooperativity of the interaction.

*Equation 4.1:* $$E_{FRET} = 1 - \left(\frac{FI_d}{FI_d^0}\right)$$

*Equation 4.3:* $$\theta = A\frac{[Cysteine]^n}{K^n + [Cysteine]^n}$$

Where $K$ is the ligand concentration at which half the DNA is unfolded ($[Cysteine]_{50}$) and $n$ is the Hill coefficient. $A$ is a constant ($A$ = start + (end-start) values). Using this method, $[Cysteine]_{50}$ = 3.5 µM was measured and a very high positive cooperativity of

*n* = 9 was calculated. This indicates that an excess of L-cysteine is needed to unfold the structure.

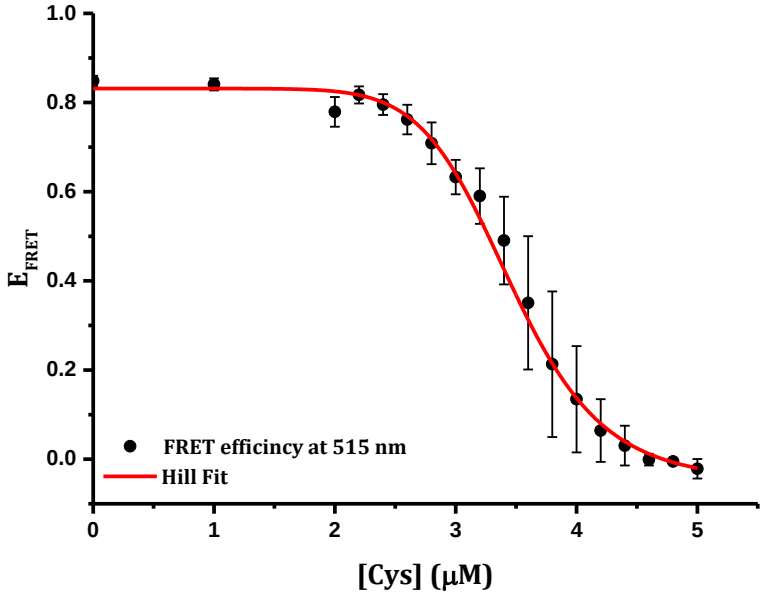



*Figure 4.14: FRET efficiency showing unfolding of silver stabilised i-motif with increasing concentration of L-cysteine. Error bars show the standard deviation across 3 repeats. [DNA] = 100 nM, [$AgNO_3$] = 4 μM, buffer = 10 mM sodium cacodylate and 5 mM NaCl. Figure adapted from Henry Day et. al.*[199] *by permission of The Royal Society of Chemistry.*

The repeat reversibility of folding which could enable switch functionality was also observed in the fluorescence at 515 nm with a slight decrease in fluorescence over time likely to be due to dilution (figure 4.15). The fluorescence of FAM was monitored at 515 nm as sequential additions of 4 μM of $AgNO_3$ and 4 μM of cysteine were added to the i-motif sequence. This resulted in a decrease in FAM fluorescence after addition of silver nitrate and an increase after addition of cysteine indicating folding and unfolding respectively.

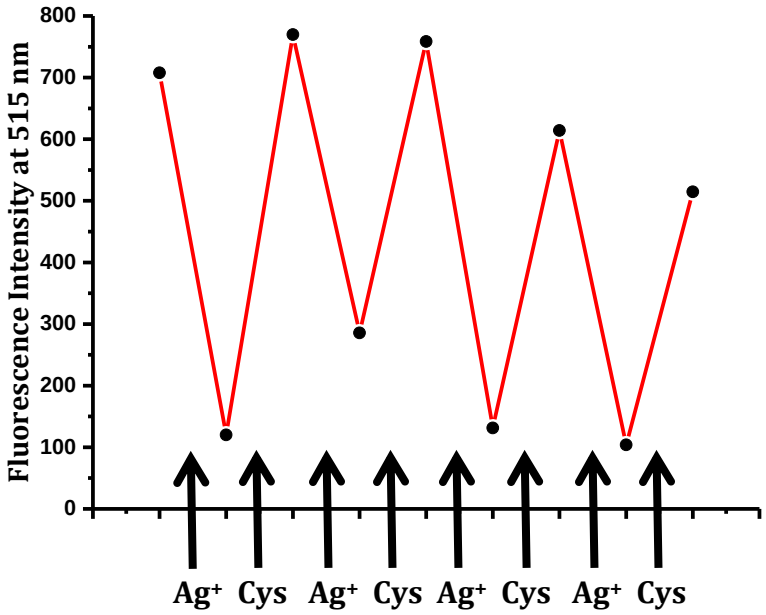



*Figure 4.15: Change in fluorescence intensity at 515 nm with repeated additions of $AgNO_3$ and L-cysteine. [DNA] = 100 nM, buffer = 10 mM sodium cacodylate and 5 mM NaCl. Figure adapted from Henry Day et. al.*[199] *by permission of The Royal Society of Chemistry.*

The FRET titration carried out previously in the Waller group indicates that the binding of silver (I) ions to the DNA is cooperative ($n$ = 3.3) giving an $[Ag^+]_{50}$ = 1.9 μM. These results show that the addition of cysteine does unfold the DNA, giving further support for the observations made in circular dichroism experiments, and is also highly cooperative ($n$ = 9). In this experiment, the concentration of cysteine at which half the silver structure was unfolded ($[Cysteine]_{50}$) was determined to be 3.5 μM. This would suggest that slightly more cysteine than silver is required to unfold the structure. This may be because the first additions of cysteine will bind to the excess $Ag^+$ ions in solution before binding to the $Ag^+$ ions that are bound to cytosine. As described in section 4.2, these FRET experiments can only demonstrate that folding is occurring but cannot give information about the type of structure formed. The circular dichroism spectra in section 4.3 suggest that an i-motif-like structure is forming but also shows some differences in the spectra compared to a normal acid stabilised i-motif therefore more evidence is needed to support the hypothesis that an i-motif-like structure is forming.

### 4.2.5 UV Difference Spectroscopy

To further investigate the type of structure formed, the change in UV absorbance of the DNA was measured with increasing concentrations of silver (I) ions. Whether folded or unfolded, DNA gives an absorbance peak at 260 nm. However, due to the difference in stacking interactions, the absorbance spectrum of folded and unfolded DNA is different. Whilst the absorbance at 260 nm changes very little, when folded, the absorbance of the i-motif is decreased between 220 and 260 nm and increased at 295 nm. This can be observed in the spectra of the human telomeric i-motif forming sequence at low temperature (20°C) and at high temperature (90°C) shown in figure 4.16.

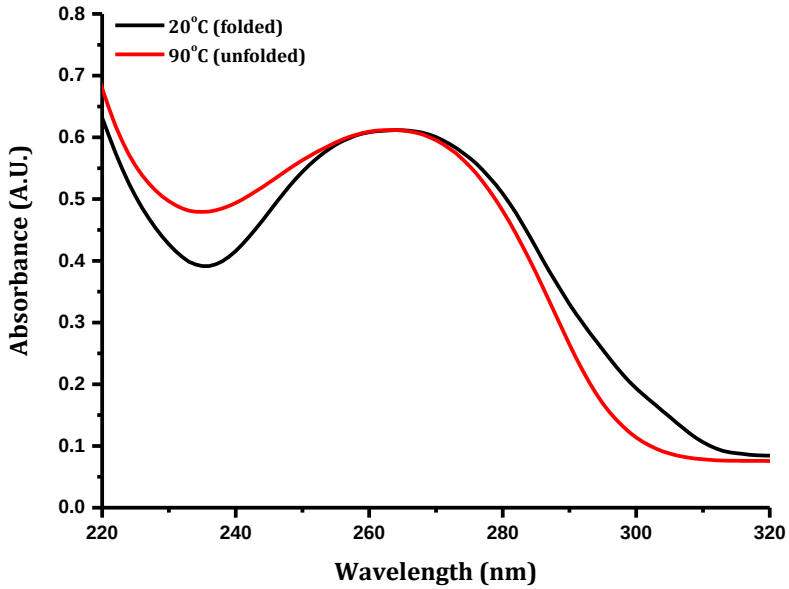



*Figure 4.16: The UV absorbance spectra of the human telomeric i-motif sequence at 20°C (folded) and 90°C (unfolded) at pH 5.5. [DNA] = 2.5 μM, buffer = 10 mM sodium cacodylate and 5 mM NaCl.*

The human telomeric sequence in its unfolded form at pH 7.4 was titrated with silver nitrate. This showed that with increasing concentrations of silver cations, the absorbance of the i-motif forming sequence increased at 260 nm with an isosbestic point at 255 nm (figure 4.17). This increase in absorbance does indicate a change in structure while the isosbestic point indicates that there are only two species involved in the absorbance, for example folded and unfolded DNA. The isosbestic point was maintained up until 6 equivalents of $Ag^+$ were added, after which point the absorbance continued to increase with increasing silver concentration but the isosbestic point is

lost. This indicates that any structural change of the DNA is likely to be completed after addition of 6 equivalents of silver and that the further increases in absorbance seen with greater concentrations are due to the silver nitrate or scattering effects.

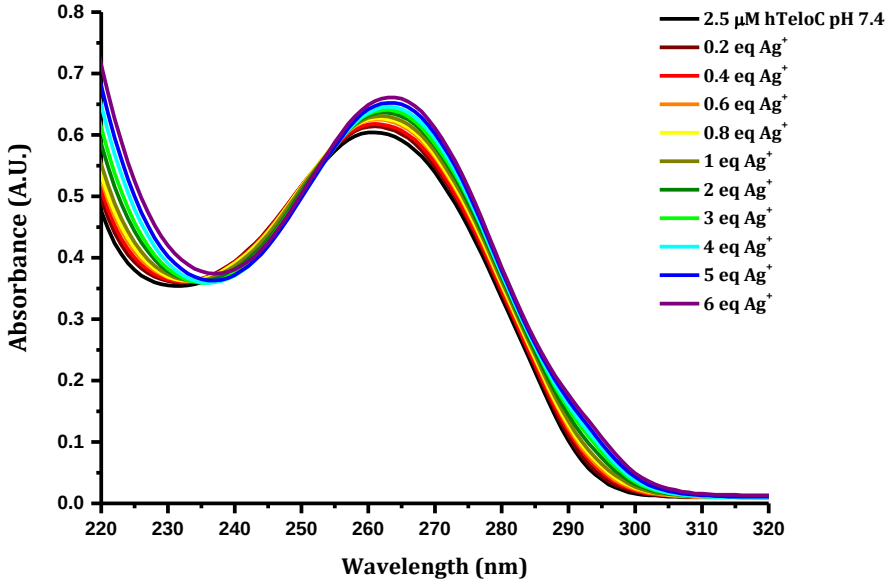



*Figure 4.17: The UV absorbance spectrum of human telomeric i-motif DNA with up to 6 equivalents of $AgNO_3$ at pH 7.4. [DNA] = 2.5 µM, Buffer = 10 mM sodium cacodylate and 5 mM NaCl.*

As mentioned, the normal folding and unfolding of the i-motif does not show an increase in absorbance at 260 nm (figure 4.16) so on first appearance, the change in UV absorbance with silver may not be indicative of a folding event. However, after an analogous titration of silver nitrate alone in buffer is subtracted from the data, the resultant spectrum (figure 4.18), does show very little change at 260 nm as well as the decrease at 220 – 260 nm and the increase at 295 nm that is expected for a folding event. The subtraction is necessary to discount the absorbance effect of silver nitrate.

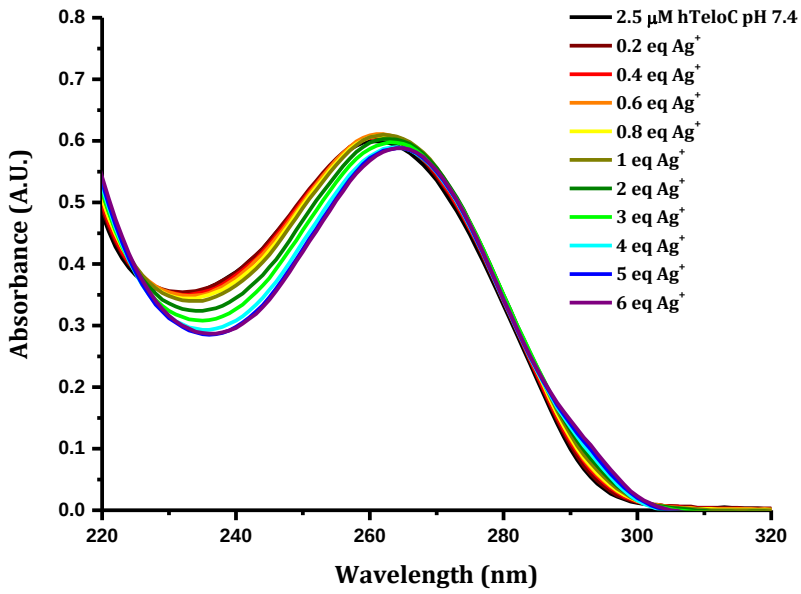



*Figure 4.18: The UV absorbance spectrum of human telomeric i-motif DNA with up to 6 equivalents of $AgNO_3$ at pH 7.4 after subtraction of an analogous titration of $AgNO_3$ in buffer. [DNA] = 2.5 μM, Buffer = 10 mM sodium cacodylate and 5 mM NaCl.*

To look at the type of folded structure forming, the UV difference spectrum was measured.[93] As described in chapter 1, in 2005 Jean-Louis Mergny and co-workers published an investigation showing that taking the difference between the UV spectra of folded and unfolded DNA gives a spectrum with a specific signal dependent on the type of DNA structure present in the solution. In Mergny *et. al.'s* work they used thermal melting in order to measure the spectra of both folded (low temperature) and unfolded (high temperature) structures.[93] Here the DNA is being folded by the addition of silver cations, therefore the technique was adapted. To generate the difference spectrum, first a titration of $AgNO_3$ alone in buffer was subtracted from the titration with DNA to discount any effects not due to binding. Then the spectrum of folded DNA (6 equivalents of $Ag^+$) was subtracted from the spectrum of unfolded DNA (0 equivalents of $Ag^+$). This gave a "silver difference spectrum" (figure 4.19a) which can be used in a similar way to circular dichroism to indicate the type of DNA structure forming with addition of silver cations.

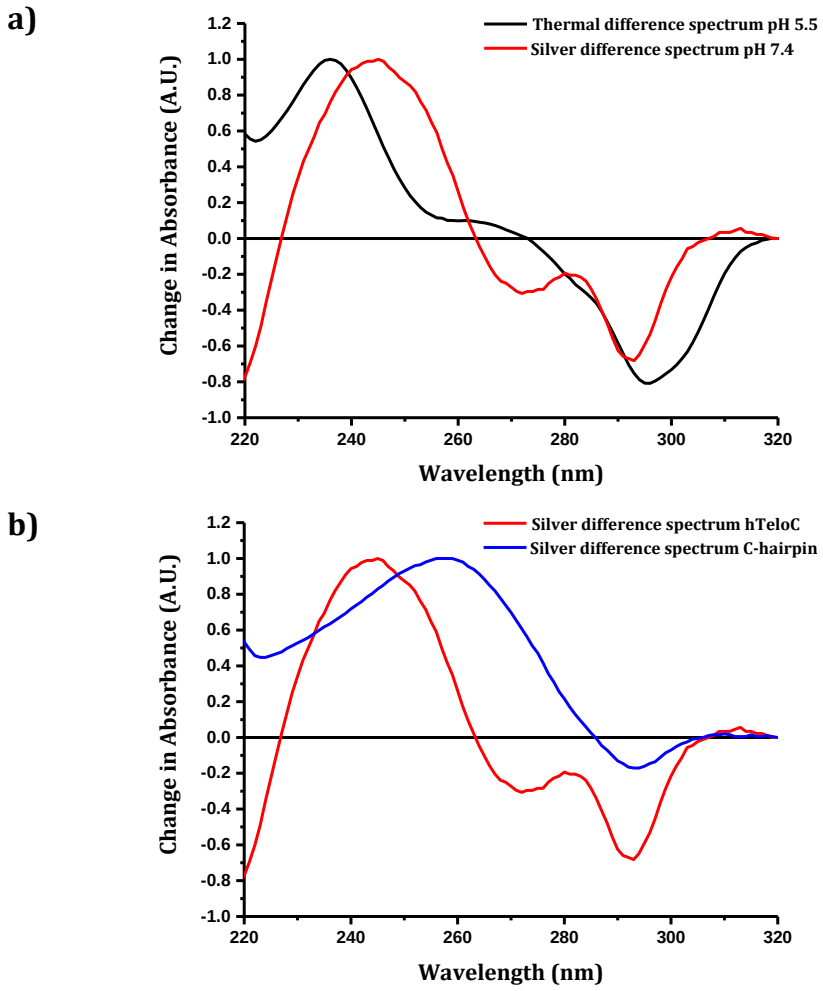



*Figure 4.19: a) "Silver difference" at pH 7.4 and thermal difference spectra at pH 5.5 of human telomeric i-motif. b) "Silver difference" spectra of human telomeric i-motif and a hairpin forming sequence at pH 7.4. [DNA] = 2.5 μM, [$AgNO_3$] = 15 μM, buffer = 10 mM sodium cacodylate and 5 mM NaCl. Figure adapted from Henry Day et. al.*[199] *by permission of The Royal Society of Chemistry.*

The silver difference spectrum has a positive peak at 245 nm and a negative peak at 293 nm and is similar to the thermal difference spectrum of the acid stabilised i-motif which also has peaks in this region as illustrated by Mergny *et al*. for a variety of different i-motif sequences.[93] This suggests that formation of an i-motif type folded structure is occurring. However, with the human telomeric sequence it is possible that the silver cations are only stabilising a hairpin folded structure rather than an i-motif like structure. To investigate this, a control UV difference experiment was carried out using the sequence 5'-d(CTCTCTTCTCTTCATTTTTCAACACAACACAC)-3'.[190] This particular sequence was chosen because it is the same as the one used by Ono and

colleagues, which is known to form a hairpin in the presence of $Ag^+$ ions and, the sequence does not have the necessary tracts of cytosine to enable i-motif formation. The control gave a significantly different spectrum (figure 4.19b) with a peak at 260 nm and a much smaller negative peak at 295 nm. This UV difference resembles those of Hoogsteen duplexes in Mergny's study,[93] however, Mergny did not look at the thermal difference spectra of hairpin DNA and, as already discussed, this particular sequence is known to form a hairpin in the presence of silver.[190] The difference between these spectra indicates that a structure other than a hairpin is formed by addition of silver (I) ions to the sequence.

These results have again shown that the human telomeric i-motif forming sequence forms a folded structure in the presence of silver (I) ions at neutral pH. This supports the results of experiments in sections 4.2 to 4.5. The UV difference spectra for silver is an application of the thermal difference spectroscopy method outlined by Mergny and co-workers. These spectra demonstrate that the folded structure is not that of a simple hairpin and shows greater similarity to the thermal difference spectrum of the i-motif. This provides greater evidence in support of the CD spectra, suggesting that the folded DNA forms an i-motif-like structure in the presence of silver.

### 4.2.6 Summary of effect of silver (I) ions on i-motif forming DNA.

In summary, a number of investigations by Akira Ono have shown that $Ag^+$ cations stabilise cytosine-cytosine base pairs by selectively binding to the N3 of cytosine. Here the hypothesis that this property could enable silver cations to stabilise i-motif formation at neutral pH has been tested with a range of biophysical experiments. The addition of silver (I) ions to the human telomeric i-motif forming sequence was shown to induce folding of the DNA in FRET melting experiments and FRET titration experiments with earlier work in the Waller group demonstrating a cooperative binding ($n$ = 3.3) with an $[Ag^+]_{50}$ = 1.9 μM. Folding was also observed in circular dichroism experiments and UV titrations with the CD and UV difference spectra providing evidence of formation of an i-motif-like structure. UV difference spectroscopy was also used to rule out the formation of a simpler hairpin structure. The differences between the CD spectrum of the folded structure and that of a normal acid stabilised i-motif were not unexpected. This is because of the effect of

incorporating larger silver ions into the structure that will distort the conformation of glycosidic bonds. This is supported by the stoichiometry of 4:1 which was measured by CD indicating that less silver (I) ions are needed than the 6 cytosine pairs expected for a normal i-motif. This suggests that either only 4 cytosine-Ag+-cytosine pairs are required for the structure to form, or that the silver cations stabilise the structure in a different way, perhaps by resting between cytosine base pairs or by forming a nanocluster which the DNA can wrap around. Finally the reversibility of this structure formation was demonstrated using cysteine to chelate the silver (I) ions. This process was also shown to be cooperative ($n$ = 9) with slightly more cysteine required to unfold the structure ($[Cysteine]_{50}$ = 3.5 μM).

The ability to reversibly fold and unfold i-motif forming sequences at neutral pH with addition of silver cations and cysteine could have a number of applications. The reversible folding of the i-motif due to changes in pH has been used in a range of nanotechnology applications and these results offer an alternative switch mechanism.[203] As well as applications focusing on the folding of the DNA, these results also offer the potential for i-motif sequences to be used in cysteine sensing applications such as those previously published based on double stranded DNA.[193,194] It is still not clear exactly where and how the silver ions are located within the structure and how the overall conformation of the DNA is different to that of an acid stabilised i-motif. Further work is needed using NMR or X-ray crystallography to get a clearer idea of what the silver induced structure looks like.

## 4.3 Screening other cations

As discussed in section 4.1.1, only a small number of mostly group 1 and 2 cations have been studied for their effect on the stability of i-motif DNA. After showing that silver cations are able to stabilise the formation of an i-motif like structure at neutral pH, the next stage in the investigation was to look at the effect of other cations on i-motif DNA. For this purpose, a selection of 26 different cations were screened using a FRET melting assay against the telomeric i-motif at pH 7.4 and pH 5.5 (table 4.1). The cations were from a variety of groups including 1, 2, and 3, the transition metals and the lanthanides and hence have a range of different charges and ionic radii. The selection also includes

cations of biological importance from the Irving-Williams series including $Mn^{2+}$, $Co^{2+}$, $Ni^{2+}$, $Cu^{2+}$ and $Zn^{2+}$. This series describes the stability of complexes of the first row transition metals and is important in a biological context as it reflects the different binding affinities of proteins for these cations.[204,205] Unfortunately $Fe^{2+}$, which is also part of the series, could not be tested as the available salt, iron (II) chloride, was not soluble enough. However, iron (III) chloride was soluble enough to be tested.

| Salt | Cation | Group |
|---|---|---|
| $AlCl_3$ | $Al^{3+}$ | 3 |
| $BaCl_2$ | $Ba^{2+}$ | 2 |
| $CaCl_2$ | $Ca^{2+}$ | 2 |
| $CdCl_2$ | $Cd^{2+}$ | TM |
| CsCl | $Cs^{+}$ | 1 |
| $CoCl_2$ | $Co^{2+}$ | TM |
| $CuCl_2$ | $Cu^{2+}$ | TM |
| $Ga(NO_3)_3$ | $Ga^{3+}$ | 3 |
| $HoCl_3$ | $Ho^{3+}$ | L |
| $In(NO_3)_3$ | $In^{3+}$ | 3 |
| $FeCl_3$ | $Fe^{3+}$ | TM |
| $PbAc_2$ | $Pb^{2+}$ | 4 |
| $MgCl_2$ | $Mg^{2+}$ | 2 |
| $MnCl_2$ | $Mn^{2+}$ | TM |
| $HgAc_2$ | $Hg^{2+}$ | TM |
| $NiCl_2$ | $Ni^{2+}$ | TM |
| KCl | $K^{+}$ | 1 |
| RbCl | $Rb^{+}$ | 1 |
| $RuCl_3$ | $Ru^{3+}$ | TM |
| $SmCl_3$ | $Sm^{3+}$ | L |
| $SrCl_2$ | $Sr^{2+}$ | 2 |
| TlAc | $Tl^{+}$ | 3 |
| $TbCl_3$ | $Tb^{3+}$ | L |
| $YbCl_3$ | $Yb^{3+}$ | L |
| $YCl_3$ | $Y^{3+}$ | TM |
| $ZnAc_2$ | $Zn^{2+}$ | TM |
| $NaAc_2$ | Acetate | Control |

*Table 4.1: Table of salts tested in the FRET melting screen. TM = Transition Metal, L = Lanthanide.*

The cations were tested at concentrations of 100 mM, 1 mM, 100 μM and 10 μM in 10 mM sodium cacodylate, 5 mM sodium chloride buffer the same as has been used previously in section 4.2. Most of the salts used were chlorides or nitrates with the exception of $Pb^{2+}$, $Hg^{2+}$, and $Zn^{2+}$, for which the acetate salt was used. Nitrate has already been shown not to effect i-motif formation earlier in this chapter (figure 4.5 & 4.6) and chloride is used throughout all our buffer conditions with no apparent effect on i-motif formation. However, the effect of the acetate anion has not been examined, therefore sodium acetate was also tested as a control. The cations were screened at pH 7.4 in order to look for induced folding of the DNA at neutral pH, similar to that seen for silver. They were also screened at pH 5.5 as the i-motif is already formed under these conditions and this enables any stabilising or destabilising effects on the i-motif to be observed.

After the initial screen, several hits were identified which seemed to cause folding at neutral pH. These were: $AlCl_3$, $CuCl_2$, $Ga(NO_3)_3$, $HoCl_3$, $SmCl_3$, $TbCl_3$, $YbCl_3$, $YCl_3$, $HgAc_2$ and $ZnAc_2$ (Table 4.2). These cations caused folding at concentrations between 0 and 100 μM and most produced base level fluorescence signals, where the DNA does not melt, at 100 μM which may indicate strong stabilisation or precipitation. The cations all have a charge of +2 or +3, enabling them to effectively counteract the negative charge of the phosphate backbone. Of particular interest are copper, mercury and zinc, which have the same linear coordination geometry as silver, suggesting that they may be able to bind to cytosine base pairs. In fact, mercury has already been shown to stabilise T-T and T-C base pairs as described in section 4.1.2.[191]

| Cation | Concentration (μM) | $T_m$ (°C) |
|---|---|---|
| $Al^{3+}$ | 100 | 65.8 |
| $Cu^{2+}$ | 10 | 57.2 |
| $Ga^{3+}$ | 100 | 72.5 |
| $Ho^{3+}$ | 10 | 81.8 |
| $Sm^{3+}$ | 10 | 63.8 |
| $Tb^{3+}$ | 10 | 78.8 |
| $Yb^{3+}$ | 10 | >95 |
| $Y^{3+}$ | 10 | 92.5 |
| $Hg^{2+}$ | 10 | 61.5 |
| $Zn^{2+}$ | 100 | 76.5 |

*Table 4.2: Melting temperatures of folded DNA as a result of different cations. [DNA] = 200 nM, buffer = 10 mM sodium cacodylate 5 mM sodium chloride. The error in the $T_m$ values is ± 0.3 °C.*

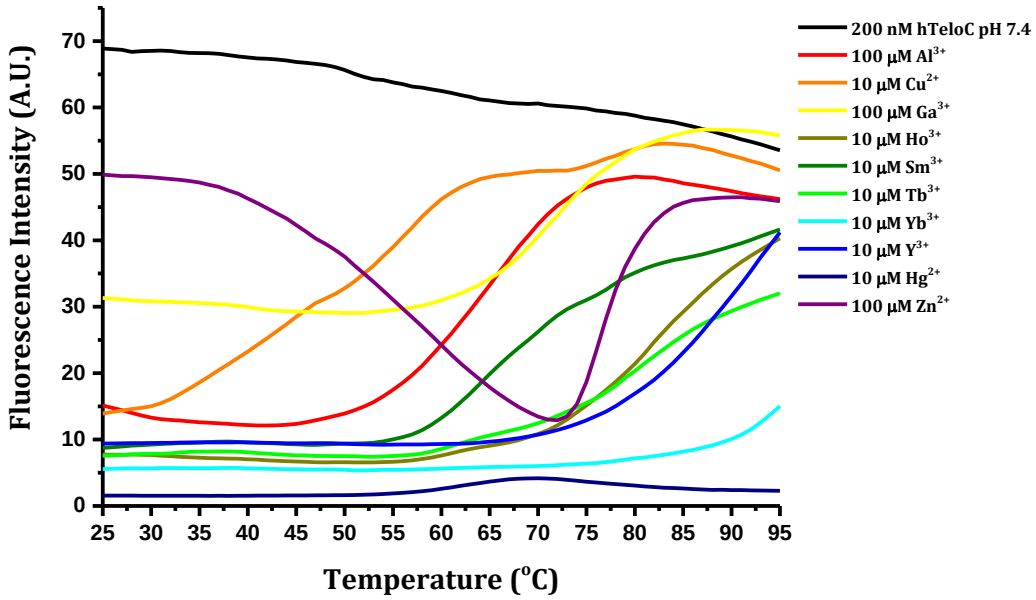



*Figure 4.20: FRET melting curves of human telomeric i-motif forming sequence at pH 7.4 in the presence of cations which caused folding of the DNA. [DNA] = 200 nM, buffer = 10 mM sodium cacodylate and 5 mM NaCl.*

Figure 4.20 shows the melting curves induced by the addition of the cations from table 4.2 to the hTeloC sequence at pH 7.4. Some cations seemed to have more of an effect, causing folding at 10 μM while others only caused folding at 100 μM. As the DNA is normally unfolded at pH 7.4, it is not possible to measure a melting temperature in the

absence of the cations, however at acidic pH, the i-motif normally has a melting temperature of 46°C. These cations have significantly higher melting temperatures suggesting they form very stable structures. Full results from the cation screen can be found in the appendix (section A.3).

The effect of the cations on the DNA was tested by circular dichroism in order to see what type of structure may be forming or if precipitation was occurring. Unfortunately, due to the higher concentration of DNA required for CD compared to FRET experiments (10 μM vs. 200 nM) testing the cations at the same concentrations as FRET (80 - 100 μM) resulted in very little change in the CD spectrum, while testing at the same number of equivalents required much higher concentrations of cation (4 – 5 mM). At these concentrations, precipitation was apparent in the presence of $Ho^{3+}$, $Sm^{3+}$, $Tb^{3+}$ and $Yb^{3+}$ but the addition of $Al^{3+}$, $Ga^{3+}$ and $Zn^{2+}$ did seem to induce i-motif formation. However, after testing the pH of the cations in the 10 mM sodium cacodylate buffer at this high concentration, the pH was not maintained when greater than 100 μM of ions were added. Therefore the observed change to i-motif structure is most likely a result of decreasing pH rather than stabilisation by the cations themselves and these were not investigated further.

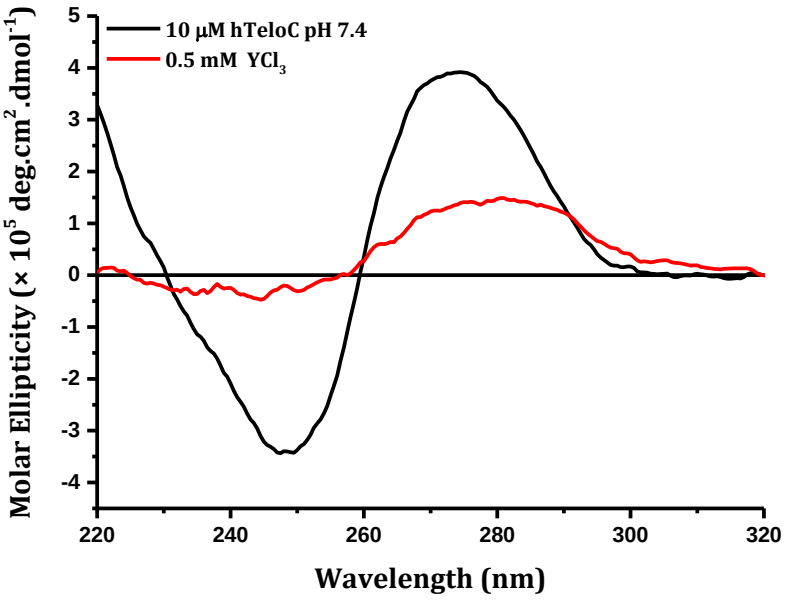



*Figure 4.21: CD spectrum of human telomeric i-motif DNA at pH 7.4 with 0.5 mM $YCl_3$. [DNA] = 10 μM, buffer = 10 mM sodium cacodylate and 5 mM NaCl.*

More promisingly, $Y^{3+}$ showed a small shift towards 280 nm in the CD spectrum at a concentration of 0.5 mM (figure 4.21) but the spectrum still indicated some precipitation. This concentration equated to 50 equivalents, the same number of equivalents at which a stabilisation effect was seen in FRET melting experiments, yet it was still below the concentration where changes in pH occurred. $Cu^{2+}$ showed a small shift at 80 μM (figure 4.22). This is a comparable concentration to that used in FRET melting experiments where an effect was seen (100 μM) but represents fewer equivalents due to the higher DNA concentration in CD. $YCl_3$ maintained a pH of 7 at 1 mM and $CuCl_2$ maintained a pH of 7.2 at 100 μM compared to pH 7.3 for the buffer initially, so these cations were taken forward for further study.

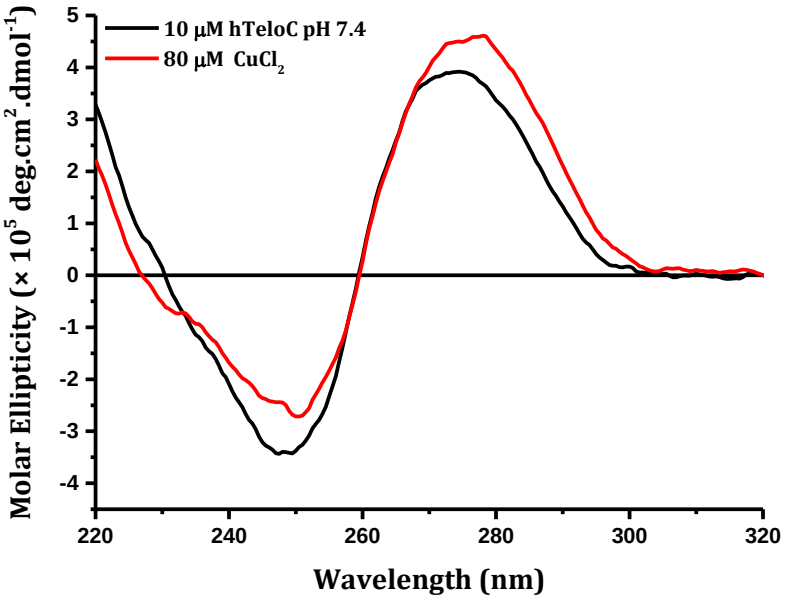



*Figure 4.22: CD spectrum of human telomeric i-motif DNA at pH 7.4 with 80 μM $CuCl_2$. [DNA] = 10 μM, buffer = 10 mM sodium cacodylate and 5 mM NaCl.*

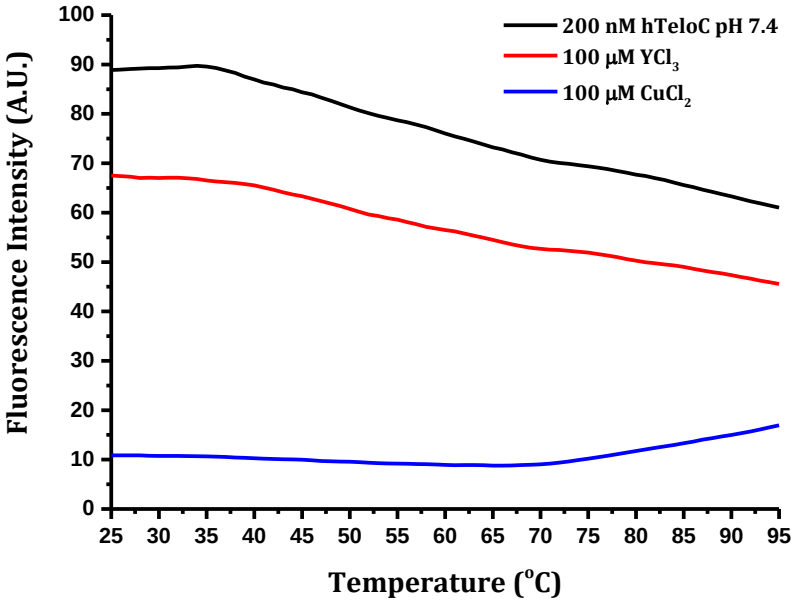



*Figure 4.23: FRET melting curves for human telomeric i-motif at pH 7.4 with 100 µM $YCl_3$ and $CuCl_2$ at pH 7.4. [DNA] = 200 nM, buffer = 50 mM sodium cacodylate.*

In order to enable testing at higher cation concentrations without changing the pH, the concentration of sodium cacodylate in the buffer was increased from 10 to 50 mM. In these buffer conditions, the pH still showed a decrease but could be maintained at neutral or above at up to 4 mM $CuCl_2$ (pH = 6.97) and 7 mM $YCl_3$ (pH = 7.02). The drawback of this change is that the higher concentration of sodium ions could mask the effect of a lower concentration of another cation. To ensure that the cations still had a stabilising effect on the i-motif sequence in the new buffer concentrations, they were re-tested by FRET melting at 100 µM (figure 4.23). Here $Y^{3+}$ no longer showed any folding effect on the DNA, whilst $Cu^{2+}$ still showed a decrease in fluorescence indicating folding.

In summary, after carrying out a FRET melting screen of 26 cations, 10 of them appeared to cause folding of the DNA at pH 7.4 ($Al^{3+}$, $Cu^{2+}$, $Ga^{3+}$, $Ho^{3+}$, $Sm^{3+}$, $Tb^{3+}$, $Yb^{3+}$, $Y^{3+}$, $Hg^{2+}$ and $Zn^{2+}$). Further testing of these hits by CD showed that several caused precipitation ($Ho^{3+}$, $Sm^{3+}$, $Tb^{3+}$ and $Yb^{3+}$), while others ($Al^{3+}$, $Ga^{3+}$ and $Zn^{2+}$) did indicate i-motif formation but this was due to a decrease in pH. Unfortunately $Hg^{2+}$ could not be tested due to the insolubility of the available salt leaving $Y^{3+}$ and $Cu^{2+}$ as the only cations which did show small shifts in the CD at concentrations that did not significantly alter the pH. In order to carry out further experiments at higher cation concentrations, the concentration of the buffer was increased to 50 mM in order to

maintain the pH above neutral. However, repeat FRET melting experiments on $Y^{3+}$ and $Cu^{2+}$ in the new buffer showed that only copper (II) ions still caused an apparent folding event. Therefore, only copper (II) ions have been taken forward for further investigation which is discussed in section 4.4.

## 4.4 The Effect of Copper (II) Ions on i-Motif Forming Sequences.

Having screened a range of different cations for stabilisation of the telomeric i-motif sequence at neutral pH, $Cu^{2+}$ was the only cation from the screen that still appeared to stabilise a folded structure without reducing the pH below neutral. To investigate this further, preliminary experiments have been carried out using FRET melting, UV difference and circular dichroism spectroscopy. FRET melting alone can indicate folding, give information about the stability of the structure and the number of structures in equilibrium, but these other experiments are needed in order to give more detailed information about the type of structure that is being formed.

### 4.4.1 FRET melting

In section 4.3, the cations were screened at concentrations of 100 mM, 1 mM, 100 μM and 10 μM in 10 mM sodium cacodylate buffer with 5 mM NaCl, showing stabilisation between 0 and 100 μM. This result was confirmed in the new 50 mM sodium cacodylate buffer by repeating the experiment with 100 μM of $Cu^{2+}$ and at this concentration the DNA did not melt within the temperature range of the experiment. Therefore a range of different concentrations of copper (II) cations were tested between 0 and 100 μM at pH 7.4 (figure 4.24) and pH 5.5 (figure 4.25) to examine whether the stabilisation shows a dose response. At pH 7.4, the data indicates the onset of folding with a decrease in fluorescence at a $Cu^{2+}$ concentration of 10 μM. At 40 μM the DNA appears to be folded at 25°C and there is an apparent melting transition at $T_m$ = 68°C which suggests the formation of a stable folded structure. At higher concentrations the fluorescence is quenched throughout the temperature range of the experiment. This could be due to high stabilisation or aggregation of the DNA. Conversely at pH 5.5, there appears to be

a slight decrease in melting temperature from 46.2°C to 45.5°C at [$CuCl_2$] = 20 μM, however, as the concentration of $CuCl_2$ is increased there is still a low fluorescence intensity. This could be as a result of stabilising an alternative structure or could again be due to precipitation.

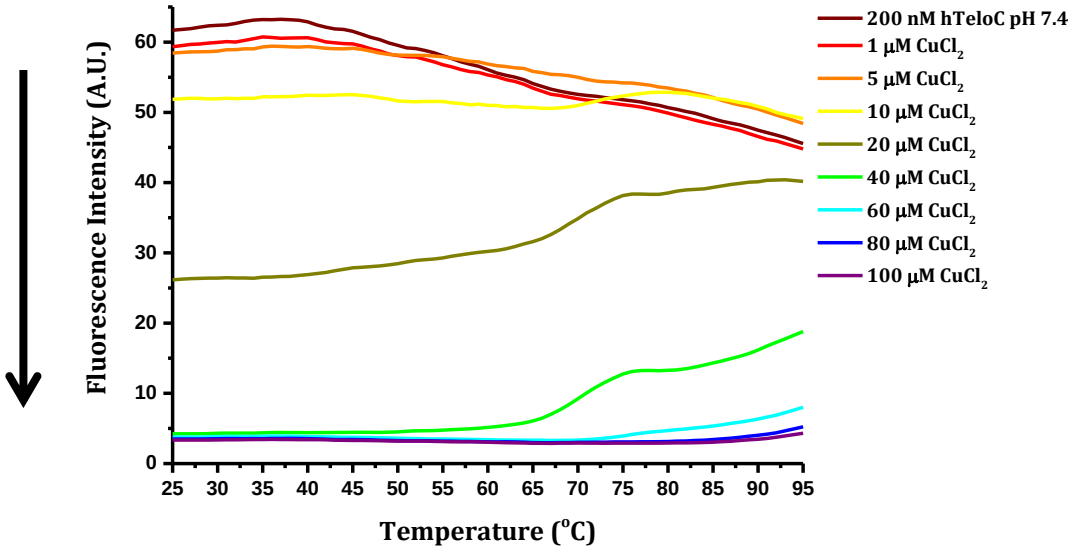



*Figure 4.24: Example FRET melting data for the human telomeric i-motif sequence at pH 7.4 with increasing concentrations of $CuCl_2$. [DNA] = 200 nM, buffer = 50 mM sodium cacodylate.*

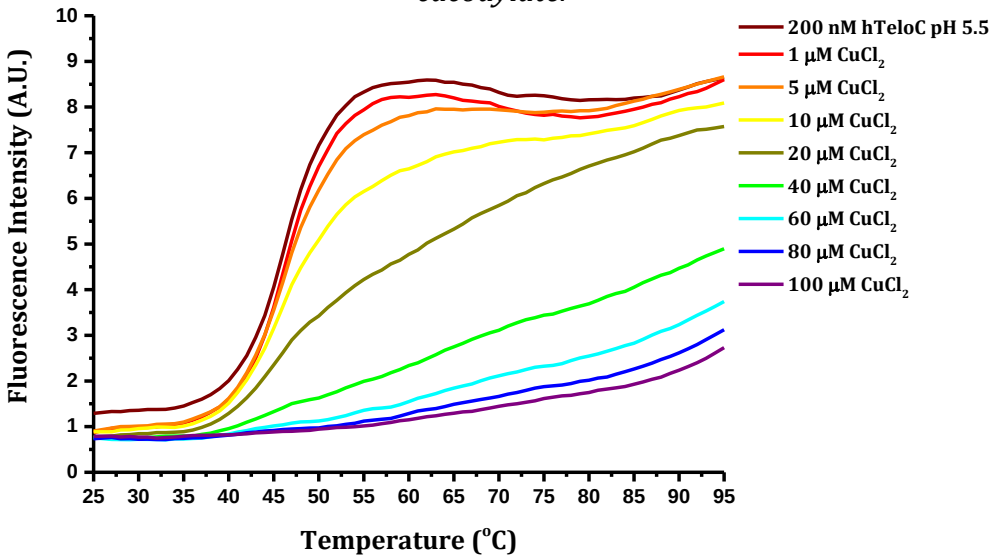



*Figure 4.25: Example FRET melting data for the human telomeric i-motif sequence at pH 5.5 with increasing concentrations of $CuCl_2$. [DNA] = 200 nM, buffer = 50 mM sodium cacodylate.*

As described in section 4.2.4, further indication of folding can be seen in FRET titration experiments. As the DNA folds, the two fluorophores come into close proximity and the

fluorescence of FAM is quenched so it is expected that there will be a decrease in the fluorescence signal for FAM at 515 nm. However the fluorescence of TAMRA at 580 nm should not decrease. However, in figure 4.26, it can be seen that the signals of both FAM and TAMRA decrease. This would suggest a precipitation effect.

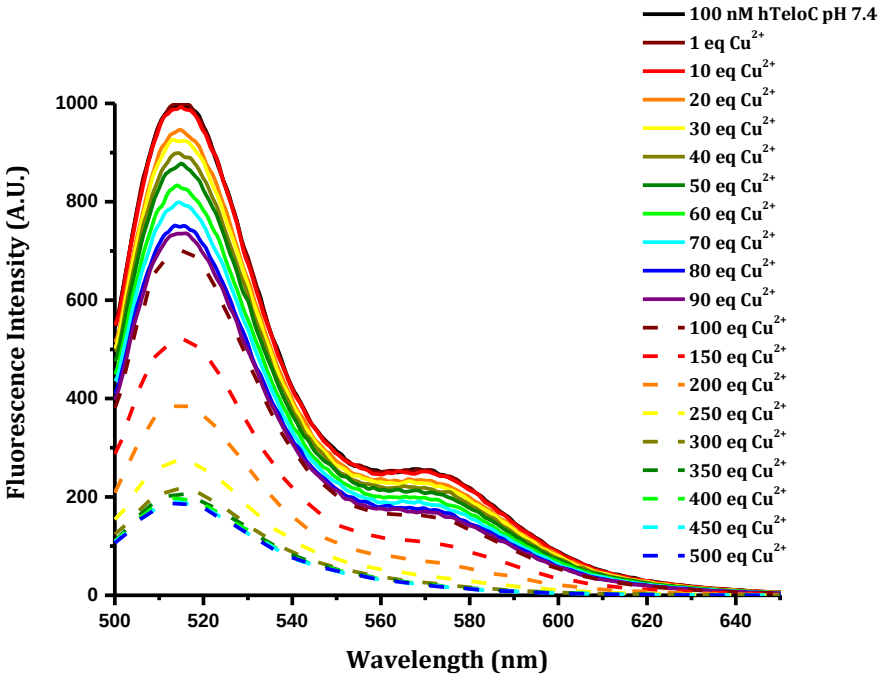



*Figure 4.26: FRET titration of human telomeric i-motif sequence with increasing concentration of $CuCl_2$ at pH 7.4. [DNA] = 100 nM, buffer = 50 mM sodium cacodylate.*

The FRET melting results indicate that at pH 7.4, the addition of increasing concentrations of copper (II) ions leads to a decrease in the fluorescence signal suggesting the formation of a folded structure. At a $Cu^{2+}$ concentration of 40 μM the folded structure shows a melting transition with a melting temperature ($T_m$) of 68°C which is much higher than an acid stabilised i-motif ($T_m$ = 46°C) implying a more stable structure. A FRET titration at pH 7.4 also indicates folding of the DNA initially but, at higher concentrations of $Cu^{2+}$ the fluorescence of both FAM and TAMRA decrease which indicates that precipitation is occurring. At pH 5.5, where the i-motif is already folded, the FRET melting experiments suggest that increasing concentrations of copper (II) ions destabilises the i-motif structure giving lower temperature melting transitions but the fluorescence remains low after this transition. There is not an obvious 2nd melting transition so it is not clear whether this effect is due to formation of another more stable structure as was seen with silver (I) ions. It is clear from these results that the copper (II) ions do interact with the i-motif structure and that they do induce

folding at neutral pH but further experiments such as UV difference and CD are required to give information about the type of structures forming.

### 4.4.2 UV Difference

Following the same procedure as described to show folding in the presence of silver cations (section 4.4), UV difference spectra were produced from UV titrations of different i-motif forming sequences in the presence of copper (II) ions. UV titrations were carried out with the human telomeric i-motif sequence as well as promoter region i-motif sequences from the c-Myc, HIF-1α and PDGF-A genes. These i-motif sequences are interesting because they are quite different to the telomeric i-motif. For example, the c-Myc i-motif (5'-CCCCACCTTCCCCACCCTCCCCACCCTCCCC-3') has longer loops (section 1.3.2), whereas the i-motifs formed from the HIF-1α (5'-CGCGCTCCCGCCCCCTCTCCCCTCCCCGCGC-3') and PDGF-A (5'-CCGCGCCCCTCCCCCGCCCCCGCCCCCGCCCCCCCCCCCCC-3') sequences have longer cytosine tracts, as described in section 1.3.2.[41,62] As described in section 4.2.5, the UV difference spectra for each sequence (figure 4.27) was generated by subtracting the titration of just $Cu^{2+}$ in buffer from the titration in DNA then subtracting the spectrum at the plateau (folded) from the starting spectrum (unfolded).

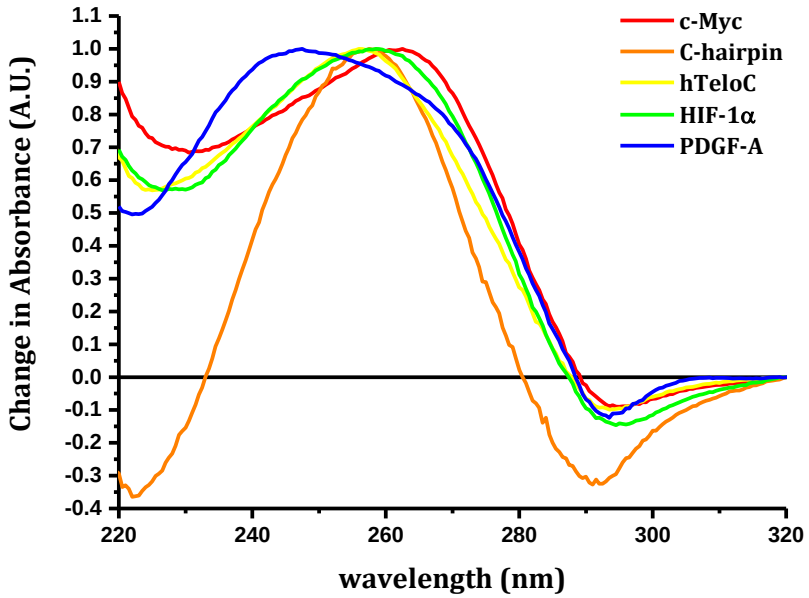



*Figure 4.27: Comparison of copper difference spectra of human telomeric, c-Myc, HIF-1α and PDGF-A i-motif sequences and the C-hairpin sequence at pH 7.4. [DNA] = 2.5 μM, [$CuCl_2$] = 250 μM, buffer = 50 mM sodium cacodylate.*

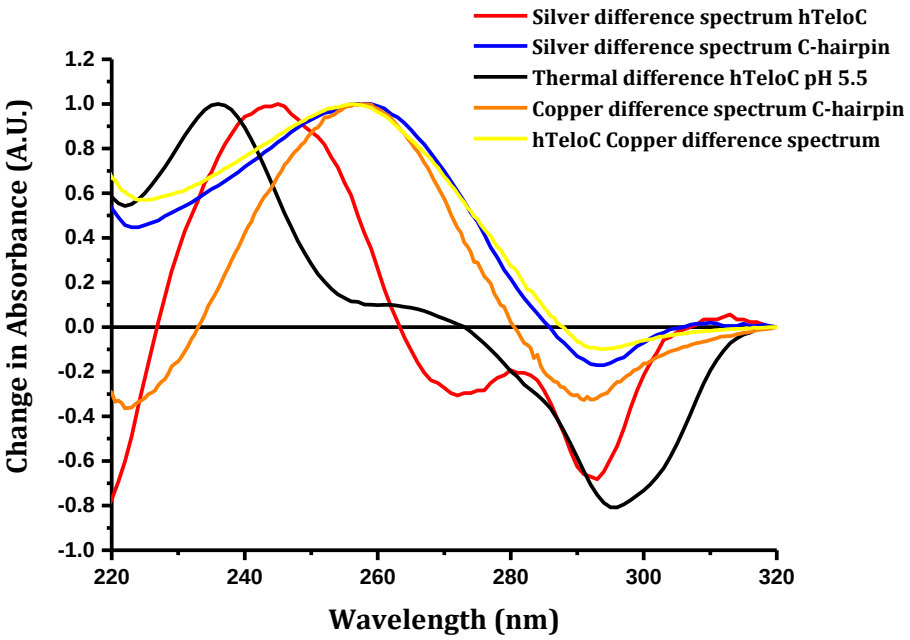



*Figure 4.28: "Silver difference" and "copper difference" at pH 7.4 and thermal difference spectra at pH 5.5 of human telomeric i-motif and "silver difference" and "copper difference" of C-hairpin at pH 7.4. [DNA] = 2.5 μM, [$AgNO_3$] = 15 μM, buffer = 10 mM sodium cacodylate and 5 mM NaCl, [$CuCl_2$] = 250 μM, buffer = 50 mM sodium cacodylate.*

These "copper (II) ion difference spectra" show a great deal of similarity suggesting that a similar structure is being formed from each sequence. The spectra are significantly different from the thermal difference and "silver difference" spectrum of the i-motif shown in section 4.4 and closely resemble the "silver difference" spectrum of the C-hairpin sequence shown in figure 4.28. As a further comparison, the "copper difference" of the C-hairpin forming sequence was measured which again showed significant similarity, sharing a peak at 260 nm. This suggests that rather than stabilising the i-motif, $Cu^{2+}$ may in fact stabilise formation of a hairpin structure instead. To investigate this further, circular dichroism was used to give a more detailed view of the possible structural changes.

### 4.4.3 Circular Dichroism

Circular dichroism was used to look at the change in structure caused by addition of copper (II) cations. The copper (II) cations were titrated into a solution of 10 μM hTeloC both at pH 7.4, to give an indication of the folding effects observed in FRET

melting experiments, and at pH 5.5 to look for any evidence of a change in structure from the i-motif.

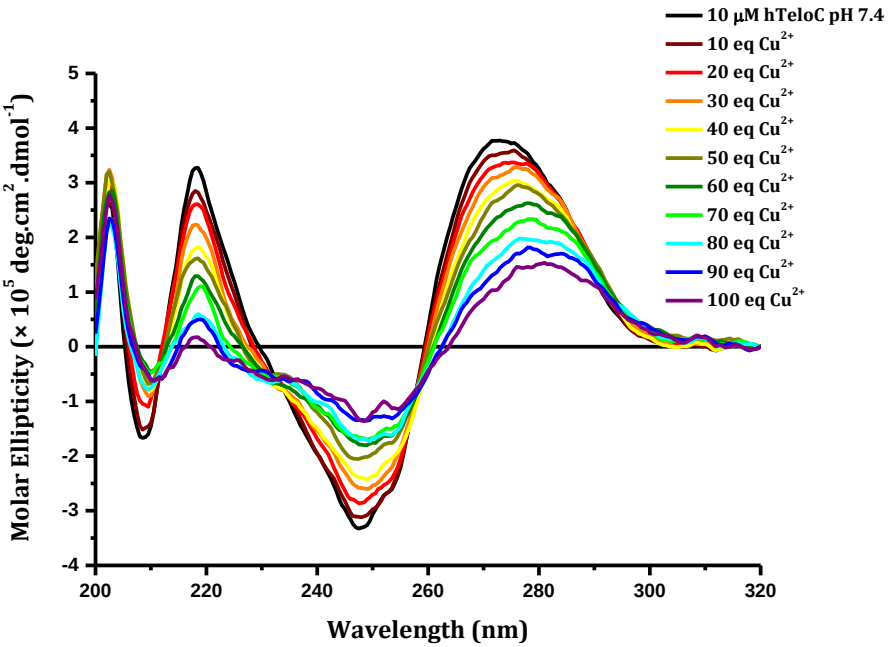



*Figure 4.29: CD spectrum of human telomeric i-motif sequence at pH 7.4 with increasing concentration of $CuCl_2$. [DNA] = 10 µM, buffer = 50 mM sodium cacodylate.*

At pH 7.4, the high concentrations of $Cu^{2+}$ required to replicate the same ratio as in the FRET melting experiments resulted in precipitation (figure 4.29). This can be seen by the decrease in molar ellipticity across the whole spectrum and by the fact that the curves themselves are less smooth, showing possible interference from precipitate. A small bathochromic shift in the positive peak is observed from 275 nm to 280 nm but the other peaks do not show any shift in wavelength. This small shift may again be due to precipitation as has been observed previously in the CD spectra of **BisA** in chapter 2.

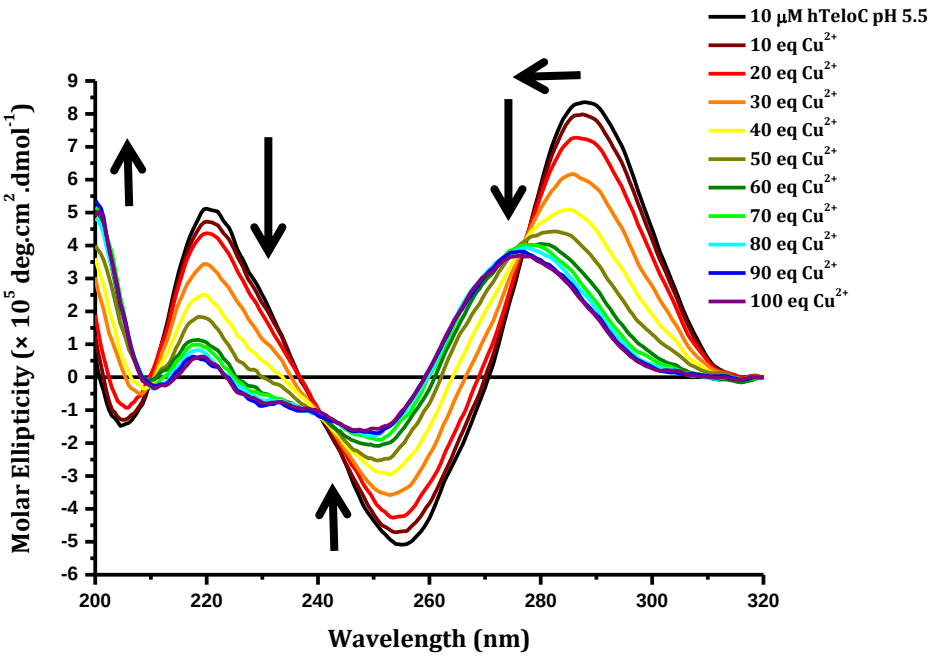



*Figure 4.30: CD spectrum of human telomeric i-motif with increasing concentration of $CuCl_2$ at pH 5.5. [DNA] = 10 µM, buffer = 50 mM sodium cacodylate.*

However, in contrast at pH 5.5 no precipitation occurred and the spectrum showed a gradual change with increasing copper concentration with the peak showing a hypsochromic shift from 288 nm to 276 nm (figure 4.30). The negative peak also showed a shift from 255 to 250 nm while the peaks at 220 and 200 nm showed a decrease and increase respectively. The spectrum also has multiple isodichroic points indicating that there are two structures in equilibrium with one another. After addition of 100 equivalents of $Cu^{2+}$, the spectrum is very similar to that of single stranded DNA. However, the FRET melting data suggests that copper does not simply unfold the DNA as this would be indicated by an increase in the fluorescence intensity with increasing copper concentration and this is not observed in the CD experiment.

By plotting the change in molar ellipticity at 288 nm with increasing concentration of $CuCl_2$ at pH 5.5 (figure 4.31), the data can be fitted to obtain the $[Cu^{2+}]_{50}$ for the change and the cooperativity of the interaction using equation 4.3. Where $K$ is the concentration at which half the structural change has occurred ($[Cu^{2+}]_{50}$) and $n$ is the Hill coefficient. $A$ is a constant (*A = start + (end-start)* values). Fitting this data gave a $[Cu^{2+}]_{50}$ = 378 µM and a Hill coefficient $n$ = 2.8 showing positive cooperativity ($n > 1$).

*Equation 4.3:* $$\theta = A\frac{[Cu^{2+}]^n}{K^n+[Cu^{2+}]^n}$$

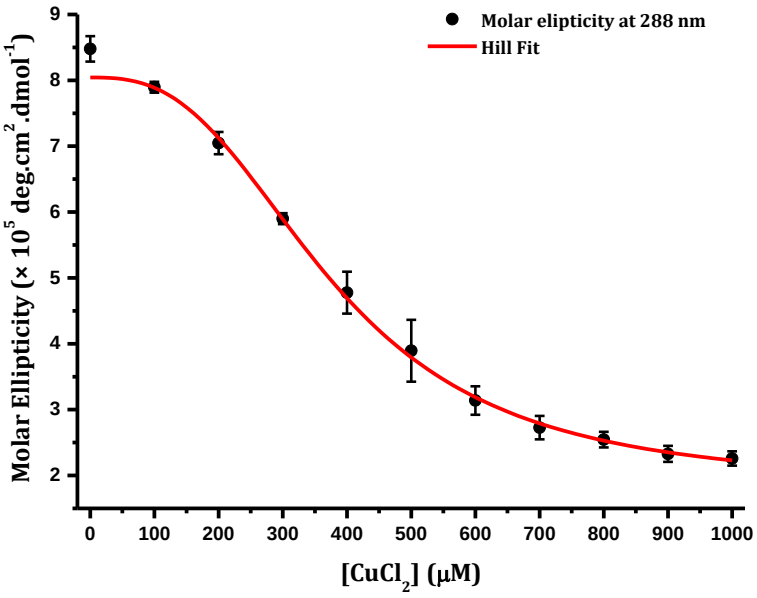



*Figure 4.31: Change in Molar ellipticity with increasing concentration of $CuCl_2$ at pH 5.5. Error bars show the standard deviation across 3 repeats. [DNA] = 10 μM, buffer = 50 mM sodium cacodylate.*

Comparing the CD spectra of the different structures (figure 4.32) shows that there is a slight difference between the spectrum of the single stranded i-motif sequence and that observed for the i-motif sequence with 100 equivalents of $Cu^{2+}$. The largest positive peak only shows a slight difference in wavelength (280 nm) compared to the random coil (275 nm). However, what is striking is that this same small difference is also observed for the C-hairpin sequence under acidic conditions, which should form a hairpin. The rest of the spectrum has a lower molar ellipticity than the single strand and is similar to some CD spectra of hairpins published in the literature.[206] Finally, there is an increase in molar ellipticity at 200 nm in both the C-hairpin sequence and in the presence of copper (II) ions, which is not observed in the spectra of either i-motif or single stranded DNA. This provides further evidence that copper cations may stabilise the formation of a hairpin structure in i-motif sequences.

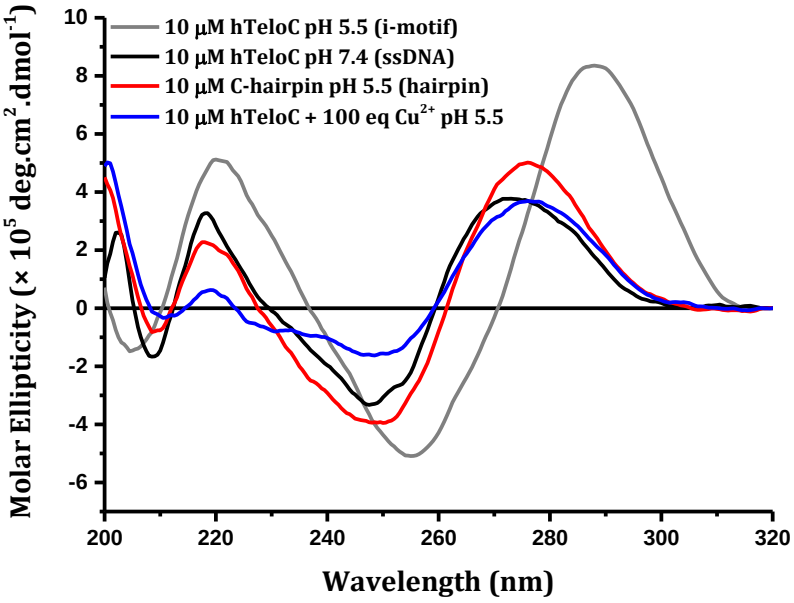



*Figure 4.32: CD spectra of different DNA structures i-motif, single stranded DNA and hairpin in comparison to hTeloC with 100 eq $CuCl_2$ at pH 5.5. [DNA] = 10 μM, buffer = 50 mM sodium cacodylate.*

These results provide further support for the results seen in sections 4.4.1 and 4.4.2. At pH 7.4 the CD titration indicates that copper (II) ions cause precipitation of the DNA but at pH 5.5 no precipitation occurs and it is clear that the addition of $Cu^{2+}$ causes the i-motif to change to a different structure. This could explain the observed destabilisation in the FRET melting experiments at pH 5.5. The resultant CD spectrum after addition of 100 equivalents of copper (II) ions shows subtle differences to the spectrum of unfolded random coil structure and shares similarities with the spectrum of the C-hairpin sequence. This offers further evidence in support of the UV difference spectra suggesting that copper (II) ions are able to promote the formation of a hairpin structure within i-motif forming oligonucleotide sequences under acidic conditions where the i-motif would ordinarily be formed.

### 4.4.4 Reversibility

In order to see whether the changes induced by copper cations were reversible, the CD sample was titrated with equal equivalents of ethylenediaminetetraacetic acid (EDTA). This resulted in complete reversal of the spectral changes and a return to the spectra previously obtained for the i-motif or single stranded structures.

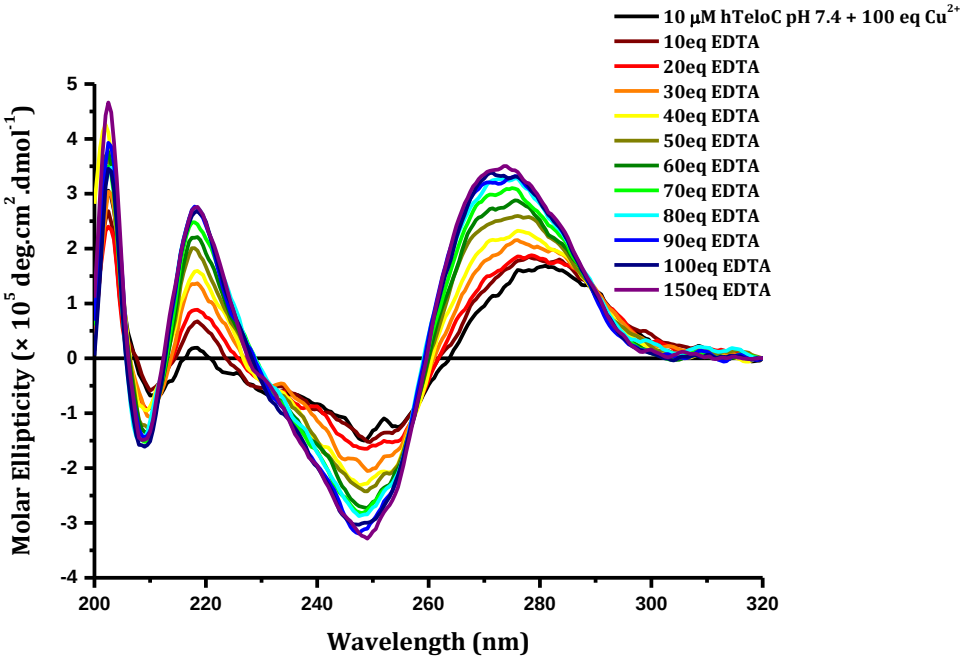



*Figure 4.33: CD titration of human telomeric i-motif sequence with 100 eq $CuCl_2$ and increasing concentration of EDTA at pH 7.4. [DNA] = 10 µM, buffer = 50 mM sodium cacodylate.*

At pH 7.4, it can be seen that there is an increase in molar ellipticity across the spectrum and there is a shift in the positive peak from 280 nm to 275 nm corresponding to the spectrum previously obtained for the random coil structure. The spectrum also becomes smoother. This indicates that addition of EDTA was able to re-dissolve the sample and afforded the single stranded DNA structure (figure 4.33).

At pH 5.5, (figure 4.34) there is a decrease in molar ellipticity at 200 nm, an increase at 220 nm, a decrease and a shift from 250 to 255 nm as well as an increase and shift from 276 to 288 nm. These changes are the opposite to those seen for copper and return the spectrum to that previously obtained for the i-motif. This suggests that the addition of EDTA is able to reverse the change induced by copper and switch the structure back from a hairpin form to an i-motif form.

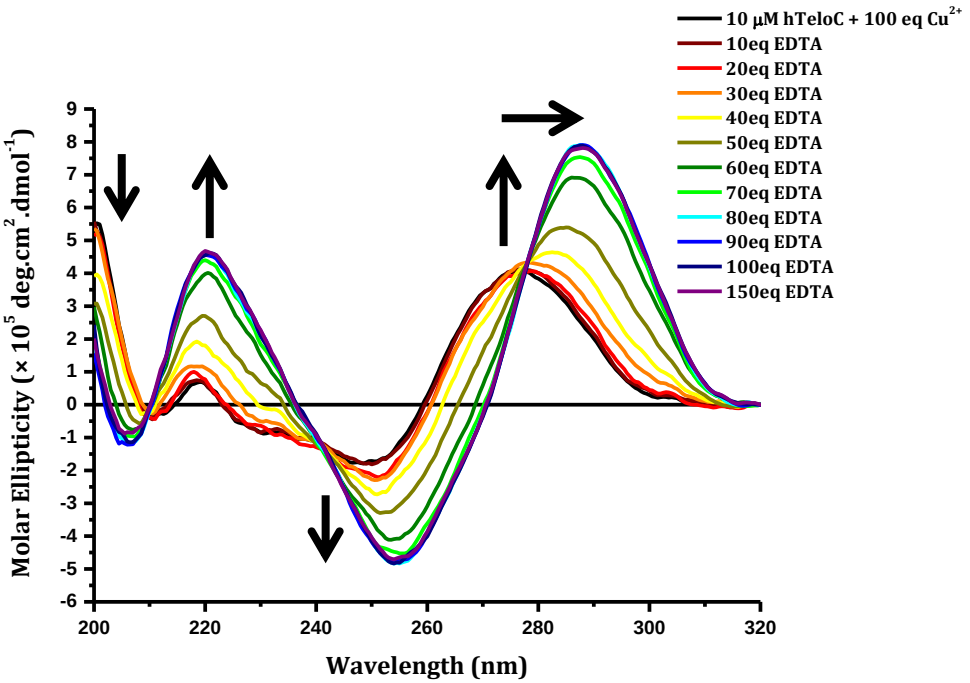



*Figure 4.34: CD titration of human telomeric i-motif sequence with 100 eq $CuCl_2$ and increasing concentration of EDTA at pH 5.5. [DNA] = 10 μM, buffer = 50 mM sodium cacodylate.*

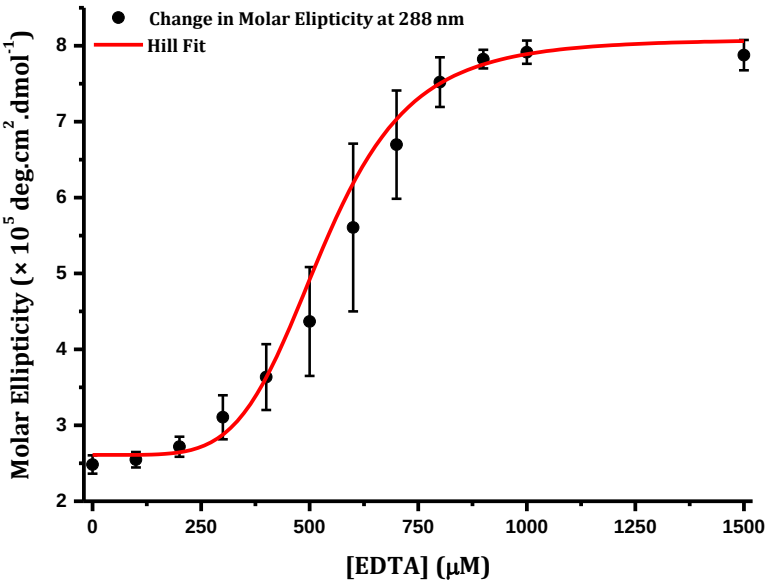



*Figure 4.35: Change in molar ellipticity with increasing concentration of EDTA at pH 5.5. Error bars show the standard deviation across 3 repeats. [DNA] = 10 μM, [$CuCl_2$] = 1000 μM, buffer = 50 mM sodium cacodylate.*

As with the data in figure 4.30, the change in molar ellipticity with increasing concentration of EDTA can be plotted to give a sigmoidal curve (figure 4.35). Fitting

this data with equation 4.3 gave a $[EDTA]_{50}$ = 531 μM and a cooperativity of $n$ = 5.2 indicating that slightly more EDTA than $Cu^{2+}$ is required for the reversal.

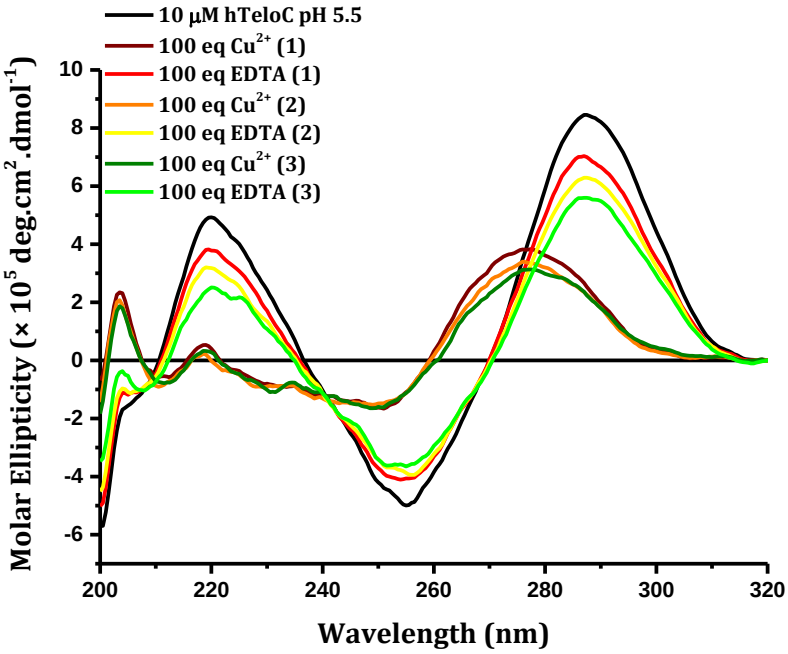



*Figure 4.36: CD spectrum of human telomeric i-motif with repeat additions of 100 eq $CuCl_2$ and 100 eq EDTA at pH 5.5. [DNA] = 10 μM, buffer = 50 mM sodium cacodylate.*

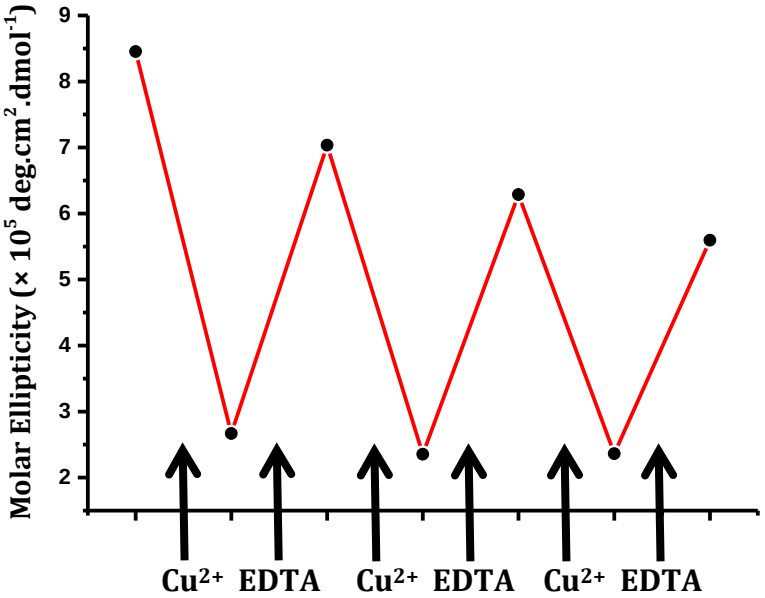



*Figure 4.37: Change in molar ellipticity at 287 nm of human telomeric i-motif with repeat additions of 100 eq $CuCl_2$ and 100 eq EDTA at pH 5.5. [DNA] = 10 μM, buffer = 50 mM sodium cacodylate.*

Figure 4.36 and 4.37 demonstrate that the DNA can be repeatedly switched from the i-motif structure to the apparent hairpin structure at pH 5.5 with repeated additions of $Cu^{2+}$ and EDTA. This is important as it offers a different kind of switchable motion to the normal folding and unfolding that could be used in nanomotor or logic gate applications.[21,22] This result is also interesting as the dynamics between i-motif and hairpin formation have been shown to be important in the work of Laurence Hurley and co-workers on the Bcl-2 promoter region but has yet to be fully understood.[69]

### 4.4.5 NMR studies on copper (II) ions with i-motif DNA

Proton NMR studies were carried out in collaboration with Dr Colin MacDonald. The change in structure between a hairpin and an i-motif has previously been demonstrated using NMR by Laurence Hurley and co-workers investigating the structure dynamics within the Bcl-2 promoter region.[69] The central imino protons which stabilise the cytosine-$H^+$-cytosine base pairs in a normal i-motif structure give a characteristic peak in $^1H$ NMR at around 15.5 ppm.[35] In Hurley and colleague's work, they showed that on moving from an i-motif to a hairpin structure, the imino proton peak is lost and a new peak forms at 12 – 13 ppm indicating the formation of a hairpin.

In order to gain further evidence of a structural transformation in i-motif DNA on addition of copper (II) ions, a 1D $^1H$ NMR experiment was carried out (Figure 4.38). The human telomeric sequence was prepared at a concentration of 10 μM in pH 5.5 50 mM sodium cacodylate buffer with 5% $D_2O$ to enable the NMR instrument to obtain a solvent lock. The spectrum was obtained over 1 hour before 100 equivalents (1 mM) of $CuCl_2$ was added and the spectrum taken again over 2 hours. Finally 100 equivalents (1 mM) of EDTA was added and the spectrum taken for a third time over 1 hour.

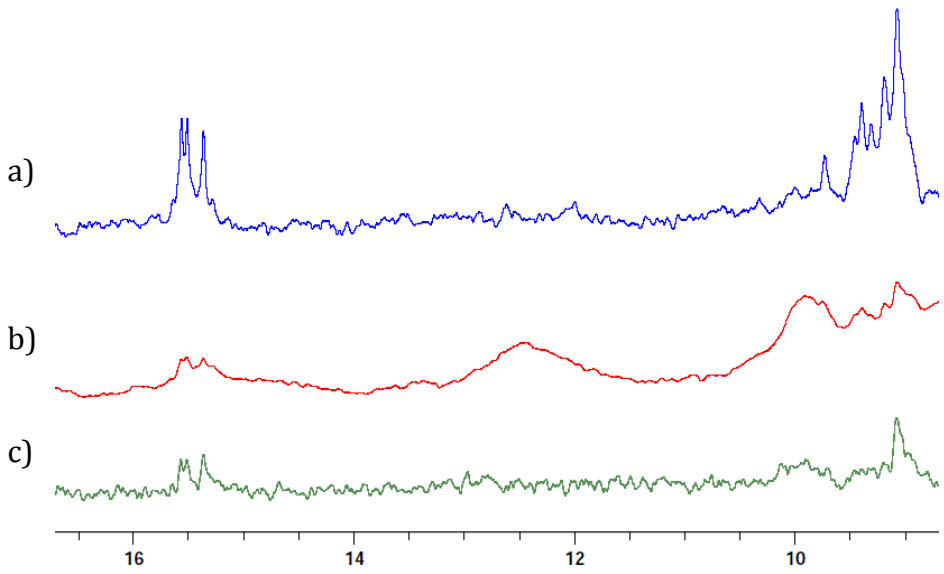



*Figure 4.38: Proton NMR of the imino proton region of human telomeric i-motif at pH 5.5. a) in the absence of $Cu^{2+}$, b) in the presence of 1 mM $Cu^{2+}$, c) in the presence of 1 mM $Cu^{2+}$ and 1 mM EDTA. The spectra have been scaled for easier comparison. [DNA] = 10 μM, buffer = 50 mM sodium cacodylate with 5% $D_2O$.*

The spectrum in the absence of copper (II) ions (figure 4.38a) clearly shows a peak at 15.5 ppm which corresponds to the central imino protons of the cytosine-$H^+$-cytosine base pairs. The peaks at 9-10 ppm represent the other protons on the bases in the sequence. After addition of 100 equivalents of $CuCl_2$ (figure 4.38) the spectrum has less resolution and the peaks become much broader. This is due to the presence of the $Cu^{2+}$ ions which are paramagnetic. However, it can clearly be seen that there is an apparent reduction in the peak at 15.5 ppm corresponding to the i-motif, while there is the formation of a new peak at 12.5 ppm. This peak indicates a change in structure and may be representative of the protons in a DNA hairpin as suggested by Hurley et. al.[69] Taken on its own, it is not certain precisely which protons this peak represents, however, in conjunction with the UV difference and circular dichroism spectra, this spectrum provides further evidence to suggest the formation of a hairpin structure. Finally after addition of 100 equivalents of EDTA, the spectrum becomes clearer, the peak at 12.5 ppm disappears while the imino proton peak at 15.5 ppm remains, but with lower intensity (figure 4.38c). This suggests that EDTA is able to reverse the changes in the spectrum induced by copper. This provides further support for the reversibility of the changes that have been demonstrated by circular dichroism.

### 4.4.6 Summary of effect of copper (II) ions on i-motif forming DNA

After it was initially identified as a hit from a screen of 26 different salts, copper (II) ions have been shown to interact with i-motif forming DNA sequences. At pH 7.4 where the DNA is unfolded, FRET melting experiments suggested that $Cu^{2+}$ caused the DNA to fold and UV difference experiments indicated that the folded structure was not an i-motif but may be a hairpin type structure as the "copper difference" spectrum shared similarities with the "silver difference" spectrum of a known silver DNA hairpin. However further FRET titration and circular dichroism experiments at this pH showed that the addition of copper (II) ions also resulted in precipitation. More interestingly, at pH 5.5 the addition of copper did not result in precipitation and, circular dichroism experiments also showed the formation of a possible hairpin structure. This was further supported by NMR experiments which showed a reduction in the imino proton peak at 15.5 ppm and the appearance of a new peak at 12.5 ppm which may be indicative of hairpin formation.[69] In both CD and NMR experiments the structural changes were shown to be reversible upon addition of EDTA.

These results are particularly exciting because the dynamics between hairpin and i-motif structures have been shown to be important in the case of the Bcl-2 promoter region but have not been investigated fully.[69] As well as this, most nanotechnology applications of i-motif DNA have relied on the structure changing from unfolded to folded, whether this be to create nanomechanical motion, or for sensing and logic gate applications.[203] The ability to reversibly change from one folded structure to a different folded structure offers the possibility of a different switch system and a different type of "mechanical motion." In combination with the existing ability of i-motif forming sequences to fold and unfold in response to changes in pH, or in response to silver and cysteine as shown in section 4.2, further work on this project could enable the development of switches or logic gates with multiple inputs and outputs using the same oligonucleotide sequence.

## 4.5 Group 1 and 2 cations destabilise the i-motif

As mentioned in the section 4.3, the cations were also tested at pH 5.5. Whilst the hits at pH 7.4 generally produced base level fluorescence signals, it was apparent that the cations from group 1 and group 2 had a destabilising effect on the i-motif (figure 4.39). Group 1 cations; potassium, rubidium and caesium as well as group 2 cations; magnesium, calcium, strontium and barium were tested as part of the screen in 10 mM sodium cacodylate buffer with 5 mM sodium chloride. At a cation concentration of 100 mM they all showed a destabilising effect on the i-motif (table 4.3). In order to try and complete the set of group 1 cations and to compare the results to previously published examples, lithium and sodium were also tested. However, francium, beryllium and radium were not tested because they were not available and all three are highly toxic, particularly in the case of francium and radium which are radioactive.

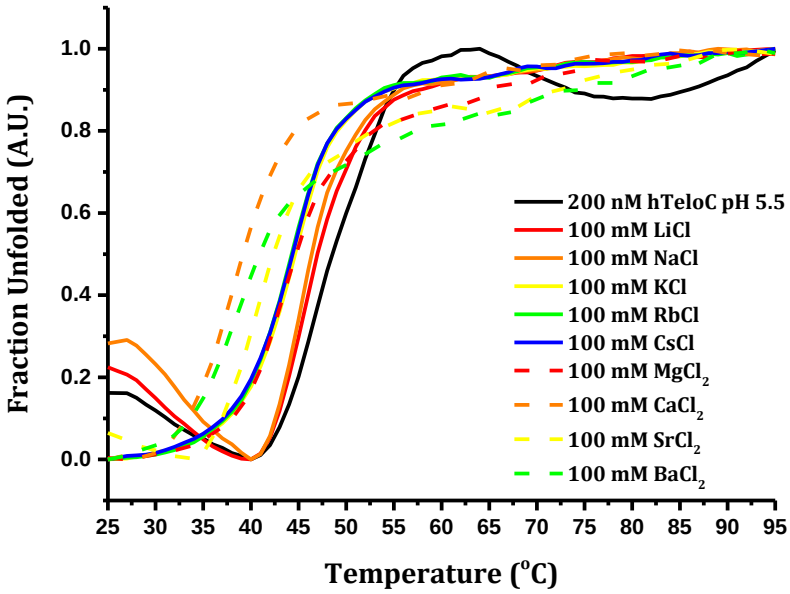



*Figure 4.39: Normalised FRET melting curves for human telomeric i-motif with 100 mM of group 1 and group 2 cations at pH 5.5. [DNA] = 200 nM, buffer = 10 mM sodium cacodylate and 5 mM NaCl.*

| [Salt] = 100 mM | $T_m$ (°C) | $\Delta T_m$ (°C) |
|---|---|---|
| LiCl | 45.8 | -0.7 |
| NaCl | 45.5 | -1.0 |
| KCl | 45.2 | -1.0 |
| RbCl | 44.2 | -2.0 |
| CsCl | 44.5 | -1.7 |
| $MgCl_2$ | 43.2 | -3.0 |
| $CaCl_2$ | 38.2 | -8.0 |
| $SrCl_2$ | 40.5 | -6.0 |
| $BaCl_2$ | 38.8 | -7.3 |

*Table 4.3: Change in melting temperature of human telomeric i-motif with 100 mM of group 1 and group 2 salts. The error in the $T_m$ values is ± 0.3 °C.*

These results give support to previous reports in the literature that $Li^+$ ions destabilise the i-motif,[52] as well as $Na^+$, $K^+$ and $Mg^{2+}$ and shows that this is also the case for other group 1 and 2 cations.[50,51] At pH 7.4 the group 1 cations have no effect on the unfolded DNA. The group 2 cations do cause a small decrease in fluorescence at 100 mM but this is not to base level suggesting it is not a folding event and there is no change with temperature. These results are only preliminary and are carried out at a pH where the i-motif is already stable. Further work could be carried out to see whether these cations promote unfolding of the i-motif using circular dichroism and FRET titrations at a pH closer to the transitional pH of the structure. This work offers important information to those working in the i-motif field about what factors affect i-motif stability particularly when choosing buffer conditions for example in the design of an i-motif screen or assay.

## 4.6 Conclusions

In the first part of this chapter, the interaction between silver cations and an i-motif forming sequence has been investigated, demonstrating that they are able to fold the DNA into an i-motif-like structure at neutral pH. This effect was observed using FRET melting and FRET titration experiments, circular dichroism and UV difference spectra summarised in section 4.2.6. Since these investigations were completed, the results have been supported by published results from other groups. Iryna Goncharova has

studied the effect of silver cations on the stability of C-C, G-G and C-G base pairs using electronic and vibrational circular dichroism.[207] Her results showed that silver (I) ions enabled the stacking of cytosine base pairs at pH 10 that resembled the spectra of stacked cytosine-$H^+$-cytosine base pairs in the i-motif. Matias Berdakin and co-workers took a different approach, comparing the structure of C-$Ag^+$-C base pairs to C-$H^+$-C base pairs in the gas phase using infrared-multiphoton dissociation spectroscopy.[208,209] Their results showed that the two types of base pair are both planar and share the same keto-amino structure with glycosidic bonds in the trans form. One difference they did observe is in the angle of the glycosidic bonds in the C-$Ag^+$-C base pair compared to the hemiprotonated pair. This difference in glycosidic bond angle may be one contributing factor to the difference in CD spectra observed in section 4.2.2 as the optical activity that is responsible for the circular dichroism spectra of DNA are a result of changes in glycosidic bond orientation. In contrast to these published examples, Ting-Ting Zhao and colleagues have looked at the interaction between i-motif and silver in the formation of silver nanoclusters at acidic pH. Their CD data indicate that at acidic pH, silver disrupts the i-motif and favours the formation of an antiparallel 4 stranded structure.[210] This may be supported by the FRET melting experiments carried out at pH 5.5 in section 4.2.1 (figure 4.4) which shows a small decrease in melting temperature for the first melting transition which is likely to be that of the i-motif, while there was a second, higher melting transition corresponding to a different structure. However, the CD spectrum they published does not resemble that observed at neutral pH in figure 4.7 as it does not have the positive peak at 240 nm and the strongly negative peak at 260 nm.

The formation of an i-motif-like structure was also shown to be reversible using cysteine to chelate the silver. These results have a number of implications. The fact that silver enables the formation of an i-motif like structure at neutral pH could enable the study of the structure and investigation of compounds interacting with the structure at neutral rather than acidic pH. The reversibility of the folding with cysteine could be used as an alternative switch mechanism for i-motif nanotechnology applications such as motors and switches which have so far relied on changes in pH to create the required structural change.[203] Since this work was carried out, the use of silver to form an i-motif-like structure has already been applied to the design of a logic gate with OR and INHIBIT functionality that relies on both $H^+$ and $Ag^+$ as inputs to the system.[211] When

combined with a fluorescent ligand that is selective for the i-motif, this work has been applied to the design of an i-motif based sensor for silver (I) ions with an estimated 0.25 nM limit of detection.[212] These results provide good evidence that addition of silver (I) ions results in a structure similar to the i-motif rather than a hairpin. Other applications in sensing of both silver and cysteine could be developed, particularly utilising the FRET effect of a dual labelled oligonucleotide, similar to examples that have already been published based on double stranded DNA.[193,194] It is worthwhile to note that whilst the results strongly indicate the formation of an i-motif type structure, they also show some differences compared to a normal acid stabilised i-motif. It is therefore not clear exactly how the silver cations stabilise this structure – are they forming cytosine-$Ag^{+}$-cytosine base pairs or are the silver ions sitting between base pairs or forming a nanocluster around which the DNA folds? To investigate this in more detail, structural studies will be required using NMR to build up a model of the structure or ideally a crystal structure would be obtained which could show precisely where the silver ions interact with the DNA.

Having investigated the effect of silver cations on the formation of i-motif DNA, a wider investigation of the effect of different cations on i-motif DNA was undertaken. A selection of 26 different cations including those from groups 1, 2 and 3, the transition metals and the lanthanides, were screened by FRET melting at pH 5.5 and pH 7.4. Initially it could be seen that the group 1 and 2 cations: $Li^{+}$, $Na^{+}$, $K^{+}$, $Rb^{+}$, $Cs^{+}$, $Mg^{2+}$, $Ca^{2+}$, $Sr^{2+}$ and $Ba^{2+}$ destabilised the i-motif at acidic pH with $\Delta T_m$'s between -0.7 and -8. This expands on previous work by Saxena,[50] Mergny,[51] and Kim.[52] Further work could be carried out using circular dichroism or FRET titrations to see if these cations promote the unfolding of the i-motif, particularly at the transitional pH of the structure. These results have important implications as several of these cations are found in biological conditions and are routinely used in buffers to vary the ionic strength. This information could be important in choosing conditions for the design of experiments on i-motif DNA.

The FRET screen at pH 7.4 identified 10 cations which seemed to show folding of the DNA. These were $Al^{3+}$, $Cu^{2+}$, $Ga^{3+}$, $Ho^{3+}$, $Sm^{3+}$, $Tb^{3+}$, $Yb^{3+}$, $Y^{3+}$, $Hg^{2+}$ and $Zn^{2+}$. Interestingly, $Cu^{2+}$, $Hg^{2+}$ and $Zn^{2+}$ share the same linear coordination geometry as $Ag^{+}$ which could enable them to facilitate C-C base pairs effectively. However when testing the cations

by CD, the majority caused precipitation of the DNA and lowered the pH at the high concentrations needed so they were not investigated further. $Cu^{2+}$ was the only cation that was investigated in greater detail as it showed some signs of change in the CD at lower concentration and still caused folding of the DNA in a higher concentration buffer that maintained the pH. $Hg^{2+}$ could not be tested due to the low solubility of the mercury salt used and would be worth investigating in the future with a more soluble salt such as $Hg(ClO_4)_2$ used by Ono and co-workers.

In the second half of this chapter the effect of copper (II) ions on i-motif forming DNA was examined in more detail. By completing an analogous set of experiments to those used on the silver (I) interaction, the results suggest that copper (II) cations are able to initiate and stabilise a change in structure from an i-motif to a hairpin at acidic pH, summarised in section 4.4.6. The structural changes observed were also demonstrated to be reversible with the addition of EDTA enabling the structure to be repeatedly switched from i-motif to hairpin form. However, as with the silver (I) stabilised structure, it is not clear from these experiments exactly how the copper ions interact with the DNA. A high concentration of $Cu^{2+}$ (1000 μM) is required to effect the changes, therefore it is unlikely that the $Cu^{2+}$ is only simply stabilising C-C base pairs, as far more copper (II) ions are needed than the number of bases present. Again further work is required using NMR or X-ray crystallography to fully characterise the copper stabilised structure. For example, NMR experiments have been used previously to identify metal ion coordination sites in the loop regions of a DNA minihairpin.[213] It is also not clear how the DNA rearranges itself when moving from the i-motif to a hairpin structure. Recently, molecular dynamics simulations have been used to model the unfolding of the i-motif,[214] therefore similar experiments could be used to shed light on the way the structure changes during this transition.

The dynamic equilibrium between hairpin and i-motif have already been shown to be important in a biological context. Laurence Hurley's group, working on the i-motif formed within the Bcl-2 promoter sequence showed that stabilisation of the hairpin resulted in down regulation and stabilisation of the i-motif resulted in up regulation of gene expression.[69] Therefore the effect of $Cu^{2+}$ could be a useful tool to study this relationship as it could be used to design *in vitro* assays to assess the binding of ligands to both the i-motif and hairpin structures formed from the same oligonucleotide

sequence. Without the use of copper to alter the structure, different sequences would be needed to form the two structures which may in itself affect the binding of potential ligands. Finally, the ability to reversibly alter the structure of DNA from the i-motif to a hairpin rather than just to and from unfolded DNA could provide useful applications in nanotechnology as it may offer a different kind of mechanical actuation. As has been described the ability of silver to form an i-motif-like structure has already been applied to the design of a logic gate.[211] With the additional reversible effects of copper (II) ions, it may now be possible to design switches and logic gates which respond to multiple input signals of protons, silver (I) ions and copper (II) ions that give multiple outputs of single strand, i-motif and hairpin structures, all with the same oligonucleotide sequence. This will be of special interest to those working in the fast growing field of DNA based logic operations and nanotechnology.

# Chapter 5: Conclusions and Future Work

## 5.1 Overall Conclusions

This thesis has described investigations into the effect of small molecule ligands and cations on i-motif DNA. In chapter 1 the current research into the biological role of i-motif structures has been reviewed highlighting the factors that show the i-motif is a promising target worthy of investigation while illustrating how little work has been done in comparison to more well established DNA structures such as the G-quadruplex. This provided the background for the work described in this thesis, which aimed to discover new chemical biological tools to interact with i-motif DNA structures. These could be used in the future to study the possible biological function of the i-motif in more detail.

In chapter 2, an existing i-motif ligand **BisA** was investigated. This ligand was chosen in order to be used as a comparison when looking for new i-motif ligands. **BisA** was one of the first i-motif ligands to be identified using a FRET melting assay but no other data was shown to support this.[100] Therefore, in this work a range of different biophysical techniques were used to look at the interaction between **BisA** and the i-motif in more detail for the first time. First the FRET melting experiments on **BisA** were repeated at pH 6.8 and after normalising the data produced comparable results to those shown by Alberti and co-workers with an apparent stabilisation of $\Delta T_m$ = 55°C. However, the raw FRET melting data showed a decrease in fluorescence signal across the temperature range of the experiment indicating that the DNA did not melt at up to 95°C or that aggregation of the DNA had occurred. To look at this in more detail circular dichroism was used. This technique has not been used to look at the interaction between **BisA** and the i-motif before. Interestingly this showed an ICD signal for **BisA** which suggests a binding event and later a binding affinity was measured using fluorescence titrations ($K_{D1}$ = 2.1 μM, $K_{D2}$ = 145.9 μM). However, the rest of the CD spectrum showed a decrease in the i-motif signals indicating condensation of the DNA. This was further supported by UV titrations showing a decrease in absorbance at 260 nm and by polyacrylamide gel electrophoresis which showed retardation and smearing of the bands associated with **BisA** binding. Finally dynamic light scattering was used to measure the formation of condensed particles, showing that bigger particles form at pH 5.5 compared to pH 6.8 and 7.4. The condensation effects were

also apparent with G-quadruplex and double stranded DNA. Control experiments were carried out with the mono-acridine (**MonoA**) and the linker molecule diethylenetriamine (**DETA**) and these did not show the same condensation effects. This indicates that it is the combination of two acridine units in the molecule, not just the positive charges, which causes this condensation effect.

After identifying the problems with the known i-motif ligand **BisA**, it unfortunately was not suitable to be used as a comparison when looking for new i-motif ligands but did highlight the issue of condensation and how this can affect the biophysical techniques used to investigate i-motif-ligand interactions. With this in mind, in chapter 3 a diverse library of 960 known biologically active compounds was screened by FRET melting to identify new i-motif ligands based on the ability of the ligand to stabilise or destabilise the structure. This resulted in 34 hits which were narrowed down to 13 after setting a criteria of $\Delta T_m$ = ±5°C. These hits were further reduced down after concentration dependence studies and circular dichroism experiments to 6 lead compounds. These were tested against different DNA structures by FRET melting and tested for binding by SPR.

From this screen two new i-motif binding compounds were identified; **tilorone** and **mitoxantrone** (figure 5.1). Both are already known DNA binding compounds but have never before been shown to interact with i-motif DNA.[163,171] The two compounds were shown to stabilise the DNA and binding affinities were measured by surface plasmon resonance in the low micromolar range ($K_D$ = 157.9 μM and 11.0 μM respectively). Since **mitoxantrone** has a much stronger binding affinity than **tilorone** it can be considered as the key hit from the screen. Whilst neither ligand showed any selectivity in their binding, **mitoxantrone** in particular provides a good starting point to develop novel analogues with better affinities and selectivities which are different to existing i-motif ligands.

In chapter 4, the focus moved on to the interaction of metal cations with i-motif DNA. It has previously been shown that silver cations stabilise cytosine-cytosine base pairs which lead to the hypothesis that silver may enable the formation of i-motif structure at neutral pH where the structure can be formed utilising cytosine-$Ag^+$-cytosine base pairs instead of the cytosine-$H^+$-cytosine base pairs which do not readily form at this

high pH. This hypothesis was tested using FRET melting which showed folding at neutral pH. CD titrations then showed formation of an i-motif like structure. FRET titrations showed the cooperativity of the interaction with an $[Ag]_{50}$ = 1.9 μM. After this, the folding was shown to be reversible using cysteine to chelate the silver, thus providing an alternative switch mechanism for control of i-motif folding at neutral pH.

Finally, the last part of chapter 4 extended the investigation of cations to screen 26 different cations from groups 1, 2 and 3 as well as the transition metals and lanthanides. This screen showed that the group 1 and 2 cations destabilise the i-motif supporting work previously carried out on Lithium, sodium, potassium and magnesium and extending the study to include other group 1 and 2 cations not previously investigated.[50–52] As well as this, copper (II) was identified for its ability to cause folding at neutral pH but as shown by CD experiments does not form an i-motif structure. In fact, it appears to favour formation of a different folded structure. A combination of UV difference, CD and NMR experiments all indicate that this structure is likely to be a hairpin. Like the effect of silver, the effect of copper was also reversible using EDTA and offers an alternative folding and unfolding mechanism which could be useful in nanotechnology applications.

## 5.2 Future Work

The most obvious priority for future work is to begin designing and testing analogues based on screen hit compounds in order to improve affinity and selectivity for the i-motif over other secondary DNA structures. Both the ligands (figure 5.1) have a core planar structure with attached aliphatic chains. This provides the opportunity to vary the number and position of the chains, the length of the chains and how many hydrogen bond donors or acceptors are present. For example, adding another two aliphatic chains to **mitoxantrone** in the 5 and 8 positions may inhibit intercalation and disfavour duplex binding thus improving selectivity. Meanwhile, increasing the length of the chains at the 1 and 4 positions and increasing the number of hydrogen bond donors and acceptors within the side chains could enable a greater degree of interaction with loop regions in the i-motif and hence improve affinity. With the methods established in this thesis, analogues can quickly be tested by FRET melting to

look for stabilisation and compared to other structures such as the G-quadruplex or duplex.

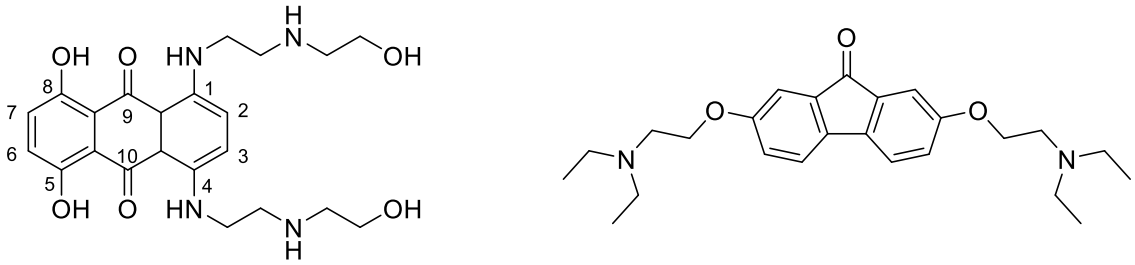


*Figure 5.1: Structures of screen hits **mitoxantrone** and **tilorone**.*

Another hit from the screen worthy of investigation is **tyrothricin**. Although its binding affinity could not be measured by SPR, it is clear from the FRET melting experiments that tyrothricin has some effect on the i-motif structure. **Tyrothricin** itself is actually a mixture of different peptides called tyrocidines and gramicidins.[162] These need to be separated using HPLC or synthesised separately so that they can be tested individually. It may be that one of the compounds has a more significant effect than the mixture as a whole or the separate compounds may have opposing effects. Several reports by Hansjürgen Ristow and colleagues have looked at the interaction between tyrocidine and gramicidin and DNA suggesting that tyrocidine forms a tight complex with DNA which is disrupted by gramicidin.[215,216] They have also shown that tyrocidine reduces torsional stress in DNA by forming a complex and altering DNA conformation with an effect on the regulation of sporulation which involves DNA replication.[179,180] These investigations were carried out before the i-motif structure was proposed but suggest that tyrocidine does have an effect on DNA structure and could be of interest in the context of the i-motif.

Aside from looking at analogues of the screen hit compounds, it is also important to look at different i-motifs. For example, **tilorone** showed only a weak affinity for the human telomeric i-motif however it may have a higher affinity for a different i-motif structure. **Tilorone's** mode of action as an antiviral agent is to induce the interferon response but it is not clear precisely how it achieves this. The promoter regions of the interferon regulatory factor genes are rich in cytosine and may well contain a potential i-motif forming sequence. The interferon regulatory factors are transcription factors which control the expression of interferon, therefore the possibility of an i-motif in the IRF promoter region could be investigated as a possible mode of action for tilorone. As

well as this specific example, there is quite a lot of variety even between the relatively small numbers of i-motif structures that have been identified. In particular they vary in their loop length and sequence which could significantly affect the binding of a ligand. It is therefore important to test the hit compounds and any analogues against a range of different i-motif structures as some may have better affinities and specificities than others. For example **IMC-48** identified by Laurence Hurley and colleagues was not only specific for i-motif but was specific for the Bcl-2 i-motif over other i-motifs such as VEGF and c-Myc.[69]

The library screened in this work was quite small in comparison to other available libraries and whilst its diversity enabled the inclusion of a wide variety of different structure types it was difficult to see any trends in the types of ligands which may bind to the i-motif. Now, having established methods for screening libraries against the i-motif, larger libraries could be screened to enable trends in i-motif binding compounds to be identified and more focused libraries of compounds analogous to the hit compounds could be screened to increase the number of chemical biological tools available to study the i-motif and look at structure activity relationships. Another drawback of this screen is that, despite initially identifying 34 interacting compounds, this resulted in only 2 main leads due to problems such as precipitation and possible micelle formation. Whilst this in itself is a reasonable hit rate from a comparatively small compound library, this could be improved by screening a library with better characteristics such as solubility.

Throughout this thesis, the primary experiment for identifying an interaction with i-motif DNA has been FRET melting, identifying compounds based on their ability to stabilise the DNA shown by an increase in melting temperature. However this precludes the possibility of compounds which bind to the i-motif but do not increase its melting temperature. A different screening method such as that used by Laurence Hurley and co-workers could be used to look at i-motif interaction by change in fluorescence.[69] By testing at the transitional pH where the i-motif and the unfolded DNA are in equilibrium, a decrease in fluorescence would indicate a shift in the equilibrium towards i-motif formation whilst an increase would suggest the opposite. Re-screening the compound library by this alternative method may provide some

additional hit compounds which can shift the equilibrium towards i-motif but did not necessarily increase the melting temperature of an already stable i-motif structure.

As described above, in chapter 4, the effect of different cations was investigated and it was shown that silver cations are able to stabilise the formation of an i-motif like structure at neutral pH. Further work in this area would be to apply this property for use as an alternative switch mechanism to fold the i-motif without the need for changing the pH. The rapid folding of the i-motif after decreasing pH has led to its use in a variety of nanotechnology applications which could now be extended to included $Ag^+$ as a trigger.[203] This has been illustrated recently with the design of OR and INHIBIT logic gates based on the i-motif structure using both $H^+$ and $Ag^+$ as inputs.[211] Here Yunhua Shi and co-workers used a fluorescent probe to detect i-motif formation in the presence of $H^+$ or $Ag^+$ or both. As well as performing logic operations, the sensing of silver (I) ions has also been shown using an i-motif forming sequence combined with a fluorescent probe, enabling detection of silver ions at nanomolar concentrations.[212] In the next part of the chapter, it was shown that unlike silver, copper cations stabilise the formation of a different type of structure within i-motif sequences which is likely to be a hairpin. This is exciting as it means that using the same oligonucleotide strand, three different structural conformations can be formed: single strand, hairpin and i-motif depending on the different concentrations of $H^+$, $Ag^+$ or $Cu^{2+}$. This is likely to be of interest to those working in the field of nanotechnology applications, enabling the design of logic gates or switches with multiple inputs and outputs.

Another potential hit from the cation screen that is worthy of further investigation is mercury. This appeared to cause folding of the DNA in FRET melting experiments at pH 7.4 with a melting temperature of 61.5°C (at $[Hg^{2+}]$ = 10 μM). Unfortunately this could not be investigated further because the salt $HgAc_2$ was not sufficiently soluble. A more soluble salt such as $Hg(ClO_4)_2$ used by Akira Ono and co-workers,[186] could enable the testing of $Hg^{2+}$ with i-motif by CD and UV difference spectroscopy to further understand its effect on i-motif DNA.

# Chapter 6: Experimental

## 6.1 General Experimental

**Chemical reagents** were used as supplied unless otherwise stated and were of general purpose grade. Solvents were purchased from either Sigma-Aldrich or Fisher.

**NMR spectra** were acquired on a Bruker Ultrashield Plus 400 MHz spectrometer and analysed using TopSpin 3.1. Deuterated solvents were used as indicated and at ambient probe temperature (300 K). The notation used for splitting patterns in $^{1}$H-NMR are: singlet (*s*), doublet (*d*), triplet (*t*), quartet (*q*), multiplet (*m*) and broad (*br*). The signals are given as δ values in ppm with coupling constants (*J*), provided in Hertz.

**IR spectra** were acquired as neat samples on a Perkin Elmer Spectrum BX FT-IR system with Spectrum v5.3.1 software.

**Mass spectra** were obtained from the EPSRCs National Mass Spectrometry Service Centre at the University of Swansea using atmospheric-pressure chemical ionisation (APCI).

**TLC** was carried out on Merck Kieselgel 60 $F_{254}$ plates and spots were visualised under UV-light.

**Flash chromatography** was carried out using Aldrich 60 silica gel at room temperature under a positive pressure of air.

**Melting Points** were measured on an Electrothermal melting point apparatus using open capillary tubes.

**Oligonucleotides** both labelled and unlabelled (table 6.1) were purchased from Eurogentec, purified by rpHPLC. Solid DNA samples were dissolved in MilliQ water to make stock solutions of 100 μM for labelled and 1 mM for unlabelled oligonucleotides. Using the extinction coefficients provided for each sequence, final concentrations were confirmed by UV absorbance at 260 nm with a Nanodrop ND-1000 spectrophotometer instrument. These stocks were then diluted in the appropriate buffers for further experiments.

| Name | Sequence 5'→3' |
|---|---|
| hTeloC | d(TAACCCTAACCCTAACCCTAACCC) |
| hTeloG | d(GGGTTAGGGTTAGGGTTAGGG) |
| DS[a] | d(TATAGCTATA-HEG(18)-TATAGCTATA) |
| c-Myc | d(TCCCCACCTTCCCCACCCTCCCCACCCTCCCCA) |
| C-hairpin | d(CTCTCTTCTCTTCATTTTTCAACACAACACAC) |
| HIF-1α | d(CGCGCTCCCGCCCCCTCTCCCCTCCCCGCGC) |
| PDGFA | d(CCGCGCCCCTCCCCCGCCCCCGCCCCCGCCCCCCCCCCCCC) |
| hTeloC$_{FRET}$ | FAM-d(TAACCCTAACCCTAACCCTAACCC)-TAMRA |
| hTeloG$_{FRET}$ | FAM-d(GGGTTAGGGTTAGGGTTAGGG)-TAMRA |
| DS$_{FRET}$ | FAM-d(TATAGCTATA-HEG(18)-TATAGCTATA)-TAMRA |
| c-Myc$_{FRET}$ | FAM-d(TCCCCACCTTCCCCACCCTCCCCACCCTCCCCA)-TAMRA |
| 18 base marker | FAM-d(GTAAAACGACGGCCAGTG) |
| hTeloC$_{Biotin}$ | Biotin-d(TAACCCTAACCCTAACCCTAACCC) |
| c-Myc$_{Biotin}$ | Biotin-d(CCTTCCCCACCCTCCCCACCCTCCCCA) |
| DS$_{Biotin}$ | Biotin-d(GGCATAGTGCGTGGGCGTTAGC) |
| DS compliment | d(GCTAACGCCCACGCACTATGCC) |

*Table 6.1: Sequences used throughout this research.* [a] *The double stranded sequence consists of two complimentary 10 base sequences linked by an 18 unit hexaethylene glycol (HEG) polymer.*

**Annealing DNA**. Where stated, samples were thermally annealed by heating to 95°C in a heat block before being allowed to cool to room temperature slowly over night. In general the sequences were annealed if forming duplex or G-quadruplex structures was required. The i-motif forming sequences were annealed at pH 6.8 but in general were not annealed at pH 5.5 because formation of the i-motif structure is rapid under the acidic conditionsand therefore was not necessary. The i-motif forming sequences were also not annealed when tested at pH 7.4 where the random coil structure was required.

**FRET melting** experiments were carried out on a QIAgen Roto-gene Q PCR machine using QIAgen PCR tubes. The following fluorescently labelled oligonucleotide sequences were used hTeloC$_{FRET}$, hTeloG$_{FRET}$, DS$_{FRET}$, c-Myc$_{FRET}$ (table 6.1). In a typical experiment, 400 nM solutions of labelled oligonucleotide sequence were made up in the appropriate buffer. The donor fluorophore FAM was 6-carboxyfluorescein, the

acceptor fluorophore TAMRA was 6-carboxytetramethylrhodamine. Ligand stock solutions were made up in either water or DMSO and then diluted to the appropriate concentrations in buffer by serial dilution. Cation stock solutions were made up in water. 10 μL of DNA solution were combined with 10 μL of ligand or cation solution and the samples heated from 25 to 95°C at a rate of 1°C per minute. The samples were excited at 470 nm and the fluorescence of FAM was monitored at 510 nm using a green filter at 1°C intervals. The final data was analysed using QIAgen Roto-gene Q-series software, Microsoft Excel and Origin Pro 9.1.

**Fluorescence titrations** were carried out on a Perkin Elmer LS55 fluorescence spectrometer using a quartz cuvette with a 10 mm path length. For FRET titrations, the appropriate FRET labelled oligonucleotide was diluted to 100 nM in either pH 5.5 or pH 7.4 buffer. The sample was excited at 490 nm and the fluorescence emission measured from 500 – 600 nm allowing the monitoring of emission peaks for both FAM (515 nm) and TAMRA (580 nm). Ligand or cation solutions were diluted in buffer and titrated into the sample in small aliquots using a Hamilton syringe or a pipette. The sample was mixed in the cuvette using a pipette and the fluorescence measured. The spectra measured are the average of 3 scans. For ligand fluorescence titrations, stock solutions of ligand were diluted in buffer and both the emission and excitation spectra were measured (200 – 800 nm) to find the most appropriate excitation and emission wavelengths. The concentration of ligand, slit width and scan speed were optimised to give the best spectrum for each ligand. Small aliquots of DNA were titrated into the sample using a Hamilton syringe or a pipette and the change in fluorescence of the ligand was measured. The data was analysed using Spekwin 32, Microsoft Excel and Origin Pro 9.1.

**Circular dichroism** experiments were carried out on a JASCO J-810 spectropolarimeter, under a constant flow of nitrogen using a quartz cuvette with a 1 mm path length. In a typical experiment, hTeloC oligonucleotide was diluted in pH 5.5, pH 6.8 or pH 7.4 buffer to a concentration of 10 μM. Sample volumes were 200 μL. Aliquots of ligands or cations were added using a Hamilton syringe or a pipette and the samples were scanned from 400 nm to 200 nm using a data pitch of 0.5 nm, a scan speed of 200 $nm.min^{-1}$, a response time of 1 s and a slit width of 2 nm. Each spectrum shown is the average of 3 scans. Each sample had a scan of the appropriate buffer

subtracted from it. The spectra were zero corrected at 320 nm and corrected for the effect of dilution. The data was analysed using Microsoft Excel and Origin Pro 9.1.

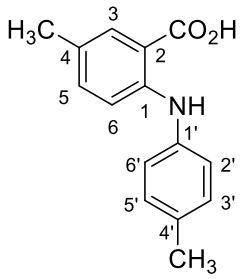
**UV Absorbance** titrations were carried out on either a Hitachi U-3010 spectrophotometer or an Agilent Technologies Cary 60 UV-Vis spectrometer as described, using a quartz cuvette with a 10 mm path length. UV melting experiments were carried out on an Agilent Technologies Cary 60 UV-Vis spectrometer with a Quantum Northwest TC 1 temperature controller. For UV absorption titrations, unlabelled oligonucleotide sequences were diluted in either pH 5.5, pH 6.8 or pH 7.4 buffer and the absorbance measured from 200 to 400 nm or from 200 to 800 nm. Small aliquots of ligand or cation were titrated into the sample using a Hamilton syringe or a pipette and the change in absorbance monitored. An analogous titration of ligand or cation in just buffer was also carried out and subtracted from the titration with DNA. Data was baseline corrected by subtraction of the spectrum of just buffer and zero corrected at either 400 or 800 nm as described. The data was analysed using Microsoft Excel and Origin Pro 9.1.

## 6.2 Experimental for Chapter 2

### 6.2.1 Synthesis of BisA and MonoA.

These steps were carried out following literature procedures.[136,149,154]

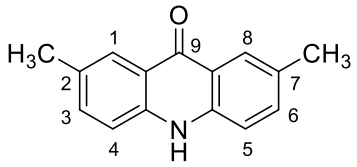
***4,4'-dimethyldiphenylamine-2-carboxylic acid (3)***




**3**

2-amino-5-methylbenzoic acid (5.013 g, 33.2 mmol, 1.1 eq), *p*-bromotoluene (5.177 g, 0.0302 mol, 1 eq), potassium carbonate (4.14 g, 30.2 mmol, 1 eq), copper iodide (95 mg) and copper powder (94.8 mg) were dissolved in n-pentanol (dried over molecular sieves, 15 mL) and was heated to reflux at 150°C with vigorous stirring overnight. The hot reaction mixture was filtered through celite and then washed thoroughly with methanol. The dark filtrate was acidified with 6 M hydrochloric acid then left in the fridge overnight. Product crystals were collected by filtration and recrystallised from ethanol. The remaining dark filtrate was reduced to dryness, re-dissolved in DCM and filtered through a short silica column (100% DCM) until the fractions were clear. The yellow fractions were combined and reduced to give the crude product which was crystallised from ethanol to give **3** as a yellow solid (4.64 g, 64% from 4 crops)
**$^1$H NMR ($CDCl_3$) $\delta_H$ ppm:** 9.05 (1H, N**H**, br s), 7.82 (1H, Ar**H**, m), 7.12 (6H, Ar**H**, m), 2.35 (3H, C**H**$_3$, s), 2.26 (3H, C**H**$_3$, s). This is consistent with the literature.[136]

***9-oxo-9,10-dihydro-2,7-dimethylacridone (8)***



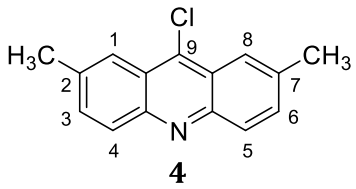

**8**

4,4'-Dimethyldiphenylamine-2-carboxylic acid **3** (2.27 g, 9.42 mmol, 1 eq) was dissolved in concentrated sulfuric acid (25 mL) and heated to reflux at 100°C for 1 hour.[154] The reaction was then allowed to cool and poured over ice water (20 mL). The yellow precipitate was filtered and washed with water. The insoluble material was collected and recrystalised from DMF to give the pure product **8** as a yellow solid (0.814 g, 39% from 3 crops)

**m.p.:** > 355°C, **$R_f$:** 0.29 (eluent: 100% DCM)

**$^1$H NMR (DMSO) $\delta_H$ ppm:** 11.59 (1H, N**H**, s), 8.02 (2H, Ar**H** 1 and 8, ap s), 7.55 (2H, Ar**H** 3 and 6, dd, $J_1$ = 8.5 Hz, $J_2$ = 2.0 Hz), 7.44 (2H, Ar**H** 4 and 5, d, $J$ = 8.5 Hz), 2.42 (6H, C**H**$_3$, s). Assignments were confirmed by 2D COSY NMR. **$^{13}$C NMR (DMSO) $\delta_C$ ppm:** 176.31 (1C, **C**O 9), 138.87 (2C, Ar**C**), 134.70 (2C, Ar**C** 1 and 8), 129.77 (2C, Ar**C**), 125.04 (2C, Ar**C** 3 and 6), 120.21 (2C, Ar**C** 2 and 7), 117.21 (2C, Ar**C** 4 and 5), 20.56 (2C, **C**H$_3$). **IR:** 2785 cm$^{-1}$ (C-H), 1637 cm$^{-1}$ (C=0), 1480 cm$^{-1}$ (C=C) **HRMS (APCI):** Calculated for $C_{15}H_{14}O_1N_1$ ([M+H$^+$]): 224.1070. Found: 224.1071.

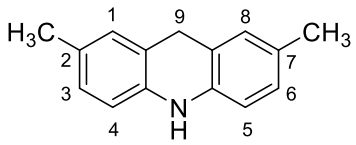
***9-Chloro-2,7-dimethylyacridine (4)***




**4**

Acridone **8** (106.7 mg, 0.478 mmol) was dissolved in thionyl chloride (1 mL) with 1 drop of DMF and heated to reflux at 90°C for 1.5 hours.[152] After which, the mixture was reduced *in vacuo*, washed with benzene and reduced to dryness *in vacuo* to give a yellow solid. The crude product was purified by flash column chromatography (100% DCM) to give the yellow product as a solid (122.7 mg, impure due to remaining DCM). **$^1$H NMR (DMSO) $\delta_H$ ppm:** 8.02 (2H, Ar**H** 4 and 5, d, $J$ = 8.8 Hz), 7.97 (2H, Ar**H** 1 and 8, ap s), 7.50 (2H, Ar**H** 3 and 6, dd, $J_1$ = 8.8 Hz, $J_2$ = 1.6 Hz), 2.50 (6H, C**H**$_3$, s) (This corresponds with the literature,[136]). **HRMS (APCI):** Calculated for $C_{15}H_{13}N_1Cl_1$ ([M+H$^+$]): 242.0731. Found: 242.0734.

***9,10-dihydro-2,7-dimethylacridine (9)***



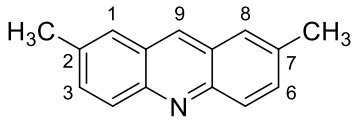

**9**

The acridone **8** (640 mg, 2.87 mmol, 1 eq) was suspended in anhydrous THF (11 mL) under nitrogen. 1 M Borane-THF complex (3.44 mL, 3.44 mmol, 1.2 eq) was added drop-wise and the reaction was heated to reflux under nitrogen overnight. 3 M aqueous hydrochloric acid solution (20.8 mL) was added before the mixture was basified with saturated sodium carbonate solution (88 mL). The mixture was extracted with DCM (3 x 50 mL), dried over magnesium sulfate and reduced *in vacuo*. The unstable product was quickly purified on a short column (eluent: 100% DCM) and reduced *in vacuo* to give a white solid (376 mg, 63%).

**m.p.:** 218°C, **$R_f$:** 0.78 (100% DCM)

**$^1$H NMR ($CDCl_3$) $\delta_H$ ppm:** 6.91 (2H, Ar**H** 1 and 8, s), 6.87 (2H, Ar**H** 3 and 6, d, *J* = 7.9 Hz), 6.55 (2H, Ar**H** 4 and 5, d, *J* = 7.9 Hz), 5.79 (1H, N**H**, br s) 3.98 (2H, C$\mathbf{H}_2$ 9, s), 2.26 (6H, C$\mathbf{H}_3$, s). Assignments were confirmed by 2D COSY NMR. **$^{13}$C NMR ($CDCl_3$) $\delta_C$ ppm:** 138.03 (2C, Ar**C**), 129.57 (2C, Ar**C**), 129.12 (2C, Ar**C**), 127.40 (2C, Ar**C**), 119.86 (2C, Ar**C**), 113.21 (2C, Ar**C**), 31.38 (1C, **C**$H_2$ 9), 20.59 (2C, **C**$H_3$). **IR:** 3400 $cm^{-1}$ (N-H), 2920 $cm^{-1}$ (C-H), 1610 $cm^{-1}$ (C=C), 1510 $cm^{-1}$ (C=C)**. HRMS (APCI):** Calculated for $C_{15}H_{16}N_1$ ([M+$H^+$]): 210.1277. Found: 210.1277.

### *2,7-dimethylacridine (5)*

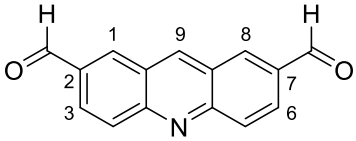



**5**

The acridane **9** (376.2 mg, 1.8 mmol, 1 eq) was stirred in ethanol (47 mL) and water (9.5 mL) with iron trichloride (1.42 g) at 50°C for 1 hour. Saturated sodium bicarbonate solution (81.47 mL) was then added and the mixture extracted with DCM (3 x 50 mL), dried over magnesium sulfate and reduced *in vacuo*. This gave the yellow product **5** (358.3 mg, 96%), which was used in the next step without purification.

**$^1$H NMR ($CDCl_3$) $\delta_H$ ppm:** 8.55 (1H, Ar**H** 9, s), 8.11 (2H, Ar**H** 4 and 5 d, *J* = 8.9 Hz), 7.73 (2H, Ar**H** 1 and 8, ap s), 7.59 (2H, Ar**H** 3 and 6, dd, $J_1$ = 8.9 Hz, $J_2$ = 1.92 Hz), 2.58 (6H, C**H**$_3$, s). This is consistent with the literature which gave **($CDCl_3$, 200 MHz):** 8.56 (1H, s), 8.12 (2H, d, *J* = 9.0 Hz), 7.73 (2H, s), 7.61 (2H, d, *J* = 9.0 Hz), 2.58 (6H, s).[136]

### *2,7-acridinedicarboxaldehyde (6)*

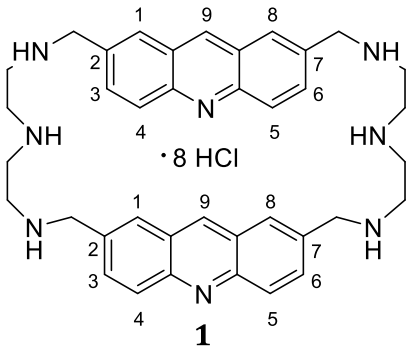



**6**

2,7-Dimethylacridine (337.4 mg, 1.63 mmol, 1 eq), selenium dioxide (397.9 mg, 3.59 mmol, 2.2 eq) and naphthalene (1.627 g, 12.7 mmol, 7.8 eq) were heated slowly to 200°C in a heat block and stirred. The reaction was monitored by TLC (100% DCM). When the starting material was exhausted (~5 hours), the mixture was dissolved in DCM and filtered through celite. The filtrate was reduced *in vacuo*, washed with hexane and reduced to dryness. Excess naphthalene was removed by sublimation at 200°C to give the pure yellow product **6** as an amorphous solid (319 mg, 83%).

**$^1$H NMR ($CDCl_3$) $\delta_H$ ppm:** 10.27 (2H, C**H**O, s), 9.17 (1H, Ar**H** 9, s), 8.60 (2H, Ar**H**, s), 8.35 (4H, Ar**H**, m). This is consistent with the literature.[136]

***2,5,8,21,24,27-Hexaaza[9,9](2,7)acridinophane (1 BisA)***

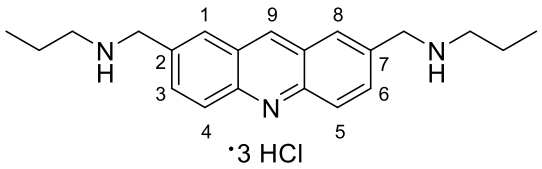



**1**

2,7-Acridinedicarboxaldehyde (100 mg, 0.426 mmol, 2 eq) was dissolved in a mixture of DCM and methanol (34.1 mL, 1:1 v/v) under nitrogen. Diethylenetriamine (45.9 µL, 43.88 mg, 0.426 mmol, 2 eq) in DCM and methanol (3.41 mL, 1:1 v/v) was added drop-wise and the reaction stirred at room temperature under nitrogen for 48 hours. The mixture was then reduced to dryness *in vacuo* at 25°C, washed with diethyl ether and redissolved in a solution of DCM and methanol (25.6 mL, 1:2 v/v). Sodium borohydride (48.3 mg, 1.278 mmol, 6 eq) was added to the solution at 0°C then stirred for 3 hours. The resultant mixture was reduced *in vacuo* and redissolved in water (7 mL) and extracted with a solution of DCM and methanol (4 × 10 mL 9:1 v/v). The combined organic extracts were dried and reduced. The residues were dissolved in water washed repeatedly with ether and DCM. 6 M hydrochloric acid (200 µL) was added and the solution reduced *in vacuo* to give the pure yellow green product as the 8 × HCl salt (53.1 mg, 28%).

**$^1$H NMR ($D_2O$) $\delta_H$ ppm:** 9.60 (2H, Ar**H**, s), 8.39 (4H, Ar**H**, s), 8.21 (4H, Ar**H**, d, *J* = 9.09 Hz), 8.09 (4H, Ar**H**, d, *J* = 9.09 Hz), 4.49 (8H, NHC**H**$_2$Ar, s), 3.40 (8H, NHC**H**$_2$, t, *J* = 6 Hz), 3.31 (8H, NHC**H**$_2$, t, *J* = 6 Hz). This compares well with the literature which was carried out on a lower sensitivity instrument **($D_2O$, 200 MHz):** 9.84 (2H, s), 8.54 (4H, s), 8.24 (8H, q), 4.57 (8H, s), 3.48 (16H, s).[136]

***2,7-(Di-n-propylaminomethyl)acridine (2 MonoA)***




**2**

2,7-Acridinedicarboxaldehyde (48 mg, 0.20 mmol, 1 eq) was dissolved in a mixture of DCM and methanol (62.5 mL, 1:1 v/v) and added dropwise to a solution of n-propylamine (0.62 g, 0.45 mL, 10.5 mmol) in methanol (6.2 mL) under argon and stirred at room temperature for 48 hours. The mixture was reduced *in vacuo* to give a yellow solid with some red/brown side products. The mixture was redissolved in methanol (100 mL), the yellow solid was soluble but the red/brown material was not soluble. Sodium borohydride (18.3 mg, 0.483 mmol, 2.3 eq) was added to the suspension at 0°C and the reaction stirred at 0°C for 30 minutes then for a further hour at room temperature until TLC showed the reaction to be complete. The mixture was reduced *in vacuo* to a yellow solid then re-dissolved in water (5 mL) and extracted with a mixture of DCM in methanol (40 mL 40:1 v/v). The organic extract was washed with water (3 × 5 mL) and the aqueous layer was extracted with DCM (3 × 15 mL). The combined organic extracts were dried over $MgSO_4$ and reduced *in vacuo*. The crude product was again re-dissolved in methanol and centrifuged to remove insoluble red material. The methanol supernatant was reduced *in vacuo* to give a yellow/brown oil. The product was washed with 1 M hydrochloric acid in methanol (5 mL) then reduced *in vacuo*. Half the **MonoA** salt (26.2 mg) was dissolved in saturated $NaHCO_3$ (2 mL) and extracted 8 times with equal volumes of DCM, dried over $MgSO_4$ and reduced *in vacuo* to give the pure yellow green solid which was re-made into the salt form by dissolving in 1 M hydrochloric acid in methanol (2 mL) and reducing to dryness *in vacuo* (24 mg, 0.056 mmol, 28%).

**$^1$H NMR ($D_2O$) $\delta_H$ ppm:** 9.59 (1H, Ar**H** 9, s), 8.42 (2H, Ar**H** 1 and 8, ap s), 8.28 (2H, Ar**H** 4 and 5, d, $J$ = 9.03 Hz), 8.12 (2H, Ar**H** 3 and 6, dd, $J_1$ = 9.03 Hz, $J_2$ = 1.83 Hz), 4.48 (4H, $NH_2C\mathbf{H}_2ArH$, s), 3.07 (4H, $CH_2C\mathbf{H}_2NH_2$, t, $J$ = 7.78 Hz), 1.69 (4H, $C\mathbf{H}_2CH_2NH_2$, m), 0.91 (6H, $C\mathbf{H}_3CH_2$, t, $J$ = 7.46 Hz). This compares well with the literature **($D_2O$, 200 MHz):** 9.78 (2H, s), 8.46 (2H, s), 8.24 (4H, q), 4.47 (4H, s), 3.03 (4H, t), 1.65 (4H, m), 0.87 (6H, t).[136]

### 6.2.2 FRET Melting Experiments

FRET melting experiments were carried out as described in section 6.1. hTeloC$_{FRET}$ oligonucleotide was used at pH 5.5 and 7.4 without annealing but at pH 6.8 was thermally annealed, hTeloG$_{FRET}$ and DS$_{FRET}$ oligonucleotide sequences were thermally annealed, all at a concentration of 400 nM. A stock solution of **BisA** was made up in water at a concentration of 10 mM and then diluted to the appropriate concentration in buffer by serial dilution. 10 μL of DNA solution were combined with 10 μL of **BisA** solution giving final concentrations of [DNA] = 200 nM and [**BisA**] = 0, 0.2, 0.4, 0.6, 0.8 and 1 μM.

### 6.2.3 Circular Dichroism Experiments

Circular dichroism experiments were carried out as described in section 6.1. 200 μL of 10 μM unlabelled hTeloC oligonucleotide was thermally annealed in either 10 mM sodium cacodylate buffer (pH 6.8), or used immediately without annealing in 10 mM sodium cacodylate 100 mM sodium chloride buffer (pH 7.4 and 5.5). 1 equivalent aliquots of **BisA** were added, up to a concentration of 100 μM using a Hamilton syringe and the samples were mixed with a pipette before being scanned from 400 nm to 220 nm.

### 6.2.4 Fluorescence Titration Experiments

Fluorescence titrations were carried out as described in section 6.1. **BisA** was diluted to a concentration of 1 μM in pH 5.5 buffer. The sample was excited at 255 nm and emission spectra taken from 265 to 500 nm with a slit width of 5 nm and a scan speed of 500 nm.min$^{-1}$. Unlabelled hTeloC DNA was titrated into the sample in 1 equivalent portions using a Hamilton syringe, the sample was mixed with a pipette and the change in fluorescence at the emission wavelength of 426 nm was measured.

### 6.2.5 FRET Titration Experiments

FRET titrations were carried out following the procedure outlined in section 6.1. Fluorescently labelled oligonucleotide hTeloC$_{FRET}$ was diluted to 100 nM and annealed

in pH 6.8 buffer containing 10 mM sodium cacodylate or diluted in pH 5.5 buffer containing 10 mM sodium cacodylate and 100 mM sodium chloride and used immediately without annealing. **BisA** was titrated into the sample in small aliquots up to a concentration of 2 μM using a Hamilton syringe. The sample was excited at 490 nm and the emission spectrum taken from 500 to 660 nm. This enabled the monitoring of the change in fluorescence of both FAM (515 nm) and TAMRA (580 nm).

### 6.2.6 UV Titration Experiments

UV titrations were carried out on a Hitachi U-3010 spectrophotometer as described in section 6.1. A matched cuvette containing buffer was used as a reference sample. Unlabelled hTeloC i-motif was diluted in 10 mM sodium cacodylate 100 mM sodium chloride buffer (pH 5.5) or thermally annealed in 10 mM sodium cacodylate buffer (pH 6.8) and diluted to a concentration of 2.5 μM and a volume of 200 μL. Aliquots of **BisA** were added with a Hamilton syringe and the sample mixed with a pipette. Initial titrations were carried out without centrifugation but in later titrations the sample was centrifuged at 6000 rpm before scanning to remove dust. The samples were scanned from 200 – 800 nm. The change in absorbance of DNA with increasing **BisA** concentration was plotted at 260 nm. An analogous titration of just **BisA** in buffer was carried out and the change in absorbance of **BisA** was plotted at 248 nm.

### 6.2.7 Polyacrylamide Gel Electrophoresis Experiments

hTeloC$_{FRET}$ labelled DNA was diluted to 1 μM in pH 5.5 10 mM sodium cacodylate 100 mM NaCl buffer without annealing. For titration gels, samples of i-motif DNA were combined with 0, 1, 5, 10, 15, 20, 25 and 30 equivalents of **BisA**, in lanes 2 to 9. Control gels were carried out showing the effect of increasing concentrations of **MonoA** and **DETA**. For these gels, samples of i-motif DNA were combined with 0, 2, 10, 20, 30, 40, 50 and 60 equivalents of **MonoA** or **DETA** in lanes 2 to 9. A comparison gel was carried out to directly compare the effect of **BisA** on i-motif, G-quadruplex and duplex DNA structures. For this gel i-motif, double stranded and G-quadruplex DNA sequences were combined with 0, 1 and 30 equivalents of **BisA** in lanes 2 to 10 respectively. The samples were transferred to a native 15% polyacrylamide gel and run at 100 Volts for 1 hour 15 minutes in pH 5.5 10 mM sodium cacodylate 100 mM sodium chloride buffer.

An 18 base single stranded sequence labelled with FAM was used as a marker in lane 1 for all the gels (table 6.1). Gels were made by combining 30% acrylamide/bis solution (5 mL) with water (4.5 mL), 10 x TBE pH 8.0 buffer (0.5 mL) and 10% ammonium persulfate (100 μL) before addition of tetramethylethylenediamine (4 μL) and being allowed to set for approximately 1 hour. The gels were imaged on a FujiFilm LAS-3000 dark box using 460 nm excitation and a Y515 filter with a ¼ second automatic exposure.

### 6.2.8 Dynamic Light Scattering Experiments

Dynamic light scattering measurements were carried out on a Malvern Zetasizer Nanoseries Nano-ZS instrument using a quartz cuvette with a 10 mm path length. For pH 5.5 experiments, unlabelled hTeloC DNA was diluted in pH 5.5 10 mM sodium cacodylate 100 mM NaCl buffer and used immediately. For pH 6.8 and pH 7.4 experiments, unlabelled hTeloC DNA was diluted in pH 6.8 10 mM sodium cacodylate or pH 7.4 10 mM sodium cacodylate 100 mM NaCl buffer and thermally annealed. In all cases, the DNA was diluted to 2.5 μM and the buffers, **BisA** solution and final sample were filtered through 0.22 μm syringe filters to remove any particulate matter. 1 equivalent aliquots of **BisA** were added with a Hamilton syringe and centrifuged at 6000 rpm for 3 minutes again to remove dust then measured. The sample was irradiated at 633 nm and the scattering was measured at an angle of 173°. The presented results are the average of 3 scans. The data was analysed using Malvern Zetasizer Software v6.34, Microsoft excel and Origin Pro 9.1.

## 6.3 Experimental for Chapter 3

### 6.3.1 Initial FRET melting screen

The initial FRET melting screen was performed in collaboration with Camille Huguin. hTeloC$_{FRET}$ labelled oligonucleotide was diluted to 220 nM and annealed in 10 mM sodium cacodylate buffer at pH 5.5 with 100 mM sodium chloride. The 960 compound library was provided as individual 1 mM stock solutions of each compound in DMSO. 2 µL of compound stock solution was combined with 18 µL of DNA to give a final concentration of [DNA] = 200 nM and [Ligand] = 100 µM. Control samples were prepared of just DNA in buffer and DMSO with DNA in buffer. The 960 compound library was tested 70 compounds at a time following the same procedure as described in section 6.1. The data was analysed using Roto-gene Q series software and Microsoft Excel. Initial Hits were repeated in 10 mM sodium cacodylate buffer at pH 5.5 with both 100 mM and 5 mM sodium chloride concentrations following the procedure outlined above and in section 6.1.

### 6.3.2 FRET dose response

Hit compounds from the initial screen were repeated against hTeloC$_{FRET}$ and compared to c-Myc$_{FRET}$ at pH 5.5, hTeloG$_{FRET}$ and DS$_{FRET}$ at pH 7.4 in 10 mM sodium cacodylate buffers with both 5 mM sodium chloride and 100 mM sodium chloride salt conditions. The experiments were carried out as previously described in section 6.1. Hit compounds were diluted in buffer by serial dilution and tested at concentrations of 1, 5, 10, 20, 40, 60, 80 and 100 µM. For the initial dose response study with hTeloC in 100 mM salt conditions, control samples were made with 0.1, 0.5, 1, 2, 4, 6, 8 and 10% DMSO. For later dose response studies a control of just 10% DMSO was used. The data was analysed using Roto-gene Q series software and Microsoft Excel.

### 6.3.3 Circular Dichroism

CD experiments were carried out as described in section 6.1. Unlabelled hTeloC oligonucleotide was diluted to a concentration of 10 µM in pH 5.5 10 mM sodium cacodylate 100 mM sodium chloride buffer. Hit compounds from the screen were

titrated into the sample using a Hamilton syringe, mixed with a pipette and were tested at concentrations of 10, 50, 100, 150 and 200 μM. If precipitation occurred at lower concentrations then higher concentrations were not tested. The samples were scanned from 220 to 400 nm.

### 6.3.4 Surface Plasmon Resonance (SPR) Experiments

SPR experiments were carried out on a Biacore T200 instrument from GE Healthcare using a series S streptavidin (SA) coated gold chip. Throughout all the SPR experiments a pH 5.5 running buffer was used containing 10 mM sodium cacodylate, 100 mM sodium chloride and 0.05% Tween 20, which was filtered and degassed before use. Biotinylated DNA hTeloC$_{Biotin}$, c-Myc$_{Biotin}$ sequences (Table 6.1) were diluted to 1 μM in running buffer. Biotinylated double stranded DNA was produced by combining a biotinylated oligonucleotide DS$_{Biotin}$ with its complimentary strand DScomp at 1 μM in pH 5.5 running buffer and thermally annealed. The chip was primed with running buffer and the detector was then normalised using the BIA normalising solution containing 70% (w/w) glycerol. The biotinylated oligonucleotides were bound to the streptavidin chip surface using the immobilisation wizard in the Biacore T200 software. This includes a conditioning of the chip surface with three 60 s washes of 1 M NaCl and 50 mM NaOH at a flow rate of 10 μL.min$^{-1}$ before injection of the DNA at the same flow rate. An approximate response of 500 RU was set as a target for immobilisation and flow cell 1 was left blank. hTeloC was injected over flow cell 2 (522.0 RU), c-Myc was injected over flow cell 3 (492.9 RU) and DS DNA was injected over flow cell 4 (507.1 RU). After immobilisation, the system (excluding the chip surface) was washed with a solution of 50% isopropyl alcohol in 50 mM NaOH and 1 M NaCl. **Mitoxantrone** and **tilorone** were tested for binding to each flow cell simultaneously using the affinity run wizard in the Biacore T200 software. Both compounds were diluted in running buffer from 10 mM stock solutions in water to concentrations of 100, 50, 25, 12.5, 6.25, 3.125, 1.56, 0.78, 0.39 and 0 μM by serial dilution. The compounds were tested by injecting each concentration starting with the lowest and finishing with the highest and measuring the response for each of the four flow cells. Samples were injected for 120 s at a flow rate of 30 μL.min$^{-1}$. After each injection the chip surface was regenerated by two 60 s injections of 1 M sodium chloride and the chip surface washed with running buffer for 60 s. The flow rate was

maintained at 30 μL.min$^{-1}$ throughout. Before the compounds were injected, 3 startup cycles were carried out, these are blank injections of just buffer, followed by 2 regeneration injections of 1 M NaCl. Each concentration was repeated with a second injection to ensure reproducibility and all experiments were carried out at 25°C. The response data was double referenced by subtracting the 3 startup cycles and injections of just buffer. Non selective binding to the chip surface was accounted for by subtracting the response from the blank flow cell. The average equilibrium response for each concentration was fitted using the affinity fit from the Biacore T200 evaluation software v2.0, assuming a 1:1 binding model. This data was then presented using Origin Pro 9.1. The $K_D$'s determined are the average of two experiments.

## 6.4 Experimental for Chapter 4

### 6.4.1 FRET Melting Experiments

FRET melting procedures were carried out in accordance with the procedure described in section 6.1. Silver nitrate and sodium nitrate stock solutions were made up at 1 M in water and were diluted in buffer containing 10 mM sodium cacodylate and 5 mM sodium chloride at either pH 5.5 or pH 7.4 by serial dilution. Silver nitrate and the control sodium nitrate were tested at concentrations of 0.5, 1, 5, 10, 25, 50 and 100 μM. For the cation screen, metal salts were also prepared in water at 1 M and diluted in the same buffer conditions. The library of salts were tested at concentrations of 10 μM, 100 μM, 1 mM and 100 mM. Dose response studies on $CuCl_2$ were carried out in 50 mM sodium cacodylate buffer at pH 5.5 and 7.4 at concentrations of 1, 5, 10, 20, 40, 60, 80 and 100 μM.

### 6.4.2 Circular Dichroism Experiments

CD titrations were carried out following the procedure described in section 6.1. Silver (I) ion experiments were carried out in 10 mM sodium cacodylate, 5 mM NaCl buffer at pH 7.4. Silver nitrate was titrated into the solution of 10 μM hTeloC by addition of 1 equivalent aliquots using a Hamilton syringe up to a concentration of 100 μM (10 eq). The sample was mixed thoroughly with a pipette and the CD spectrum taken from 220 to 400 nm. Copper (II) ion experiments were carried out in 50 mM sodium cacodylate buffer at pH 5.5 and 7.4. Copper (II) chloride was titrated into the sample in 10 equivalent aliquots up to a concentration of 1 mM (100 eq), mixed thoroughly and scanned from 200 to 400 nm.

### 6.4.3 Reversibility with Cysteine and EDTA

The reversibility of DNA folding with silver by addition of cysteine was carried out using CD as described in section 6.1. $AgNO_3$ was titrated into the sample in 1 equivalent steps by addition of 0.2 μL aliquots using a Hamilton syringe. After addition of 10 equivalents of $AgNO_3$, L-cysteine was titrated into the sample in 1 equivalent steps by

addition of 0.2 μL aliquots. The repeat reversibility of folding was shown by sequential addition of 10 equivalents of $AgNO_3$ (2 μL) and 10 equivalents of L-cysteine (2 μL) to 10 μM hTeloC in pH 7.4 buffer. The change in CD was monitored at 250 and 261 nm.

Analogous experiments were carried out to reverse the effects of $Cu^{2+}$ using ethylenediaminetetraacetic acid (EDTA). $CuCl_2$ was titrated into the sample in 10 equivalent steps by addition of 2 μL aliquots. After addition of 100 equivalents of $CuCl_2$, EDTA was titrated into the sample in 10 equivalent steps (0.2 μL) up to a concentration of 1.5 mM (150 eq). The repeat reversibility was demonstrated by sequential addition of 100 equivalents of $CuCl_2$ (20 μL) and 100 equivalents of EDTA (2 μL) at pH 5.5. The change in CD was monitored at 287 nm.

### 6.4.4 Continuous Variation Binding (Job's) Analysis

Continuous variation binding analysis was carried out using circular dichroism spectroscopy to measure the stoichiometry of the silver-DNA interaction. Samples of DNA and $AgNO_3$ were diluted in pH 7.4 buffer, keeping the total concentration of DNA + $AgNO_3$ constant at 20 μM but varying the individual concentrations of DNA and $AgNO_3$ between 0 and 20 μM, with the ratio of $Ag^+$:DNA changing in 1 equivalent steps. A second set of samples were prepared with the same concentrations of DNA but without $Ag^+$ to be used as a control and the molar ellipticity of both sets of data were measured. The spectra of the DNA only control samples were subtracted from the DNA + $Ag^+$ samples to give the change in molar ellipticity due to $Ag^+$. The change in CD signal at 275 nm was plotted versus the mole fraction of DNA to give a job plot. The inflection point of the job plot corresponds to the stoichiometry and was determined by a linear fit of the two parts of the graph and solving their equations simultaneously to find the intersection point using Microsoft Excel.

Repeat 1:

$$y = -4.4943x + 1.3168\,, \quad y = -0.2772x + 0.5142$$

$$\therefore (4.4943 - 0.2772)x = 1.3168 - 0.5142$$

$$\therefore x = \frac{0.8026}{4.2171} = 0.19$$

$$\therefore [DNA] = 20 \times 0.19 = 3.8$$

$$\therefore [Ag^+] = 20 - 3.8 = 16.2$$

$$\therefore Stoichiometry = \frac{16.2}{3.8} = 4.3$$

Repeat 2:

$$y = -2.9638x + 1.1213\,, \quad y = -0.7506x + 0.6695$$

$$\therefore (2.9638 - 0.7506)x = 1.1213 - 0.6695$$

$$\therefore x = \frac{0.4518}{2.2132} = 0.20$$

$$\therefore [DNA] = 20 \times 0.20 = 4$$

$$\therefore [Ag^+] = 20 - 4 = 16$$

$$\therefore Stoichiometry = \frac{16}{4} = 4$$

Average:

$$y = -3.0986x + 1.1419\,, \quad y = -0.5527x + 0.6222$$

$$\therefore (3.0986 - 0.5527)x = 1.1419 - 0.6222$$

$$\therefore x = \frac{0.5197}{2.5459} = 0.20$$

$$\therefore [DNA] = 20 \times 0.20 = 4$$

$$\therefore [Ag^+] = 20 - 4 = 16$$

$$\therefore Stoichiometry = \frac{16}{4} = 4$$

### 6.4.5 UV Titrations and Difference Experiments

UV titrations with $Ag^+$ were carried out on a Hitachi U-3010 spectrophotometer. Either hTeloC or C-hairpin unlabelled oligonucleotides were diluted to 2.5 μM in pH 7.4 10 mM sodium cacodylate 5 mM sodium chloride buffer and used immediately without annealing. Up to 6 equivalents of $AgNO_3$ was titrated into the sample in aliquots using a Hamilton syringe and the UV absorbance spectra measured from 200 – 400 nm at room temperature. An equivalent titration of $AgNO_3$ in buffer alone was also recorded. Both titrations were zero corrected at 400 nm and the absorbance spectra of the titration of $AgNO_3$ in buffer was subtracted from the spectra of the titration of $AgNO_3$ in DNA. The UV difference spectrum for silver induced folding was calculated by subtracting the folded spectrum (in the presence of 6 equivalents $AgNO_3$) from the

unfolded spectrum (DNA in the absence of $AgNO_3$). The resultant spectrum was zero corrected at 320 nm and normalised to give a maximum change in absorbance of 1 as described in the literature.[93] UV titrations with $Cu^{2+}$ were carried out on an Agilent Technologies Cary 60 UV-Vis spectrometer following the same procedure with the HIF-1α, PDGF-A, hTeloC and C-hairpin sequences. The data was analysed using Microsoft Excel and Origin Pro 9.1.

### 6.4.6 FRET Titration Experiments

Reverse FRET titrations using cysteine were carried out following the procedure outlined in section 6.1. Fluorescently labelled oligonucleotide $hTeloC_{FRET}$ was diluted to 100 nM in pH 7.4 buffer containing 10 mM sodium cacodylate and 5 mM sodium chloride and used immediately without annealing. 5 μM $AgNO_3$ was added immediately to fold the DNA before L-cysteine was titrated into the sample in small aliquots up to a concentration of 5 μM using a Hamilton syringe. The sample was excited at 490 nm and the emission spectrum taken from 500 to 650 nm. This enabled the monitoring of the change in fluorescence of both FAM (515 nm) and TAMRA (580 nm). The FRET efficiency ($E_{FRET}$) was calculated for each of 3 repeat experiments using equation 6.1. Where $FI_d^0$ is the fluorescence intensity of FAM in the absence of TAMRA. Repeat reversibility was shown by consecutive addition of 5 μM $AgNO_3$ and 5 μM cysteine.

*Equation 6.1:* $$E_{FRET} = 1 - \frac{FI_d}{FI_d^0}$$

FRET titrations were also carried out with $CuCl_2$ following the same procedure. $hTeloC_{FRET}$ was diluted to 100 nM in 50 mM sodium cacodylate buffer at pH 7.4 before $CuCl_2$ was titrated into the sample in small aliquots. First 1 equivalent then in 10 equivalent portions up to 100 eq and finally in 50 equivalent portions up to a final concentration of 500 eq (50 μM).

### 6.4.7 Biophysical NMR Experiments on Copper (II) ions with hTeloC

Biophysical $^1H$ NMR experiments were carried out in collaboration with Dr. Colin MacDonald using a Bruker Avance III 800 MHz spectrometer equipped with an HCN

inverse triple resonance z-gradient probe. The resonance signal of the solvent water was suppressed using the 1D Watergate sequence consisting of a symmetrical 3 – τ – 9 – τ – 19 pulse train inversion element with the interpulse delay time (τ) set to achieve a maximum in the imino proton region. The solvent resonance which was minimized was set on-resonance at the transmitter offset. The hTeloC oligonucleotide sequence was diluted to a concentration of 10 μM in pH 5.5 50 mM sodium cacodylate buffer containing 5% $D_2O$. The spectrum of hTeloC alone was measured over 1 hour after which 1 mM of $CuCl_2$ was added and the spectrum acquired over 2 hours. Finally 1 mM EDTA was added and the spectrum taken again for 1 hour. The spectra were acquired and processed using Bruker's TopSpin software package Bruker Biospin for NMR data analysis.

# Chapter 7: References

(1) Watson, J.; Crick, F. *Nature* **1953**, *171*, 737–738.

(2) Choi, J.; Majima, T. *Chem. Soc. Rev.* **2011**, *40*, 5893–5909.

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

(4) Mathad, R. I.; Hatzakis, E.; Dai, J.; Yang, D. *Nucleic Acids Res.* **2011**, *39*, 9023–9033.

(5) Neidle, S.; Balasubramanian, S. *Quadruplex Nucleic Acids*; RSC Publishing, Cambridge, 2006.

(6) Huppert, J. L. *FEBS J.* **2010**, *277*, 3452–3458.

(7) Biffi, G.; Tannahill, D.; McCafferty, J.; Balasubramanian, S. *Nat. Chem.* **2013**, *5*, 182–186.

(8) Rodriguez, R.; Miller, K. M.; Forment, J. V; Bradshaw, C. R.; Nikan, M.; Britton, S.; Oelschlaegel, T.; Xhemalce, B.; Balasubramanian, S.; Jackson, S. P. *Nat. Chem. Biol.* **2012**, *8*, 301–310.

(9) Henderson, A.; Wu, Y.; Huang, Y. C.; Chavez, E. A.; Platt, J.; Johnson, F. B.; Brosh, R. M.; Sen, D.; Lansdorp, P. M. *Nucleic Acids Res.* **2014**, *42*, 860–869.

(10) Balasubramanian, S.; Hurley, L. H.; Neidle, S. *Nat. Rev. Drug Discov.* **2011**, *10*, 261–275.

(11) Waller, Z. A. E.; Sewitz, S. a; Hsu, S.-T. D.; Balasubramanian, S. *J. Am. Chem. Soc.* **2009**, *131*, 12628–12633.

(12) González, V.; Guo, K.; Hurley, L.; Sun, D. *J. Biol. Chem.* **2009**, *284*, 23622–23635.

(13) Dexheimer, T. S.; Carey, S. S.; Zuohe, S.; Gokhale, V. M.; Hu, X.; Murata, L. B.; Maes, E. M.; Weichsel, A.; Sun, D.; Meuillet, E. J.; Montfort, W. R.; Hurley, L. H. *Mol Cancer Ther.* **2009**, *8*, 1363–1377.

(14) Brázda, V.; Hároníková, L.; Liao, J. C. C.; Fojta, M. *Int. J. Mol. Sci.* **2014**, *15*, 17493–17517.

(15) Gehring, K.; Leroy, J.-L.; Guéron, M. *Nature* **1993**, *363*, 561–565.

(16) Day, H. A.; Pavlou, P.; Waller, Z. A. E. *Bioorg. Med. Chem.* **2014**, *22*, 4407–4418.

(17) Modi, S.; Swetha, M. G.; Goswami, D.; Gupta, G. D.; Mayor, S.; Krishnan, Y. *Nat. Nanotechnol.* **2009**, *4*, 325–330.

(18) Surana, S.; Bhat, J. M.; Koushika, S. P.; Krishnan, Y. *Nat. Commun.* **2011**, *2*, article number 340.

(19) Elbaz, J.; Wang, Z.-G.; Orbach, R.; Willner, I. *Nano Lett.* **2009**, *9*, 4510–4514.

(20) Yang, Y.; Liu, G.; Liu, H.; Li, D.; Fan, C.; Liu, D. *Nano Lett.* **2010**, *10*, 1393–1397.

(21) Li, T.; Ackermann, D.; Hall, A. M.; Famulok, M. *J. Am. Chem. Soc.* **2012**, *134*, 3508–3516.

(22) Liu, D.; Balasubramanian, S. *Angew. Chem. Int. Ed. Engl.* **2003**, *42*, 5734–5736.

(23) Wang, Z.-G.; Elbaz, J.; Willner, I. *Nano Lett.* **2011**, *11*, 304–309.

(24) Wang, W.; Liu, H.; Liu, D.; Xu, Y.; Yang, Y.; Zhou, D. *Langmuir* **2007**, *23*, 11956–11959.

(25) Sharma, J.; Chhabra, R.; Yan, H.; Liu, Y. *Chem. Commun.* **2007**, 477–479.

(26) Marsh, R. E.; Bierstedt, R.; Eichhorn, E. L. *Acta Cryst.* **1962**, *15*, 310–316.

(27) Lyamichev, V. I.; Mirkin, S. M.; Danilevskaya, O. N.; Voloshin, O. N.; Balatskaya, S. V.; Dobrynin, V. N.; Filippov, S. A.; Frank-Kamenetskii, M. D. *Nature* **1989**, *339*, 634–637.

(28) Catasti, P.; Chen, X.; Deaven, L. L.; Moyzis, R. K.; Bradbury, E. M.; Gupta, G. *J. Mol. Biol.* **1997**, *272*, 369–382.

(29) Kang, C.; Berger, I.; Lockshin, C.; Ratliff, R.; Moyzis, R.; Rich, A. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 3874–3878.

(30) Cai, L.; Chen, L.; Raghavan, S.; Ratliff, R.; Moyzis, R.; Rich, A. *Nucleic Acids Res.* **1998**, *26*, 4696–4705.

(31) Chen, L.; Cai, L.; Zhang, X.; Rich, A. *Biochemistry* **1994**, *33*, 13540–13546.

(32) Weil, J.; Min, T.; Yang, C.; Wang, S.; Sutherland, C.; Sinha, N.; Kang, C. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1999**, *55*, 422–429.

(33) Kang, C. H.; Berger, I.; Lockshin, C.; Ratliff, R.; Moyzis, R.; Rich, A. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 11636–11640.

(34) Han, X.; Leroy, J.; Guéron, M. *J. Mol. Biol.* **1998**, *278*, 949–965.

(35) Phan, A. T.; Guéron, M.; Leroy, J. L. *J. Mol. Biol.* **2000**, *299*, 123–144.

(36) Kanaori, K.; Maeda, A.; Kanehara, H.; Tajima, K.; Makino, K. *Biochemistry* **1998**, *37*, 12979–12986.

(37) Guéron, M.; Leroy, J. L. *Curr. Opin. Struct. Biol.* **2000**, *10*, 326–331.

(38) Nonin-Lecomte, S.; Leroy, J. L. *J. Mol. Biol.* **2001**, *309*, 491–506.

(39) Kanaori, K.; Shibayama, N.; Gohda, K.; Tajima, K.; Makino, K. *Nucleic Acids Res.* **2001**, *29*, 831–840.

(40) Brooks, T. A.; Kendrick, S.; Hurley, L. *FEBS J.* **2010**, *277*, 3459–3469.

(41) Brazier, J. A.; Shah, A.; Brown, G. D. *Chem. Commun.* **2012**, *48*, 10739–10741.

(42) Zhou, J.; Wei, C.; Jia, G.; Wang, X.; Feng, Z.; Li, C. *Mol. Biosyst.* **2010**, *6*, 580–586.

(43) Miyoshi, D.; Sugimoto, N. *Biochimie* **2008**, *90*, 1040–1051.

(44) Miyoshi, D.; Matsumura, S.; Nakano, S.-I.; Sugimoto, N. *J. Am. Chem. Soc.* **2004**, *126*, 165–169.

(45) Cui, J.; Waltman, P.; Le, V. H.; Lewis, E. A. *Molecules* **2013**, *18*, 12751–12767.

(46) Bhavsar-jog, Y. P.; Dornshuld, E. Van; Brooks, T. A.; Tschumper, G. S.; Wadkins, R. M. *Biochemistry* **2014**, *53*, 1586–1594.

(47) Rajendran, A.; Nakano, S.; Sugimoto, N. *Chem. Commun.* **2010**, *46*, 1299–1301.

(48) König, S. L. B.; Huppert, J. L.; Sigel, R. K. O.; Evans, A. C. *Nucleic Acids Res.* **2013**, *41*, 7453–7461.

(49) Sun, D.; Hurley, L. H. *J. Med. Chem.* **2009**, *52*, 2863–2874.

(50) Saxena, S.; Bansal, A.; Kukreti, S. *Arch. Biochem. Biophys.* **2008**, *471*, 95–108.

(51) Mergny, J.; Lacroix, J. L.; Han, X.; Leroy, J.; Helene, C. *J. Am. Chem. Soc.* **1995**, *117*, 8887–8898.

(52) Kim, S. E.; Lee, I.-B.; Hyeon, C.; Hong, S.-C. *J. Phys. Chem. B* **2014**, *118*, 4753–4760.

(53) Dai, J.; Carver, M.; Yang, D. *Biochimie* **2008**, *90*, 1172–1183.

(54) Mergny, J.-L.; Riou, J.-F.; Mailliet, P.; Teulade-Fichou, M.-P.; Gilson, E. *Nucleic Acids Res.* **2002**, *30*, 839–865.

(55) Kim, N. W.; Piatyszek, M. A.; Prowse, K. R.; Harley, C. B.; West, M. D.; Ho, P. L. C.; Coviello, G. M.; Wright, W. E.; Weinrich, S. L.; Shay, J. W. *Science.* **1994**, *266*, 2011–2015.

(56) Sun, D.; Thompson, B.; Cathers, B. E.; Salazar, M.; Kerwin, S. M.; Trent, J. O.; Jenkins, T. C.; Neidle, S.; Hurley, L. H. *J. Med. Chem.* **1997**, *40*, 2113–2116.

(57) Maji, B.; Bhattacharya, S. *Chem. Commun.* **2014**, *50*, 6422–6438.

(58) Leroy, J. L.; Guéron, M.; Mergny, J. L.; Hélène, C. *Nucleic Acids Res.* **1994**, *22*, 1600–1606.

(59) Chen, Y.; Qu, K.; Zhao, C.; Wu, L.; Ren, J.; Wang, J.; Qu, X. *Nat. Commun.* **2012**, *3*, article number 1074.

(60) Marsich, E.; Piccini, A.; Xodo, L. E.; Manzini, G. *Nucleic Acids Res.* **1996**, *24*, 4029–4033.

(61) Lacroix, L.; Liénard, H.; Labourier, E.; Djavaheri-Mergny, M.; Lacoste, J.; Leffers, H.; Tazi, J.; Hélène, C.; Mergny, J. L. *Nucleic Acids Res.* **2000**, *28*, 1564–1575.

(62) Simonsson, T.; Pribylova, M.; Vorlickova, M. *Biochem. Biophys. Res. Commun.* **2000**, *278*, 158–166.

(63) González, V.; Hurley, L. H. *Annu. Rev. Pharmacol. Toxicol.* **2010**, *50*, 111–129.

(64) Berberich, S. J.; Postel, E. H. *Oncogene* **1995**, *10*, 2343–2347.

(65) Kendrick, S.; Akiyama, Y.; Hecht, S. M.; Hurley, L. H. *J. Am. Chem. Soc.* **2009**, *131*, 17667–17676.

(66) Vaux, D. L.; Cory, S.; Adams, J. M. *Nature* **1988**, *335*, 440–442.

(67) Joensuu, H.; Pyikkanen, L.; Toikkanen, S. *Am. J. Pathol.* **1994**, *145*, 1191–1198.

(68) Bar-Am, O.; Weinreb, O.; Amit, T.; Youdim, M. B. H. *FASEB J.* **2005**.

(69) Kendrick, S.; Kang, H.-J.; Alam, M. P.; Madathil, M. M.; Agrawal, P.; Gokhale, V.; Yang, D.; Hecht, S. M.; Hurley, L. H. *J. Am. Chem. Soc.* **2014**, *136*, 4161–4171.

(70) Kang, H.-J.; Kendrick, S.; Hecht, S. M.; Hurley, L. H. *J. Am. Chem. Soc.* **2014**, *136*, 4172–4185.

(71) Xu, Y.; Sugiyama, H. *Nucleic Acids Res.* **2005**, *49*, 177–178.

(72) Xu, Y.; Sugiyama, H. *Nucleic Acids Res.* **2006**, *34*, 949–954.

(73) Guo, K.; Gokhale, V.; Hurley, L. H.; Sun, D. *Nucleic Acids Res.* **2008**, *36*, 4598–4608.

(74) Manzini, G.; Yathindra, N.; Xodo, L. E. *Nucleic Acids Res.* **1994**, *22*, 4634–4640.

(75) Bucek, P.; Jaumot, J.; Aviñó, A.; Eritja, R.; Gargallo, R. *Chem. Eur. J.* **2009**, *15*, 12663–12671.

(76) Bucek, P.; Gargallo, R.; Kudrev, A. *Anal. Chim. Acta* **2010**, *683*, 69–77.

(77) Benabou, S.; Ferreira, R.; Aviñó, A.; González, C.; Lyonnais, S.; Solà, M.; Eritja, R.; Jaumot, J.; Gargallo, R. *Biochim. Biophys. Acta* **2014**, *1840*, 41–52.

(78) Guo, K.; Pourpak, A.; Beetz-Rogers, K.; Gokhale, V.; Sun, D.; Hurley, L. H. *J. Am. Chem. Soc.* **2007**, *129*, 10220–10228.

(79) Jolad, V. V; Murad, F. K.; Arnold, J. R. P.; Fisher, J. *Org. Biomol. Chem.* **2005**, *3*, 2234–2236.

(80) Dhakal, S.; Yu, Z.; Konik, R.; Cui, Y.; Koirala, D.; Mao, H. *Biophys. J.* **2012**, *102*, 2575–2584.

(81) Fojtík, P.; Vorlícková, M. *Nucleic Acids Res.* **2001**, *29*, 4684–4690.

(82) Huppert, J. L.; Balasubramanian, S. *Nucleic Acids Res.* **2005**, *33*, 2908–2916.

(83) Todd, A. K.; Johnston, M.; Neidle, S. *Nucleic Acids Res.* **2005**, *33*, 2901–2907.

(84) Huppert, J. L.; Balasubramanian, S. *Nucleic Acids Res.* **2007**, *35*, 406–413.

(85) Cornuel, J. F.; Moraillon, A.; Guéron, M. *Biochimie* **2002**, *84*, 279–289.

(86) Cui, Y.; Koirala, D.; Kang, H.; Dhakal, S.; Yangyuoru, P.; Hurley, L. H.; Mao, H. *Nucleic Acids Res.* **2014**, 1–10.

(87) Backe, P. H.; Messias, A. C.; Ravelli, R. B. G.; Sattler, M.; Cusack, S. *Structure* **2005**, *13*, 1055–1067.

(88) Banerjee, K.; Wang, M.; Cai, E.; Fujiwara, N.; Baker, H.; Cave, J. W. *Nat. Commun.* **2014**, *5*, Article No. 5769.

(89) Monchaud, D.; Teulade-Fichou, M.-P. *Org. Biomol. Chem.* **2008**, *6*, 627–636.

(90) Murat, P.; Singh, Y.; Defrancq, E. *Chem. Soc. Rev.* **2011**, *40*, 5293–5307.

(91) Liebert, M. A.; Mergny, J.; Lacroix, L. *Oligonucleotides* **2003**, *13*, 515–537.

(92) Fedoroff, O. Y.; Rangan, A.; Chemeris, V. V; Hurley, L. H. *Biochemistry* **2000**, *39*, 15083–15090.

(93) Mergny, J.-L.; Li, J.; Lacroix, L.; Amrane, S.; Chaires, J. B. *Nucleic Acids Res.* **2005**, *33*, e138.

(94) Campbell, I. D. *Biophysical Techniques*; 1st ed.; Oxford University Press, 2012; pp. 185-199

(95) Chirio-Lebrun, M.-C.; Prats, M. *Biochem. Educ.* **1998**, *26*, 320–323.

(96) Mergny, J. L.; Maurizot, J. C. *Chembiochem* **2001**, *2*, 124–132.

(97) Guyen, B.; Schultes, C. M.; Hazel, P.; Mann, J.; Neidle, S. *Org. Biomol. Chem.* **2004**, *2*, 981–988.

(98) Mergny, J. L. *Biochemistry* **1999**, *38*, 1573–1581.

(99) Rahman, K. M.; Tizkova, K.; Reszka, A. P.; Neidle, S.; Thurston, D. E. *Bioorganic Med. Chem. Lett.* **2012**, *22*, 3006–3010.

(100) Alberti, P.; Ren, J.; Teulade-Fichou, M.-P.; Guittat, L.; Riou, J.-F.; Chaires, J. B.; Hélène, C.; Vigneron, J.-P.; Lehn, J.-M.; Mergny, J.-L. *J. Biomol. Struct. Dyn.* **2001**, *19*, 505–513.

(101) Campbell, I. D. *Biophysical Techniques*; 1st ed.; Oxford University Press, 2012; pp. 176-182

(102) Kypr, J.; Kejnovská, I.; Renciuk, D.; Vorlícková, M. *Nucleic Acids Res.* **2009**, *37*, 1713–1725.

(103) Campbell, I. D. *Biophysical Techniques*; 1st ed.; Oxford University Press, 2012; pp. 111-116

(104) Redman, J. E. *Methods* **2007**, *43*, 302–312.

(105) Zeng, Z.-X.; Zhao, Y.; Hao, Y.-H.; Tan, Z. *J. Mol. Recognit.* **2005**, *18*, 267–271.

(106) Xia, H.; Hou, Y.; Ngai, T.; Zhang, G. *J. Phys. Chem. B* **2010**, *114*, 775–779.

(107) Li, X.; Peng, Y.; Ren, J.; Qu, X. *Proc. Natl. Acad. Sci.* **2006**, *103*, 19658–19663.

(108) Zhao, C.; Song, Y.; Ren, J.; Qu, X. *Biomaterials* **2009**, *30*, 1739–1745.

(109) Benabou, S.; Aviñó, a.; Eritja, R.; González, C.; Gargallo, R. *RSC Adv.* **2014**, *4*, 26956.

(110) Hellman, L. M.; Fried, M. G. *Nat. Protoc.* **2007**, *2*, 1849–1861.

(111) Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 11593–11598.

(112) Grand, C. L.; Han, H.; Muñoz, R. M.; Weitman, S.; Von Hoff, D. D.; Hurley, L. H.; Bearss, D. J. *Mol. Cancer Ther.* **2002**, *1*, 565–573.

(113) Martino, L.; Pagano, B.; Fotticchia, I.; Neidle, S.; Giancola, C. *J. Phys. Chem. B* **2009**, *113*, 14779–14786.

(114) Fernández, S.; Eritja, R.; Aviñó, A.; Jaumot, J.; Gargallo, R. *Int. J. Biol. Macromol.* **2011**, *49*, 729–736.

(115) Li, X.; Peng, Y.; Qu, X. *Nucleic Acids Res.* **2006**, *34*, 3670–3676.

(116) Peng, Y.; Li, X.; Ren, J.; Qu, X. *Chem. Commun.* **2007**, 5176–5178.

(117) Zhao, C.; Ren, J.; Qu, X. *Chem. Eur. J.* **2008**, *14*, 5435–5439.

(118) Balasubramanian, S.; Neidle, S. *Curr. Opin. Chem. Biol.* **2009**, *13*, 345–353.

(119) Chen, X.; Zhou, X.; Han, T.; Wu, J.; Zhang, J.; Guo, S. *ACS Nano* **2013**, *7*, 531–537.

(120) Zhou, X.; Zhang, Y.; Wang, C.; Wu, X.; Yang, Y.; Zheng, B.; Wu, H.; Guo, S.; Zhang, J. *ACS Nano* **2012**, *6*, 6592–6599.

(121) Cohen, G. M. *Biochem. J.* **1997**, *326*, 1–16.

(122) Wang, L.; Wu, Y.; Chen, T.; Wei, C. *Int. J. Biol. Macromol.* **2013**, *52*, 1–8.

(123) Xue, L.; Ranjan, N.; Arya, D. P. *Biochemistry* **2011**, *50*, 2838–2849.

(124) Kong, D.-M.; Ma, Y.-E.; Wu, J.; Shen, H.-X. *Chem. Eur. J.* **2009**, *15*, 901–909.

(125) Ma, D.-L.; Kwan, M. H.-T.; Chan, D. S.-H.; Lee, P.; Yang, H.; Ma, V. P.-Y.; Bai, L.-P.; Jiang, Z.-H.; Leung, C.-H. *Analyst* **2011**, *136*, 2692–2696.

(126) Zhou, J.; Amrane, S.; Korkut, D. N.; Bourdoncle, A.; He, H.-Z.; Ma, D.-L.; Mergny, J.-L. *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 7742–7746.

(127) Zhang, X. Y.; Luo, H. Q.; Li, N. B. *Anal. Biochem.* **2014**, *455*, 55–59.

(128) Xu, H.; Zhang, H.; Qu, X. *J. Inorg. Biochem.* **2006**, *100*, 1646–1652.

(129) Shi, S.; Geng, X.; Zhao, J.; Yao, T.; Wang, C.; Yang, D.; Zheng, L.; Ji, L. *Biochimie* **2010**, *92*, 370–377.

(130) Shi, S.; Zhao, J.; Geng, X.; Yao, T.; Huang, H.; Liu, T.; Zheng, L.; Li, Z.; Yang, D.; Ji, L. *Dalt. Trans.* **2010**, *39*, 2490–2493.

(131) Lee, I. J.; Yi, J. W.; Kim, B. H. *Chem. Commun.* **2009**, 5383–5385.

(132) Lee, I. J.; Kim, B. H. *Chem. Commun.* **2012**, *48*, 2074–2076.

(133) Park, J. W.; Seo, Y. J.; Kim, B. H. *Chem. Commun.* **2014**, *50*, 52–54.

(134) El-Sayed, A. A.; Pedersen, E. B.; Khaireldin, N. A. *Nucleosides. Nucleotides Nucleic Acids* **2012**, *31*, 872–879.

(135) Dembska, A.; Rzepecka, P.; Juskowiak, B. *J. Fluoresc.* **2013**, *23*, 807–812.

(136) Teulade-Fichou, M.-P.; Vigneron, J.; Lehn, J.-M. *Supramol. Chem.* **1995**, *5*, 139–147.

(137) Bahr, M.; Gabelica, V.; Granzhan, A.; Teulade-Fichou, M.-P.; Weinhold, E. *Nucleic Acids Res.* **2008**, *36*, 5000–5012.

(138) Slama-Schwok, A.; Teulade-Fichou, M.-P.; Vigneron, J.-P.; Taillandier, E.; Lehn, J.-M. *J. Am. Chem. Soc.* **1995**, *117*, 6822–6830.

(139) Slama-Schwok, A.; Peronnet, F.; Hantz-Brachet, E.; Taillandier, E.; Teulade-Fichou, M.-P.; Vigneron, J.-P.; Best-Belpomme, M.; Lehn, J.-M. *Nucleic Acids Res.* **1997**, *25*, 2574–2581.

(140) Blacker, J.; Teulade-Fichou, M.-P.; Vigneron, J.-P.; Fauquet, M.; Lehn, J.-M. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 601–606.

(141) Berthet, N.; Michon, J.; Lhomme, J.; Teulade-fichou, M.; Vigneron, J. P.; Lehn, J.-M. *Chem. Eur. J.* **1999**, *5*, 3625–3630.

(142) Jourdan, M.; Garcia, J.; Lhomme, J.; Teulade-Fichou, M. P.; Vigneron, J. P.; Lehn, J. M. *Biochemistry* **1999**, *38*, 14205–14213.

(143) Teulade-Fichou, M. P.; Fauquet, M.; Baudoin, O.; Vigneron, J. P.; Lehn, J. M. *Bioorg. Med. Chem.* **2000**, *8*, 215–222.

(144) David, A.; Bleimling, N.; Beuck, C.; Lehn, J.-M.; Weinhold, E.; Teulade-Fichou, M.-P. *Chembiochem* **2003**, *4*, 1326–1331.

(145) Bahr, M.; Gabélica, V.; Granzhan, A.; Weinhold, E.; Teulade-Fichou, M.-P. *Nucleic Acids Symp. Ser. (Oxf).* **2008**, 109–110.

(146) Ren, J.; Chaires, J. B. *Biochemistry* **1999**, *38*, 16067–16075.

(147) Krupp, G.; Khne, K.; Tamm, S.; Klapper, W.; Heidorn, K.; Rott, A.; Parwaresch, R. *Nucleic Acids Res.* **1997**, *25*, 919–921.

(148) Fajkus, J. *Clin. Chim. Acta.* **2006**, *371*, 25–31.

(149) Desbois, N.; Szollosi, A.; Maisonial, A.; Weber, V.; Moreau, E.; Teulade, J.-C.; Chavignon, O.; Blache, Y.; Chezal, J. M. *Tetrahedron Lett.* **2009**, *50*, 6894–6896.

(150) Sperotto, E.; van Klink, G. P. M.; van Koten, G.; de Vries, J. G. *Dalt. Trans.* **2010**, *39*, 10338–10351.

(151) Jones, G. O.; Liu, P.; Houk, K. N.; Buchwald, S. L. *J. Am. Chem. Soc.* **2010**, *132*, 6205–6213.

(152) Atwell, G. J.; Cain, B. F.; Baguley, B. C.; Finlay, G. J.; Denny, W. A. *J. Med. Chem.* **1984**, *27*, 1481–1485.

(153) Doddrell, D. .; Pegg, D. .; Bendall, M. . *J. Magn. Reson.* **1982**, *48*, 323–327.

(154) Howell, L. a; Howman, A.; O'Connell, M. a; Mueller, A.; Searcey, M. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 5880–5883.

(155) Campbell, I. D. *Biophysical Techniques*; 1st ed.; Oxford University Press, 2012; pp. 11-14

(156) Weiss, J. N. *FASEB J.* **1997**, *11*, 835–841.

(157) Maniatis, T.; Jeffrey, A.; Van deSande, H. *Biochemistry* **1975**, *14*, 3787–3794.

(158) Nobbmann, U.; Connah, M.; Fish, B.; Varley, P.; Gee, C.; Mulot, S.; Chen, J.; Zhou, L.; Lu, Y.; Sheng, F.; Yi, J.; Harding, S. E. *Biotechnol. Genet. Eng. Rev.* **2007**, *24*, 117–128.

(159) Rajski, S. R.; Williams, R. M. *Chem. Rev.* **1998**, *98*, 2723–2796.

(160) Malo, N.; Hanley, J. A.; Cerquozzi, S.; Pelletier, J.; Nadon, R. *Nat. Biotechnol.* **2006**, *24*, 167–175.

(161) Diller, D. J.; Hobbs, D. W. *J. Med. Chem.* **2004**, *47*, 6373–6383.

(162) Munyuki, G.; Jackson, G. E.; Venter, G. a; Kövér, K. E.; Szilágyi, L.; Rautenbach, M.; Spathelf, B. M.; Bhattacharya, B.; van der Spoel, D. *Biochemistry* **2013**, *52*, 7798–7806.

(163) Nishimura, T.; Okobira, T.; Kelly, A. M.; Shimada, N.; Takeda, Y.; Sakurai, K. *Biochemistry* **2007**, *46*, 8156–8163.

(164) Reed, J. E.; Neidle, S.; Vilar, R. *Chem. Commun.* **2007**, 4366–4368.

(165) Dash, J.; Shirude, P. S.; Hsu, S. T. D.; Balasubramanian, S. *J. Am. Chem. Soc.* **2008**, *130*, 15950–15956.

(166) Smith, P. J.; Morgan, S. A.; Fox, M. E.; Watson, J. V. *Biochem. Pharmacol.* **1990**, *40*, 2069–2078.

(167) Precigoux, G.; Courseille, C.; Geoffre, S.; Hospital, M. *Acta Crystallogr. Sect. B* **1979**, *35*, 3070–3072.

(168) Green, N. M. *Adv. Protein Chem.* **1975**, *29*, 85–133.

(169) Rich, R. L.; Myszka, D. G. *J. Mol. Recognit.* **2010**, *23*, 1–64.

(170) Parker, B. S.; Buley, T.; Evison, B. J.; Cutts, S. M.; Neumann, G. M.; Iskander, M. N.; Phillips, D. R. *J. Biol. Chem.* **2004**, *279*, 18814–18823.

(171) Wu, C.-C.; Li, Y.-C.; Wang, Y.-R.; Li, T.-K.; Chan, N.-L. *Nucleic Acids Res.* **2013**, *41*, 10630–10640.

(172) Huang, H.-S.; Chen, I.-B.; Huang, K.-F.; Lu, W.-C.; Shieh, F.-Y.; Huang, Y.-Y.; Huang, F.-C.; Lin, J.-J. *Chem. Pharm. Bull.* **2007**, *55*, 284–292.

(173) Stringfellow, D. A.; Glasgow, L. A. *Antimicrob. Agents Chemother.* **1972**, *2*, 73–78.

(174) Krueger, R. F.; Mayer, G. D. *Science.* **1970**, *169*, 1213–1214.

(175) Zhou, D.; Tuo, W.; Hu, H.; Xu, J.; Chen, H.; Rao, Z.; Xiao, Y.; Hu, X.; Liu, P. *Eur. J. Med. Chem.* **2013**, *64*, 432–441.

(176) Lee, C.-C.; Chang, D.-M.; Huang, K.-F.; Chen, C.-L.; Chen, T.-C.; Lo, Y.; Guh, J.-H.; Huang, H.-S. *Bioorg. Med. Chem.* **2013**, *21*, 7125–7133.

(177) Sims, S. H.; Cha, Y.; Romine, M. F.; Gao, P.; Gottlieb, K.; Deisseroth, A. B. *Mol. Cell. Biol.* **1993**, *13*, 690–702.

(178) Harada, H.; Takahashi, E.; Itoh, S.; Harada, K.; Hori, T. A.; Taniguchi, T. *Mol. Cell. Biol.* **1994**, *14*, 1500–1509.

(179) Bohg, A.; Ristow, H. *Eur. J. Biochem.* **1987**, *170*, 253–258.

(180) Bohg, A.; Ristow, H. *Eur. J. Biochem.* **1986**, *160*, 587–591.

(181) Johannsen, S.; Megger, N.; Böhme, D.; Sigel, R. K. O.; Müller, J. *Nat. Chem.* **2010**, *2*, 229–234.

(182) Yamaguchi, H.; Sebera, J.; Kondo, J.; Oda, S.; Komuro, T.; Kawamura, T.; Dairaku, T.; Kondo, Y.; Okamoto, I.; Ono, A.; Burda, J. V; Kojima, C.; Sychrovský, V.; Tanaka, Y. *Nucleic Acids Res.* **2014**, *42*, 4094–4099.

(183) Clever, G. H.; Kaul, C.; Carell, T. *Angew. Chem. Int. Ed. Engl.* **2007**, *46*, 6226–6236.

(184) Clever, G. H.; Shionoya, M. *Coord. Chem. Rev.* **2010**, *254*, 2391–2402.

(185) Müller, J. *Eur. J. Inorg. Chem.* **2008**, *2008*, 3749–3763.

(186) Miyake, Y.; Togashi, H.; Tashiro, M.; Yamaguchi, H.; Oda, S.; Kudo, M.; Tanaka, Y.; Kondo, Y.; Sawa, R.; Fujimoto, T.; Machinami, T.; Ono, A. *J. Am. Chem. Soc.* **2006**, *128*, 2172–2173.

(187) Tanaka, Y.; Oda, S.; Yamaguchi, H.; Kondo, Y.; Kojima, C.; Ono, A. *J. Am. Chem. Soc.* **2007**, *129*, 244–245.

(188) Kondo, J.; Yamada, T.; Hirose, C.; Okamoto, I.; Tanaka, Y.; Ono, A. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 2385–2388.

(189) Urata, H.; Yamaguchi, E.; Funai, T.; Matsumura, Y.; Wada, S.-I. *Angew. Chem. Int. Ed. Engl.* **2010**, *49*, 6516–6519.

(190) Ono, A.; Cao, S.; Togashi, H.; Tashiro, M.; Fujimoto, T.; Machinami, T.; Oda, S.; Miyake, Y.; Okamoto, I.; Tanaka, Y. *Chem. Commun.* **2008**, 4825–4827.

(191) Ono, A.; Torigoe, H.; Tanaka, Y.; Okamoto, I. *Chem. Soc. Rev.* **2011**, *40*, 5855–5866.

(192) Funai, T.; Miyazaki, Y.; Aotani, M.; Yamaguchi, E.; Nakagawa, O.; Wada, S.; Torigoe, H.; Ono, A.; Urata, H. *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 6464–6466.

(193) Xie, W. Y.; Huang, W. T.; Li, N. B.; Luo, H. Q. *Analyst* **2011**, *136*, 4130–4133.

(194) Xie, W. Y.; Huang, W. T.; Li, N. B.; Luo, H. Q. *Chem. Commun.* **2012**, *48*, 82–84.

(195) Sengupta, B.; Springer, K.; Buckman, J. G.; Story, S. P.; Abe, O. H.; Hasan, Z. W.; Prudowsky, Z. D.; Rudisill, S. E.; Degtyareva, N. N.; Petty, J. T. *J. Phys. Chem. C* **2009**, *113*, 19518–19524.

(196) Fu, Y.; Zhang, J.; Chen, X.; Huang, T.; Duan, X.; Li, W.; Wang, J. *J. Phys. Chem. C* **2011**, *115*, 10370–10379.

(197) Wu, J.; Fu, Y.; He, Z.; Han, Y.; Zheng, L.; Zhang, J.; Li, W. *J. Phys. Chem. B* **2012**, *116*, 1655–1665.

(198) Li, W.; Liu, L.; Fu, Y.; Sun, Y.; Zhang, J.; Zhang, R. *Photochem. Photobiol. Sci.* **2013**, *12*, 1864–1872.

(199) Day, H. A.; Huguin, C.; Waller, Z. A. E. *Chem. Commun.* **2013**, *49*, 7696–7698.

(200) Yamaguchi, T.; Lindqvist, O.; Boyce, J. B.; Claeson, T. *Acta Chem. Scand. A* **1984**, *38*, 423–428.

(201) Jenkins, T. C. *Methods Mol. Biol.* **1997**, *90*, 195–218.

(202) Tang, C.-X.; Bu, N.-N.; He, X.-W.; Yin, X.-B. *Chem. Commun.* **2011**, *47*, 12304–12306.

(203) Dong, Y.; Yang, Z.; Liu, D. *Acc. Chem. Res.* **2014**, *47*, 1853–1860.

(204) Irving, H.; Williams, R. J. P. *J. Chem. Soc.* **1953**, 3192–3210.

(205) Tottey, S.; Waldron, K. J.; Firbank, S. J.; Reale, B.; Bessant, C.; Sato, K.; Cheek, T. R.; Gray, J.; Banfield, M. J.; Dennison, C.; Robinson, N. J. *Nature* **2008**, *455*, 1138–1142.

(206) Bomholt, N.; Filichev, V. V; Pedersen, E. B. *Org. Biomol. Chem.* **2011**, *9*, 4527–4534.

(207) Goncharova, I. *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.* **2014**, *118*, 221–227.

(208) Berdakin, M.; Steinmetz, V.; Maitre, P.; Pino, G. A. *J. Phys. Chem. A* **2014**, *118*, 3804–3809.

(209) Berdakin, M.; Feraud, G.; Dedonder-Lardeux, C.; Jouvet, C.; Pino, G. A. *J. Phys. Chem. Lett.* **2014**, *5*, 2295–2301.

(210) Zhao, T.-T.; Chen, Q.-Y.; Yang, H. *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.* **2014**, *137C*, 66–69.

(211) Shi, Y.; Sun, H.; Xiang, J.; Chen, H.; Yang, Q.; Guan, A.; Li, Q.; Yu, L.; Tang, Y. *Chem. Commun.* **2014**, *50*, 15385–15388.

(212) Shi, Y.; Sun, H.; Xiang, J.; Yu, L.; Yang, Q.; Li, Q.; Guan, A.; Tang, Y. *Anal. Chim. Acta* **2014**, *857*, 79–84.

(213) Fiala, R.; Spackova, N.; Foldynova-Trantirkova, S.; Sponer, J.; Sklenar, V.; Trantirek, L. *J. Am. Chem. Soc.* **2011**, *133*, 13790–13793.

(214) Smiatek, J.; Heuer, A. *RSC Adv.* **2014**, *4*, 17110–17113.

(215) Chakraborty, T.; Hansen, J.; Schazschneider, B.; Ristow, H. *Eur. J. Biochem.* **1978**, *90*, 261–270.

(216) Hansen, J.; Pschorn, W.; Ristow, H. *Eur. J. Biochem.* **1982**, *126*, 279–284.

# Appendix

# A.1 FRET Melting Dose Response Curves For Screen Hits

*Example FRET melting curves for human telomeric i-motif with increasing concentration of Ligand. [DNA] = 200 nM hTeloC, buffer = 10 mM sodium cacodylate and 100 mM NaCl at pH 5.5.*

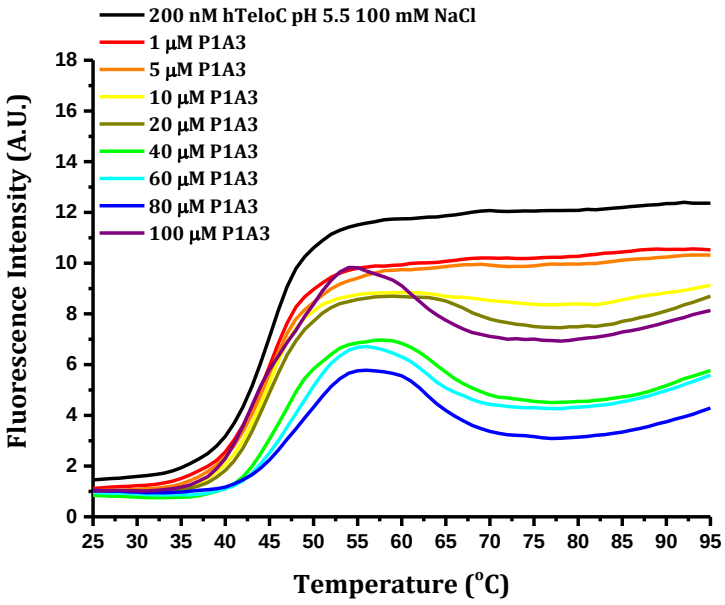



*Figure A.1.1:* ***Chlorhexidine****.*

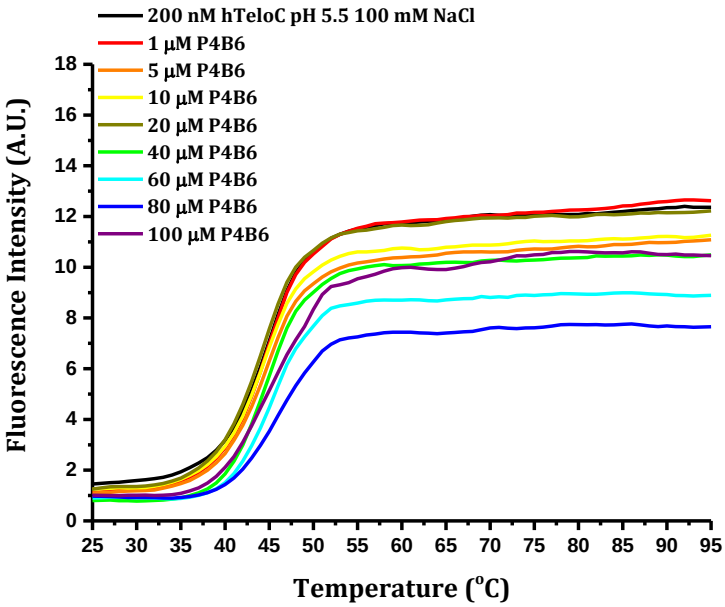



*Figure A.1.2:* ***Tamoxifen citrate****.*

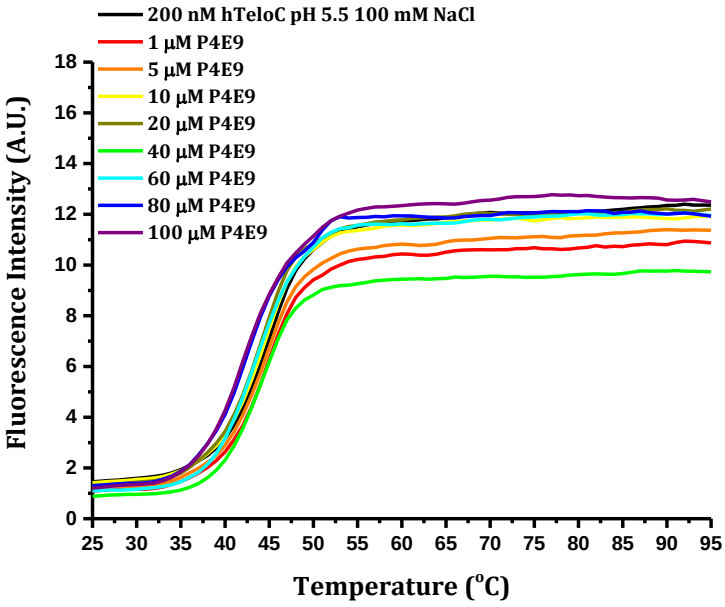



*Figure A.1.3:* ***Hexetidine.***

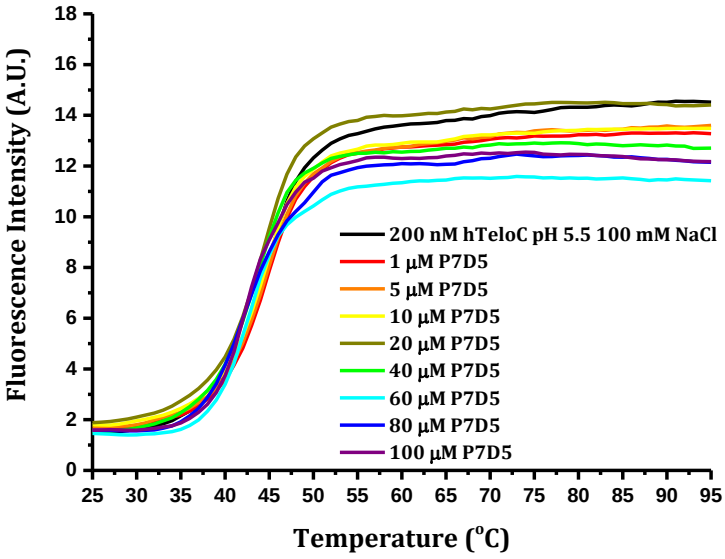



*Figure A.1.4:* ***Bromocriptine mesylate.***

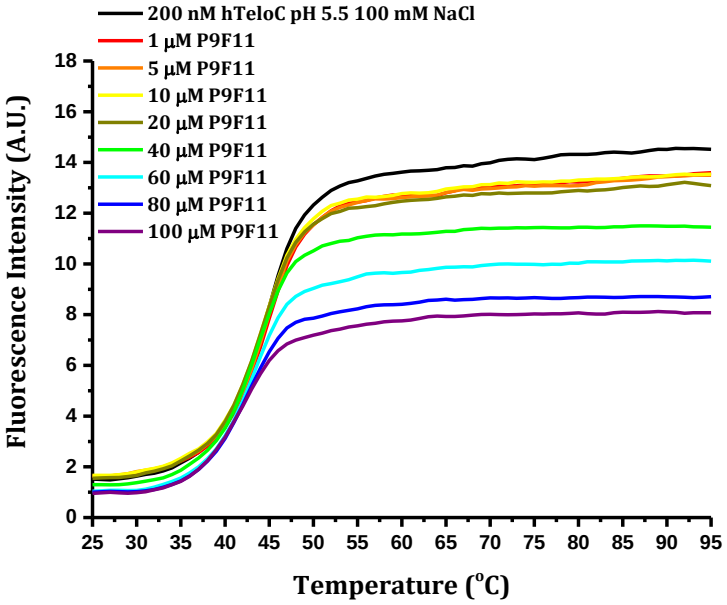



*Figure A.1.5:* ***Pararosaniline pamoate.***

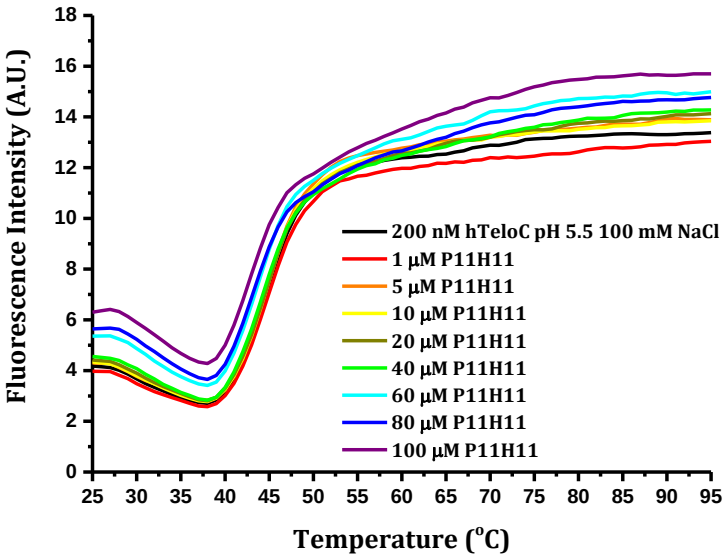



*Figure A.1.6:* ***Tilorone hydrochloride.***

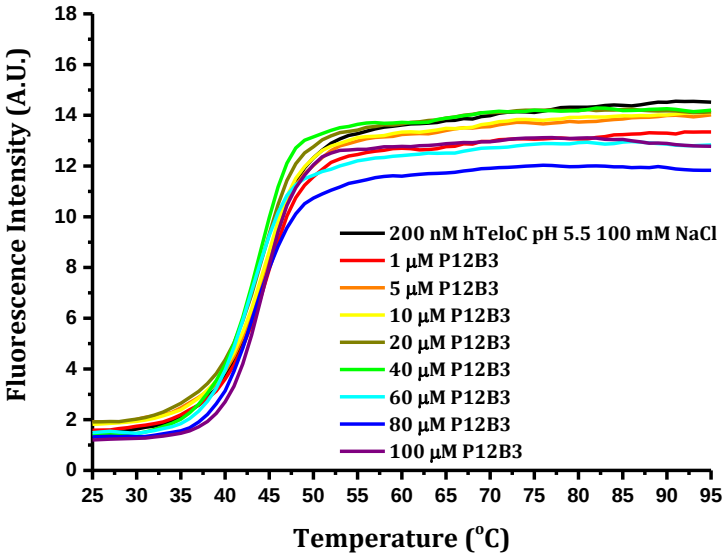



*Figure A.1.7:* ***Aiodarone hydrochloride****.*

## A.2 Further Example SPR Binding Curves

*Example binding curves for **mitoxantrone** and **tilorone**. Running buffer = 10 mM sodium cacodylate, 100 mM NaCl and 0.05% tween 20 at pH 5.5.*

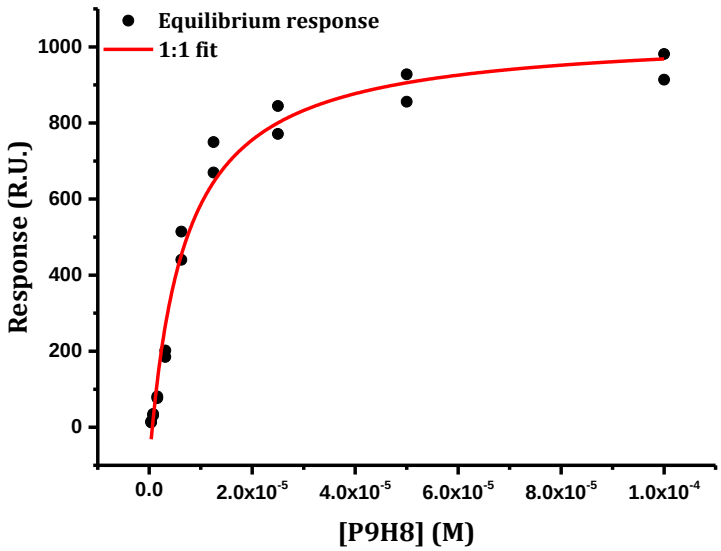



*Figure A.2.1: **Mitoxantrone** with the human telomeric i-motif.*

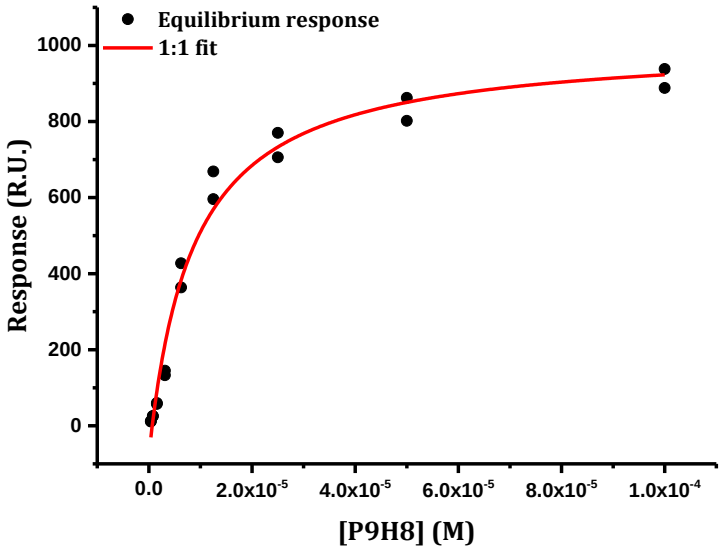



*Figure A.2.2: **Mitoxantrone** with the c-Myc promoter i-motif.*

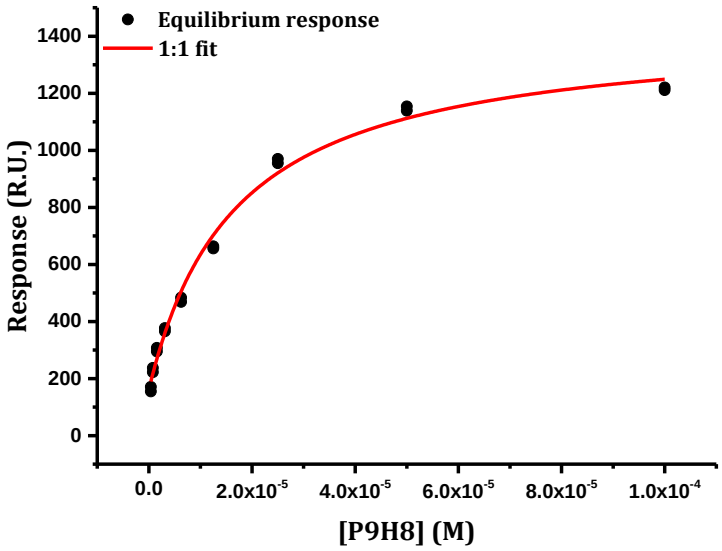



*Figure A.2.3:* ***Mitoxantrone*** *with duplex DNA.*

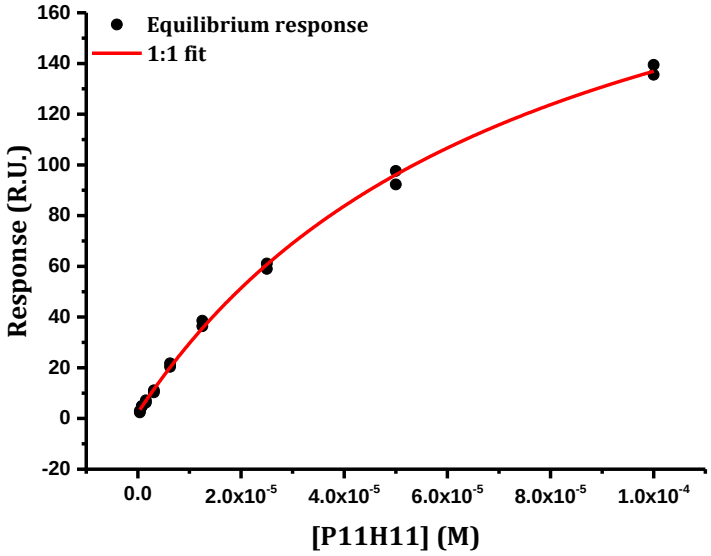



*Figure A.2.4:* ***Tilorone*** *with the human telomeric i-motif.*

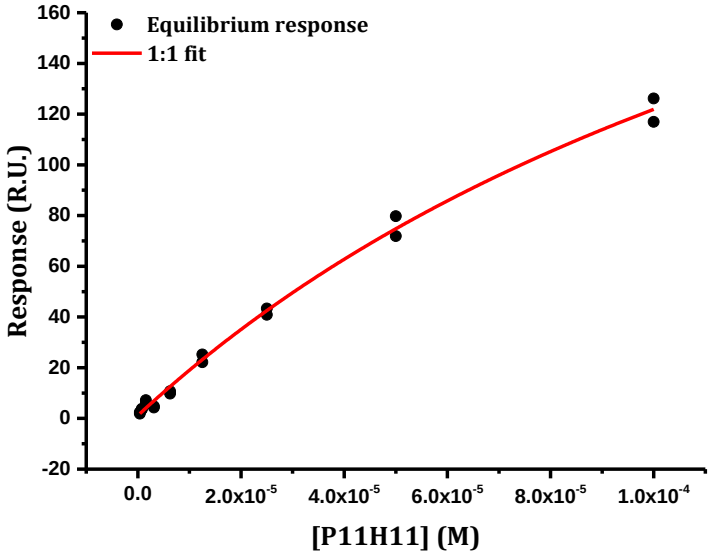



*Figure A.2.5:* ***Tilorone*** *with the c-Myc promoter i-motif.*

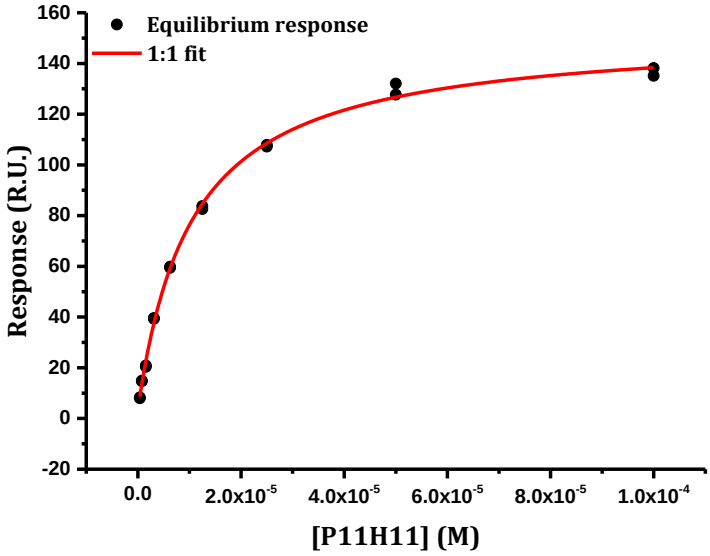



*Figure A.2.6:* ***Tilorone*** *with duplex DNA.*

## A.3 FRET Melting Curves For The Cation Screen

*FRET melting curves for human telomeric i-motif with increasing concentrations of different metal salts. [DNA] = 200 nM hTeloC, buffer = 10 mM sodium cacodylate and 5 mM NaCl at pH 7.4 (left) and pH 5.5 (right).*

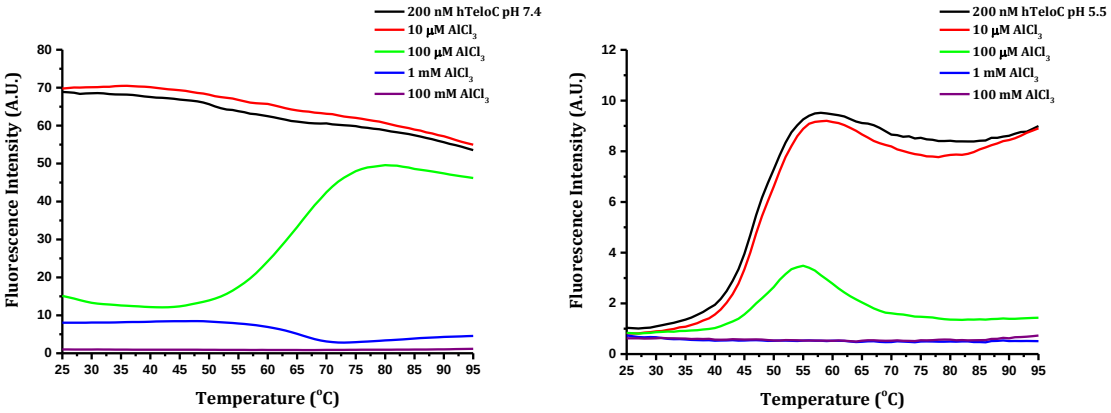



*Figure A.3.1: $AlCl_3$*

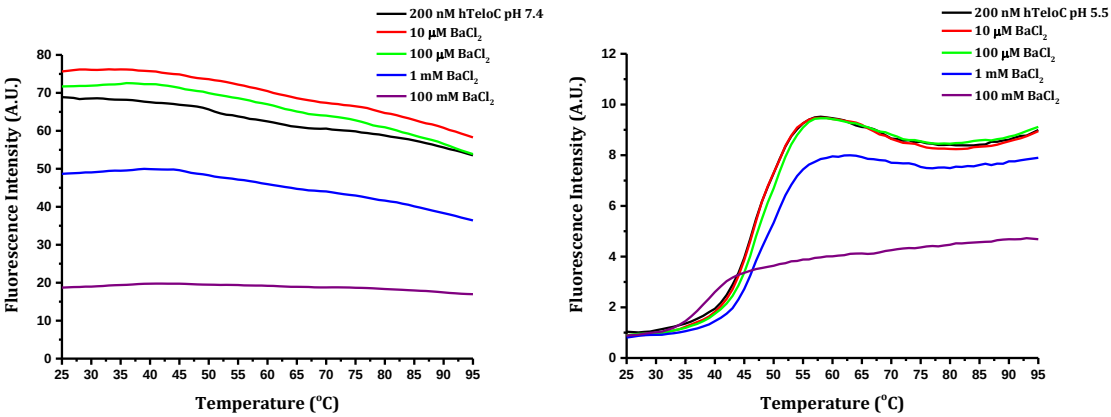



*Figure A.3.2: $BaCl_2$*

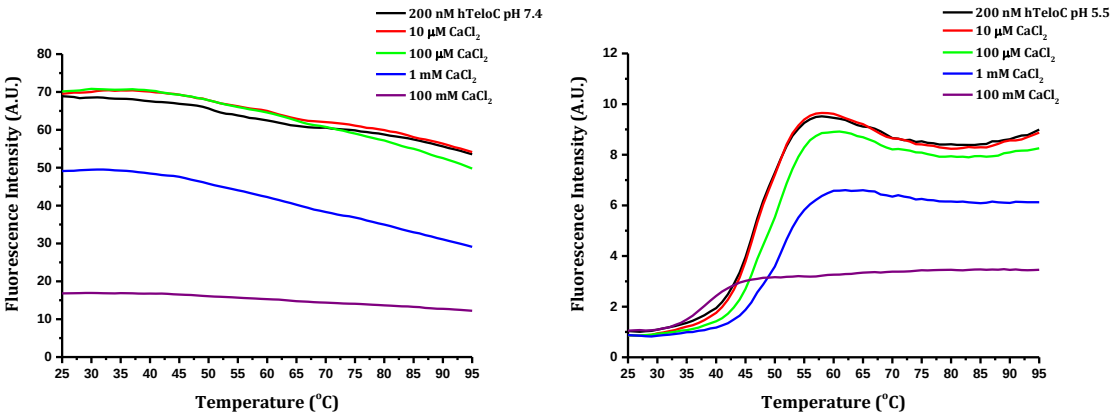



*Figure A.3.3: $CaCl_2$*

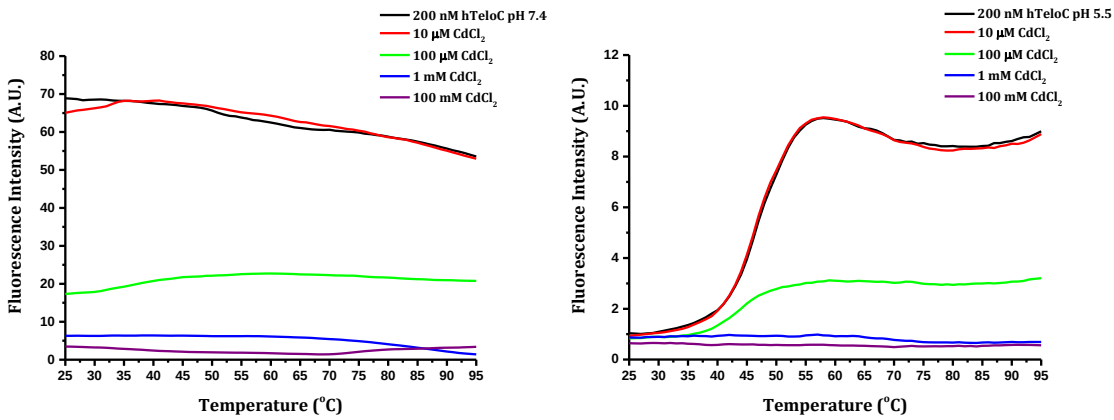



*Figure A.3.4: $CdCl_2$*

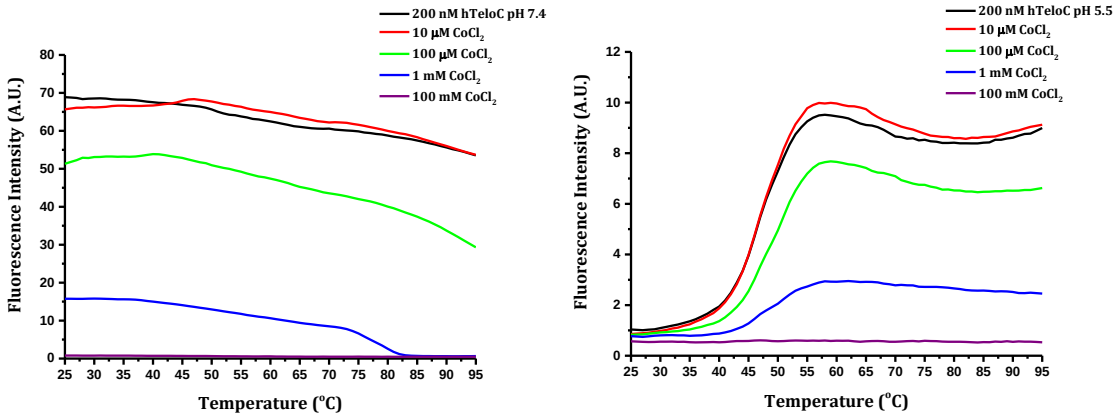



*Figure A.3.5: $CoCl_2$*

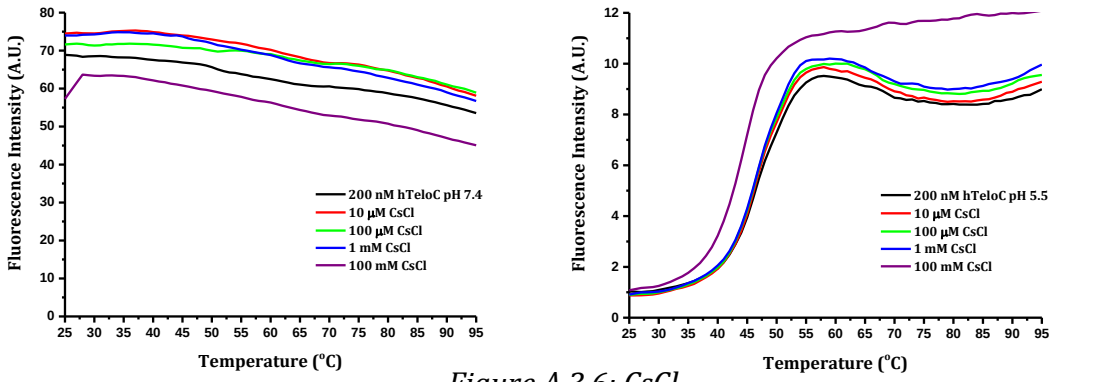



*Figure A.3.6: CsCl*

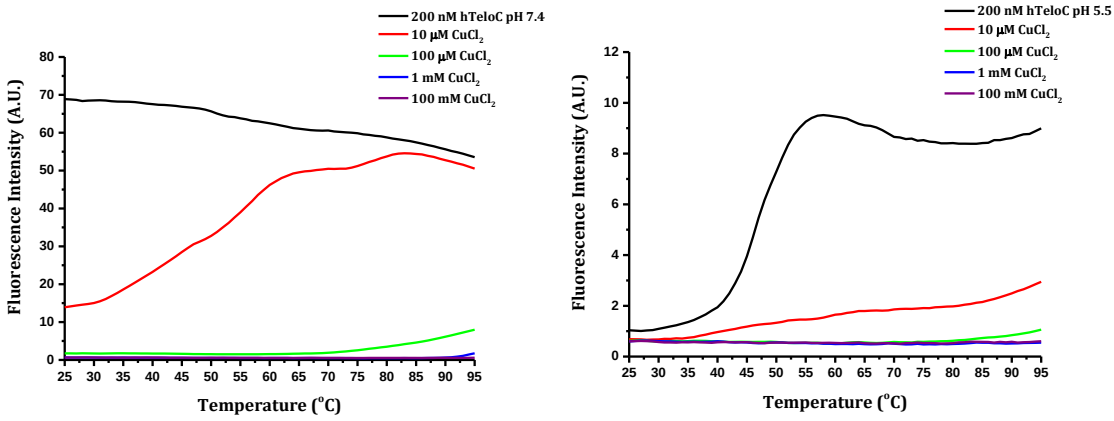



*Figure A.3.7: $CuCl_2$*

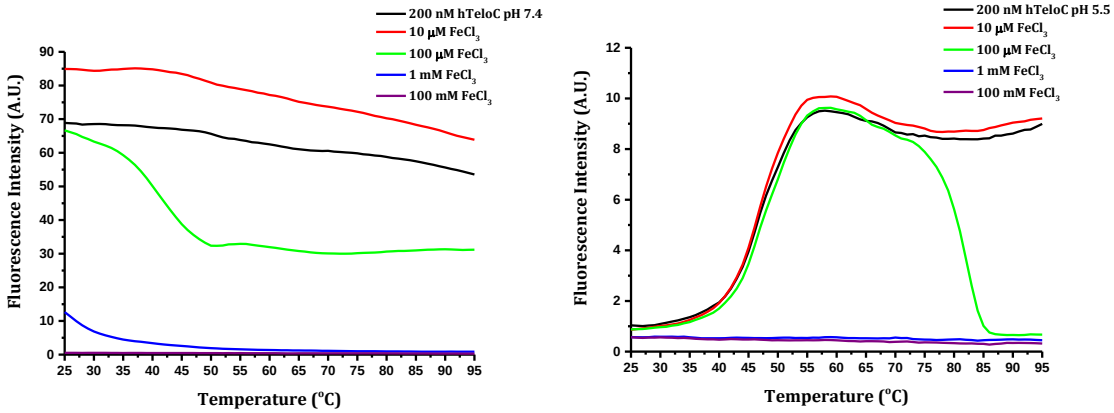



*Figure A.3.8: $FeCl_3$.*

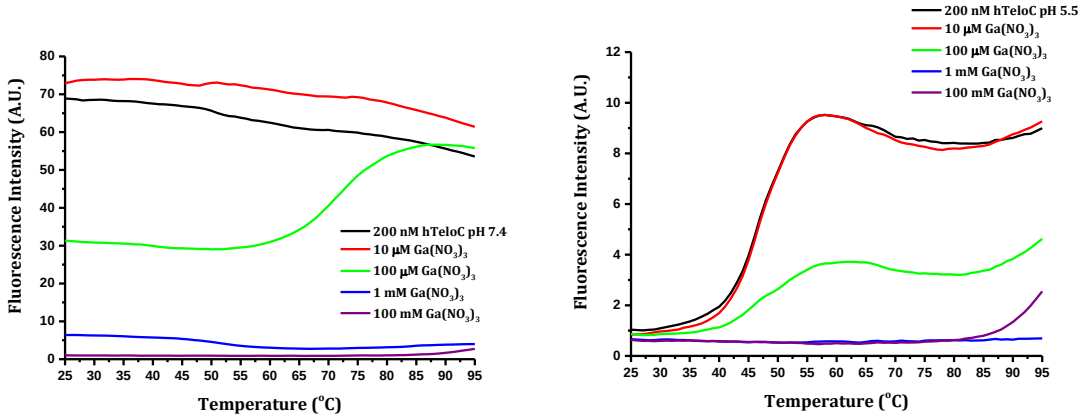



*Figure A.3.9: $Ga(NO_3)_3$.*

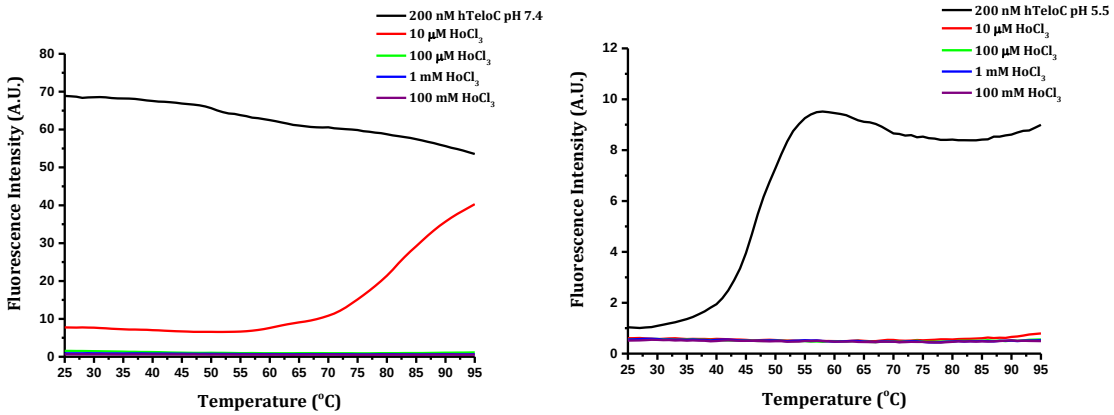



*Figure A.3.10: $HoCl_3$.*

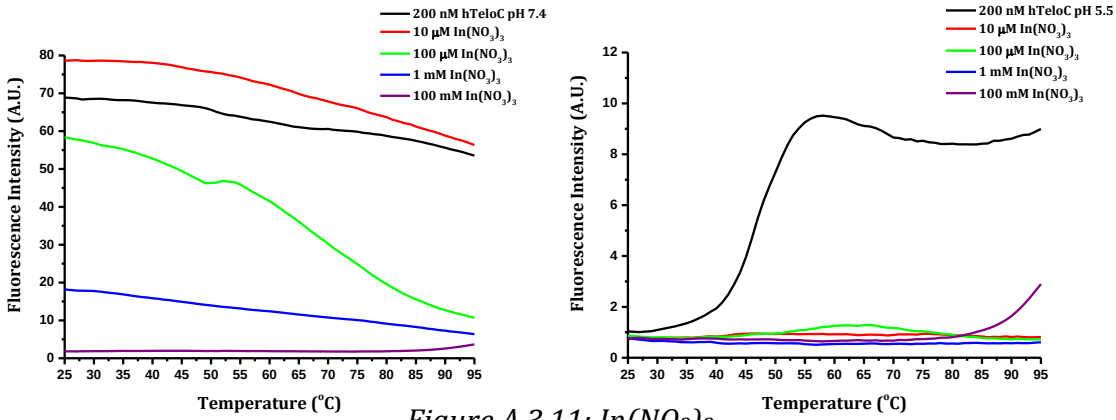



*Figure A.3.11: $In(NO_3)_3$.*

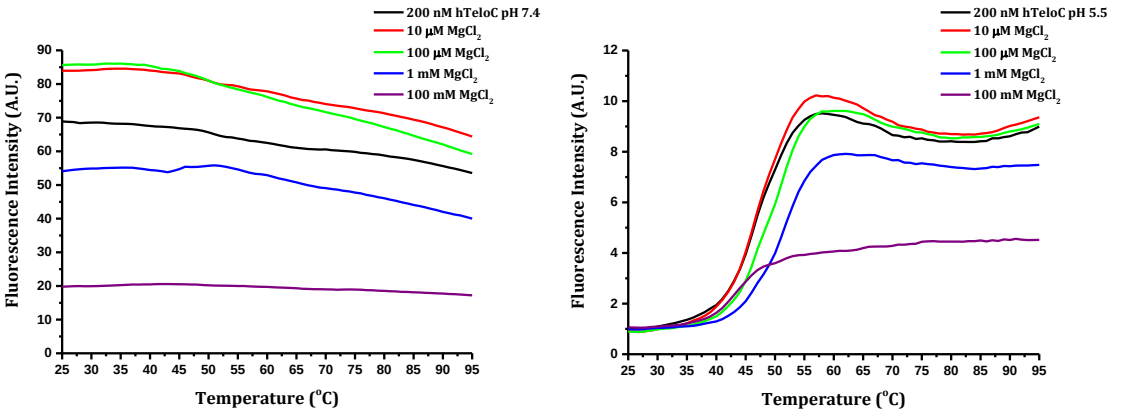



*Figure A.3.12: $MgCl_2$.*

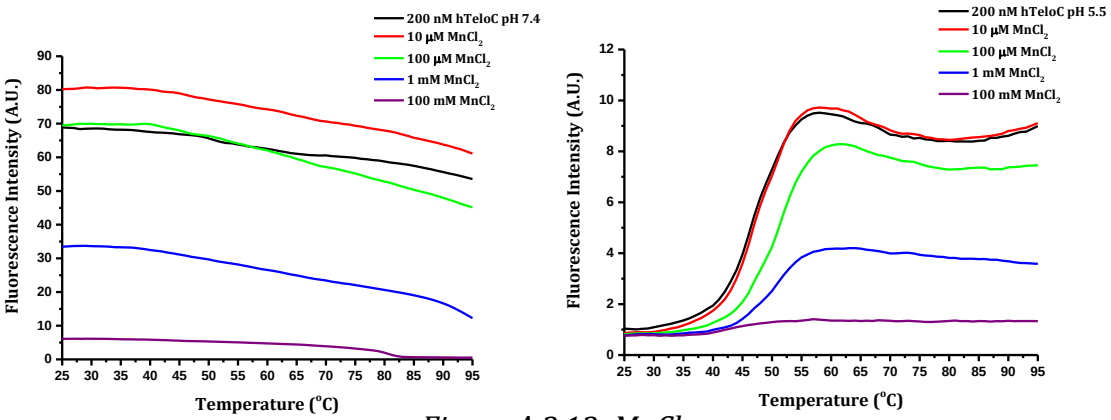



*Figure A.3.13: $MnCl_2$.*

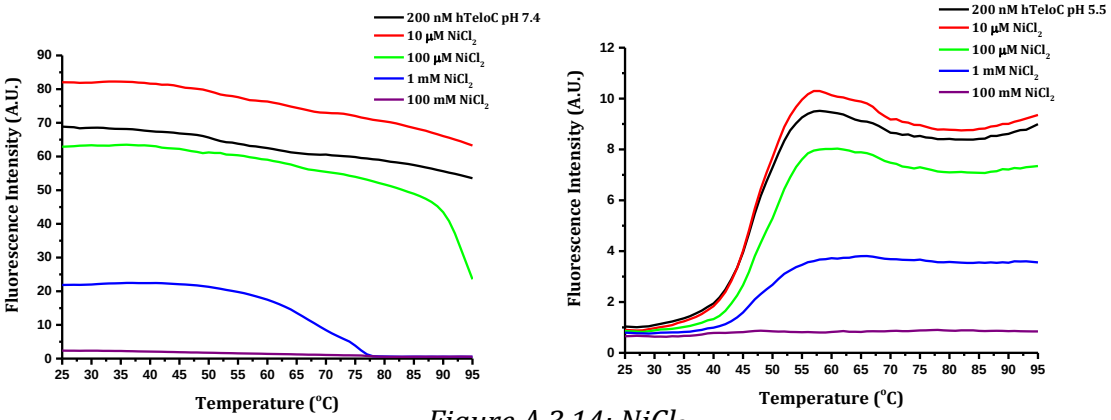



*Figure A.3.14: $NiCl_2$.*

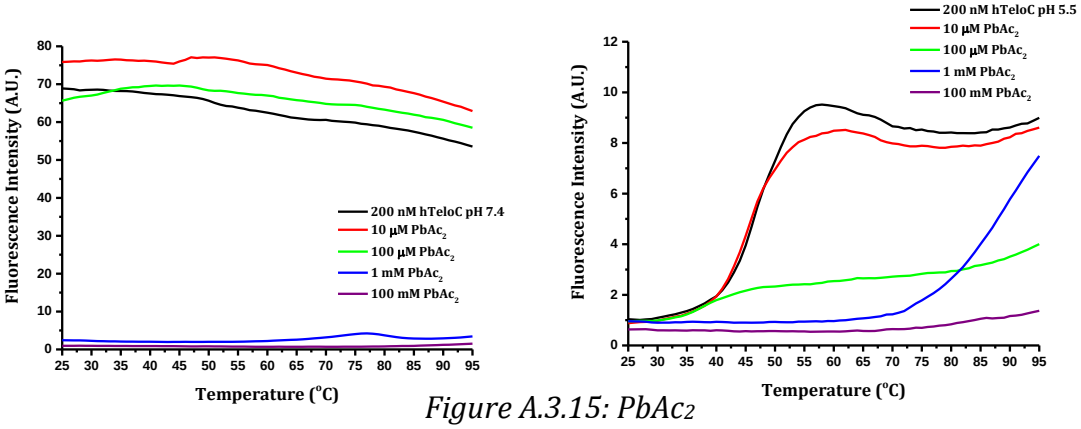



*Figure A.3.15: $PbAc_2$*

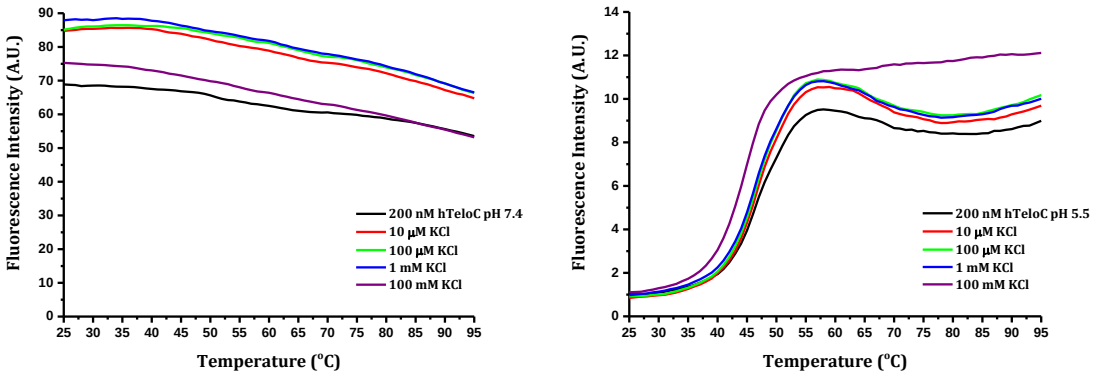



*Figure A.3.16: KCl.*

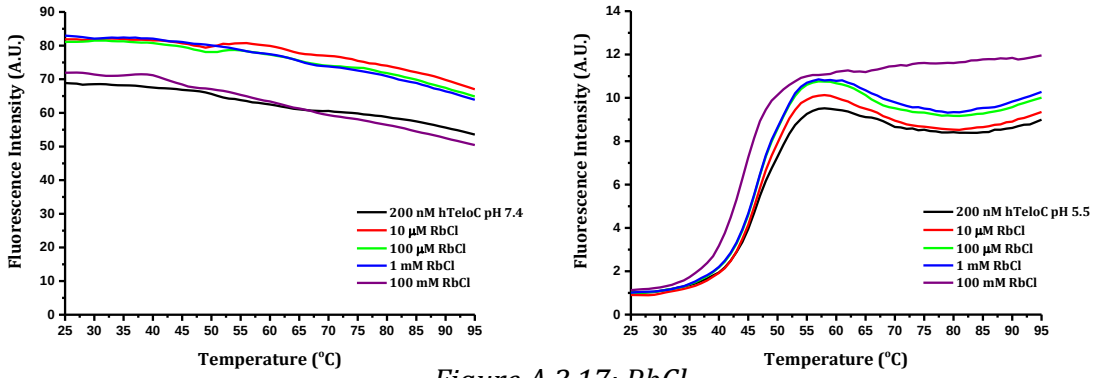



*Figure A.3.17: RbCl.*

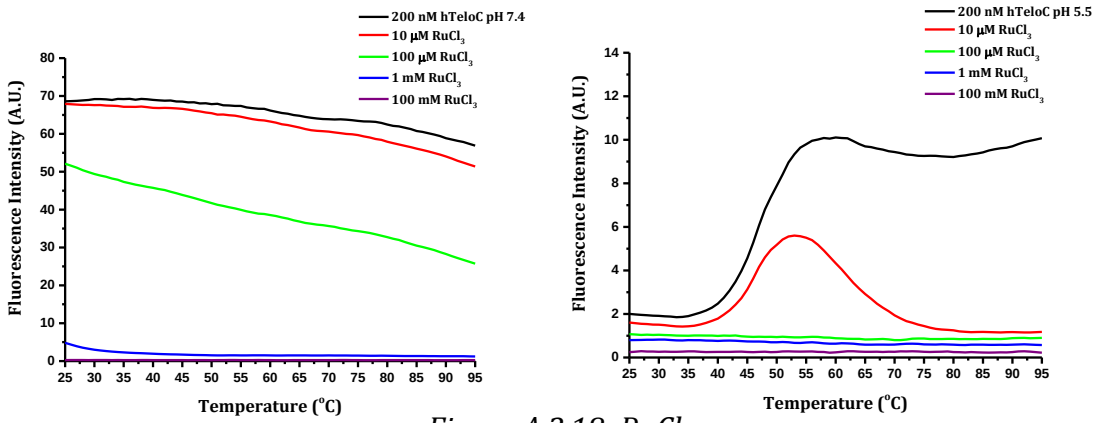



*Figure A.3.18: $RuCl_3$.*

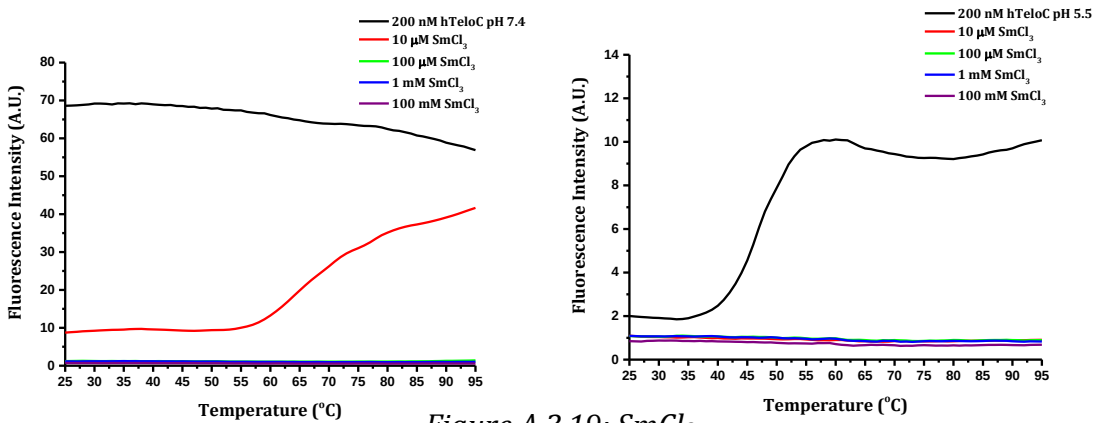



*Figure A.3.19: $SmCl_3$.*

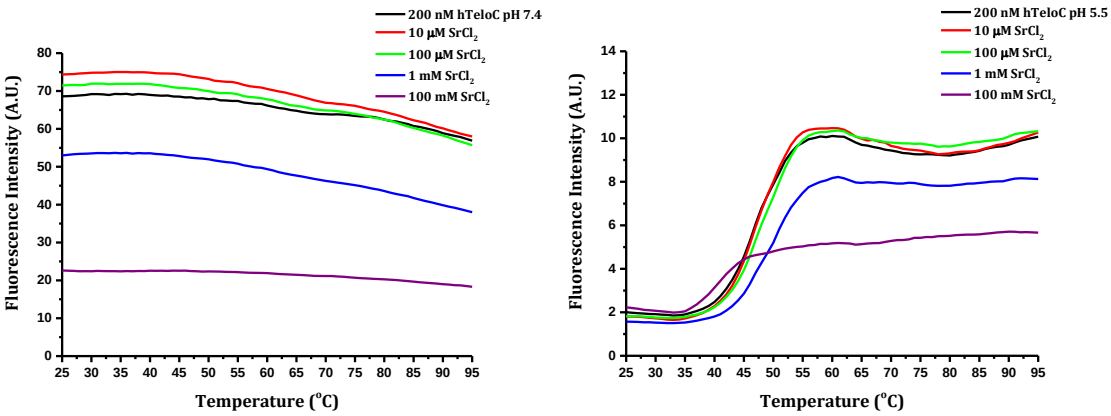



*Figure A.3.20: $SrCl_2$.*

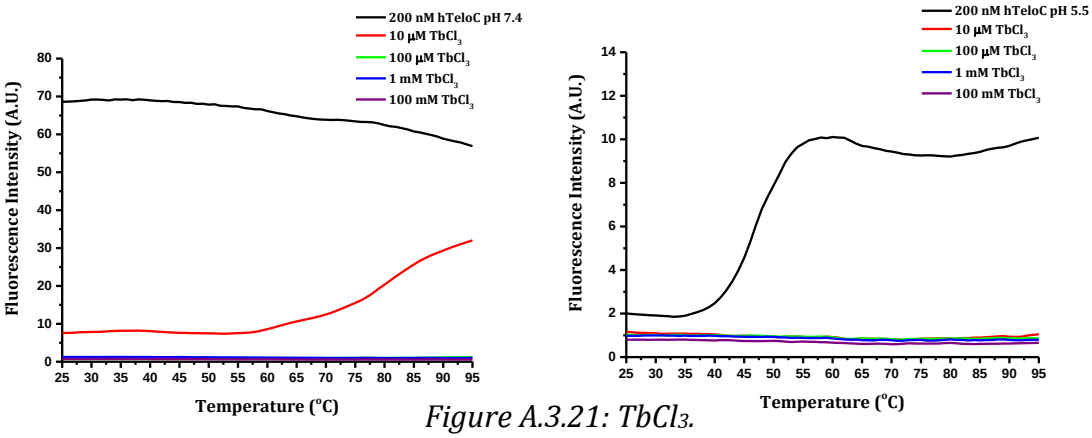



*Figure A.3.21: $TbCl_3$.*

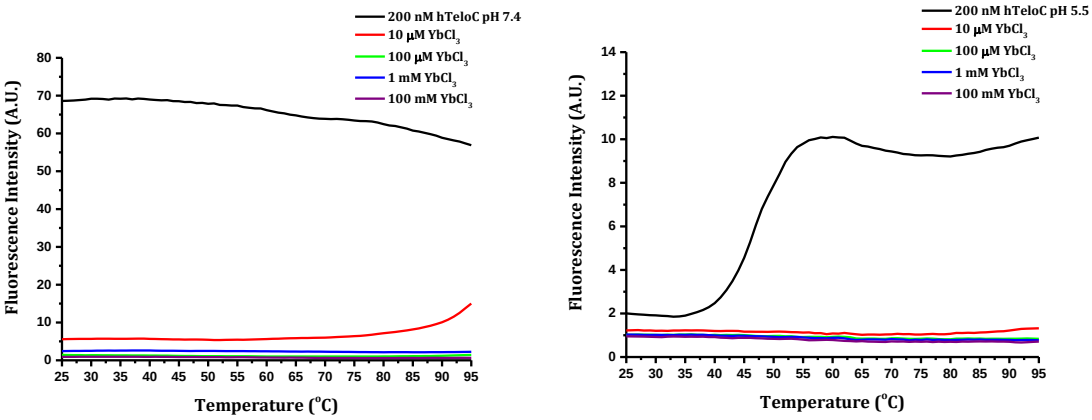



*Figure A.3.22: $YbCl_3$.*

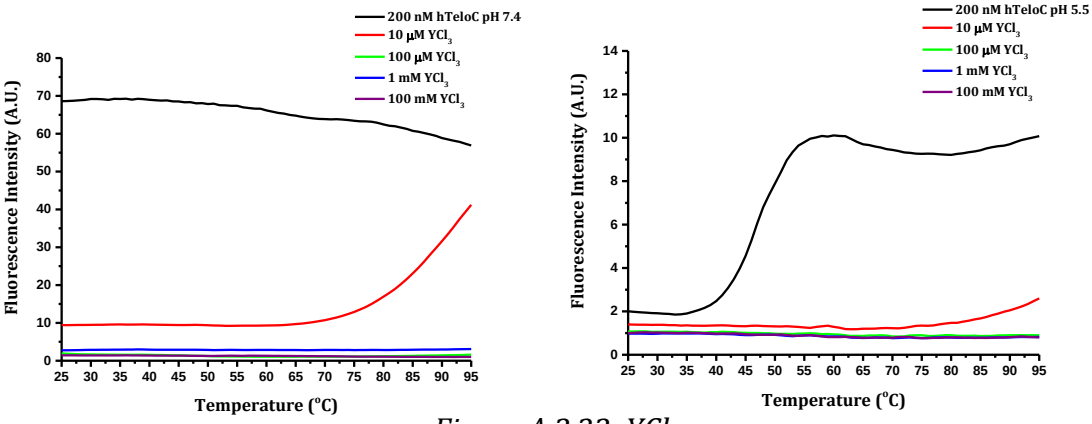



*Figure A.3.23: $YCl_3$.*

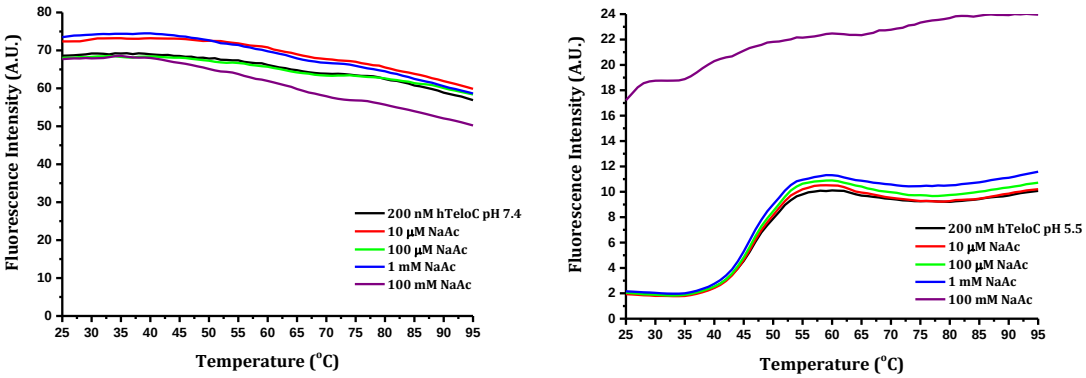



*Figure A.3.24: NaAc.*

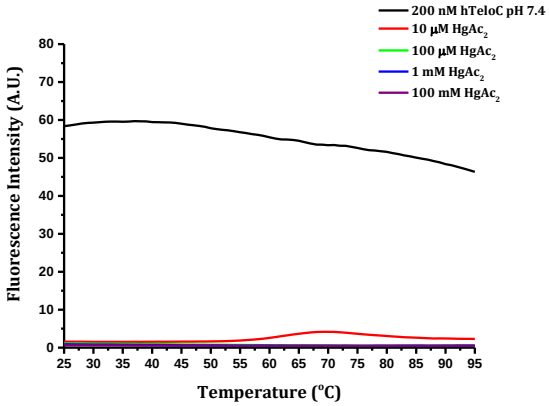



*Figure A.3.25: $HgAc_2$ pH 7.4.*

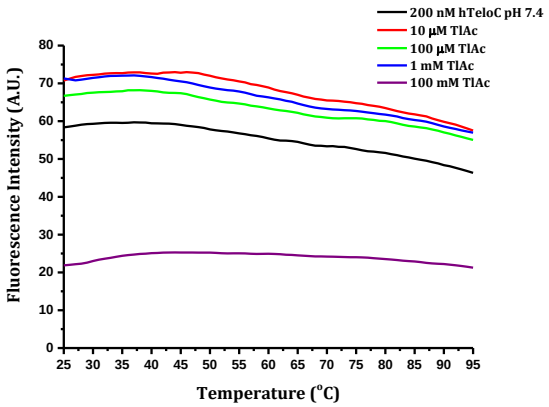



*Figure A.3.26: TlAc pH 7.4.*

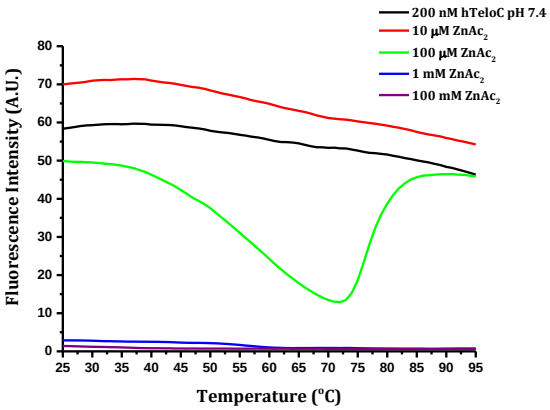



*Figure A.3.27: $ZnAc_2$ pH 7.4.*

Publications based upon work within this thesis can be found at the following links:

http://pubs.rsc.org/en/Content/ArticleLanding/2013/CC/c3cc43495h#!divAbstract

http://www.sciencedirect.com/science/article/pii/S0968089614004064