

# **Synthesis of Novel Biological Nanostructures**

**Mahmoud Amgad Salaheldin Abdelhamid**

A thesis submitted for the degree of Doctor of Philosophy

University of East Anglia

School of Pharmacy

May 2019

© “This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that use of any information derived therefrom must be in accordance with current UK Copyright Law. In addition, any quotation or extract must include full attribution.”

---

## **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 are given, this work is original and has been carried out by the author alone.

Mahmoud A. S. Abdelhamid

May 2019



---

## Acknowledgements

Pursuing a PhD had been a dream of mine long before I truly understood what it meant. Knowing now what was required of me to complete this monumental undertaking, I also know that it absolutely would not have been possible without the support of many people to whom I owe a great deal of thanks.

I must begin by sincerely thanking my supervisors, Dr Zoë Waller and Dr Andrew Gates, for allowing me to perform this research in their laboratories, and for their invaluable guidance throughout this journey. I would also like to thank the Eastern ARC for funding my work; as well as the University of East Anglia for being a wonderful place where I have learned so much over the past 9 years. I also extend my gratitude to all the collaborators I had the pleasure of working with during my PhD, at the University, on the wider Norwich Research Park and across the world.

Next, I am extremely grateful to my family for everything they have done for me, and put up with from me, for the past 27 years. My father and mother, Amgad & Hanaa, for their undying love and support which has made me into the man I am today. My sister, Shaymaa, for loving me far more than I deserve, and whom I love far more than I show. My brother, Moustafa, no matter if you're 30 feet or 3000 miles away from me I know you will always be by my side.

I was extremely fortunate to have worked in a group with some great people who all helped me accomplish the dream of completing my PhD. First thanks go to the members of the Waller research group there when I began: Qiran and Elisé, who taught me so much when I was starting out. Over time I myself became one of the group's veterans and had the pleasure of working with the newer members of the group: Phil, Zara, Emily, Summer, Ying, Alex and Rupesh. As well as some researchers who were with us for a short time: Jessica, Rouven, Mimi, Bob and Zuzana. Apart from being colleagues I am blessed to have those of you whom I consider great friends of mine, and others more than that still.

Outside of the research group I also made some very cherished friends who I hope to stay close to as we all continue along our divergent life paths: Quim & Sara, Andrew & Zoë G., Gloria, Ryan, Remy, and Hui. Although I may be terrible at keeping in touch, you all mean a great deal to me and always will.

---

Finally, I must return and give due consideration to the immeasurable gratitude I owe to Dr Zoë Waller. Zoë, while my PhD journey was predominantly a positive one you were there for me when it wasn't, believing in my ability to do that which I myself did not believe I could do. In so many ways this achievement would not have been possible without you. Having you as a supervisor, a mentor and a friend has been the greatest privilege; I look forward to continuing to learn from you and to working together for a long time to come.

Mahmoud Abdelhamid

---

## Abstract

*Chapter 1* is a general introduction to the field surrounding quadruplex DNA with a focus on the i-motif, a quadruplex DNA structure.

*Chapter 2* is split into three sections and explores the interactions of cations with quadruplex DNA. The first section discusses what was already known about the effects of cations on the G-quadruplex and i-motif structures. The second section describes the effects observed upon addition of a range of cations to the i-motif. The final section describes the discovery and characterisation of a redox-dependent system for the control of i-motif DNA structure using copper cations.

*Chapter 3* explores the interaction of small molecules with quadruplex DNA. The first section discusses reports in the literature relating to the effects of ligands on the G-quadruplex and the i-motif. The second describes our efforts using a variety of screening methodologies to identify a ligand that is specific for i-motif. The third describes how 'specific' G-quadruplex ligands have been found to interact with the i-motif.

*Chapter 4* begins with an introduction on the use of DNA as a functional material. This is followed by a discussion of DNA tetrahedra and introduces the proposed design of a novel biological nanostructure in an attempt to exploit quadruplex DNA to impart further functionality to DNA nanoarchitectures. The third section discusses the efforts in synthesising the structure.

*Chapter 5* offers a general discussion about this body of work and explores avenues for the future development of it.

*Chapter 6* describes the experimental procedures used in *Chapters 2, 3 and 4*.

---

# Table of Contents

<i>Declaration</i> .....	<i>ii</i>
<i>Acknowledgements</i> .....	<i>iii</i>
<i>Abstract</i> .....	<i>v</i>
<i>Table of Contents</i> .....	<i>vi</i>
<i>List of Figures</i> .....	<i>ix</i>
<i>List of Tables</i> .....	<i>xvii</i>
<i>List of Abbreviations</i> .....	<i>xix</i>
<i>Published work within this thesis</i> .....	<i>xx</i>
<b>Chapter 1: Introduction</b> .....	<b>1</b>
1.1 DNA SECONDARY STRUCTURE.....	2
1.2 QUADRUPLEX DNA.....	4
1.2.1 <i>The G-Quadruplex</i> .....	6
1.3 THE I-MOTIF .....	7
1.4 DNA AS A FUNCTIONAL MATERIAL.....	10
1.4.1 <i>Quadruplex DNA Nanotechnology</i> .....	12
1.5 TECHNIQUES FOR STUDYING QUADRUPLEXES.....	14
1.5.1 <i>Ultraviolet (UV) Spectroscopy</i> .....	14
1.5.2 <i>Circular Dichroism (CD) Spectroscopy</i> .....	14
1.5.3 <i>Nuclear Magnetic Resonance (NMR) Spectroscopy</i> .....	15
1.5.4 <i>Fluorescence Resonance Energy Transfer (FRET) and FRET-melting</i> .....	15
1.5.5 <i>Fluorescent Indicator Displacement (FID) Assays</i> .....	16
1.5.6 <i>Surface Plasmon Resonance (SPR)</i> .....	16
1.5.7 <i>Gel Electrophoresis</i> .....	17
1.5.8 <i>Molecular Modelling and Dynamics</i> .....	17
1.5 CHALLENGES AND OPPORTUNITIES IN I-MOTIF RESEARCH .....	18
1.6 AIM AND OBJECTIVES .....	19
<b>Chapter 2: Cations</b> .....	<b>21</b>
2.1 CATIONS AND QUADRUPLEXES.....	22
2.2 INVESTIGATING THE EFFECTS OF CATIONS ON I-MOTIF DNA.....	23
2.2.1 <i>Tricky Temperature</i> .....	23
2.2.2 <i>Pesky pH</i> .....	26
2.3 REDOX-DEPENDENT CONTROL OF I-MOTIF DNA STRUCTURE USING COPPER CATIONS .....	29
2.3.1 <i>Biophysical Characterisation of the Interaction between Cu<sup>+</sup> and hTeloC</i> .....	30
2.3.2 <i>Computational Investigation of the Interaction between Cu<sup>+</sup> and hTeloC</i> .....	33
2.3.3 <i>Reversibility and Redox Sensitivity of the Cu<sup>+</sup>-stabilised i-Motif</i> .....	36

<b>Chapter 3: Ligands .....</b>	<b>45</b>
3.1 LIGANDS AND QUADRUPLICES .....	46
3.2 I-MOTIF LIGAND HUNT .....	47
3.2.1 Surface Plasmon Resonance (SPR) Screen.....	52
3.2.2 Pesky pH Strikes SPR .....	57
3.2.3 FRET-melting .....	61
3.2.4 NMR .....	64
3.2.5 The Curious Case of 143491 .....	66
3.3 THE EFFECTS OF G-QUADRUPLIX LIGANDS ON I-MOTIF DNA .....	70
3.3.1 <i>i</i> -Motif ‘Home Advantage’ – Acidic pH.....	70
3.3.2 <i>i</i> -Motif ‘Plays Away’ – Neutral pH.....	73
<b>Chapter 4: DNA Nanotechnology .....</b>	<b>82</b>
4.1 DNA TETRAHEDRA.....	83
4.1.1 Reconfigurable DNA Tetrahedra .....	85
4.2 THE QUADRUPLIX TETRAHEDRON ( $Q_{TET}$ ).....	87
4.2.1 $Q_{TET}$ Quadruplexes .....	88
4.2.2 $Q_{TET}$ Synthesis .....	91
4.2.3 Seeing the Switch .....	94
<b>Chapter 5: Discussion and Future Work .....</b>	<b>105</b>
5.1 DISCUSSION .....	106
5.2 FUTURE WORK.....	107
<b>Chapter 6: Experimental.....</b>	<b>112</b>
6.1 GENERAL EXPERIMENTAL.....	113
6.2 EXPERIMENTAL FOR CHAPTER 2 .....	116
6.3 EXPERIMENTAL FOR CHAPTER 3 .....	119
6.4 EXPERIMENTAL FOR CHAPTER 4 .....	120
<b>References.....</b>	<b>122</b>
<b>Appendix.....</b>	<b>146</b>
A.1 FRET-MELTING DATA FOR INITIAL LIGAND IDENTIFICATION.....	147
A.2 % $R_{MAX}$ FROM SPR SCREEN AT pH 5.5.....	155
A.3 RESPONSE VS CONCENTRATION PLOTS FROM SPR AT pH 5.5.....	156
A.4 RESPONSE VS CONCENTRATION PLOTS FROM SPR AT pH 7.0.....	160
A.5 $\Delta T_M$ OF FRET-LABELLED DNA WITH NSC LIGANDS .....	165
A.6 FRET-MELTING CURVES OF DNA WITH NSC LIGANDS .....	166

---

A.7 NMR SPECTRA OF NSC LIGANDS WITH DAP .....	169
A.8 FID ASSAYS AT ACIDIC PH .....	171
A.9 CD MELTING OF I-MOTIF WITH LIGANDS AT NEUTRAL PH .....	172
A.10 UV MELTING OF I-MOTIF WITH LIGANDS AT NEUTRAL PH .....	174
A.11 FID ASSAY OF I-MOTIF WITH LIGANDS AT NEUTRAL PH.....	176
A.12 LIGAND STRUCTURES REFERENCE SHEETS .....	177
A.13 PUBLICATIONS.....	180

---

## List of Figures

### Chapter 1: Introduction

- Figure 1.1.1* Watson and Crick base pairing and the structure of B-form DNA. PDB ID: 1BNA.
- Figure 1.1.2* Examples of some alternative DNA secondary structures. A) A-DNA (PDB ID: 440D). B) Hairpin (PDB ID: 1AC7). C) Triplex (PDB ID: 1D3X). D) Z-DNA (PDB ID: 4OCB). E) Four-way junction (PDB ID: 467D).
- Figure 1.2.1* (Top) Hoogsteen hydrogen bonding of G-tetrad with stabilising cation in the core (left) and a hemi-protonated cytosine-cytosine+ base pair (right). (Bottom) G-quadruplex (PDB ID: 143D; left) and i-motif (PDB ID: 1ELN; right). Nucleic Acid Database (NDB) colouring: adenine is red, thymine is blue, guanine is green, cytosine is yellow.
- Figure 1.2.2* Illustration of biological processes in cells which have been found to potentially be regulated by G-quadruplexes.
- Figure 1.3.1* Monomeric (PDB ID: 1ELN), dimeric (PDB ID: 2MRZ) and tetrameric (PDB ID: 1YBL) i-motifs.
- Figure 1.4.1* Illustration of the design space expansion and the growth of interest in structural DNA nanotechnology. Reprinted from *Current Opinion in Biotechnology*, 24, Beikko Linko, Hendrik Dietz, The enabled state of DNA nanotechnology, 555-561, Copyright (2013), with permission from Elsevier.71
- Figure 1.4.1.1* Representation of the operation cycle of the aptamer-based molecular machine in the presence of thrombin. Reprinted from *Angewandte Chemie International Edition*, 43, Wendy U. Dittmer, Andreas Reuter, Friedrich C. Simmel, A DNA-Based Machine That Can Cyclically Bind and Release Thrombin, 3550-3553, Copyright (2004), with permission from Wiley.

### Chapter 2: Cations

- Figure 2.2.1.1 CD spectra of 10  $\mu\text{M}$  hTeloC in 10 mM sodium cacodylate buffer at room temperature. Dotted lines represent spectra of samples at different pH values measured without storage at low temperature. Solid lines represent spectra of samples at pH 7.0 measured after storage at 4°C for the indicated time.
- Figure 2.2.2.1 CD titrations of 10  $\mu\text{M}$  hTeloC in 10 mM sodium cacodylate buffer at pH 7.0 (black) with increasing cation concentration in 10-fold increments from 1  $\mu\text{M}$  to 10 mM (coloured) and the change in pH of 10 mM sodium cacodylate buffer from an initial pH of 7 with addition of cation. Aluminium(III) – green, chromium(II) – blue, and vanadium(II) – sand.
- Figure 2.3.1 Hemi-protonated C·C<sup>+</sup> base pair (left) and proposed C·C base pair stabilised by Cu<sup>+</sup>.
- Figure 2.3.1.1 ‘Cu<sup>+</sup>-difference’ spectra using 125  $\mu\text{M}$  of Cu<sup>+</sup> to form the final conformations at pH 7.4 (red) and pH 5.5 (black).
- Figure 2.3.1.2 CD spectra of 10  $\mu\text{M}$  hTeloC in 50 mM sodium cacodylate buffer (black) at pH 7.4 (left) with titration up to 50  $\mu\text{M}$  Cu<sup>+</sup> (red) and at pH 5.5 (right) with titration up to 150  $\mu\text{M}$  Cu<sup>+</sup> (red).
- Figure 2.3.1.3 CD spectra (left) of 10  $\mu\text{M}$  hTeloC at pH 5.5 (dashed black) with addition of 150  $\mu\text{M}$  Cu<sup>+</sup> (red) or 1 mM Cu<sup>2+</sup> (black) and the ellipticity at 288 nm of 10  $\mu\text{M}$  hTeloC (right) in 50 mM sodium cacodylate buffer at pH 5.5 with titration up to 195  $\mu\text{M}$  Cu<sup>+</sup> (red) and in 10 mM sodium cacodylate buffer at pH 5.5 with titration up to 1 mM Cu<sup>2+</sup> (black). Error bars show standard deviation across three repeats.
- Figure 2.3.1.4 Job plot of hTeloC and Cu<sup>+</sup> in 50 mM sodium cacodylate buffer at pH 5.5. The black and red symbols represent the points used for fitting the respective linear best fits to determine the intercept.
- Figure 2.3.2.1 Structures of the C·Cu<sup>+</sup>·C (left) and C·Cu<sup>2+</sup>·C (right) base pairs, determined by DFT [TPSS-D3(BJ)/def2-TZVP] geometry optimisation. Colour code: C: cyan, H: white, N: blue, O: red, Cu: gold.
- Figure 2.3.2.2 Model of hTeloC i-motif structure stabilised by protonation of cytosine residues, snapshot from the end of a 200 ns simulation



---

(left). Model of the *i*-motif structure stabilised by  $\text{Cu}^+$  ions derived from the protonated model by geometry optimisation with the PM6-D3H4 method.

Figure 2.3.3.1 CD spectra (left) of  $10\ \mu\text{M}$  hTeloC with  $150\ \mu\text{M}$   $\text{Cu}^+$  at pH 5.5 (black) with titration up to  $300\ \mu\text{M}$  DETC (red) and the ellipticity at 288 nm (right) of  $10\ \mu\text{M}$  hTeloC in 50 mM sodium cacodylate buffer at pH 5.5 with titration up to  $300\ \mu\text{M}$  DETC. Error bars show standard deviation across three repeats.

Figure 2.3.3.2  $^1\text{H}$  NMR of  $10\ \mu\text{M}$  hTeloC in 50 mM sodium cacodylate buffer pH 5.5 with 5%  $\text{D}_2\text{O}$  (green), with addition of  $150\ \mu\text{M}$   $\text{Cu}^+$  (red), and addition of  $150\ \mu\text{M}$   $\text{Cu}^+$  and  $540\ \mu\text{M}$  DETC (blue).

Figure 2.3.3.3 CD spectra of  $10\ \mu\text{M}$  hTeloC in 50 mM sodium cacodylate buffer at pH 5.5 with  $150\ \mu\text{M}$   $\text{Cu}^+$  scanned immediately after addition (solid lines) and 3 hours later (dashed lines); (red) sample exposed to the air; (black) sample maintained in anoxic environment.

Figure 2.3.3.4 CD spectra of  $10\ \mu\text{M}$  hTeloC in 50 mM sodium cacodylate buffer at pH 5.5 (green); addition of  $150\ \mu\text{M}$   $\text{Cu}^{2+}$  (blue); addition of  $150\ \mu\text{M}$  sodium ascorbate (pink). For comparison, the CD spectra of  $10\ \mu\text{M}$  hTeloC in 50 mM sodium cacodylate buffer at pH 5.5 after addition of  $150\ \mu\text{M}$   $\text{Cu}^+$  (black).

Figure 2.3.3.5 CD spectra of  $10\ \mu\text{M}$  hTeloC in 50 mM sodium cacodylate buffer at pH 5.5 with  $150\ \mu\text{M}$   $\text{Cu}^{2+}$  immediately after addition of  $150\ \mu\text{M}$  sodium ascorbate and every 10 minutes for 240 minutes (black to red).

Figure 2.3.3.6 Fluorescence intensity at  $25^\circ\text{C}$  normalised using values in the absence of  $\text{Cu}^+$  at pH 5.5 as 1 (folded) and at pH 7.4 (unfolded).  $200\ \text{nM}$  hTeloCFRET in 10 mM sodium cacodylate buffer at pH 5.5 (black) and at pH 7.4 (red). Error bars show standard deviation across three repeats.

Figure 2.3.3.7 Ellipticity at 288 nm of  $10\ \mu\text{M}$  hTeloC at pH 5.5 with  $150\ \mu\text{M}$   $\text{Cu}^{2+}$  as a function of time with three additions of  $150\ \mu\text{M}$  sodium ascorbate under ambient conditions.

Figure 2.3.3.8 CD spectra of a single sample of  $10\ \mu\text{M}$  hTeloC at pH 5.5 (black); addition of  $1\ \text{mM}$   $\text{Cu}^{2+}$  (blue); addition of  $150\ \mu\text{M}$  sodium

---

ascorbate (purple); after 4 h exposure to air (green); chelation using 1 mM EDTA (pink).

Figure 2.3.3.9 Illustration of proposed system for the pH and copper-redox-dependent control of the structure of the i-motif forming DNA sequence hTeloC.

### Chapter 3: Ligands

Figure 3.1.1 Structures of the BCL2 i-motif ligands IMC-48 (left) and IMC-76 (right).

Figure 3.1.2 Structure of the c-myc i-motif selective ligand B19.

Figure 3.2.1 Structures of the 38 compounds selected for further investigation.

Figure 3.2.1.1 Response vs ligand concentration plots for 60339 with hTeloC, c-Myc and double stranded DNA.

Figure 3.2.2.1 Response vs Injection Number for flow cell 1 (black) and flow cell 2 (red).

Figure 3.2.2.2 Corrected Response vs Time where flow cell 1 response is subtracted from flow cell 2 response. Dashed orange lines indicate start and end of injection. Left is pH 7.0 alone. Right shows injection of buffer at different pH values. Colour scale transitions from green to red starting at pH 8 going down to pH 5 in 0.25-unit intervals; pH 7.0 is black.

Figure 3.2.3.1  $\Delta T_m$  of 200 nM FRET-labelled DNA (black square ■ = DAP, red circle ● = DS, blue triangle ▲ = hTeloG) in 10 mM sodium cacodylate buffer at pH 7.0 with 0 to 128  $\mu$ M ligand.

Figure 3.2.4.1 Imino region of 1D 1H NMR spectra of 100  $\mu$ M DAP in 140 mM NaPO<sub>4</sub>, 10 mM MgCl<sub>2</sub>, 5 mM KCl buffer at pH 7.0 with 10% D<sub>2</sub>O at 25°C with increasing concentration of 143491.

Figure 3.2.4.2 Imino region of 1D 1H NMR spectra of 100  $\mu$ M DAP in 140 mM NaPO<sub>4</sub>, 10 mM MgCl<sub>2</sub>, 5 mM KCl buffer at pH 7.0 with 10% D<sub>2</sub>O at 25°C showing effect of incubation with 5 equivalents of 143491 over time.

Figure 3.2.5.1 CD spectra of 10  $\mu$ M DAP (left) and ATXN2L (right) with titration up to 50  $\mu$ M 143491 in 10 mM sodium cacodylate buffer at pH 7.0.

---

*Figure 3.2.5.2 Normalised Ellipticity at 288 nm (top) and normalised absorbance at 295 nm (bottom) of 10  $\mu$ M DAP (left) and ATXN2L (right) without ligand (black) and with 5 equivalents (50  $\mu$ M) 143491 (red). Experiments performed at pH 7.0 in 10 mM sodium cacodylate buffer.*

*Figure 3.2.5.3 Spectra overlap showing shifts in imino proton signals from 1D 1H NMR spectra of 100  $\mu$ M ATXN2L in 140 mM NaPO<sub>i</sub>, 10 mM MgCl<sub>2</sub>, 5 mM KCl buffer at pH 7.0 with 10% D<sub>2</sub>O at 25°C with increasing concentration of 143491.*

*Figure 3.2.5.4 Imino region of 1D 1H NMR spectra of 100  $\mu$ M DNA in 140 mM NaPO<sub>i</sub>, 10 mM MgCl<sub>2</sub>, 5 mM KCl buffer with 10% D<sub>2</sub>O at 25°C with increasing concentration of 143491. JAZF and PDGF-A at pH 7.0 and hTeloC at pH 6.0.*

*Figure 3.3.1.1 Structures of the ligands investigated at acidic pH.*

*Figure 3.3.2.1 Structure of TmPyP4.*

*Figure 3.3.2.2 CD spectra of 10  $\mu$ M ATXN2L with titration up to 50  $\mu$ M ligand in 10 mM sodium cacodylate buffer at pH 7.0.*

*Figure 3.3.2.3 CD spectra of 10  $\mu$ M DAP with titration up to 50  $\mu$ M ligand in 10 mM sodium cacodylate buffer at pH 7.0.*

## **Chapter 4: DNA Nanotechnology**

*Figure 4.1.1 Example NUPACK workflow: A) Dot-parens-plus notation: each unpaired base is represented by a dot, each base pair by matching parentheses, and each nick between strands by a plus; and DU+ notation: a duplex of length x base pairs is represented by Dx and an unpaired region of length x nucleotides is represented by Ux. Each duplex is followed immediately by the substructure that is enclosed by the duplex. If this substructure includes more than one element, parentheses are used to denote scope. A nick between strands is specified by a plus.<sup>166</sup> B) Visual preview of target design entered. C) Computed base identities for target design. D) Equilibrium probability of the base identities forming the target design. E) Depiction of target structure adopted shown using the ideal helical geometry of B-form helices for DNA.*

- 
- Figure 4.1.2 *Illustration of design of DNA tetrahedron from four oligonucleotides consisting of complementary segments (colour coordinated) which hybridise to form the target tetrahedron structure.*
- Figure 4.2.1.1 *Models depicting the i-motif (left) and G-quadruplex (right) of the same DNA sequence from which the double helix (centre) is formed. Black dashed lines show distance between oxygens on the 5' and 3' ends of the molecules. NDB colouring: adenine is red, thymine is blue, guanine is green, cytosine is yellow.*
- Figure 4.2.1 *Proposed multi-state QTET model; orange segments represent added quadruplex forming sequences which can form their respective structure and consequently alter the size of the structure in response to triggers.*
- Figure 4.2.1.1 *CD of candidate quadruplex forming sequences for QTET. A) i-Motif forming sequences at pH 8.03 (dotted) and pH 4.94 (solid). B) G-quadruplex forming sequences without KCl (dotted) and with 100 mM KCl (solid). C) Sequence pairs annealed together. All at 10  $\mu$ M in 10 mM Tris, 5 mM MgCl<sub>2</sub> buffer.*
- Figure 4.2.1.2 *Analysis of double helix formation by QTET quadruplex forming sequences using native PAGE.*
- Figure 4.2.2.1 *Analysis of rapid tetrahedral formation using native PAGE. A) Tetrahedron with no quadruplex forming segments. B) QTET with quadruplex segment in single edge. C) QTET with all six edges containing quadruplex segments.*
- Figure 4.2.2.2 *Analysis of QTET formation with (C2T2)3C2 – (G2A2)3G2 quadruplex segments under different conditions using native PAGE.*
- Figure 4.2.3.1 *Analysis of quadruplex formation using S1' and S2' sequences using native PAGE.*
- Figure 4.2.3.2 *AFM images showing the result of annealing the four constituent strands of the tetrahedron in the presence (top) and absence (bottom) of MgCl<sub>2</sub>.*
- Figure 4.2.3.3 *DLS data of from Goodman's tetrahedron (red) and QTET with quadruplex forming segments in one edge (green) and in all edges (blue).*

---

*Figure 4.2.3.4 Models depicting a DNA tetrahedron. Top-down view from single vertex (left). Side-on view with three vertices from a single face in the same plane (right). Each of the four constituent oligonucleotides is indicated by a single colour.*

*Figure 4.2.3.5 QTET models: top-down view from single vertex (top). Side-on view with three vertices from a single face in the same plane (bottom). 'Open' state with edges formed from double helical DNA (left). 'Closed' state with quadruplex structures formed in the edges (right). Each of the four constituent oligonucleotides is indicated by a single colour; with the quadruplex forming segments in orange. Inset shows magnified central region of edge with the G-quadruplex (green) and the i-motif (yellow) formed.*

*Figure 4.2.3.6 Single edge of QTET with shown with the i-motif (yellow) and the G-quadruplex (green) structures formed (centre) and as part of the double helix (top and bottom). Black dashed lines show distance between oxygens on the 5' and 3' ends of the molecules.*

*Figure 4.2.3.7 CD spectra of A) i-motif forming segment (C4T)3C4 and S1'-(C4T)3C4 at pH 8 (dotted) and pH 5 (solid). B) G-quadruplex forming segment (G4A)3G4 and S2'-(G4A)3G4 without KCl (dotted) and with 100 mM KCl (solid). C) Tetrahedra and QTET under different conditions: pH 8 (solid), pH 8 + 100 mM KCl (dotted), pH 5 (dashed), pH 5 + 100 mM KCl (dot-dash).*

## **Chapter 5: Discussion and Future Work**

*Figure 5.2.1 Ligandable pockets (red) identified on hTeloC i-motif model using fpocket. Panels on the focus on each ligandable pocket with the i-motif's loops in cyan and the cytosine tracts in magenta.*

## **Appendix**

*Figure A.9.1 Normalised ellipticity at 288 nm of 10  $\mu$ M ATXN2L without ligand (black) and with five equivalents (50  $\mu$ M) ligand (red).*

*Figure A.9.2 Normalised ellipticity at 288 nm of 10  $\mu$ M DAP without ligand (black) and with five equivalents (50  $\mu$ M) ligand (red).*

- 
- Figure A.10.1 Normalised absorbance at 295 nm of 10  $\mu$ M ATXN2L without ligand (black) and with five equivalents (50  $\mu$ M) ligand (red).*
- Figure A.10.2 Normalised absorbance at 295 nm of 10  $\mu$ M DAP without ligand (black) and with five equivalents (50  $\mu$ M) ligand (red).*
- Figure A.11.1 Dose-response curves from FID assays of DAP (1  $\mu$ M) at pH 7.0 in 10 mM sodium cacodylate buffer with ligands. Error bars show standard deviation across three repeats.*

---

## List of Tables

### Chapter 3: Ligands

- Table 3.2.1.1  $\%R_{max}$  calculated for each ligand-DNA pair. [ligand] = 50  $\mu$ M  
Running buffer: 10 mM sodium cacodylate, 100 mM NaCl, 0.05%  
tween at pH 5.5.
- Table 3.2.1.2  $K_D$  calculated for each ligand-DNA pair at pH 5.5.  $\times$  = not  
determinable.
- Table 3.2.2.1  $K_D$  calculated for each ligand-DNA pair at pH 7.0.  $\times$  = not  
determinable.
- Table 3.2.5.1  $T_m$  of DAP and ATXN2L measured by CD and UV in the absence  
and presence of 5 equivalents of **143491**.
- Table 3.3.1.1  $\Delta T_m$  of ligand-DNA pairs measured by CD and UV melting using 5  
equivalents of each ligand. All experiments were performed in  
duplicate and  $\Delta T_m$  values are reported as the mean. Errors were  
 $\pm 0.5^\circ\text{C}$ .  $\times$  = not determinable.
- Table 3.3.1.2  $DC_{50}$  values for ligand-DNA pairs at pH 4.3 and pH 5.7 determined  
using the FID assay. Experiments performed in triplicate and  $DC_{50}$   
values are reported as the 50% displacement value calculated  
from fitted dose response curves. Standard errors are calculated  
using R-square values from the statistics on the data fit.
- Table 3.3.2.1  $T_m$  is the midpoint of the transition from each CD melting  
experiment, and SE is calculated using R-square values from the  
statistics on the data fit.  $\Delta T_m$  is the difference between the  $T_m$  of  
the DNA in the presence of 5 equivalents of each ligand and the  
DNA on its own.  $\times$  = not determined.
- Table 3.3.2.2  $T_m$  is the midpoint of the transition from each melting experiment.  
 $\times$  = not determined. % Hypochromicity is the decrease in ellipticity  
observed at 288 nm upon the addition of five equivalents of each  
ligand.
- Table 3.3.2.3  $T_m$  is the midpoint of the transition from each UV melting  
experiment, and SE is calculated using R-square values from the  
statistics on the data fit.  $\Delta T_m$  is the difference between the  $T_m$  of

---

*the DNA in the presence of 5 equivalents of each ligand and the DNA on its own. x = not determined.*

*Table 3.3.2.4 Ligand DC<sub>50</sub> values for DAP determined using FID assay. Experiments performed at least in triplicate, and DC<sub>50</sub> values are reported as the 50% displacement value calculated from fitted dose-response curves. Standard errors are calculated using R-square values from the statistics on the data fit. x = not determined.*

## **Chapter 4: DNA Nanotechnology**

*Table 4.2.3.1 Hydrodynamic diameter of Goodman's tetrahedron, Q<sub>TET</sub> with quadruplex forming segment in one edge and in all edges under different annealing conditions measured by DLS.*

## **Chapter 6: Experimental**

*Table 6.1.1 Custom oligonucleotide sequences used throughout this research.*

*Table 6.1.2 Custom oligonucleotides sequences with end-modifications used throughout this research for FRET and SPR experiments.*

*Table 6.1.3 Custom oligonucleotide sequences used in this research as potential quadruplex segments for Q<sub>TET</sub>.*

## **Appendix**

*Table A.2.1 %R<sub>max</sub> calculated for each ligand-DNA pair; colour coded scale transitioning from green to red indicating lowest to highest %R<sub>max</sub> for each DNA structure respectively. [ligand] = 50 µM Running buffer: 10 mM sodium cacodylate, 100 mM NaCl, 0.05% tween at pH 5.5.*



---

## List of Abbreviations

<b>Å</b>	Ångstrom
<b>A</b>	Adenine
<b>AFM</b>	Atomic force microscopy
<b>Ag</b>	Silver
<b>C·C<sup>+</sup></b>	cytosine-cytosine <sup>+</sup>
<b>C</b>	Cytosine
<b>CD</b>	Circular dichroism
<b>CEITEC</b>	Central European Institute of Technology
<b>CryoEM</b>	Cryo-electron microscopy
<b>Cu<sup>+</sup></b>	Cuprous ion
<b>Cu<sup>2+</sup></b>	Cupric ion
<b>DETC</b>	Diethyldithiocarbamate
<b>DMSO</b>	Dimethyl sulfoxide
<b>DNA</b>	Deoxyribonucleic acid
<b>EDTA</b>	Ethylenediaminetetraacetic acid
<b>EtOH</b>	Ethanol
<b>FAM</b>	6-Carboxyfluorescein
<b>FID</b>	Fluorescent indicator displacement
<b>FRET</b>	Förster resonance energy transfer
<b>G</b>	Guanine
<b>HEG</b>	Hexaethylene glycol
<b>HPLC</b>	High performance liquid chromatography
<b>MD</b>	Molecular dynamics
<b>MeOH</b>	Methanol
<b>NDB</b>	Nucleic Acids Database
<b>NMR</b>	Nuclear magnetic resonance
<b>NUPACK</b>	The Nucleic Acid Package
<b>PAGE</b>	Polyacrylamide gel electrophoresis
<b>PDB</b>	Protein Data Bank
<b>PEG</b>	Polyethylene glycol
<b>Q<sub>TET</sub></b>	Quadruplex tetrahedron
<b>SPR</b>	Surface plasmon resonance
<b>T</b>	Thymine
<b><i>T<sub>m</sub></i></b>	Melting temperature
<b><math>\Delta T_m</math></b>	Change in melting temperature
<b>TO</b>	Thiazole orange
<b>TDS</b>	Thermal difference spectra
<b>TAMRA</b>	6-Carboxytetramethyl rhodamine
<b>UV</b>	Ultraviolet

---

## Published work within this thesis

- (1) Abdelhamid, M. A. S.; Fábíán, L.; MacDonald, C. J.; Cheesman, M. R.; Gates, A. J.; Waller, Z. A. E. Redox-Dependent Control of i-Motif DNA Structure Using Copper Cations. *Nucleic Acids Res.* **2018**, 46 (12), 5886–5893.
- (2) Pagano, A.; Iaccarino, N.; Abdelhamid, M. A. S.; Brancaccio, D.; Garzarella, E. U.; Di Porzio, A.; Novellino, E.; Waller, Z. A. E.; Pagano, B.; Amato, J.; *et al.* Common G-Quadruplex Binding Agents Found to Interact With i-Motif-Forming DNA: Unexpected Multi-Target-Directed Compounds. *Front. Chem.* **2018**, 6 (July), 1–13.
- (3) Abdelhamid, M. A. S.; Gates, A. J.; Waller, Z. A. E. Destabilization of i-Motif DNA at Neutral PH by G-Quadruplex Ligands. *Biochemistry* **2018**, acs.biochem.8b00968.



---

# **Chapter 1: Introduction**

## 1.1 DNA Secondary Structure

A remarkable suggestion was made in 1953 when James Watson and Francis Crick, relying on the experimental results of Rosalind Franklin, assimilated nearly a century's worth of research efforts to propose that DNA exists in a three-dimensional double-helical conformation with the bases on the inside of the helix and the phosphates on the outside. In their seminal paper they presented their ground-breaking deduction that these two helices are held together by pairs of bases, consisting of a purine and a pyrimidine base, and that only specific combinations of these pairs would be able to bond together to form this right-handed structure, concluding that those pairs must be adenine with thymine and guanine with cytosine.<sup>1</sup>

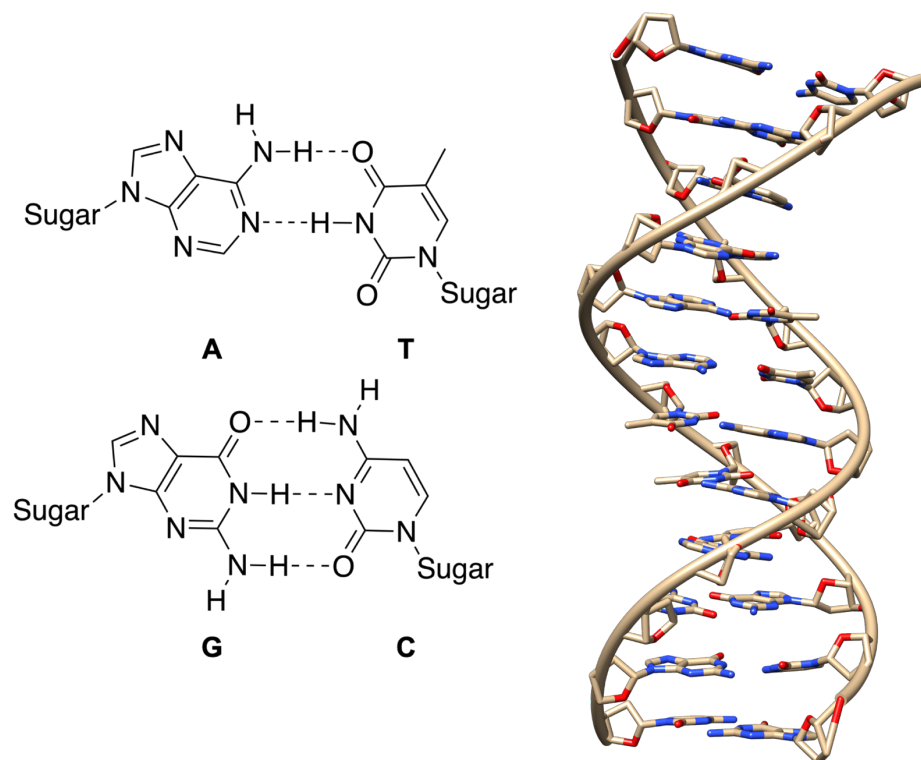
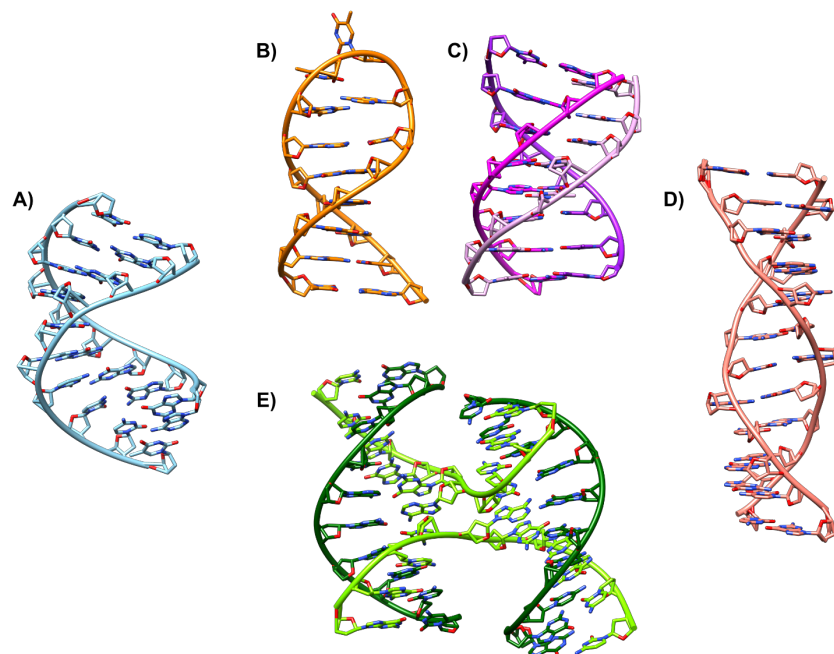


Figure 1.1.1 Watson and Crick base pairing and the structure of B-form DNA. PDB ID: 1BNA.

This historic discovery was a revolutionary moment in the determination of the future direction of biological studies. The debate between a vitalist approach to biology, where the *whole organism* was the object of study, and a reductionist approach, where the objective was to decipher the *fundamentals* of systems, was

effectively eliminated as this discovery propelled the field in the direction of reductionism and molecular biology.<sup>2</sup> As a result, a veritable wealth of information has been discovered in the intervening time about the fundamental aspects of the structure and function of DNA. The original double helix model proposed by Watson and Crick was found to be the most common conformation in living cells and was labelled B-DNA.<sup>3</sup> Further conformations of double helical DNA have since been identified including: A-DNA, a shorter and wider conformer, and Z-DNA, a left-handed helix.<sup>4</sup>

It is now also known that under particular conditions the strands of DNA that make up these helices can adopt a multitude of different structures such as hairpins, triplexes, cruciforms, junctions and quadruplexes. These different structural conformations considerably increase the capacity of DNA to contain and coordinate the information which it holds. The ability of DNA to form these structures is finely regulated by the environment and the nucleotide sequence, and conformational changes result in effects on the DNA's physical, physiological and pathological characteristics.<sup>5</sup>



*Figure 1.1.2 Examples of some alternative DNA secondary structures. A) A-DNA (PDB ID: 440D). B) Hairpin (PDB ID: 1AC7). C) Triplex (PDB ID: 1D3X). D) Z-DNA (PDB ID: 4OCB). E) Four-way junction (PDB ID: 467D).*

---

## 1.2 Quadruplex DNA

Particular research efforts of late have been devoted to studying the characteristics of quadruplex DNA structures. Two principal quadruplex structures have come to the fore as targets for investigation owing to an array of interesting properties: the G-quadruplex and the i-motif. The G-quadruplex is formed by sequences containing stretches of consecutive guanine nucleotides, and the i-motif is dependent on a sequence with cytosine base repeats, both with interspersed loop-forming regions.

The G-quadruplex is a tetrameric structure, the formation of which requires the presence of cations. Four guanine bases form a tetrad through Hoogsteen hydrogen bonding, these tetrads are then able to stack thereby forming a channel-like region in which the localisation of cations is necessary to stabilise the quadruplex.<sup>6</sup> The i-motif on the other hand derives its name from the intercalation of the cytosine bases that occurs, as it is the only known nucleic acid structure formed by the systematic intercalation of bases.<sup>7</sup> i-Motif formation has been associated with the necessity of acidic conditions due to the intercalation requiring a hemi-protonated cytosine-cytosine<sup>+</sup> (C·C<sup>+</sup>) base pair, however it has also been shown to form at neutral pH.<sup>8</sup>

While it has been assumed that these structures would be similarly localised in the genome, hypothetically in the same place on opposite strands due to the complementarity of guanine and cytosine,<sup>9,10</sup> the different requirements for the formation of each quadruplex make it so that this is not necessarily the case.<sup>11</sup> Some evidence suggests that the concurrent formation of both structures on complementary strands, which are capable of forming either structure independently, is prevented by steric hindrance between the quadruplexes.<sup>12</sup> Although examples also exist where the quadruplexes, when slightly offset, can co-exist on complementary strands; with some evidence even that the formation of the first encourages the formation of the second.<sup>13</sup>

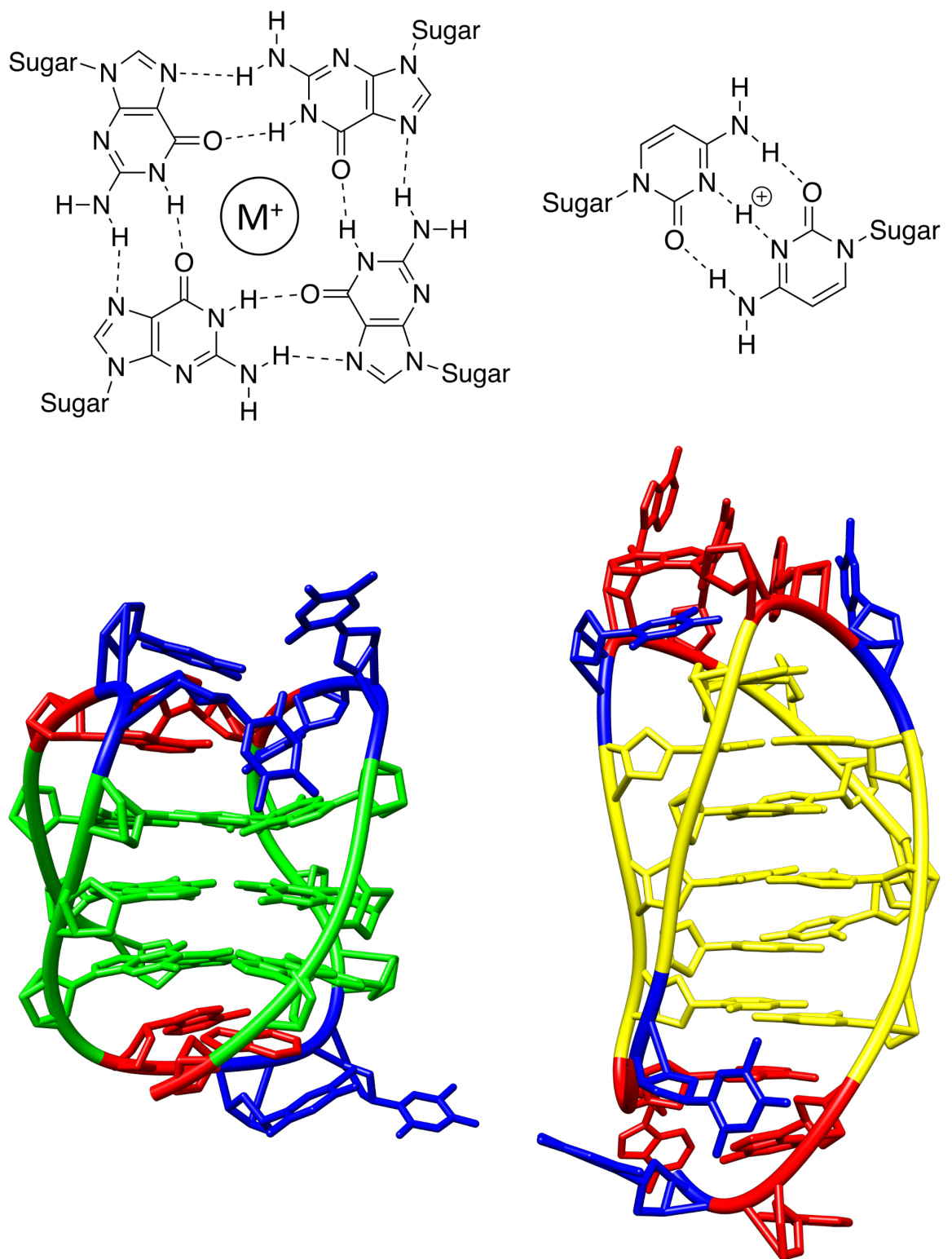


Figure 1.2.1 (Top) Hoogsteen hydrogen bonding of G-tetrad with stabilising cation in the core (left) and a hemi-protonated cytosine-cytosine<sup>+</sup> base pair (right). (Bottom) G-quadruplex (PDB ID: 143D; left) and i-motif (PDB ID: 1ELN; right). Nucleic Acid Database (NDB) colouring: adenine is red, thymine is blue, guanine is green, cytosine is yellow.



---

### 1.2.1 The G-Quadruplex

The more prominent quadruplex of the two in terms of research output is undoubtedly the G-quadruplex. The arrangement of the quartet of Hoogsteen hydrogen bonded guanine bases that form the tetrads of the G-quadruplex was originally proposed in 1962 after observations that concentrated solutions of guanylic acid formed gels.<sup>14</sup> In the 1980s it was found that specific segments of telomeric DNA formed the same structure *in vitro*.<sup>15,16</sup> Since then, the G-quadruplex has transformed from an *in vitro* structural curiosity to an important *in vivo* regulator of biological function.<sup>17</sup> An already substantial body of evidence for the significance of the G-quadruplex is growing and garnering further interest due to the implications raised by the location and effects being discovered.

Various studies have attempted to quantify the potential regions in the human genome that can form G-quadruplexes. Initial *in silico* methods suggested >375,000 sequences that have the potential to form a G-quadruplex,<sup>18,19</sup> while a high-throughput sequencing of the human genome 'G4-seq' identified >700,000.<sup>20</sup> More recent work by Mergny and co-workers discusses the challenges surrounding the accurate determination of the number G-quadruplexes that form *in vivo*, and suggests a two to tenfold higher number of G-quadruplex forming sequences in the human genome than previously proposed, potentially reaching 8 million.<sup>21</sup>

More important than determining the abundance of G-quadruplexes is evaluating the implications of their existence on biological processes. The locations of potential G-quadruplex forming sequences from the genomes of divergent organisms have been found to be non-random and conserved: evolutionary pressure has directed the retention of G-quadruplex forming sequences to specific functional regions.<sup>22</sup> In the human genome, G-quadruplexes were found to be particularly enriched in the telomeric region<sup>23</sup> and in the promoter regions of genes,<sup>24</sup> Additionally, G-quadruplexes have been found in bacteria,<sup>25</sup> and in human RNA and DNA viruses.<sup>26,27</sup> They have been shown to play important regulatory roles in biology,<sup>17,28</sup> and to be involved in cellular pathways for gene expression<sup>29</sup> including transcription<sup>30</sup>, translation<sup>31</sup> and epigenetics,<sup>32,33</sup> and in the maintenance of genome integrity.<sup>34</sup>

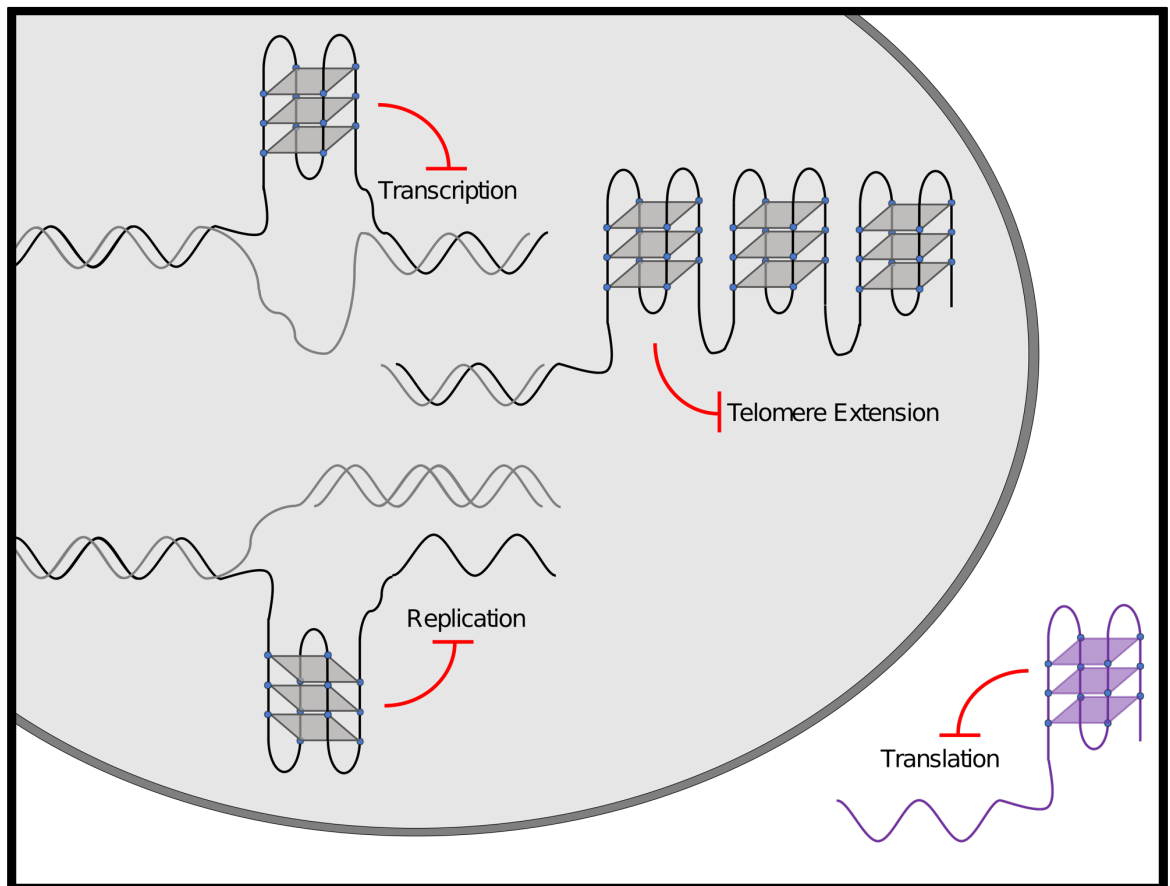


Figure 1.2.2 Illustration of biological processes in cells which have been found to potentially be regulated by G-quadruplexes.

### 1.3 The i-Motif

A similar, albeit more limited in number, body of evidence has been accumulating for the 'other' quadruplex structure, the i-motif.<sup>35</sup> As far as has been revealed thus far, the location and effects of the i-motif are apparently just as interesting as those of the G-quadruplex. Unfortunately, the number of sequences in the genome with the propensity to form i-motif is not clearly known. The primary cause for the discrepancy in research output was the i-motif's requirement for hemi-protonated cytosine base pairs giving rise to the assumption that it necessitated acidic pH and would thus not be physiologically relevant.<sup>36</sup> This was later shown to clearly not be the case and several examples of genomic DNA sequences which form i-motif structures at neutral pH have been presented.<sup>8,37</sup>

The term *i-motif* was coined in 1993 by Gehring, Leroy and Guéron, who, using a combination of NMR, PAGE and molecular dynamics (MD), presented the structure for the tetrameric complex formed by four strands of 5'-d(TC<sub>5</sub>). They found that the multimeric structure formed readily at higher concentrations (> 1 μM) and acidic pH. The quadruplex they observed was highly symmetrical and stabilised by hemiprotonated C·C<sup>+</sup> base pairs. They determined that the structure was composed of two parallel duplexes intercalated in an antiparallel orientation to form the *i-motif*.<sup>36</sup>

Subsequently, much has been learned about the diverse nature of the *i-motif* structure. The first examples identified, such as that discovered by Gehring *et al.*, were intermolecular structures from four separate cytosine-rich strands (tetramers).<sup>36,38,39</sup> This was shortly followed by the identification of examples formed by the intercalation of two hairpin structures (dimers)<sup>40</sup> and of intramolecular *i-motifs* (monomers),<sup>41</sup> all of which are stabilised by the same C·C<sup>+</sup> pairing.

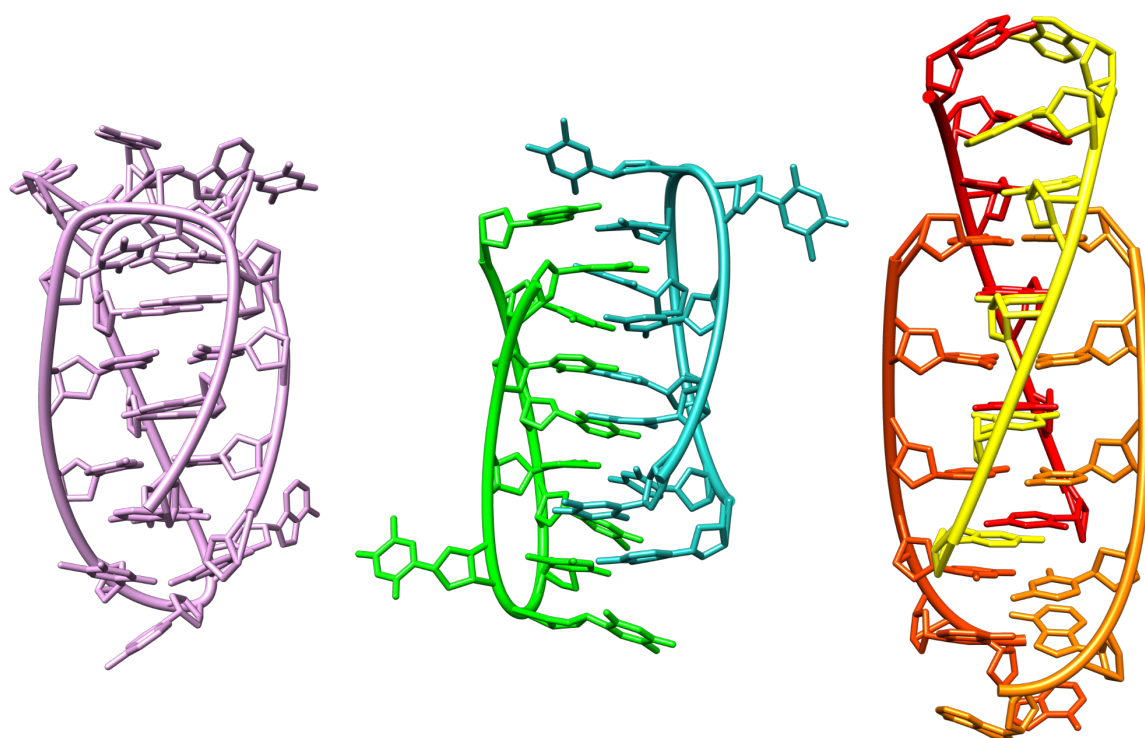


Figure 1.3.1 Monomeric (PDB ID: 1ELN), dimeric (PDB ID: 2MRZ) and tetrameric (PDB ID: 1YBL) *i-motifs*.

---

Despite the suggestion by Gehring *et al.* that, while the tetrameric structure they reported was not stable at neutral pH, sequences containing cytosines might be able to adopt the i-motif structure at “physiological pH and be biologically important,”<sup>36</sup> scepticism surrounding the biological relevance of a structure that appeared to only exist at acidic pH delayed investigation into the biological functions of the i-motif.<sup>11</sup> Consequently, a large proportion of i-motif research focused on using it as a material in nanotechnological applications. Pursuits in this direction yielded a great deal of information about the structure of the quadruplex as well as a variety of very interesting applications.<sup>42</sup>

One of the earliest examples of a potential biological role for i-motifs was a result of interesting effects being uncovered during the concurrent investigation of the G-quadruplex. Human telomeric DNA, as well as that of several other organisms, is composed of a disparate distribution of guanines and cytosines, on opposing strands. In humans, this distribution consists of repeats of 5'-d(TTAGGG) and the complementary 5'-d(CCCTAA),<sup>43</sup> which form G-quadruplex and i-motif respectively.<sup>44</sup> The telomeric region of DNA and a host of associated proteins are essential in the faithful replication of cells' genetic material and protection against DNA damage, thereby ensuring genomic stability and integrity.<sup>45</sup> Dysfunctional maintenance of telomeric DNA, due to elevated levels of the enzyme telomerase's function, has been linked with 85-90% of cancers.<sup>46,47</sup> As such, the existence of both the G-quadruplex and the i-motif in this region heightened interest into the role the quadruplexes could play in inhibiting telomerase activity.

In 1991 Zahler *et al.* showed that telomerase activity could be inhibited by stabilisation of, what they referred to at the time as, the G-quartet by potassium ions<sup>48</sup> (N.B. this precedes the discovery of the i-motif). What followed in this example, illustrative of the overall state of the 'competing' quadruplex fields, was a rapidly growing arsenal of G-quadruplex stabilising ligands that inhibited telomerase expression,<sup>49</sup> which by 2009 included three that had been tested *in vivo*.<sup>50</sup> By contrast, the first example of telomerase inhibition by stabilisation of the i-motif was not reported until 2012. Even this example did not represent a clear 'victory' for the i-motif: single-walled carbon nanotubes were found to cause human telomeric duplex dissociation and while they did in-fact stabilise the i-motif, it was

---

hypothesised by Qu and co-workers that the ensuing formation of G-quadruplexes on the complementary strand was responsible for telomerase inhibition.<sup>51</sup>

Nevertheless, evidence for a biological role for the i-motif continued to grow. Examples of i-motifs in the promoter regions of genes were presented, alongside evidence of their involvement in the control of the transcription of those genes.<sup>52,53</sup> Proteins were found that potentially bound to the i-motif, or with the C-rich single strand,<sup>54,55</sup> with the most notable example being the *BCL2* activating transcription factor hnRNP LL.<sup>56</sup> As with the G-quadruplex before it,<sup>34,57,58</sup> the i-motif began to be considered as a physiologically relevant target.<sup>55,56,59</sup>

While evidence refuting the view that acidic pH was a prerequisite for i-motif formation has been available for some time,<sup>37</sup> the most significant contributions to the acceptance of the i-motif as a physiologically relevant structure were only presented in the past two years. I believe three works in particular will stand at the helm of a fortune of biological research related to the i-motif that is to come. The first, presented by Waller and co-workers, is the identification of genomic DNA sequences which form i-motif at neutral pH.<sup>8</sup> The second, from Trantirek's group, provided the first *in vivo* evidence for the persistence of transfected i-motifs in the nuclei of living human cells.<sup>60</sup> Finally, and most prominently, is Christ's work on the generation and characterization of the i-motif selective antibody (iMab) which showed that the i-motif is formed natively in the nuclei of human cells.<sup>61</sup>

## 1.4 DNA as a Functional Material

With the abundance of information about DNA that is now readily available it is astonishing to realise that this monumental amount of knowledge has almost exclusively been amassed in just over half a century. General agreement that organisms' complex genetic information was held in their DNA rather than their proteins was only reached as recently as the 1950s.<sup>62</sup> Before then, it seemed far more likely that proteins, with their 20 constituent amino acids and greater complexity, rather than DNA, a simple molecule with merely four subunits, contained the "information of life."<sup>62</sup>

---

In the past few decades we have gathered invaluable information from studying nature's extraordinary ability to use the four DNA bases to essentially coordinate all of life. Armed with this knowledge, inspired researchers have thought of ways to exploit the properties of DNA for applications outside of the context of biological systems. Several properties of DNA enable it to be a suitable material for designing highly structured architectures.<sup>63</sup>

- (i) the capacity for self-recognition and self-assembly allows for the 'simple' design of very complex systems
- (ii) technological advances have made it relatively easy to synthesise and amplify
- (iii) it is electroconductive
- (iv) it can be used to store information
- (v) it can recognise and interact with other molecules

Another important characteristic property of DNA is its size: with the 2 nm diameter of the double helix making it the ideal size to design architectures at the nanoscale.

The approach to the design of nanoscale architectures can generally be classified as either "top-down" or "bottom-up."<sup>64</sup> In the top-down approach, the starting point is a relatively large structure (generally microscale) and this is manipulated on the molecular or atomic level to produce the desired nanoarchitecture.<sup>65</sup> In the bottom-up approach, within which most DNA nanotechnology lies, system design begins at the molecular level and constituent molecules are assembled in a manner which exploits their individual interactivities to build the nanoarchitecture.<sup>66,67</sup> The first report of a nanoscale architecture built using DNA was in 1991 by the "Founder of DNA Nanotechnology"<sup>68</sup> Ned Seeman, who had managed to build a "cube-like object."<sup>69</sup> Seeman later explained that from his conception of the idea of building a 3D-connected object from DNA in 1980, it took more than ten years before the first DNA cube was able to be made in the laboratory by Junghuei Chen.<sup>68,69</sup> Enormous progress has been made since then (Figure 4.1.1), and excitingly per Seeman's own assessment in 2015, this field remains in its infancy<sup>70</sup>

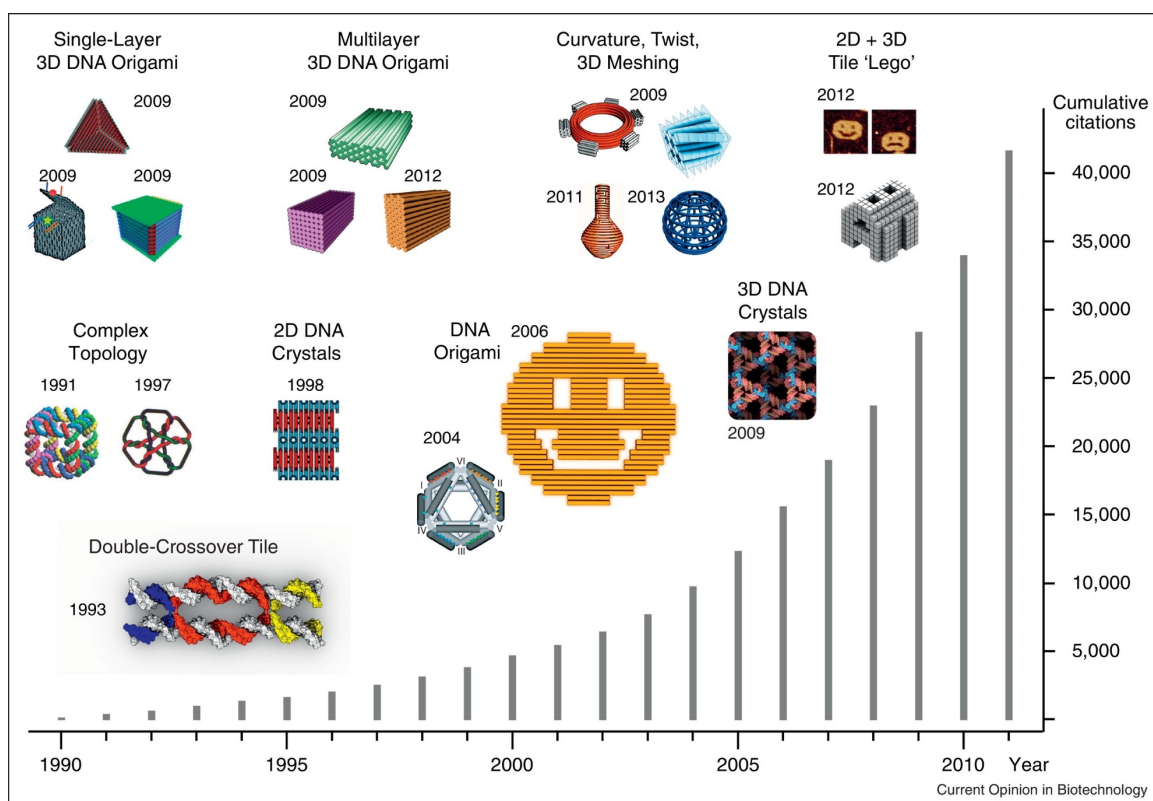


Figure 1.4.1 Illustration of the design space expansion and the growth of interest in structural DNA nanotechnology. Reprinted from *Current Opinion in Biotechnology*, 24, Beikko Linko, Hendrik Dietz, *The enabled state of DNA nanotechnology*, 555-561, Copyright (2013), with permission from Elsevier.<sup>71</sup>

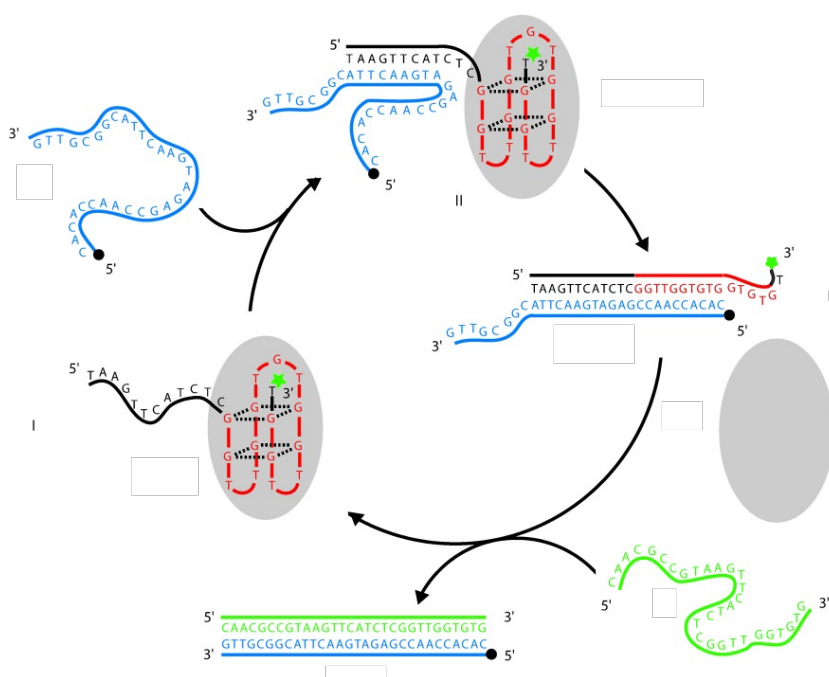
### 1.4.1 Quadruplex DNA Nanotechnology

Notwithstanding the aforementioned properties of DNA which make it a suitable material for building DNA nanoassemblies, there are some limitations to nanostructures designed using only the classical Watson-Crick base pairs such as: susceptibility to enzymatic degradation, poor resistance to heat and/or denaturing reagents, flexibility and deformability, and low sensitivity to chemical stimulation. The incorporation of quadruplexes into DNA based nanoarchitectures has the potential to overcome some of these limitations inherent to B-DNA helices.<sup>72</sup>

Both quadruplex structures have the advantage in terms of enabling the building of more rigid structures: B-DNA duplexes are too flexible to enable the building of stable nanostructures.<sup>72</sup> On the other hand, the extended backbone of i-motifs and

the stacking of the quartets in G-quadruplexes improve rigidity and stiffness relative to the duplex.<sup>73</sup> G-quadruplexes have been also found to be resistant to conditions which typically denature DNA,<sup>74</sup> and the formation of both quadruplexes is favoured under conditions of molecular crowding which, again, destabilises B-DNA.<sup>75–78</sup> In addition to these benefits, both quadruplexes have the upper hand in their ability to be more responsive to chemical stimuli: both have been characterised to have specific responses and changes in formation and stability with different cations,<sup>6,79–82</sup> and the i-motif has “exquisite” sensitivity to pH.<sup>72,83,84</sup>

Many examples of nanoarchitectures incorporating a quadruplex have been developed including sensors,<sup>85–87</sup> walkers,<sup>88</sup> molecular motors and beacons<sup>67</sup> and logic gates;<sup>89,90</sup> as well as systems incorporating both quadruplexes.<sup>91</sup> An example application for a nanodevice which uses a G-quadruplex is shown in Figure 1.4.1.1 below. This “molecular machine” can be used to go through cycles of binding and releasing the human blood-clotting factor  $\alpha$ -thrombin in a controllable manner.<sup>92</sup>



*Figure 1.4.1.1 Representation of the operation cycle of the aptamer-based molecular machine in the presence of thrombin. Reprinted from *Angewandte Chemie International Edition*, 43, Wendy U. Dittmer, Andreas Reuter, Friedrich C. Simmel, *A DNA-Based Machine That Can Cyclically Bind and Release Thrombin*, 3550-3553, Copyright (2004), with permission from Wiley.*



---

## 1.5 Techniques for Studying Quadruplexes

A variety of techniques can be employed to examine quadruplexes, the choice of which is largely dependent on the system under investigation. What follows is a brief introduction and discussion on the techniques used in the investigations presented in this thesis.

### 1.5.1 Ultraviolet (UV) Spectroscopy

UV spectroscopy can be used to monitor the interaction between quadruplexes and cations or ligands. When using this technique, attention must be paid to the wavelengths under examination when making observations about the different quadruplexes, especially in the presence of ligands which may have their own characteristic absorptions. Using UV spectroscopy and performing melting experiments is also a valuable tool in gathering information on the stability of quadruplexes under different conditions. For the i-motif, absorbance at 260 nm may be used, but typically 295 nm is used to monitor the stability of the quadruplex and determine a melting temperature for the secondary structure.<sup>93</sup> This is done by mathematical fitting of the observed hypochromicity at this wavelength in response to the increase in temperature denaturing the structure. The midpoint of the transition observed can then be calculated and this is reported as the melting temperature. 'UV-difference' spectra can also be calculated and used to further elucidate the effect of a given variable on the structure of quadruplexes. Valuable work was done by Mergny *et al.* that showed that the thermal difference spectra (TDS) of a given "structural family are strikingly similar."<sup>94</sup> Thus, the presence of a major positive TDS peak at  $239 \pm 1$  nm and a negative peak at  $\sim 295$  nm can also be used to determine the adoption of the i-motif structure by a sequence in a given set of conditions.

### 1.5.2 Circular Dichroism (CD) Spectroscopy

Using circularly polarised light characteristic signals can be observed in the spectra which can be used to determine the presence or absence of the quadruplex

---

structures. CD spectroscopy offers an upgrade on UV spectroscopy in the examination of quadruplex structures; it can be used to differentiate between parallel (dominant positive peak at 260 nm) and anti-parallel (negative band at 260nm and positive band at 295nm) G-quadruplexes, and to determine the adoption of i-motif by a given sequence (characteristic spectrum with a dominant positive peak at 288 nm).<sup>95</sup> CD can also be used to monitor the effects of changing conditions, or the addition of cations or ligands, to quadruplexes. Also, as with UV, melting experiments can be performed to gather further information on the thermal stability of the system under investigation.

### **1.5.3 Nuclear Magnetic Resonance (NMR) Spectroscopy**

NMR is a powerful technique and can be used to provide complex atomic-level details on the most intricate molecular interactions to aid in the determination of a quadruplex's structure.<sup>96</sup> It can also however be used in a more fundamental fashion to establish the existence, and determine the type, of structure present in solution. Using 1D <sup>1</sup>H NMR, the hemi-protonated C<sup>+</sup>-C base pairs in an i-motif give rise to characteristic imino proton resonance at ~15.5 ppm.<sup>97</sup> These signals can be used to monitor the i-motif, and how the structure responds to the introduction of different elements to its environment.

### **1.5.4 Fluorescence Resonance Energy Transfer (FRET) and FRET-melting**

FRET is a physical phenomenon where the fluorescence of one fluorophore (the donor), when in sufficient proximity, can transfer through non-radiative means its energy to another (the acceptor). This system is employed in the study of quadruplexes by the attachment of two fluorophores at the 5' and 3' ends of a quadruplex forming sequence, in the case of the work presented here these were 6-carboxyfluorescein (FAM) and 6-carboxytetramethylrhodamine (TAMRA) respectively. Using this system, the fluorescence emission spectra can be measured across a range (500 – 750 nm) to monitor both fluorophores and gather information about the distance between them. Alternatively, by exploiting the FRET phenomenon and designating the acceptor as a 'quencher' of the donor's

---

fluorescence when they are sufficiently close to one another, the fluorescence of the donor can be monitored independently.<sup>98</sup> Thus, in an appropriately designed system, when the quadruplex is formed the donor's fluorescence is quenched by the proximate fluorophore. In FRET-melting experiments subsequent thermal denaturation of the quadruplex will result in the fluorophores moving apart and an increase in the donor's fluorescence. The resulting transition can be monitored to determine a melting temperature for the system under investigation.

### **1.5.5 Fluorescent Indicator Displacement (FID) Assays**

FID assays rely on the enhancement of the fluorescence of a 'probe' when bound to a DNA secondary structure. The interaction of the probe thiazole orange (TO) was characterised with the i-motif,<sup>99</sup> and it has been used to perform competition assays with the structure to measure the relative affinities of ligands which can competitively displace the probe.<sup>100,101</sup> TO is first allowed to equilibrate with the quadruplex structure to allow for the fluorescence enhancement to reach a steady state. Aliquots of ligands which can potentially bind to the structure are then added, and if they bind to the structure, and that binding results in displacement of the probe, their relative affinity for the structure can be determined.

### **1.5.6 Surface Plasmon Resonance (SPR)**

SPR is a technique that allows determination of interactions between target structures which are immobilised on the surface of a chip and samples of analytes in solution. Biotinylated quadruplex-forming DNA can be immobilised on the surface of a streptavidin-coated gold SPR chip. Polarised light shone through a prism is reflected off the surface opposite to that on which the DNA is immobilised: plasmon resonance from the gold on the chip's surface results in the reflection of light at two different angles which are monitored by a detector. When a sample is introduced that binds to the DNA on the surface of the chip this affects the resonance of the gold plasmons and consequently the angle of reflected light monitored by the detector.<sup>102</sup> Real-time monitoring of sensorgrams which are produced as a result of changes in this angle allows for the detection of binding events. Adequate experimental design and some mathematical analysis of the

---

changes in the sensorgram allow the determination of binding affinity and kinetics data for the system under investigation.

### **1.5.7 Gel Electrophoresis**

Gel electrophoresis can be used to determine the formation of higher-order multimeric structures by DNA sequences, or to study their formation of monomeric secondary structures. The migration of a sample through a gel is determined by its size, mass and charge. Generally, species with the smallest mass and the highest charge will move through a gel most rapidly. However, a species of the same mass, and the same global charge, will migrate more rapidly through a gel if it adopts a more compact structure, such as a quadruplex. Gel electrophoresis can be performed under 'native' conditions, i.e. conditions that should not disrupt any higher-order architectures that may have formed. Or, depending on the characteristic under investigation, denaturing conditions that will remove such structures. Using gel electrophoresis with chemical footprinting methods using radio-label DNA sequences has provided a powerful method to examine which bases are involved in the formation of a quadruplex structure.<sup>12</sup> Unfortunately the radioactive element of this method precludes it from being widely adopted by most researchers. The use of fluorescently labelled DNA as an alternative has recently been reported and would be technically more accessible.<sup>103</sup>

### **1.5.8 Molecular Modelling and Dynamics**

Molecular modelling and dynamics simulations have been essential in understanding the structure of quadruplexes, and especially the i-motif. As the pursuit of an i-motif crystal structure continues solution techniques have been combined with increasingly powerful and constantly developing MD techniques to propose structures for the unimolecular i-motif.<sup>82,96,104–106</sup> Only very recently have force field parameters become available that enable dynamic simulation, as opposed to fixed parametrization, of the protonated cytosines at the core of the i-motif.<sup>107,108</sup> These parameters are a considerable improvement however they remain inadequate: while they have enabled MD measurements showing an enhanced stability for the hemi-protonated i-motif structure, when allowed to run

---

for 'extended' periods of time (~40 ns) the structure spontaneously unfolds under conditions which have been shown otherwise to retain the i-motif.<sup>108</sup> While virtual drug screening has been performed with quadruplexes since the 1980's, owing to the complexities of modelling the non-canonical DNA structures, there was undoubtedly a preference for biophysical high-throughput screening methods which offered more success. Recently more success has been achieved with virtual drug screens targeting the G-quadruplex,<sup>109</sup> however no such example has yet been reported for the i-motif.

## 1.5 Challenges and Opportunities in i-Motif Research

Many of the challenges facing the 'ugly duckling'<sup>35</sup> quadruplex (the i-motif) simultaneously present rewarding opportunities for whomever can resolve them; such should be the approach to all challenges in research.

One of the most significant obstacles since the discovery of the i-motif is the scepticism surrounding its biological relevance. The opportunity to address this issue has been, and continues to be, tackled by many research groups and this scepticism, while at one point understandable, now stands contrary to substantial evidence. The opportunity that lies in investigating the biological effects of i-motifs is the impact those who do so stand to potentially make in controlling diseases or developing therapeutics using the i-motif. The role of the G-quadruplex as a therapeutic target has been the subject of much discussion, spurred by the progression of candidates into preclinical testing and clinical trials.<sup>57,110</sup> The opportunity for a small molecule therapeutic that targets i-motif is a monumental one, the development for use in the clinic of such a therapy is a pursuit that will undoubtedly take some time yet, however the opportunity presented by this challenge is a tantalising one.

Interestingly, one opportunity that has already been seized as a result of the doubts surrounding the i-motif's biological relevance is its use as a material for nanotechnological applications.<sup>42</sup> One of the earliest examples, reported in 2003, was a proton-fuelled i-motif nanomotor.<sup>111</sup> Since then, the i-motif has been exploited in the design of hundreds of pH-driven nanomachines<sup>42,78</sup> including an

---

example of a light-driven pH-jump system<sup>112</sup> and a DNA nanomachine that can map spatial and temporal pH changes in living cells.<sup>113</sup> Studies towards the development of the i-motif for nanotechnological applications have provided significant advances in the understanding of fundamental aspects of the quadruplex's structure, its dynamics,<sup>114–116</sup> and also about the various i-motifs formed by different sequences.<sup>8,117,118</sup> This information has enabled fine-tuning of the properties of these types of devices, and while crucial to the development of these noteworthy applications, it also provides valuable knowledge to researchers exploring the biological role of the i-motif.

Finally, there are gaps in our knowledge surrounding the i-motif that exist which present further challenges to our ability to fully understand and study the structure. Perhaps most prominent among these is the absence of a structure for an unmodified intramolecular i-motif. Intramolecular i-motif solution structures do exist, however due to the overall global symmetry of the quadruplex<sup>119</sup> these have necessitated chemical modification to enable the determination of the structure. Similarly, no X-ray crystallographic data exists for the intramolecular i-motif, which would provide an even higher level of resolution and consequently further information about the structure. While efforts are being made in this regard the ability to 'easily' determine the structure of an i-motif formed by a given sequence remains elusive. This is compounded by the dynamic nature of the i-motif and a debate surrounding the existence of several possible structures in equilibrium for a given sequence in a given environment.<sup>80,82,120</sup>

## **1.6 Aim and Objectives**

The overall aim of this project was to synthesise a novel DNA-based biological nanostructure whose formation, and conformation, could be controlled by exploiting the characteristic properties of both DNA quadruplexes; work on this nanostructure is presented in Chapter 4. What precedes that is the presentation of the research that was performed in order to learn more about the i-motif, this was necessary to generate the knowledge needed to incorporate and control the i-motif in the nanostructure. This primary aim was directed by two objectives: the first was to understand how different cations affected the i-motif so they could be used to

---

control the structure, this is presented in Chapter 2. The second objective was to examine ligands for the same purpose, and this is presented in Chapter 3.

---

## **Chapter 2: Cations**



---

## 2.1 Cations and Quadruplexes

Metal ions are inextricably linked to nucleic acid chemistry and, amongst a plethora of other functions, play a fundamental role in the formation and maintenance of the structure of DNA.<sup>121–123</sup> Nucleic acids are polyanionic by nature: the phosphate groups that link nucleosides together each carry a negative charge that needs to be shielded to enable the formation of stable secondary structures. In biological systems the counterions associated with nucleic acids are normally polyamines,<sup>124–126</sup> or mono- or divalent metal cations.<sup>127</sup> The cations found in physiological environments have been shown to play an essential role in the formation and the maintenance of all higher-order DNA structures including double and triple helices, three-way and four-way junctions, and quadruplexes.<sup>128</sup>

Cations play a particularly important role in G-quadruplexes as the structure cannot exist in the absence of stabilising cations at their core. While a monovalent metal ion is normally required for the formation of the G-quadruplex, examples exist of divalent cations enabling it as well.<sup>129</sup> Changes in the concentration and/or identity of the cation associated with a given G-quadruplex can have significant effects on the stability and indeed the conformation of the structure; with an example G-quadruplex in the presence of  $K^+$  having been shown to have a melting temperature up to 40°C higher than in  $Na^+$ .<sup>130</sup> Astonishingly, changes in only the loop-forming region of that sequence, not the quartets of the quadruplex, can result in it becoming almost insensitive to this change in the cation with its melting temperature only shifting by 1°C under the different conditions.<sup>130</sup> Furthermore, a diverse array of topologies can be adopted by G-quadruplexes, this has also been shown to be controlled by the nature of the cationic environment.<sup>131</sup> The relationship between G-quadruplexes and cations is multivariate however, it has generally been established that the order in which cations are able to stabilise G-quadruplexes is:  $K^+ > Na^+, NH_4^+, Rb^+ \gg Li^+, Cs^+$ .<sup>6</sup>

The relationship between i-motif forming DNA sequences and cations has not been explored as comprehensively. In 1995 Mergny *et al.* showed that increasing the concentration of NaCl from 0 to 100 mM decreased the thermal stability of the i-motifs under investigation over a range of pH values; further increases in the

---

concentration to 300 mM had no effect on the melting temperature of the structures.<sup>41</sup> They also noted that addition of 5 mM MgCl<sub>2</sub>, CaCl<sub>2</sub>, ZnCl<sub>2</sub> or 10 mM LiCl or KCl to their buffer conditions (which contained 100 mM NaCl) did not affect the melting temperature of one of the oligonucleotides under investigation.<sup>41</sup> They attributed these effects to the ionic strength of the solution rather than a direct interaction between the cation and one of the components of the i-motif structure.

In terms of a direct interaction between the i-motif and a cation, Kim *et al.* reported that sub-molar concentrations (50 – 500 mM) of Li<sup>+</sup> had a destabilising effect on the i-motif from the human telomeric region. In this example, the destabilisation by lithium is ascribed to its size: while it is sufficiently small to fit in the ‘pocket’ between two cytosines in an i-motif it cannot duplicate the stabilisation of the structure due to the hydrogen bonding offered by the proton in the C·C<sup>+</sup> base pair.<sup>132</sup> Thus far, the only cation that has been shown to enable i-motif formation is Ag<sup>+</sup>.<sup>80</sup> Another interaction between a cation and an i-motif forming sequence is that of Cu<sup>2+</sup>, it was shown to be capable of re-folding an i-motif forming sequence into a hairpin structure, even competing with the acid-stabilised i-motif at low pH.<sup>81</sup>

## 2.2 Investigating the Effects of Cations on i-Motif DNA

With only limited reports in the literature on the interactions between i-motif forming DNA sequences and cations, and these few examples showing interesting results which could be applied to the control of the i-motif function in the nanostructure to be developed, further investigation into the effects of cations on i-motif DNA were warranted. A list of cations was selected to be tested, based initially on their availability and absence of prior investigation in the literature, these were: aluminium(III), cadmium(II), calcium(II), chromium(II), cobalt(II), copper(II), iron(II), magnesium(II), manganese(II), molybdenum(III), nickel(II), tungsten(IV), vanadium(II), and zinc(II).

### 2.2.1 Tricky Temperature

The i-motif sequence chosen for these investigations was the well-studied human telomeric i-motif (hTeloC) the sequence for this written in the 5' to 3' direction, as

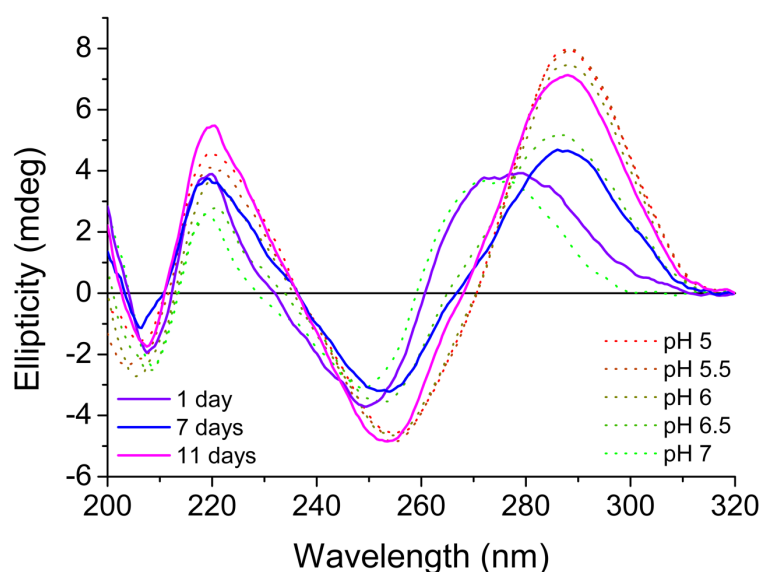
---

for all sequences in this work, is (TA<sub>2</sub>C<sub>3</sub>)<sub>4</sub>. This sequence had been reported to have a transitional pH of 6.11 at room temperature in 10 mM sodium cacodylate buffer with 100 mM sodium chloride.<sup>133</sup> This is the pH at which the population is evenly split between the folded and unfolded states. A lower pH would result in the sequence being predominantly folded into the i-motif and the converse true at higher pH. The aim was to initially use circular dichroism spectroscopy to search for a cation that would induce the formation of the i-motif structure by this sequence at pH 7.0. However, peculiarities arose when after some time the control sample, which was measured each time prior to cation addition, appeared to have folded itself into an i-motif. The sample had been prepared as described commonly in the literature. It was confirmed that the correct buffer was used, and that the pH had not changed. Once the oligonucleotide was diluted to the desired concentration the sample had been held in a heating block at 95°C for five minutes and then allowed to cool to room temperature slowly overnight to ensure uniformity of the population and allow the adoption of the equilibrium structure under those conditions. Finally, the measurements were being collected at room temperature, which was close to the average ~20°C.

The fact that measurements were being taken at room temperature was an important variable that was confounding explanation. All the variables were controlled so that the i-motif forming sequence would be unfolded and the only change that was taking place over time was in the temperature. The samples were stored in a refrigerator (~4°C) between measurements, but enough time was allowed for them to return to room temperature before measurement. This was because the human telomeric i-motif has been shown to form at neutral and slightly alkaline pH at low temperature.<sup>134</sup> However, the melting temperature for the structure at pH 7.0 was 13°C,<sup>134</sup> and in the literature it was evident that above this temperature the structure returned to an unfolded random coil. Therefore, it was assumed that temporary cold storage would have no effect on measurements taken at room temperature.

Accounting for the possibility that some unknown error had been made, new samples were prepared and stored in the refrigerator overnight and when measured the next day at room temperature they appeared to be unfolded.

Eventually however, the same apparently spontaneous i-motif formation reappeared. What transpired to be the likely cause of this was a previously unreported phenomenon. While the i-motif forming sequence behaved initially in a manner congruent with what was reported in the literature, keeping the i-motif forming sequence at a low temperature for extended periods of time enabled the i-motif structure adopted in the low temperature environment to persist at higher temperatures than before.



*Figure 2.2.1.1 CD spectra of 10  $\mu$ M hTeloC in 10 mM sodium cacodylate buffer at room temperature. Dotted lines represent spectra of samples at different pH values measured without storage at low temperature. Solid lines represent spectra of samples at pH 7.0 measured after storage at 4°C for the indicated time.*

Unfortunately, as these observations were initially presumed to be caused by an unidentifiable sample preparation error, and the focus was to find cations that induced i-motif folding, this has not been followed up. A systematic preparation of samples and maintenance at low temperature for increasing periods of time, although time-consuming, would be simple enough to perform. In combination with CD melting experiments (a facility that was not available at the time), and/or UV and FRET melting, this could produce a sufficiently robust dataset to allow the dissemination of this phenomenon to those involved in i-motif research.

---

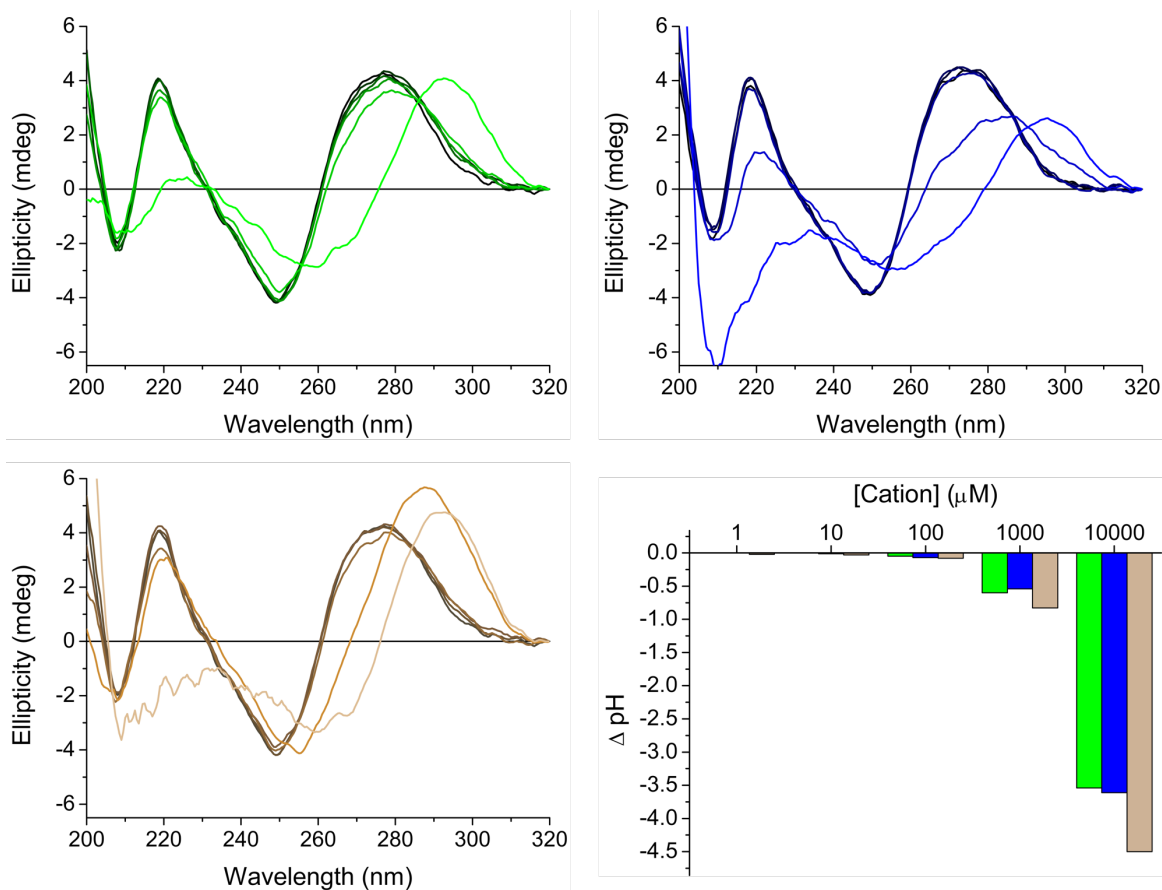
On its own this phenomenon does not present a particularly astounding development, however, there does exist a potential application for this phenomenon. i-Motifs have been used to design a DNA hydrogel which forms at acidic pH. Subunits of DNA are designed in such a way that they have one half of a bi-molecular i-motif forming sequence at their extremities. When the pH is lowered the i-motif forms causing the DNA to aggregate and form a gel, the pH required to form the hydrogel was 5 which precluded potential biomedical applications.<sup>135</sup> Other DNA-based hydrogels have since been developed exploiting other systems but, while they are elegant solutions, these systems rely on more complex interactions and some necessitate multiple steps, or the addition of exogenous components, for hydrogel formation.<sup>136,137</sup> With adequate design, this phenomenon can be exploited to design i-motif based hydrogels with two potential benefits: firstly, formation of the hydrogel at neutral pH and maintenance at higher temperatures could enable biomedical applications; and secondly, hydrogels based on such a system can be designed so that they spontaneously disintegrate after exposure to body temperature once a designated period of time from application has elapsed.

### **2.2.2 Pesky pH**

Once the mystery of the self-forming i-motif was resolved, the investigation into the effects of the cations could proceed. As mentioned, the aim was to find a cation that would induce the formation of the i-motif by the hTeloC sequence at pH 7.0. Initial screens were performed by titrating each cation in steps increasing the concentration ten-fold with each subsequent addition from 1  $\mu$ M to 10 mM. As can be seen in Figure 2.2.2.1, some apparently interesting results were found for aluminium(III), chromium(II) and vanadium(II).

These results indicated that some form of structural reconfiguration was occurring in each of these cases. Additionally, the bathochromic shift observed in the peak ellipticity around 277 nm was congruent with adoption of an i-motif-like structure. Initial excitement about these results was diminished once it was realised that the higher concentration of cations required before a change is observed could suggest an alteration of the pH was occurring. Unfortunately, this turned out to be

the case for the three cations which were the ones to have shown interesting results. Control titrations of the aluminium(III), chromium(II) and vanadium(II) showed that they decreased the pH of 10 mM sodium cacodylate buffer, initially at pH 7.0, by 3.54, 3.61 and 4.50 units, respectively.



**Figure 2.2.2.1** CD titrations of 10  $\mu\text{M}$  hTeloC in 10 mM sodium cacodylate buffer at pH 7.0 (black) with increasing cation concentration in 10-fold increments from 1  $\mu\text{M}$  to 10 mM (coloured) and the change in pH of 10 mM sodium cacodylate buffer from an initial pH of 7 with addition of cation. Aluminium(III) – green, chromium(II) – blue, and vanadium(II) – sand.

Under these altered conditions, well below the transitional pH of hTeloC, it can be expected that the i-motif (or another alternative secondary structure) may form. The remaining cations, which did not change the pH, nor did they show any interesting results in the original CD experiments, were excluded from further investigation. However, to confirm whether the three possibly interesting results observed thus far were due to the respective cation or the change in pH, the buffer

---

was changed to 50 mM sodium cacodylate and the pH control experiment repeated. The higher sodium cacodylate concentration increased the buffering capacity sufficiently so that the pH (except for 10 mM vanadium(II)) remained within 0.20 units of 7.0. Unfortunately, under these new conditions which controlled the pH, no appreciable differences were observed in the CD spectra with hTeloC.

General consensus in the literature is that increasing cationic concentration destabilises the i-motif. Nevertheless, with the knowledge that  $\text{Ag}^+$ ,<sup>80</sup> and  $\text{Cu}^{2+}$ ,<sup>81</sup> showed interesting results of their own these investigations into further cations were performed. The aim of the experiments described here was to find a cation that could induce the formation of the i-motif above its transitional pH. While this original list of candidates failed to do so, this does not necessarily exclude that these cations may have a stabilising effect on the i-motif. It is possible that at, or below, the transitional pH for a given sequence one or more of these cations could have a stabilising effect on the structure which could be observed using melting experiments. However, as this would not be directly applicable to the control of the DNA nanostructure to be designed, they were not performed.

I believe however that these experiments are worth mentioning, and, in the right context, worth doing, as the possible results of these investigations should not be discounted. Alberti *et al.* reported on the “exception that confirms the rule,”<sup>6</sup> A G-quadruplex structure which, contrary to all previous observations – which had led to the establishment of the “rules” – was more stable in sodium than in potassium. Examples such as these are why investigation into these cations could continue, although in a different direction. Any interesting results which may be discovered that stand contrary to the rules should be incentive enough to do so. By exploiting anomalies such as these, members within a class of structures which have their own unique characteristics and behaviours, truly sophisticated architectures can be designed which incorporate these members and exploit their differentiated responses to the benefit of the overall design.

## 2.3 Redox-dependent Control of i-Motif DNA Structure Using Copper Cations

As the screening of the cations discussed in section 2.2 continued to unfold without the discovery of a particularly noteworthy or promising candidate, deeper investigation of the literature yielded a new candidate worthy of investigation:  $\text{Cu}^+$ . Computational studies had found that  $\text{Cu}^+$  can act as a substitute for  $\text{H}^+$  to support formation of cytosine dimers with a similar conformation to the hemi-protonated base pair found in i-motif DNA.<sup>138</sup> Oomens and co-workers used infrared ion spectroscopy in combination with density functional theory (DFT) calculations to show that cytosine monomers in the presence of  $\text{Cu}^+$  form C-Cu<sup>+</sup>-C structures, with an analogous orientation to the hemi-protonated C-dimers at the core of the i-motif.<sup>138</sup> In contrast to alkali metal ions, that induced a different dimer conformation which sacrificed hydrogen-bonding interactions between bases for improved chelation of the metal cation, the C-Cu<sup>+</sup>-C dimer complex was proposed to be stable.<sup>138</sup>

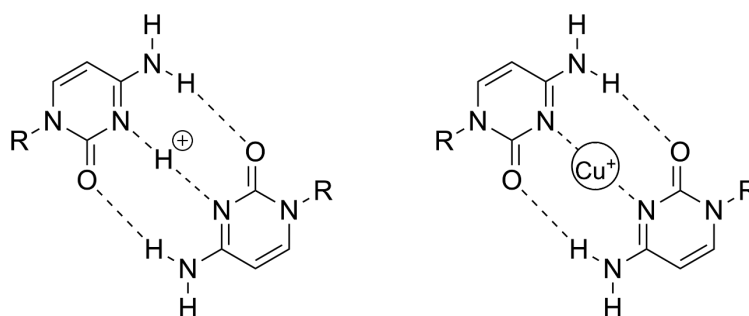


Figure 2.3.1 Hemi-protonated C-C<sup>+</sup> base pair (left) and proposed C-C base pair stabilised by Cu<sup>+</sup>.

Thus, given the requirement for C-C base pairs in the i-motif, the effects of  $\text{Cu}^+$  on an i-motif forming DNA sequence were investigated. This was in order to determine whether  $\text{Cu}^+$  could induce the formation of a secondary structure in DNA, by stabilising C-C base pairs in hTeloC. As discussed previously, hTeloC is predominantly unfolded at physiological pH, but capable of forming an i-motif at acidic pH (pH < 6). One important consideration was that  $\text{Cu}^+$  in solution is well known to oxidise readily to  $\text{Cu}^{2+}$  when exposed to  $\text{O}_2$ ,<sup>139,140</sup> therefore all



experiments were performed under strict anoxic conditions in an N<sub>2</sub> atmosphere (containing < 0.5 ppm O<sub>2</sub>).

### 2.3.1 Biophysical Characterisation of the Interaction between Cu<sup>+</sup> and hTeloC

The UV absorbance profile of DNA is dependent on its conformation. Therefore, UV spectroscopy can be used to elucidate whether DNA is folded or unfolded, and to reveal the existence of higher-order secondary structure(s).<sup>44</sup> UV-difference spectra are used to identify and characterise the behaviour of the secondary structure in response to experimental conditions.<sup>94,141</sup> “Cu<sup>+</sup> difference” spectra for hTeloC were measured at pH 5.5 and pH 7.4, where the structure is an i-motif or unfolded DNA respectively. The resulting spectra display a positive signal at 260 nm at both pH values, and a negative signal at 295 nm at pH 7.4, both consistent with when the i-motif formed by decreasing the pH.<sup>142</sup> These results indicate that the final configuration of the secondary structure adopted at either pH is similar, and that at pH 7.4 a more substantial reconfiguration is necessary to form the final structure.

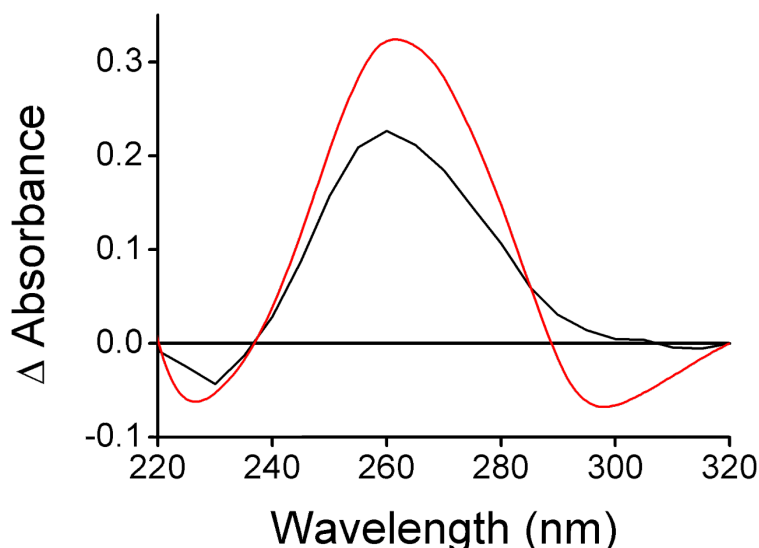
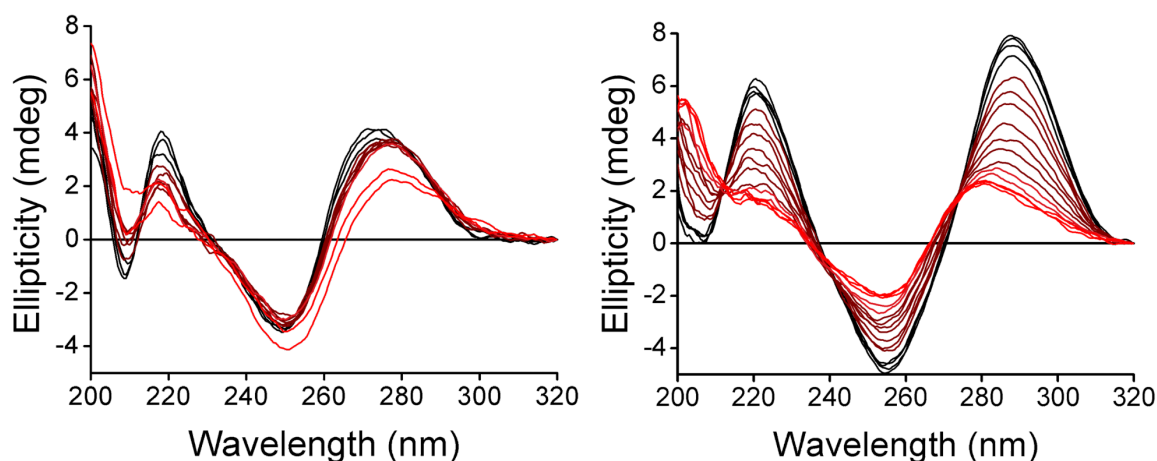


Figure 2.3.1.1 ‘Cu<sup>+</sup>-difference’ spectra using 125 μM of Cu<sup>+</sup> to form the final conformations at pH 7.4 (red) and pH 5.5 (black).

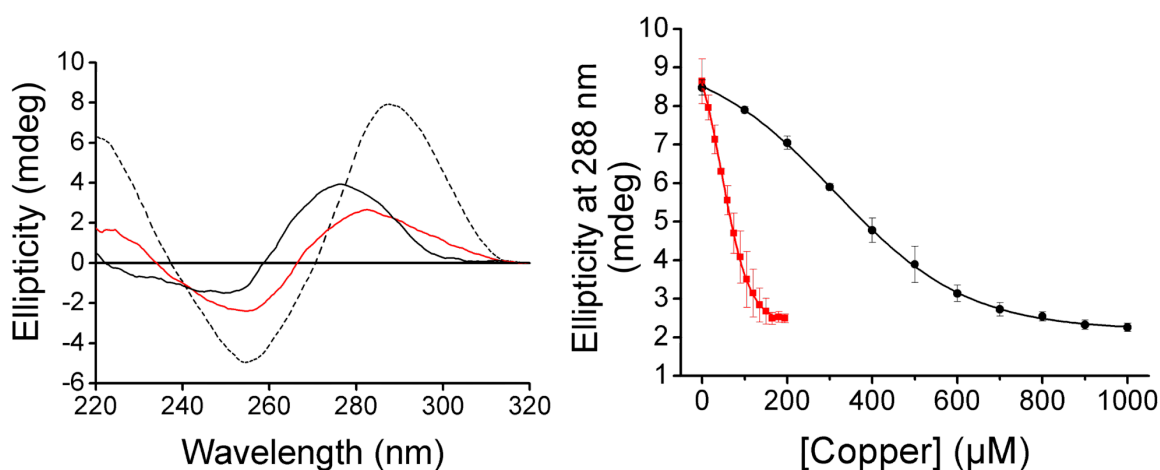
CD spectroscopy was used to further characterise the structure adopted by hTeloC in the presence of  $\text{Cu}^+$ . The CD spectrum of hTeloC at pH 7.4 has a positive peak at 270 nm and a negative peak at 250 nm, indicative of a primarily unfolded population of oligonucleotide.<sup>143</sup> Sequential addition of  $\text{Cu}^+$  up to 5 equivalents (i.e., 50  $\mu\text{M}$  final) resulted in a bathochromic shift in the positive peak from 270 nm to 278 nm, while the position of the negative peak at 250 nm remained constant. Further addition of  $\text{Cu}^+$  at this pH however, resulted in visible precipitation of the  $\text{Cu}^+$ -DNA complex and consequent deterioration of the CD signal. At pH 5.5, hTeloC is already folded into an i-motif with a characteristic positive peak at 288 nm and negative peak at 255 nm.<sup>143</sup> Under these conditions, titration of  $\text{Cu}^+$  up to 19.5 equivalents (195  $\mu\text{M}$   $\text{Cu}^+$ ) led to a hypsochromic shift of the positive peak from 288 nm to 283 nm, and a decrease in the amplitude of the negative peak at 255 nm. In contrast to the precipitation observed at pH 7.4, the  $\text{Cu}^+$ -DNA complex at acidic pH was completely soluble beyond the concentration where no further changes are observed (150  $\mu\text{M}$   $\text{Cu}^+$ ). The changes observed at pH 7.4 and 5.5 are consistent with a  $\text{Cu}^+$  induced reconfiguration of the structure. Crucially, post- $\text{Cu}^+$  addition and at both pH values, the spectra are practically superimposable indicating that a similar final structure is adopted regardless of the initial pH.



*Figure 2.3.1.2 CD spectra of 10  $\mu\text{M}$  hTeloC in 50 mM sodium cacodylate buffer (black) at pH 7.4 (left) with titration up to 50  $\mu\text{M}$   $\text{Cu}^+$  (red) and at pH 5.5 (right) with titration up to 150  $\mu\text{M}$   $\text{Cu}^+$  (red).*

Given the spectroscopic changes previously reported with  $\text{Cu}^{2+}$  and hTeloC, the possibility that the structure adopted in the presence of  $\text{Cu}^+$  may also display

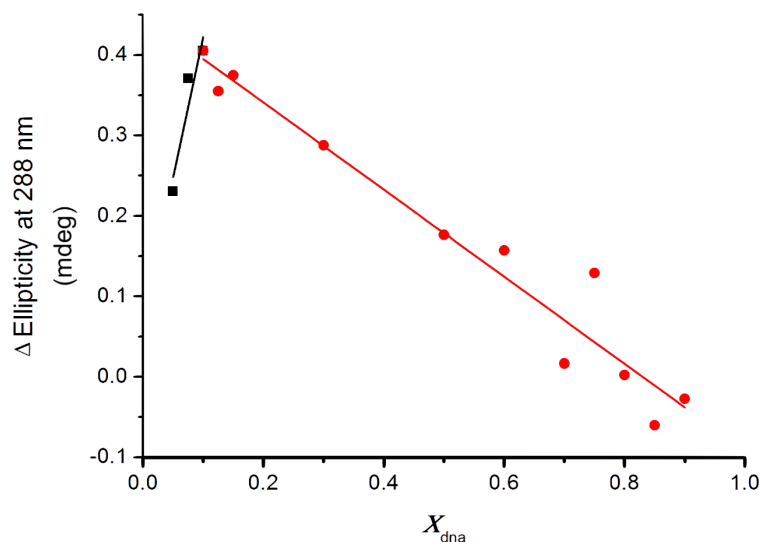
hairpin-like character<sup>81</sup> was explored, and the different copper-DNA complexes were compared using CD at pH 5.5. At this pH, in the absence of copper the CD spectrum of hTeloC has a positive peak at 288 nm indicative of i-motif structure. Addition of either Cu<sup>+</sup> or Cu<sup>2+</sup> resulted in a hypsochromic-shift consistent with an alteration in the structure of the DNA. Addition of Cu<sup>2+</sup> shifts this peak to 276 nm compared to only 283 nm when Cu<sup>+</sup> is added. The negative peak at 255 nm also undergoes a hypsochromic-shift to 250 nm in the presence of Cu<sup>2+</sup>, while the peak position does not shift at all when Cu<sup>+</sup> is added. In addition to the spectroscopic differences observed using the different oxidation states of copper, the half-cation concentrations also vary by an order of magnitude. A value of 46 (±3) μM was determined for the [Cu<sup>+</sup>]<sub>50</sub>, while the [Cu<sup>2+</sup>]<sub>50</sub> was comparatively higher at 382 (±14) μM. This strongly suggests that the Cu<sup>2+</sup>-DNA complex is different to the Cu<sup>+</sup>-DNA complex.



*Figure 2.3.1.3 CD spectra (left) of 10 μM hTeloC at pH 5.5 (dashed black) with addition of 150 μM Cu<sup>+</sup> (red) or 1 mM Cu<sup>2+</sup> (black) and the ellipticity at 288 nm of 10 μM hTeloC (right) in 50 mM sodium cacodylate buffer at pH 5.5 with titration up to 195 μM Cu<sup>+</sup> (red) and in 10 mM sodium cacodylate buffer at pH 5.5 with titration up to 1 mM Cu<sup>2+</sup> (black). Error bars show standard deviation across three repeats.*

The relatively high concentration of Cu<sup>2+</sup> required to form the hairpin was suggested to be because the Cu<sup>2+</sup> did not affect the structure by interacting directly with the bases in the oligonucleotide, but rather by shifting the equilibrium to the hairpin structure via an interaction with the sugar-phosphate backbone.<sup>81</sup> In the

case of  $\text{Cu}^+$ , continuous variation binding analysis determined the stoichiometry of  $\text{Cu}^+$  to DNA to be 9:1. This ratio further suggests that  $\text{Cu}^+$  is interacting with hTeloC in a different manner to  $\text{Cu}^{2+}$ , possibly via direct mediation of an interaction between bases in the sequence, perhaps consistent with the model proposed by Oomens: one  $\text{Cu}^+$  for each C·C base pair and, additionally, one for each loop.



*Figure 2.3.1.4 Job plot of hTeloC and  $\text{Cu}^+$  in 50 mM sodium cacodylate buffer at pH 5.5. The black and red symbols represent the points used for fitting the respective linear best fits to determine the intercept.*

### 2.3.2 Computational Investigation of the Interaction between $\text{Cu}^+$ and hTeloC

In collaboration with Dr László Fábíán from the School of Pharmacy at the University of East Anglia, models were created and optimised by DFT computational methods with cytosine base pairs stabilised by either ion (Figure 2.3.2.1). This was in order to compare the viability of C·C base pairs stabilised by  $\text{Cu}^+$  and  $\text{Cu}^{2+}$ .<sup>144</sup> The results revealed a planar C· $\text{Cu}^{2+}$ ·C complex similar to the model of Oomens, but with a symmetrical structure and both the N and O atoms from a single face of each cytosine moiety coordinating the metal ion. Notably, the interaction energy between the two bases and the cation was sensitive to the redox-state of the metal and was significantly larger for  $\text{Cu}^{2+}$  (+1,700 kJ/mol) than for  $\text{Cu}^+$  (+650 kJ/mol). However, the experimental hydration enthalpies of the ions suggested that hydration was more energetically favourable to base pair formation

involving  $\text{Cu}^{2+}$  (+2,100 kJ/mol), compared to  $\text{Cu}^+$  (+593 kJ/mol);<sup>145</sup> which was consistent with the different behaviour observed experimentally with these cations.

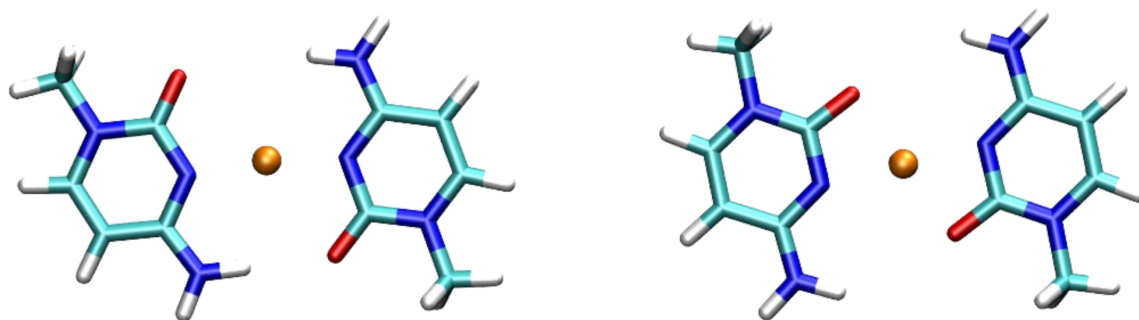
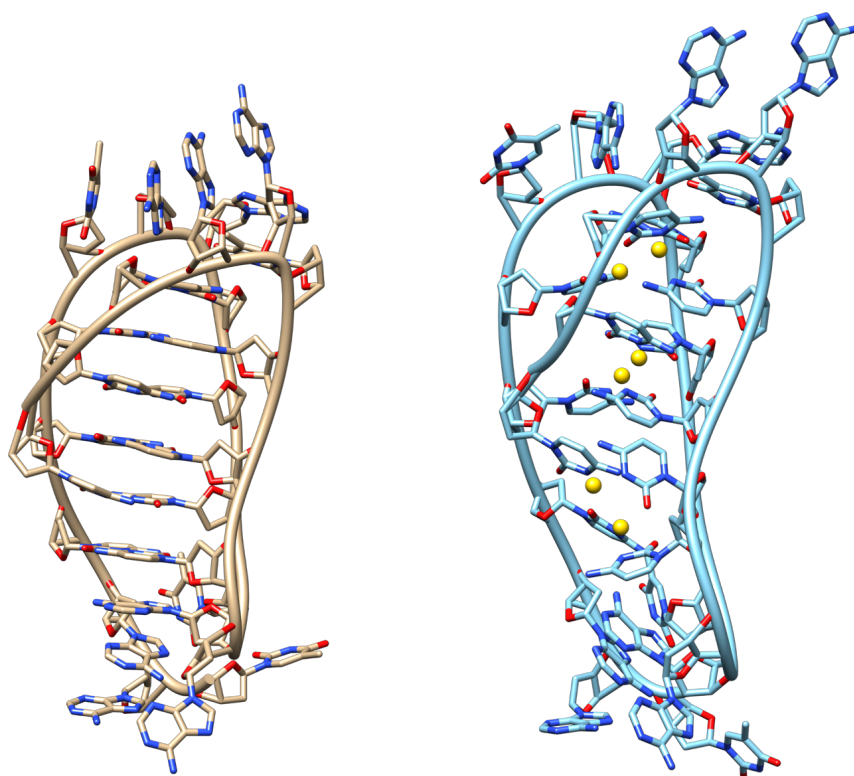


Figure 2.3.2.1 Structures of the  $\text{C}\cdot\text{Cu}^+\cdot\text{C}$  (left) and  $\text{C}\cdot\text{Cu}^{2+}\cdot\text{C}$  (right) base pairs, determined by DFT [TPSS-D3(BJ)/def2-TZVP] geometry optimisation. Colour code: C: cyan, H: white, N: blue, O: red, Cu: gold.

The CD spectra of the DNA at both pH values in the presence of  $\text{Cu}^+$  lay somewhere between that of unfolded DNA and i-motif DNA. Addition of  $\text{Cu}^+$  at pH 5.5 resulted in what could be interpreted as a slight unfolding of the i-motif, consistent with the structure expanding to accommodate the  $\text{Cu}^+$  cations, which are significantly larger than the protons which were previously stabilising the  $\text{C}\cdot\text{C}^+$  base pairs. The potential folded i-motif structures were investigated in more detail using molecular modelling. First, a model of protonated hTeloC was created on the basis of the reported NMR structure from a similar sequence.<sup>96</sup> The manually modified structure was optimised and then relaxed in a 200 ns explicit solvent MD simulation. To match the experimentally measured 9:1  $\text{Cu}^+$ :DNA stoichiometry, nine  $\text{Cu}^+$  cations were added to this relaxed model manually.<sup>146</sup> Six  $\text{Cu}^+$  ions were placed at the geometric midpoints between the N3 atoms of matching cytosine groups after the deprotonation of that position. The three additional cations were placed in the loop regions. The geometry of this initial  $\text{Cu}^+$ -DNA complex structure was optimised using the semi-empirical PM6-D3H4 method.<sup>147,148</sup> In the optimised structure the  $\text{Cu}^+$  ions showed a preference to interact with more than two bases, thereby breaking the planarity of the  $\text{C}\cdot\text{Cu}^+\cdot\text{C}$  units. Nevertheless, the overall folded structure was retained. To confirm these observations, a stack of six  $\text{C}\cdot\text{Cu}^+\cdot\text{C}$  base pairs capped at both ends with the nearest molecular fragments was extracted from the initial  $\text{Cu}^+$ -DNA complex and optimised using a DFT [TPSS-

D3(BJ)/def2-SV(P)] method. Both the semi-empirical and DFT calculations confirmed the preference of  $\text{Cu}^+$  ions to interact with more than two bases. Full exploration of the folding with  $\text{Cu}^+$  would require derivation and fitting of specific  $\text{Cu}^+$  force field parameters, which was beyond the scope of this work. Nevertheless, the computational modelling indicated the acid-stabilised and copper-stabilised i-motif structures were slightly different, which would explain the spectroscopic differences observed between these two species. Thus, it appeared that the planar base pairing ( $\text{C}\cdot\text{Cu}^+\cdot\text{C}$ ) model may only be true for cytosine monomers. When the cytosines form part of a larger secondary structure, the interactions are more complex which gives rise to a slightly different i-motif structure, as supported by the spectroscopic data.



*Figure 2.3.2.2 Model of hTeloC i-motif structure stabilised by protonation of cytosine residues, snapshot from the end of a 200 ns simulation (left). Model of the i-motif structure stabilised by  $\text{Cu}^+$  ions derived from the protonated model by geometry optimisation with the PM6-D3H4 method.*

### 2.3.3 Reversibility and Redox Sensitivity of the Cu<sup>+</sup>-stabilised i-Motif

The effects of Cu<sup>2+</sup> on the structure of hTeloC DNA were previously shown to be reversible using the chelator ethylenediaminetetraacetic acid (EDTA).<sup>81</sup> To determine if a similar reversibility could be achieved with Cu<sup>+</sup>, the high-affinity chelator diethyldithiocarbamate (DETC) was used. Titration of DETC into hTeloC at pH 5.5 reversed the effects of the Cu<sup>+</sup> addition and the structure reverted to that of the acid-stabilised i-motif. hTeloC with Cu<sup>+</sup> had the positive peak at 283 nm, incremental titration of the chelator DETC resulted in a red-shift of the peak until it returned to the position of the acid-stabilised i-motif peak at 288 nm. The negative peak in the presence of Cu<sup>+</sup> at 255 nm did not shift its position but the amplitude of the signal increased to be more consistent with that of the original acid-stabilised i-motif.

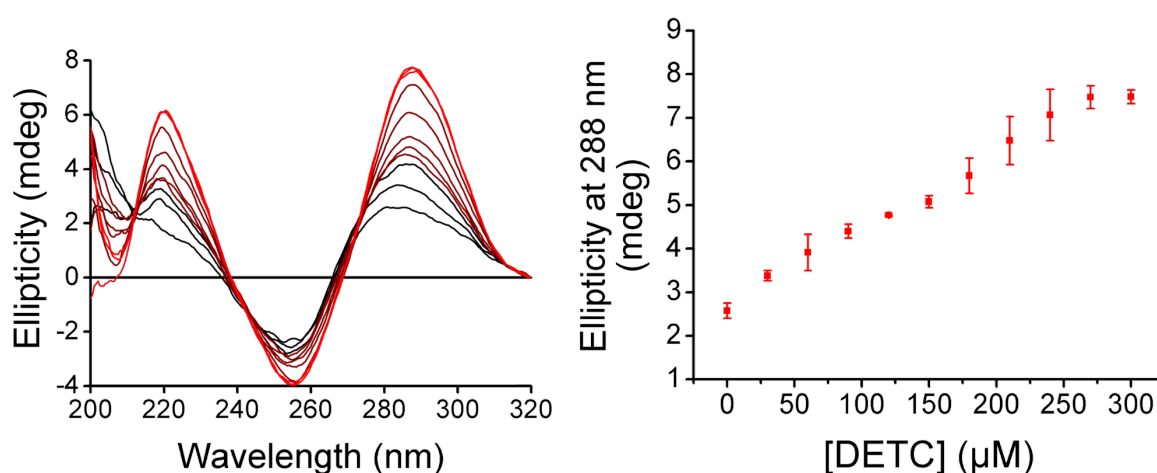


Figure 2.3.3.1 CD spectra (left) of 10 μM hTeloC with 150 μM Cu<sup>+</sup> at pH 5.5 (black) with titration up to 300 μM DETC (red) and the ellipticity at 288 nm (right) of 10 μM hTeloC in 50 mM sodium cacodylate buffer at pH 5.5 with titration up to 300 μM DETC. Error bars show standard deviation across three repeats.

Further experiments to examine the mode of copper binding, and reversibility, to hTeloC were performed using <sup>1</sup>H NMR. At pH 5.5, imino proton signals can be observed at 15.5 ppm and are characteristic of the C-C<sup>+</sup> base pairs in an acid-stabilised i-motif.<sup>97</sup> On addition of Cu<sup>+</sup>, these signals disappeared, consistent



with  $\text{Cu}^+$  replacing the protons in that position. This is also consistent with the model proposed by Oomens and co-workers.<sup>138</sup> Furthermore, no additional signals appeared, ruling out a hairpin conformation with additional Watson-Crick base pairing, as was seen with  $\text{Cu}^{2+}$ .<sup>81</sup> Subsequent addition of the chelator DETC caused the NMR spectrum to return to that of the acid-stabilised i-motif, with chelation of  $\text{Cu}^+$  resulting in the reappearance of the imino proton signal at 15.5 ppm.

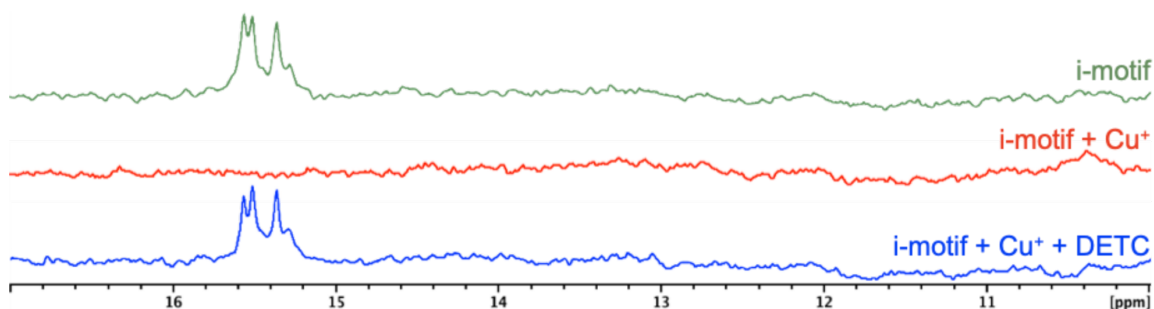


Figure 2.3.3.2  $^1\text{H}$  NMR of  $10\ \mu\text{M}$  hTeloC in  $50\ \text{mM}$  sodium cacodylate buffer pH 5.5 with 5%  $\text{D}_2\text{O}$  (green), with addition of  $150\ \mu\text{M}$   $\text{Cu}^+$  (red), and addition of  $150\ \mu\text{M}$   $\text{Cu}^+$  and  $540\ \mu\text{M}$  DETC (blue).

Thus far, all experiments were carried out under stringent anoxic conditions to prevent the oxidation of  $\text{Cu}^+$  that would occur in the open air. To examine a possible further dimension to this system and determine if a redox-linked structural rearrangement would be observed *in situ*, the  $\text{Cu}^+$ -DNA complex was exposed to air to allow the metal to be oxidised. The  $\text{Cu}^+$ -i-motif can be formed by adding  $150\ \mu\text{M}$  of  $\text{Cu}^+$  and, as  $\text{Cu}^+$  is a high-affinity ligand compared with  $\text{Cu}^{2+}$ , complete oxidation to  $\text{Cu}^{2+}$  would yield a cation concentration below the  $[\text{Cu}^{2+}]_{50}$  ( $382 \pm 14\ \mu\text{M}$ ). As a result, there would not be enough  $\text{Cu}^{2+}$  to form the hairpin. To test this hypothesis,  $\text{Cu}^+$  was added to hTeloC at pH 5.5 and the sample was split into two. One sample was maintained in an anoxic environment whilst the other was exposed to the open air. The CD spectrum for each condition was subsequently measured and the one which had been exposed to oxygen reverted almost completely to the acid-stabilised i-motif: the positive peak moved from 281 nm to 286 nm, and the amplitude of the negative peak at 255 nm increased, as observed when the  $\text{Cu}^+$  was chelated with DETC. In contrast, the sample maintained in the anoxic environment remained essentially unchanged.



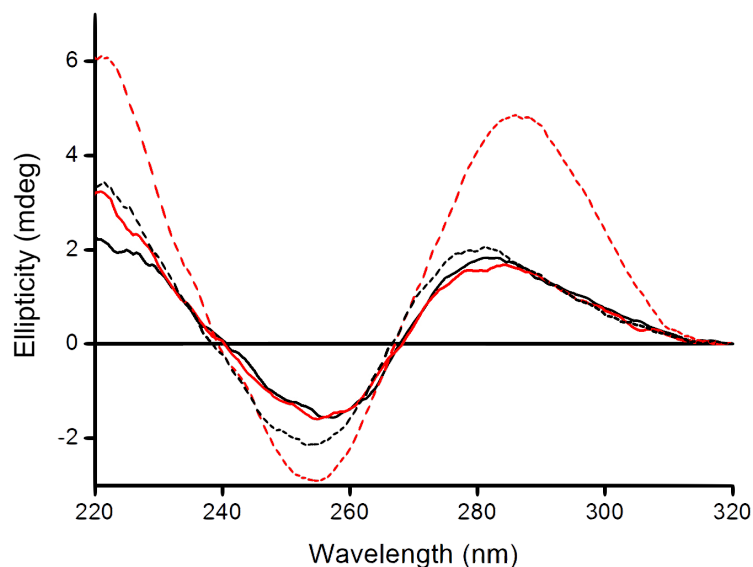


Figure 2.3.3.3 CD spectra of 10  $\mu\text{M}$  hTeloC in 50 mM sodium cacodylate buffer at pH 5.5 with 150  $\mu\text{M}$   $\text{Cu}^+$  scanned immediately after addition (solid lines) and 3 hours later (dashed lines); (red) sample exposed to the air; (black) sample maintained in anoxic environment.

Having discovered that this system was oxygen-responsive and that oxidation of the  $\text{Cu}^+$  resulted in the restoration of the acid-stabilised i-motif structure we were interested in determining whether this transition was possible in the opposite direction; i.e. whether it would be possible to reduce  $\text{Cu}^{2+}$  *in situ* to form the  $\text{Cu}^+$ -i-motif. To explore this, the well-established reaction between  $\text{Cu}^{2+}$  and sodium ascorbate, where  $\text{Cu}^{2+}$  is reduced to  $\text{Cu}^+$ , was used.<sup>149</sup> 150  $\mu\text{M}$   $\text{Cu}^{2+}$  was added to the acid-stabilised i-motif and, as this concentration was below the  $[\text{Cu}^{2+}]_{50}$ , no structural change was observed using CD. Subsequent addition of 150  $\mu\text{M}$  sodium ascorbate resulted in the successful formation of the  $\text{Cu}^+$ -i-motif complex observed previously when  $\text{Cu}^+$  was added under anoxic conditions.

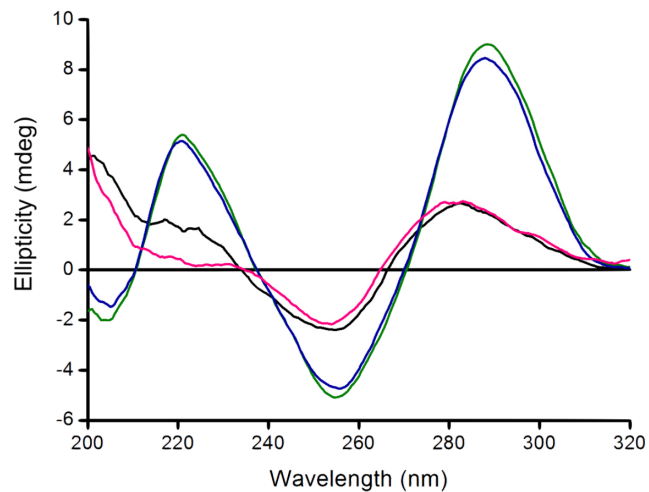


Figure 2.3.3.4 CD spectra of 10  $\mu\text{M}$  hTeloC in 50 mM sodium cacodylate buffer at pH 5.5 (green); addition of 150  $\mu\text{M}$   $\text{Cu}^{2+}$  (blue); addition of 150  $\mu\text{M}$  sodium ascorbate (pink). For comparison, the CD spectra of 10  $\mu\text{M}$  hTeloC in 50 mM sodium cacodylate buffer at pH 5.5 after addition of 150  $\mu\text{M}$   $\text{Cu}^+$  (black).

When one equivalent of sodium ascorbate was used, exposure of the sample to air led to a slow process of oxidation and return to the acid-stabilised i-motif structure, while adding an excess of sodium ascorbate allowed for the prolonged maintenance of the concentration of  $\text{Cu}^+$  and the corresponding  $\text{Cu}^+$  stabilised i-motif structure.

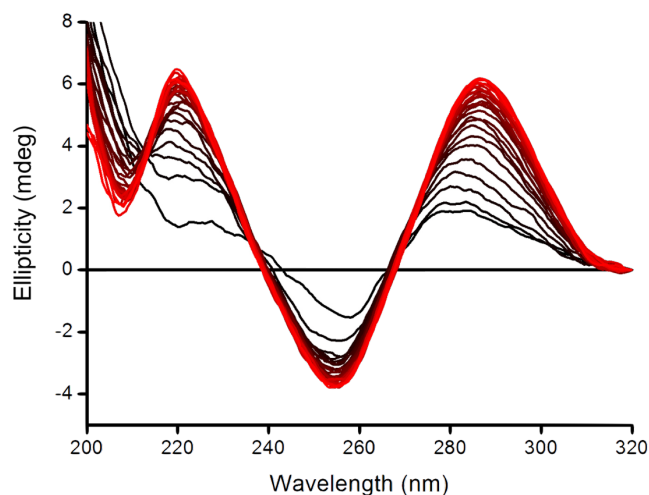


Figure 2.3.3.5 CD spectra of 10  $\mu\text{M}$  hTeloC in 50 mM sodium cacodylate buffer at pH 5.5 with 150  $\mu\text{M}$   $\text{Cu}^{2+}$  immediately after addition of 150  $\mu\text{M}$  sodium ascorbate and every 10 minutes for 240 minutes (black to red).

---

The ability to maintain the Cu<sup>+</sup>-i-motif structure for several hours in the open air enabled the performance of FRET-based DNA melting experiments using 2 equivalents of sodium ascorbate to observe the folding behaviour using the dual-labelled sequence hTeloC<sub>FRET</sub>. The fluorescence intensity at 25°C was used to determine the fraction of the DNA that is folded in the presence of increasing concentrations of Cu<sup>+</sup>. In good agreement with the UV and CD data, addition of Cu<sup>+</sup> to hTeloC<sub>FRET</sub> at pH 5.5 did not affect the proportion of the population of the DNA that was folded. Further ruling out unfolding and supporting the observation that the folded conformations of the proton-stabilised and Cu<sup>+</sup>-stabilised i-motif are similar. Conversely, at pH 7.4 addition of Cu<sup>+</sup> results in folding of the sequence into a secondary structure that brings the two ends of the sequence into sufficient proximity for FRET to occur. Additionally, using this technique we were able to determine a  $T_m$  which was calculated as the midpoint temperature of the transition from the folded to the unfolded structure. At pH 7.4 increasing Cu<sup>+</sup> concentration led to an increase in  $T_m$  until 15 equivalents (3 μM) at which point it was 65°C, and after which no further change was observed. This is in agreement with the 15 equivalents of Cu<sup>+</sup> required to fold the DNA as determined by CD. An increase in  $T_m$  was also observed at pH 5.5, however the temperature required to unfold completely was greater than 95°C, the limit of the instrument, therefore an accurate determination of the  $T_m$  was not possible.

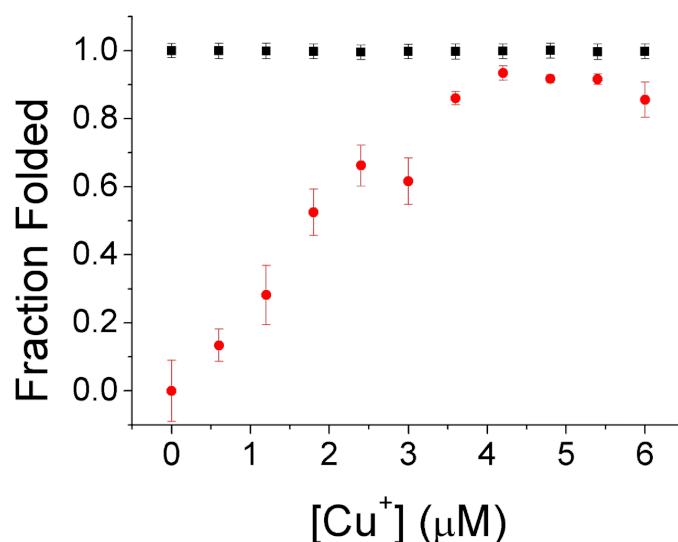
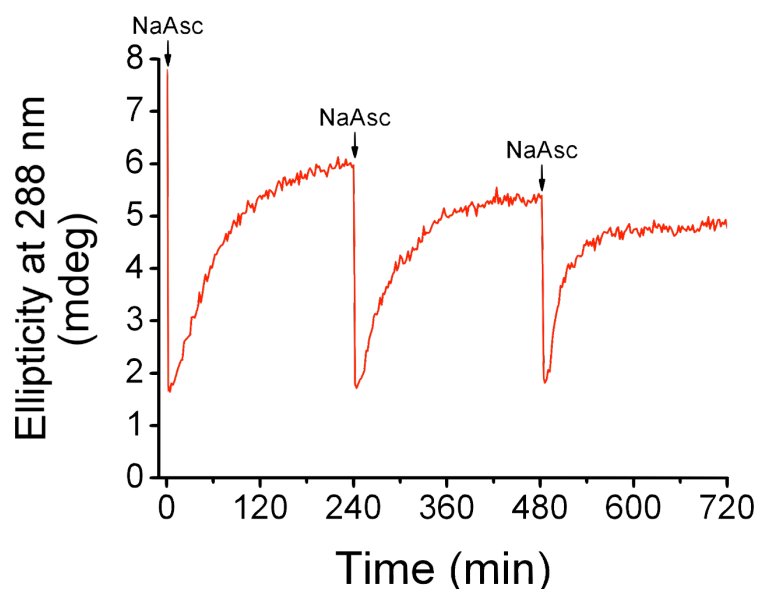


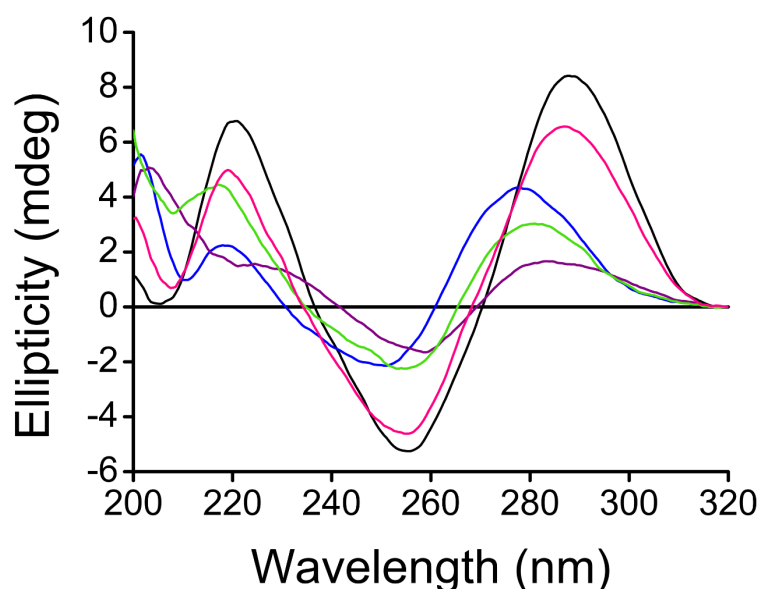
Figure 2.3.3.6 Fluorescence intensity at 25°C normalised using values in the absence of Cu<sup>+</sup> at pH 5.5 as 1 (folded) and at pH 7.4 (unfolded). 200 nM hTeloC<sub>FRET</sub> in 10 mM sodium cacodylate buffer at pH 5.5 (black) and at pH 7.4 (red). Error bars show standard deviation across three repeats.

Multiple iterations of the conformational change induced by Cu<sup>2+</sup> on hTeloC were reported to be possible by repeated chelation and metalation.<sup>81</sup> From a nanotechnology perspective, the potential ability to have a conformational change of the structure controlled by redox-cycling the metal was very encouraging. Thus, the ability for similar repeat switching between the Cu<sup>+</sup>-stabilised and the acid-stabilised i-motif structures was investigated. To truly test the versatility of this system, rather than perform repeated addition and chelation of Cu<sup>+</sup>, a single addition of Cu<sup>2+</sup> was added and this was reduced *in situ* repeatedly to Cu<sup>+</sup> using sodium ascorbate. The structural reconfiguration in response to the oxidation state of the copper was observed by monitoring the molar ellipticity at 288 nm as a function of time. Figure 2.3.3.7 shows the results of this experiment with three successive additions of the reducing agent successfully resulting in adoption of the Cu<sup>+</sup>-stabilised i-motif structure and the oxidation to Cu<sup>2+</sup> over time similarly resulting in the return to the acid-stabilised i-motif prior to the next sodium ascorbate addition.



*Figure 2.3.3.7 Ellipticity at 288 nm of 10  $\mu\text{M}$  hTeloC at pH 5.5 with 150  $\mu\text{M}$   $\text{Cu}^{2+}$  as a function of time with three additions of 150  $\mu\text{M}$  sodium ascorbate under ambient conditions.*

Having established the redox dependent coordination of copper by the i-motif forming DNA sequence hTeloC, it was hypothesised that this system could act as a continuous redox sensitive cycle, allowing for dynamic movement between the various structural conformations adopted under the different conditions. The final step was to determine whether it was possible to convert the  $\text{Cu}^{2+}$  hairpin structure to the  $\text{Cu}^+$ -stabilised i-motif structure. As can be seen in Figure 2.3.3.8, addition of 1 mM  $\text{Cu}^{2+}$  to a sample of hTeloC at pH 5.5 forms the hairpin structure and subsequent reduction to  $\text{Cu}^+$  using 150  $\mu\text{M}$  sodium ascorbate successfully forms the previously observed  $\text{Cu}^+$ -DNA i-motif, even in the presence of excess  $\text{Cu}^{2+}$ . This was a fortunate but predictable outcome due to the difference of an order of magnitude between the binding affinities of the different oxidation states of copper. Leaving the same sample in the open air over time resulted in conversion back to the  $\text{Cu}^{2+}$  stabilised hairpin structure. Finally, addition of 1 mM EDTA chelated the  $\text{Cu}^{2+}$  and the sample returned to its initial configuration as an acid-stabilised i-motif.



*Figure 2.3.3.8 CD spectra of a single sample of 10  $\mu\text{M}$  hTeloC at pH 5.5 (black); addition of 1 mM  $\text{Cu}^{2+}$  (blue); addition of 150  $\mu\text{M}$  sodium ascorbate (purple); after 4 h exposure to air (green); chelation using 1 mM EDTA (pink).*

The discovery of this interaction demonstrated that  $\text{Cu}^+$  could be used to fold an i-motif forming DNA sequence into a  $\text{Cu}^+$ -stabilised i-motif structure. This process could be reversed by chelation of the metal, or by oxidation of  $\text{Cu}^+$  to  $\text{Cu}^{2+}$ . This was the first, and as of this writing the only, example of redox-dependent control of DNA secondary structure. This work realises that a series of alternative conformational switches for i-motif forming DNA sequences are possible using different conditions, without changing the pH. A summary of the transitions possible is conveyed in Figure 2.3.3.9, illustrating the pH and redox sensitive control of the structural conformation of the i-motif forming DNA sequence hTeloC in the presence of copper. The dynamics of this system could be applied to develop dual oxygen and pH-sensitive nanomachines, logic gates or sensors based on i-motif DNA.

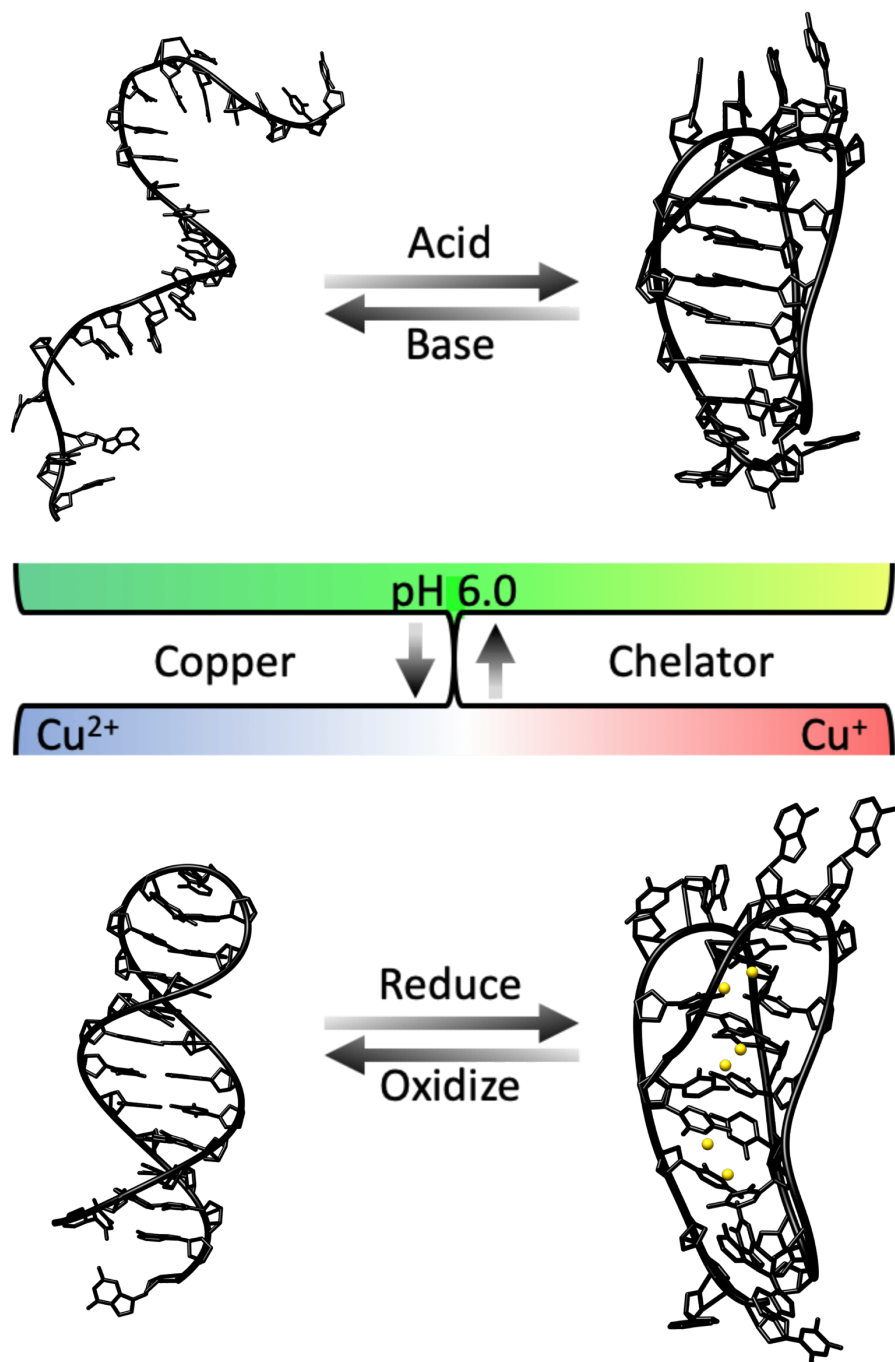


Figure 2.3.3.9 Illustration of proposed system for the pH and copper-redox-dependent control of the structure of the i-motif forming DNA sequence hTeloC.

---

## **Chapter 3: Ligands**



### 3.1 Ligands and Quadruplexes

Historically, DNA was the first target for anti-cancer drugs and remains the mainstay of most treatment regimes.<sup>150</sup> Given their positions in the genome, the potential therapeutic impact of targeting quadruplex DNA structures to control homeostatic or pathogenic processes is vast.<sup>57,58</sup> Ground-breaking recent work enabled the *in vivo* visualization of both quadruplex structures in the nuclei of human cells. The G-quadruplex was shown first in 2013 by Balasubramanian and co-workers,<sup>151</sup> followed by the i-motif in 2018 by Christ and co-workers.<sup>61</sup> The confirmation of the existence of both quadruplexes in cells further increased the significance of finding ligands that can interact with them.

As introduced in section 1.5, one of the challenges facing the i-motif is the wealth of small molecule G-quadruplex ligands compared to the dearth of ligands that selectively target the i-motif. One interesting reported example of an i-motif stabilising ligand is that of **IMC-48**.<sup>120</sup> Hurley and co-workers reported that the cytosine-rich sequence from the *BCL2* promoter could form a hairpin or an i-motif, and that the equilibrium between the two structures could be controlled by using two ligands identified through screening methods. Shifting the population to i-motif by **IMC-48** activated transcription, while a shift to the hairpin by **IMC-76** repressed it.<sup>120</sup>

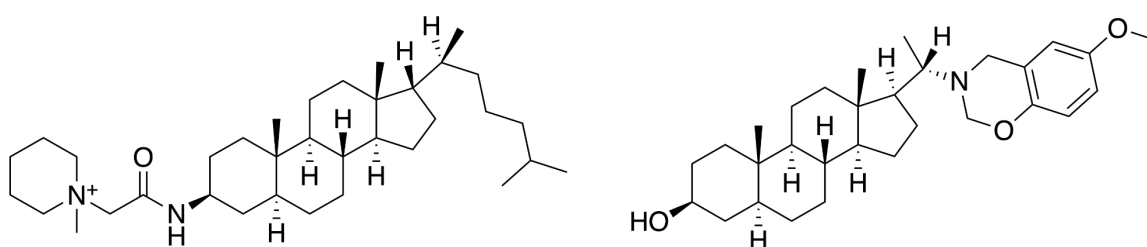


Figure 3.1.1 Structures of the *BCL2* i-motif ligands **IMC-48** (left) and **IMC-76** (right).

Relatively few other reports in the literature exist for i-motif binding ligands, and those ligands found to have affinity for i-motif have lacked selectivity; additionally having good affinity for G-quadruplex or duplex DNA.<sup>11,35,152</sup> Therefore, the need to identify a small molecule that stabilises the i-motif in a selective manner is

needed; not only to better understand the structure, but also to enable potentially meaningful application as a future therapeutic. It must be noted that a recent article has been published that suggests **B19**, one of a series of acridone derivatives synthesised to target i-motif, can selectively bind the c-myc i-motif. Crucially, this compound reportedly did not show affinity for G-quadruplex or duplex DNA. Additionally, it repressed c-myc transcription and expression thereby leading to tumour cell death.<sup>153</sup> This latest report only serves to provide further support to the i-motif hypothesis and will encourage research in this pursuit.

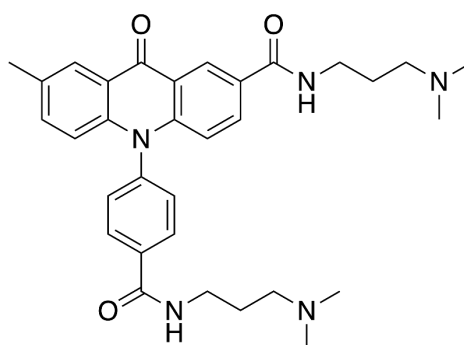


Figure 3.1.2 Structure of the c-myc i-motif selective ligand **B19**.

### 3.2 i-Motif Ligand Hunt

Finding a small molecule compound that can specifically stabilise or induce i-motif structure had potential significance on several fronts:

- (i) it could serve as a tool that can be used to further our understanding of i-motif structure and function
- (ii) with the emerging evidence for physiological and/or pathological roles for the i-motif it had the potential to be developed for diagnostic or therapeutic applications
- (iii) it could be used to control the i-motif functionality in the DNA based nanostructure being developed. With these potential applications in mind, the hunt commenced for an i-motif specific ligand.

While limited reports existed in the literature for ligand interactions with the i-motif the Waller research group had been investigating the interaction of small molecules with different i-motif forming sequences for some time. Medium throughput FRET-

---

melting experiments (as described in section 1.5.4) were one of the primary methods used by the group previously to explore the interaction of the 1584 compounds in the National Cancer Institute (NCI) Diversity Set VI library with a variety of i-motif forming DNA sequences as well as a double stranded DNA sequence and a G-quadruplex sequence. The data collected contained information on the change in melting temperature ( $\Delta T_m$ ) of the structure under examination induced by the ligands in the NCI library. Using this data, a list of small molecules was selected based on their ability to increase the melting temperature of one of the investigated i-motif forming sequences and, where available, for having a smaller stabilising effect on the melting temperature of the double stranded structure and/or the G-quadruplex (details in Appendix A.1). Additional compounds were included based on a variety of other factors including results from other experiments or for having a pharmacophore of interest, as well as mitoxantrone (NSC 301739) as a positive control.<sup>152</sup> The compounds in this combined list (Figure 3.2.1) were investigated further by a variety of methods including more comprehensive FRET-melting experiments, SPR, FID, CD, UV and NMR.

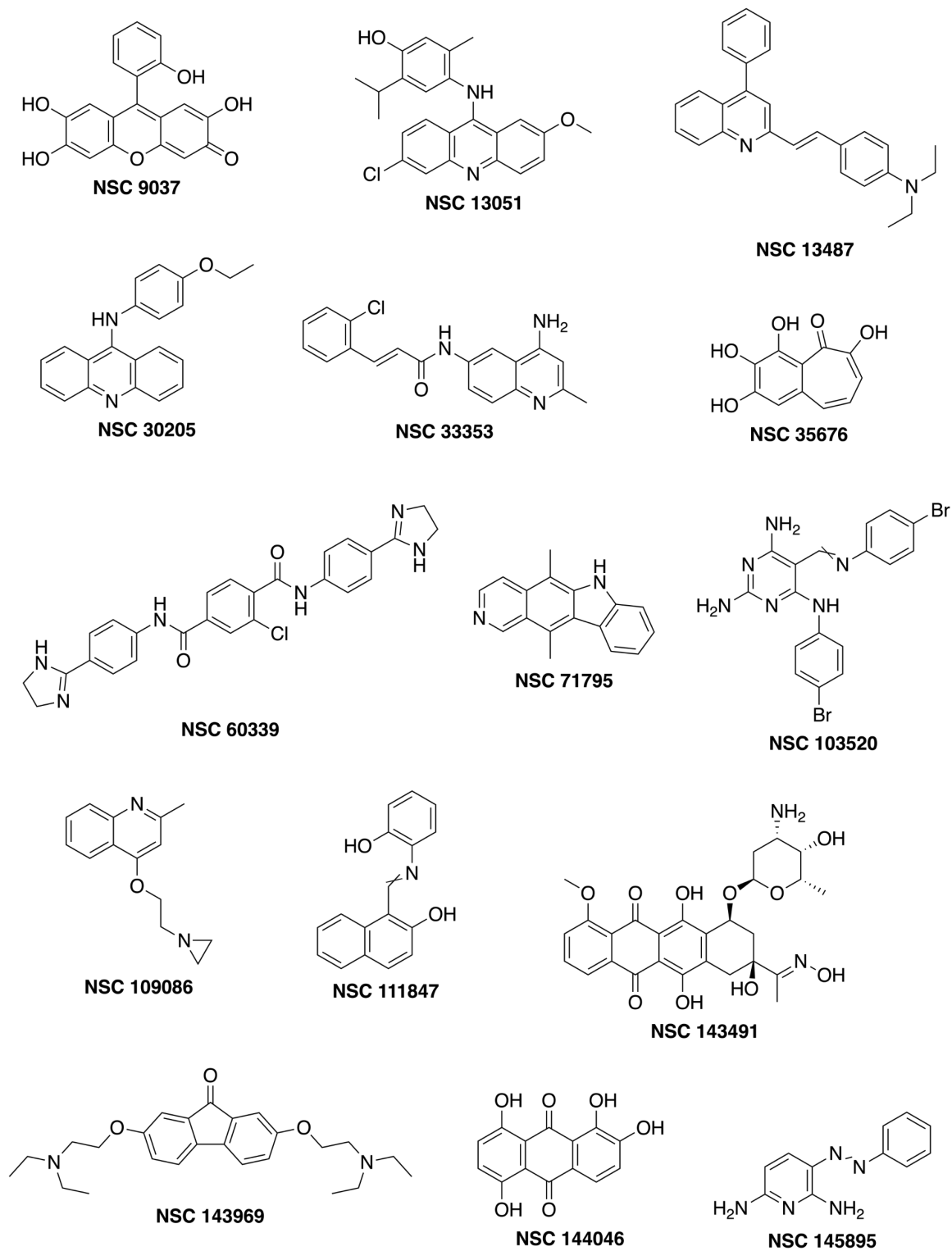


Figure 3.2.1 Structures of the 38 compounds selected for further investigation.

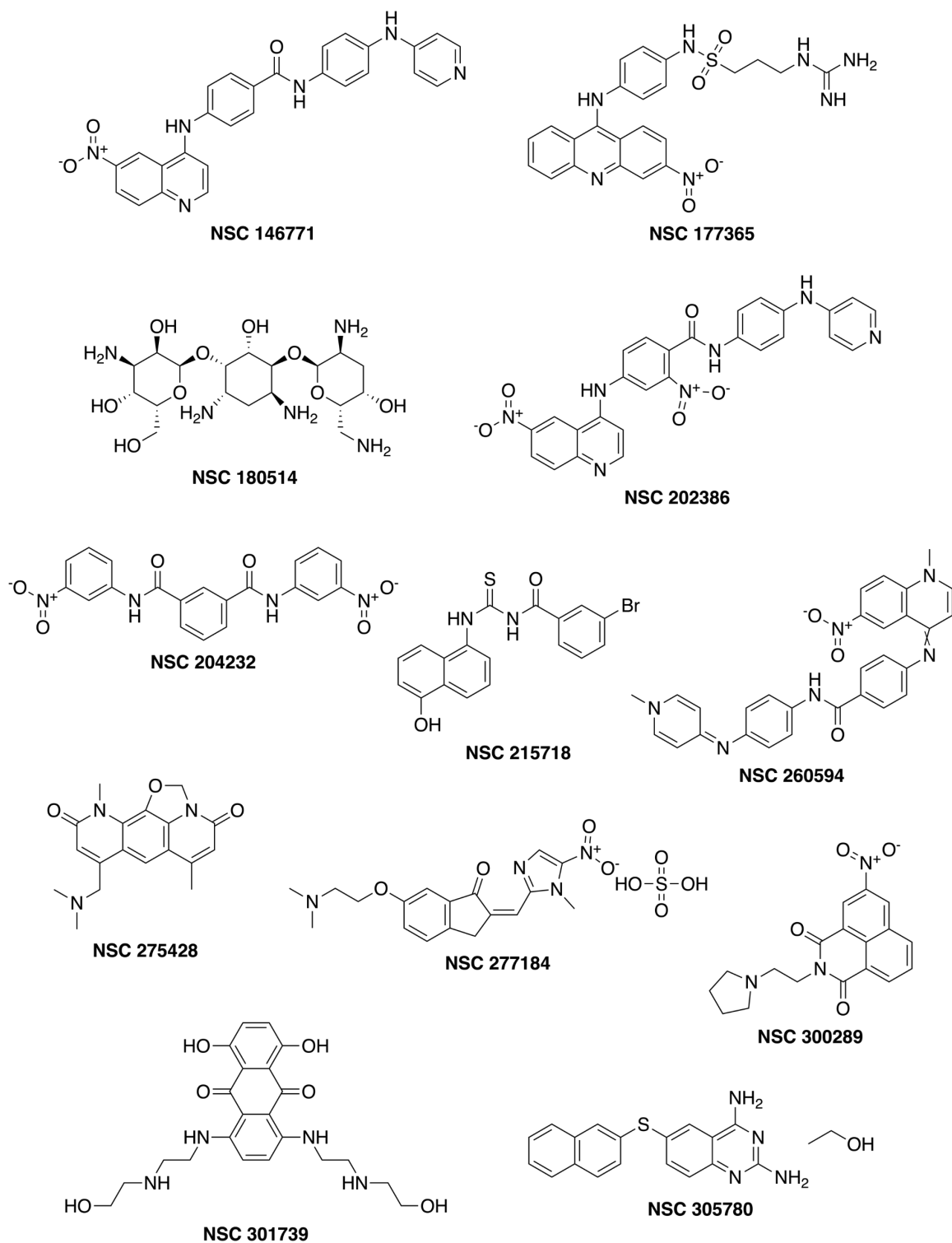


Figure 3.2.1 Structures of the 38 compounds selected for further investigation.

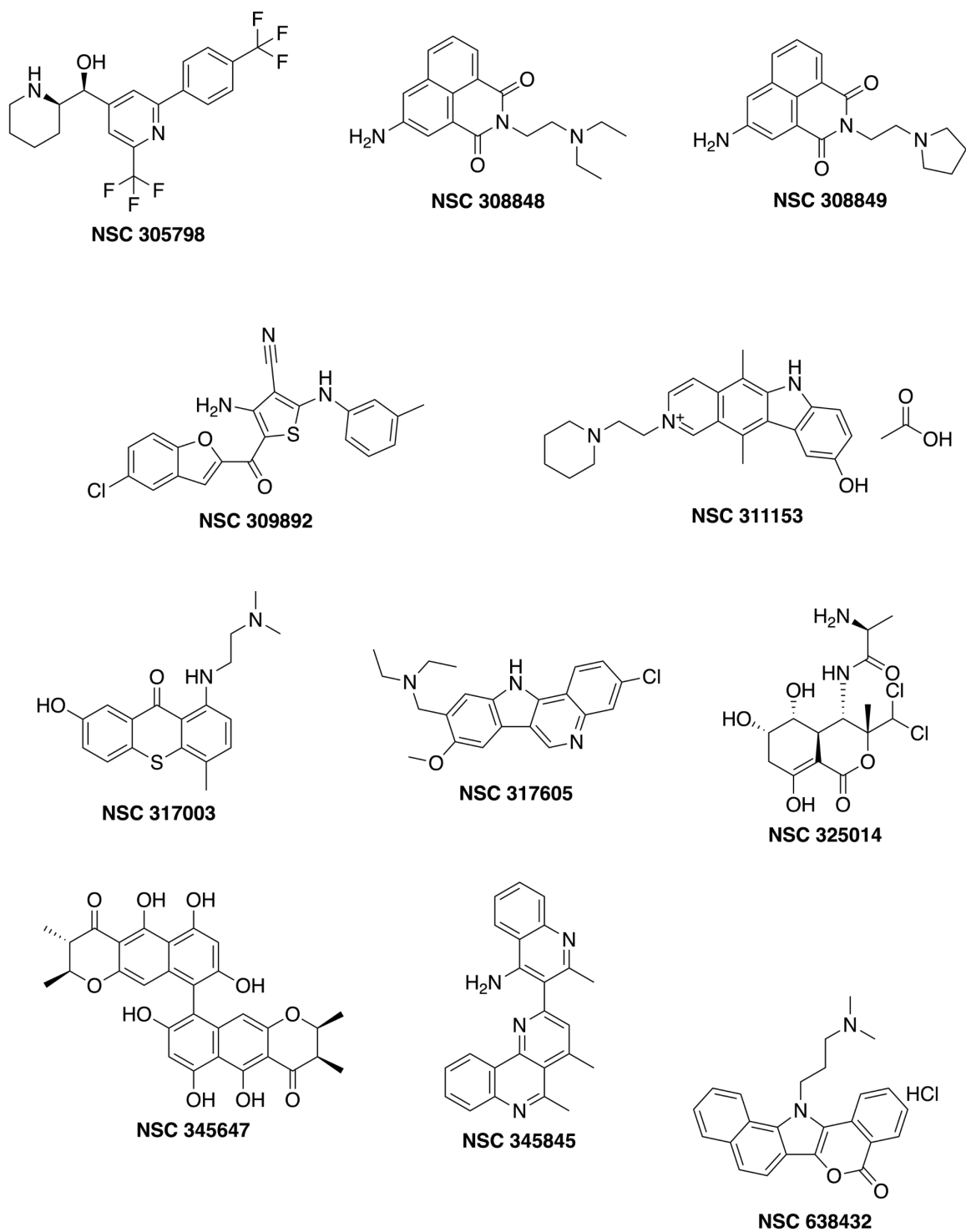


Figure 3.2.1 Structures of the 38 compounds selected for further investigation.

---

### 3.2.1 Surface Plasmon Resonance (SPR) Screen

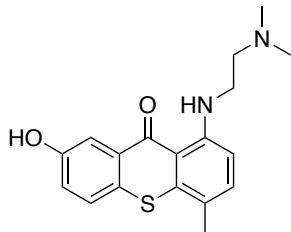
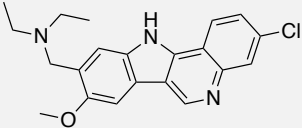
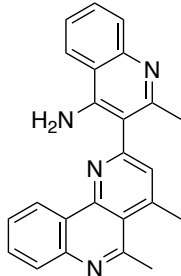
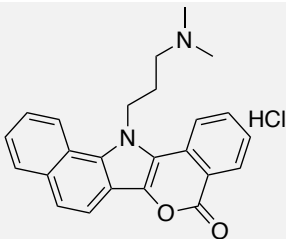
As introduced in section 1.5.6, SPR can be used to determine binding affinity and kinetics data for ligands with DNA. Initially however, it was used in a higher throughput manner to detect binding events between this list of ligands and three 5'-biotinylated DNA sequences immobilised on the chip surface: the hTeloC i-motif, the c-Myc i-motif: (TC<sub>4</sub>AC<sub>2</sub>T<sub>2</sub>C<sub>4</sub>A(C<sub>3</sub>TC<sub>4</sub>A)<sub>2</sub>) and a double stranded sequence DS<sub>biotin</sub>: (G<sub>2</sub>CATAGTGCGTG<sub>3</sub>CGT<sub>2</sub>AGC) and its complement: DS<sub>comp</sub>. The SPR work described here was performed with the help of Dr Clare Stevenson from the John Innes Centre. The response (R) observed in sensorgrams is proportional to the mass on the surface of the SPR chip. Therefore, if a compound binds to the immobilised DNA this results in a change in the mass on the surface and which is reflected by a change in the observed response. The maximal response (R<sub>max</sub>) describes the capacity of the surface to accommodate a given ligand, taking into account the level of DNA immobilised on the chip. This can be calculated using Equation 3.2.1.1:

$$\text{Equation 3.2.1.1: } R_{max} = \frac{Mw_{ligand}}{Mw_{DNA}} \times R_{DNA}$$

R<sub>DNA</sub> is the response measured upon immobilisation of each DNA structure to the surface of the chip and, for simplicity, this equation assumes a 1:1 binding model of ligand to DNA. The first SPR screen was performed at pH 5.5 using a single concentration (50 µM) of each of the 38 ligands, and a theoretical R<sub>max</sub> was calculated for each ligand-DNA pair. The response from each binding event was then measured and the %R<sub>max</sub> calculated wherein the response observed was presented as a percentage of the theoretical R<sub>max</sub>. From this data 11 ligands were selected to be tested further, the results for the chosen 11 are presented in Table 3.2.1.1 (full results in Appendix A.2):

Ligand (NSC)	Structure	%R <sub>max</sub> (%)		
		hTeloC	c-Myc	DS
9037		-43.91	210.65	-1498.5
13487		-211	172.11	-603.44
60339		353.2	545.27	-510.2
143491		368.49	373.99	-53.13
260594		294.89	344.74	-427.53
308848		145.19	100.75	-165.54
311153		308.63	746.38	-2637.1



<b>317003</b>		263.98	425.61	-695.47
<b>317605</b>		385.06	617.58	-202.54
<b>345845</b>		186.9	281.84	-325.77
<b>638432</b>		260.97	420.62	-557.13

*Table 3.2.1.1 %R<sub>max</sub> calculated for each ligand-DNA pair. [ligand] = 50 μM  
Running buffer: 10 mM sodium cacodylate, 100 mM NaCl, 0.05% tween at pH 5.5.*

While some of the other ligands had higher %R<sub>max</sub> values, these ligands were selected on the basis of them showing potential selectivity for the i-motif(s) over the double stranded DNA. They all had negative %R<sub>max</sub> values for the double stranded structure, indicating very low levels of binding or a possible duplex denaturing effect, while at the same time having positive %R<sub>max</sub> values for the i-motifs (with the exception of **9037** and **13487** which had an apparent further selectivity for the c-Myc i-motif).

Further SPR experiments were performed using a concentration range (0 to 500 μM) in order to determine the dissociation constant (K<sub>D</sub>) for these 11 ligands with the DNA structures. In order for the K<sub>D</sub> to be determined accurately a variety of factors must be taken into consideration. The quality of a sensorgram is one of the most important factors: ideally there should be an increase in the response (R)

---

upon injection of a ligand, followed by a plateau and then a decrease in the response, and an eventual return to the baseline response level, as the ligand dissociates from the DNA. There should also exist a relationship between the concentration of the ligand injected and the magnitude of the response observed: this second factor is essential in the determination of binding affinities or dissociation constants. Provided these quality control criteria are met the  $K_D$  can be determined: using Equation 3.2.1.2 where  $R_{eq}$  is the response at equilibrium, the association constant ( $K_A$ ) can be determined and the dissociation constant calculated as:  $K_D = \frac{1}{K_A}$

$$\text{Equation 3.2.1.2: } R_{eq} = \frac{K_A[\text{ligand}]R_{max}}{1+K_A[\text{ligand}]}$$

In this way, an attempt was made to determine the dissociation constants for the 11 compounds, where the criteria were met. Figure 3.2.1.1 show example data of the plots generated in the attempt to fit this data and calculate the  $K_D$  for ligand **60339** with the three different DNA secondary structures, for hTeloC and c-Myc dissociation constants were determined as  $0.22 \pm 0.04 \mu\text{M}$  and  $0.34 \pm 0.04 \mu\text{M}$ , respectively, while in the case of double stranded DNA the data failed the software's quality control analysis necessary to determine a  $K_D$  (this was the case for all 11 ligands with double stranded DNA). The case of **60339** was a good example of the binding data being in agreement with the  $\%R_{max}$  data, although the remaining ligands did show some variability. **9037** and **13487** were also both consistent with the  $\%R_{max}$  data: dissociation constants could not be calculated for hTeloC or double stranded DNA in these cases, as they both had negative  $\%R_{max}$  values (response vs concentration plots for all 11 ligands in Appendix A.3). Unfortunately, with these two ligands the  $K_D$  values calculated for the c-Myc i-motif had a margin of error of practically 100% ( $K_D$  data in Table 3.2.1.2).

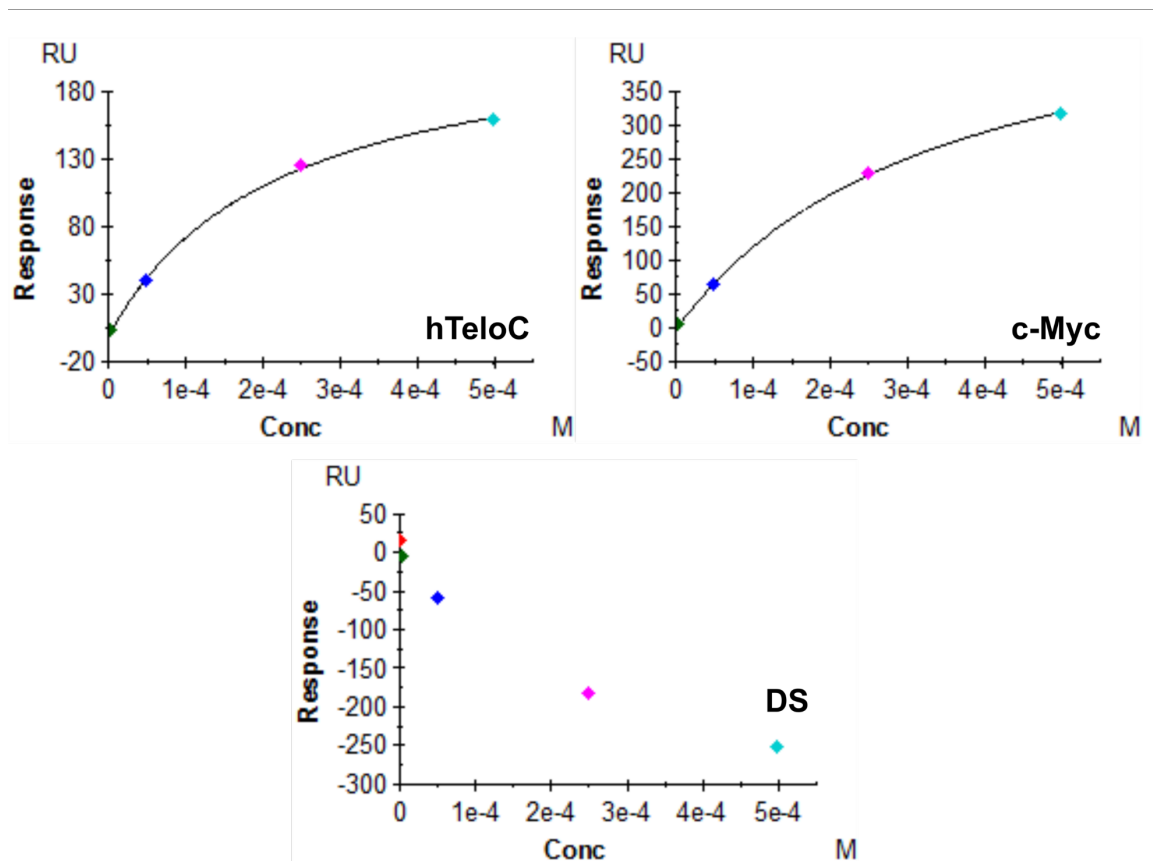


Figure 3.2.1.1 Response vs ligand concentration plots for 60339 with hTeloC, c-Myc and double stranded DNA.

Ligand (NSC)	$K_D$ ( $\mu\text{M}$ )		
	hTeloC	c-Myc	DS
9037	x	$0.30 \pm 0.31$	x
13487	x	$0.05 \pm 0.03$	x
60339	$0.22 \pm 0.04$	$0.34 \pm 0.04$	x
143491	$0.23 \pm 0.05$	$0.33 \pm 0.09$	x
260594	x	$0.11 \pm 0.13$	x
308848	$0.15 \pm 0.01$	$0.30 \pm 0.03$	x
311153	x	$0.01 \pm 0.01$	x
317003	$0.07 \pm 0.05$	$0.10 \pm 0.04$	x
317605	$0.19 \pm 0.03$	$0.28 \pm 0.04$	x
345845	x	x	x
638432	x	x	x

Table 3.2.1.2  $K_D$  calculated for each ligand-DNA pair at pH 5.5. x = not determinable.

---

### 3.2.2 Pesky pH Strikes SPR

While it might have been possible to use these results to decide which ligands to progress through for further investigation a few concerns at the time prompted a reassessment of the results prior to further elimination. The increase in the response being observed was relatively low: the interaction being monitored between the ligand and the DNA was essentially at the lower limit of detection for SPR, which is normally used in this way for monitoring the interaction of much larger complexes. This meant that any factor which adversely affected the signal to noise ratio was having a significant impact on the results. Additionally, the negative responses being observed and consequent negative %R<sub>max</sub> values raised further questions. To immobilise the DNA on the surface of the streptavidin coated SPR chip the DNA sequences were biotinylated. Thus, by exploiting the extremely high affinity of biotin for streptavidin ( $K_D \sim 0.01 \text{ pM}$ )<sup>154</sup> the DNA could easily be attached to the chip's surface. The surface of the SPR chip was comprised of four separate flow cells the first of which was kept blank, hTeloC was immobilised on the second, c-Myc on the third and the double stranded DNA on the fourth. The first flow cell was left blank intentionally so that it could be used as a control. The response from the first flow cell was subtracted from the response of each of the other cells to account for any non-specific binding to any of the components on the chip's surface. Example data for the cause of the negative responses is shown in Figure 3.2.2.1: at the higher ligand concentrations the binding to the 'blank' flow cell 1 is greater than that of flow cell 2. Consequently, when the response was corrected for the non-specific binding to flow cell 1 the results became negative.

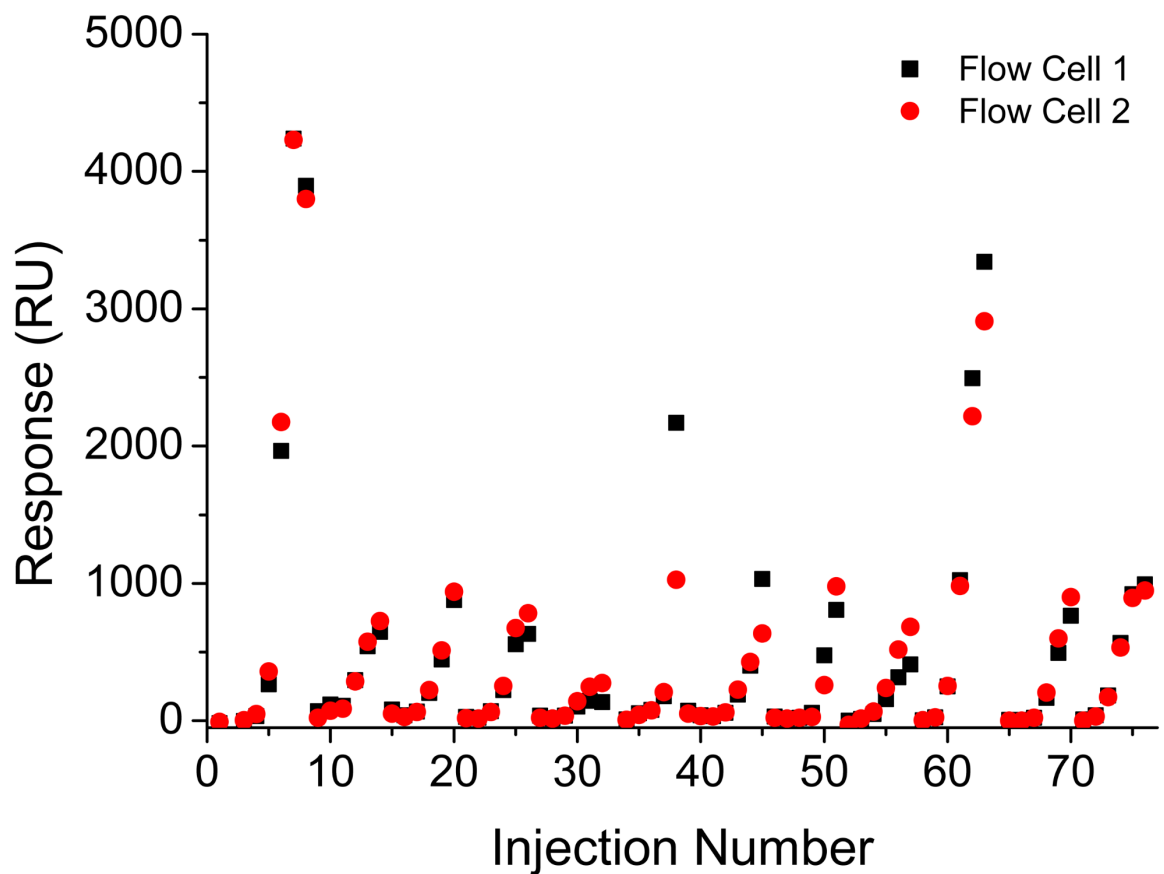
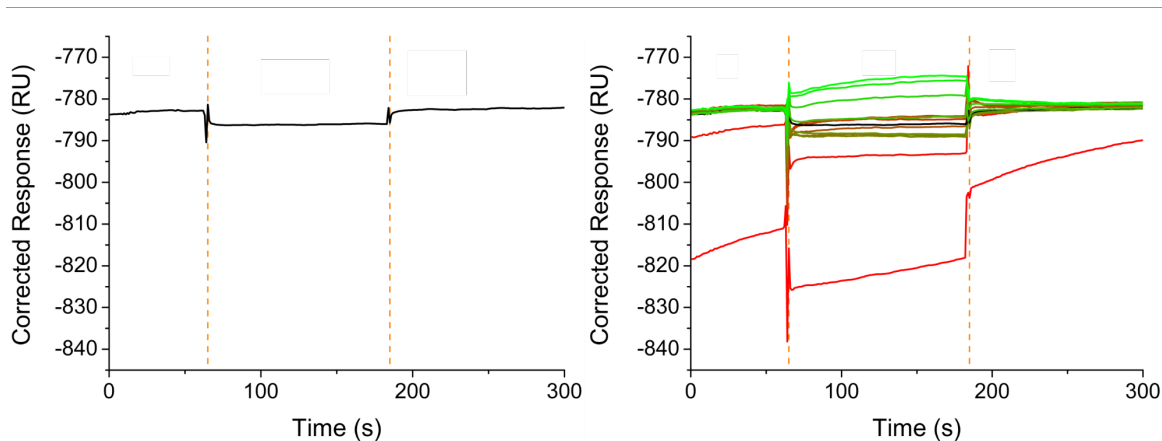


Figure 3.2.2.1 Response vs Injection Number for flow cell 1 (black) and flow cell 2 (red).

In considering the possible causes for these observations it was decided to also investigate whether the acidic pH being used to stabilise the i-motif (pH 5.5) was having an effect on the results. Figure 3.2.2.2 shows the results of this investigation where buffer alone was injected over the chip at different pH values. In SPR, a stock buffer is prepared and allowed to flow over the cells continuously to establish an equilibrium response level. Whenever anything is then injected, this is observable in the response from the sensorgrams as a result of turbulence introduced by the injection, even if the injection contains an aliquot of the same buffer at which the baseline response has equilibrated. What would then be expected to occur in the scenario of injecting identical buffer is a return to baseline response levels as the injected buffer flows over the cells. When the injection is complete (as determined by a preselected injection time window) there is another observable response as the flow from the stock buffer is re-established.



*Figure 3.2.2.2 Corrected Response vs Time where flow cell 1 response is subtracted from flow cell 2 response. Dashed orange lines indicate start and end of injection. Left is pH 7.0 alone. Right shows injection of buffer at different pH values. Colour scale transitions from green to red starting at pH 8 going down to pH 5 in 0.25-unit intervals; pH 7.0 is black.*

What was observed here raised some further concerns as the baseline unexpectedly did not return to the equilibrium level after the original buffer flow was restored. While a differential response in the sensorgrams during the flow of the injected buffer at different pHs over the flow cells could be explained by rearrangement of the DNA structures present on the surface, this should not have resulted in the observed warping of the baseline that occurred at low pH; especially considering that what is shown in Figure 3.2.2.2 is the corrected response observed after subtracting the response from flow cell 1 which was designed to be a control.

As a result of these concerns, it was decided to perform the SPR again at neutral pH. Twenty-five of the original 38 ligands were chosen for the neutral pH screen, the 13 chosen for elimination had either shown a higher % $R_{max}$  for double stranded DNA over the i-motif forming sequences, or a low % $R_{max}$  across the three structures. Two new i-motif forming sequences were selected for this screen which could form the structure at pH 7.0; these were ATXN2L: C<sub>24</sub>, and DAP: (C<sub>5</sub>G)<sub>4</sub>C<sub>5</sub>, which both had a reported transitional pH of 7.0.<sup>8</sup> These sequences were biotinylated and immobilised on the surface of a new chip and SPR testing using a concentration range (0 to 500  $\mu$ M) was performed to attempt to determine  $K_D$

values for these ligands with the new i-motifs. As before, where the quality control criteria were met  $K_D$  values were calculated and are presented in Table 3.2.2.1, and the response vs concentration plots generated in the attempt to determine these values are presented in Appendix A.4.

Ligand (NSC)	$K_D$ ( $\mu\text{M}$ )	
	ATXN2L	DAP
9037	0.04	0.37
13487	x	x
35676	x	0.20
60339	0.05	0.19
71795	0.03	0.16
103520	x	x
143491	0.18	0.02
146771	0.06	0.02
202386	0.30	0.02
204232	0.00	0.02
215718	x	x
260594	0.21	0.25
275428	x	x
277184	x	x
300289	x	x
305798	1.14	x
308848	0.05	0.03
308849	0.43	x
309892	x	0.19
311153	0.15	x
317003	x	525000.00
325014	x	129000.00
345647	x	0.32
345845	x	x
638432	x	x

Table 3.2.2.1  $K_D$  calculated for each ligand-DNA pair at pH 7.0. x = not determinable.

---

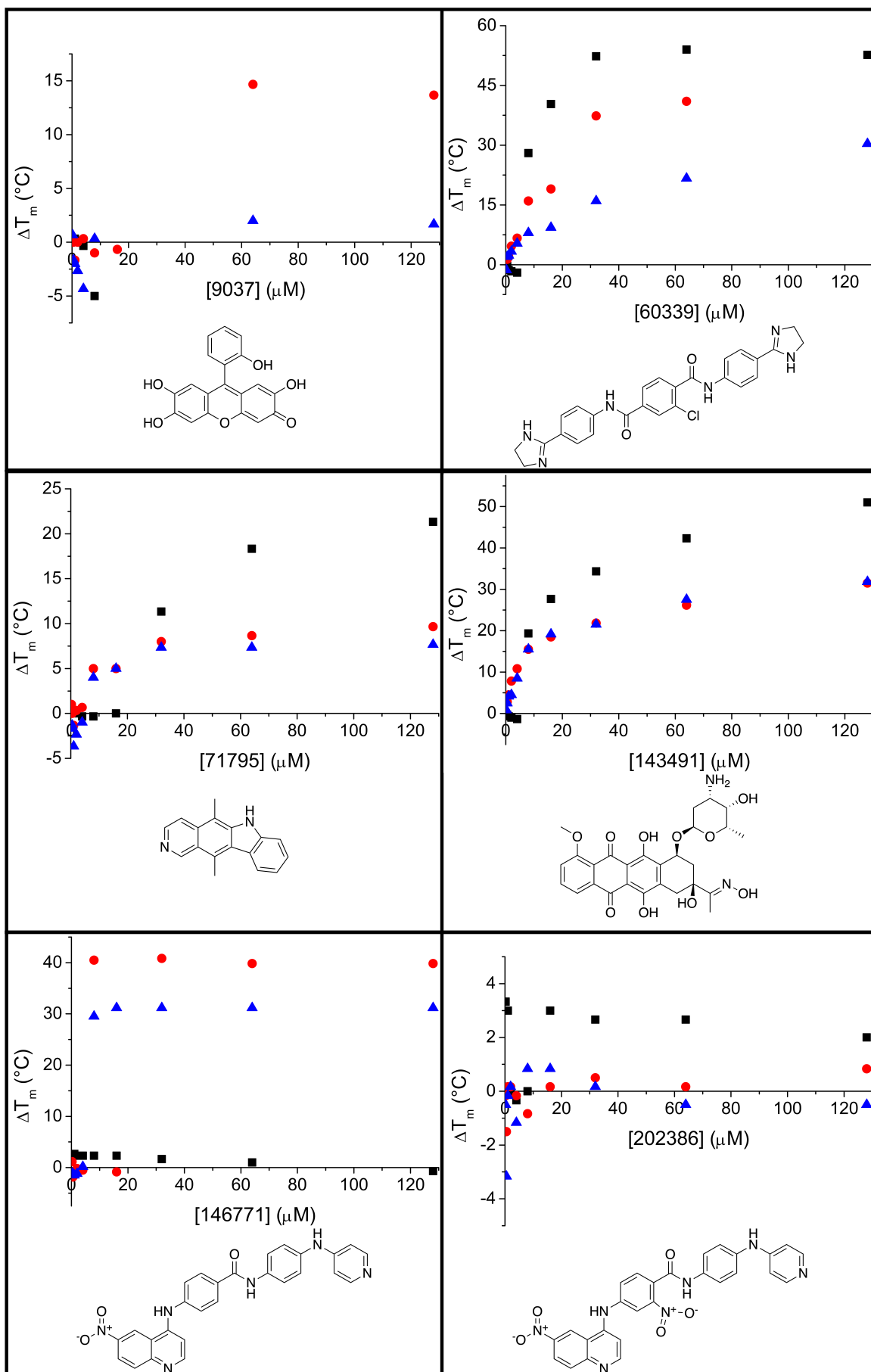
While further data was able to be collected under the new neutral conditions some of the aforementioned problems from the screen performed at acidic pH persisted. It appeared that some of this may have been inescapable due to the relatively small scale of the interaction being monitored lying near the boundary of detection by SPR. Nevertheless, it was rationalised that those ligands which had yielded positive results warranted further investigation due to the ability to detect their interaction with the DNA despite the limitations inherent to the method.

### 3.2.3 FRET-melting

The SPR results provided an idea of the ability of the ligands to bind to the DNA however they did not give an indication of their effect on the stability of the structures under examination. Eight ligands were chosen from the SPR screens for investigation using FRET-melting to determine their effects on the thermal stability of the DNA. It was decided that these experiments would be performed at neutral pH and thus DAP was selected as the candidate i-motif forming sequence as it had a higher melting temperature than ATXN2L.<sup>8</sup> To determine whether the ligands showed potential selectivity in their ability to stabilise the i-motif over other DNA structures a construct designed to mimic double stranded DNA was also used which consisted of two complementary 10-base sequences linked by an 18-unit hexaethylene glycol (HEG) polymer: DS<sub>FRET</sub> (FAM-TATAGCTATA-HEG(18)-TATAGCTATA-TAMRA). Finally, the guanine rich sequence from the human telomere: hTeloG (G<sub>3</sub>T<sub>2</sub>A)<sub>3</sub>G<sub>3</sub> was chosen as the candidate G-quadruplex.

The FRET-melting experiments were performed over a concentration range from 0 to 128  $\mu$ M and the change in melting temperature ( $\Delta T_m$ ) induced was calculated for each ligand-DNA pair at each concentration. The change in melting temperature results are shown in Figure 3.2.3.1 and the temperature values are provided in Appendix A.5. The FRET-melting curves from which the  $\Delta T_m$  values were derived are presented in Appendix A.6.





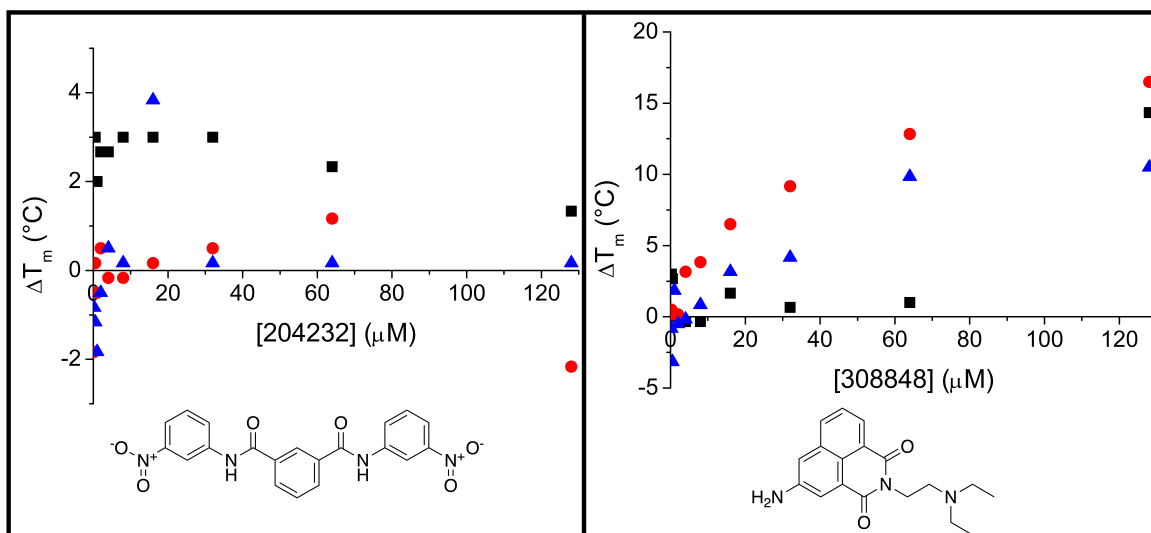


Figure 3.2.3.1  $\Delta T_m$  of 200 nM FRET-labelled DNA (black square ■ = DAP, red circle ● = DS, blue triangle ▲ = hTeloG) in 10 mM sodium cacodylate buffer at pH 7.0 with 0 to 128  $\mu\text{M}$  ligand.

Based solely on the change in melting temperature values it seemed that some of these results were promising. Three ligands, **60339**, **71795** and **143491**, appeared to stabilise DAP to a greater extent than the double stranded or G-quadruplex structures. However, the relatively large increase in the melting temperature in these instances was being observed at ratios where the ligands were in huge excess compared to the DNA. The highest ligand concentration used was 128  $\mu\text{M}$  which represents a 640-fold excess compared to the DNA, the concentration of which was 200 nM. The concern was that this stabilisation, particularly the differential stabilisation of the i-motif, was not observed at the lower ligand concentrations which represent more equivalent ratios of ligand to DNA. It would be expected that if there was a genuine ligand-DNA binding interaction, especially if it was a specific interaction with a particular structure, that it would occur at ratios where the ligand is equivalent to the DNA, or only in a small excess representative of possible multiple binding sites. The observation of stabilisation when the ligands were in such a large excess suggested that some non-specific interaction might have been responsible for the change in melting temperature; potentially explained by a reduced solubility of the ligand-DNA complex, or an interaction with the fluorescent tags rather than with the DNA.

---

### 3.2.4 NMR

To probe this relationship further, NMR experiments were carried out on the same eight NSC ligands from the FRET-melting during a research visit to the group of Lukáš Trantírek at the Central European Institute of Technology (CEITEC) in the Czech Republic. Trantírek's group have expertise with using NMR to study the i-motif and used it to provide the first *in vivo* evidence for the i-motif in the nuclei of living cells.<sup>60</sup> The ligands were titrated against DAP and spectra were measured in the absence of ligand followed by a gradual increase in the proportion of ligand. The i-motif structure was monitored as a function of the intensity of the imino proton signal in the NMR spectra at ~15.5 ppm.<sup>56,60</sup> In contrast to the NMR described in chapter 2, where spectra were collected over two hours using 10 µM DNA, to enable collection of spectra in a more reasonable time (~15 minutes) the concentration of the DNA used here was 100 µM. Unfortunately, this meant that these experiments were plagued by poor solubility. Most of the ligands could not be measured at a DNA:ligand ratio higher than 1:2 as visible precipitate formed in the NMR tube and the quality of the signal deteriorated substantially. Furthermore, from the data that was able to be collected, it appeared that the ligands, with the exception of **9037** and **308848**, destabilised the i-motif. When accounting for the destabilisation of the i-motif induced by an amount of DMSO equivalent to that added to achieve each DNA:ligand ratio, **9037** appeared to have no effect on the structure and **308848** had a very slight stabilising effect that was within error (NMR spectra in Appendix A.7). As with the other tested ligands, **143491** appeared to destabilise the i-motif in a concentration-dependent manner as observed by the disappearance of the imino proton signals at ~15.5 ppm which represented the loss of C·C<sup>+</sup> base pairs which form the i-motif. However, in contrast to the other ligands, addition of **143491** resulted in the appearance of signals with chemical shifts between 12-14 ppm which correspond to the formation of Watson-Crick base pairs (Figure 3.2.4.1).<sup>120</sup> This data suggested that the DAP sequence, which consisted of only cytosines and guanines, may have been adopting an alternative conformation, possibly some form of hairpin stabilised by C·G base pairs.

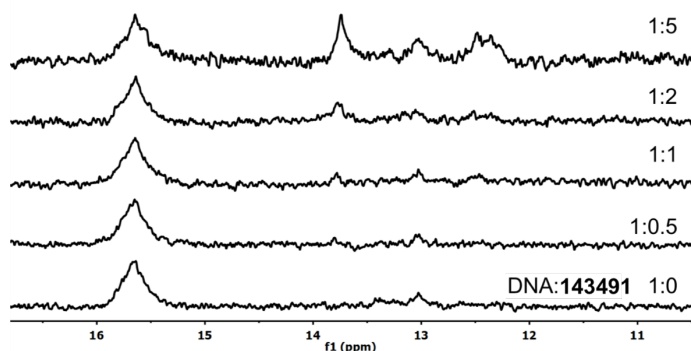


Figure 3.2.4.1 Imino region of 1D  $^1\text{H}$  NMR spectra of 100  $\mu\text{M}$  DAP in 140 mM  $\text{NaPO}_i$ , 10 mM  $\text{MgCl}_2$ , 5 mM KCl buffer at pH 7.0 with 10%  $\text{D}_2\text{O}$  at 25°C with increasing concentration of **143491**.

Unfortunately, even in the case of **143491** solubility continued to present an issue, precluding measurements with more than five equivalents of ligand. However, an interesting observation was made where, in the presence of five equivalents of **143491**, the proportion of DAP folded into the i-motif conformation continued to diminish over time while the proportion of the alternative hairpin conformation increased (Figure 3.2.4.2). It should be noted that the imino proton signals from the  $\text{C}\cdot\text{C}^+$  base pairs at ~15.5 ppm did not disappear completely and thus it was possible that the alternative conformation which was being adopted consisted of a combination of both the  $\text{C}\cdot\text{C}^+$  base pairs from the i-motif with additional canonical Watson-Crick  $\text{C}\cdot\text{G}$  base pairs.

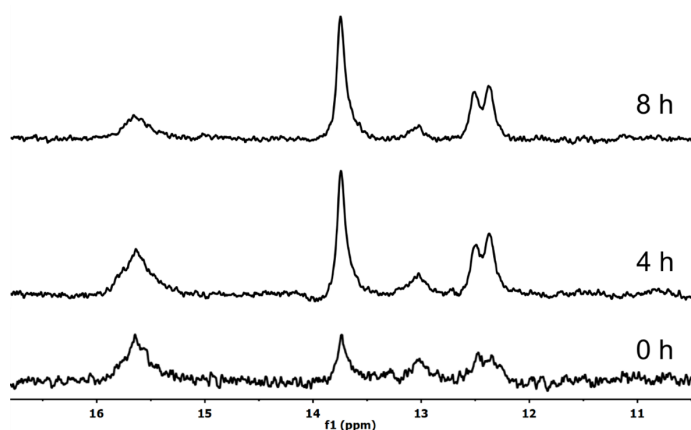
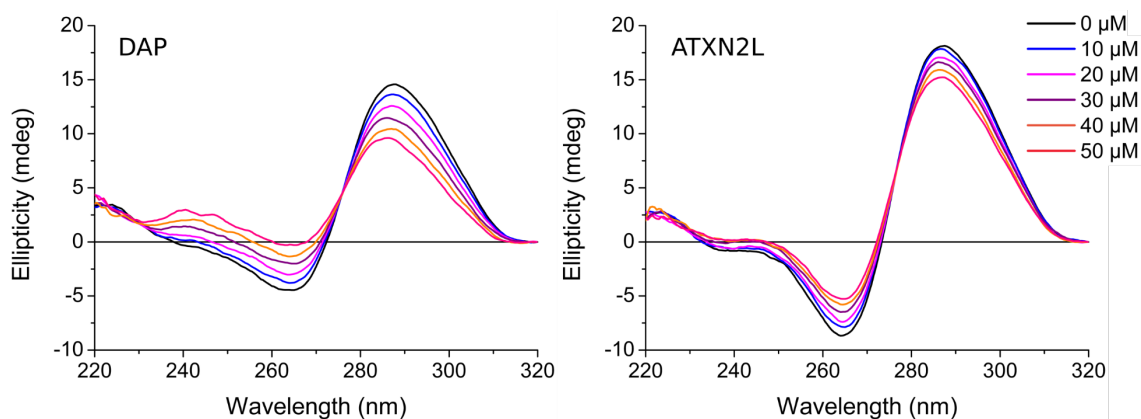


Figure 3.2.4.2 Imino region of 1D  $^1\text{H}$  NMR spectra of 100  $\mu\text{M}$  DAP in 140 mM  $\text{NaPO}_i$ , 10 mM  $\text{MgCl}_2$ , 5 mM KCl buffer at pH 7.0 with 10%  $\text{D}_2\text{O}$  at 25°C showing effect of incubation with 5 equivalents of **143491** over time.

### 3.2.5 The Curious Case of 143491

Following the research visit, **143491** was investigated further using CD titrations and CD and UV melting experiments. In addition to DAP, it was decided to also investigate this ligand's effects on the ATXN2L i-motif as, while both sequences have 24 cytosines, the four guanines interspersed in the DAP sequence are absent in ATXN2L. The lack of these guanines meant that the formation of C·G base pairs was impossible and thus there was an interest in determining what effect, if any, **143491** would have on the ATXN2L i-motif. Monitoring the titration of the ligand by CD into the two i-motifs showed a similar effect: in both instances a minor hypsochromic shift of the peak at ~288 nm is observed in conjunction with a slight hypochromic shift across the spectra. Interestingly, while the changes observed were the same, the extent of these changes was greater for DAP than for ATXN2L (Figure 3.2.5.1).



*Figure 3.2.5.1 CD spectra of 10  $\mu$ M DAP (left) and ATXN2L (right) with titration up to 50  $\mu$ M **143491** in 10 mM sodium cacodylate buffer at pH 7.0.*

For the CD titrations the ligand concentration was increased up to 50  $\mu$ M which, in keeping with the NMR experiments, represented a DNA:ligand ratio of 1:5. The spectrum of ATXN2L in the presence of five equivalents of **143491** appeared to still be i-motif with the position of the quintessential large positive peak at ~288 nm and the negative signal at ~265 nm similarly being maintained. In the case of DAP while the positions of the peaks did not change in a dramatic fashion, there was a more pronounced change in the intensity of the signals being observed resulting in an almost complete disappearance of the negative signal at ~265 nm.

To determine the effects of **143491** on the stability of the i-motifs CD and UV melting experiments were performed and melting temperatures calculated for the structures in the presence and absence of five equivalents of the ligand.

	DAP		ATXN2L	
	$T_m$ (°C)			
	CD melting	UV melting	CD melting	UV melting
DNA	38.69 ± 0.27	38.04 ± 0.25	35.50 ± 0.29	35.07 ± 0.18
DNA + <b>143491</b> 1:5	32.64 ± 0.38	31.86 ± 0.41	33.42 ± 0.30	32.76 ± 0.32
$\Delta T_m$ (°C)	-6.05	-6.18	-2.08	-2.31

Table 3.2.5.1  $T_m$  of DAP and ATXN2L measured by CD and UV in the absence and presence of 5 equivalents of **143491**.

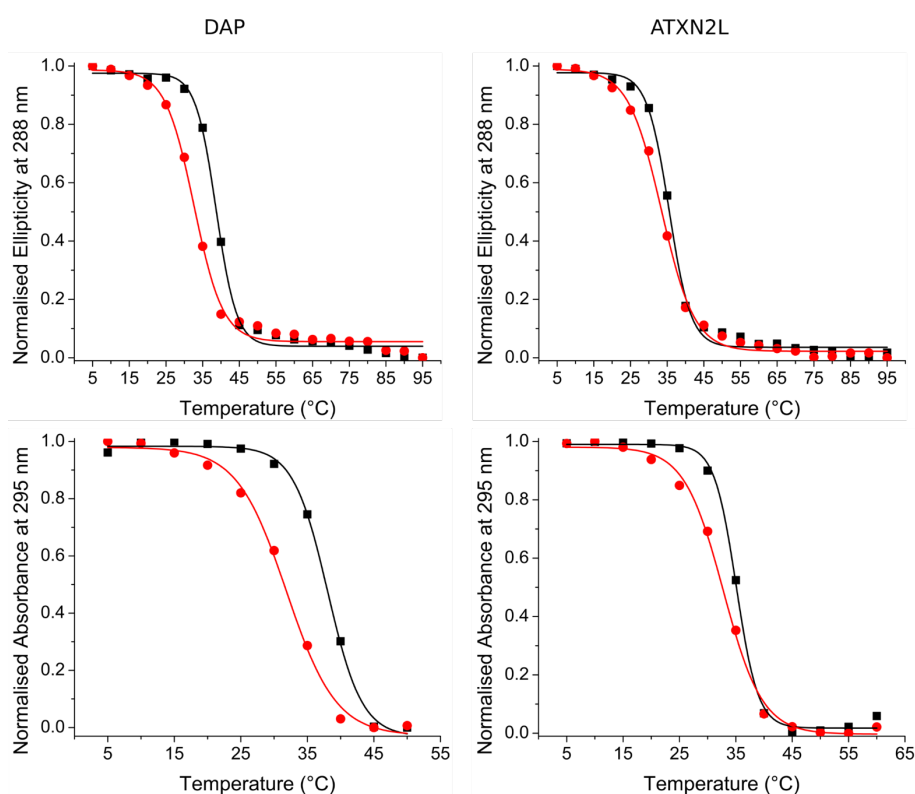
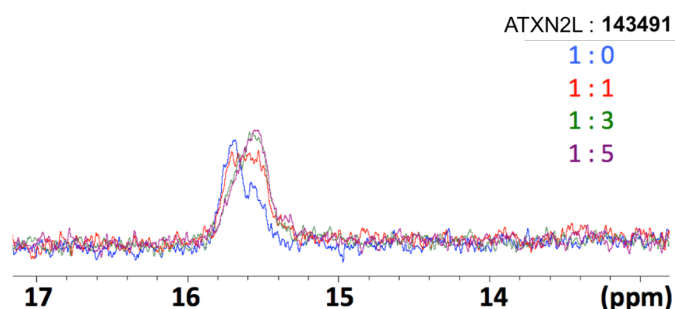


Figure 3.2.5.2 Normalised Ellipticity at 288 nm (top) and normalised absorbance at 295 nm (bottom) of 10  $\mu$ M DAP (left) and ATXN2L (right) without ligand (black) and with 5 equivalents (50  $\mu$ M) **143491** (red). Experiments performed at pH 7.0 in 10 mM sodium cacodylate buffer.

From this data it could be seen that **143491** destabilised the i-motif formed by both sequences and, in keeping with the trend observed from the CD titrations, the extent of the ligand's effect were more pronounced with DAP: the decrease in melting temperature for DAP in the presence of the ligand was  $\sim 6^{\circ}\text{C}$  whereas it for ATXN2L the decrease was only  $\sim 2^{\circ}\text{C}$ .

Further NMR experiments were then carried out by Martin Gajarský from the Trantírek group to see what effects **143491** had on ATXN2L. Unsurprisingly, because ATXN2L is made up of only cytosines, no signals were observed between 12-14 ppm which would have corresponded to the Watson-Crick pairs which were induced with DAP. What was somewhat surprising however was that, contrary to what had been observed previously with all the ligands, there was no apparent reduction in the imino proton signals involved in the C·C<sup>+</sup> base pairs which formed the i-motif (Figure 3.2.5.3).



*Figure 3.2.5.3 Spectra overlap showing shifts in imino proton signals from 1D <sup>1</sup>H NMR spectra of 100  $\mu\text{M}$  ATXN2L in 140 mM NaPO<sub>i</sub>, 10 mM MgCl<sub>2</sub>, 5 mM KCl buffer at pH 7.0 with 10% D<sub>2</sub>O at 25°C with increasing concentration of **143491**.*

Additionally, there were changes in the positions of the signals in this region as the ligand was added. The cause of these changes was suggested to range from the ligand binding to the i-motif resulting in a simple alteration of the chemical environment of the imino protons, to a possible reconfiguration of the original i-motif conformation to another. In an attempt to further elucidate what effect **143491** was having on the i-motif structures three more i-motif forming sequences were examined: JAZF (C<sub>8</sub>G(C<sub>5</sub>G)<sub>2</sub>C<sub>3</sub>TC<sub>6</sub>), PDGF-A (C<sub>2</sub>GCGC<sub>4</sub>T(C<sub>5</sub>G)<sub>3</sub>C<sub>13</sub>) and hTeloC. The JAZF and PDGF-A sequences contained guanines and addition of the ligand

resulted a reduction of the C-C<sup>+</sup> imino proton signals and the appearance of the same Watson-Crick imino signals as observed with DAP (Figure 3.2.5.4). The hTeloC sequence did not contain guanines and, unlike ATXN2L, there was a clear destabilisation of the i-motif as seen by the decrease in the C-C<sup>+</sup> imino proton signals at ~15.5 ppm. Also, additional signals did appear in the region typical for Watson-Crick base pairs, although in this case these signals did not correspond to the those observed with the sequences containing guanine.

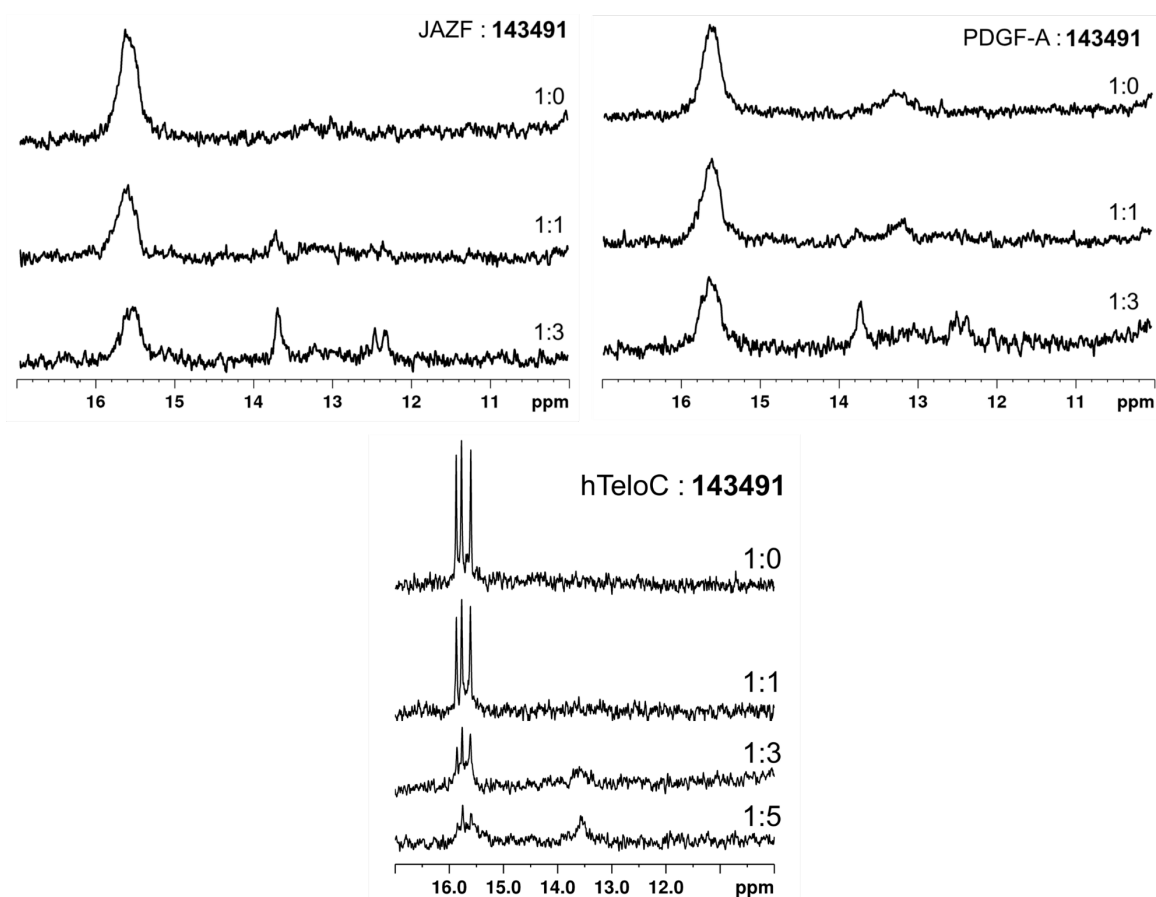


Figure 3.2.5.4 Imino region of 1D <sup>1</sup>H NMR spectra of 100 μM DNA in 140 mM NaPO<sub>i</sub>, 10 mM MgCl<sub>2</sub>, 5 mM KCl buffer with 10% D<sub>2</sub>O at 25°C with increasing concentration of **143491**. JAZF and PDGF-A at pH 7.0 and hTeloC at pH 6.0.

Unfortunately, while further experiments were performed in an attempt to understand this interaction better, with each change in sequence or environmental variable **143491** appeared to have different effects. These studies were started in an attempt to find a ligand which stabilised the i-motif. Consequently, while this



---

ligand was having interesting behaviours – ranging from induction of non-C·C<sup>+</sup> base pairs in the loops, to conformational rearrangement of the i-motifs, and with apparent sensitivity to the composition of the sequence – it was clearly not stabilising the i-motif in any of these instances and was thus not examined further.

### 3.3 The Effects of G-quadruplex Ligands on i-Motif DNA

Exasperation from the lack of selectivity being observed between the two quadruplexes from the screens discussed in the previous section led to the decision to investigate G-quadruplex selective ligands in an attempt to understand what gave them selectivity for the G-quadruplex over the i-motif. Several studies had identified ligands which were presented as having specific binding to the G-quadruplex, however these were not originally tested on the complementary i-motif structure. Additionally, while working on developing the i-motif specific antibody iMab, Christ and co-workers found that the antibody used to confirm the *in vivo* existence of the G-quadruplex in cells was also able to interact with the i-motif.<sup>61</sup> As a result of the omission of the i-motif in the identification of ‘specific’ G-quadruplex ligands, it was possible that these reported ligands might have had an as yet undiscovered i-motif-stabilising effect as well.

#### 3.3.1 i-Motif ‘Home Advantage’ – Acidic pH

Initial efforts to explore the interaction of G-quadruplex ligands with the i-motif were performed in collaboration with the group of Antonio Randazzo at the University of Naples Federico II in Italy. Six G-quadruplex ligands were selected: **Berberine**,<sup>155</sup> **BRACO-19**,<sup>156</sup> **Mitoxantrone**,<sup>157</sup> **Phen-DC3**,<sup>158</sup> **Pyridostatin**,<sup>159</sup> and **RHPS4**<sup>160</sup> (Figure 3.3.1.1). Experiments were performed at pH 4.3 and pH 5.7 to explore the interaction of the six ligands with an i-motif and a G-quadruplex using CD, UV, NMR, FRET and FID.

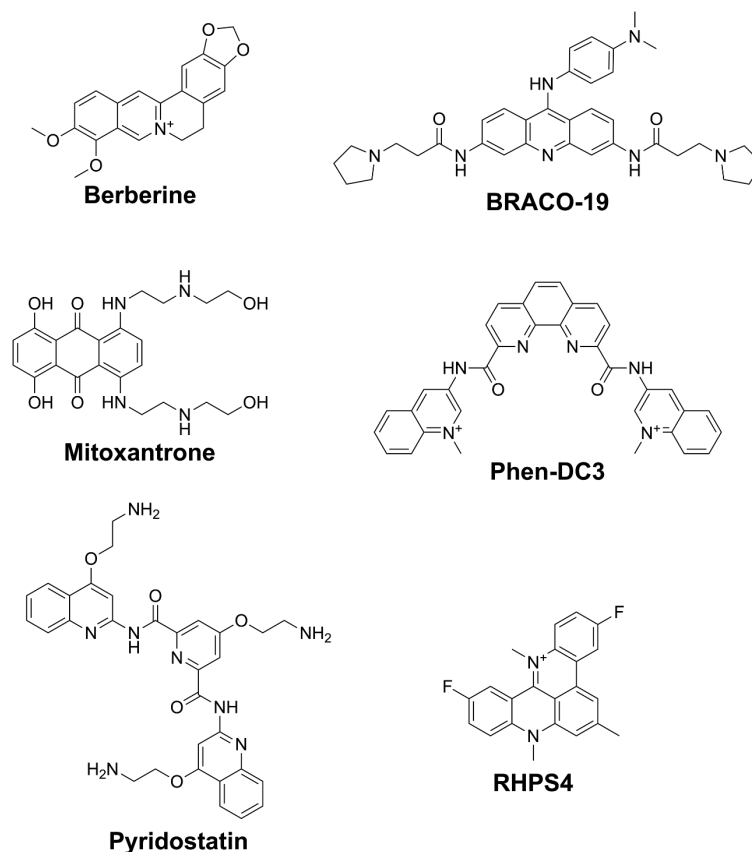


Figure 3.3.1.1 Structures of the ligands investigated at acidic pH.

The biophysical experiments confirmed that the ligands all interacted with both quadruplex structures.<sup>100</sup> It was decided that the concentration of the ligand would not be increased beyond five molar equivalents of ligand with respect to the DNA. This was rationalised as potentially representing an excess of ligand to any potential binding sites on the structure. Under these conditions, the CD and UV melting experiments performed by our collaborators showed that, as expected, all the ligands resulted in stabilisation of the G-quadruplex at both pH values. Interestingly, also at both pH values, it was found that while **Berberine**, **Pyridostatin** and **RHPS4** did not have a significant effect on the melting temperature of the i-motif, **BRACO-19**, **Mitoxantrone** and **Phen-DC3** significantly destabilised the structure (Table 3.3.1.1).<sup>100</sup>

	$\Delta T_m$ (°C) – CD melting				$\Delta T_m$ (°C) – UV melting				
	i-motif		G-quadruplex		i-motif		G-quadruplex		
	pH	4.3	5.7	4.3	5.7	4.3	5.7	4.3	5.7
<b>Berberine</b>		-2.5	-0.8	13.4	12.4	1.2	0.9	11.7	15.3
<b>BRACO-19</b>		-13.4	-9.2	12.4	8.9	-17.6	-6.5	x	x
<b>Mitoxantrone</b>		-4.8	-9.9	7.7	4.5	-16.6	-1.0	4.1	0.6
<b>Phen-DC3</b>		-13.4	-6.8	x	14.4	-17.1	-6.3	x	5.3
<b>Pyridostatin</b>		-2.8	0.8	12.9	8.8	1.1	0.2	9.3	5.5
<b>RHPS4</b>		-1.0	-0.3	22.0	20.6	0.7	0.1	x	x

*Table 3.3.1.1  $\Delta T_m$  of ligand-DNA pairs measured by CD and UV melting using 5 equivalents of each ligand. All experiments were performed in duplicate and  $\Delta T_m$  values are reported as the mean. Errors were  $\pm 0.5^\circ\text{C}$ . x = not determinable.*

Following on from this, using an FID assay the concentration at which each of the ligands displaced 50% of the fluorescent probe ( $\text{DC}_{50}$ ) was determined (Table 3.3.1.2), **RHPS4** was excluded as its fluorescence profile overlapped with the assay parameters. As this assay relied on displacement, rather than occupation of all potential binding sites, the concentration range used for each ligand was selected to provide sufficient data points above and below the  $\text{DC}_{50}$  to allow for reliable fitting of the data. Aliquots of the ligands were titrated in triplicate against both quadruplex structures and the  $\text{DC}_{50}$  values calculated from dose-response curves fitted to this data (Appendix A.8).

	$\text{DC}_{50}$ ( $\mu\text{M}$ )				
	pH	i-motif		G-quadruplex	
		4.3	5.7	4.3	5.7
<b>Berberine</b>		$30.38 \pm 0.02$	$1.46 \pm 0.01$	$3.32 \pm 0.03$	$1.26 \pm 0.01$
<b>BRACO-19</b>		$0.66 \pm 0.00$	$0.87 \pm 0.00$	$0.26 \pm 0.01$	$0.50 \pm 0.00$
<b>Mitoxantrone</b>		$0.70 \pm 0.00$	$1.34 \pm 0.01$	$0.54 \pm 0.00$	$0.95 \pm 0.01$
<b>Phen-DC3</b>		$0.97 \pm 0.01$	$0.95 \pm 0.00$	$0.26 \pm 0.00$	$0.39 \pm 0.00$
<b>Pyridostatin</b>		$9.09 \pm 0.01$	$18.02 \pm 0.06$	$3.15 \pm 0.00$	$9.42 \pm 0.04$

*Table 3.3.1.2  $\text{DC}_{50}$  values for ligand-DNA pairs at pH 4.3 and pH 5.7 determined using the FID assay. Experiments performed in triplicate and  $\text{DC}_{50}$  values are reported as the 50% displacement value calculated from fitted dose response curves. Standard errors are calculated using R-square values from the statistics on the data fit.*

---

The results of the FID assay showed that the five ligands which could be tested all bound to both quadruplex structures. While all six ligands showed a slightly higher affinity for the G-quadruplex over the i-motif, this assay further confirmed that at these pH values these 'G-quadruplex ligands' all also interacted with the i-motif. Despite these results, the fear that the i-motif would continue to be dismissed persisted: it could be argued that the acidic pH at which these interactions were studied was not biologically relevant. Consequently, the i-motif could continue to be excluded from future ligand-quadruplex studies aimed at biological or therapeutic applications. To address this concern, it was decided to perform a further examination of G-quadruplex ligands using i-motifs at neutral pH.

### 3.3.2 i-Motif 'Plays Away' – Neutral pH

Owing to the nature of the i-motif generally being more stable at acidic pH almost all previous efforts to examine the interaction between small molecules and the i-motif have been performed under acidic pH conditions.<sup>100</sup> While the effect of pH on the sequence under investigation is discussed at length, rarely is consideration given to the effect of pH on the ligand itself. Clearly each small molecule will be affected by the pH in a different manner depending on the functional groups that it contains. This will consequently have an effect on each given ligand's interaction with different DNA structures under different conditions. Therefore, if the aim is to develop applications for these compounds in biological conditions, it is important to examine their interaction with the different quadruplex DNA structures at neutral pH as well. With this in mind, the same six G-quadruplex binding compounds described in the previous section, in addition to **TmPyP4** (Figure 3.3.2.1),<sup>161</sup> were used to explore how these ligands interacted with two candidate i-motif forming sequences at neutral pH.

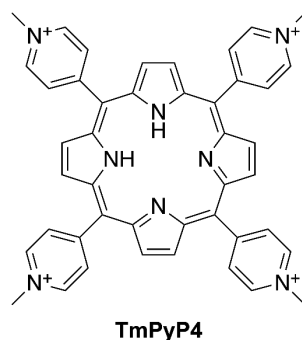


Figure 3.3.2.1 Structure of **TmPyP4**.

Three techniques were used to monitor the interactions of these ligands with the ATXN2L and DAP i-motifs at neutral pH: CD titrations were used to determine the effects of the compounds on i-motif structure, CD and UV melting were used to observe changes in thermal stability upon addition of the ligand, and an FID assay was used to compare the affinities of the ligands. Only DAP was used in the FID experiments as ATXN2L is a poor candidate for this technique due to the low fluorescence enhancement observed after equilibration of the fluorescent probe. While using a lower pH would have resolved this issue, to allow meaningful comparisons we consistently use 10 mM sodium cacodylate buffer at pH 7.0 to better address the scepticism surrounding the physiological relevance of the i-motif.

CD was used first to confirm that DAP and ATXN2L formed i-motif structures under our experimental conditions at pH 7.0. Both sequences showed the characteristic positive peak at 288 nm and a negative one at 260 nm. This was followed by titration of each of the ligands in 1 equivalent (10  $\mu$ M) increments up to 5 equivalents (50  $\mu$ M) as described before. At these concentrations, **Berberine** and **Pyridostatin** did not have significant effects on the spectra of either i-motif, with **Pyridostatin** resulting in only a minimal hypochromic shift of the peak at 288 nm. By contrast, **BRACO-19**, **Mitoxantrone**, **Phen-DC3** and **RHPS4** resulted, to varying extents, in considerably greater hypochromic shifts of both the positive band at 288 nm and the negative band at 260 nm, as well as a small hypsochromic shift of the positive band at 288 nm, consistent with an apparent unfolding effect.<sup>162</sup> The DAP i-motif appears to be affected by these ligands to a greater extent compared with ATXN2L; this was most apparent when observing the negative band at 260 nm which came very close to zero and almost lost its negative

---

character in the cases of **BRACO-19** and **Phen-DC3**. The effects of **TmPyP4** were most apparent, as it resulted in the most significant hypochromic shifts of the bands at 288 and 260 nm for both sequences. Interestingly, while the bands at 288 nm initially exhibited a hypsochromic shift similar to that seen with the other ligands as their concentration was increased, with **TmPyP4** at 40 and 50  $\mu\text{M}$  a bathochromic shift was then observed. This data (Figure 3.3.2.2 and 3.3.2.3) suggested that these ligands were interacting with the i-motif structures formed by ATXN2L and DAP. The hypochromicity observed suggested that the ligands were inducing a partial unfolding the i-motif structure. While **TmPyP4** appeared to do the same at the lower concentrations, the subsequent bathochromic shift suggested that perhaps another structure was then adopted.

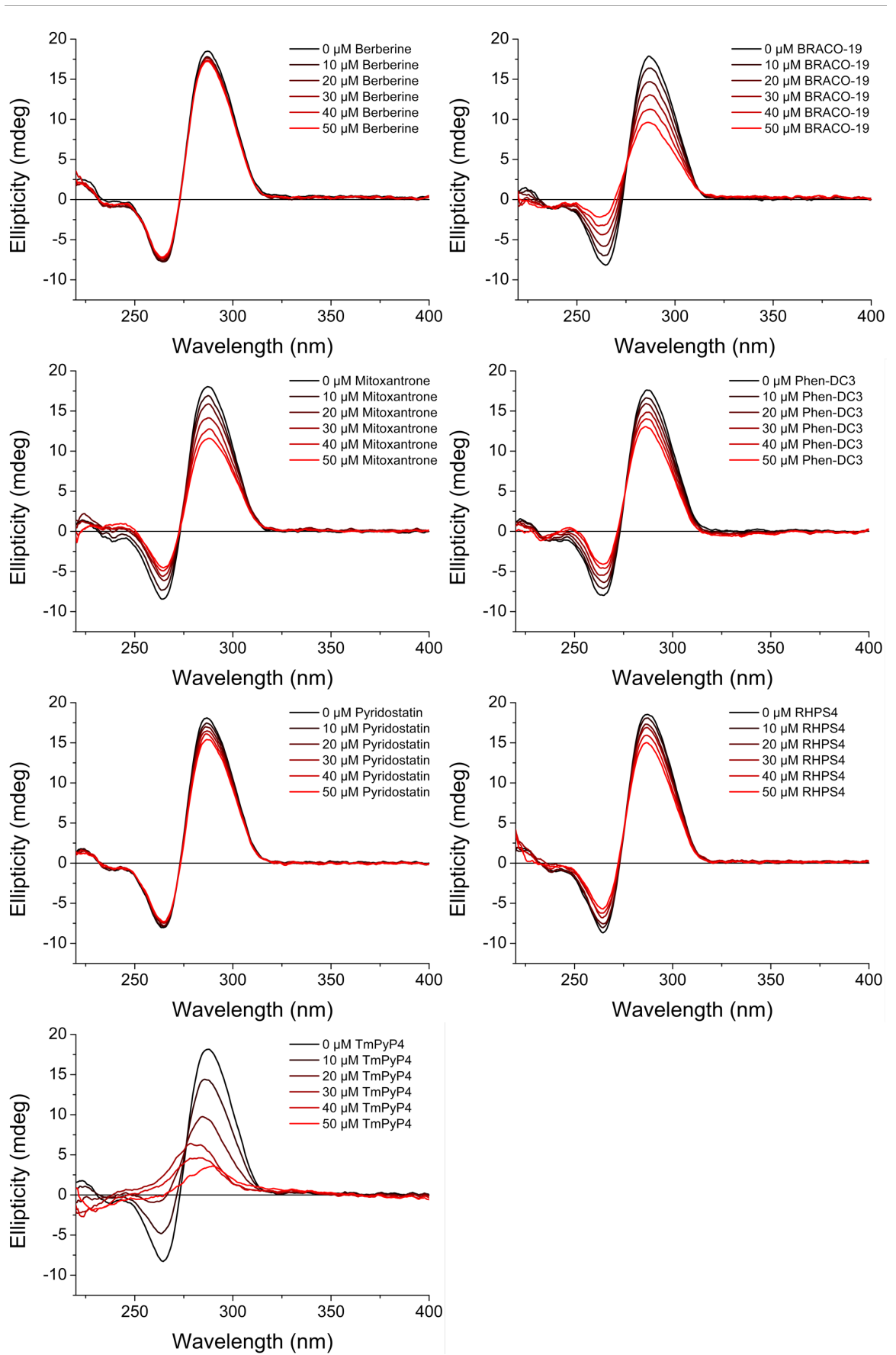


Figure 3.3.2.2 CD spectra of 10  $\mu\text{M}$  ATXN2L with titration up to 50  $\mu\text{M}$  ligand in 10 mM sodium cacodylate buffer at pH 7.0.

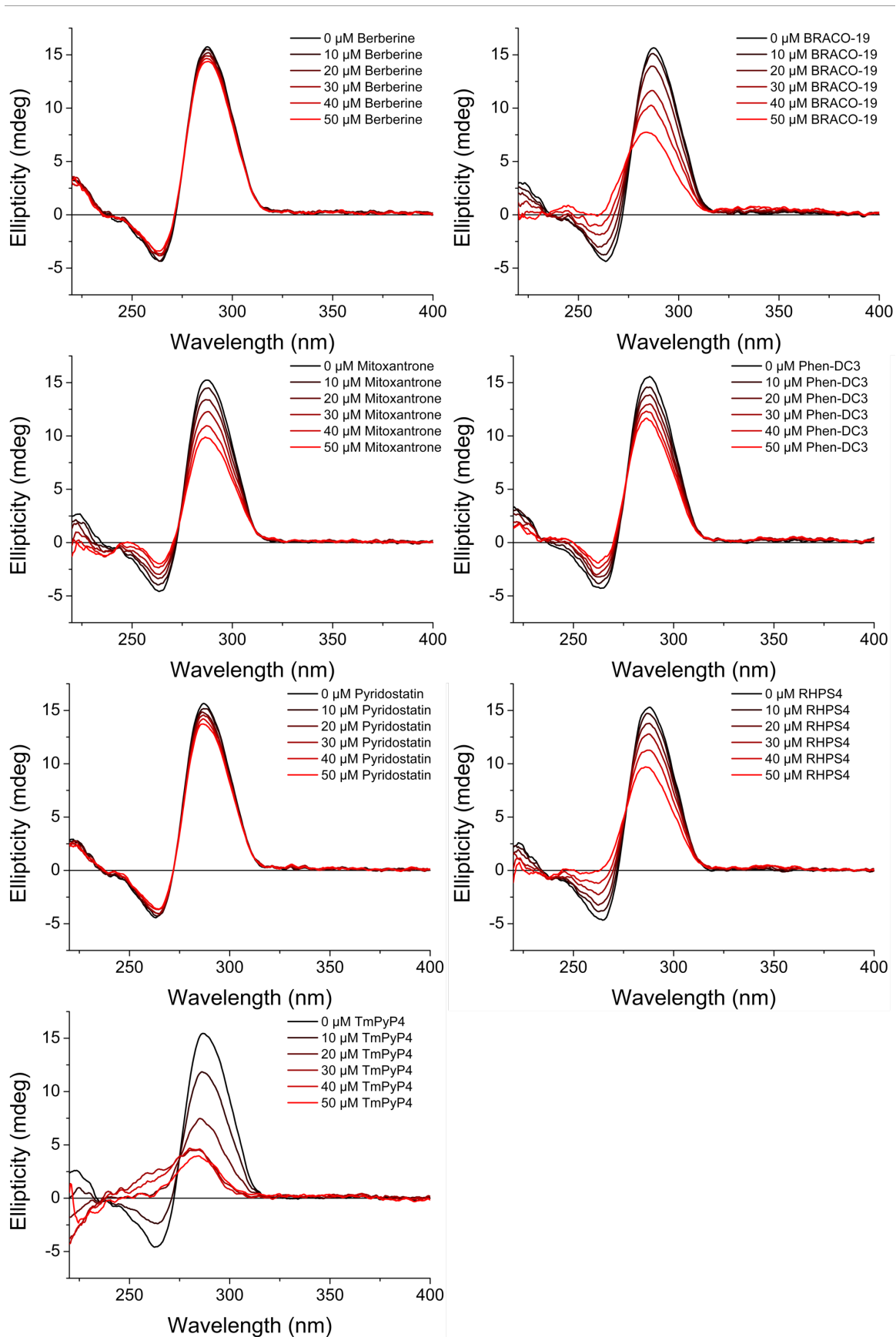


Figure 3.3.2.3 CD spectra of 10  $\mu\text{M}$  DAP with titration up to 50  $\mu\text{M}$  ligand in 10 mM sodium cacodylate buffer at pH 7.0.



To assess the effects of the ligands on the stability of the i-motif structures, the melting temperatures for ATXN2L and DAP were determined in the absence of ligands using both CD and UV melting experiments. The results from both techniques were in very good agreement with one another: using CD to monitor ellipticity at 288 nm the  $T_m$  for ATXN2L was determined to be  $36.9 \pm 0.2^\circ\text{C}$  and for DAP it was  $39.7 \pm 0.1^\circ\text{C}$ ; using UV and monitoring absorbance at 295 nm the  $T_m$  for ATXN2L was determined to be  $36.0 \pm 0.02^\circ\text{C}$  and for DAP it was  $38.6 \pm 0.2^\circ\text{C}$ . Apart from **TmPyP4**, all the ligands showed a destabilisation of the i-motif structure; with the destabilisation of DAP being more pronounced than ATXN2L (Table 3.3.2.1, Appendix A.9). Due to the significant changes in the CD spectra of the i-motifs in the presence of **TmPyP4**, a  $T_m$  could not be determined using this method.

Ligand	$\Delta T_m$ ( $^\circ\text{C}$ )	
	ATXN2L	DAP
<b>Berberine</b>	$-1.5 \pm 0.3$	$-3.9 \pm 0.3$
<b>BRACO-19</b>	$-6.4 \pm 0.3$	$-7.3 \pm 0.7$
<b>Mitoxantrone</b>	$-1.0 \pm 0.4$	$-7.1 \pm 0.3$
<b>Phen-DC3</b>	$-1.4 \pm 0.3$	$-3.0 \pm 0.4$
<b>Pyridostatin</b>	$-0.8 \pm 0.3$	$-1.8 \pm 0.3$
<b>RHPS4</b>	$-1.4 \pm 0.3$	$-6.8 \pm 0.4$
<b>TmPyP4</b>	x	x

Table 3.3.2.1  $T_m$  is the midpoint of the transition from each CD melting experiment, and SE is calculated using R-square values from the statistics on the data fit.  $\Delta T_m$  is the difference between the  $T_m$  of the DNA in the presence of 5 equivalents of each ligand and the DNA on its own. x = not determined.

Additionally, it was found that the extent of destabilisation of the i-motif observed using CD melting in the presence of the ligands correlated very well with the intensity of the hypochromicity observed in the titrations ( $r = -0.7691$ ,  $p = 0.0035$ , Table 3.3.2.2).

	$T_m$ (°C)	% Hypochromicity
ATXN2L + 50 $\mu$ M Berberine	35.4	0.0709
ATXN2L + 50 $\mu$ M BRACO-19	30.5	0.4614
ATXN2L + 50 $\mu$ M Mitoxantrone	35.9	0.3557
ATXN2L + 50 $\mu$ M Phen-DC3	35.5	0.2686
ATXN2L + 50 $\mu$ M Pyridostatin	36.1	0.1432
ATXN2L + 50 $\mu$ M RHPS4	35.5	0.1972
ATXN2L + 50 $\mu$ M TmPyP4	x	0.806
DAP + 50 $\mu$ M Berberine	35.8	0.0852
DAP + 50 $\mu$ M BRACO-19	32.4	0.5244
DAP + 50 $\mu$ M Mitoxantrone	32.6	0.357
DAP + 50 $\mu$ M Phen-DC3	36.7	0.2639
DAP + 50 $\mu$ M Pyridostatin	37.9	0.1225
DAP + 50 $\mu$ M RHPS4	32.9	0.3719
DAP + 50 $\mu$ M TmPyP4	x	0.77
<b>Pearson correlation coefficient</b>		-0.76909
<b>P-value</b>		0.00346

Table 3.3.2.2  $T_m$  is the midpoint of the transition from each melting experiment.  $x$  = not determined. % Hypochromicity is the decrease in ellipticity observed at 288 nm upon the addition of five equivalents of each ligand.

In contrast to the CD melting experiments, the results from UV melting were less straightforward to interpret. Monitoring the ellipticity at 288 nm versus temperature using CD is a direct reporter of the melting of the i-motif structure.<sup>163</sup> On the other hand, changes in absorbance at 295 nm are not unique to i-motif.<sup>94</sup> In the cases of **Mitoxantrone**, **Pyridostatin** and **TmPyP4** with the ATXN2L sequence, an increase of the  $T_m$  is observed possibly due to the stabilisation of a proportion of the DNA that has formed an alternative secondary structure, or an otherwise altered unfolding process reflected in the changes in the slopes of the melting curves. This result is not completely surprising as reports in the literature have shown using some techniques that **Mitoxantrone** can stabilise i-motif forming sequences<sup>152</sup> while others showed destabilization.<sup>100</sup> Otherwise, a similar destabilisation pattern is observed for all seven ligands with DAP, and for the remaining ligands (**Berberine**, **BRACO-19**, **Phen-DC3** and **TmPyP4**) with ATXN2L. Again, the stability of the DAP i-motif was more significantly affected than ATXN2L (Table 3.3.2.3, Appendix A.10).

Ligand	$\Delta T_m$ (°C)	
	ATXN2L	DAP
<b>Berberine</b>	-1.17 ± 0.1	-3.4 ± 0.3
<b>BRACO-19</b>	-1.50 ± 0.1	-11.6 ± 0.3
<b>Mitoxantrone</b>	3.84 ± 0.2	-2.3 ± 0.4
<b>Phen-DC3</b>	-0.61 ± 0.1	-4.4 ± 0.8
<b>Pyridostatin</b>	0.87 ± 0.3	-1.5 ± 0.3
<b>RHPS4</b>	-7.48 ± 0.1	-8.7 ± 0.3
<b>TmPyP4</b>	4.48 ± 0.1	-16.0 ± 0.3

Table 3.3.2.3  $T_m$  is the midpoint of the transition from each UV melting experiment, and SE is calculated using R-square values from the statistics on the data fit.  $\Delta T_m$  is the difference between the  $T_m$  of the DNA in the presence of 5 equivalents of each ligand and the DNA on its own. x = not determined.

Finally, to be able to compare the relative affinities of the ligands to DAP an FID assay was used to determine the concentration at which each ligand could displace 50% of the fluorescent thiazole orange probe (Table 3.3.2.4, Appendix A.11). As mentioned previously, the fluorescence profile of **RHPS4** overlapped with the parameters used in this assay and as such it was excluded from analysis.

Ligand	DC <sub>50</sub> (μM)
<b>Berberine</b>	27.86 ± 0.23
<b>BRACO-19</b>	0.57 ± 0.00
<b>Mitoxantrone</b>	0.62 ± 0.00
<b>Phen-DC3</b>	1.29 ± 0.02
<b>Pyridostatin</b>	14.61 ± 0.17
<b>RHPS4</b>	x
<b>TmPyP4</b>	0.16 ± 0.00

Table 3.3.2.4 Ligand DC<sub>50</sub> values for DAP determined using FID assay. Experiments performed at least in triplicate, and DC<sub>50</sub> values are reported as the 50% displacement value calculated from fitted dose-response curves. Standard errors are calculated using R-square values from the statistics on the data fit. x = not determined.

From these results it can be clearly seen that the seven G-quadruplex binding compounds examined all interacted with two i-motif sequences at neutral pH. This furthered highlighted the issue first proposed by Christ and co-workers during their efforts to develop the i-motif interacting antibody iMab: it is essential to examine

---

both the G-quadruplex and the i-motif when reporting on a molecule that appears to bind specifically to a given DNA secondary structure. These results also further demonstrated that the purported necessity for acidic pH when working with i-motif is false and that i-motifs which fold at neutral pH ought to be considered as potentially physiologically relevant structures with the same confidence as G-quadruplexes.

---

## **Chapter 4: DNA Nanotechnology**

---

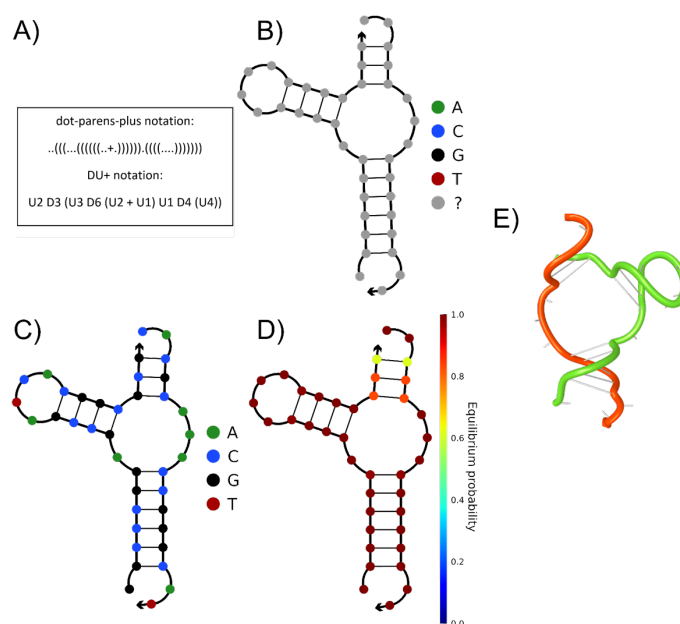
## 4.1 DNA Tetrahedra

As introduced in section 1.6, the research efforts presented in chapters 2 and 3 above were undertaken with the goal of finding and characterising further triggers which could be used to reconfigure the structure of a biological nanoarchitecture whose size could be contextually controlled, relying for the first time on the incorporation of quadruplexes into the edges of a DNA nanostructure.

The inspiration for this undertaking came from work presented by Goodman *et al.* which showed that, in stark contrast to the decade-long journey to arrive at the first DNA cube, they were able to assemble in seconds DNA tetrahedra with near-quantitative yield of a single diastereomer.<sup>164,165</sup> It does not escape my notice that it is perhaps unfair to compare directly the entire timeline, from the conception of the idea to the synthesis of the first DNA cube, with the rapid assembly reported by Goodman. However, the technologies available to, and in certain instances developed by, Goodman do make this comparison not entirely unreasonable. The bottom-up approach to the design of DNA nanotechnologies has been propelled forward by the development of software that enabled, in a sense, the incorporation of aspects from the top-down approach.<sup>166</sup>

With a target structure in mind, researchers can now use software such as NUPACK to determine the sequence or sequences of DNA required to build it.<sup>166</sup> The predecessor to NUPACK, NANEV<sup>167</sup> is what enabled Goodman to synthesise a complex 3D DNA nanoarchitecture in such a uniform and rapid manner.<sup>164</sup> The power of NUPACK is derived from its assimilation of decades of knowledge about DNA to present two tools that aid in the synthesis of DNA based nanoarchitectures. The first tool, *Analysis*, is underpinned by algorithms<sup>168–170</sup> which, when given input variables relating to the base composition of the sequences, the number of nucleic acid strands, their concentrations, and their environment, produce results about the identity and concentrations of the structures that could possibly form based on minimum free energy probabilities of the base-pairings that could occur between those strands.<sup>166</sup> The second tool, *Design*, essentially does the opposite of this: following specific but straightforward rules,<sup>166</sup> a target secondary structure at a given temperature is provided as the input along with the number of strands that it

should be composed of. Algorithms<sup>171–174</sup> then compute sequences that can form the desired target secondary structure, and present them along with the average percentage of nucleotides which will be incorrectly paired at the set equilibrium temperature relative to the target structure.<sup>166</sup> Used in concert, these two tools provide an extremely powerful system to design nucleic acid nanostructures. Furthermore, the NUPACK software is very versatile and allows results to be passed back and forth between the *Analysis* and *Design* tools to increase the likelihood that a particular target structure will be formed under the specified conditions.



**Figure 4.1.1** *Example NUPACK workflow: A) Dot-parens-plus notation: each unpaired base is represented by a dot, each base pair by matching parentheses, and each nick between strands by a plus; and DU+ notation: a duplex of length  $x$  base pairs is represented by  $Dx$  and an unpaired region of length  $x$  nucleotides is represented by  $Ux$ . Each duplex is followed immediately by the substructure that is enclosed by the duplex. If this substructure includes more than one element, parentheses are used to denote scope. A nick between strands is specified by a plus.<sup>166</sup> B) Visual preview of target design entered. C) Computed base identities for target design. D) Equilibrium probability of the base identities forming the target design. E) Depiction of target structure adopted shown using the ideal helical geometry of B-form helices for DNA.*

Using NANEV, Goodman was able to work on developing a DNA tetrahedron which assembled rapidly and uniformly from four oligonucleotides. They additionally reported on its purification, and ligation of the nicks between the four strands to form one contiguous structure resistant to enzymatic digestion (Figure 4.1.2).<sup>164</sup> This resulted in many more examples of DNA polyhedra being designed and synthesised;<sup>175–179</sup> with examples being built by a single strand of DNA<sup>180</sup> and others where the polyhedra are used to assemble higher-order multimeric complexes.<sup>178,181</sup> Research in this area has also progressed from simply the design and synthesis of DNA tetrahedra, to some fascinating applications: they have been used as carriers to encapsulate and release fluorescent biopolymers<sup>182</sup> and proteins,<sup>183</sup> in the tracking and modulation of cell entry pathways,<sup>184</sup> in targeting tumours and allowing *in vivo* imaging,<sup>185</sup> and as theranostics.<sup>186,187</sup>

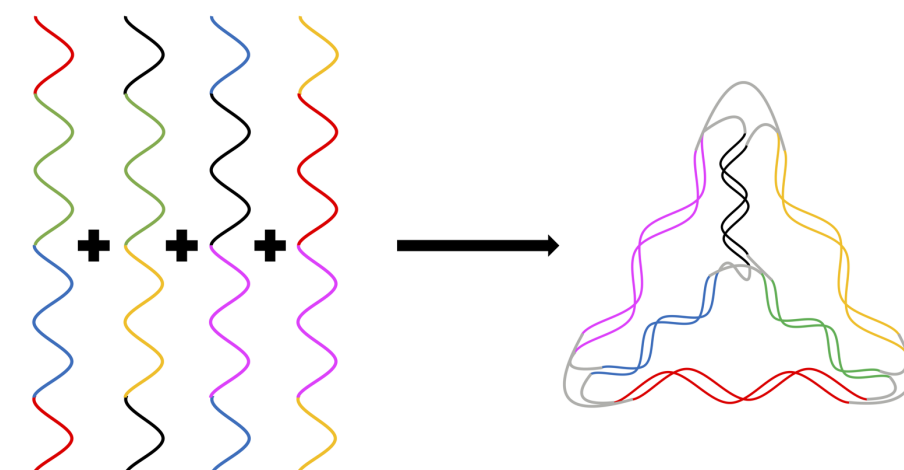


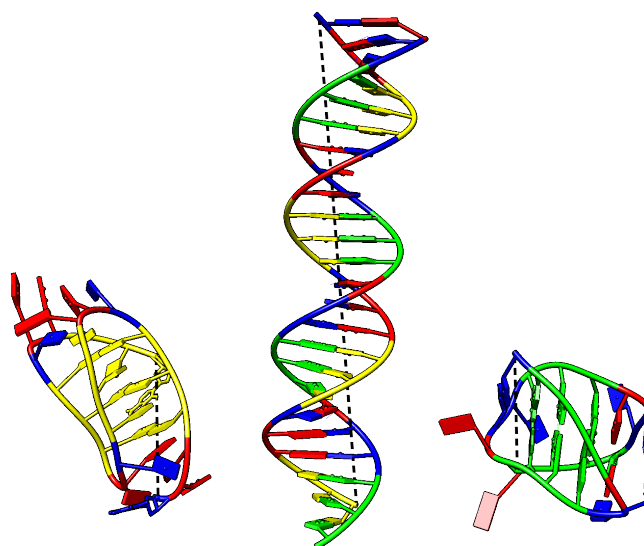
Figure 4.1.2 Illustration of design of DNA tetrahedron from four oligonucleotides consisting of complementary segments (colour coordinated) which hybridise to form the target tetrahedron structure.

#### 4.1.1 Reconfigurable DNA Tetrahedra

A properly characterised i-motif forming sequence can respond predictably and rapidly to changes in pH, which is easy to modulate, compared to the G-quadruplex which is generally insensitive to pH<sup>12,44,188</sup> and requires triggers which are more difficult to remove post-addition or to recycle.<sup>90</sup> The benefit afforded by the ability to efficiently cycle an i-motif forming DNA sequence between the unfolded and the quadruplex states without introducing by-products is what has led to the i-motif



being used so extensively in nanotechnological applications;<sup>90,189</sup> and the DNA tetrahedron is no exception to this trend. The fact that the i-motif formed by a DNA sequence will be more condensed than either the single-stranded random coil or the double helix formed with the complementary strand has been combined with a DNA tetrahedron to develop an architecture that can “shrink and stretch”. Wang *et al.* designed a tetrahedron with an i-motif forming sequence protruding from one of its vertices and showed that this structure’s size changed in response to changes in pH.<sup>190</sup> While this structure combined a DNA quadruplex with a tetrahedron, the size of the tetrahedron itself was not affected by the quadruplex structure and could not be altered by it.



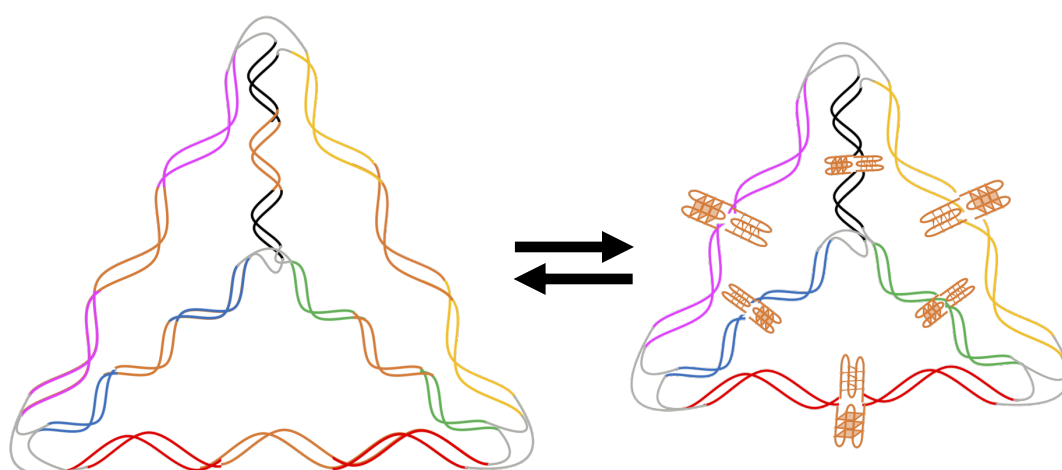
*Figure 4.2.1.1 Models depicting the i-motif (left) and G-quadruplex (right) of the same DNA sequence from which the double helix (centre) is formed. Black dashed lines show distance between oxygens on the 5' and 3' ends of the molecules. NDB colouring: adenine is red, thymine is blue, guanine is green, cytosine is yellow.*

A more notable example, and the first of a DNA tetrahedron whose size could be contextually controlled, was presented by Goodman *et al.* in 2008.<sup>177</sup> They designed a system where a segment in the middle of one (or two) of the sides of the tetrahedron formed a hairpin which stuck out of its respective edge rather than being part of the double-helix which ran along the six edges. The length of the edge containing the hairpin, when the hairpin is formed, is less than if that entire segment

was part of a double-helix, thereby reducing the length of that edge of the tetrahedron (by a factor of three for a 10-nucleotide long segment).<sup>177</sup> Using a 'fuel' strand that is complementary to the hairpin sequence, that segment could be induced to form a double-helix, thereby increasing the size of that side of the tetrahedron. This could then be reversed by the introduction of an 'antifuel' strand which is complementary to the 'fuel' strand and has a higher affinity to it than the hairpin forming segment, thus resulting in the displacement of the 'fuel' strand and consequently the hairpin in the tetrahedron would reform.<sup>177</sup>

## 4.2 The Quadruplex Tetrahedron ( $Q_{TET}$ )

In the hope of providing a further avenue for the development of nanoscale DNA tetrahedra we set out to design a system which incorporated quadruplexes into the sides of the structure: the  $Q_{TET}$ . As illustrated in Figure 4.2.1, by incorporating G-quadruplex and i-motif forming sequences into the stretches of DNA making up the edges of the tetrahedron, when the quadruplex structures are formed there will be a reduction in the length of those edges, and consequently the size of the tetrahedron. Thereby allowing the  $Q_{TET}$  to change its size in response to its environment.



*Figure 4.2.1 Proposed multi-state  $Q_{TET}$  model; orange segments represent added quadruplex forming sequences which can form their respective structure and consequently alter the size of the structure in response to triggers.*

---

Incorporating the quadruplex functionality into the edges of the nanocage would provide three primary benefits over the reconfigurable tetrahedron presented by Goodman:

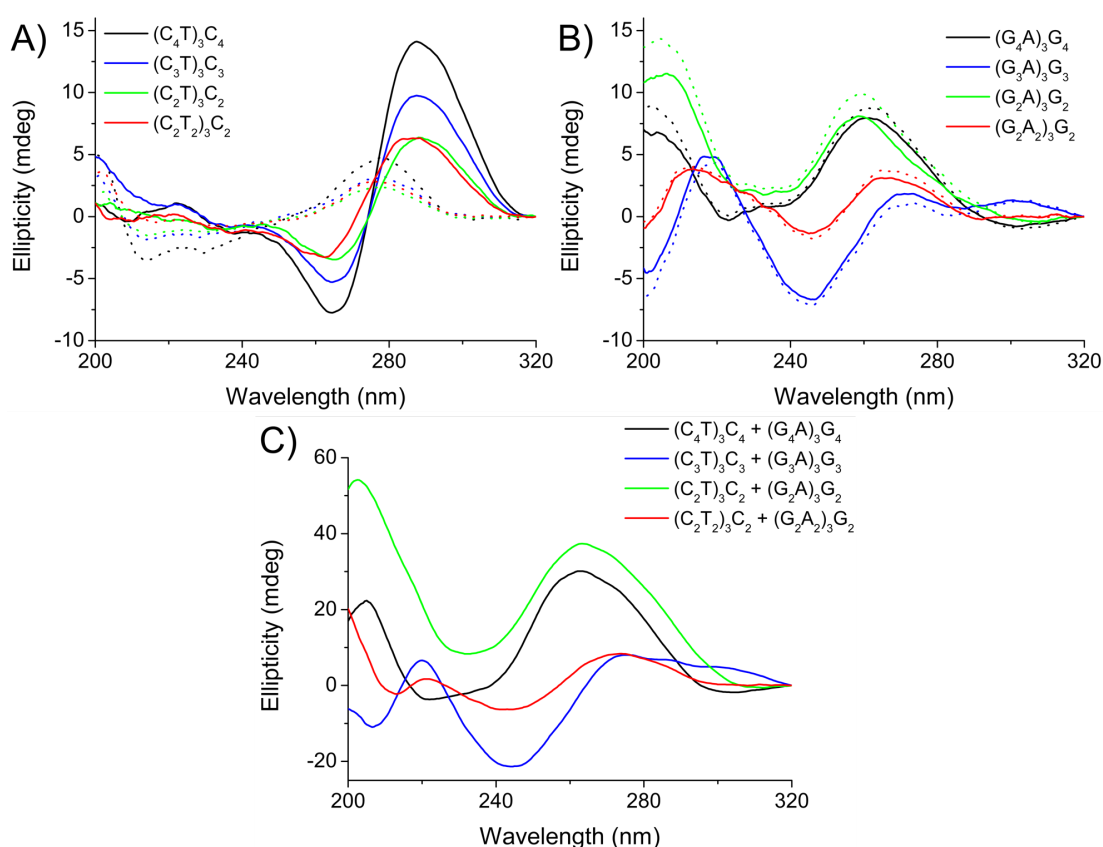
- (i) the time taken for adequately designed quadruplexes to form will be measured in fractions of a second<sup>162,191</sup> as opposed to thousands of seconds needed to reach the equilibrium between the competing complementary strands and the hairpin<sup>177</sup>
- (ii) depending on the triggers used, this system could be much more amenable to continuous cycling without the generation of significant by-products
- (iii) with DNA tetrahedra already showing potential in the development of viable theranostics,<sup>186</sup> a system responsive to triggers inherently available in different physiological environments<sup>86</sup> is preferable to one that requires repeated addition of further DNA strands.

#### 4.2.1 Q<sub>TET</sub> Quadruplexes

The attempt to incorporate quadruplex forming sequences, into a nanoscale DNA tetrahedron presented a multifaceted challenge. The ability to reliably predict how DNA sequences would assemble based on the complementarity of their bases was the cornerstone of the ability to design, synthesise and develop applications for nanoscale DNA polyhedra. Ironically, this presented the first hurdle in the design of the Q<sub>TET</sub>. The sequences chosen for the quadruplex forming segments needed to be long enough so that they could form their respective structures while simultaneously not being so long so as to overcome the intricate complementarity necessary for the assembly of the rest of the structure. The Q<sub>TET</sub> was to be built into the scaffold of the tetrahedron reported by Goodman *et al.* which was made up of four complementary strands.<sup>164</sup> In Goodman's structure each strand was designed so as to be comprised of three segments which had complementary segments running through one of the other three strands. The edges of the tetrahedron were each 20 base pairs long, and the vertex of each edge was formed by an adenine-adenine mismatch.<sup>164</sup>

The first step in designing the Q<sub>TET</sub> was to determine the base composition of the sequences to be used for the quadruplex forming segments of the structure. It was decided that these segments needed to be made up of a pair of complementary sequences to allow formation of the double helix for the 'open' conformation of the Q<sub>TET</sub>. Additionally, one of each complementary pair needed to be able to form an i-motif and the other a G-quadruplex to enable, in a controllable manner, reconfiguration to the 'closed' conformation. Four pairs of sequences were chosen and initially investigated using CD to determine their ability to form the requisite structures:

- (i) (C<sub>4</sub>T)<sub>3</sub>C<sub>4</sub> and (G<sub>4</sub>A)<sub>3</sub>G<sub>4</sub>
- (ii) (C<sub>3</sub>T)<sub>3</sub>C<sub>3</sub> and (G<sub>3</sub>A)<sub>3</sub>G<sub>3</sub>
- (iii) (C<sub>2</sub>T)<sub>3</sub>C<sub>2</sub> and (G<sub>2</sub>A)<sub>3</sub>G<sub>2</sub>
- (iv) (C<sub>2</sub>T<sub>2</sub>)<sub>3</sub>C<sub>2</sub> and (G<sub>2</sub>A<sub>2</sub>)<sub>3</sub>G<sub>2</sub>



**Figure 4.2.1.1** CD of candidate quadruplex forming sequences for Q<sub>TET</sub>. A) i-Motif forming sequences at pH 8.03 (dotted) and pH 4.94 (solid). B) G-quadruplex forming sequences without KCl (dotted) and with 100 mM KCl (solid). C) Sequence pairs annealed together. All at 10 μM in 10 mM Tris, 5 mM MgCl<sub>2</sub> buffer.

From the CD data it could be seen that all four i-motif forming sequences appear to have adopted the desired quadruplex structure in an acidic pH environment. They all exhibited the quintessential signal maxima at  $\sim 288$  nm and the negative character at  $\sim 265$  nm.<sup>143</sup> For the G-quadruplexes,  $(G_4A)_3G_4$ ,  $(G_2A)_3G_2$  and  $(G_2A_2)_3G_2$  appeared to have adopted the parallel-stranded G-quadruplex topology indicated by the signal maxima at  $\sim 260$  nm and, more so for  $(G_2A_2)_3G_2$  than the other two, the negative ellipticity at  $\sim 240$  nm. On the other hand,  $(G_3A)_3G_3$  appeared to have possibly adopted a mixed-type parallel/anti-parallel quadruplex topology; exhibiting some parallel character as described for the other three sequences, albeit to a lesser extent, as well as having a second positive signal at  $\sim 290$  nm.<sup>143,192</sup> Assessing the formation of the double helix with CD was less straightforward as, while B-form DNA will generally have positive bands between 260 and 280 nm and a negative band at  $\sim 245$  nm, sequence composition is known to markedly affect the CD signal. Also, peak intensities for B-form DNA are relatively small as a result of the base pairs being perpendicular to the axis of the double helix.<sup>143</sup> Nevertheless, from the CD data it appeared that the  $(C_2T_2)_3C_2$  and  $(G_2A_2)_3G_2$  pairing seemed to fit this criteria most closely for double helix formation.

To confirm whether the candidate quadruplex forming pairs could form the double helix native PAGE was used to observe the behaviour of the pairs when annealed in the buffer conditions that would be used for  $Q_{TET}$  formation (Figure 4.2.1.2).

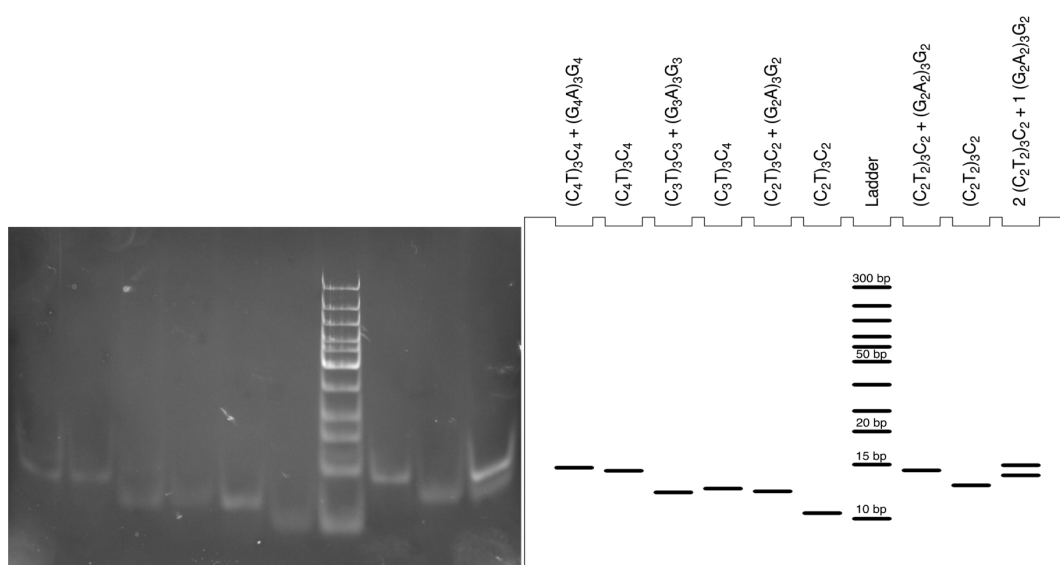


Figure 4.2.1.2 Analysis of double helix formation by  $Q_{TET}$  quadruplex forming sequences using native PAGE.

---

Using this technique, it could be seen that when the pairs were annealed together in a 1:1 stoichiometric ratio they migrated through the gel as a single species. The bands travelled as would be expected for a double helix of their respective sizes and generally showed a slightly slower migration to the single C-rich strand run alone in the neighbouring lane as a control. For further confirmation two equivalents of  $(C_2T_2)_3C_2$  were annealed with one equivalent of its complement in lane 10. The existence of two species could clearly be seen here, the upper band corresponding to the double helix in lane 8 and the lower one to the excess unpaired C-rich strand as in lane 9.

From this data it was decided to proceed with the attempted synthesis of the  $Q_{TET}$  using the  $(C_4T)_3C_4 - (G_4A)_3G_4$  pair and the  $(C_2T_2)_3C_2 - (G_2A_2)_3G_2$  pair. Both of these pairs appeared to be adopting the desired structures under the different conditions. Specifically, the first pair was chosen as it provided the longest quadruplex forming segment: the longer this segment the greater the size differential that could be achieved by adopting the quadruplex structure. However, when assessed with NUPACK, the incorporation of this quadruplex forming pair into the strands making up the tetrahedron resulted in the highest probability of this segment forcing the strands out of adopting the overall tetrahedron structure required and instead resulting in base-pairings between this quadruplex pair outside of the framework of the tetrahedron. The other pair was chosen for essentially being on the opposite end of this spectrum. This shorter segment had the lowest probability of disrupting the overall desired architecture while still being able to form a double helix and the quadruplexes.

#### **4.2.2 $Q_{TET}$ Synthesis**

With the quadruplex forming segments chosen it was decided to synthesise the  $Q_{TET}$  using the  $(C_2T_2)_3C_2 - (G_2A_2)_3G_2$  pair to have the best chance of the overall structure forming. Two versions of the  $Q_{TET}$  were synthesised following Goodman's rapid assembly procedure: the first where only one edge contained the quadruplex segments and the second where the quadruplex forming segments were in all six edges of the tetrahedron.

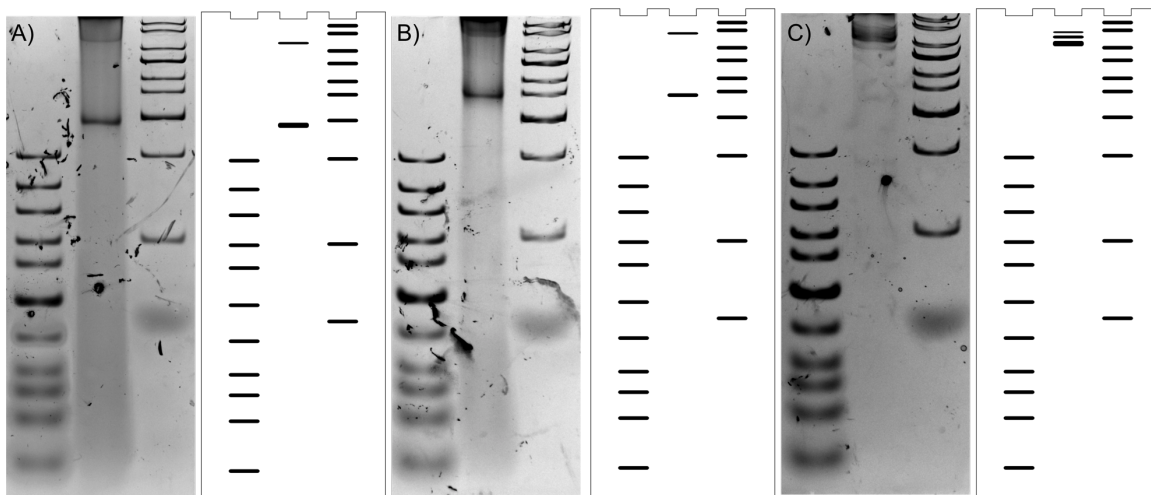


Figure 4.2.2.1 Analysis of rapid tetrahedral formation using native PAGE. A) Tetrahedron with no quadruplex forming segments. B)  $Q_{TET}$  with quadruplex segment in single edge. C)  $Q_{TET}$  with all six edges containing quadruplex segments.

After some experimentation with the annealing conditions it was found that, while the  $Q_{TET}$  could form following the same rapid assembly reported by Goodman (Figure 4.2.2.1), cooling the structure in a more controlled manner at a rate of  $-1^{\circ}\text{C}/\text{min}$  resulted in a more complete adoption of the tetrahedron's structure by the four constituent strands. The value of rapidly assembling this structure in under five minutes, as opposed to a couple of hours, was drastically outweighed by the uniformity of the product and the alleviation of the need for a gel purification step.

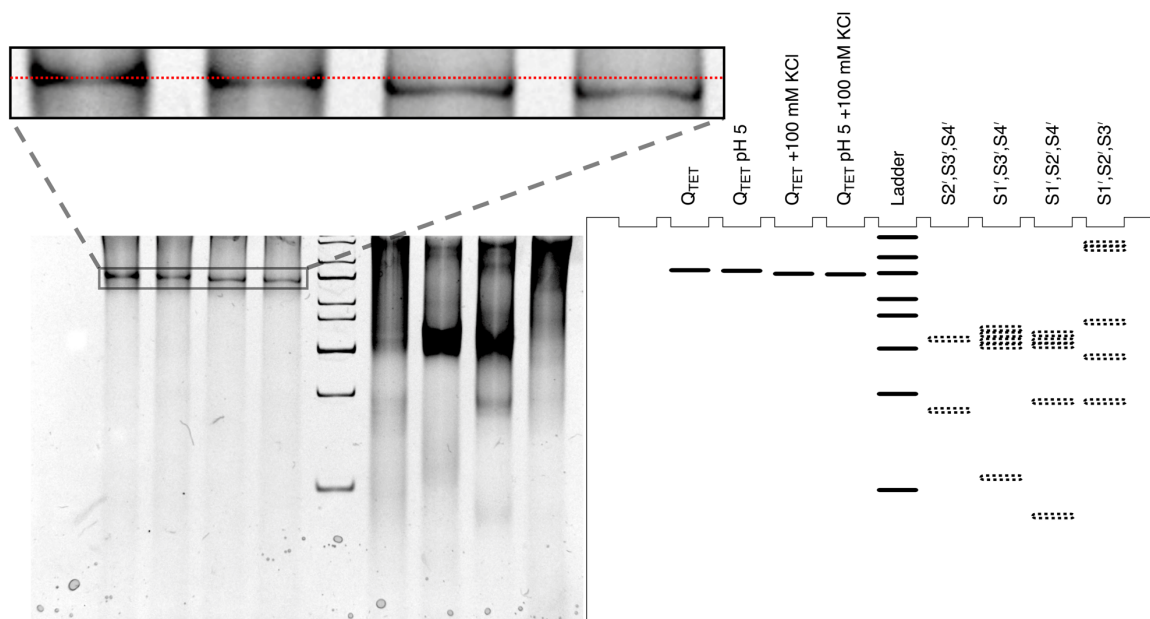


Figure 4.2.2.2 Analysis of  $Q_{TET}$  formation with  $(C_2T_2)_3C_2 - (G_2A_2)_3G_2$  quadruplex segments under different conditions using native PAGE.

The results of this optimised synthesis of the  $Q_{TET}$  can be seen in Figure 4.2.2.2. A single dominant species was apparent in the four lanes (2-5) in which the product of annealing all four constituent strands was run. The effect of omitting one of each of the constituent strands can be seen in the final four lanes (7-10); where the resultant mixtures of species can be clearly contrasted with the uniformity of combining all four. In an attempt to see whether the reconfiguration of the structure was possible, and whether the resultant size change could be seen using this technique, the buffer conditions were adjusted in an attempt to induce the quadruplex structures' formation. The pH was lowered to 5.0 to induce i-motif formation (lane 3), 100 mM KCl was added to the buffer to induce G-quadruplex formation (lane 4), and both of these changes were implemented together to induce the formation of both quadruplexes (lane 5). In the inset in Figure 4.2.2.2 it is apparent that with the addition of KCl to the annealing buffer the  $Q_{TET}$  had migrated slightly further through the gel (lane 4), and to a lesser extent possibly even for the  $Q_{TET}$  at acidic pH (lane 3), and ever so slightly further still for the combination of these factors (lane 5). While this provided evidence that the  $Q_{TET}$  could be formed by these strands it only offered indirect, and not particularly conclusive, evidence for the formation of the quadruplexes and their consequent alteration of the size of the  $Q_{TET}$ , and consequently the second hurdle in designing the  $Q_{TET}$  was uncovered.



### 4.2.3 Seeing the Switch

The second hurdle encountered in designing the  $Q_{TET}$  was being able to identify whether the quadruplexes were forming in response to triggers. To try and assess whether the difference observed in Figure 4.2.2.2 was meaningful, a similar experiment was carried out with two of the component strands: S1' and S2'. S1' was to represent the i-motif's formation as it contained two stretches of the i-motif forming segment, and S2' would be used as an indicator of the behaviour of the G-quadruplex using the same reasoning.

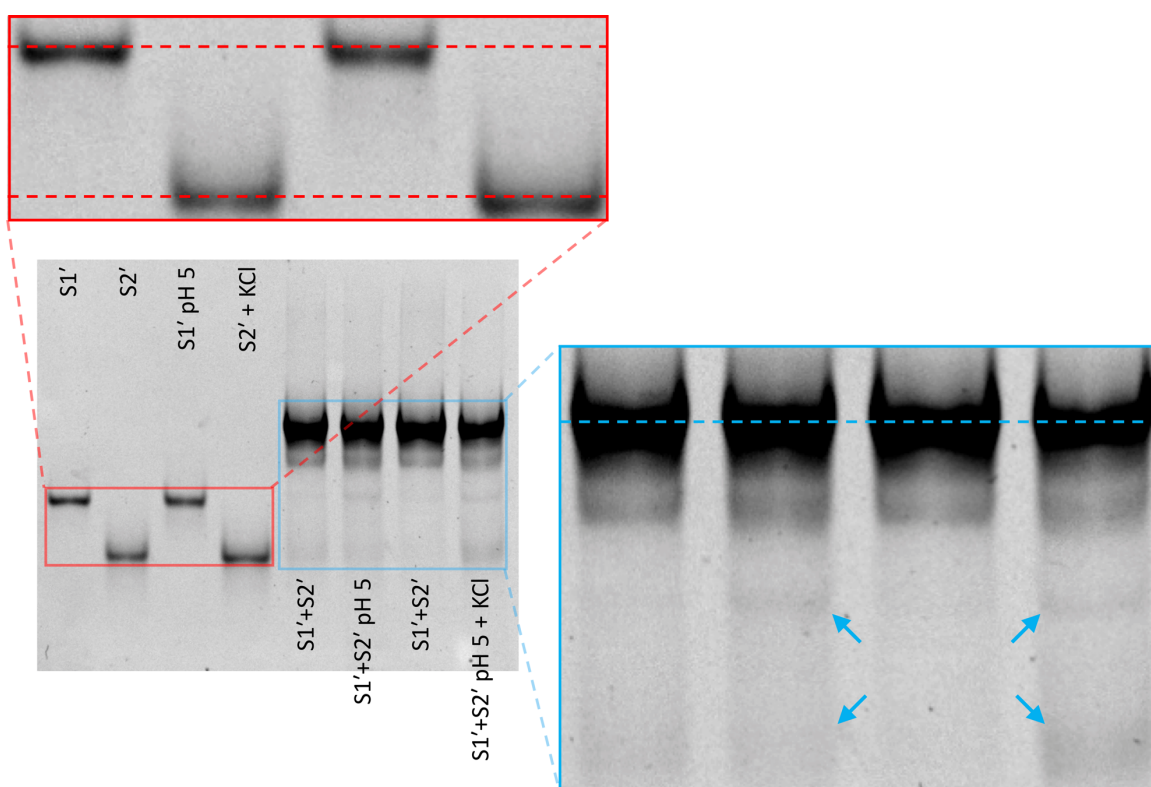


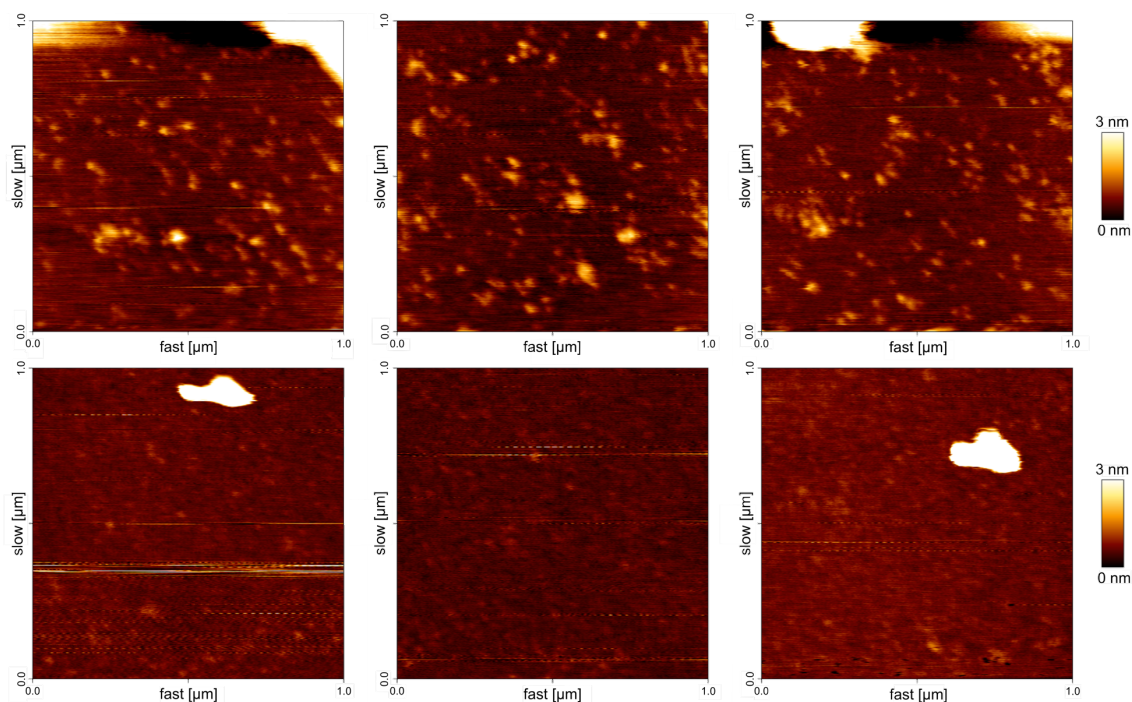
Figure 4.2.3.1 Analysis of quadruplex formation using S1' and S2' sequences using native PAGE.

The red insert from Figure 4.2.3.1 shows that compared to the  $Q_{TET}$  in Figure 4.2.2.2 there is an even smaller difference, if any at all, in the distance travelled by the single strands under their respective quadruplex forming conditions. A similar lack of response in the primary bands for the duplexes can be seen in the blue insert. Interestingly however, the addition of KCl appeared to have increased the proportion of the bands corresponding to the single stranded species (indicated by

---

the arrows). This possibly could have been caused by the formation of the G quadruplex in the complementary segment of these two strands resulting in the duplex around it dissociating. However, each of the single strands used here were 105 bases long; 42 of those bases were complementary and corresponded to three quadruplex forming segments and of the remaining bases only 20 were designed to be complementary and not to form a quadruplex. This meant that the majority of the bases forming the two stranded species seen here were from the introduced quadruplex forming segments. Consequently, if the formation of the quadruplexes were significantly disrupting the pairing of the two non-quadruplex containing complementary segments, it would be expected to see much less of the two-stranded species in the three lanes furthest to the right. While native PAGE has been used previously to identify both the G-quadruplex and the i-motif, the electrophoretic mobility of the structures has been found to be less than straightforward to predict. With evidence that in some cases intramolecular G-quadruplexes, as expected to form in this system, are indistinguishable from the unfolded forms of the same oligonucleotides.<sup>193</sup>

Following on from the unsuccessful attempts to find convincing evidence using native PAGE it was decided to attempt to use atomic force microscopy (AFM) to visualise the structures through an in-house collaboration with Dr Andrew Round. During the experiments performed to optimise the annealing conditions for the tetrahedra, it was found that combining the four constituent strands in the 10 mM Tris, 5 mM MgCl<sub>2</sub> buffer would result in the formation of the tetrahedron structure even in the absence of a thermal annealing step. However, using a buffer system consisting of only 10 mM Tris, even when employing thermal annealing, the structure would not form. This was exploited for the first trial of AFM imaging where the results of annealing the four constituent strands in the presence and absence of magnesium were imaged (Figure 4.2.3.2).



*Figure 4.2.3.2 AFM images showing the result of annealing the four constituent strands of the tetrahedron in the presence (top) and absence (bottom) of  $MgCl_2$ .*

Unfortunately, the images collected using AFM did not provide the means to overcome the second hurdle. While a clear difference was visible between the images in the upper and lower panels in Figure 4.2.3.2, and Dr Round was confident in attributing the species observed to tetrahedra and single stranded DNA respectively, attempts to perform volume analysis to enable quantification of the features of these species failed. The AFM performed here was relatively low resolution with the panels in Figure 4.2.3.2 showing a  $1 \mu m^2$  area, whereas the high-resolution AFM images published by Goodman and co-workers were focussed in on a  $30 nm^2$  area.<sup>164</sup> While the technique clearly has the potential to provide much more accurate information this would not have been possible without an investment into upgraded hardware for the microscope, including purchasing new cantilevers and supersharp tips. Consequently, it was decided to first explore other methods to attempt to observe the reconfiguration of the  $Q_{TET}$ .

The next technique employed in the attempt to determine whether the quadruplex structures were forming, and whether they had the desired effect of changing the size of the  $Q_{TET}$  was dynamic light scattering (DLS).

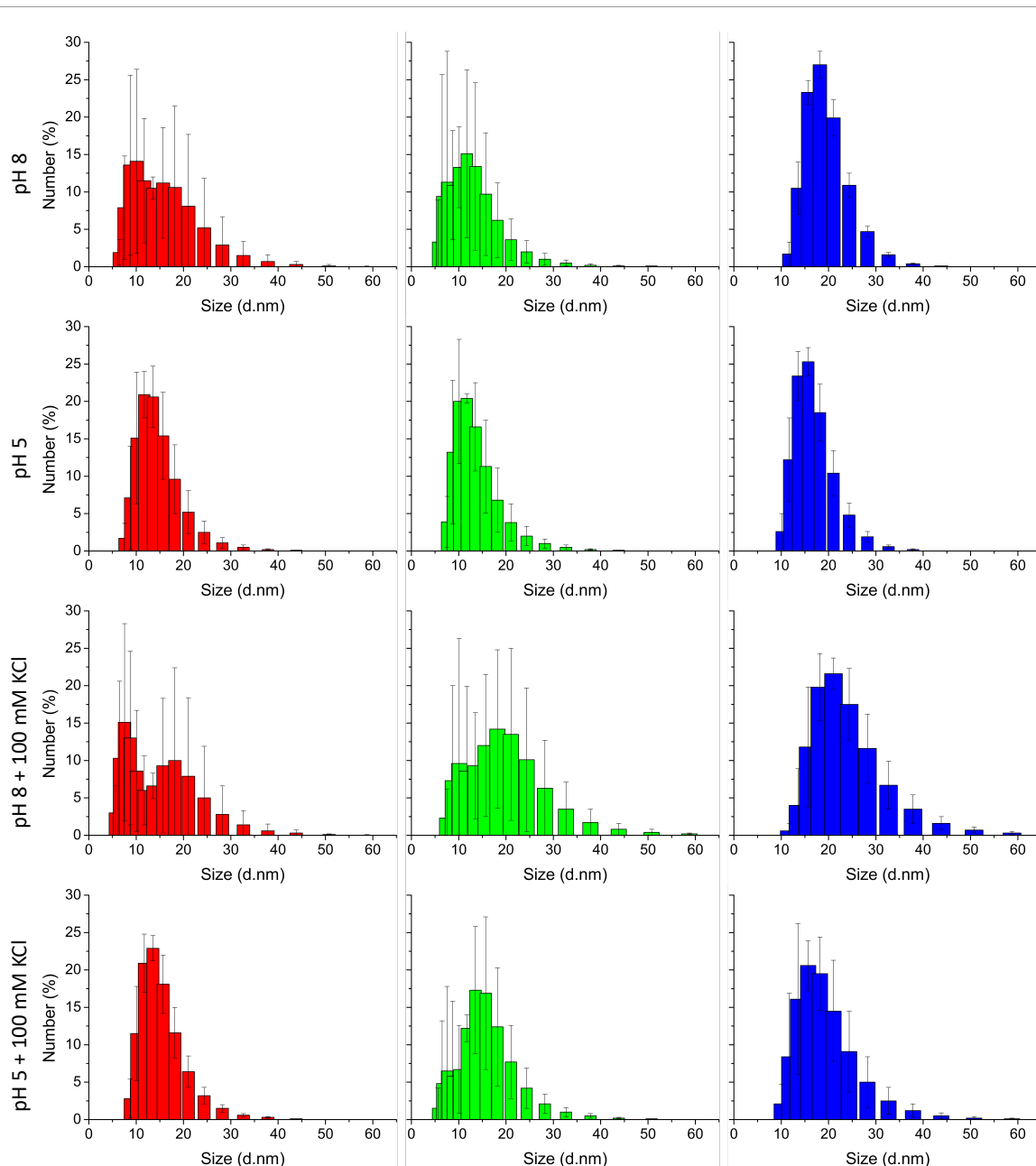


Figure 4.2.3.3 DLS data of from Goodman's tetrahedron (red) and  $Q_{TET}$  with quadruplex forming segments in one edge (green) and in all edges (blue).

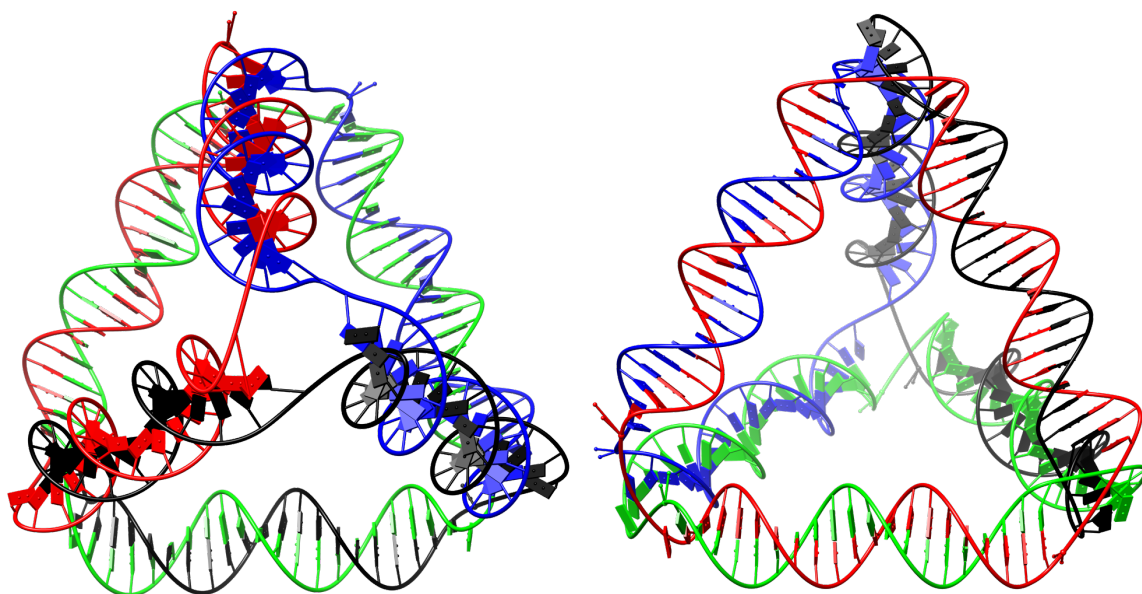
Annealing conditions	Size (d.nm)		
	Tetrahedron	$Q_{TET}$ (one edge)	$Q_{TET}$ (all edges)
pH 8	$12.01 \pm 4.68$	$14.48 \pm 4.87$	$18.74 \pm 4.28$
pH 5	$12.53 \pm 3.90$	$15.07 \pm 4.86$	$17.48 \pm 4.21$
pH 8 100 mM KCl	$20.50 \pm 6.13$	$19.23 \pm 6.31$	$25.56 \pm 7.84$
pH 5 100 mM KCl	$15.90 \pm 4.51$	$16.80 \pm 5.39$	$18.11 \pm 5.78$

Table 4.2.3.1 Hydrodynamic diameter of Goodman's tetrahedron,  $Q_{TET}$  with quadruplex forming segment in one edge and in all edges under different annealing conditions measured by DLS.

---

The DLS data showed that the tetrahedra were forming as designed and were approximately of the size expected. Each edge of Goodman's tetrahedron is made up of a 20-basepair double helix which should stretch approximately 7 nm;<sup>164</sup> the 34-base double helix formed by the edges with the introduced quadruplex segment should stretch approximately 12 nm. It is expected that the hydrodynamic diameter measured by DLS will not match exactly the size of the nanoarchitecture as it is the radius of a hypothetical sphere which would diffuse in the same manner as the particle being measured.<sup>194</sup> This difference is a result of the particle being non-spherical, solvated and dynamic – hence 'hydrodynamic' diameter. Unfortunately, the problem of being unable to determine whether the quadruplexes were forming and if they were changing the tetrahedron's size remained unresolved. Goodman's tetrahedron lacked the quadruplex functionality and was used as a control to monitor the effect of the different buffers on the hydrodynamic diameter of the structure. The changes in the apparent size of the control tetrahedron under the different annealing conditions essentially mirrored those observed with the functionalised structures, thereby rendering the size changes observed with the Q<sub>TET</sub> meaningless.

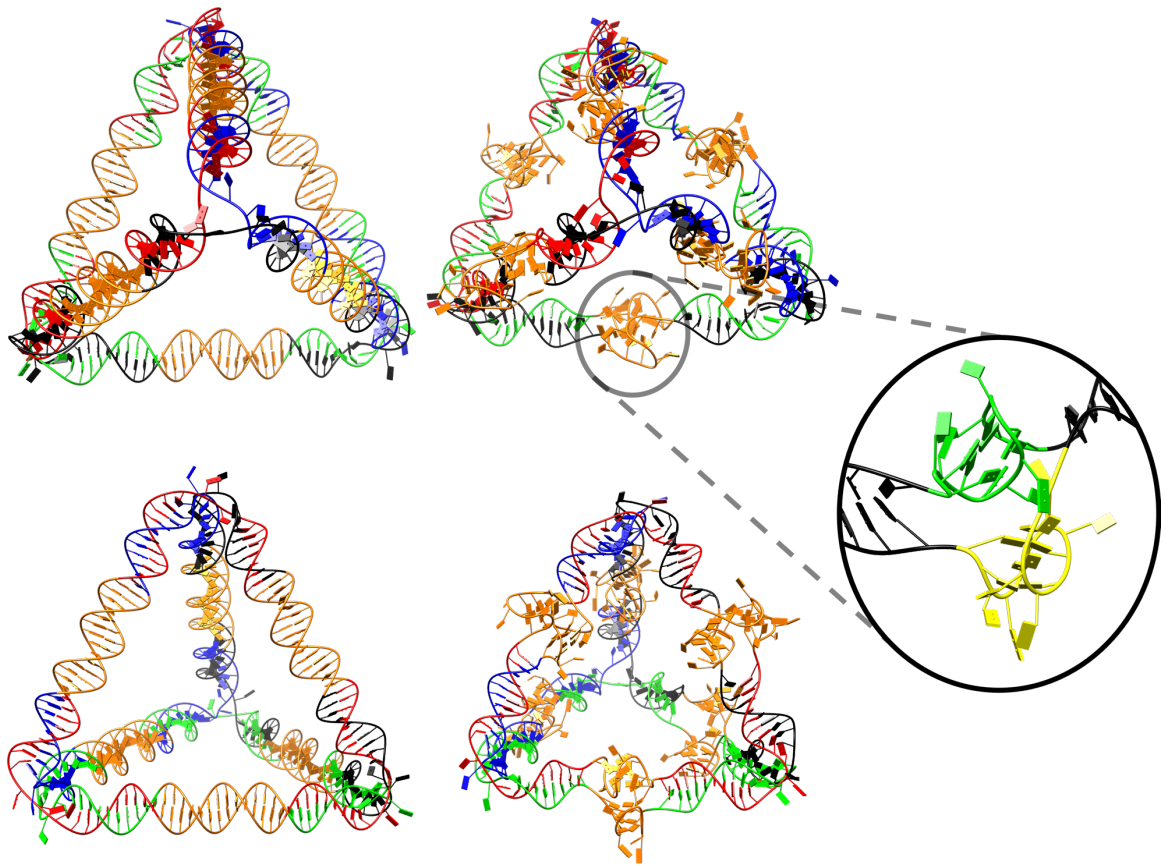
In an attempt to better understand the potential size difference which could thus far not be seen experimentally, it was decided to build some models of the tetrahedra. Help with this computational work was sought once again from Dr László Fábián, and Dr Marco Cominetti, both from within the School of Pharmacy at the University of East Anglia. The first models were built true to the sequences reported by Goodman,<sup>164</sup> to mimic the control tetrahedron used in the experiments thus far, and without any quadruplex functionality.



*Figure 4.2.3.4 Models depicting a DNA tetrahedron. Top-down view from single vertex (left). Side-on view with three vertices from a single face in the same plane (right). Each of the four constituent oligonucleotides is indicated by a single colour.*

Following on from this, models were built to represent the  $Q_{TET}$  (Figure 4.2.3.5) and the limitation of the available technology unfortunately once again rendered drawing meaningful conclusions from these efforts difficult without the further dedication of an inordinate amount of time. One of the first problems encountered when attempting to perform MD simulations on these models was the difficulty in building them in such a way that the concatenation of the four constituent strands was compatible with available force field parameters. That issue was resolved by building the models as four continuous single stranded triangles then introducing the required base-pairing between the complementary segments, as opposed to six end-to-end connected duplexes. This enabled the energy minimisation of the 'open' state of the  $Q_{TET}$  using the AMBER parmbsc0 force field,<sup>195</sup> and the elimination of all clashes from the model.

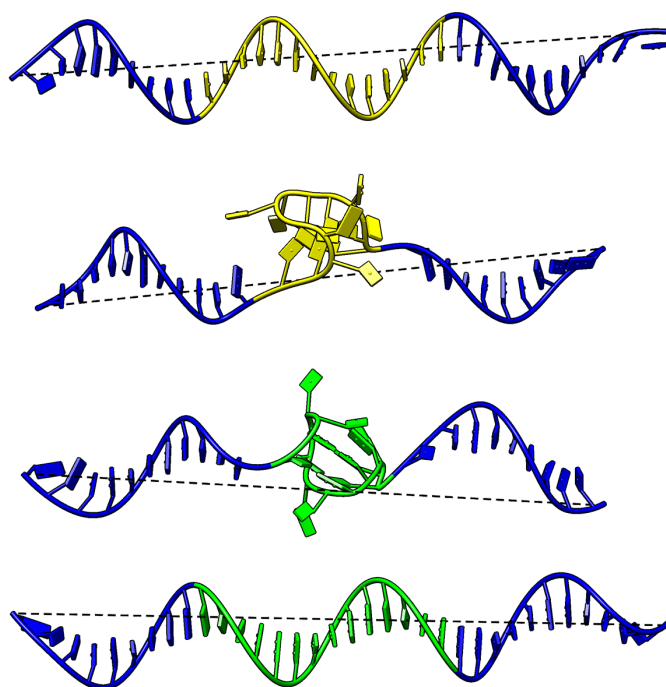




*Figure 4.2.3.5 QTET models: top-down view from single vertex (top). Side-on view with three vertices from a single face in the same plane (bottom). 'Open' state with edges formed from double helical DNA (left). 'Closed' state with quadruplex structures formed in the edges (right). Each of the four constituent oligonucleotides is indicated by a single colour; with the quadruplex forming segments in orange. Inset shows magnified central region of edge with the G-quadruplex (green) and the i-motif (yellow) formed.*

As for the 'closed' state of the  $Q_{TET}$ , it was decided that while recent advances had begun to make it possible to perform some simulations with DNA quadruplexes,<sup>196–199</sup> the additional complexity of combining this with the simulation of the rest of the tetrahedron was unwarranted. Instead, representative quadruplexes of the same length (14 nucleotides) were used as proxies for the  $(C_2T_2)_3C_2 - (G_2A_2)_3G_2$  pair, and with the quadruplex atoms fixed the energy minima of the duplex segments of the  $Q_{TET}$  were recalculated. While there were inherent limitations to this method, namely it not having been designed to resolve extended distortions or structural

problems, it did potentially shed some light on the reason for the inability to discern a size difference between the two states of the cage. From these models, the change in length of the edges of the  $Q_{TET}$  was less than 2 nm: the i-motif segment length changed from 11.75 to 10.16 nm, and the G-quadruplex segment length from 11.51 to 10.27 nm. With the quadruplexes formed the anticipated reduction in the length of each edge of the  $Q_{TET}$  was not realised. What was observed was a warping of the double helix flanking the quadruplex segments and only a minor reduction in the total length of the edges of the  $Q_{TET}$ .



*Figure 4.2.3.6 Single edge of  $Q_{TET}$  with shown with the i-motif (yellow) and the G-quadruplex (green) structures formed (centre) and as part of the double helix (top and bottom). Black dashed lines show distance between oxygens on the 5' and 3' ends of the molecules.*

This observation agrees with reports on the destabilising effects of quadruplexes on their neighbouring duplex,<sup>200</sup> and brought to the fore the inverse problem: what effects were the duplex segments having on the ability of the quadruplex segments to form their respective structures? Different reports have found that, under the right conditions, some G-quadruplexes and i-motifs are able to form in competition with their duplex whereas other have postulated that the duplex must first be denatured for quadruplex formation to be possible.<sup>201–204</sup> Generally, these reports



note that these conflicting observations are likely due to the specific sequences involved in the adoption of the different structures; and note that this will also have a role to play in the behaviour of any given system. As a result of this it was decided to measure the CD spectra of two of the component strands of the  $Q_{TET}$ , and of the structure itself, under the quadruplex triggering conditions to see what effect the flanking bases had on the formation of the quadruplex structures.

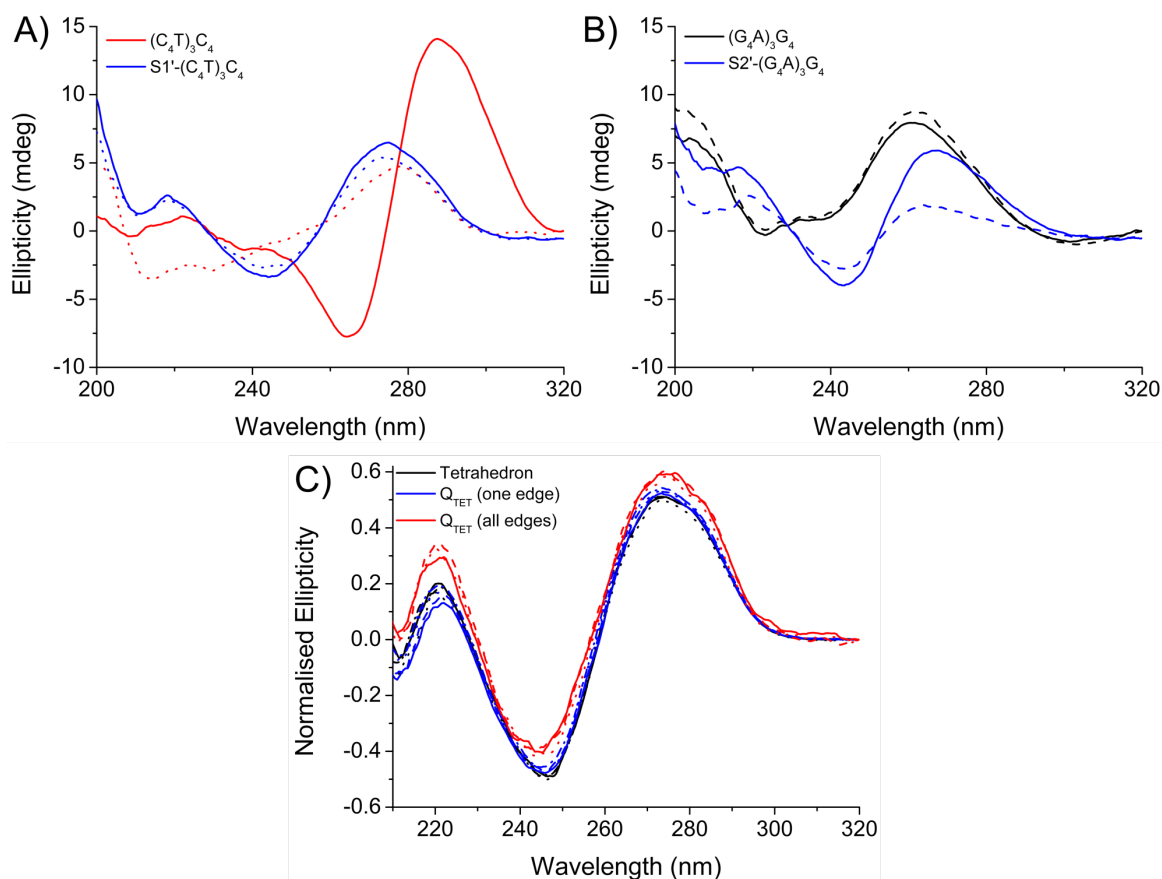


Figure 4.2.3.7 CD spectra of A) *i*-motif forming segment  $(C_4T)_3C_4$  and  $S1'-(C_4T)_3C_4$  at pH 8 (dotted) and pH 5 (solid). B) G-quadruplex forming segment  $(G_4A)_3G_4$  and  $S2'-(G_4A)_3G_4$  without KCl (dotted) and with 100 mM KCl (solid). C) Tetrahedra and  $Q_{TET}$  under different conditions: pH 8 (solid), pH 8 + 100 mM KCl (dotted), pH 5 (dashed), pH 5 + 100 mM KCl (dot-dash).

From these results it could be seen that the quadruplex forming segment of  $S1'$  does not form the *i*-motif. The spectra of  $S2'$  clearly shows a sensitivity to potassium and with a spectrum showing signals near the characteristic G-quadruplex regions of a peak at  $\sim 260$  nm and a negative signal at  $\sim 245$  nm it

---

was possible that the quadruplex structure was forming in this case.<sup>205</sup> However, when the four strands of the tetrahedron and both varieties of the Q<sub>TET</sub> are annealed together there were no conformational responses to the reduced pH or to added KCl in any of the possible combinations.

It appeared that the most plausible explanation for it being impossible to observe the reconfiguration of the Q<sub>TET</sub> using the different methodologies reported here was that no reconfiguration was occurring. The apparent equilibrium between the quadruplex structures and the duplex was too far in favour of the duplex structure in this system. Using this knowledge, I believe that by initially focussing on designing such a nanoarchitecture with one of the two quadruplexes would be more conducive to success in this endeavour. The segment designed to be complementary to the quadruplex of choice would benefit from containing base mismatches to destabilise the duplex and increase the probability of forming the quadruplex. Also, the bases directly flanking the quadruplex segment should be mismatched pairs to further decrease competition between the competing topologies and increase the flexibility of those regions to accommodate the formation of the quadruplex. It should also be noted that while the Q<sub>TET</sub> was envisioned to become a single contiguous polyhedron after ligation of the 'nicks' between strands, this step was never reached. This meant the quadruplexes at the extremes of the four constituent strands could not form simply in competition with the duplex. The ligation of the structure is practically guaranteed to further discourage the formation of the quadruplex.<sup>201</sup>

While I believe it to be possible to create a DNA nanoarchitecture whose size can be contextually controlled using quadruplexes, one crucial factor must be considered before making an attempt at this: there needs to be access to a more facile and conclusive method to determining the three-dimensional structure of any designed nanoarchitecture. Apart from those discussed already, a substantial amount of time and effort were dedicated to optimisation of the conditions for the formation of the Q<sub>TET</sub>, to its purification, and to other fruitless attempts at visualising the structure using different techniques such as transmission/scanning electron microscopy. This concern is shared by researchers in the field and the solution is cryo-electron microscopy (cryoEM). Namba *et al.* discuss how "confirmation of

---

accurate 3D construction [of DNA nanostructures] is particularly challenging” and present their accomplishment in using cryoEM and image analysis software to resolve the structure of a 7 nm DNA tetrahedron to nearly 10 Å resolution.<sup>206</sup> They were able to use cryoEM to not only confirm the 3D structure but also to discriminate between two structurally similar diastereomers of a DNA tetrahedron.<sup>206</sup> Access to cryoEM hardware and expertise, in addition to the continued advancements being made with the technology, will be indispensable for those progressing the field of DNA nanotechnology.

---

## **Chapter 5: Discussion and Future Work**

---

## 5.1 Discussion

At its inception, the aim of my PhD research was to design and synthesise a novel biological nanostructure whose size could be contextually controlled by exploiting two DNA quadruplex structures: the G-quadruplex and the i-motif. Ultimately, the synthesis of this reconfigurable structure was unsuccessful, but during the process of working towards this aim I was able to make a contribution to our understanding of these quadruplexes.

The investigations performed in chapter 2 aimed to expand on our knowledge of the effects of cations on i-motif DNA. This was of particular interest due to the wide-ranging uses for both in nanotechnological applications. When first attempting to explore this relationship naïveté in the handling of the i-motif led to the serendipitous discovery that it may be possible to increase the transitional pH and the thermal stability of the structure by prolonged maintenance at low temperature. Once the initial oversights were eliminated and experience was gained in working with the i-motif an intricate redox-dependent system for the control of the i-motif using copper was found and characterised using a whole host of techniques including CD, UV, FRET, NMR, and MD. The discovery of this system provided the field for the first time the ability to control DNA secondary structure in response to changes in the redox state of a metal; and opened the door to potentially developing systems which are sensitive to oxygen as well as pH.

Small molecule ligands were the next avenue explored to attempt to expand the arsenal of tools to control, in a specific manner, the i-motif structure. A variety of techniques were once again used to explore the interaction of the i-motif with these ligands including SPR, FRET, CD, UV and NMR. At first, it seemed from high-throughput FRET and SPR screens that there may have been some promising candidates. However, when moving from high-throughput methods to more meticulous methods to characterise these interactions the promise of these ligands disintegrated. Specificity between the quadruplexes continued to be elusive and, in an attempt to learn from the more mature G-quadruplex field, seven well documented 'G-quadruplex specific' ligands were chosen for investigation with the i-motif. They were found to have an unexpected, and previously unreported, ability

---

to interact with both quadruplex structures. This finding held true in both the proverbial home of the i-motif – at acidic pH, and in an environment more relevant for biological applications – neutral pH. Given the undeniable march of researchers towards developing ligands with the quadruplex structures as therapeutic targets two important factors for consideration were raised by these results: (i) the i-motif is not only relevant at acidic pH and it has now been shown to exist, and be affected by ligands, at near physiological conditions; (ii) with the quadruplexes' possible opposing functions in the control of transcription,<sup>100</sup> and with the apparent opposing effects of the same ligands on these structures, the interplay between the two quadruplexes necessitates the examination of the i-motif hand-in-hand with the G-quadruplex when hoping to find ligands that are 'specific' for either quadruplex.

The unavoidable fact that my work towards designing and synthesising the Q<sub>TET</sub> have as yet been unsuccessful does not haunt me as the knowledge and skills that I have gained on that journey have been invaluable. In addition to enabling me to complete the work described here, I plan to use what I have learned to continue to advance research in the i-motif field, and that of the wider quadruplex nucleic acids.

## 5.2 Future Work

To that end, one crucial element that will propel the i-motif field forward is the development of methodologies to collect more detailed information about the structure at the single molecule level. Crystallographic and NMR methods have as yet failed to provide the detailed structural information possible using these techniques for an unmodified intramolecular i-motif. It is known that C·C<sup>+</sup> pairs stabilise the core of the structure, however we lack the ability to determine which cytosines in a given i-motif forming sequence are involved in this pairing, and how the remaining bases in that sequence behave to form that particular i-motif (given a sequence with the potential to form more than one). Work has been done using a combination of i-motif sequences with the capacity for the i-motif to fold in only one way, and others with myriad possible combinations to form the structure (a stark example being the ATXN2L sequence used in this work which is composed of 24 cytosines). By repurposing the fluorescence enhancement observed upon the addition of thiazole orange from the FID assay, initial results have shown that

---

a correlation exists between the level of enhancement observed and different determinable characteristics of the i-motifs. Algorithms are now being designed to further this so that a set of rules can be built and used to learn more about the characteristics of 'undefined' i-motifs.

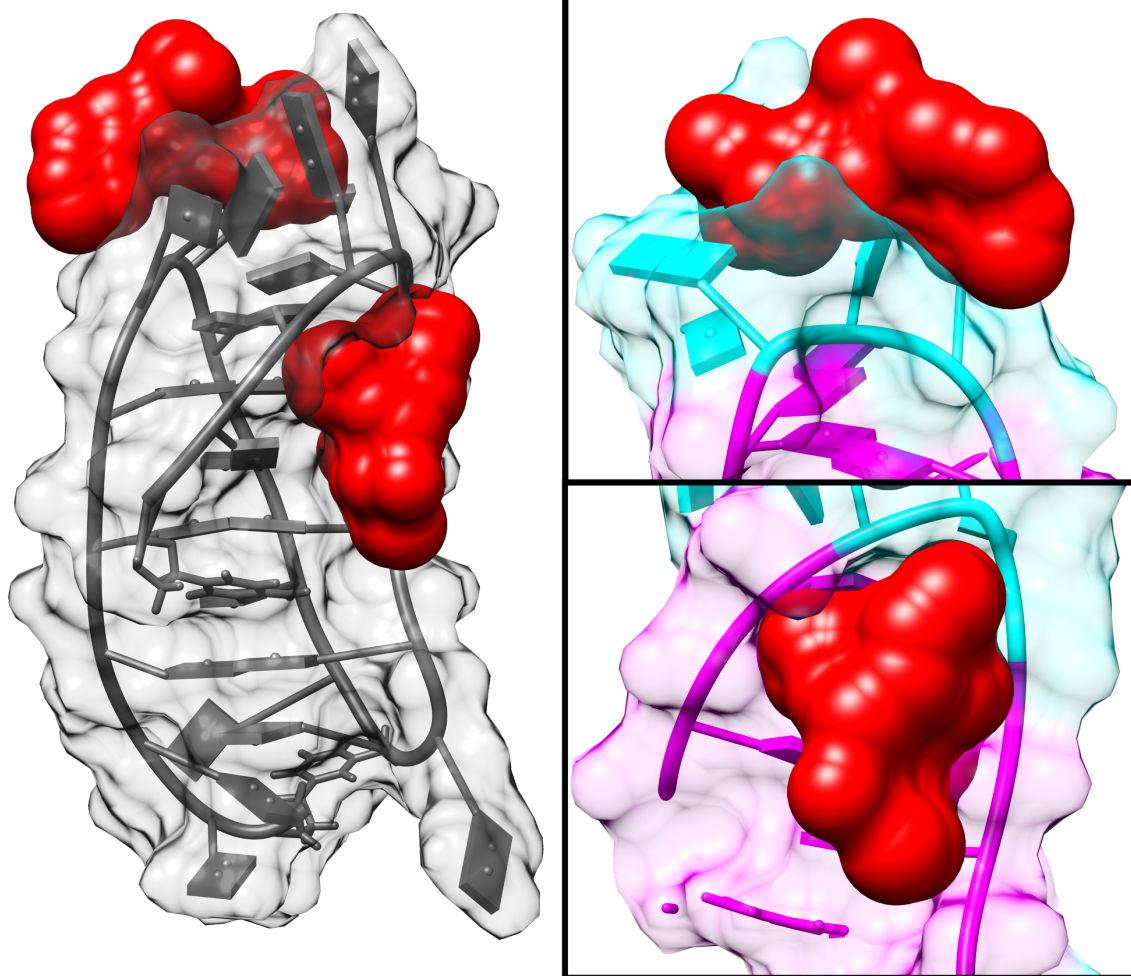
In terms of more specific work which can be carried out from the results described herein it has already been mentioned that the observations discussed in the *Tricky Temperature* section need to be consolidated. Once this phenomenon is confirmed it can be applied to create time and temperature dependent hydrogels. Combining this with the redox-dependent control of the i-motif described at the end of *Chapter 2* to further provide these hydrogels with oxygen sensitivity will result in a truly remarkable system and demonstrate the first practical application for the Cu<sup>+</sup>-stabilised i-motif.

As for the ligand work in *Chapter 3*, the discovery that reported G-quadruplex specific ligands interact with the i-motif has led to a reassessment of what ought to be done in this field. Prior to tackling the issue of intra-quadruplex selectivity, it is perhaps necessary to find a ligand that can truly stabilise the i-motif and increases the melting temperature of the structure. The two examples of i-motif stabilising ligands discussed in *Chapter 3* have, in my opinion, not been successful at doing this. Using CD melting **IMC-48** reportedly increased the melting point of the *BCL2* i-motif by "about 1°C."<sup>120</sup> The effect of **B19** on the melting temperature of an i-motif (Fpy27T) was determined using FRET-melting experiments and was reported to increase the melting temperature of that i-motif by 11.5°C.<sup>153</sup> Intriguingly, this 'selective' ligand also increased the melting temperature of a G-quadruplex (Fpu22T) by 2.7°C and of double-stranded DNA (F10T) by 1.1°C. Thus, what was touted as a stabilising effect for **IMC-48** was not considered a significant stabilisation between **B19** and the non-i-motif structures. A further concern about the published **B19** results is that, in my experience, ligands resulting in an apparent stabilisation of the i-motif observed in FRET-melting experiments do not appear to stabilise this structure when different methods are used. A good reported example of this is **Mitoxantrone** which has been shown by FRET-melting to stabilise the hTeloC i-motif,<sup>152</sup> but when tested using CD and UV-melting, and NMR, results in destabilisation.<sup>100,101</sup>

---

As alluded to at the end of section 3.2.3, what is likely to be causing these potential false-positive results is an interaction between the fluorescent tags used in the FRET-melting experiments rather than a genuine stabilisation of the quadruplex structure. Essentially, the results of FRET experiments alone only indicate that the two fluorescent tags are in proximity and any stabilisation observed should not then be attributed directly to stabilisation of the i-motif structure; these results should be confirmed by different techniques such as UV or CD melting, or NMR experiments. The primary advantage of FRET-melting was that it enabled the screening of many more samples using much lower concentrations of DNA than what is needed for the other methods. However, the high incidence of false-positive results was discouraging, and consequently it was decided to employ *in silico* screening methods in an attempt to improve the success rate of identifying promising i-motif ligands. The molecular model of hTeloC built for the copper work was used to perform docking experiments with the same 1584 ligands from the NCI library used in *Chapter 3*. Initial results were encouraging with some of the best binders identified by the docking corresponding to ligands which had been identified through high-throughput biophysical techniques as having an interaction with the i-motif. These efforts were further enhanced however when it was decided to use fpocket, a pocket detection package designed to identify 'ligandable' pockets in proteins,<sup>207</sup> on the i-motif model (Figure 5.2.1). Ligandable pockets are defined as areas on the structure which, based on a variety of complex factors, are most easy for a small molecule to bind to *in vitro*.<sup>208</sup>





*Figure 5.2.1 Ligandable pockets (red) identified on hTeloC i-motif model using fpocket. Panels on the focus on each ligandable pocket with the i-motif's loops in cyan and the cytosine tracts in magenta.*

By then focussing the docking of the NCI library ligands on these regions of the structure, an improvement in the docking results was observed with a better correspondence between the docking results and previously collected experimental data. This has provided a new factor to consider when selecting ligands to test using the more accurate, but also time and resource intensive, biophysical methods. Using these docking results, CD titration and melting experiments are being used to examine the effects of these ligands on the i-motif. This has potentially led to the discovery of some i-motif stabilising compounds, as measured by CD, and further experiments will be carried out to determine the effects of these ligands on the i-motif. With these promising results we will continue to build upon the knowledge presented here, and by carrying out the future work

---

that has been discussed I hope to be able to further our collective understanding of the fascinating intricacies of the quadruplex DNA structure: the i-motif.

---

## **Chapter 6: Experimental**

## 6.1 General Experimental

**Chemical reagents** were general purpose grade and, unless otherwise indicated, purchased from Sigma-Aldrich, Alfa Aesar, or Thermo Fisher and used without further purification.

**Oligonucleotides** were purchased dried and purified using reverse phase HPLC from Eurogentec. Dry DNA samples were dissolved in ultrapure water to prepare stock solutions at approximately 100  $\mu$ M for labelled and 1 mM for unlabelled oligonucleotides. Using the extinction coefficients provided by the manufacturer the concentrations of the stock solutions were confirmed from their UV absorbance at 260 nm with a NanoDrop ND-1000 spectrophotometer.

Name	Sequence 5' $\rightarrow$ 3'
hTeloC	TAACCCTAACCCCTAACCCCTAACCC
c-Myc	TCCCCACCTTCCCCACCCTCCCCACCCTCCCCA
ATXN2L	CCCCCCCCCCCCCCCCCCCCCCCCCCCC
DAP	CCCCCGCCCCCGCCCCCGCCCCCGCCCC
hTeloG	GGGTAGGGTTAGGGTTAGGG
JAZF	CCCCCCCCGCCCGCCCCCGCCCCCGCCCTCCCC
PDGF-A	CCGCGCCCCTCCCCCGCCCCCGCCCCCGCCCCCCCCCCCC

*Table 6.1.1 Custom oligonucleotide sequences used throughout this research.*

Name	Sequence modification-5' $\rightarrow$ 3'-modification
hTeloC <sub>FRET</sub>	FAM-TAACCCCTAACCCCTAACCCCTAACCC-TAMRA
DS <sub>FRET</sub>	FAM-TATAGCTATA-HEG(18)-TATAGCTATA-TAMRA
hTeloC <sub>biotin</sub>	Biotin-TAACCCCTAACCCCTAACCCCTAACCC
hTeloG <sub>biotin</sub>	Biotin-GGGTAGGGTTAGGGTTAGGG
c-Myc <sub>biotin</sub>	Biotin-TCCCCACCTTCCCCACCCTCCCCACCCTCCCCA
DS <sub>biotin</sub>	Biotin-GGCATAGTGCGTGGGCGTTAGC
DS <sub>comp</sub>	GCTAACGCCACGCACTATGCC
ATXN2L <sub>biotin</sub>	Biotin-CCCCCCCCCCCCCCCCCCCCCCCC
DAP <sub>biotin</sub>	Biotin-CCCCCGCCCCCGCCCCCGCCCCCGCCCC

*Table 6.1.2 Custom oligonucleotides sequences with end-modifications used throughout this research for FRET and SPR experiments.*

Name	Sequence 5' → 3'
(C <sub>4</sub> T) <sub>3</sub> C <sub>4</sub>	CCCCTCCCCTCCCCTCCCC
(G <sub>4</sub> A) <sub>3</sub> G <sub>4</sub>	GGGGAGGGGAGGGGAGGGG
(C <sub>3</sub> T) <sub>3</sub> C <sub>3</sub>	CCCTCCCTCCCTCCC
(G <sub>3</sub> A) <sub>3</sub> G <sub>3</sub>	GGGAGGGAGGGAGGG
(C <sub>2</sub> T) <sub>3</sub> C <sub>2</sub>	CCTCCTCCTCC
(G <sub>2</sub> A) <sub>3</sub> G <sub>2</sub>	GGAGGAGGAGG
(C <sub>2</sub> T <sub>2</sub> ) <sub>3</sub> C <sub>2</sub>	CCTTCCTTCCTTCC
(G <sub>2</sub> A <sub>2</sub> ) <sub>3</sub> G <sub>2</sub>	GGAAGGAAGGAAGG

Table 6.1.3 Custom oligonucleotide sequences used in this research as potential quadruplex segments for Q<sub>TET</sub>.

Sequences for tetrahedra (5' → 3'), colour coordination indicates complementary segments:

Literature Tetrahedron:

<b>S1</b>	AGG CAG TTG AG	A	CGA ACA TTC CTA AGT CTG AA	A	TTT ATC ACC CGC CAT AGT AG	A	CGT ATC ACC
<b>S2</b>	CTT GCT ACA CG	A	TTC AGA CTT AGG AAT GTT CG	A	CAT GCG AGG GTC CAA TACCG	A	CGA TTA CAG
<b>S3</b>	GGT GAT AAA	A	CG TGT AGC AAG CTG TAA TCG	A	CG GGA AGA GCA TGC CCA TCC	A	CT ACT ATG GCG
<b>S4</b>	CCT CGC ATG	A	CT CAA CTG CCT GGT GAT ACG	A	GG ATG GGC ATG CTC TTC CCG	A	CG GTA TTG GAC

Q<sub>TET</sub> (one edge):

<b>S1'</b>	AGG CAG TTG AG	A	CGA ACA TTC CC TTC CTT CCT TCC CTA AGT CTG AA	A	TTT ATC ACC CGC CAT AGT AG	A	CGT ATC ACC
<b>S2'</b>	CTT GCT ACA CG	A	TTC AGA CTT AG GGA AGG AAG GAA GG G AAT GTT CG	A	CAT GCG AGG GTC CAA TACCG	A	CGA TTA CAG
<b>S3'</b>	GGT GAT AAA	A	CG TGT AGC AAG CTG TAA TCG	A	CG GGA AGA GCA TGC CCA TCC	A	CT ACT ATG GCG
<b>S4'</b>	CCT CGC ATG	A	CT CAA CTG CCT GGT GAT ACG	A	GG ATG GGC ATG CTC TTC CCG	A	CG GTA TTG GAC

Q<sub>TET</sub> (all edges):

<b>S1'</b>	CC TTC CTT CCT TCC AGG CAG TTG AG	A	CGA ACA TTC CC TTC CTT CCT TCC CTA AGT CTG AA	A	TTT ATC ACC GGA AGG AAG GAA GG CGC CAT AGT AG	A	CGT ATC ACC
<b>S2'</b>	CC TTC CTT CCT TCC CTT GCT ACA CG	A	TTC AGA CTT AG GGA AGG AAG GAA GG G AAT GTT CG	A	CAT GCG AGG GGA AGG AAG GAA GG GTC CAA TAC CG	A	CGA TTA CAG
<b>S3'</b>	CC TTC CTT CCT TCC GGT GAT AAA	A	CG TGT AGC AAG GGA AGG AAG GAA GG CTG TAA TCG	A	CG GGA AGA GCA CC TTC CTT CCT TCC TGC CCA TCC	A	CT ACT ATG GCG
<b>S4'</b>	CC TTC CTT CCT TCC CCT CGC ATG	A	CT CAA CTG CCT GGA AGG AAG GAA GG GGT GAT ACG	A	GG ATG GGC A GGA AGG AAG GAA GG TG CTC TTC CCG	A	CG GTA TTG GAC

---

**DNA annealing**, unless otherwise indicated, was performed either in a heating block, or in an Applied Biosystems Veriti 96 well thermal cycler. Using the heating block, samples were placed into the block at 95°C for five minutes and then it was powered off and the samples allowed to slowly cool down to room temperature. Using the thermal cycler samples were heated to 95°C and held at that temperature for 5 minutes. The temperature was then reduced at a rate of 1°C/min to 20°C.

**Circular dichroism (CD) and UV spectroscopy**, unless otherwise indicated, were performed using a Jasco J-810 spectropolarimeter using a 1 mm path length quartz cuvette. Scans were performed with a scanning speed of 200 nm/min, response time of 1 s, 0.5 nm pitch and 2 nm bandwidth. The spectra are an accumulation of either 3 or 4 scans and are zero corrected at the longest wavelength of the measurement. Melting experiments were performed using the same measurement parameters while heating the sample at a rate of 1°C/min within the desired temperature range and measuring at 5°C intervals. The temperature at which 50% of the thermal denaturation had taken place ( $T_m$ ) was calculated using Origin data analysis software to plot normalised ellipticity at 288 nm against temperature for the CD melts and normalised absorbance at 295 nm against temperature for the UV melts. These data were fitted with sigmoidal curves and the equations of the curves solved for  $y = 0.5$  to give the  $T_m$  values.

**FRET-melting** experiments were performed using 200 nM DNA concentration in QIAGEN strip-tubes with a final volume of 20  $\mu$ L per sample. Fluorescence melting curves were acquired in a QIAGEN Rotor-Gene Q-series PCR machine. Measurements were made with excitation at 483 nm and detection at 533 nm, with final analysis of the data carried out using QIAGEN Rotor-Gene Q-series software and Origin.

**FID** experiments were performed on a BMG CLARIOstar plate reader using Corning 96-Well Solid Black Flat Bottom plates. A 10 mM stock solution of TO was prepared in DMSO and dilution to 2  $\mu$ M in the appropriate buffer. 90  $\mu$ L of the 2  $\mu$ M TO solution were added to each well and fluorescence emission at 450 nm measured with excitation at 430 nm; this was normalised to 0% representing background fluorescence. 1  $\mu$ L of 90  $\mu$ M DNA was then added, shaken using

---

double orbital shaking at 700 rpm in the plate reader for 15 seconds, and allowed to equilibrate for 15 minutes. After equilibration, fluorescence emission was measured once again and normalised to 100% representing maximum fluorescence enhancement from the TO probe binding the DNA secondary structure. 0.9  $\mu\text{L}$  aliquots of ligand were titrated into each well (in triplicate) and measured as before. Fluorescence measurements after ligand addition were normalised between the 0 and 100% levels determined per the respective well. Percentage TO displacement was calculated as the difference between the normalised 100% fluorescence level and the normalised fluorescence measured after each ligand addition. The concentration for each ligand at which 50% of the TO was displaced was calculated using Origin data analysis software to plot percentage TO displacement against ligand concentration. These data were fitted with dose-response curves and the equations of the curves were solved for  $y = 50$  to determine the  $\text{DC}_{50}$  values.

## 6.2 Experimental for Chapter 2

Experiments with  $\text{Cu}^+$  requiring anoxic conditions were performed in a UNIlab Plus glove box workstation under nitrogen with the oxygen level maintained below 0.5 ppm. Unless otherwise indicated, cation solutions were prepared by dissolving the solid metal salts in ultrapure water to the required concentrations.  $\text{Cu}^+$  solutions were prepared by dissolving solid copper chloride in a 0.1 M HCl and 1 M NaCl solution to double the desired concentration and then diluted to the final concentration using 2x sodium cacodylate buffer. Solutions of  $\text{Cu}^{2+}$ , DETC and sodium ascorbate were prepared by dissolving solid  $\text{Cu}_2\text{SO}_4$ , DETC or sodium ascorbate in ultrapure water, respectively.

**CD spectroscopy** was performed as described above with a variation in the cuvette: the 1 mm path length quartz cuvette that was used had a neck, and a silicone stopper was used to exclude air on transfer from the glove box to the spectropolarimeter. 10  $\mu\text{M}$  hTeloC in 50 mM sodium cacodylate buffer at the desired pH was used with a 200  $\mu\text{L}$  sample volume.  $\text{Cu}^+$  titrations were performed in 0.5  $\mu\text{L}$  aliquots to yield the desired concentrations. For the chelator titration 150  $\mu\text{M}$   $\text{Cu}^+$  was added via a 1  $\mu\text{L}$  addition from a 30 mM stock solution, DETC

---

was then added in 0.5  $\mu\text{L}$  aliquots to the desired concentration as above. Control spectra of DETC in buffer and DETC with hTeloC were also measured to confirm DETC itself had no effect on the spectra or the conformation of the DNA. For the redox experiments, each component was added via a 1  $\mu\text{L}$  addition from a stock solution prepared at the concentration needed to yield the desired concentration. Control spectra of sodium ascorbate in buffer and sodium ascorbate with hTeloC were also measured to confirm that sodium ascorbate itself had no effect on the spectra or the conformation of the DNA. Titration experiments were performed at least in triplicate and processing of the data was carried out using Origin.

**UV spectroscopy** experiments were performed on an Agilent Technologies Cary 4000 UV-Vis spectrophotometer and recorded using an open-top screw-cap 10 mm quartz cuvette with a silicone rubber septum to exclude air on transfer from the glove box to the spectrophotometer. Samples (1 mL) were diluted to give 2.5  $\mu\text{M}$  hTeloC in 10 mM sodium cacodylate buffer at the desired pH.  $\text{Cu}^+$  was added in 0.5  $\mu\text{L}$  aliquots, and mixed, using a pipette to the desired concentration. Spectra were recorded over a wavelength range of 200 to 400 nm at room temperature and zero corrected at 400 nm.

**Molecular modelling and dynamics** began by building initial models of  $\text{C}\cdot\text{Cu}^+\cdot\text{C}$  and  $\text{C}\cdot\text{Cu}^{2+}\cdot\text{C}$  using the program Avogadro.<sup>146</sup> These were optimized by the DFT methods B3LYP-D3/def2-TZVP and TPSS-D3(BJ)/def2-TZVP using the NWChem package.<sup>144</sup> For the i-motif model, the starting point for was the NMR solution structure of a sequence very similar to hTeloC: (CCCTAA5mCCCTAACCCUAACCCT), PDB entry 1EL2.<sup>96</sup> Its model was manually edited in Avogadro<sup>146</sup> to match the hTeloC sequence used in the biophysical experiments. A molecular dynamics run was performed to relax the modified structure in order to allow minor conformational changes in response to the alteration of the sequence.

The AmberTools 16 package<sup>209</sup> was used to assign force field parameters to the model and to add 17 charge-compensating sodium ions. For DNA the OL15 parameter set was used,<sup>210–212</sup> for the ions the parameters of Li, Song and Merz<sup>213</sup> together with the TIP3P water model.<sup>214</sup> The protonated cytosines were modelled



---

using parameters taken from analogous atoms in neutral residues, with the exception of partial charges. These were fitted to reproduce the electrostatic potential of a geometry optimized B3LYP/6-31G\* model of the protonated nucleoside. Only the charges of the protonated cytosine and the C1'-H1' atoms of the sugar were taken from the RESP fitting procedure, where their sum was constrained to give a molecular charge of +1 when used in combination with standard force field charges for the rest of the sugar fragment.

After minimisation using an implicit solvent model with sander,<sup>209</sup> the model was transferred to the Gromacs package<sup>215</sup> using a locally modified version of the `amber_to_gmx_by_XL` script.<sup>216</sup> Calculated energies between the two packages agreed to within 0.1%. The model was solvated in a cubic box with a 63.37 Å axis (8057 water molecules). The positions of the water molecules and sodium ions were energy minimized and then equilibrated using constant volume and temperature (NVT, T = 300 K) molecular dynamics for 100 ps, followed by a constant pressure and temperature simulation (NPT, p = 0.1 MPa) for 200 ps. In subsequent equilibration runs (1 ns NVT, 2 ns NPT) the coordinates of the DNA as well as the solvent molecules and counterions were allowed to change. The length of the production run was 200 ns. Each simulation used a time step of 2 fs.

**<sup>1</sup>H NMR** experiments were performed using a Bruker Avance III 800 MHz spectrometer equipped with an HCN inverse triple resonance z-gradient probe. Aqueous solutions were prepared with the addition of 5% D<sub>2</sub>O to enable field/frequency lock. Solvent suppression of the water resonance was achieved using a 1D Watergate sequence employing a symmetrical 3-τ-9-τ-19 pulse train inversion element. The solvent resonance, which was minimized, was set on-resonance at the transmitter offset and the interpulse delay time (τ) was adjusted to achieve an excitation maximum in the imino proton region of interest. hTeloC was diluted to a concentration of 10 μM in 50 mM sodium cacodylate buffer at pH 5.5 containing 5% D<sub>2</sub>O. The spectrum of hTeloC alone was measured over 1 hour after which 150 μM of Cu<sup>+</sup> was added and the subsequent spectrum acquired over 2 hours. Finally, excess DETC (540 μM) was added and the spectrum acquired again for 1 hour. NMR spectra were acquired and processed using

---

Bruker's TopSpin™ software package (v3.1.7 Bruker Biospin) for NMR data analysis.

**FRET-melting** was enabled once the ability to maintain the oxidation state of the  $\text{Cu}^+$  was characterised. Here, hTeloC<sub>FRET</sub> was prepared as a 400 nM solution in 10 mM sodium cacodylate buffer at the respective pH and thermally annealed. Samples were prepared by aliquoting 10  $\mu\text{L}$  of the annealed DNA, followed by addition of  $\text{Cu}^{2+}$  solution and 2x sodium ascorbate solution to give the desired  $\text{Cu}^+$  concentration range across the samples, and made up with 10 mM sodium cacodylate buffer to a final volume of 20  $\mu\text{L}$ .

### 6.3 Experimental for Chapter 3

**SPR experiments** were performed using the GE Healthcare Biacore T200 with a series S streptavidin (SA) coated gold chip (Biacore SA-chip). A 10 mM sodium cacodylate, 100 mM sodium chloride and 0.05% Tween running buffer was used at the required pH. Initial screen was performed at pH 5.5 with all the compounds measured at 50  $\mu\text{M}$  from 1 mM stocks in DMSO and a corresponding amount of DMSO (5%) in the running buffer. Subsequent testing was performed over a concentration range from 0 to 500  $\mu\text{M}$  with the ligands prepared from a 10 mM stock in 100% DMSO. The stock was diluted to 0.5 mM in running buffer to give a 5% DMSO concentration, further serial dilutions were performed using buffer containing DMSO to maintain the concentration of DMSO at 5% for all concentration. Additionally, solvent correction was performed using a series of DMSO concentrations from 4.5 to 5.8% in running buffer.

**FRET-melting** experiments were performed as described above however, for the control melting curves for the DNA without ligand DMSO was added so that the concentration of DMSO was maintained the same for all samples.

**$^1\text{H}$  NMR** performed in Brno were collected on a 700 MHz Bruker Avance III HD spectrometer using the p3919gp pulse program modified to suppress DMSO signal. Samples were 500  $\mu\text{L}$  and DNA concentration was 100  $\mu\text{M}$  in 140 mM

---

NaPO<sub>i</sub> buffer with 10 mM MgCl<sub>2</sub>, 5 mM KCl and 10% D<sub>2</sub>O at 25°C. Ligand additions were performed from 10 mM stock solutions in DMSO. Spectra were corrected for the effect of DMSO using an equivalent titration of DMSO without ligand.

## 6.4 Experimental for Chapter 4

**DNA annealing** in this section was performed using the Applied Biosystems Veriti 96 well thermal cycler. Samples were always first heated to 95°C and held at that temperature for 5 minutes. For the initial rapid synthesis of the tetrahedron, samples were then cooled directly to 4°C (~30 seconds). This rapid cooling was then abandoned in favour of the more uniform structure formation resulting from a non-rapid procedure where the temperature was reduced at a rate of 1°C/min to 4°C.

**Native PAGE** was performed using 6-10% gels (19:1 acrylamide:bisacrylamide mixture) in 1x TAE buffer and stained with Stains-All for visualisation. Images were acquired on an ImageQuant LAS 4000.

**AFM** experiments were carried out using quantitative imaging mode or tapping mode using a JPK Nanowizard III. Samples of either the four constituent strands in 10 mM Tris buffer combined without an annealing step, or the annealed tetrahedron in 10 mM Tris, 5 mM MgCl<sub>2</sub> buffer were diluted to 10 nM and incubated for 30 minutes on freshly cleaved mica. To ensure attachment of the sample to the surface NiCl<sub>2</sub> was added to a final concentration of 2 nM. 1x1 micron images were collected with a height scale of 3 nm (black to white). The background colour was adjusted to be approximately equal in all cases to enable direct comparison of images. Force spectra were exported and analysed using JPK's data processing software.

**DLS** measurements were carried out on a Malvern Zetasizer Nanoseries Nano-ZS instrument using micro range BRAND X100 UV cuvettes. 100 µL samples were annealed at 250 nM in the respective buffer conditions for the desired conformation. The dispersant refractive index was measured for each buffer condition and subsequent samples measurements corrected accordingly. Samples

---

were irradiated at 633 nm and the scattering measured at an angle of 173°. Results show the average and standard deviation of three measurements of the size distribution by number. Data was analysed using Malvern Zetasizer software v7.13 and Origin data analysis software.

**Molecular modelling** began by building models of the tetrahedra using UCSF Chimera software version 1.13.1. Sequences were built as B-form DNA helices using the sequences of each of the component strands of the tetrahedra. Vertex bases were then modified where required from thymine to adenine. The model was then manually repositioned to resemble the supposed tetrahedral structure. pdb files were then exported and atom properties were manually edited to enable concatenation of the component strands at their 5' and 3' ends. Energy minimisation of the structure was performed using the built-in 'Minimize Structure' function of chimera initially using 1000 steepest descent steps and 100 conjugate gradient steps using 1 Å step sizes. This was followed by 100 steepest descent steps and 100 conjugate gradient steps using the default 0.02 Å step size until all clashes were eliminated. Clashes were defined per Chimera's default settings as a Van der Waals overlap  $\geq 0.6$  Å.

---

## References

- 1) Watson, J. D.; Crick, F. H. C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **1953**, *171* (4356), 737–738. <https://doi.org/10.1038/171737a0>.
- 2) Brigandt, I.; Love, A. Reductionism in Biology. In *The Stanford Encyclopedia of Philosophy*; Zalta, E. N., Ed.; Metaphysics Research Lab, Stanford University, 2017.
- 3) Scofield, M. Nucleic Acids. In *xPharm: The Comprehensive Pharmacology Reference*; Elsevier, 2007; pp 1–15. <https://doi.org/10.1016/B978-008055232-3.60059-5>.
- 4) Pray, L. A. Discovery of DNA Structure and Function: Watson and Crick. *Nat. Educ.* **2008**.
- 5) Hud, N. V.; Plavec, J. A Unified Model for the Origin of DNA Sequence-Directed Curvature. *Biopolymers* **2003**, *69* (1), 144–158. <https://doi.org/10.1002/bip.10364>.
- 6) Saintomé, C.; Amrane, S.; Mergny, J.-L.; Alberti, P. The Exception That Confirms the Rule: A Higher-Order Telomeric G-Quadruplex Structure More Stable in Sodium than in Potassium. *Nucleic Acids Res.* **2016**, *44* (6), 2926–2935. <https://doi.org/10.1093/nar/gkw003>.
- 7) Amato, J.; Iaccarino, N.; Randazzo, A.; Novellino, E.; Pagano, B. Noncanonical DNA Secondary Structures as Drug Targets: The Prospect of the i-Motif. *ChemMedChem* **2014**, *9* (9), 2026–2030. <https://doi.org/10.1002/cmdc.201402153>.
- 8) Wright, E. P.; Huppert, J. L.; Waller, Z. A. E. Identification of Multiple Genomic DNA Sequences Which Form I-Motif Structures at Neutral PH. *Nucleic Acids Res.* **2017**, *45* (6), 2951–2959. <https://doi.org/10.1093/nar/gkx090>.
- 9) Sun, D.; Hurley, L. H. The Importance of Negative Superhelicity in Inducing the Formation of G-Quadruplex and i-Motif Structures in the c-Myc Promoter: Implications for Drug Targeting and Control of Gene Expression. *J. Med. Chem.* **2009**, *52* (9), 2863–2874. <https://doi.org/10.1021/jm900055s>.
- 10) Kaiser, C. E.; Van Ert, N. A.; Agrawal, P.; Chawla, R.; Yang, D.; Hurley, L. H. Insight into the Complexity of the I-Motif and G-Quadruplex DNA Structures Formed in the KRAS Promoter and Subsequent Drug-Induced Gene Repression. *J. Am. Chem. Soc.* **2017**, *139* (25), 8522–8536.

---

<https://doi.org/10.1021/jacs.7b02046>.

- (11) Day, H. A.; Pavlou, P.; Waller, Z. A. E. I-Motif DNA: Structure, Stability and Targeting with Ligands. *Bioorg. Med. Chem.* **2014**, *22* (16), 4407–4418. <https://doi.org/10.1016/j.bmc.2014.05.047>.
- (12) Dhakal, S.; Yu, Z.; Konik, R.; Cui, Y.; Koirala, D.; Mao, H. G-Quadruplex and i-Motif Are Mutually Exclusive in ILPR Double-Stranded DNA. *Biophys. J.* **2012**, *102* (11), 2575–2584. <https://doi.org/10.1016/j.bpj.2012.04.024>.
- (13) Cui, Y.; Kong, D.; Ghimire, C.; Xu, C.; Mao, H. Mutually Exclusive Formation of G-Quadruplex and i-Motif Is a General Phenomenon Governed by Steric Hindrance in Duplex DNA. *Biochemistry* **2016**, *55* (15), 2291–2299. <https://doi.org/10.1021/acs.biochem.6b00016>.
- (14) Gellert, M.; Lipsett, M. N.; Davies, D. R. Helix Formation by Guanylic Acid. *Proc. Natl. Acad. Sci.* **1962**, *48* (12), 2013–2018. <https://doi.org/10.1073/pnas.48.12.2013>.
- (15) Sundquist, W. I.; Klug, A. Telomeric DNA Dimerizes by Formation of Guanine Tetrads between Hairpin Loops. *Nature* **1989**, *342* (6251), 825–829. <https://doi.org/10.1038/342825a0>.
- (16) Lipps, H. J.; Gruissem, W.; Prescott, D. M. Higher Order DNA Structure in Macronuclear Chromatin of the Hypotrichous Ciliate *Oxytricha Nova*. *Proc. Natl. Acad. Sci. U. S. A.* **1982**, *79* (8), 2495–2499. <https://doi.org/10.1073/pnas.79.8.2495>.
- (17) Rhodes, D.; Lipps, H. J. G-Quadruplexes and Their Regulatory Roles in Biology. *Nucleic Acids Res.* **2015**, *43* (18), 8627–8637. <https://doi.org/10.1093/nar/gkv862>.
- (18) Huppert, J. L.; Balasubramanian, S. Prevalence of Quadruplexes in the Human Genome. *Nucleic Acids Res.* **2005**, *33* (9), 2908–2916. <https://doi.org/10.1093/nar/gki609>.
- (19) Todd, A. K.; Johnston, M.; Neidle, S. Highly Prevalent Putative Quadruplex Sequence Motifs in Human DNA. *Nucleic Acids Res.* **2005**, *33* (9), 2901–2907. <https://doi.org/10.1093/nar/gki553>.
- (20) Chambers, V. S.; Marsico, G.; Boutell, J. M.; Di Antonio, M.; Smith, G. P.; Balasubramanian, S. High-Throughput Sequencing of DNA G-Quadruplex Structures in the Human Genome. *Nat. Biotechnol.* **2015**, *33* (8), 877–881. <https://doi.org/10.1038/nbt.3295>.

- 
- (21) Bedrat, A.; Lacroix, L.; Mergny, J.-L. Re-Evaluation of G-Quadruplex Propensity with G4Hunter. *Nucleic Acids Res.* **2016**, *44* (4), 1746–1759. <https://doi.org/10.1093/nar/gkw006>.
- (22) Frees, S.; Menendez, C.; Crum, M.; Bagga, P. S. QGRS-Conserve: A Computational Method for Discovering Evolutionarily Conserved G-Quadruplex Motifs. *Hum. Genomics* **2014**, *8* (1), 1–13. <https://doi.org/10.1186/1479-7364-8-8>.
- (23) Lam, E. Y. N.; Beraldi, D.; Tannahill, D.; Balasubramanian, S. G-Quadruplex Structures Are Stable and Detectable in Human Genomic DNA. *Nat. Commun.* **2013**, *4*, 1796. <https://doi.org/10.1038/ncomms2792>.
- (24) Huppert, J. L.; Balasubramanian, S. G-Quadruplexes in Promoters throughout the Human Genome. *Nucleic Acids Res.* **2007**, *35* (2), 406–413. <https://doi.org/10.1093/nar/gkl1057>.
- (25) Beaume, N.; Pathak, R.; Yadav, V. K.; Kota, S.; Misra, H. S.; Gautam, H. K.; Chowdhury, S. Genome-Wide Study Predicts Promoter-G4 DNA Motifs Regulate Selective Functions in Bacteria: Radioresistance of *D. Radiodurans* Involves G4 DNA-Mediated Regulation. *Nucleic Acids Res.* **2013**, *41* (1), 76–89. <https://doi.org/10.1093/nar/gks1071>.
- (26) Norseen, J.; Johnson, F. B.; Lieberman, P. M. Role for G-Quadruplex RNA Binding by Epstein-Barr Virus Nuclear Antigen 1 in DNA Replication and Metaphase Chromosome Attachment. *J. Virol.* **2009**, *83* (20), 10336–10346. <https://doi.org/10.1128/JVI.00747-09>.
- (27) Sundquist, W. I.; Heaphy, S. Evidence for Interstrand Quadruplex Formation in the Dimerization of Human Immunodeficiency Virus 1 Genomic RNA. *Proc. Natl. Acad. Sci.* **1993**, *90* (8), 3393–3397. <https://doi.org/10.1073/pnas.90.8.3393>.
- (28) Lipps, H. J.; Rhodes, D. G-Quadruplex Structures: In Vivo Evidence and Function. *Trends Cell Biol.* **2009**, *19* (8), 414–422. <https://doi.org/10.1016/j.tcb.2009.05.002>.
- (29) Waller, Z. A. E.; Sewitz, S. A.; Hsu, S. D.; Balasubramanian, S. A Small Molecule That Disrupts G-Quadruplex DNA Structure and Enhances Gene Expression. *J. Am. Chem. Soc.* **2009**, *131* (35), 12628–12633. <https://doi.org/10.1021/ja901892u>.
- (30) Farhath, M. M.; Thompson, M.; Ray, S.; Sewell, A.; Balci, H.; Basu, S. G-



- 
- Quadruplex-Enabling Sequence within the Human Tyrosine Hydroxylase Promoter Differentially Regulates Transcription. *Biochemistry* **2015**, *54* (36), 5533–5545. <https://doi.org/10.1021/acs.biochem.5b00209>.
- (31) Xiao, S.; Zhang, J.-Y.; Zheng, K.-W.; Hao, Y.-H.; Tan, Z. Bioinformatic Analysis Reveals an Evolutional Selection for DNA:RNA Hybrid G-Quadruplex Structures as Putative Transcription Regulatory Elements in Warm-Blooded Animals. *Nucleic Acids Res.* **2013**, *41* (22), 10379–10390. <https://doi.org/10.1093/nar/gkt781>.
- (32) Baral, A.; Kumar, P.; Pathak, R.; Chowdhury, S. Emerging Trends in G-Quadruplex Biology – Role in Epigenetic and Evolutionary Events. *Mol. Biosyst.* **2013**, *9* (7), 1568. <https://doi.org/10.1039/c3mb25492e>.
- (33) Guilbaud, G.; Murat, P.; Recolin, B.; Campbell, B. C.; Maiter, A.; Sale, J. E.; Balasubramanian, S. Local Epigenetic Reprogramming Induced by G-Quadruplex Ligands. *Nat. Chem.* **2017**, *9* (11), 1110–1117. <https://doi.org/10.1038/nchem.2828>.
- (34) Murat, P.; Balasubramanian, S. Existence and Consequences of G-Quadruplex Structures in DNA. *Curr. Opin. Genet. Dev.* **2014**, *25* (1), 22–29. <https://doi.org/10.1016/j.gde.2013.10.012>.
- (35) Abou Assi, H.; Garavís, M.; González, C.; Damha, M. J. I-Motif DNA: Structural Features and Significance to Cell Biology. *Nucleic Acids Res.* **2018**, *46* (16), 8038–8056. <https://doi.org/10.1093/nar/gky735>.
- (36) Gehring, K.; Leroy, J.-L.; Guéron, M. A Tetrameric DNA Structure with Protonated Cytosine-Cytosine Base Pairs. *Nature* **1993**, *363* (6429), 561–565. <https://doi.org/10.1038/363561a0>.
- (37) Brazier, J. A.; Shah, A.; Brown, G. D. I-Motif Formation in Gene Promoters: Unusually Stable Formation in Sequences Complementary to Known G-Quadruplexes. *Chem. Commun.* **2012**, *48* (87), 10739. <https://doi.org/10.1039/c2cc30863k>.
- (38) Hardin, C. C.; Corregan, M.; Brown, B. A.; Frederick, L. N. Cytosine-Cytosine+ Base Pairing Stabilizes DNA Quadruplexes and Cytosine Methylation Greatly Enhances the Effect. *Biochemistry* **1993**, *32* (22), 5870–5880.
- (39) Kang, C.; Berger, I.; Lockshin, C.; Ratliff, R.; Moyzis, R.; Rich, A. Stable Loop in the Crystal Structure of the Intercalated Four-Stranded Cytosine-Rich

- Metazoan Telomere. *Proc. Natl. Acad. Sci.* **1995**, *92* (9), 3874–3878. <https://doi.org/10.1073/pnas.92.9.3874>.
- (40) Catasti, P.; Chen, X.; Deaven, L. L.; Moyzis, R. K.; Bradbury, E. M.; Gupta, G. Cytosine-Rich Strands of the Insulin Minisatellite Adopt Hairpins with Intercalated Cytosine + ·Cytosine Pairs 1 Edited by I. Tinoco. *J. Mol. Biol.* **1997**, *272* (3), 369–382. <https://doi.org/10.1006/jmbi.1997.1248>.
- (41) Mergny, J.-L.; Lacroix, L.; Han, X.; Leroy, J.-L.; Helene, C. Intramolecular Folding of Pyrimidine Oligodeoxynucleotides into an I-DNA Motif. *J. Am. Chem. Soc.* **1995**, *117* (35), 8887–8898. <https://doi.org/10.1021/ja00140a001>.
- (42) Dong, Y.; Yang, Z.; Liu, D. DNA Nanotechnology Based on I-Motif Structures. *Acc. Chem. Res.* **2014**, *47* (6), 1853–1860. <https://doi.org/10.1021/ar500073a>.
- (43) Wellinger, R. J.; Sen, D. The DNA Structures at the Ends of Eukaryotic Chromosomes. *Eur. J. Cancer* **1997**, *33* (5), 735–749. [https://doi.org/10.1016/S0959-8049\(97\)00067-1](https://doi.org/10.1016/S0959-8049(97)00067-1).
- (44) Phan, A. T.; Mergny, J.-L. Human Telomeric DNA: G-Quadruplex, i-Motif and Watson-Crick Double Helix. *Nucleic Acids Res.* **2002**, *30* (21), 4618–4625. <https://doi.org/10.1093/nar/gkf597>.
- (45) Lu, W.; Zhang, Y.; Liu, D.; Songyang, Z.; Wan, M. Telomeres - Structure, Function, and Regulation. *Exp. Cell Res.* **2013**, *319* (2), 133–141. <https://doi.org/10.1016/j.yexcr.2012.09.005>.
- (46) Shay, J. W.; Bacchetti, S. A Survey of Telomerase Activity in Human Cancer. *Eur. J. Cancer* **1997**, *33* (5), 787–791. [https://doi.org/10.1016/S0959-8049\(97\)00062-2](https://doi.org/10.1016/S0959-8049(97)00062-2).
- (47) Cesare, A. J.; Reddel, R. R. Alternative Lengthening of Telomeres: Models, Mechanisms and Implications. *Nat. Rev. Genet.* **2010**, *11* (5), 319–330. <https://doi.org/10.1038/nrg2763>.
- (48) Zahler, A. M.; Williamson, J. R.; Cech, T. R.; Prescott, D. M. Inhibition of Telomerase by G-Quartet DNA Structures. *Nature* **1991**, *350* (6320), 718–720. <https://doi.org/10.1038/350718a0>.
- (49) Maji, B.; Bhattacharya, S. Advances in the Molecular Design of Potential Anticancer Agents via Targeting of Human Telomeric DNA. *Chem. Commun.* **2014**, *50* (49), 6422–6438. <https://doi.org/10.1039/c4cc00611a>.

- 
- (50) Balasubramanian, S.; Neidle, S. G-Quadruplex Nucleic Acids as Therapeutic Targets. *Curr. Opin. Chem. Biol.* **2009**, *13* (3), 345–353. <https://doi.org/10.1016/j.cbpa.2009.04.637>.
- (51) Chen, Y.; Qu, K.; Zhao, C.; Wu, L.; Ren, J.; Wang, J.; Qu, X. Insights into the Biomedical Effects of Carboxylated Single-Wall Carbon Nanotubes on Telomerase and Telomeres. *Nat. Commun.* **2012**, *3* (1), 1074. <https://doi.org/10.1038/ncomms2091>.
- (52) 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.; et al. NM23-H2 May Play an Indirect Role in Transcriptional Activation of c-Myc Gene Expression but Does Not Cleave the Nuclease Hypersensitive Element III1. *Mol. Cancer Ther.* **2009**, *8* (5), 1363–1377. <https://doi.org/10.1158/1535-7163.MCT-08-1093>.
- (53) Kendrick, S.; Akiyama, Y.; Hecht, S. M.; Hurley, L. H. The I-Motif in the Bcl-2 P1 Promoter Forms an Unexpectedly Stable Structure with a Unique 8:5:7 Loop Folding Pattern. *J. Am. Chem. Soc.* **2009**, *131* (48), 17667–17676. <https://doi.org/10.1021/ja9076292>.
- (54) Marsich, E.; Piccini, A.; Xodo, L. E.; Manzini, G. Evidence for a HeLa Nuclear Protein That Binds Specifically to the Single-Stranded d(CCCTAA)<sub>n</sub> Telomeric Motif. *Nucleic Acids Res.* **1996**, *24* (20), 4029–4033. <https://doi.org/10.1093/nar/24.20.4029>.
- (55) Brooks, T. A.; Kendrick, S.; Hurley, L. H. Making Sense of G-Quadruplex and i-Motif Functions in Oncogene Promoters. *FEBS J.* **2010**, *277* (17), 3459–3469. <https://doi.org/10.1111/j.1742-4658.2010.07759.x>.
- (56) Kang, H.; Kendrick, S.; Hecht, S. M.; Hurley, L. H. The Transcriptional Complex Between the BCL2 I-Motif and HnRNP LL Is a Molecular Switch for Control of Gene Expression That Can Be Modulated by Small Molecules. *J. Am. Chem. Soc.* **2014**, *136* (11), 4172–4185. <https://doi.org/10.1021/ja4109352>.
- (57) Balasubramanian, S.; Hurley, L. H.; Neidle, S. Targeting G-Quadruplexes in Gene Promoters: A Novel Anticancer Strategy? *Nat. Rev. Drug Discov.* **2011**, *10* (4), 261–275. <https://doi.org/10.1038/nrd3428>.
- (58) Hänsel-Hertsch, R.; Di Antonio, M.; Balasubramanian, S. DNA G-Quadruplexes in the Human Genome: Detection, Functions and Therapeutic

- Potential. *Nat. Rev. Mol. Cell Biol.* **2017**, *18* (5), 279–284. <https://doi.org/10.1038/nrm.2017.3>.
- (59) Debnath, M.; Ghosh, S.; Chauhan, A.; Paul, R.; Bhattacharyya, K.; Dash, J. Preferential Targeting of I-Motifs and G-Quadruplexes by Small Molecules. *Chem. Sci.* **2017**, *8* (11), 7448–7456. <https://doi.org/10.1039/C7SC02693E>.
- (60) Dzatko, S.; Krafcikova, M.; Hänsel-Hertsch, R.; Fessler, T.; Fiala, R.; Loja, T.; Krafcik, D.; Mergny, J.-L.; Foldynova-Trantirkova, S.; Trantirek, L. Evaluation of the Stability of DNA I-Motifs in the Nuclei of Living Mammalian Cells. *Angew. Chemie Int. Ed.* **2018**, *57* (8), 2165–2169. <https://doi.org/10.1002/anie.201712284>.
- (61) Zeraati, M.; Langley, D. B.; Schofield, P.; Moye, A. L.; Rouet, R.; Hughes, W. E.; Bryan, T. M.; Dinger, M. E.; Christ, D. I-Motif DNA Structures Are Formed in the Nuclei of Human Cells. *Nat. Chem.* **2018**, *10* (6), 631–637. <https://doi.org/10.1038/s41557-018-0046-3>.
- (62) Amos, M. DNA: The Molecule of Life. In *Theoretical and Experimental DNA Computation*; Natural Computing Series; Springer-Verlag: Berlin/Heidelberg, 2005; pp 5–21. <https://doi.org/10.1007/3-540-28131-2>.
- (63) Seeman, N. C. DNA in a Material World. *Nature* **2003**, *421* (6921), 427–431. <https://doi.org/10.1038/nature01406>.
- (64) Lin, C.; Liu, Y.; Yan, H. Designer DNA Nanoarchitectures. *Biochemistry* **2009**, *48* (8), 1663–1674. <https://doi.org/10.1021/bi802324w>.
- (65) Veneziano, R.; Ratanalert, S.; Zhang, K.; Zhang, F.; Yan, H.; Chiu, W.; Bathe, M.; Veneziano, R. Designer Nanoscale DNA Assemblies Programmed from the Top Down. **2016**. <https://doi.org/10.1126/science.aaf4388>.
- (66) Douglas, S. M.; Dietz, H.; Liedl, T.; Högberg, B.; Graf, F.; Shih, W. M. Self-Assembly of DNA into Nanoscale Three-Dimensional Shapes. *Nature* **2009**, *459* (7245), 414–418. <https://doi.org/10.1038/nature08016>.
- (67) Krishnan, Y.; Simmel, F. C. Nucleic Acid Based Molecular Devices. *Angew. Chemie Int. Ed.* **2011**, *50* (14), 3124–3156. <https://doi.org/10.1002/anie.200907223>.
- (68) Weiss, P. S. A Conversation with Prof. Ned Seeman: Founder of DNA Nanotechnology. *ACS Nano* **2008**, *2* (6), 1089–1096. <https://doi.org/10.1021/nn800316c>.

- 
- (69) Chen, J.; Seeman, N. C. Synthesis from DNA of a Molecule with the Connectivity of a Cube. *Nature* **1991**, *350* (6319), 631–633. <https://doi.org/10.1038/350631a0>.
- (70) Jones, M. R.; Seeman, N. C.; Mirkin, C. A. Programmable Materials and the Nature of the DNA Bond. *Science* (80-. ). **2015**, *347* (6224), 1260901–1260901. <https://doi.org/10.1126/science.1260901>.
- (71) Linko, V.; Dietz, H. The Enabled State of DNA Nanotechnology. *Curr. Opin. Biotechnol.* **2013**, *24* (4), 555–561. <https://doi.org/10.1016/j.copbio.2013.02.001>.
- (72) Mergny, J.-L.; Sen, D. DNA Quadruple Helices in Nanotechnology. *Chem. Rev.* **2019**, *119*, acs.chemrev.8b00629. <https://doi.org/10.1021/acs.chemrev.8b00629>.
- (73) Zhuravel, R.; Stern, A.; Fardian-Melamed, N.; Eidelstein, G.; Katrivas, L.; Rotem, D.; Kotlyar, A. B.; Porath, D. Advances in Synthesis and Measurement of Charge Transport in DNA-Based Derivatives. *Adv. Mater.* **2018**, *30* (41), 1706984. <https://doi.org/10.1002/adma.201706984>.
- (74) Arimondo, P. B. Interaction of Human DNA Topoisomerase I with G-Quartet Structures. *Nucleic Acids Res.* **2000**, *28* (24), 4832–4838. <https://doi.org/10.1093/nar/28.24.4832>.
- (75) Cui, J.; Waltman, P.; Le, V. H.; Lewis, E. a. The Effect of Molecular Crowding on the Stability of Human C-MYC Promoter Sequence I-Motif at Neutral PH. *Molecules* **2013**, *18* (10), 12751–12767. <https://doi.org/10.3390/molecules181012751>.
- (76) Bhavsar-Jog, Y. P.; Van Dornshuld, E.; Brooks, T. A.; Tschumper, G. S.; Wadkins, R. M. Epigenetic Modification, Dehydration, and Molecular Crowding Effects on the Thermodynamics of i-Motif Structure Formation from C-Rich DNA. *Biochemistry* **2014**, *53* (10), 1586–1594. <https://doi.org/10.1021/bi401523b>.
- (77) Saxena, S.; Joshi, S.; Shankaraswamy, J.; Tyagi, S.; Kukreti, S. Magnesium and Molecular Crowding of the Cosolutes Stabilize the I-Motif Structure at Physiological PH. *Biopolymers* **2017**, *107* (7), e23018. <https://doi.org/10.1002/bip.23018>.
- (78) Alberti, P.; Bourdoncle, A.; Saccà, B.; Lacroix, L.; Mergny, J.-L. DNA Nanomachines and Nanostructures Involving Quadruplexes. *Org. Biomol.*

- 
- Chem.* **2006**, *4* (18), 3383. <https://doi.org/10.1039/b605739j>.
- (79) Martín-Hidalgo, M.; García-Arriaga, M.; González, F.; Rivera, J. M. Tuning Supramolecular G-Quadruplexes with Mono- and Divalent Cations. *Supramol. Chem.* **2015**, *27* (3), 174–180. <https://doi.org/10.1080/10610278.2014.924626>.
- (80) Day, H. A.; Huguin, C.; Waller, Z. A. E. Silver Cations Fold I-Motif at Neutral PH. *Chem. Commun.* **2013**, *49* (70), 7696. <https://doi.org/10.1039/c3cc43495h>.
- (81) Day, H. A.; Wright, E. P.; MacDonald, C. J.; Gates, A. J.; Waller, Z. A. E. Reversible DNA I-Motif to Hairpin Switching Induced by Copper(II) Cations. *Chem. Commun.* **2015**, *51* (74), 14099–14102. <https://doi.org/10.1039/C5CC05111H>.
- (82) Abdelhamid, M. A. S.; Fábíán, L.; MacDonald, C. J.; Cheesman, M. R.; Gates, A. J.; Waller, Z. A. E. Redox-Dependent Control of i-Motif DNA Structure Using Copper Cations. *Nucleic Acids Res.* **2018**, *46* (12), 5886–5893. <https://doi.org/10.1093/nar/gky390>.
- (83) Chen, C.; Pu, F.; Huang, Z.; Liu, Z.; Ren, J.; Qu, X. Stimuli-Responsive Controlled-Release System Using Quadruplex DNA-Capped Silica Nanocontainers. *Nucleic Acids Res.* **2011**, *39* (4), 1638–1644. <https://doi.org/10.1093/nar/gkq893>.
- (84) Kaushik, M.; Kaushik, S.; Roy, K.; Singh, A.; Mahendru, S.; Kumar, M.; Chaudhary, S.; Ahmed, S.; Kukreti, S. A Bouquet of DNA Structures: Emerging Diversity. *Biochem. Biophys. Reports* **2016**, *5*, 388–395. <https://doi.org/10.1016/j.bbrep.2016.01.013>.
- (85) Shi, Y.; Sun, H.; Xiang, J.; Yu, L.; Yang, Q.; Li, Q.; Guan, A.; Tang, Y. I-Motif-Modulated Fluorescence Detection of Silver(I) with an Ultrahigh Specificity. *Anal. Chim. Acta* **2015**, *857*, 79–84. <https://doi.org/10.1016/j.aca.2014.12.002>.
- (86) Surana, S.; Bhat, J. M.; Koushika, S. P.; Krishnan, Y. An Autonomous DNA Nanomachine Maps Spatiotemporal PH Changes in a Multicellular Living Organism. *Nat. Commun.* **2011**, *2* (May), 340. <https://doi.org/10.1038/ncomms1340>.
- (87) Wang, M.; Zhang, G.; Zhang, D. Enzyme-Driven i-Motif DNA Folding for Logic Operations and Fluorescent Biosensing. *Chem. Commun.* **2015**, *51*

- (18), 3812–3815. <https://doi.org/10.1039/C4CC09905B>.
- (88) Wang, Z.-G.; Elbaz, J.; Willner, I. DNA Machines: Bipedal Walker and Stepper. *Nano Lett.* **2011**, *11* (1), 304–309. <https://doi.org/10.1021/nl104088s>.
- (89) Amir, Y.; Ben-Ishay, E.; Levner, D.; Ittah, S.; Abu-Horowitz, A.; Bachelet, I. Universal Computing by DNA Origami Robots in a Living Animal. *Nat. Nanotechnol.* **2014**, *9* (5), 353–357. <https://doi.org/10.1038/nnano.2014.58>.
- (90) Yan, Y. Y.; Tan, J. H.; Lu, Y. J.; Yan, S. C.; Wong, K. Y.; Li, D.; Gu, L. Q.; Huang, Z. S. G-Quadruplex Conformational Change Driven by PH Variation with Potential Application as a Nanoswitch. *Biochim. Biophys. Acta - Gen. Subj.* **2013**, *1830* (10), 4935–4942. <https://doi.org/10.1016/j.bbagen.2013.06.019>.
- (91) Li, T.; Ackermann, D.; Hall, A. M.; Famulok, M. Input-Dependent Induction of Oligonucleotide Structural Motifs for Performing Molecular Logic. *J. Am. Chem. Soc.* **2012**, *134* (7), 3508–3516. <https://doi.org/10.1021/ja2108883>.
- (92) Dittmer, W. U.; Reuter, A.; Simmel, F. C. A DNA-Based Machine That Can Cyclically Bind and Release Thrombin. *Angew. Chemie Int. Ed.* **2004**, *43* (27), 3550–3553. <https://doi.org/10.1002/anie.200353537>.
- (93) Mergny, J.-L.; Lacroix, L. Analysis of Thermal Melting Curves. *Oligonucleotides* **2003**, *13* (6), 515–537. <https://doi.org/10.1089/154545703322860825>.
- (94) Mergny, J.-L.; Li, J.; Lacroix, L.; Amrane, S.; Chaires, J. B. Thermal Difference Spectra: A Specific Signature for Nucleic Acid Structures. *Nucleic Acids Res.* **2005**, *33* (16), e138. <https://doi.org/10.1093/nar/gni134>.
- (95) Vorlíčková, M.; Kejnovská, I.; Bednářová, K.; Renčiuk, D.; Kypr, J. Circular Dichroism Spectroscopy of DNA: From Duplexes to Quadruplexes. *Chirality* **2012**, *24* (9), 691–698. <https://doi.org/10.1002/chir.22064>.
- (96) Phan, A. T.; Guéron, M.; Leroy, J. L. The Solution Structure and Internal Motions of a Fragment of the Cytidine-Rich Strand of the Human Telomere. *J. Mol. Biol.* **2000**, *299* (1), 123–144. <https://doi.org/10.1006/jmbi.2000.3613>.
- (97) Dai, J.; Ambrus, A.; Hurley, L. H.; Yang, D. A Direct and Nondestructive Approach To Determine the Folding Structure of the I-Motif DNA Secondary Structure by NMR. *J. Am. Chem. Soc.* **2009**, *131* (17), 6102–6104. <https://doi.org/10.1021/ja900967r>.

- (98) Mergny, J.-L. Fluorescence Energy Transfer as a Probe for Tetraplex Formation: The i-Motif. *Biochemistry* **1999**, 38 (5), 1573–1581. <https://doi.org/10.1021/bi982208r>.
- (99) Sheng, Q.; Neaverson, J. C.; Mahmoud, T.; Stevenson, C. E. M.; Matthews, S. E.; Waller, Z. A. E. Identification of New DNA I-Motif Binding Ligands through a Fluorescent Intercalator Displacement Assay. *Org. Biomol. Chem.* **2017**, 24. <https://doi.org/10.1039/C7OB00710H>.
- (100) Pagano, A.; Iaccarino, N.; Abdelhamid, M. A. S.; Brancaccio, D.; Garzarella, E. U.; Di Porzio, A.; Novellino, E.; Waller, Z. A. E.; Pagano, B.; Amato, J.; et al. Common G-Quadruplex Binding Agents Found to Interact With i-Motif-Forming DNA: Unexpected Multi-Target-Directed Compounds. *Front. Chem.* **2018**, 6 (July), 1–13. <https://doi.org/10.3389/fchem.2018.00281>.
- (101) Abdelhamid, M. A. S.; Gates, A. J.; Waller, Z. A. E. Destabilization of I-Motif DNA at Neutral PH by G-Quadruplex Ligands. *Biochemistry* **2019**, 58 (4), 245–249. <https://doi.org/10.1021/acs.biochem.8b00968>.
- (102) Redman, J. E. Surface Plasmon Resonance for Probing Quadruplex Folding and Interactions with Proteins and Small Molecules. *Methods* **2007**, 43 (4), 302–312. <https://doi.org/10.1016/j.ymeth.2007.05.008>.
- (103) Stevens, A. J.; Kennedy, H. L.; Kennedy, M. A. Fluorescence Methods for Probing G-Quadruplex Structure in Single- and Double-Stranded DNA. *Biochemistry* **2016**, 55 (26), 3714–3725. <https://doi.org/10.1021/acs.biochem.6b00327>.
- (104) Cheatham, T. E.; Brooks, B. R.; Kollman, P. A. Molecular Modeling of Nucleic Acid Structure: Energy and Sampling. In *Current Protocols in Nucleic Acid Chemistry*; 2001; pp 7.8.1-7.8.20. <https://doi.org/10.1002/0471142700.nc0709s05>.
- (105) Bergonzo, C.; Galindo-Murillo, R.; Cheatham, T. E. Molecular Modeling of Nucleic Acid Structure: Electrostatics and Solvation. *Curr. Protoc. Nucleic Acid Chem.* **2013**, 55 (1), 7.9.1-7.9.27. <https://doi.org/10.1002/0471142700.nc0709s55>.
- (106) Sponer, J.; Spacková, N. Molecular Dynamics Simulations and Their Application to Four-Stranded DNA. *Methods* **2007**, 43 (4), 278–290. <https://doi.org/10.1016/j.ymeth.2007.02.004>.
- (107) Ivani, I.; Dans, P. D.; Noy, A.; Pérez, A.; Faustino, I.; Hospital, A.; Walther,



- 
- J.; Andrio, P.; Goñi, R.; Balaceanu, A.; et al. Parmbsc1: A Refined Force Field for DNA Simulations. *Nat. Methods* **2016**, *13* (1), 55–58. <https://doi.org/10.1038/nmeth.3658>.
- (108) Panczyk, T.; Wolski, P. Molecular Dynamics Analysis of Stabilities of the Telomeric Watson-Crick Duplex and the Associated i-Motif as a Function of PH and Temperature. *Biophys. Chem.* **2018**, *237* (February), 22–30. <https://doi.org/10.1016/j.bpc.2018.03.006>.
- (109) Monsen, R. C.; Trent, J. O. G-Quadruplex Virtual Drug Screening: A Review. *Biochimie* **2018**, *152*, 134–148. <https://doi.org/10.1016/j.biochi.2018.06.024>.
- (110) Xu, H.; Di Antonio, M.; McKinney, S.; Mathew, V.; Ho, B.; O’Neil, N. J.; Santos, N. Dos; Silvester, J.; Wei, V.; Garcia, J.; et al. CX-5461 Is a DNA G-Quadruplex Stabilizer with Selective Lethality in BRCA1/2 Deficient Tumours. *Nat. Commun.* **2017**, *8* (1), 14432. <https://doi.org/10.1038/ncomms14432>.
- (111) Liu, D.; Balasubramanian, S. A Proton-Fuelled DNA Nanomachine. *Angew. Chemie Int. Ed.* **2003**, *42* (46), 5734–5736. <https://doi.org/10.1002/anie.200352402>.
- (112) Liu, H.; Xu, Y.; Li, F.; Yang, Y.; Wang, W.; Song, Y.; Liu, D. Light-Driven Conformational Switch of i-Motif DNA. *Angew. Chemie - Int. Ed.* **2007**, *46* (14), 2515–2517. <https://doi.org/10.1002/anie.200604589>.
- (113) Modi, S.; M. G., S.; Goswami, D.; Gupta, G. D.; Mayor, S.; Krishnan, Y. A DNA Nanomachine That Maps Spatial and Temporal PH Changes inside Living Cells. *Nat. Nanotechnol.* **2009**, *4* (5), 325–330. <https://doi.org/10.1038/nnano.2009.83>.
- (114) Dhakal, S.; Schonhoft, J. D.; Koirala, D.; Yu, Z.; Basu, S.; Mao, H. Coexistence of an ILPR I-Motif and a Partially Folded Structure with Comparable Mechanical Stability Revealed at the Single-Molecule Level. *J. Am. Chem. Soc.* **2010**, *132* (26), 8991–8997. <https://doi.org/10.1021/ja100944j>.
- (115) Choi, J.; Kim, S.; Tachikawa, T.; Fujitsuka, M.; Majima, T. PH-Induced Intramolecular Folding Dynamics of i-Motif DNA. *J. Am. Chem. Soc.* **2011**, *133* (40), 16146–16153. <https://doi.org/10.1021/ja2061984>.
- (116) Lieblein, A. L.; Buck, J.; Schlepckow, K.; Fürtig, B.; Schwalbe, H. Time-Resolved NMR Spectroscopic Studies of DNA i-Motif Folding Reveal Kinetic

- Partitioning. *Angew. Chemie Int. Ed.* **2012**, *51* (1), 250–253. <https://doi.org/10.1002/anie.201104938>.
- (117) Gurung, S. P.; Schwarz, C.; Hall, J. P.; Cardin, C. J.; Brazier, J. A. The Importance of Loop Length on the Stability of I-Motif Structures. *Chem. Commun. (Camb.)* **2015**, *51* (26), 5630–5632. <https://doi.org/10.1039/c4cc07279k>.
- (118) McKim, M.; Buxton, A.; Johnson, C.; Metz, A.; Sheardy, R. D. Loop Sequence Context Influences the Formation and Stability of the I-Motif for DNA Oligomers of Sequence (CCCXXX)<sub>4</sub>, Where X = A and/or T, under Slightly Acidic Conditions. *J. Phys. Chem. B* **2016**, *120* (31), 7652–7661. <https://doi.org/10.1021/acs.jpcc.6b04561>.
- (119) Lieblein, A. L.; Krämer, M.; Dreuw, A.; Fürtig, B.; Schwalbe, H. The Nature of Hydrogen Bonds in Cytidine...H+...Cytidine DNA Base Pairs. *Angew. Chemie Int. Ed.* **2012**, *51* (17), 4067–4070. <https://doi.org/10.1002/anie.201200549>.
- (120) Kendrick, S.; Kang, H.-J.; Alam, M. P.; Madathil, M. M.; Agrawal, P.; Gokhale, V.; Yang, D.; Hecht, S. M.; Hurley, L. H. The Dynamic Character of the BCL2 Promoter I-Motif Provides a Mechanism for Modulation of Gene Expression by Compounds That Bind Selectively to the Alternative DNA Hairpin Structure. *J. Am. Chem. Soc.* **2014**, *136* (11), 4161–4171. <https://doi.org/10.1021/ja410934b>.
- (121) Barragán, E.; Gordillo, B.; Vargas, G.; Velazco, L. The Role of Cobalt, Copper, Nickel, and Zinc in the DNA Replication Inhibitory Activity Ofp-Aminophenyl Triphenylporphyrin. *Appl. Organomet. Chem.* **2004**, *18* (7), 311–317. <https://doi.org/10.1002/aoc.649>.
- (122) Clever, G. H.; Shionoya, M. Metal–Base Pairing in DNA. *Coord. Chem. Rev.* **2010**, *254* (19–20), 2391–2402. <https://doi.org/10.1016/j.ccr.2010.04.014>.
- (123) Clever, G. H.; Kaul, C.; Carell, T. DNA-Metal Base Pairs. *Angew. Chemie Int. Ed.* **2007**, *46* (33), 6226–6236. <https://doi.org/10.1002/anie.200701185>.
- (124) Waller, Z. A. E.; Pinchbeck, B. J.; Buguth, B. S.; Meadows, T. G.; Richardson, D. J.; Gates, A. J. Control of Bacterial Nitrate Assimilation by Stabilization of G-Quadruplex DNA. *Chem. Commun.* **2016**, *52* (92), 13511–13514. <https://doi.org/10.1039/C6CC06057A>.
- (125) Howell, L. A.; Waller, Z. A. E.; Bowater, R.; O'Connell, M.; Searcey, M. A

- 
- Small Molecule That Induces Assembly of a Four Way DNA Junction at Low Temperature. *Chem. Commun.* **2011**, 47 (29), 8262. <https://doi.org/10.1039/c1cc12922h>.
- (126) Waller, Z. A. E.; Howell, L. A.; Macdonald, C. J.; O'Connell, M. A.; Searcey, M. Identification and Characterisation of a G-Quadruplex Forming Sequence in the Promoter Region of Nuclear Factor (Erythroid-Derived 2)-like 2 (Nrf2). *Biochem. Biophys. Res. Commun.* **2014**, 447 (1), 128–132. <https://doi.org/10.1016/j.bbrc.2014.03.117>.
- (127) Müller, J. Functional Metal Ions in Nucleic Acids. *Metallomics* **2010**, 2 (5), 318–327. <https://doi.org/10.1039/c000429d>.
- (128) Hsiao, C.; Tannenbaum, M.; VanDeusen, H.; Hershkovitz, E.; Perng, G.; Tannenbaum, A.; Williams, L. D. *Complexes of Nucleic Acids with Group I and II Cations*; 2008.
- (129) Kankia, B. I.; Marky, L. A. Folding of the Thrombin Aptamer into a G-Quadruplex with Sr<sup>2+</sup>: Stability, Heat, and Hydration. *J. Am. Chem. Soc.* **2001**, 123 (44), 10799–10804. <https://doi.org/10.1021/ja010008o>.
- (130) Guédin, A.; Gros, J.; Alberti, P.; Mergny, J.-L. How Long Is Too Long? Effects of Loop Size on G-Quadruplex Stability. *Nucleic Acids Res.* **2010**, 38 (21), 7858–7868. <https://doi.org/10.1093/nar/gkq639>.
- (131) Bhattacharyya, D.; Mirihana Arachchilage, G.; Basu, S. Metal Cations in G-Quadruplex Folding and Stability. *Front. Chem.* **2016**, 4 (September), 1–14. <https://doi.org/10.3389/fchem.2016.00038>.
- (132) Kim, S. E.; Lee, I. B.; Hyeon, C.; Hong, S. C. Destabilization of I-Motif by Submolar Concentrations of a Monovalent Cation. *J. Phys. Chem. B* **2014**, 118 (18), 4753–4760. <https://doi.org/10.1021/jp500120d>.
- (133) Wright, E. P.; Lamparska, K.; Smith, S. S.; Waller, Z. A. E. Substitution of Cytosine with Guanylurea Decreases the Stability of I-Motif DNA. *Biochemistry* **2017**, 56 (36), 4879–4883. <https://doi.org/10.1021/acs.biochem.7b00628>.
- (134) Zhou, J.; Wei, C.; Jia, G.; Wang, X.; Feng, Z.; Li, C. Formation of I-Motif Structure at Neutral and Slightly Alkaline PH. *Mol. BioSyst.* **2010**, 6 (3), 580–586. <https://doi.org/10.1039/B919600E>.
- (135) Cheng, E.; Xing, Y.; Chen, P.; Yang, Sun, Y.; Zhou, D.; Xu, T.; Fan, Q.; Liu, D. A PH-Triggered, Fast-Responding DNA Hydrogel. *Angew. Chemie - Int.*

- 
- Ed.* **2009**, *48* (41), 7660–7663. <https://doi.org/10.1002/anie.200902538>.
- (136) Li, J.; Mo, L.; Lu, C.-H.; Fu, T.; Yang, H.-H.; Tan, W. Functional Nucleic Acid-Based Hydrogels for Bioanalytical and Biomedical Applications. *Chem. Soc. Rev.* **2016**, *45* (5), 1410–1431. <https://doi.org/10.1039/C5CS00586H>.
- (137) Liu, J. Oligonucleotide-Functionalized Hydrogels as Stimuli Responsive Materials and Biosensors. *Soft Matter* **2011**, *7* (15), 6757. <https://doi.org/10.1039/c1sm05284e>.
- (138) Gao, J.; Berden, G.; Rodgers, M. T.; Oomens, J. Interaction of Cu<sup>+</sup> with Cytosine and Formation of I-Motif-like C–M + –C Complexes: Alkali versus Coinage Metals. *Phys. Chem. Chem. Phys.* **2016**, *18* (10), 7269–7277. <https://doi.org/10.1039/C6CP00234J>.
- (139) Yuan, X.; Pham, A. N.; Xing, G.; Rose, A. L.; Waite, T. D. Effects of PH, Chloride, and Bicarbonate on Cu(I) Oxidation Kinetics at Circumneutral PH. *Environ. Sci. Technol.* **2012**, *46* (3), 1527–1535. <https://doi.org/10.1021/es203394k>.
- (140) Jhaveri, A. S.; Sharma, M. M. Kinetics of Absorption of Oxygen in Aqueous Solutions of Cuprous Chloride. *Chem. Eng. Sci.* **1967**, *22* (1), 1–6. [https://doi.org/10.1016/0009-2509\(67\)80097-6](https://doi.org/10.1016/0009-2509(67)80097-6).
- (141) Mergny, J.-L.; Lacroix, L. UV Melting of G-Quadruplexes. In *Current Protocols in Nucleic Acid Chemistry*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2009; pp 17.1.1-17.1.15. <https://doi.org/10.1002/0471142700.nc1701s37>.
- (142) Benabou, S.; Aviñó, A.; Eritja, R.; González, C.; Gargallo, R. Fundamental Aspects of the Nucleic Acid I-Motif Structures. *RSC Adv.* **2014**, *4* (51), 26956. <https://doi.org/10.1039/c4ra02129k>.
- (143) Kypr, J.; Kejnovska, I.; Rencik, D.; Vorlickova, M. Circular Dichroism and Conformational Polymorphism of DNA. *Nucleic Acids Res.* **2009**, *37* (6), 1713–1725. <https://doi.org/10.1093/nar/gkp026>.
- (144) Valiev, M.; Bylaska, E. J.; Govind, N.; Kowalski, K.; Straatsma, T. P.; Van Dam, H. J. J.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T. L.; et al. NWChem: A Comprehensive and Scalable Open-Source Solution for Large Scale Molecular Simulations. *Comput. Phys. Commun.* **2010**, *181* (9), 1477–1489. <https://doi.org/10.1016/j.cpc.2010.04.018>.
- (145) Smith, D. W. Ionic Hydration Enthalpies. *J. Chem. Educ.* **1977**, *54* (9), 540.

---

<https://doi.org/10.1021/ed054p540>.

- (146) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. *J. Cheminform.* **2012**, *4* (1), 17. <https://doi.org/10.1186/1758-2946-4-17>.
- (147) Řezáč, J.; Hobza, P. Advanced Corrections of Hydrogen Bonding and Dispersion for Semiempirical Quantum Mechanical Methods. *J. Chem. Theory Comput.* **2012**, *8* (1), 141–151. <https://doi.org/10.1021/ct200751e>.
- (148) Stewart, J. J. P. MOPAC 2016. Stewart Computational Chemistry, Colorado Springs, CO, USA 2016.
- (149) Liang, L.; Astruc, D. The Copper(I)-Catalyzed Alkyne-Azide Cycloaddition (CuAAC) “Click” Reaction and Its Applications. An Overview. *Coord. Chem. Rev.* **2011**, *255* (23–24), 2933–2945. <https://doi.org/10.1016/j.ccr.2011.06.028>.
- (150) Kohn, K. W. Beyond DNA Cross-Linking: History and Prospects of DNA-Targeted Cancer Treatment-Fifteenth Bruce F. Cain Memorial Award Lecture. *Cancer Res.* **1996**, *56* (24), 5533–5546.
- (151) Biffi, G.; Tannahill, D.; McCafferty, J.; Balasubramanian, S. Quantitative Visualization of DNA G-Quadruplex Structures in Human Cells. *Nat. Chem.* **2013**, *5* (3), 182–186. <https://doi.org/10.1038/nchem.1548>.
- (152) Wright, E. P.; Day, H. A.; Ibrahim, A. M.; Kumar, J.; Boswell, L. J. E.; Huguin, C.; Stevenson, C. E. M.; Pors, K.; Waller, Z. A. E. Mitoxantrone and Analogues Bind and Stabilize I-Motif Forming DNA Sequences. *Sci. Rep.* **2016**, *6* (1), 39456. <https://doi.org/10.1038/srep39456>.
- (153) Shu, B.; Cao, J.; Kuang, G.; Qiu, J.; Zhang, M.; Zhang, Y.; Wang, M.; Li, X.; Kang, S.; Ou, T.-M.; et al. Syntheses and Evaluation of New Acridone Derivatives for Selective Binding of Oncogene C- Myc Promoter i-Motifs in Gene Transcriptional Regulation. *Chem. Commun.* **2018**, *54* (16), 2036–2039. <https://doi.org/10.1039/C8CC00328A>.
- (154) Michael Green, N. Avidin and Streptavidin. In *Methods in Enzymology*; 1990; Vol. 184, pp 51–67. [https://doi.org/10.1016/0076-6879\(90\)84259-J](https://doi.org/10.1016/0076-6879(90)84259-J).
- (155) Franceschin, M.; Rossetti, L.; D’Ambrosio, A.; Schirripa, S.; Bianco, A.; Ortaggi, G.; Savino, M.; Schultes, C.; Neidle, S. Natural and Synthetic G-Quadruplex Interactive Berberine Derivatives. *Bioorg. Med. Chem. Lett.*

- 
- 2006**, *16* (6), 1707–1711. <https://doi.org/10.1016/j.bmcl.2005.12.001>.
- (156) Gowan, S. M.; Harrison, J. R.; Patterson, L.; Valenti, M.; Read, M. A.; Neidle, S.; Kelland, L. R. A G-Quadruplex-Interactive Potent Small-Molecule Inhibitor of Telomerase Exhibiting in Vitro and in Vivo Antitumor Activity. *Mol. Pharmacol.* **2002**, *61* (5), 1154–1162. <https://doi.org/10.1124/mol.61.5.1154>.
- (157) Huang, H.-S.; Chen, I.-B.; Huang, K.-F.; Lu, W.-C.; Shieh, F.-Y.; Huang, Y.-Y.; Huang, F.-C.; Lin, J.-J. Synthesis and Human Telomerase Inhibition of a Series of Regioisomeric Disubstituted Amidoanthraquinones. *Chem. Pharm. Bull. (Tokyo)*. **2007**, *55* (2), 284–292. <https://doi.org/10.1248/cpb.55.284>.
- (158) De Cian, A.; DeLemos, E.; Mergny, J.-L.; Teulade-Fichou, M.-P.; Monchaud, D. Highly Efficient G-Quadruplex Recognition by Bisquinolinium Compounds. *J. Am. Chem. Soc.* **2007**, *129* (7), 1856–1857. <https://doi.org/10.1021/ja067352b>.
- (159) Rodriguez, R.; Müller, S.; Yeoman, J. A.; Trentesaux, C.; Riou, J.-F.; Balasubramanian, S. A Novel Small Molecule That Alters Shelterin Integrity and Triggers a DNA-Damage Response at Telomeres. *J. Am. Chem. Soc.* **2008**, *130* (47), 15758–15759. <https://doi.org/10.1021/ja805615w>.
- (160) Izbicka, E.; Wheelhouse, R. T.; Raymond, E.; Davidson, K. K.; Lawrence, R. A.; Sun, D.; Windle, B. E.; Hurley, L. H.; Von Hoff, D. D. Effects of Cationic Porphyrins as G-Quadruplex Interactive Agents in Human Tumor Cells. *Cancer Res.* **1999**, *59* (3), 639–644. <https://doi.org/10.1126/science.7605428>.
- (161) Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H. Direct Evidence for a G-Quadruplex in a Promoter Region and Its Targeting with a Small Molecule to Repress c-MYC Transcription. *Proc. Natl. Acad. Sci.* **2002**, *99* (18), 11593–11598. <https://doi.org/10.1073/pnas.182256799>.
- (162) Chen, C.; Li, M.; Xing, Y.; Li, Y.; Joedecke, C.-C.; Jin, J.; Yang, Z.; Liu, D. Study of PH-Induced Folding and Unfolding Kinetics of the DNA i-Motif by Stopped-Flow Circular Dichroism. *Langmuir* **2012**, *28* (51), 17743–17748. <https://doi.org/10.1021/la303851a>.
- (163) Li, W.; Miyoshi, D.; Nakano, S.; Sugimoto, N. Structural Competition Involving G-Quadruplex DNA and Its Complement. *Biochemistry* **2003**, *42* (40), 11736–11744. <https://doi.org/10.1021/bi034168j>.
- (164) Goodman, R. P.; Schaap, I. A. T.; Tardin, C. F.; Erben, C. M.; Berry, R. M.;

- 
- Schmidt, C. F.; Turberfield, A. J. Rapid Chiral Assembly of Rigid DNA Building Blocks for Molecular Nanofabrication. *Science* **2005**, *310* (5754), 1661–1665. <https://doi.org/10.1126/science.1120367>.
- (165) Goodman, R. P.; Berry, R. M.; Turberfield, A. J. The Single-Step Synthesis of a DNA Tetrahedron. *Chem. Commun.* **2004**, *44* (12), 1372. <https://doi.org/10.1039/b402293a>.
- (166) Zadeh, J. N.; Steenberg, C. D.; Bois, J. S.; Wolfe, B. R.; Pierce, M. B.; Khan, A. R.; Dirks, R. M.; Pierce, N. A. NUPACK: Analysis and Design of Nucleic Acid Systems. *J. Comput. Chem.* **2011**, *32* (1), 170–173. <https://doi.org/10.1002/jcc.21596>.
- (167) Goodman, R. P. NANEV: A Program Employing Evolutionary Methods for the Design of Nucleic Acid Nanostructures. *Biotechniques* **2005**, *38* (4), 548–550. <https://doi.org/10.2144/05384BM06>.
- (168) Dirks, R. M.; Pierce, N. A. A Partition Function Algorithm for Nucleic Acid Secondary Structure Including Pseudoknots. *J. Comput. Chem.* **2003**, *24* (13), 1664–1677. <https://doi.org/10.1002/jcc.10296>.
- (169) Dirks, R. M.; Pierce, N. A. An Algorithm for Computing Nucleic Acid Base-Pairing Probabilities Including Pseudoknots. *J. Comput. Chem.* **2004**, *25* (10), 1295–1304. <https://doi.org/10.1002/jcc.20057>.
- (170) Dirks, R. M.; Bois, J. S.; Schaeffer, J. M.; Winfree, E.; Pierce, N. A. Thermodynamic Analysis of Interacting Nucleic Acid Strands. *SIAM Rev.* **2007**, *49* (1), 65–88. <https://doi.org/10.1137/060651100>.
- (171) Zadeh, J. N.; Wolfe, B. R.; Pierce, N. A. Nucleic Acid Sequence Design via Efficient Ensemble Defect Optimization. *J. Comput. Chem.* **2011**, *32* (3), 439–452. <https://doi.org/10.1002/jcc.21633>.
- (172) Dirks, R. M.; Lin, M.; Winfree, E.; Pierce, N. A. Paradigms for Computational Nucleic Acid Design. *Nucleic Acids Res.* **2004**, *32* (4), 1392–1403. <https://doi.org/10.1093/nar/gkh291>.
- (173) Wolfe, B. R.; Pierce, N. A. Sequence Design for a Test Tube of Interacting Nucleic Acid Strands. *ACS Synth. Biol.* **2015**, *4* (10), 1086–1100. <https://doi.org/10.1021/sb5002196>.
- (174) Wolfe, B. R.; Porubsky, N. J.; Zadeh, J. N.; Dirks, R. M.; Pierce, N. A. Constrained Multistate Sequence Design for Nucleic Acid Reaction Pathway Engineering. *J. Am. Chem. Soc.* **2017**, *139* (8), 3134–3144.

---

<https://doi.org/10.1021/jacs.6b12693>.

- (175) Pinheiro, A. V.; Han, D.; Shih, W. M.; Yan, H. Challenges and Opportunities for Structural DNA Nanotechnology. *Nat. Nanotechnol.* **2011**, *6* (12), 763–772. <https://doi.org/10.1038/nnano.2011.187>.
- (176) Sadowski, J. P.; Calvert, C. R.; Zhang, D. Y.; Pierce, N. A.; Yin, P. Developmental Self-Assembly of a DNA Tetrahedron. *ACS Nano* **2014**, *8* (4), 3251–3259. <https://doi.org/10.1021/nn4038223>.
- (177) Goodman, R. P.; Heilemann, M.; Doose, S.; Erben, C. M.; Kapanidis, A. N.; Turberfield, A. J. Reconfigurable, Braced, Three-Dimensional DNA Nanostructures. *Nat. Nanotechnol.* **2008**, *3* (2), 93–96. <https://doi.org/10.1038/nnano.2008.3>.
- (178) He, Y.; Ye, T.; Su, M.; Zhang, C.; Ribbe, A. E.; Jiang, W.; Mao, C. Hierarchical Self-Assembly of DNA into Symmetric Supramolecular Polyhedra. *Nature* **2008**, *452* (7184), 198–201. <https://doi.org/10.1038/nature06597>.
- (179) Zhang, C.; Su, M.; He, Y.; Zhao, X.; Fang, P. -a.; Ribbe, A. E.; Jiang, W.; Mao, C. Conformational Flexibility Facilitates Self-Assembly of Complex DNA Nanostructures. *Proc. Natl. Acad. Sci.* **2008**, *105* (31), 10665–10669. <https://doi.org/10.1073/pnas.0803841105>.
- (180) Li, Z.; Wei, B.; Nangreave, J.; Lin, C.; Liu, Y.; Mi, Y.; Yan, H. A Replicable Tetrahedral Nanostructure Self-Assembled from a Single DNA Strand. *J. Am. Chem. Soc.* **2009**, *131* (36), 13093–13098. <https://doi.org/10.1021/ja903768f>.
- (181) Xing, S.; Jiang, D.; Li, F.; Li, J.; Li, Q.; Huang, Q.; Guo, L.; Xia, J.; Shi, J.; Fan, C.; et al. Constructing Higher-Order DNA Nanoarchitectures with Highly Purified DNA Nanocages. *ACS Appl. Mater. Interfaces* **2015**, *7* (24), 13174–13179. <https://doi.org/10.1021/am505592e>.
- (182) Banerjee, A.; Bhatia, D.; Saminathan, A.; Chakraborty, S.; Kar, S.; Krishnan, Y. Controlled Release of Encapsulated Cargo from a DNA Icosahedron Using a Chemical Trigger. *Angew. Chemie Int. Ed.* **2013**, *52* (27), 6854–6857. <https://doi.org/10.1002/anie.201302759>.
- (183) Erben, C. M.; Goodman, R. P.; Turberfield, A. J. Single-Molecule Protein Encapsulation in a Rigid DNA Cage. *Angew. Chemie Int. Ed.* **2006**, *45* (44), 7414–7417. <https://doi.org/10.1002/anie.200603392>.



- 
- (184) Liang, L.; Li, J.; Li, Q.; Huang, Q.; Shi, J.; Yan, H.; Fan, C. Single-Particle Tracking and Modulation of Cell Entry Pathways of a Tetrahedral DNA Nanostructure in Live Cells. *Angew. Chemie Int. Ed.* **2014**, *53* (30), 7745–7750. <https://doi.org/10.1002/anie.201403236>.
- (185) Jiang, D.; Sun, Y.; Li, J.; Li, Q.; Lv, M.; Zhu, B.; Tian, T.; Cheng, D.; Xia, J.; Zhang, L.; et al. Multiple-Armed Tetrahedral DNA Nanostructures for Tumor-Targeting, Dual-Modality in Vivo Imaging. *ACS Appl. Mater. Interfaces* **2016**, *8* (7), 4378–4384. <https://doi.org/10.1021/acsami.5b10792>.
- (186) Pei, H.; Zuo, X.; Zhu, D.; Huang, Q.; Fan, C. Functional DNA Nanostructures for Theranostic Applications. *Acc. Chem. Res.* **2014**, *47* (2), 550–559. <https://doi.org/10.1021/ar400195t>.
- (187) Zhang, F.; Nangreave, J.; Liu, Y.; Yan, H. Structural DNA Nanotechnology: State of the Art and Future Perspective. *J. Am. Chem. Soc.* **2014**, *136* (32), 11198–11211. <https://doi.org/10.1021/ja505101a>.
- (188) Chang, S.; Kilic, T.; Lee, C.; Avci, H.; Bae, H.; Oskui, S.; Jung, S.; Shin, S.; Kim, S. Reversible Redox Activity by Ion-PH Dually Modulated Duplex Formation of i-Motif DNA with Complementary G-DNA. *Nanomaterials* **2018**, *8* (4), 226. <https://doi.org/10.3390/nano8040226>.
- (189) Lu, C.-H.; Ceconello, A.; Willner, I. Recent Advances in the Synthesis and Functions of Reconfigurable Interlocked DNA Nanostructures. *J. Am. Chem. Soc.* **2016**, *138* (16), 5172–5185. <https://doi.org/10.1021/jacs.6b00694>.
- (190) Wang, P.; Xia, Z.; Yan, J.; Liu, X.; Yao, G.; Pei, H.; Zuo, X.; Sun, G.; He, D. A Study of PH-Dependence of Shrink and Stretch of Tetrahedral DNA Nanostructures. *Nanoscale* **2015**, *7* (15), 6467–6470. <https://doi.org/10.1039/C5NR00757G>.
- (191) Gray, R. D.; Chaires, J. B. Kinetics and Mechanism of K<sup>+</sup>- and Na<sup>+</sup>-Induced Folding of Models of Human Telomeric DNA into G-Quadruplex Structures. *Nucleic Acids Res.* **2008**, *36* (12), 4191–4203. <https://doi.org/10.1093/nar/gkn379>.
- (192) Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. Quadruplex DNA: Sequence, Topology and Structure. *Nucleic Acids Res.* **2006**, *34* (19), 5402–5415. <https://doi.org/10.1093/nar/gkl655>.
- (193) Sun, D.; Hurley, L. H. Biochemical Techniques for the Characterization of G-Quadruplex Structures: EMSA, DMS Footprinting, and DNA Polymerase

- 
- Stop Assay. In *G-Quadruplex DNA: Methods and Protocols*; Baumann, P., Ed.; Humana Press: Berlin/Heidelberg, 2010; Vol. 608, pp 65–79. [https://doi.org/10.1007/978-1-59745-363-9\\_5](https://doi.org/10.1007/978-1-59745-363-9_5).
- (194) Stetefeld, J.; McKenna, S. A.; Patel, T. R. Dynamic Light Scattering: A Practical Guide and Applications in Biomedical Sciences. *Biophys. Rev.* **2016**, *8* (4), 409–427. <https://doi.org/10.1007/s12551-016-0218-6>.
- (195) Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of  $\alpha/\gamma$  Conformers. *Biophys. J.* **2007**, *92* (11), 3817–3829. <https://doi.org/10.1529/biophysj.106.097782>.
- (196) Wolski, P.; Nieszporek, K.; Panczyk, T. G-Quadruplex and I-Motif Structures within the Telomeric DNA Duplex. A Molecular Dynamics Analysis of Protonation States as Factors Affecting Their Stability. *J. Phys. Chem. B* **2019**, *123* (2), 468–479. <https://doi.org/10.1021/acs.jpccb.8b11547>.
- (197) Haider, S.; Neidle, S. Molecular Modeling and Simulation of G-Quadruplexes and Quadruplex-Ligand Complexes. In *G-Quadruplex DNA: Methods and Protocols*; Baumann, P., Ed.; Humana Press: Totowa, NJ, 2010; pp 17–37. [https://doi.org/10.1007/978-1-59745-363-9\\_2](https://doi.org/10.1007/978-1-59745-363-9_2).
- (198) Galindo-Murillo, R.; Robertson, J. C.; Zgarbová, M.; Šponer, J.; Otyepka, M.; Jurečka, P.; Cheatham, T. E. Assessing the Current State of Amber Force Field Modifications for DNA. *J. Chem. Theory Comput.* **2016**, *12* (8), 4114–4127. <https://doi.org/10.1021/acs.jctc.6b00186>.
- (199) Yang, C.; Kulkarni, M.; Lim, M.; Pak, Y. In Silico Direct Folding of Thrombin-Binding Aptamer G-Quadruplex at All-Atom Level. *Nucleic Acids Res.* **2017**, *45* (22), 12648–12656. <https://doi.org/10.1093/nar/gkx1079>.
- (200) König, S. L. B.; Huppert, J. L.; Sigel, R. K. O.; Evans, A. C. Distance-Dependent Duplex DNA Destabilization Proximal to G-Quadruplex/ i -Motif Sequences. *Nucleic Acids Res.* **2013**, *41* (15), 7453–7461. <https://doi.org/10.1093/nar/gkt476>.
- (201) Sekibo, D. A. T.; Fox, K. R. The Effects of DNA Supercoiling on G-Quadruplex Formation. *Nucleic Acids Res.* **2017**, *45* (21), 12069–12079. <https://doi.org/10.1093/nar/gkx856>.
- (202) Arora, A.; Nair, D. R.; Maiti, S. Effect of Flanking Bases on Quadruplex Stability and Watson-Crick Duplex Competition. *FEBS J.* **2009**, *276* (13),

- 3628–3640. <https://doi.org/10.1111/j.1742-4658.2009.07082.x>.
- (203) Kumar, N. The Effect of Osmolytes and Small Molecule on Quadruplex-WC Duplex Equilibrium: A Fluorescence Resonance Energy Transfer Study. *Nucleic Acids Res.* **2005**, *33* (21), 6723–6732. <https://doi.org/10.1093/nar/gki961>.
- (204) Hardin, C. C.; Perry, A. G.; White, K. Thermodynamic and Kinetic Characterization of the Dissociation and Assembly of Quadruplex Nucleic Acids. *Biopolymers* **2000**, *56* (3), 147–194. [https://doi.org/10.1002/1097-0282\(2000/2001\)56:3<147::AID-BIP10011>3.0.CO;2-N](https://doi.org/10.1002/1097-0282(2000/2001)56:3<147::AID-BIP10011>3.0.CO;2-N).
- (205) del Villar-Guerra, R.; Trent, J. O.; Chaires, J. B. G-Quadruplex Secondary Structure Obtained from Circular Dichroism Spectroscopy. *Angew. Chemie Int. Ed.* **2018**, *57* (24), 7171–7175. <https://doi.org/10.1002/anie.201709184>.
- (206) Kato, T.; Goodman, R. P.; Erben, C. M.; Turberfield, A. J.; Namba, K. High-Resolution Structural Analysis of a DNA Nanostructure by CryoEM. *Nano Lett.* **2009**, *9* (7), 2747–2750. <https://doi.org/10.1021/nl901265n>.
- (207) Le Guilloux, V.; Schmidtke, P.; Tuffery, P. Fpocket: An Open Source Platform for Ligand Pocket Detection. *BMC Bioinformatics* **2009**, *10* (1), 168. <https://doi.org/10.1186/1471-2105-10-168>.
- (208) Vukovic, S.; Huggins, D. J. Quantitative Metrics for Drug–Target Ligandability. *Drug Discov. Today* **2018**, *23* (6), 1258–1266. <https://doi.org/10.1016/j.drudis.2018.02.015>.
- (209) Case, D. A.; Cerutti, D. S.; III, T. E. C.; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Greene, D.; Homeyer, N.; et al. AMBER 2017. University of California, San Francisco 2017.
- (210) Krepl, M.; Zgarbová, M.; Stadlbauer, P.; Otyepka, M.; Banáš, P.; Koča, J.; Cheatham, T. E.; Jurečka, P.; Šponer, J. Reference Simulations of Noncanonical Nucleic Acids with Different  $\chi$  Variants of the AMBER Force Field: Quadruplex DNA, Quadruplex RNA, and Z-DNA. *J. Chem. Theory Comput.* **2012**, *8* (7), 2506–2520. <https://doi.org/10.1021/ct300275s>.
- (211) Zgarbová, M.; Luque, F. J.; Šponer, J.; Cheatham, T. E.; Otyepka, M.; Jurečka, P. Toward Improved Description of DNA Backbone: Revisiting Epsilon and Zeta Torsion Force Field Parameters. *J. Chem. Theory Comput.* **2013**, *9* (5), 2339–2354. <https://doi.org/10.1021/ct400154j>.
- (212) Zgarbová, M.; Šponer, J.; Otyepka, M.; Cheatham, T. E.; Galindo-Murillo, R.;

- 
- Jurečka, P. Refinement of the Sugar–Phosphate Backbone Torsion Beta for AMBER Force Fields Improves the Description of Z- and B-DNA. *J. Chem. Theory Comput.* **2015**, *11* (12), 5723–5736. <https://doi.org/10.1021/acs.jctc.5b00716>.
- (213) Li, P.; Song, L. F.; Merz, K. M. Systematic Parameterization of Monovalent Ions Employing the Nonbonded Model. *J. Chem. Theory Comput.* **2015**, *11* (4), 1645–1657. <https://doi.org/10.1021/ct500918t>.
- (214) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79* (2), 926–935. <https://doi.org/10.1063/1.445869>.
- (215) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindah, E. Gromacs: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25. <https://doi.org/10.1016/j.softx.2015.06.001>.
- (216) Li, X. `amber_to_gmx_by_XL` [https://github.com/recoli/amber\\_to\\_gmx\\_by\\_XL](https://github.com/recoli/amber_to_gmx_by_XL) (accessed Aug 4, 2017).

---

# Appendix

## A.1 FRET-melting Data for Initial Ligand Identification

Sample Information		NSC
		CAS
$\Delta T_m$ (°C)	i-motif	(CCG) <sub>16</sub> pH 7.4
		cMyc pH 6.6
		Hif1- $\alpha$ pH 7.2 (50uM ligand)
		Hif1- $\alpha$ pH 7.4 (20uM ligand)
		hTeloC pH 6.0 (a)
		hTeloC pH 6.0 (b)
		hTeloC pH 7.2
		hTeloC pH 7.4 (a)
		hTeloC pH 7.4 (b)
		ILPR pH 6.0
		ILPR pH 7.2
		G-quadruplex
	hTeloG pH 7.4 (a)	
	hTeloG pH 7.4 (b)	
	Double stranded	dsDNA pH 7.2
		dsDNA pH 7.4 (a)
		dsDNA pH 7.4 (b)

9037	13051	13487	30205
3569-82-2	5427-55-4	5431-11-8	x
x	11.22	0.00	23.00
x	15.83	x	10.61
x	14.28	x	16.28
x	13.22	x	x
-1.47	35.78, 25.44	-0.72	x
x	25.33	x	2.00
0.00	11.17	0.00	x
x	15.33	x	11.67
x	22.67	x	x
-14.67	32.50	1.66	-1.00
-32.67	-2.50	-32.67	x
-21.23	19.33	-21.23	x
x	1.67	x	4.33
x	14.20	x	x
-0.80	0.11	1.53	x
x	-0.33	x	2.67
x	-3.33	x	4.33

Sample Information		NSC
		CAS
$\Delta T_m$ (°C)	i-motif	(CCG) <sub>16</sub> pH 7.4
		cMyc pH 6.6
		Hif1- $\alpha$ pH 7.2 (50uM ligand)
		Hif1- $\alpha$ pH 7.4 (20uM ligand)
		hTeloC pH 6.0 (a)
		hTeloC pH 6.0 (b)
		hTeloC pH 7.2
		hTeloC pH 7.4 (a)
		hTeloC pH 7.4 (b)
		ILPR pH 6.0
		ILPR pH 7.2
	G-quadruplex	hTeloG pH 7.2
		hTeloG pH 7.4 (a)
		hTeloG pH 7.4 (b)
	Double stranded	dsDNA pH 7.2
		dsDNA pH 7.4 (a)
		dsDNA pH 7.4 (b)

33353	35676	60339	71795
x	569-77-7	70-09-7	519-23-3
31.67	x	30.22	25.67
34.28	4.17	3.39	17.28
32.28	x	53.28	9.28
x	x	x	x
41.48	x	10.57	23.13
50.00	27.33	12.33	28.67
41.42, 50.60, 54.75	x	22.80	16.33
x	-1.33	18.67	13.67
x	x	x	x
41.67	x	7.50	21.83
5.98, 20.92, 28.45	x	1.08	-11.75
37.18	x	21.57	5.03
33.33	4.33	20.33	2.67
x	x	x	x
20.33	x	6.70	4.80
3.33	0.00	4.67	2.33
28.67	x	7.67	3.67

Sample Information		NSC
		CAS
$\Delta T_m$ (°C)	i-motif	(CCG) <sub>16</sub> pH 7.4
		cMyc pH 6.6
		Hif1- $\alpha$ pH 7.2 (50uM ligand)
		Hif1- $\alpha$ pH 7.4 (20uM ligand)
		hTeloC pH 6.0 (a)
		hTeloC pH 6.0 (b)
		hTeloC pH 7.2
		hTeloC pH 7.4 (a)
		hTeloC pH 7.4 (b)
		ILPR pH 6.0
		ILPR pH 7.2
	G-quadruplex	hTeloG pH 7.2
		hTeloG pH 7.4 (a)
		hTeloG pH 7.4 (b)
	Double stranded	dsDNA pH 7.2
		dsDNA pH 7.4 (a)
		dsDNA pH 7.4 (b)

103520	109086	111847	143491	146771
18814-03-4	x	x	34610-60-1	50440-30-7
0.33	11.89	0.33	38.56	53.33
x	-8.94	x	20.61	x
20.28	x	x	40.28	53.28
x	2.33	x	x	x
0.89	-0.33	2.55	17.35	-13.80
x	40.67	x	38.67	x
0.00	x	0.00	26.33	0.00
x	12.67	x	25.33	x
x	4.67	x	x	x
11.00	0.00	0.33	19.50	51.67
<5	x	9.88	4.67	-32.67
2.67	x	15.22	20.37	42.63
x	10.67	x	22.67	x
x	7.87	x	x	x
0.22	x	2.82	13.50	27.72
x	1.00	x	11.00	x
x	0.67	x	13.00	36.00



Sample Information		NSC
		CAS
$\Delta T_m$ (°C)	i-motif	(CCG) <sub>16</sub> pH 7.4
		cMyc pH 6.6
		Hif1- $\alpha$ pH 7.2 (50uM ligand)
		Hif1- $\alpha$ pH 7.4 (20uM ligand)
		hTeloC pH 6.0 (a)
		hTeloC pH 6.0 (b)
		hTeloC pH 7.2
		hTeloC pH 7.4 (a)
		hTeloC pH 7.4 (b)
		ILPR pH 6.0
		ILPR pH 7.2
		G-quadruplex
	hTeloG pH 7.4 (a)	
	hTeloG pH 7.4 (b)	
	Double stranded	dsDNA pH 7.2
		dsDNA pH 7.4 (a)
		dsDNA pH 7.4 (b)

177365	202386	204232	215718
63345-17-5	53221-75-3	x	x
29.78	7.67	1.33	x
x	x	x	x
53.28	53.28	37.28	18.28
x	3.22	x	x
10.52	2.00, 17.22	-0.62	x
x	x	x	x
21.67	8.17	0.00	x
x	x	x	x
x	15.75	x	x
1.17	6.83	-0.89	x
-7.42	<5	-19.13, 0.85, 30.88	x
8.23	2.89	6.68, 10.32, 39.00	x
x	x	x	x
x	6.47	x	x
5.57	0.67	3.62	x
x	x	x	x
8.33	0.00	x	x

Sample Information		NSC
		CAS
$\Delta T_m$ (°C)	i-motif	(CCG) <sub>16</sub> pH 7.4
		cMyc pH 6.6
		Hif1- $\alpha$ pH 7.2 (50uM ligand)
		Hif1- $\alpha$ pH 7.4 (20uM ligand)
		hTeloC pH 6.0 (a)
		hTeloC pH 6.0 (b)
		hTeloC pH 7.2
		hTeloC pH 7.4 (a)
		hTeloC pH 7.4 (b)
		ILPR pH 6.0
		ILPR pH 7.2
		G-quadruplex
	hTeloG pH 7.4 (a)	
	hTeloG pH 7.4 (b)	
	Double stranded	dsDNA pH 7.2
		dsDNA pH 7.4 (a)
		dsDNA pH 7.4 (b)

260594	275428	277184	300289
x	x	55078-51-8	54824-20-3
53.33	14.44	14.00	33.89
23.94	x	x	x
53.28	x	20.28	15.28
x	0.00	0.56	x
16.18	0.44	0.44	x
22.67	x	x	x
52.78	x	x	x
60.67	x	x	x
x	1.17	4.67	x
18.67	0.67	2.33	-1.00
27.05	x	x	x
44.87	x	x	x
50.00	x	x	x
x	1.73	14.47	x
23.70	1.56	x	x
26.67	x	x	x
28.67	0.33	1.33	7.67

Sample Information		NSC
		CAS
$\Delta T_m$ (°C)	i-motif	(CCG) <sub>16</sub> pH 7.4
		cMyc pH 6.6
		Hif1- $\alpha$ pH 7.2 (50uM ligand)
		Hif1- $\alpha$ pH 7.4 (20uM ligand)
		hTeloC pH 6.0 (a)
		hTeloC pH 6.0 (b)
		hTeloC pH 7.2
		hTeloC pH 7.4 (a)
		hTeloC pH 7.4 (b)
		ILPR pH 6.0
		ILPR pH 7.2
		G-quadruplex
	hTeloG pH 7.4 (a)	
	hTeloG pH 7.4 (b)	
	Double stranded	dsDNA pH 7.2
		dsDNA pH 7.4 (a)
		dsDNA pH 7.4 (b)

305780	308848	308849	309892
x	69408-82-8	69408-83-9	x
4.78	17.89	22.22	x
18.72	x	x	x
6.28	13.28	11.28	46.28
x	10.33	x	x
35.78	-0.33	x	x
24.33	x	x	x
8.83	x	x	x
12.33	x	x	x
x	4.61	x	x
11.56	1.67	0.00	x
-3.50	x	x	x
16.56	x	x	x
11.33	x	x	x
x	2.78	x	x
1.11	x	x	x
0.33	x	x	x
1.00	0.67	5.67	x

Sample Information		NSC
		CAS
$\Delta T_m$ (°C)	i-motif	(CCG) <sub>16</sub> pH 7.4
		cMyc pH 6.6
		Hif1- $\alpha$ pH 7.2 (50uM ligand)
		Hif1- $\alpha$ pH 7.4 (20uM ligand)
		hTeloC pH 6.0 (a)
		hTeloC pH 6.0 (b)
		hTeloC pH 7.2
		hTeloC pH 7.4 (a)
		hTeloC pH 7.4 (b)
		ILPR pH 6.0
		ILPR pH 7.2
		G-quadruplex
	hTeloG pH 7.4 (a)	
	hTeloG pH 7.4 (b)	
	Double stranded	dsDNA pH 7.2
		dsDNA pH 7.4 (a)
		dsDNA pH 7.4 (b)

311153	317003	317605	325014
81531-60-4	80568-29-2	34374-22-6	72615-20-4
39.22	27.33	30.78	x
14.39	10.06	2.50	x
53.28	16.28	30.28	53.28
x	x	x	x
30.72	17.33	11.00	x
39.33	13.89	4.33	x
34.33	3.17	5.83	x
26.00	13.00	13.00	x
x	x	x	x
34.89	13.78	10.33	x
7.83	-1.50	0.00	x
30.67	18.11	18.11	x
30.67	16.00	15.00	x
x	x	x	x
12.62	4.22	9.11	x
8.67	4.00	7.67	x
12.00	5.67	9.67	x

Sample Information		NSC
		CAS
$\Delta T_m$ (°C)	i-motif	(CCG) <sub>16</sub> pH 7.4
		cMyc pH 6.6
		Hif1- $\alpha$ pH 7.2 (50uM ligand)
		Hif1- $\alpha$ pH 7.4 (20uM ligand)
		hTeloC pH 6.0 (a)
		hTeloC pH 6.0 (b)
		hTeloC pH 7.2
		hTeloC pH 7.4 (a)
		hTeloC pH 7.4 (b)
		ILPR pH 6.0
		ILPR pH 7.2
		G-quadruplex
	hTeloG pH 7.4 (a)	
	hTeloG pH 7.4 (b)	
	Double stranded	dsDNA pH 7.2
		dsDNA pH 7.4 (a)
		dsDNA pH 7.4 (b)

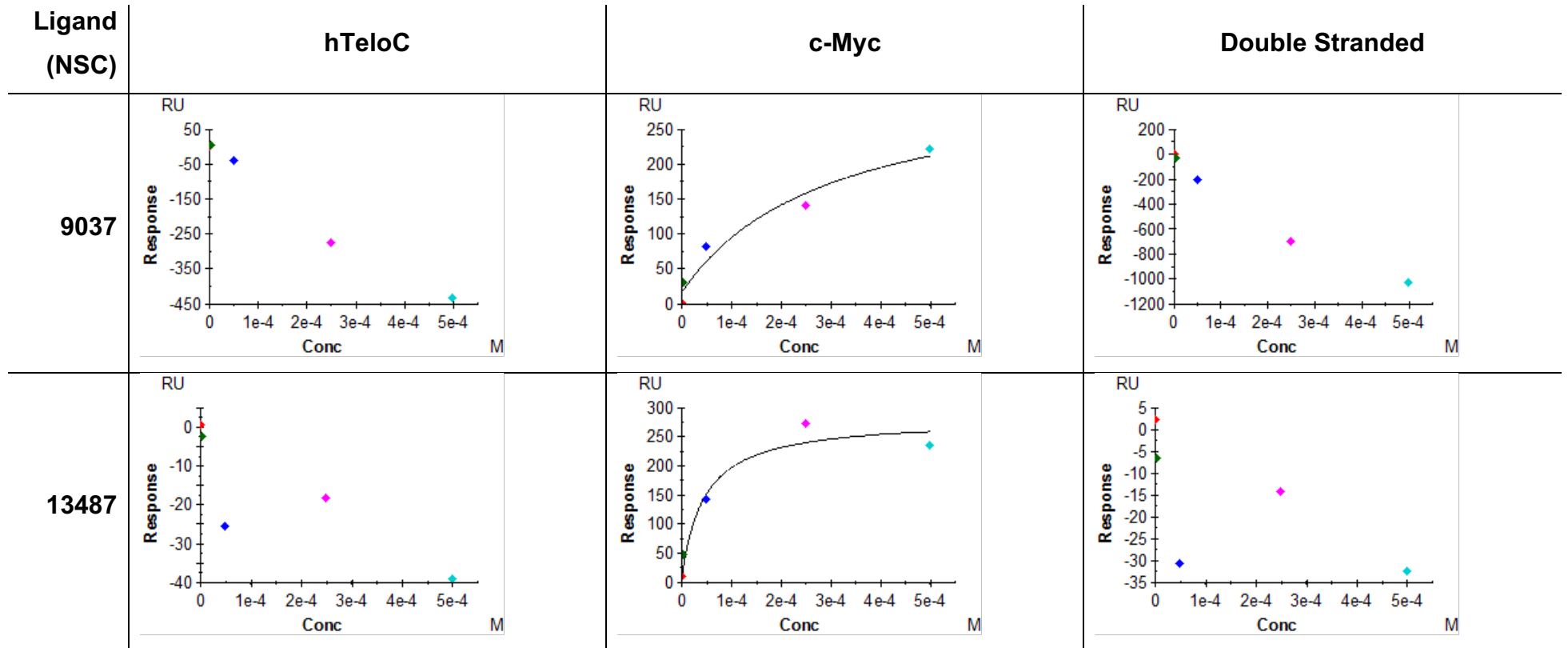
345647	345845	638432
x	71993-16-3	x
x	16.67	18.33
8.72	-0.17	x
x	2.28	4.28
x	1.00	x
x	0.78	16.55, 45.00
1.33	1.00	x
x	x	8.00
43.67	15.00	x
x	10.08	63.33
x	0.67	15.67
x	x	-12.87
x	x	10.50
0.67	1.00	x
x	4.00	7.60
x	1.22	-0.52
-0.33	1.00	x
x	1.00	0.33

## A.2 %R<sub>max</sub> from SPR Screen at pH 5.5

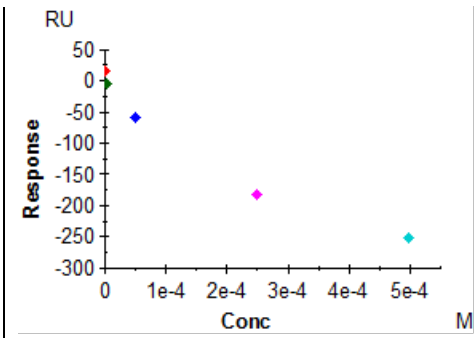
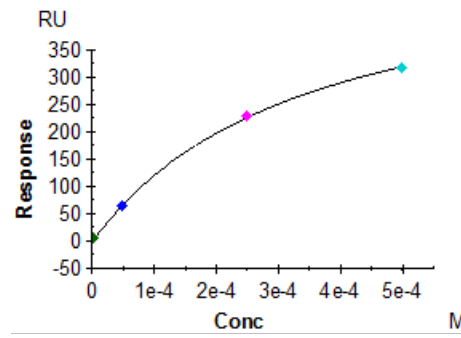
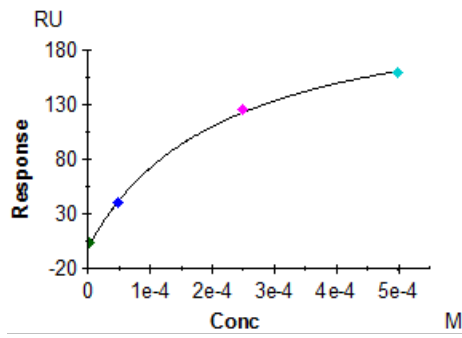
Ligand (NSC)	%R <sub>max</sub> (%)		
	hTeloC	c-Myc	DS
9037	-43.91	210.65	-1498.58
13051	16.79	97.93	189.11
13487	-211	172.11	-603.44
30205	231.36	323.51	624.74
33353	1948.85	3098.77	5984.21
35676	87.15	142.48	275.16
60339	353.2	545.27	-510.2
71795	3036.61	3745.53	7233.2
103520	-65.43	42.43	-119.46
109086	587.92	576.28	1112.88
111847	-43.12	17.78	-321.95
143491	368.49	373.99	-53.13
143969	115.99	120.78	233.25
144046	-0.37	35.78	69.1
145895	37.82	70.93	136.97
146771	-124.82	-10.88	-2255.31
177365	16.69	21.02	40.59
180514	325.34	419.93	810.96
202386	-1417.88	-2419.92	-4673.24
204232	-42.03	21.79	-42.08
215718	-240.09	28.69	-624.01
260594	294.89	344.74	-427.53
275428	5.56	-32.27	-162.93
277184	7.96	-8.07	-185.45
300289	12.14	-11.56	-176.37
301739	2125.38	2534.4	4894.33
305780	105.85	132.24	255.38
305798	-6.71	-19.08	-166.75
308848	145.19	100.75	-165.54
308849	26.3	-1.64	-255.06
309892	-47.78	-25.12	-166.51
311153	308.63	746.38	-2637.1
317003	263.98	425.61	-695.47
317605	385.06	617.58	-202.54
325014	27.79	19.81	-7.01
345647	-7.9	34.02	65.71
345845	186.9	281.84	-325.77
638432	260.97	420.62	-557.13

Table A.2.1 %R<sub>max</sub> calculated for each ligand-DNA pair; colour coded scale transitioning from green to red indicating lowest to highest %R<sub>max</sub> for each DNA structure respectively. [ligand] = 50 μM Running buffer: 10 mM sodium cacodylate, 100 mM NaCl, 0.05% tween at pH 5.5.

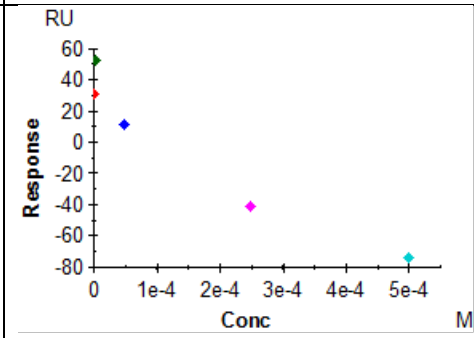
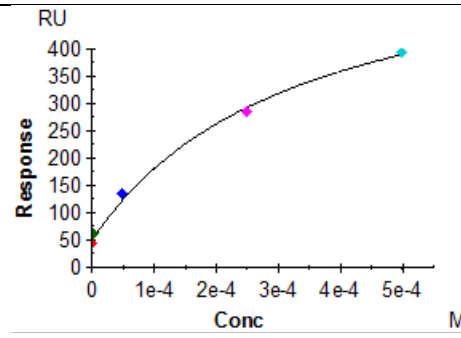
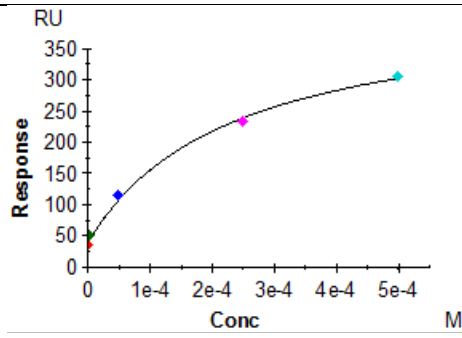
### A.3 Response vs Concentration Plots from SPR at pH 5.5



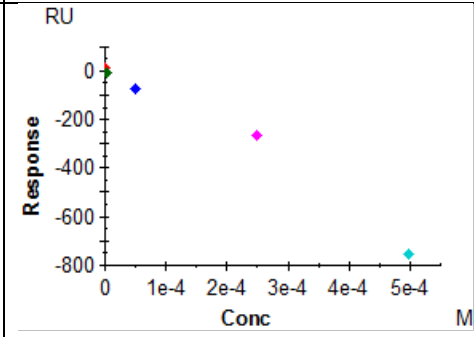
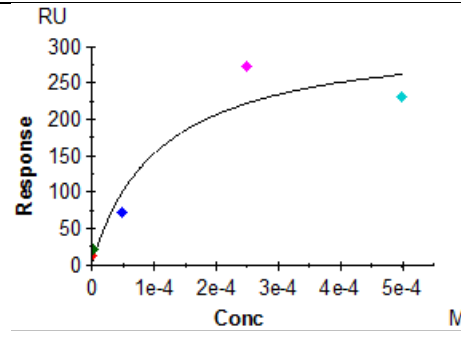
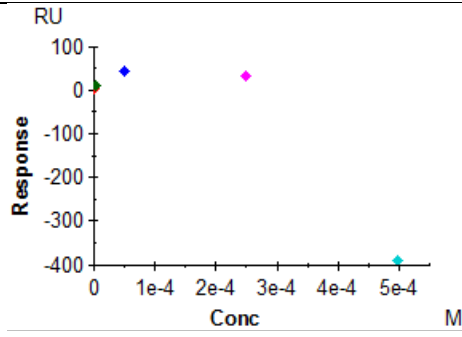
60339



143491

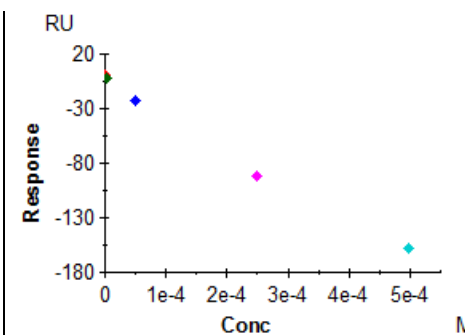
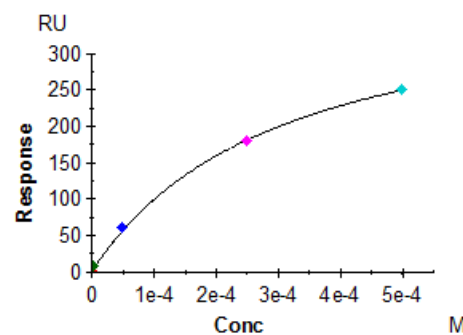
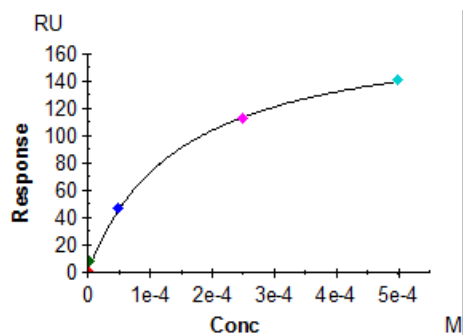


260594

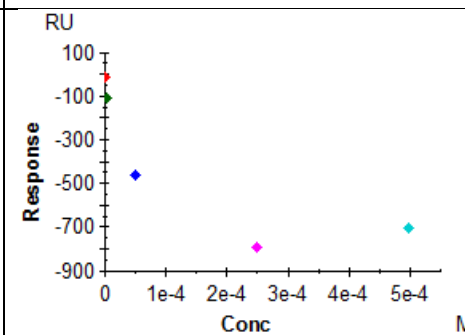
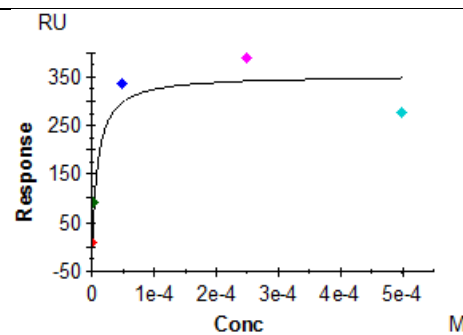
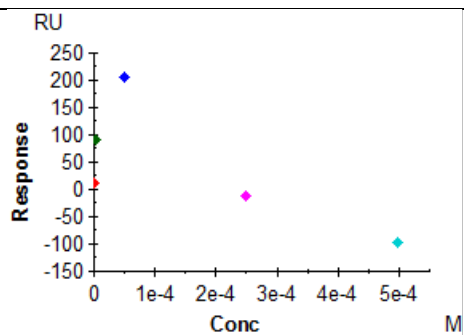




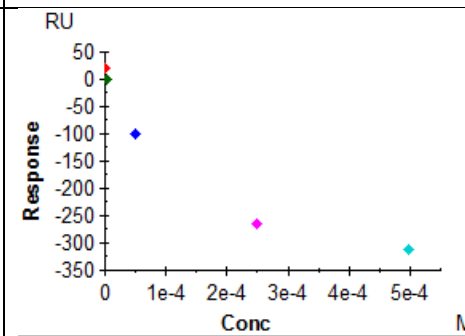
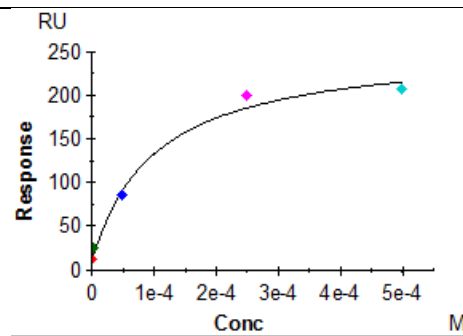
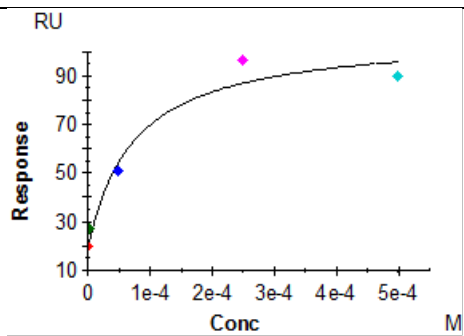
308848



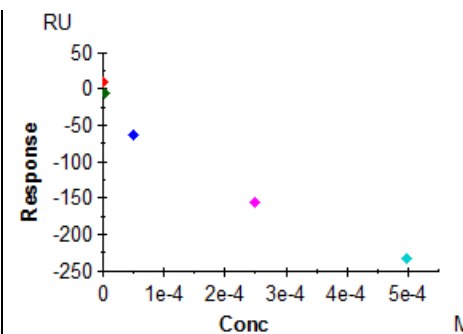
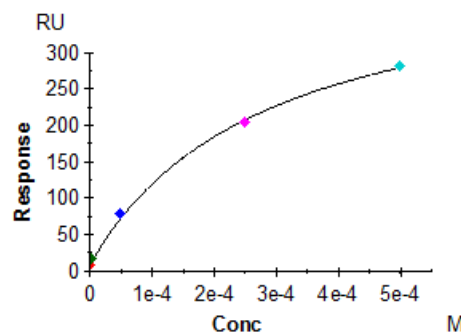
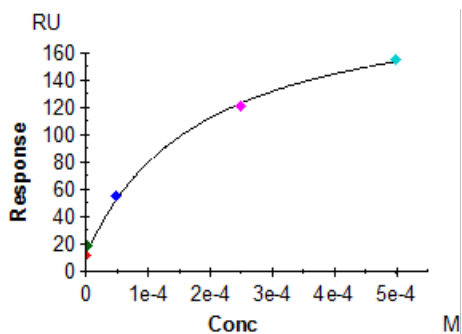
311153



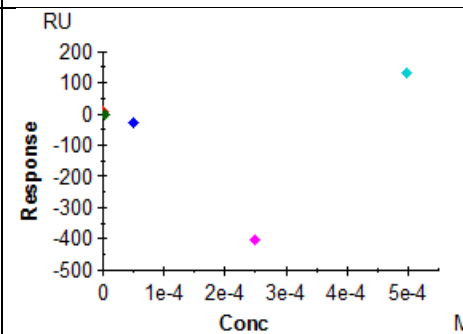
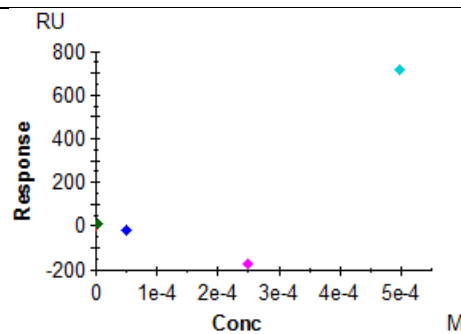
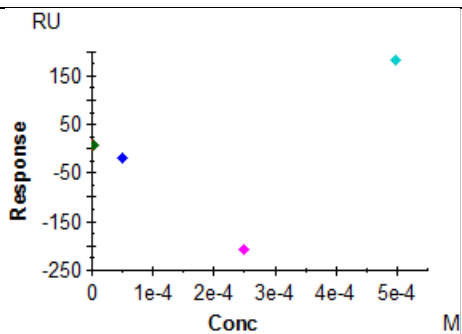
317003



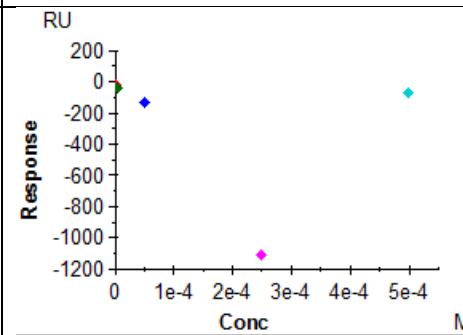
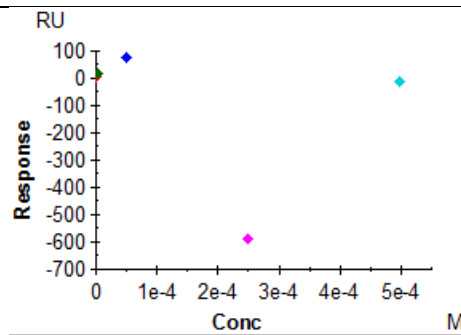
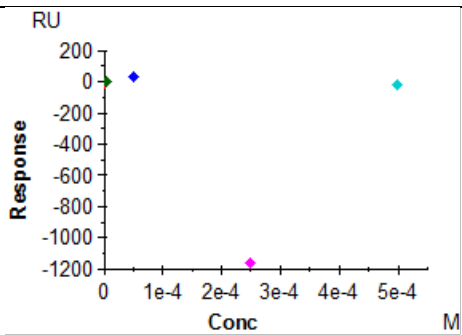
317605



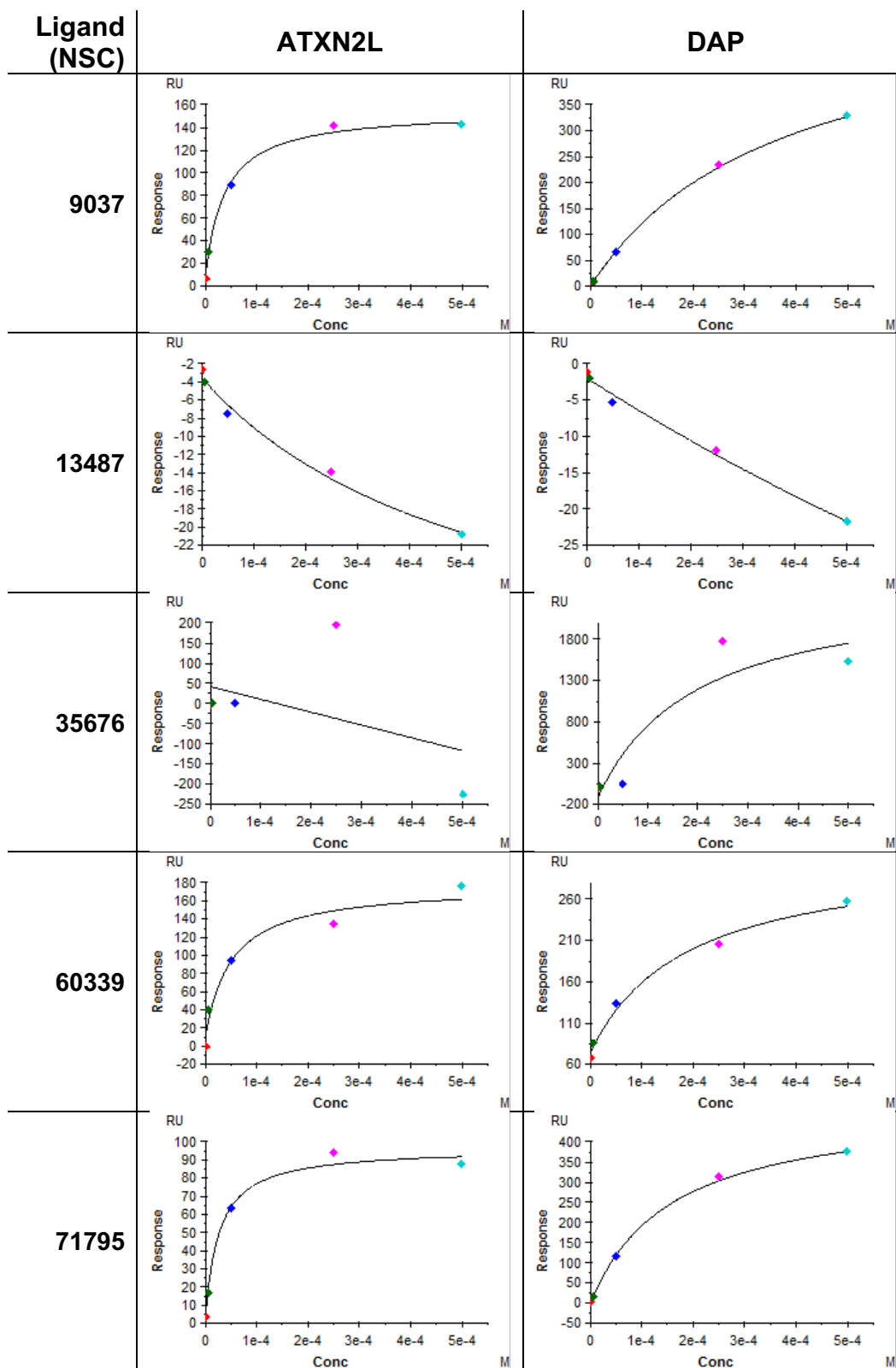
345845

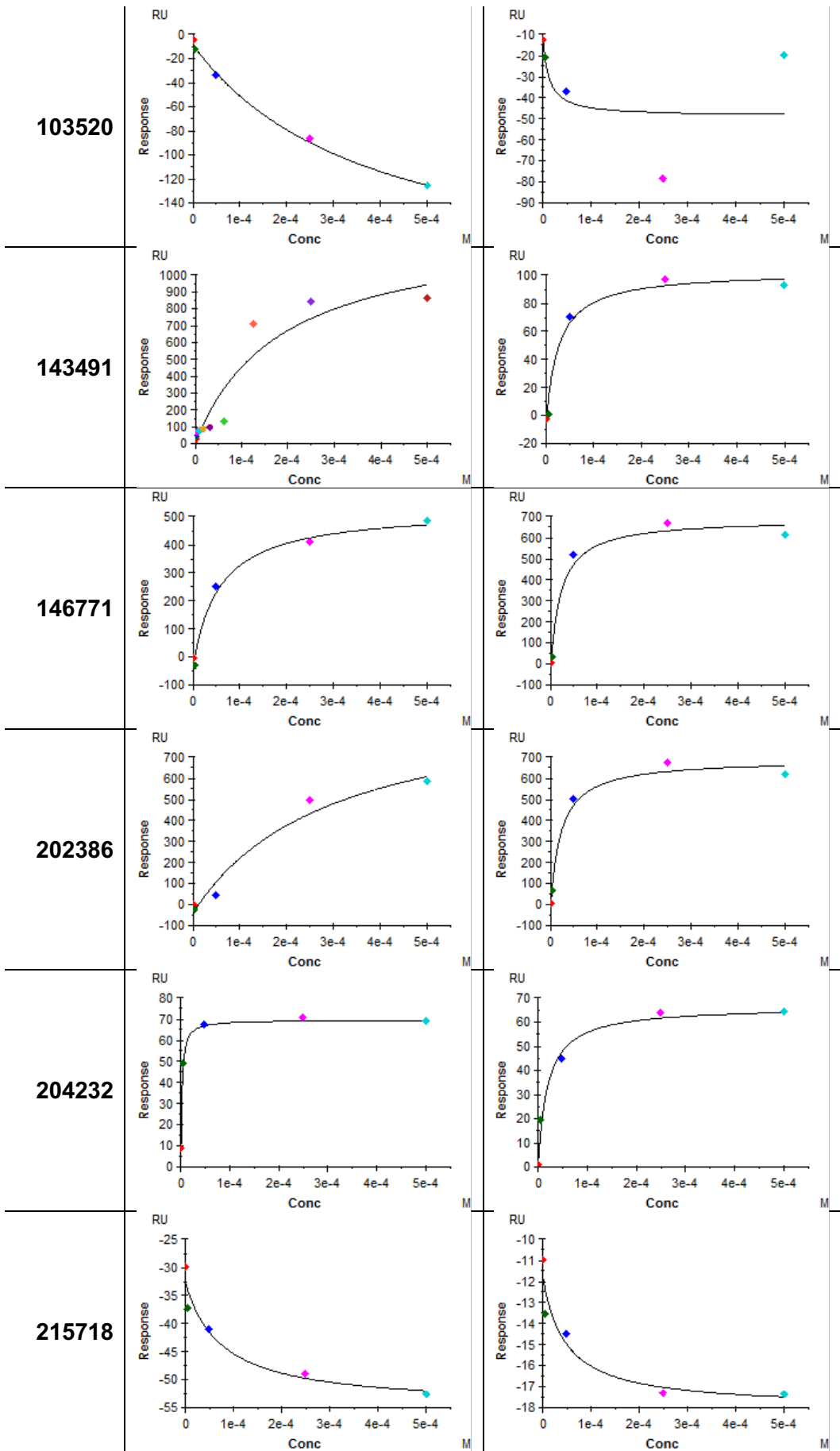


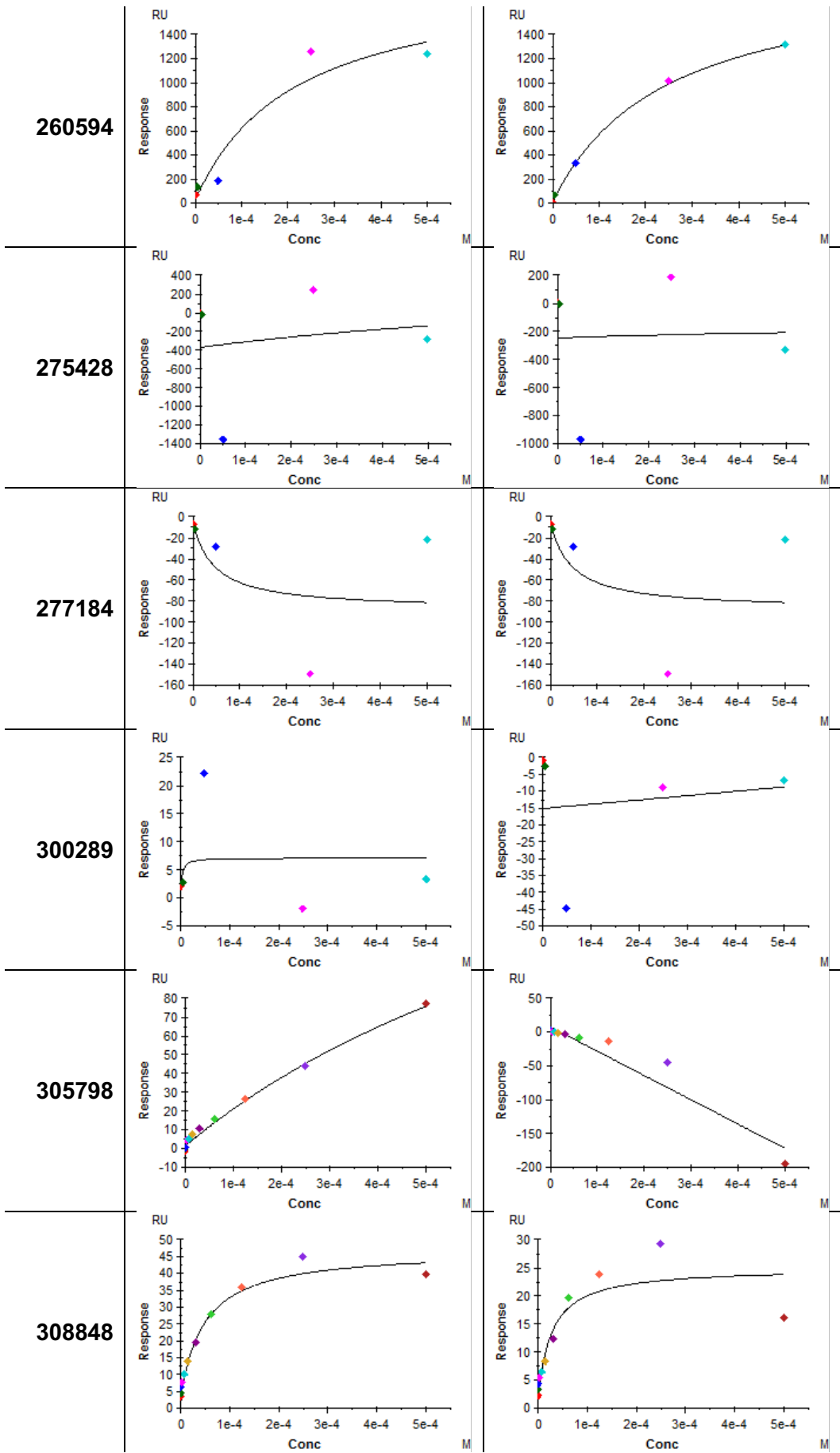
638432

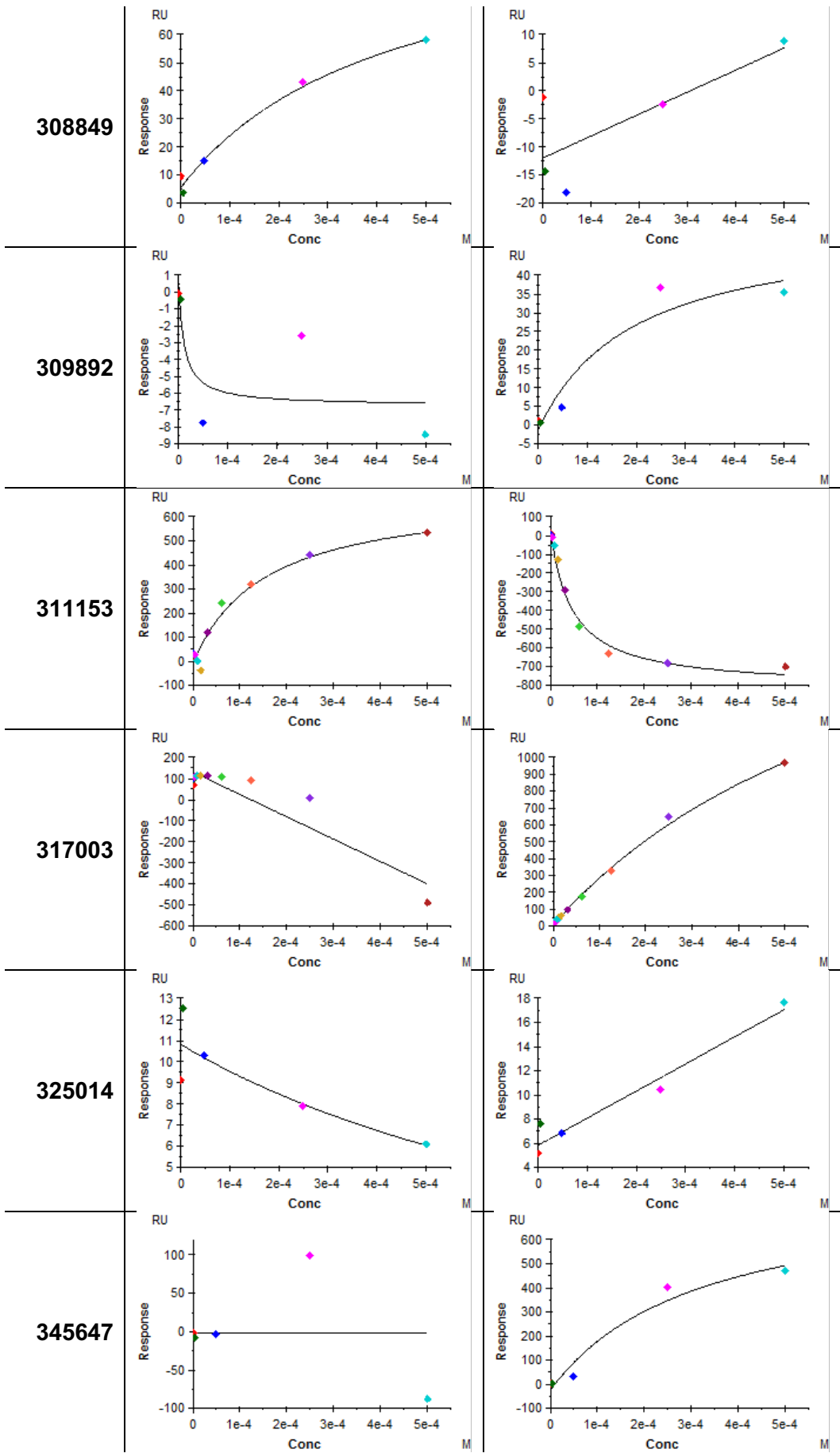


## A.4 Response vs Concentration Plots from SPR at pH 7.0

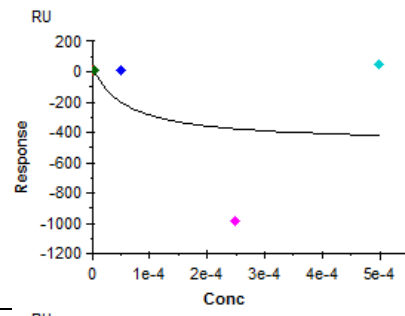
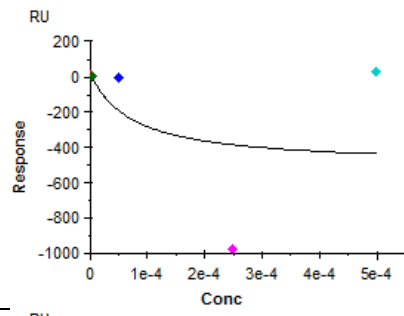




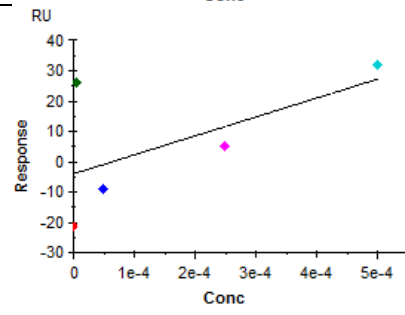
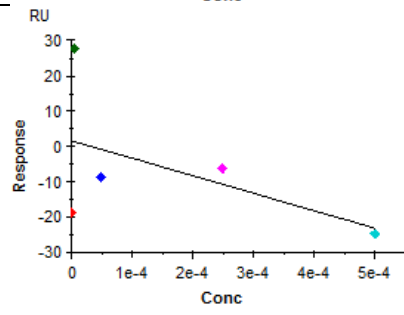




345845



638432

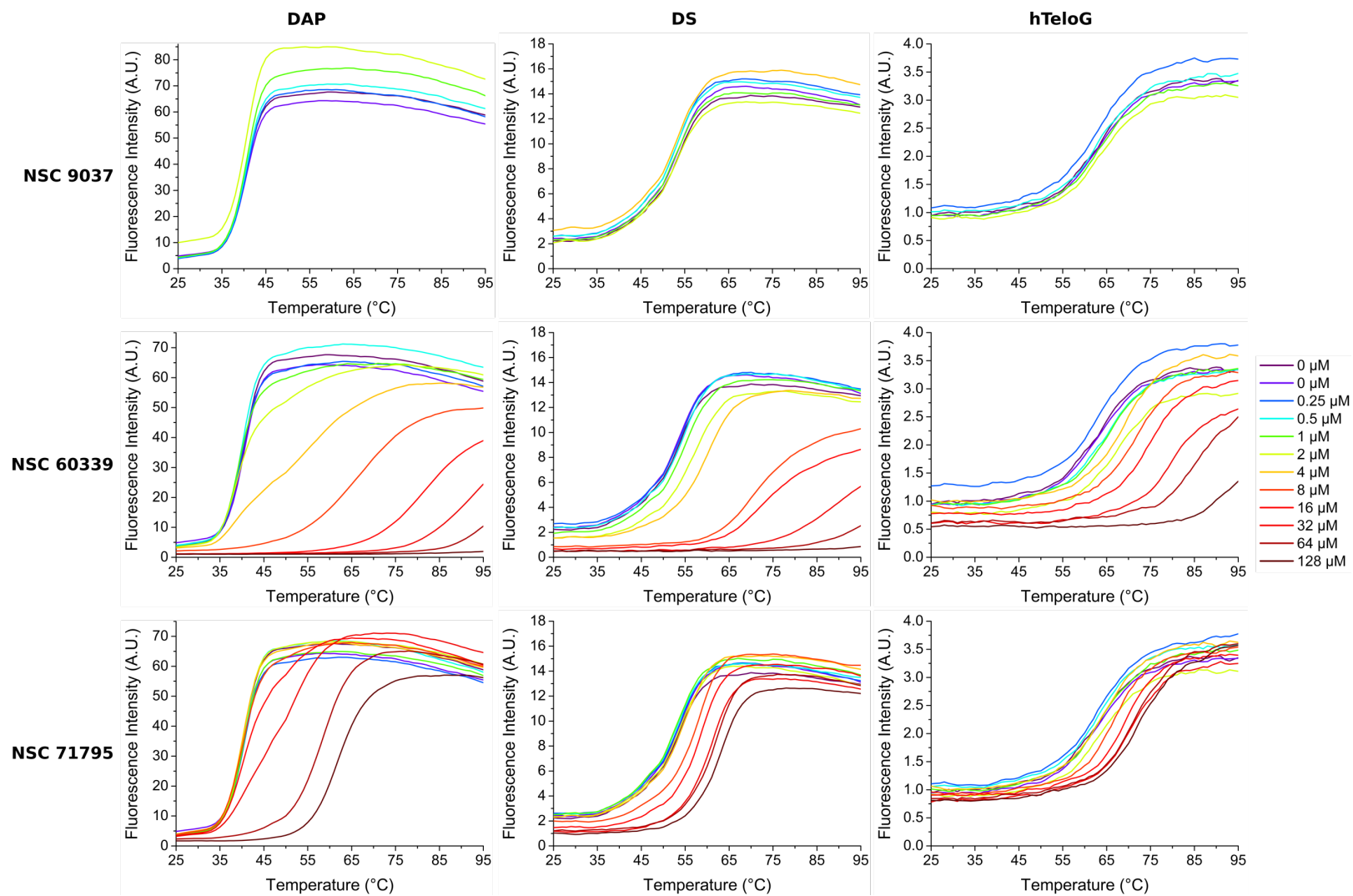


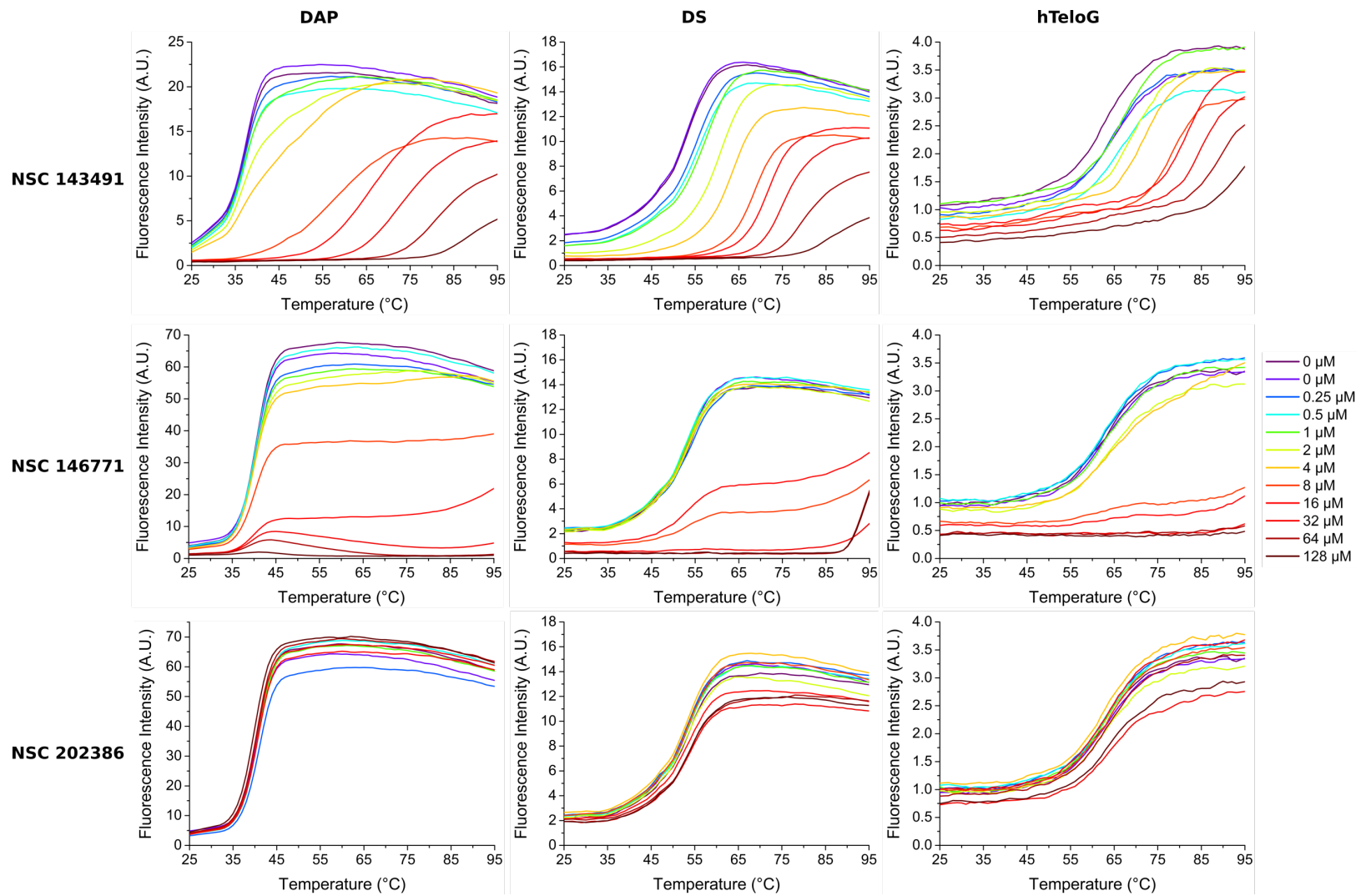
## A.5 $\Delta T_m$ of FRET-labelled DNA with NSC Ligands

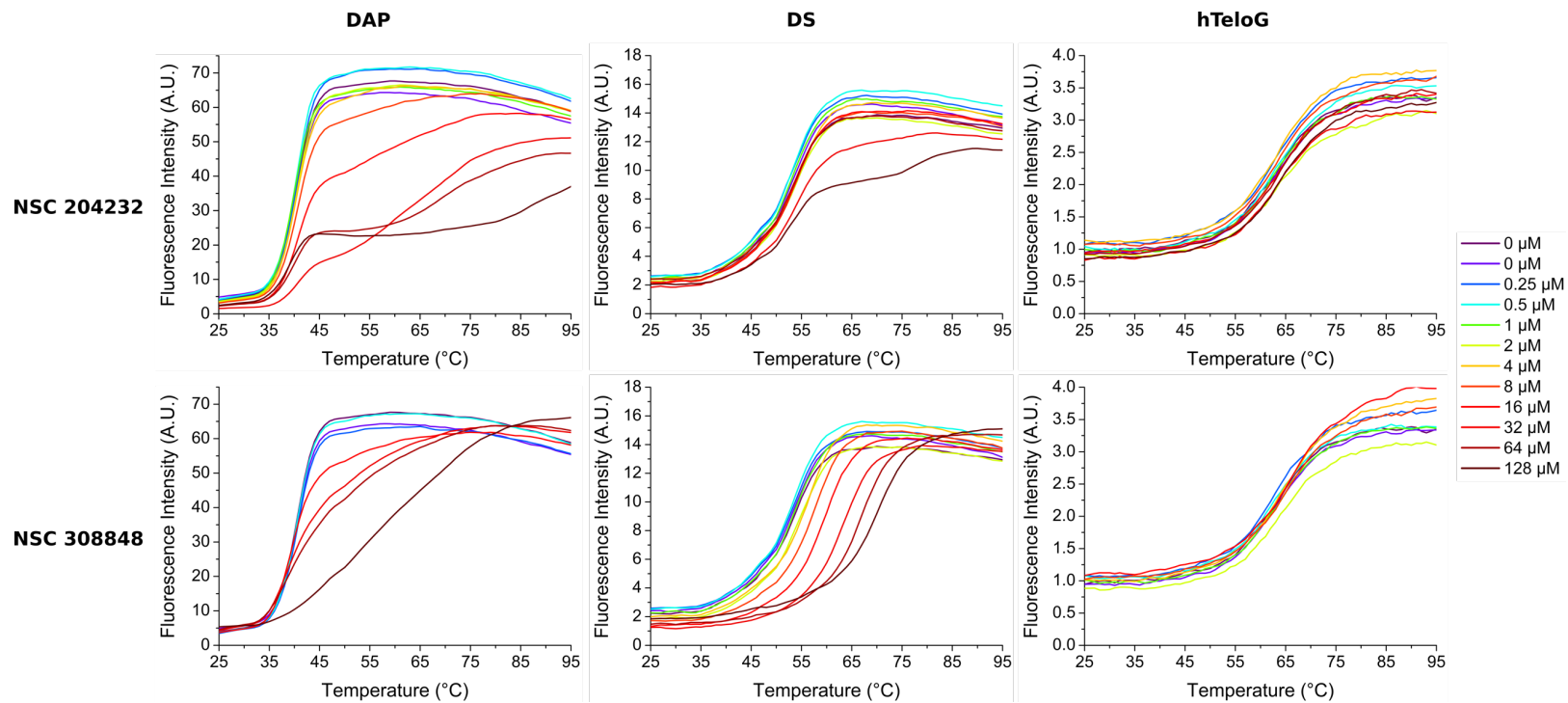
Ligand (NSC)	DNA	$\Delta T_m$ (°C)									
		0.25 $\mu$ M	0.5 $\mu$ M	1 $\mu$ M	2 $\mu$ M	4 $\mu$ M	8 $\mu$ M	16 $\mu$ M	32 $\mu$ M	64 $\mu$ M	128 $\mu$ M
9037	DAP	0.33	0.00	0.33	0.00	-0.33	-5.00	x	x	x	x
	DS	0.00	0.00	-1.67	0.00	0.33	-1.00	-0.67	-12.33	14.67	13.67
	hTeloG	0.67	-1.67	-2.00	-2.67	-4.33	0.33	-63.50	-22.00	2.00	1.67
60339	DAP	-0.33	-0.33	-1.33	-1.67	-2.00	28.00	40.33	52.33	54.00	52.67
	DS	1.33	1.33	2.00	4.67	6.67	16.00	19.00	37.33	41.00	-52.83
	hTeloG	-1.33	2.00	2.33	3.33	5.33	8.00	9.33	16.00	21.67	30.33
71795	DAP	0.00	0.00	0.00	0.00	-0.33	-0.33	0.00	11.33	18.33	21.33
	DS	1.00	0.00	-1.33	0.33	0.67	5.00	5.00	8.00	8.67	9.67
	hTeloG	-1.67	-1.33	-3.67	-2.33	-1.00	4.00	5.00	7.33	7.33	7.67
143491	DAP	-0.33	-0.33	-0.67	-1.00	-1.33	19.33	27.67	34.33	42.33	51.00
	DS	2.83	2.83	4.50	7.83	10.83	15.50	18.50	21.83	26.17	31.50
	hTeloG	0.83	2.50	4.17	4.50	8.50	15.50	19.17	21.50	27.50	31.83
146771	DAP	2.67	2.67	2.67	2.33	2.33	2.33	2.33	1.67	1.00	-0.67
	DS	1.17	-1.83	-1.50	-0.17	-0.50	40.50	-0.83	40.83	39.83	39.83
	hTeloG	-1.17	-1.17	-1.50	-1.17	0.17	29.50	31.17	31.17	31.17	31.17
202386	DAP	3.33	3.00	3.00	0.00	-0.33	0.00	3.00	2.67	2.67	2.00
	DS	-0.17	-1.50	0.17	0.17	-0.17	-0.83	0.17	0.50	0.17	0.83
	hTeloG	-0.50	-3.17	-0.17	0.17	-1.17	0.83	0.83	0.17	-0.50	-0.50
204232	DAP	3.00	3.00	2.00	2.67	2.67	3.00	3.00	3.00	2.33	1.33
	DS	-1.83	0.17	-0.50	0.50	-0.17	-0.17	0.17	0.50	1.17	-2.17
	hTeloG	-0.83	-1.17	-1.83	-0.50	0.50	0.17	3.83	0.17	0.17	0.17
308848	DAP	3.00	2.67	0.00	0.00	-0.33	-0.33	1.67	0.67	1.00	14.33
	DS	0.50	0.17	0.17	0.17	3.17	3.83	6.50	9.17	12.83	16.50
	hTeloG	-0.83	-3.17	1.83	-0.50	-0.17	0.83	3.17	4.17	9.83	10.50



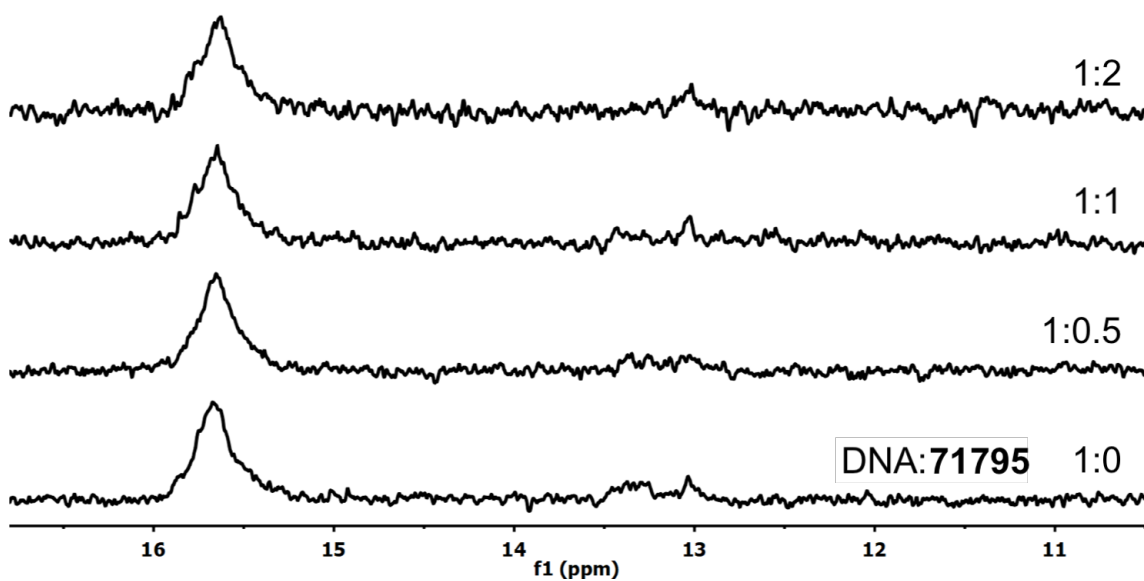
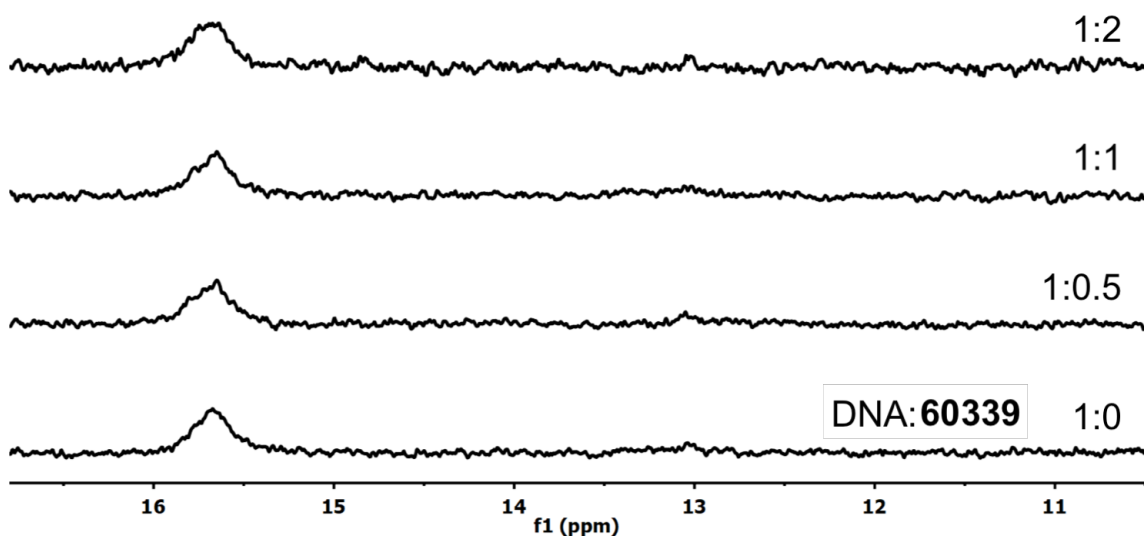
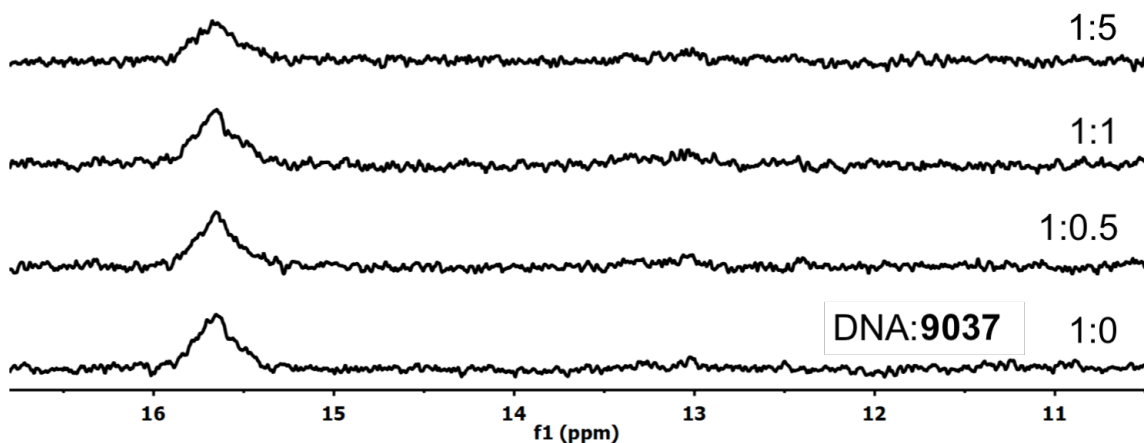
## A.6 FRET-melting curves of DNA with NSC Ligands

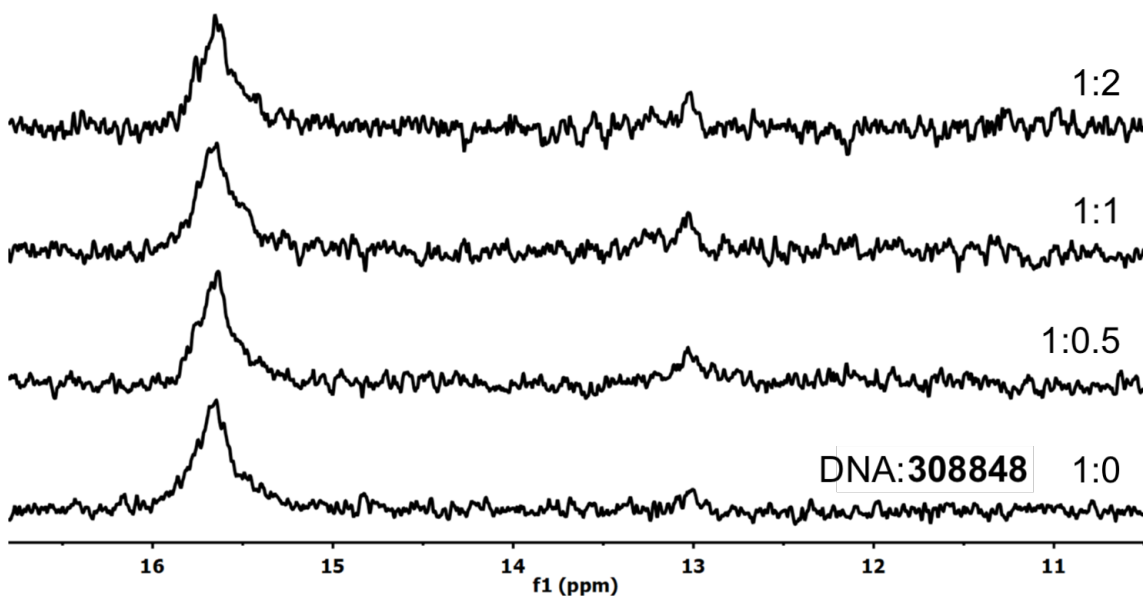
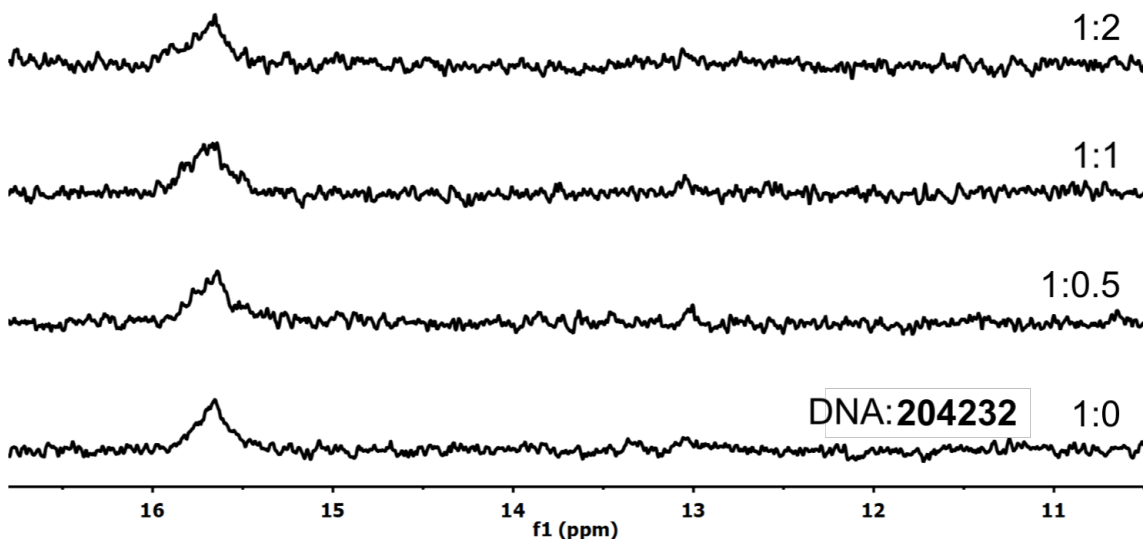
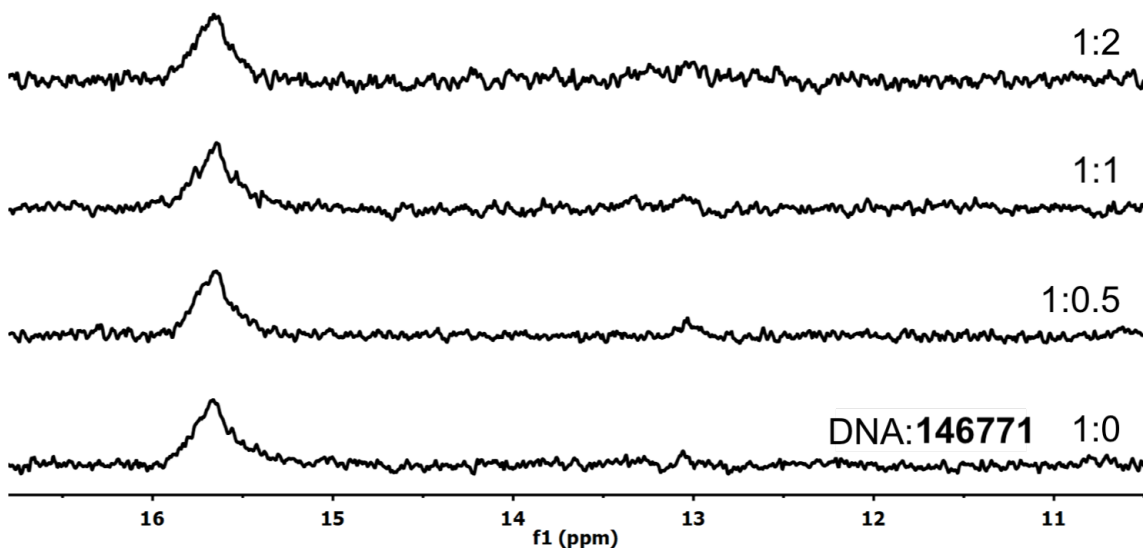






## A.7 NMR Spectra of NSC Ligands with DAP





## A.8 FID Assays at Acidic pH

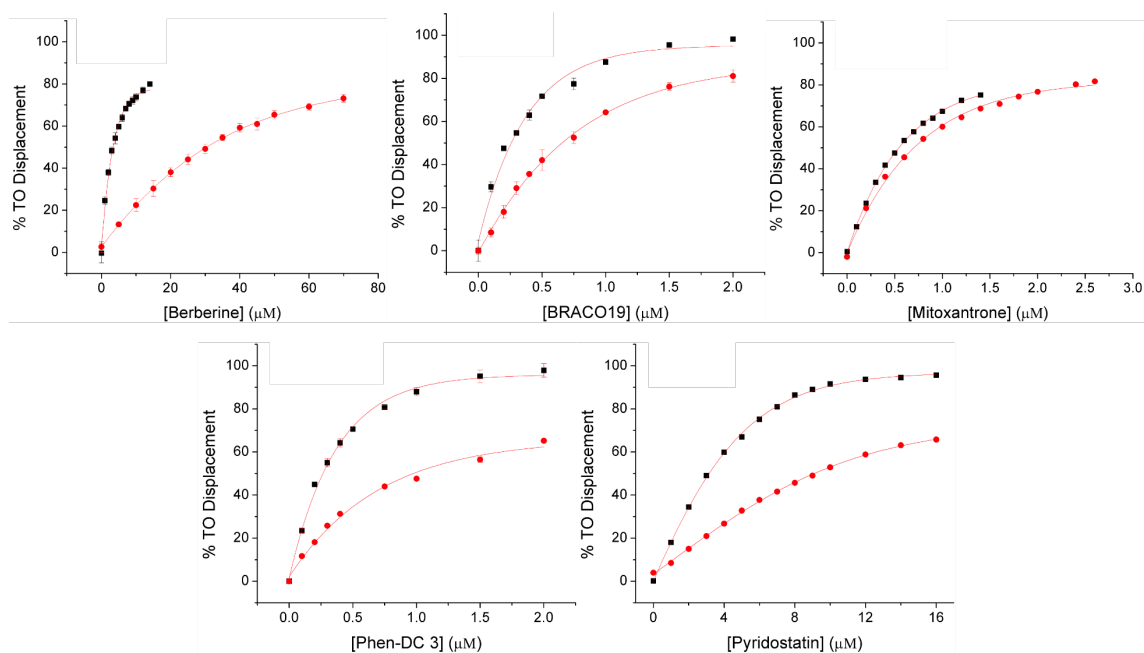


Figure A.8.1 Dose response curves from FID assays of G-quadruplex (1  $\mu\text{M}$ ) in 10 mM  $\text{KH}_2\text{PO}_4$  (black) and i-motif (1  $\mu\text{M}$ ) in 10 mM  $\text{NaH}_2\text{PO}_4$  (red) with different ligands at pH 4.3.

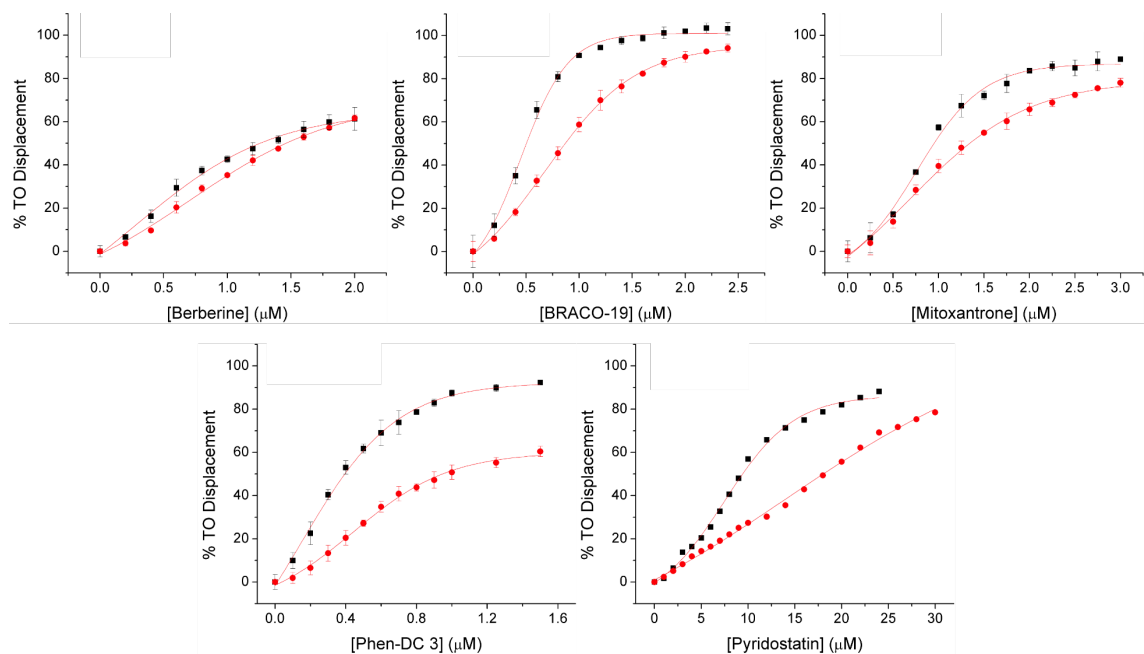


Figure A.8.2 Dose response curves from FID assays of G-quadruplex (1  $\mu\text{M}$ ) in 10 mM  $\text{KH}_2\text{PO}_4$  (black) and i-motif (1  $\mu\text{M}$ ) in 10 mM  $\text{NaH}_2\text{PO}_4$  (red) with different ligands at pH 5.7.

## A.9 CD Melting of i-Motif with Ligands at Neutral pH

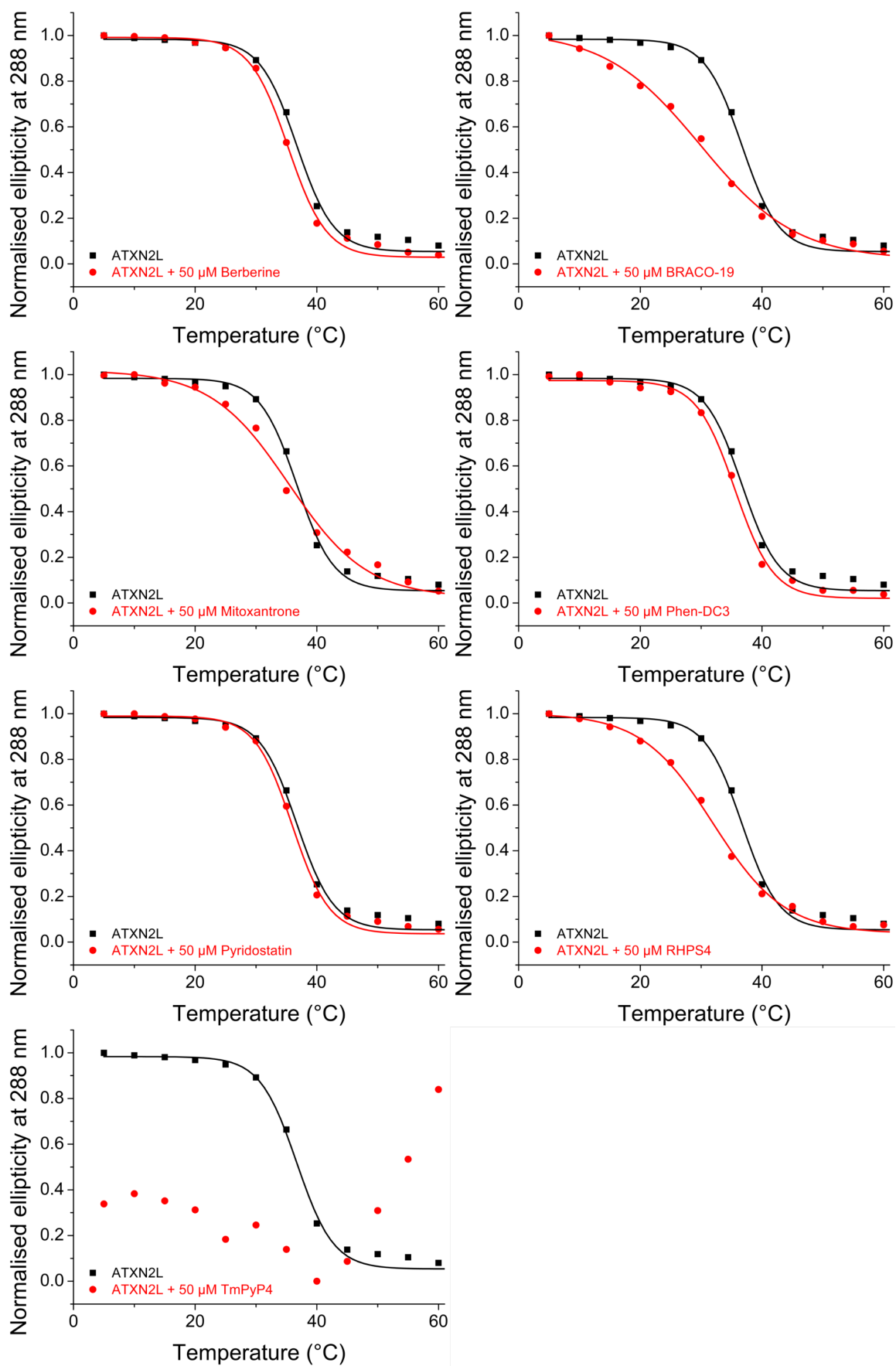


Figure A.9.1 Normalised ellipticity at 288 nm of 10  $\mu$ M ATXN2L without ligand (black) and with five equivalents (50  $\mu$ M) ligand (red).

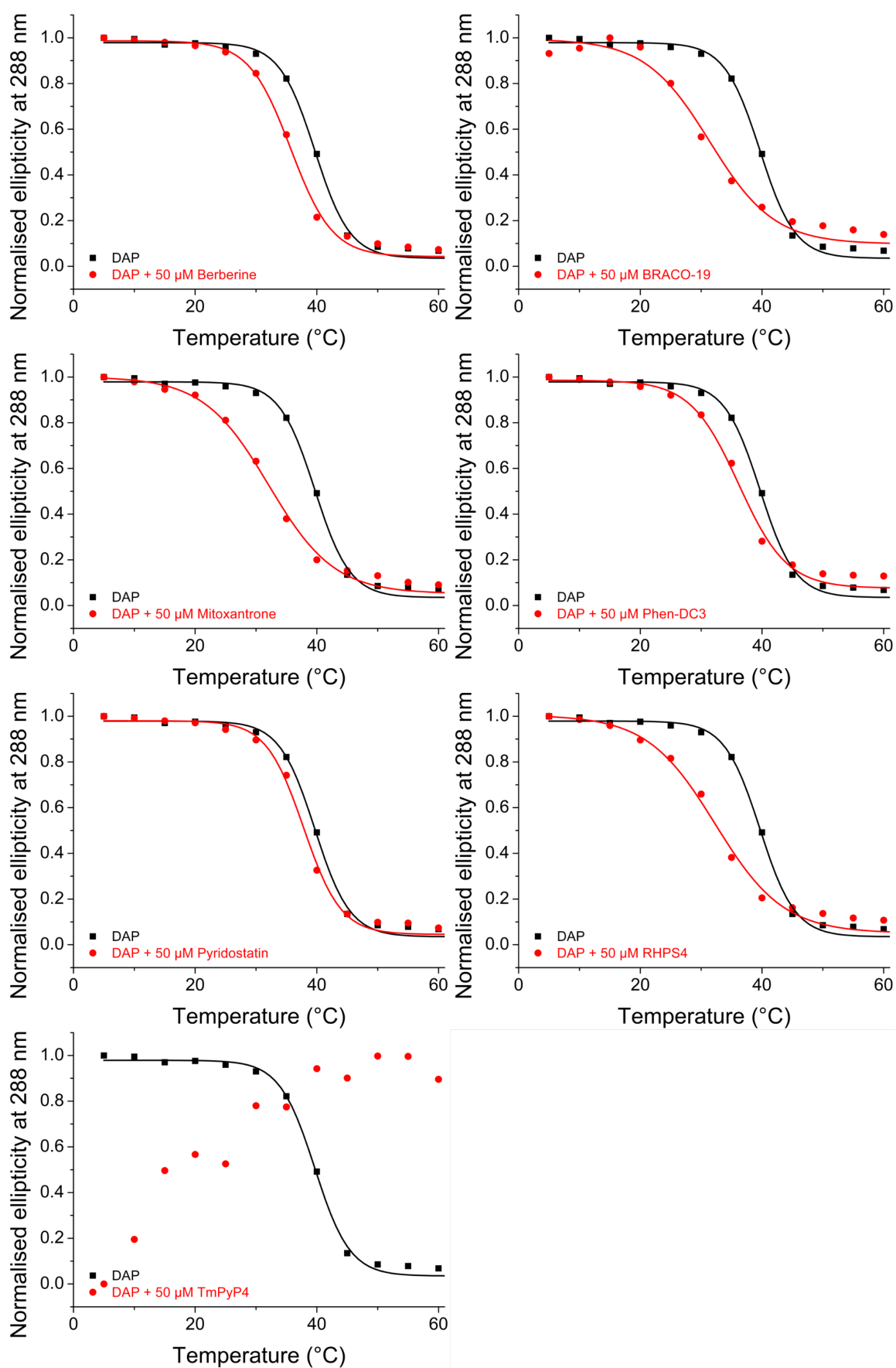


Figure A.9.2 Normalised ellipticity at 288 nm of 10 μM DAP without ligand (black) and with five equivalents (50 μM) ligand (red).



## A.10 UV Melting of i-Motif with Ligands at Neutral pH

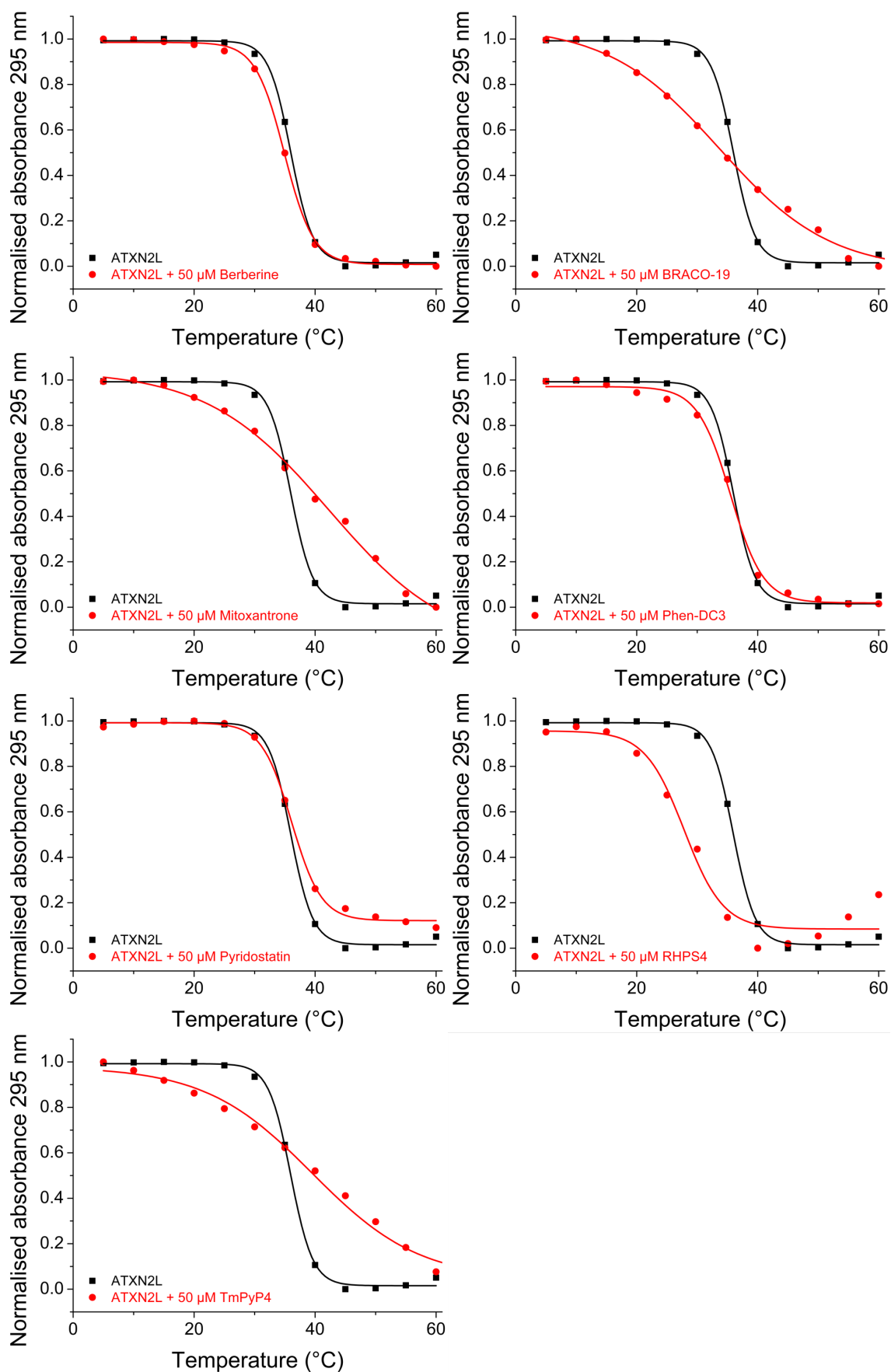


Figure A.10.1 Normalised absorbance at 295 nm of 10 μM ATXN2L without ligand (black) and with five equivalents (50 μM) ligand (red).

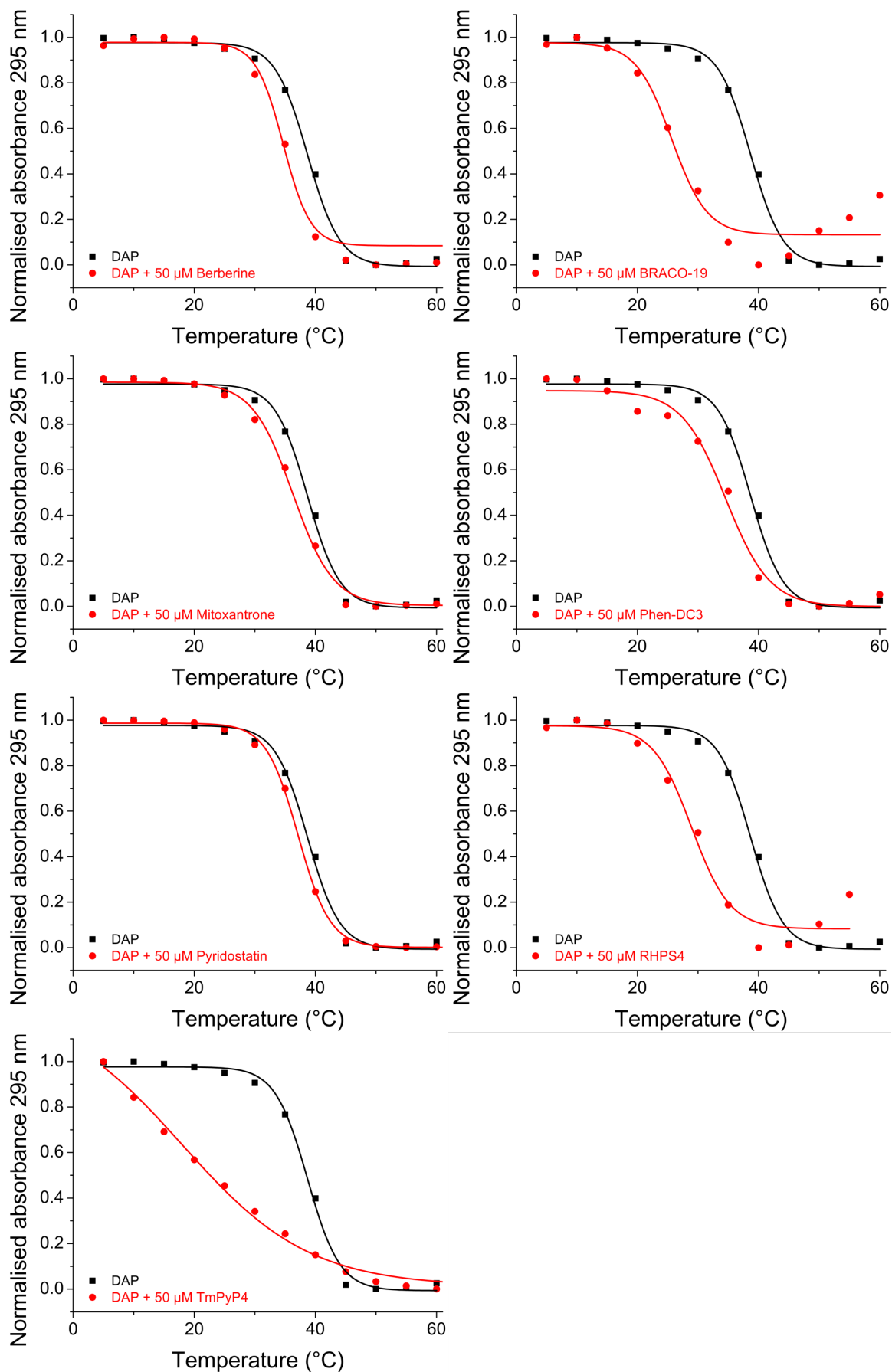


Figure A.10.2 Normalised absorbance at 295 nm of 10  $\mu\text{M}$  DAP without ligand (black) and with five equivalents (50  $\mu\text{M}$ ) ligand (red).

## A.11 FID Assay of i-Motif with Ligands at Neutral pH

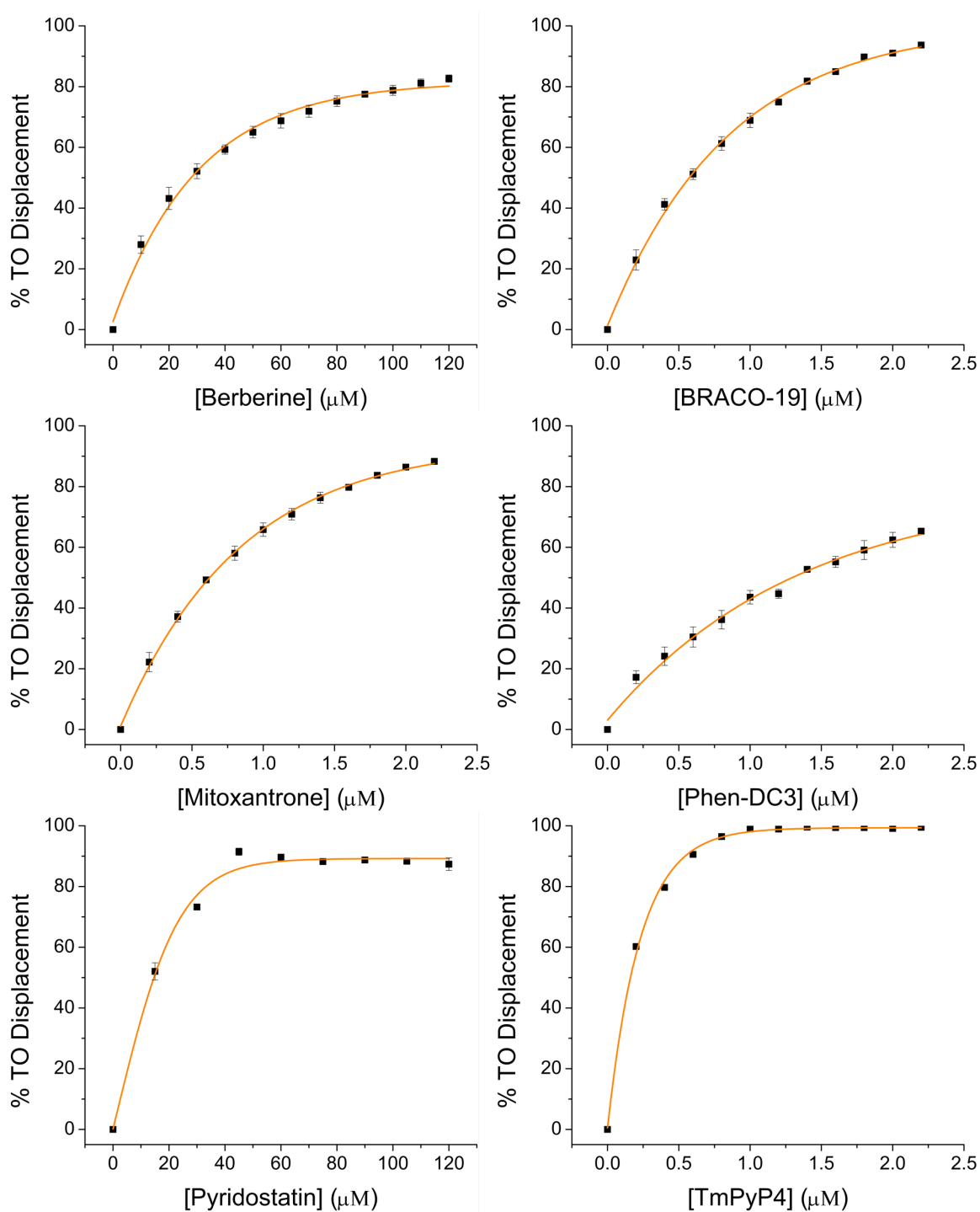
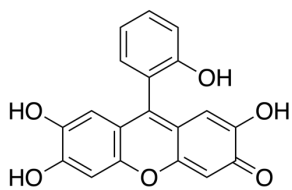
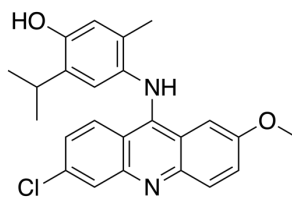


Figure A.11.1 Dose-response curves from FID assays of DAP ( $1 \mu\text{M}$ ) at pH 7.0 in 10 mM sodium cacodylate buffer with ligands. Error bars show standard deviation across three repeats.

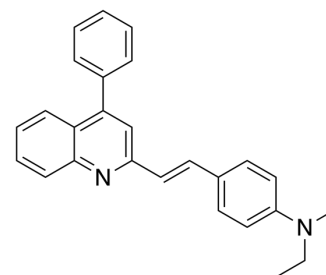
## A.12 Ligand Structures Reference Sheets



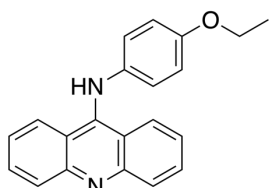
NSC 9037



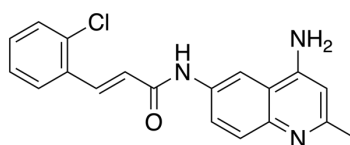
NSC 13051



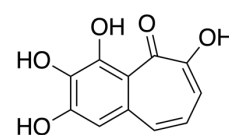
NSC 13487



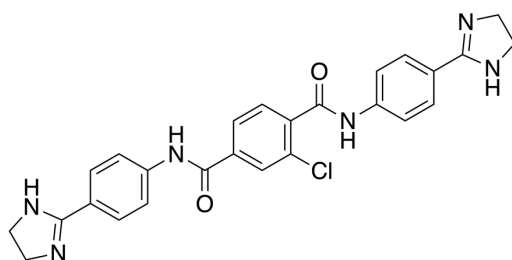
NSC 30205



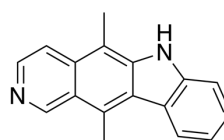
NSC 33353



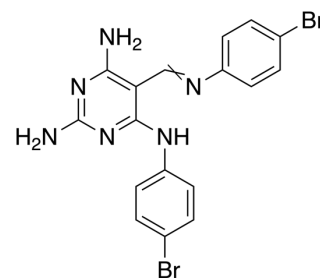
NSC 35676



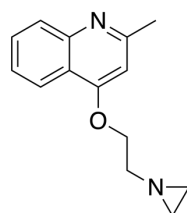
NSC 60339



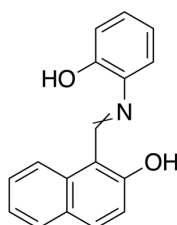
NSC 71795



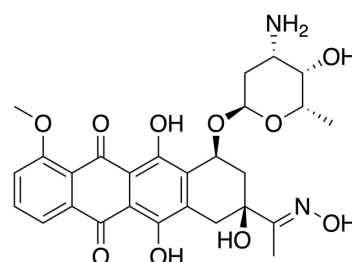
NSC 103520



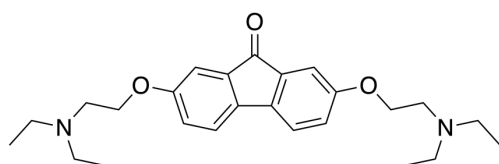
NSC 109086



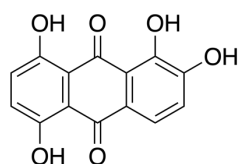
NSC 111847



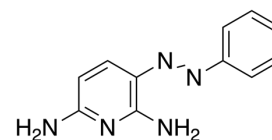
NSC 143491



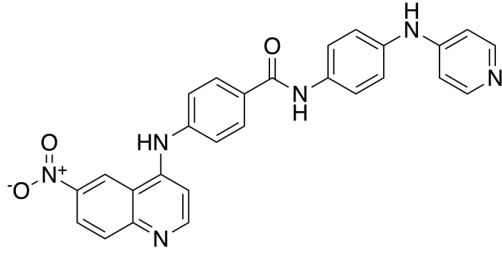
NSC 143969



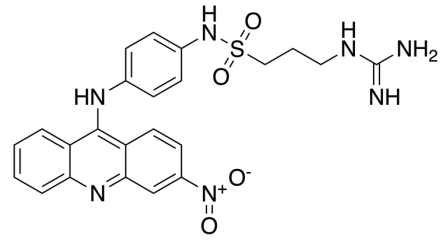
NSC 144046



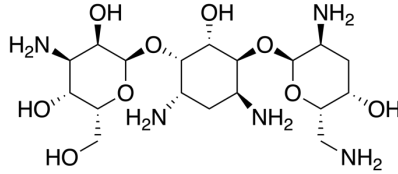
NSC 145895



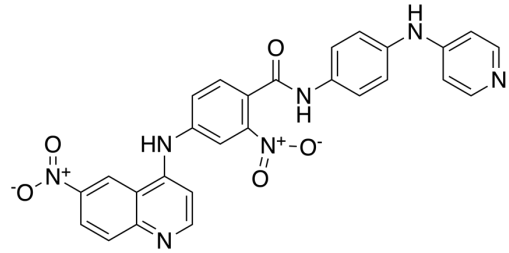
NSC 146771



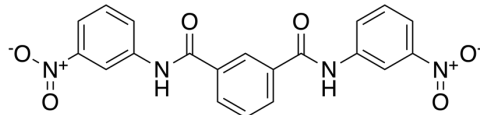
NSC 177365



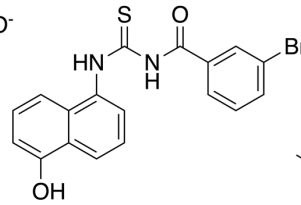
NSC 180514



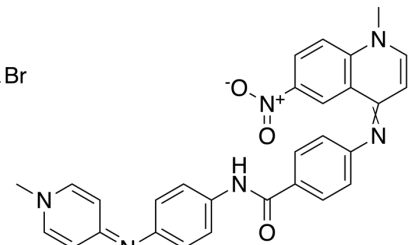
NSC 202386



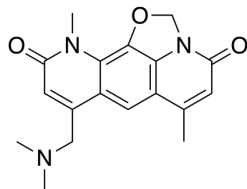
NSC 204232



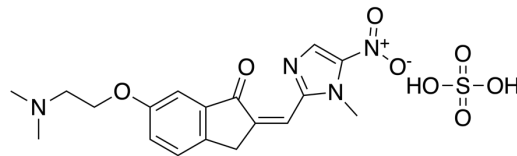
NSC 215718



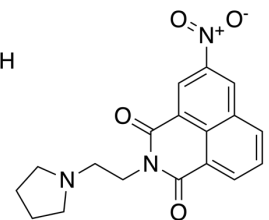
NSC 260594



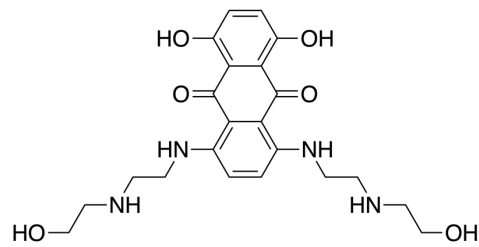
NSC 275428



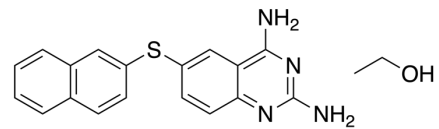
NSC 277184



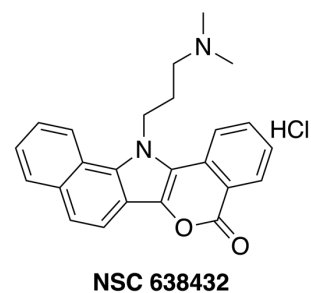
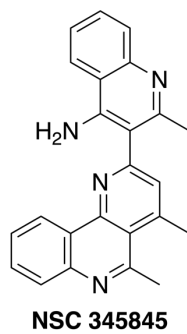
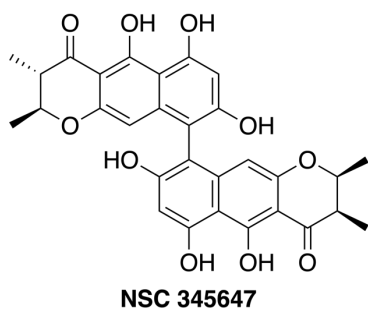
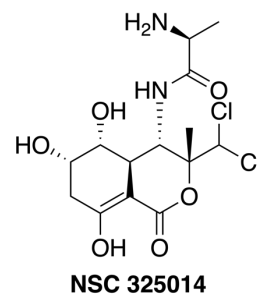
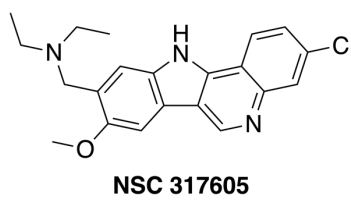
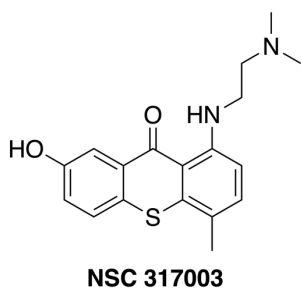
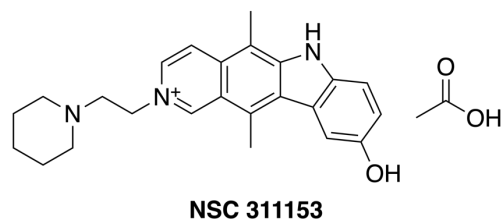
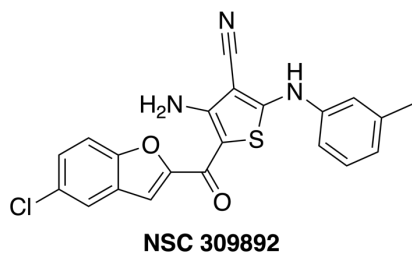
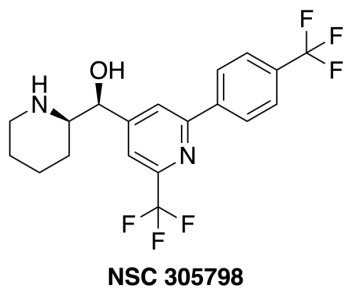
NSC 300289



NSC 301739



NSC 305780



---

## A.13 Publications

## Redox-dependent control of i-Motif DNA structure using copper cations

Mahmoud A.S. Abdelhamid<sup>1,2</sup>, László Fábrián<sup>1</sup>, Colin J. MacDonald<sup>1,3</sup>, Myles R. Cheesman<sup>2,3</sup>, Andrew J. Gates<sup>2,4,\*</sup> and Zoë A.E. Waller<sup>1,2,\*</sup>

<sup>1</sup>School of Pharmacy, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK, <sup>2</sup>Centre for Molecular and Structural Biochemistry, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK, <sup>3</sup>School of Chemistry, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK and <sup>4</sup>School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK

Received January 05, 2018; Revised April 25, 2018; Editorial Decision April 26, 2018; Accepted May 03, 2018

### ABSTRACT

**Previous computational studies have shown that Cu<sup>+</sup> can act as a substitute for H<sup>+</sup> to support formation of cytosine (C) dimers with similar conformation to the hemi-protonated base pair found in i-motif DNA. Through a range of biophysical methods, we provide experimental evidence to support the hypothesis that Cu<sup>+</sup> can mediate C–C base pairing in i-motif DNA and preserve i-motif structure. These effects can be reversed using a metal chelator, or exposure to ambient oxygen in the air that drives oxidation of Cu<sup>+</sup> to Cu<sup>2+</sup>, a comparatively weak ligand. Herein, we present a dynamic and redox-sensitive system for conformational control of an i-motif forming DNA sequence in response to copper cations.**

### INTRODUCTION

Substantial interest and research has been devoted to studying the characteristics of the many non-canonical secondary structures that can be adopted by DNA. The ability of DNA to assume different conformations is controlled by the specific sequence of bases and the local environment (1). One such structure is the i-motif that forms from cytosine-rich DNA when two parallel duplexes containing cytosine repeats intercalate to form a quadruplex structure stabilized by C<sup>+</sup>–C base pairing (Figure 1) (2). These structures may readily form in acidic conditions where the N3 of cytosine can be protonated; subsequent intercalation and formation of the i-motif occurs rapidly (3). In 2003, this property was utilized to create the first proton-fuelled i-motif nanomotor (4). Since then, the i-motif has been exploited in the design of hundreds of pH-driven nanomachines (5,6) including an example of a light-driven pH-jump system (7) and a DNA nanomachine that can map spatial and temporal pH changes in living cells (8). There have been significant ad-

vances in the understanding i-motif structure and dynamics (9–11) and also the sequences (12–14) which can enable fine-tuning of the properties of these types of devices.

In addition to pH, Hg<sup>2+</sup> (15) and Ag<sup>+</sup> (16) have been used as alternative triggers for i-motif formation, but generally studies into the influence of cations on i-motifs are limited. It is known that alkali metals destabilize the structures (17,18) and our own work has revealed that Cu<sup>2+</sup> is capable of re-folding i-motif forming sequences into a hairpin structure, even competing with acid-stabilized i-motif at low pH (19). Recently, Oomens and co-workers used infrared ion spectroscopy in combination with density functional theory (DFT) calculations to show that cytosine monomers in the presence of Cu<sup>+</sup> form C–Cu<sup>+</sup>–C structures, analogous to the hemi-protonated C-dimers at the core of the i-motif (Figure 1C) (20). In contrast to alkali metal ions, that induce a different dimer conformation which sacrifices hydrogen-bonding interactions between bases for improved chelation of the metal cation, the C–Cu<sup>+</sup>–C dimer complex was proposed to be stable (20). Given the requirement for C–C base pairs in i-motif, building from this work we decided to investigate the effects of Cu<sup>+</sup> on an i-motif forming DNA sequence.

### MATERIALS AND METHODS

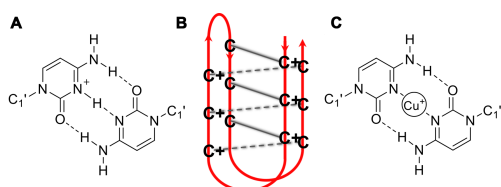
#### General experimental

Oligonucleotides hTeloC and hTeloC<sub>FRET</sub> were purchased from Eurogentec and were HPLC purified. Solid DNA samples were initially dissolved as a stock solution in MilliQ water (100 μM for labeled and 1 mM for unlabeled); further dilutions were carried out in the respective sodium cacodylate buffer. Samples were thermally annealed in a heat block at 95°C for 5 min and left to cool slowly to room temperature overnight.

Experiments requiring anoxic conditions were performed under nitrogen in a UNILab Plus glove box workstation with

\*To whom correspondence should be addressed. Tel: +44 1603 59 1972; Fax: +44 1603 59 2003; Email: z.waller@uea.ac.uk  
Correspondence may also be addressed to Andrew J. Gates. Tel: +44 1603 59 2931; Fax: +44 1603 59 2250; Email: a.gates@uea.ac.uk





**Figure 1.** (A) Hemi-protonated C<sup>+</sup>-C base pair (B) i-motif structure (C) Proposed C-C base pair stabilized by Cu<sup>+</sup>.

the oxygen level maintained below 0.5 ppm. Solutions of Cu<sup>+</sup> were prepared in the glove box by dissolving solid CuCl in 0.1 M HCl and 1 M NaCl solution to double the desired concentration and then diluted to the final concentration using 2× sodium cacodylate buffer (final dilution using buffer prevents pH change on cation addition during experiments). Solutions of Cu<sup>2+</sup>, diethyldithiocarbamate (DETC) and sodium ascorbate were prepared by dissolving solid Cu<sub>2</sub>SO<sub>4</sub>, DETC or sodium ascorbate in MilliQ water, respectively.

#### UV spectroscopy

UV spectroscopy experiments were performed on an Agilent Technologies Cary 4000 UV-Vis spectrophotometer and recorded using an open-top screw-cap 10 mm quartz cuvette with a silicone rubber septum to exclude air on transfer from the glove box to the spectrophotometer. Samples (1 ml) were diluted to 2.5 μM hTeloC in 10 mM sodium cacodylate buffer at the desired pH. Cu<sup>+</sup> was added in 0.5 μl aliquots, and mixed, using a pipette to the desired concentration. Spectra were recorded over a wavelength range of 400–200 nm at room temperature in the absence of cation then after each addition, and zero corrected at 400 nm. The difference spectra at either pH were calculated by subtraction of the final folded spectrum, in the presence of Cu<sup>+</sup>, from the spectrum in the absence of Cu<sup>+</sup>.

#### Circular dichroism

Circular dichroism (CD) spectra were recorded on a Jasco J-810 spectropolarimeter using a 1 mm path length quartz cuvette with a neck and a silicone stopper to exclude air on transfer from the glove box to the spectropolarimeter. hTeloC was diluted to 10 μM in 50 mM sodium cacodylate buffer at the desired pH to a total volume of 200 μl. The scans were performed at room temperature over a wavelength range of 200–320 nm with a scanning speed of 200 nm/min, response time of 1 s, 0.5 nm pitch and 2 nm bandwidth. A blank sample containing only buffer was treated in the same manner and subtracted from the collected data. Cu<sup>+</sup> was added in 0.5 μl aliquots, and mixed, using a pipette to the desired concentration. For the chelator titration 150 μM Cu<sup>+</sup> was added via a 1 μl addition from a 30 mM stock solution, DETC was then added in 0.5 μl aliquots to the desired concentration as above. Control spectra of DETC in buffer and DETC with hTeloC were also measured to confirm DETC itself had no effect on the spectra or the conformation of the DNA. The CD spectra represent an average

of three scans and are zero corrected at 320 nm. For the redox experiments, each component was added via a 1 μl addition from a stock solution prepared at the concentration needed to yield the desired concentration. Control spectra of sodium ascorbate in buffer and sodium ascorbate with hTeloC were also measured to confirm that sodium ascorbate itself had no effect on the spectra or the conformation of the DNA. Titration experiments were performed at least in triplicate and processing of the data was carried out using Origin.

#### FRET melting

The labeled oligonucleotide hTeloC<sub>FRET</sub> (5'-FAM-[TAA-CCC]<sub>4</sub>-TAMRA-3'; donor fluorophore FAM is 6-carboxyfluorescein; acceptor fluorophore TAMRA is 6-carboxytetramethylrhodamine) was prepared as a 400 nM solution in 10 mM sodium cacodylate buffer at the respective pH and thermally annealed. Strip-tubes (QIAGEN) were prepared by aliquoting 10 μl of the annealed DNA, followed by addition of Cu<sup>2+</sup> solution and 2× sodium ascorbate solution to give the desired Cu<sup>+</sup> concentration range across the samples, and made up with 10 mM sodium cacodylate buffer to a final volume of 20 μl. Fluorescence melting curves were acquired in a QIAGEN Rotor-Gene Q-series polymerase chain reaction machine. Measurements were made with excitation at 483 nm and detection at 533 nm. Experiments were performed at each pH in triplicate with final analysis of the data carried out using QIAGEN Rotor-Gene Q-series software and Origin.

#### <sup>1</sup>H NMR

<sup>1</sup>H nuclear magnetic resonance (NMR) experiments were performed using a Bruker Avance III 800 MHz spectrometer equipped with an HCN inverse triple resonance z-gradient probe. Aqueous solutions were prepared with the addition of 5% D<sub>2</sub>O to enable field/frequency lock. Solvent suppression of the water resonance was achieved using a 1D Watergate sequence employing a symmetrical 3-τ-9-τ-19 pulse train inversion element. The solvent resonance, which was minimized, was set on-resonance at the transmitter offset and the interpulse delay time (τ) was adjusted to achieve an excitation maximum in the imino proton region of interest. hTeloC was diluted to a concentration of 10 μM in 50 mM sodium cacodylate buffer at pH 5.5 containing 5% D<sub>2</sub>O. The spectrum of hTeloC alone was measured over 1 h after which 150 μM of Cu<sup>+</sup> was added and the subsequent spectrum acquired over 2 h. At last, excess DETC (540 μM) was added and the spectrum acquired again for 1 h. NMR spectra were acquired and processed using Bruker's TopSpin™ software package (v3.1.7 Bruker Biospin) for NMR data analysis.

#### Computational methods

The starting point for our i-motif structural model was the PDB entry: 1EL2 (21), which was manually edited (22) to match the hTeloC sequence. This model, stabilized by six C<sup>+</sup>-C base pairs, was relaxed using a 200 ns explicit solvent molecular dynamics run to allow minor conformational changes in response to the altered sequence. The force

field consisted of the OL15 parameters for DNA (23–25), TIP3P (26) model for water and Li, Song and Merz parameters for the ions (27). Partial charges for the protonated C<sup>+</sup> bases were obtained by the RESP fitting procedure (see Supplementary Data). After equilibration, the 200 ns molecular dynamics simulation was performed at constant pressure and temperature (NPT,  $P = 0.1$  MPa,  $T = 300$  K) by using the Gromacs package (28).

The model of the C–Cu<sup>+</sup>–C i-motif was created from the final snapshot of the molecular dynamics run. Six Cu<sup>+</sup> ions were inserted between matching cytosine groups, replacing the H<sup>+</sup> ions, and three in the TAA loop regions. Optimization of this initial model with the force field described above gave an unexpected result with the Cu<sup>+</sup> ions moving out of the planes of the cytosine rings. To validate this finding, we turned to the semi-empirical method PM6-D3H4 (29,30), which we had found to give good approximations to DFT results on isolated C–Cu<sup>+</sup>–C complexes (Supplementary Table S1). Optimization of the initial model with the Cu<sup>+</sup> ions inserted by the PM6-D3H4 method gave the final structural model. Unfortunately, DFT-based optimization of the complete i-motif structure was not computationally feasible. As another approximation, a stack of six C–Cu<sup>+</sup>–C base pairs and the neighboring T and A residues were extracted from the initial model and optimized by the DFT method TPSS-D3(BJ)/def2-SV(P). Both the semi-empirical and DFT models featured the non-planar C–Cu<sup>+</sup>–C linkages.

Interaction energies between Cu<sup>+</sup> or Cu<sup>2+</sup> and the surrounding DNA residues (Supplementary Tables S1 and 2) were estimated both by using planar C–Cu–C complexes and by taking a fragment from the above DFT optimized stack, which consisted of two Cu ions and five bases (with additional geometry optimization to locate an energy minimum for Cu<sup>2+</sup>). The interaction energies were calculated by dispersion corrected DFT methods with def2-TZVP basis and include counterpoise correction for the basis set superposition error. All DFT calculations were performed by using the NWChem package (31).

## RESULTS AND DISCUSSION

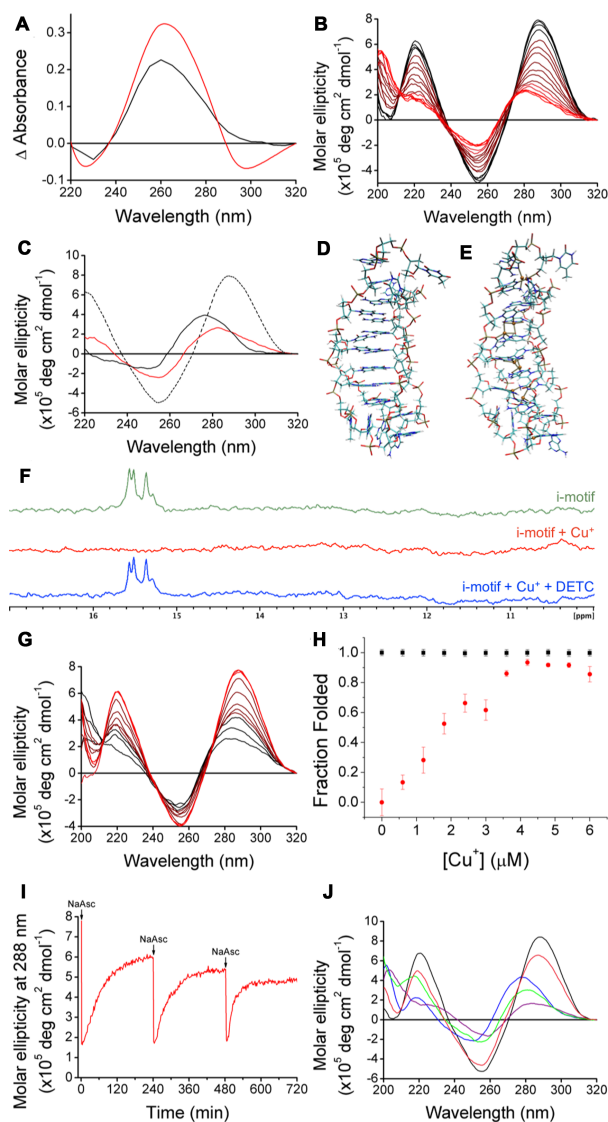
To investigate if Cu<sup>+</sup> can induce the formation of a secondary structure in DNA, we used the cytosine-rich human telomeric DNA sequence hTeloC 5'-[TAA-CCC]<sub>4</sub>-3' that is predominantly unfolded at physiological pH, but capable of forming i-motif at acidic pH (pH < 6), Cu<sup>+</sup> in solution is well known to oxidize readily to Cu<sup>2+</sup> when exposed to O<sub>2</sub> (32,33), therefore all experiments were performed at ambient temperature under strict anoxic conditions in an N<sub>2</sub> atmosphere (containing < 0.5 ppm O<sub>2</sub>). The UV absorbance profile of DNA is dependent on its conformation. Therefore, UV spectroscopy can be used to elucidate whether DNA is folded or unfolded, and to reveal the existence of higher-order secondary structure(s) (34). UV-difference spectra are used to identify and characterize the behavior of the secondary structure in response to experimental conditions (35,36). 'Cu<sup>+</sup> difference' spectra for hTeloC were measured at pH 5.5 and pH 7.4, where the structure is an i-motif or unfolded DNA, respectively. The resulting spectra (Figure 2A) display a positive signal at 260

nm at both pH values, and a negative signal at 295 nm at pH 7.4, both consistent with when the i-motif formed by decreasing the pH (37). These results indicate that the final configuration of the secondary structure adopted at either pH is similar, and that at pH 7.4 a more substantial reconfiguration is necessary to form the final structure.

To further characterize the structure adopted by hTeloC in the presence of Cu<sup>+</sup> we employed CD spectroscopy. The CD spectrum of hTeloC at pH 7.4 has a positive peak at 270 nm and a negative peak at 250 nm, indicative of a primarily unfolded population of oligonucleotide (38). Sequential addition of Cu<sup>+</sup> up to five equivalents (i.e. 50 μM final) resulted in a bathochromic shift in the positive peak from 270 to 278 nm, while the position of the negative peak at 250 nm remained constant (Supplementary Figure S1). However, further addition of Cu<sup>+</sup> at this pH resulted in visible precipitation of the Cu<sup>+</sup>–DNA complex and consequent deterioration of the CD signal. At pH 5.5, hTeloC is already folded into an i-motif with a characteristic positive peak at 288 nm and negative peak at 255 nm (38). Under these conditions, titration of Cu<sup>+</sup> up to 19.5 equivalents (195 μM Cu<sup>+</sup>) led to a hypsochromic shift of the positive peak from 288 to 283 nm, and a decrease in the amplitude of the negative peak at 255 nm (Figure 2B). In contrast to the precipitation observed at pH 7.4, the Cu<sup>+</sup>–DNA complex at acidic pH was completely soluble beyond the concentration where no further changes are observed (150 μM Cu<sup>+</sup>). The changes observed at pH 7.4 and 5.5 are consistent with a Cu<sup>+</sup> induced reconfiguration of the structure. Crucially, post-Cu<sup>+</sup> addition and at both pH values, the spectra are practically superimposable indicating that a similar final structure is adopted regardless of pH.

Given the spectroscopic changes previously observed with Cu<sup>2+</sup> and hTeloC, the possibility that the structure adopted in the presence of Cu<sup>+</sup> may also display hairpin-like character (19) was explored, and the different copper–DNA complexes were compared using CD at pH 5.5. At this pH, in the absence of copper the CD spectrum of hTeloC has a positive peak at 288 nm indicative of i-motif structure. Addition of either Cu<sup>+</sup> or Cu<sup>2+</sup> resulted in a hypsochromic-shift consistent with an alteration in the structure of the DNA. Addition of Cu<sup>2+</sup> shifts this peak to 276 nm compared to only 283 nm when Cu<sup>+</sup> is added. The negative peak at 255 nm also undergoes a hypsochromic-shift to 250 nm in the presence of Cu<sup>2+</sup>, while the peak position does not shift at all when Cu<sup>+</sup> is added (Figure 2C). This strongly suggests that the Cu<sup>2+</sup>–DNA complex is different to the Cu<sup>+</sup>–DNA complex.

In addition to the spectroscopic differences observed using the different oxidation states of copper, the half-cation concentrations also vary by an order of magnitude. A value of 46 (±3) μM was determined for the [Cu<sup>+</sup>]<sub>50</sub>, while the [Cu<sup>2+</sup>]<sub>50</sub> was comparatively higher at 382 (±14) μM (Supplementary Figure S2). We previously suggested that the relatively high concentration of Cu<sup>2+</sup> required to form the hairpin was because the Cu<sup>2+</sup> did not affect the structure by interacting directly with the bases in the oligonucleotide, but rather by shifting the equilibrium to the hairpin structure via stabilization of the sugar–phosphate backbone (19). In the case of Cu<sup>+</sup>, continuous variation binding analysis determined the stoichiometry of Cu<sup>+</sup> to DNA to be 9:1



**Figure 2.** (A) ‘Cu<sup>+</sup>-difference’ spectra using 125 μM of Cu<sup>+</sup> to form the final conformations at pH 7.4 (red) and pH 5.5 (black). (B) CD spectra of 10 μM hTeloC at pH 5.5 (black) with titration up to 150 μM Cu<sup>+</sup> (red). (C) CD spectra of 10 μM hTeloC at pH 5.5 (dashed black), after addition of 150 μM Cu<sup>+</sup> (red) or 1 mM Cu<sup>2+</sup> (black). (D) Model of i-motif structure stabilized by Cu<sup>+</sup> ions, derived from (D) by geometry optimization with the PM6-D3H4 method. (E) Model of the i-motif structure stabilized by Cu<sup>+</sup> ions, snapshot from the end of the 200 ns simulation. (F) <sup>1</sup>H NMR of (green) 10 μM hTeloC in 50 mM sodium cacodylate buffer pH 5.5 with 5% D<sub>2</sub>O; (red) addition of 150 μM Cu<sup>+</sup>; (blue) addition of 150 μM Cu<sup>+</sup> and 540 μM DETC. (G) CD spectra of 10 μM hTeloC with 150 μM Cu<sup>+</sup> at pH 5.5 (black) with titration up to 300 μM DETC (red). (H) Fluorescence intensity at 25°C normalized using values in the absence of Cu<sup>+</sup> at pH 5.5 as 1 (folded) and at pH 7.4 as 0 (unfolded). A total of 200 nM hTeloC<sub>FRET</sub> in 10 mM buffer at pH 5.5 (black) and at pH 7.4 (red). Error bars show standard deviation across three repeats. (I) Change in molar ellipticity at 288 nm of 10 μM hTeloC at pH 5.5 with 150 μM Cu<sup>2+</sup> as a function of time with three additions of 150 μM sodium ascorbate under ambient conditions. (J) CD spectra of single sample of 10 μM hTeloC at pH 5.5 (black); addition of 1 mM Cu<sup>2+</sup> (blue); addition of 150 μM sodium ascorbate (purple); after 4 h exposure to air (green); chelation using 1 mM EDTA (pink).

(Supplementary Figure S3). This ratio further suggests that  $\text{Cu}^+$  is interacting differently to  $\text{Cu}^{2+}$ , possibly via direct mediation of an interaction between bases in the sequence, perhaps consistent with the model proposed by Oomens: one  $\text{Cu}^+$  for each C–C base pair and additionally, one for each loop.

To compare the viability of C–C base pairs stabilized by  $\text{Cu}^+$  and  $\text{Cu}^{2+}$ , models similar to the one shown in Figure 1C were created with both ions and optimized by DFT computational methods (Supplementary Figure S8) (31). The results revealed a planar C– $\text{Cu}^{2+}$ –C complex similar to the model of Oomens, but with a symmetrical structure and both the N and O atoms from a single face of each cytosine moiety coordinating the metal ion. Notably, the interaction energy between the two bases and the cation is sensitive to the redox-state of the metal, and is significantly larger for  $\text{Cu}^{2+}$  (–1700 kJ/mol) than for  $\text{Cu}^+$  (–650 kJ/mol). However, the experimental hydration enthalpies of the ions suggest that hydration is more energetically favorable to base pair formation involving  $\text{Cu}^{2+}$  (–2100 kJ/mol), compared to  $\text{Cu}^+$  (–593 kJ/mol) (39); which are consistent with the different behavior observed experimentally with these cations.

The CD spectra of the DNA at both pH values in the presence of  $\text{Cu}^+$  lay somewhere between that of unfolded DNA and i-motif DNA. Addition of  $\text{Cu}^+$  at pH 5.5 resulted in what appears to be slight unfolding of the i-motif, consistent with the structure expanding to accommodate the  $\text{Cu}^+$  cations, which are significantly larger than the protons which were previously stabilizing the C<sup>+</sup>–C base pairs. The potential folded i-motif structures were investigated in more detail using molecular modeling. First, a model of protonated hTeloC was created on the basis of the reported NMR structure from a similar sequence (Figure 2D) (21). The manually modified structure was optimized and then relaxed in a 200 ns explicit solvent molecular dynamics simulation (Supplementary Figure S9).  $\text{Cu}^+$  ions were added to this relaxed model manually (22). Six  $\text{Cu}^+$  ions were placed at the geometric midpoints between the N3 atoms of matching cytosine groups and three additional ions were placed in the TAA loop regions. The geometry of this initial  $\text{Cu}^+$ –DNA complex structure was optimized using the semi-empirical PM6-D3H4 method (Figure 2E and Supplementary Figure S10) (29,30). In the optimized structure the  $\text{Cu}^+$  ions showed a preference to interact with more than two bases, thereby breaking the planarity of the C– $\text{Cu}^+$ –C units. Nevertheless, the overall folded structure was retained. To confirm these observations, a stack of six C– $\text{Cu}^+$ –C base pairs capped at both ends with the nearest molecular fragments was extracted from the initial  $\text{Cu}^+$ –DNA complex and optimized using a DFT [TPSS-D3(BJ)/def2-SV(P)] method (Supplementary Figure S11). Both the semi-empirical and DFT calculations confirmed the preference of  $\text{Cu}^+$  ions to interact with more than two bases. Full exploration of the folding with  $\text{Cu}^+$  would require derivation and fitting of specific  $\text{Cu}^+$  force field parameters, which is beyond the scope of this work. Nevertheless, the computational modeling indicates the acid-stabilized and copper-stabilized i-motif structures are slightly different, which would explain the spectroscopic differences observed between these two species. We suggest the planar base pairing (C– $\text{Cu}^+$ –C) model may only be true

for cytosine monomers. When the cytosines form part of a larger secondary structure, the interactions are more complex which gives rise to a slightly different i-motif structure, as supported by our spectroscopic data.

We have previously shown that the effects of  $\text{Cu}^{2+}$  on the structure of hTeloC DNA can be reversed using a chelator (19). To determine if a similar reversibility can be achieved with  $\text{Cu}^+$ , the high-affinity chelator DETC was used in this work. Titration of DETC into hTeloC at pH 5.5 reverses the effects of the  $\text{Cu}^+$  addition and the structure reverts to that of the acid-stabilized i-motif. hTeloC with  $\text{Cu}^+$  had the positive peak at 283 nm, incremental titration of the chelator DETC resulted in a red-shift of the peak until it returned to the position of the acid-stabilized i-motif peak at 288 nm (Figure 2G). The negative peak in the presence of  $\text{Cu}^+$  at 255 nm did not shift its position but the amplitude of the signal increased to be more consistent with that of the original acid-stabilized i-motif.

Further experiments to examine the mode of copper binding to hTeloC were performed using <sup>1</sup>H NMR. At pH 5.5, imino proton signals can be observed at 15.5 ppm and are characteristic of the C<sup>+</sup>–C base pairs in an acid-stabilized i-motif (40). On addition of  $\text{Cu}^+$ , these signals disappeared, consistent with  $\text{Cu}^+$  replacing the protons in that position. This is in-line with the model proposed by Oomens and co-workers (20). Furthermore, no additional signals appeared, ruling out a hairpin conformation with additional Watson–Crick base pairing, as was seen with  $\text{Cu}^{2+}$  (19). The addition of the chelator DETC caused the NMR spectrum to return to that of the acid-stabilized i-motif, with chelation of  $\text{Cu}^+$  resulting in the reappearance of the imino proton signal at 15.5 ppm (Figure 2F).

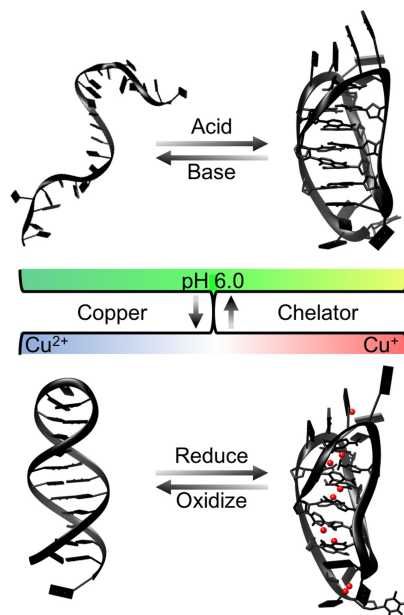
Thus far, all experiments were carried out under stringent anoxic conditions to prevent the oxidation of  $\text{Cu}^+$  that would occur in the open air. We were interested in exploring whether redox-linked structural rearrangement would be observed *in situ* if the  $\text{Cu}^+$ –DNA complex was exposed to air and the metal oxidized. The  $\text{Cu}^+$ -i-motif can be formed by adding 150  $\mu\text{M}$  of  $\text{Cu}^+$  and, as  $\text{Cu}^+$  is a high-affinity ligand compared with  $\text{Cu}^{2+}$ , complete oxidation to  $\text{Cu}^{2+}$  yields a cation concentration below the  $[\text{Cu}^{2+}]_{50}$  ( $382 \pm 14 \mu\text{M}$ ). As a result, there would not be enough  $\text{Cu}^{2+}$  to fully stabilize the hairpin. To test this hypothesis,  $\text{Cu}^+$  was added to hTeloC at pH 5.5 and the sample was split into two. One sample was maintained in an anoxic environment while the other was exposed to the open air. The CD spectrum for each condition was subsequently measured and the one which had been exposed to oxygen reverted almost completely to the acid-stabilized i-motif: the positive peak moved from 281 to 286 nm, and the amplitude of the negative peak at 255 nm increased, as observed when the  $\text{Cu}^+$  was chelated with DETC. In contrast, the sample maintained in the anoxic environment remained essentially unchanged (Supplementary Figure S4).

Having discovered that this system was oxygen-responsive and that oxidation of the  $\text{Cu}^+$  resulted in the restoration of the acid-stabilized i-motif structure we were interested in determining whether this transition was possible in the opposite direction; i.e. whether it would be possible to reduce  $\text{Cu}^{2+}$  *in situ* to form the  $\text{Cu}^+$ -i-motif. To explore this, the well-established reaction between  $\text{Cu}^{2+}$

and sodium ascorbate where  $\text{Cu}^{2+}$  is reduced to  $\text{Cu}^+$  was used (41). 150  $\mu\text{M}$   $\text{Cu}^{2+}$  was added to the acid-stabilized i-motif and, as this concentration is below the  $[\text{Cu}^{2+}]_{50}$ , no structural change was observed using CD. Subsequent addition of 150  $\mu\text{M}$  sodium ascorbate resulted in the successful formation of the  $\text{Cu}^+$ -i-motif complex observed previously when  $\text{Cu}^+$  was added under anoxic conditions (Supplementary Figure S5). When one equivalent of sodium ascorbate was used, exposure of the sample to air led to a slow process of oxidation and return to the acid-stabilized i-motif structure (Supplementary Figure S6), while adding an excess of sodium ascorbate allowed for the prolonged maintenance of the concentration of  $\text{Cu}^+$  and the corresponding  $\text{Cu}^+$  stabilized i-motif structure.

This ability to maintain the  $\text{Cu}^+$ -i-motif structure for several hours in the open air allowed us to perform FRET-based DNA melting experiments using two equivalents of sodium ascorbate and observe the folding behavior using the dual-labeled sequence hTeloC<sub>FRET</sub> 5'-FAM-[TAA-CCC]<sub>4</sub>-TAMRA-3'. The fluorescence signal at 25°C was used to determine the fraction of the DNA that is folded in the presence of increasing concentrations of  $\text{Cu}^+$ . In good agreement with the UV and CD data, addition of  $\text{Cu}^+$  to hTeloC<sub>FRET</sub> at pH 5.5 did not affect the proportion of the population of the DNA that was folded. Further ruling out unfolding and supporting the observations that the folded conformations of the proton-stabilized and  $\text{Cu}^+$ -stabilized i-motif are similar. Conversely, at pH 7.4 addition of  $\text{Cu}^+$  results in folding of the sequence into a secondary structure that brings the two ends of the sequence into sufficient proximity for FRET to occur (Figure 2H). Additionally, using this technique we were able to determine a  $T_m$  which was calculated as the midpoint temperature of the transition from the folded to the unfolded structure. At pH 7.4 increasing  $\text{Cu}^+$  concentration led to an increase in  $T_m$  until 15 equivalents (3  $\mu\text{M}$ ) at which point it was 65°C, and after which no further change was observed. This is in agreement with the 15 equivalents of  $\text{Cu}^+$  required to fold the DNA as determined by CD. An increase in  $T_m$  was also observed at pH 5.5, however the temperature required to unfold completely is above 95°C, the limit of the instrument, therefore an accurate determination of the  $T_m$  was not possible (Supplementary Figure S7).

Having previously shown that multiple iterations of the conformational change by repeated chelation and metalation were possible using  $\text{Cu}^{2+}$  (19), we wanted to determine whether similar repeat switching was possible between the  $\text{Cu}^+$ -stabilized and the acid-stabilized i-motif structures. From a nanotechnology perspective, the potential ability to have a conformational change of the structure controlled by redox-cycling the metal was very encouraging. To truly test the versatility of this system, rather than perform repeat additions of  $\text{Cu}^+$ , we decided to perform a single addition of  $\text{Cu}^{2+}$  and reduce this *in situ* repeatedly to  $\text{Cu}^+$  using sodium ascorbate. The structural reconfiguration in response to the oxidation state of the copper was observed by monitoring the molar ellipticity at 288 nm as a function of time. Figure 2I shows the results of this experiment with three successive additions of the reducing agent successfully resulting in adoption of the  $\text{Cu}^+$ -stabilized i-motif structure, and ox-



**Figure 3.** Illustration of proposed system for the pH and copper-redox-dependent control of the structure of the i-motif forming DNA sequence hTeloC.

dation to  $\text{Cu}^{2+}$  over time similarly resulting in the return to the acid-stabilized i-motif prior to the next sodium ascorbate addition.

Having established the redox-dependent coordination of copper by the i-motif forming DNA sequence hTeloC, we hypothesized that this system could act as a continuous redox-sensitive cycle, allowing for dynamic movement between the various structural conformations adopted under the different conditions. The final step was to determine whether it was possible to convert the  $\text{Cu}^{2+}$  hairpin structure to the  $\text{Cu}^+$ -stabilized i-motif structure. As can be seen in Figure 2J, addition of 1 mM  $\text{Cu}^{2+}$  to a sample of hTeloC at pH 5.5 forms the hairpin structure and subsequent reduction to  $\text{Cu}^+$  using 150  $\mu\text{M}$  sodium ascorbate successfully forms the previously observed  $\text{Cu}^+$ -DNA i-motif, even in the presence of excess  $\text{Cu}^{2+}$ ; which is predictable due to the difference of an order of magnitude between the binding affinities of the different oxidation states of copper. Leaving the same sample in the open air over time resulted in conversion back to the  $\text{Cu}^{2+}$  stabilized hairpin structure. At last, addition of 1 mM ethylenediaminetetraacetic acid (EDTA) chelated the  $\text{Cu}^{2+}$  and the sample returned to its initial configuration as an acid-stabilized i-motif (Figure 2J). A summary of the transitions possible is conveyed in Figure 3, illustrating the proposed pH and redox sensitive control of the structural conformation of the i-motif forming DNA sequence hTeloC in the presence of copper.



This research demonstrates that  $\text{Cu}^+$  can be used to fold an i-motif forming DNA sequence into a  $\text{Cu}^+$ -stabilized i-motif structure. This process can be reversed by chelation of the metal, or by oxidation of  $\text{Cu}^+$  to  $\text{Cu}^{2+}$ . To our knowledge, this is the first example of redox-sensitive control of DNA secondary structure. This work realizes that a series of alternative conformational switches for i-motif forming DNA sequences are possible using different conditions, without changing the pH. The dynamics of this system could be applied to develop dual oxygen and pH-sensitive nanomachines, logic gates or sensors based on i-motif DNA.

#### SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

#### ACKNOWLEDGEMENTS

Research presented here was carried out on the high performance computing cluster supported by the Research and Specialist Computing Support service at UEA.

#### FUNDING

Eastern Academic Research Consortium Studentship (to M.A.S.A.); Biotechnology and Biological Sciences Research Council [BB/M00256X/1 to A.J.G., BB/L02229X/1 to Z.A.E.W.]. Funding for open access charge: University of East Anglia from Research Councils UK.

*Conflict of interest statement.* None declared.

#### REFERENCES

- Choi, J. and Majima, T. (2011) Conformational changes of non-B DNA. *Chem. Soc. Rev.*, **40**, 5893–5909.
- Gehring, K., Leroy, J.-L. and Guéron, M. (1993) A tetrameric DNA structure with protonated cytosine-cytosine base pairs. *Nature*, **363**, 561–565.
- Chen, C., Li, M., Xing, Y., Li, Y., Joedecke, C.-C., Jin, J., Yang, Z. and Liu, D. (2012) Study of pH-induced folding and unfolding kinetics of the DNA i-motif by stopped-flow circular dichroism. *Langmuir*, **28**, 17743–17748.
- Liu, D. and Balasubramanian, S. (2003) A proton-fuelled DNA nanomachine. *Angew. Chem. Int. Ed. Engl.*, **42**, 5734–5736.
- Alberti, P., Bourdoncle, A., Sacc, B., Lacroix, L. and Mergny, J.-L. (2006) DNA nanomachines and nanostructures involving quadruplexes. *Org. Biomol. Chem.*, **4**, 3383–3383.
- Dong, Y., Yang, Z. and Liu, D. (2014) DNA Nanotechnology based on i-motif structures. *Acc. Chem. Res.*, **47**, 1853–1860.
- Liu, H., Xu, Y., Li, F., Yang, Y., Wang, W., Song, Y. and Liu, D. (2007) Light-Driven conformational switch of i-motif DNA. *Angew. Chem. Int. Ed. Engl.*, **46**, 2515–2517.
- Modi, S., M.G.S., Goswami, D., Gupta, G.D., Mayor, S. and Krishnan, Y. (2009) A DNA nanomachine that maps spatial and temporal pH changes inside living cells. *Nat. Nanotechnol.*, **4**, 325–330.
- Dhaka, S., Schonhoff, J.D., Koirala, D., Yu, Z., Basu, S. and Mao, H. (2010) Coexistence of an ILPR i-motif and a partially folded structure with comparable mechanical stability revealed at the single-molecule level. *J. Am. Chem. Soc.*, **132**, 8991–8997.
- Choi, J., Kim, S., Tachikawa, T., Fujitsuka, M. and Majima, T. (2011) pH-induced intramolecular folding dynamics of i-motif DNA. *J. Am. Chem. Soc.*, **133**, 16146–16153.
- Lieblein, A.L., Buck, J., Schlepckow, K., Furtig, B. and Schwalbe, H. (2012) Time-resolved NMR spectroscopic studies of DNA i-motif folding reveal kinetic partitioning. *Angew. Chem. Int. Ed. Engl.*, **51**, 250–253.
- Gurung, S.P., Schwarz, C., Hall, J.P., Cardin, C.J. and Brazier, J.A. (2015) The importance of loop length on the stability of i-motif structures. *Chem. Commun.*, **51**, 5630–5632.
- McKim, M., Buxton, A., Johnson, C., Metz, A. and Sheardy, R.D. (2016) Loop sequence context influences the formation and stability of the i-motif for DNA oligomers of sequence (CCCXX)<sub>4</sub>, where X = A and/or T, under slightly acidic conditions. *J. Phys. Chem. B*, **120**, 7652–7661.
- Wright, E.P., Huppert, J.L. and Waller, Z.A.E. (2017) Identification of multiple genomic DNA sequences which form i-motif structures at neutral pH. *Nucleic Acids Res.*, **45**, 2951–2959.
- Wang, Z.G., Elbaz, J. and Willner, I. (2011) DNA machines: bipedal walker and stepper. *Nano Lett.*, **11**, 304–309.
- Day, H.A., Huguin, C. and Waller, Z.A.E. (2013) Silver cations fold i-motif at neutral pH. *Chem. Commun.*, **49**, 7696–7698.
- Kim, S.E., Lee, I.B., Hyeon, C. and Hong, S.C. (2014) Destabilization of i-motif by submolar concentrations of a monovalent cation. *J. Phys. Chem. B*, **118**, 4753–4760.
- Mergny, J.-L., Lacroix, L., Han, X., Leroy, J.-L. and Helene, C. (1995) Intramolecular folding of pyrimidine oligodeoxynucleotides into an i-DNA Motif. *J. Am. Chem. Soc.*, **117**, 8887–8898.
- Day, H.A., Wright, E.P., MacDonald, C.J., Gates, A.J. and Waller, Z.A.E. (2015) Reversible DNA i-motif to hairpin switching induced by copper(II) cations. *Chem. Commun.*, **51**, 14099–14102.
- Gao, J., Berden, G., Rodgers, M.T. and Oomens, J. (2016) Interaction of  $\text{Cu}^+$  with cytosine and formation of i-motif-like C-M(+)-C complexes: alkali versus coinage metals. *Phys. Chem. Chem. Phys.*, **18**, 7269–7277.
- Phan, A.T., Guéron, M. and Leroy, J.-L. (2000) The solution structure and internal motions of a fragment of the cytidine-rich strand of the human telomere. *J. Mol. Biol.*, **299**, 123–144.
- Hanwell, M.D., Curtis, D.E., Lonie, D.C., Vandermeersch, T., Zurek, E. and Hutchison, G.R. (2012) Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminform.*, **4**, 17.
- Krepl, M., Zgarbová, M., Stadlbauer, P., Otyepka, M., Banáš, P., Koča, J., Cheatham, T.E., Jurečka, P. and Šponer, J. (2012) Reference simulations of noncanonical nucleic acids with Different  $\chi$  Variants of the AMBER Force Field: Quadruplex DNA, Quadruplex RNA, and Z-DNA. *J. Chem. Theory Comput.*, **8**, 2506–2520.
- Zgarbová, M., Luque, F.J., Šponer, J., Cheatham, T.E., Otyepka, M. and Jurečka, P. (2013) Toward improved description of DNA Backbone: Revisiting epsilon and zeta torsion force field parameters. *J. Chem. Theory Comput.*, **9**, 2339–2354.
- Zgarbová, M., Šponer, J., Otyepka, M., Cheatham, T.E., Galindo-Murillo, R. and Jurečka, P. (2015) Refinement of the Sugar-Phosphate backbone torsion beta for AMBER force fields improves the description of Z- and B-DNA. *J. Chem. Theory Comput.*, **11**, 5723–5736.
- Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W. and Klein, M.L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, **79**, 926–935.
- Li, P., Song, L.F. and Merz, K.M. (2015) Systematic parameterization of monovalent ions employing the nonbonded model. *J. Chem. Theory Comput.*, **11**, 1645–1657.
- Abraham, M.J., Murtola, T., Schulz, R., Páll, S., Smith, J.C., Hess, B. and Lindahl, E. (2015) Gromacs: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, **1–2**, 19–25.
- Rezáč, J. and Hobza, P. (2012) Advanced corrections of hydrogen bonding and dispersion for semiempirical quantum mechanical methods. *J. Chem. Theory Comput.*, **8**, 141–151.
- MOPAC2016 and Stewart, J.J.P. (2016) *Stewart Computational Chemistry*. Colorado Springs, CO, USA. [HTTP://OpenMOPAC.net](http://OpenMOPAC.net).
- Valiev, M., Bylaska, E.J., Govind, N., Kowalski, K., Straatsma, T.P., Van Dam, H.J.J., Wang, D., Nieplocha, J., Apra, E., Windus, T.L. et al. (2010) NWChem: a comprehensive and scalable open-source solution for large scale molecular simulations. *Comput. Phys. Commun.*, **181**, 1477–1489.
- Yuan, X., Pham, A.N., Xing, G., Rose, A.L. and Waite, T.D. (2012) Effects of pH, chloride, and bicarbonate on  $\text{Cu}(\text{I})$  Oxidation kinetics at circumneutral pH. *Environ. Sci. Technol.*, **46**, 1527–1535.

33. Jhaveri, A.S. and Sharma, M.M. (1967) Kinetics of absorption of oxygen in aqueous solutions of cuprous chloride. *Chem. Eng. Sci.*, **22**, 1–6.
34. Phan, A.T. (2002) Human telomeric DNA: G-quadruplex, i-motif and Watson-Crick double helix. *Nucleic Acids Res.*, **30**, 4618–4625.
35. Mergny, J.-L. and Lacroix, L. (2009) UV Melting of G-Quadruplexes. *Current Protocols in Nucleic Acid Chemistry*. John Wiley & Sons, Vol. **37**, pp. 17.1.1–17.1.15.
36. Mergny, J.-L., Li, J., Lacroix, L., Amrane, S. and Chaires, J.B. (2005) Thermal difference spectra: a specific signature for nucleic acid structures. *Nucleic Acids Res.*, **33**, e138–e138.
37. Benabou, S., Aviñó, A., Eriñja, R., González, C. and Gargallo, R. (2014) Fundamental aspects of the nucleic acid i-motif structures. *RSC Adv.*, **4**, 26956–26956.
38. Kypr, J., Kejnovská, I., Renčuk, D. and Vorlíčková, M. (2009) Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res.*, **37**, 1713–1725.
39. Smith, D.W. (1977) Ionic hydration enthalpies. *J. Chem. Educ.*, **54**, 540–540.
40. Dai, J., Ambrus, A., Hurley, L.H. and Yang, D. (2009) A direct and nondestructive approach to determine the folding structure of the I-motif DNA secondary structure by NMR. *J. Am. Chem. Soc.*, **131**, 6102–6104.
41. Liang, L. and Astruc, D. (2011) The copper(I)-catalyzed alkyne-azide cycloaddition (CuAAC) “click” reaction and its applications: an overview. *Coord. Chem. Rev.*, **255**, 2933–2945.



# Common G-Quadruplex Binding Agents Found to Interact With i-Motif-Forming DNA: Unexpected Multi-Target-Directed Compounds

Alessia Pagano<sup>1†</sup>, Nunzia Iaccarino<sup>1†</sup>, Mahmoud A. S. Abdelhamid<sup>2</sup>, Diego Brancaccio<sup>1</sup>, Emanuele U. Garzarella<sup>1</sup>, Anna Di Porzio<sup>1</sup>, Ettore Novellino<sup>1</sup>, Zoë A. E. Waller<sup>2,3</sup>, Bruno Pagano<sup>1</sup>, Jussara Amato<sup>1\*</sup> and Antonio Randazzo<sup>1\*</sup>

## OPEN ACCESS

### Edited by:

Stefano Alcaro,  
Consejo Superior de Investigaciones  
Científicas (CSIC), Spain  
Università degli Studi Magna Graecia di  
Catanzaro, Italy

### Reviewed by:

Juan Carlos Morales,  
Consejo Superior de Investigaciones  
Científicas (CSIC), Spain  
Janez Plavec,  
National Institute of Chemistry  
Slovenia, Slovenia  
Claudia Sissi,  
Università degli Studi di Padova, Italy

### \*Correspondence:

Jussara Amato  
jussara.amato@unina.it  
Antonio Randazzo  
antonio.randazzo@unina.it

<sup>†</sup> These authors have contributed  
equally to this work.

### Specialty section:

This article was submitted to  
Medicinal and Pharmaceutical  
Chemistry,  
a section of the journal  
Frontiers in Chemistry

**Received:** 27 April 2018

**Accepted:** 22 June 2018

**Published:** 24 July 2018

### Citation:

Pagano A, Iaccarino N,  
Abdelhamid MAS, Brancaccio D,  
Garzarella EU, Di Porzio A, Novellino E,  
Waller ZAE, Pagano B, Amato J and  
Randazzo A (2018) Common  
G-Quadruplex Binding Agents Found  
to Interact With i-Motif-Forming DNA:  
Unexpected Multi-Target-Directed  
Compounds. *Front. Chem.* 6:281.  
doi: 10.3389/fchem.2018.00281

<sup>1</sup> Department of Pharmacy, University of Naples Federico II, Naples, Italy, <sup>2</sup> School of Pharmacy, University of East Anglia, Norwich Research Park, Norwich, United Kingdom, <sup>3</sup> Centre for Molecular and Structural Biochemistry, University of East Anglia, Norwich Research Park, Norwich, United Kingdom

G-quadruplex (G4) and i-motif (iM) are four-stranded non-canonical nucleic acid structural arrangements. Recent evidences suggest that these DNA structures exist in living cells and could be involved in several cancer-related processes, thus representing an attractive target for anticancer drug discovery. Efforts toward the development of G4 targeting compounds have led to a number of effective bioactive ligands. Herein, employing several biophysical methodologies, we studied the ability of some well-known G4 ligands to interact with iM-forming DNA. The data showed that the investigated compounds are actually able to interact with both DNA *in vitro*, thus acting *de facto* as multi-target-directed agents. Interestingly, while all the compounds stabilize the G4, some of them significantly reduce the stability of the iM. The present study highlights the importance, when studying G4-targeting compounds, of evaluating also their behavior toward the i-motif counterpart.

**Keywords:** G-quadruplex, i-motif, Berberine, BRACO-19, Mitoxantrone, Phen-DC3, Pyridostatin, RHPS4

## INTRODUCTION

GC-rich nucleic acids are able to form a variety of non-canonical secondary structures (Zhao et al., 2010; Cerofolini et al., 2014). The best studied of these are G-quadruplexes (G4s), four stranded alternative nucleic acid secondary structures formed from guanine-rich DNA or RNA composed of stacked tetrads of guanines formed by Hoogsteen hydrogen bonding (Burge et al., 2006). Sequences which can form G4s are prevalent within regulatory regions of the genome, particularly within the promoter region of genes (Huppert and Balasubramanian, 2007; Chambers et al., 2015; Bedrat et al., 2016). Good evidence has been provided to support the hypothesis that G4s exist in human cells (Biffi et al., 2013), play a role in human diseases (Haeusler et al., 2014; Maizels, 2015) and can be targeted with ligands to modulate biological functions (Siddiqui-Jain et al., 2002; Lam et al., 2013; Zizza et al., 2016).

More recently, increasing interest is being paid to the i-motif (iM) structure, another four stranded structure which can form in sequences rich in cytosine, composed of two intercalated hairpins, stabilized by hemi-protonated cytosine-cytosine<sup>+</sup> (C-C<sup>+</sup>) base pairs (Gehring et al., 1993). Putative iM forming sequences also occur throughout the genome (Wright et al., 2016b;



Fleming et al., 2017), typically opposing regions which can form G4s, though the sequence requirements for stable formation are somewhat different. Studies on the iM were previously limited based on the assumption that because they are stabilized in slightly acidic conditions they were not physiologically relevant, despite a solid foundation of data indicating that these structures are detectable at neutral pH *in vitro* (Mergny et al., 1995). Over the years, this assumption has been challenged with examples of iMs which can form at neutral pH, at low temperature (Zhou et al., 2010), under conditions of negative superhelicity (Sun and Hurley, 2009), and molecular crowding (Rajendran et al., 2010) conditions. Further examples of sequences which are naturally stable at neutral pH have been found in the genome, initially by investigating sequences which oppose G4s (Brazier et al., 2012), but multiple other examples have followed (Wright et al., 2016b; Fleming et al., 2017; Mir et al., 2017). Just recently further evidence for i-motif formation *in vivo* has been provided by in cell NMR experiments (Dzatko et al., 2018) and the discovery of an antibody that binds i-motif specifically in the nuclei of human cells (Zeraati et al., 2018). The C-rich regions of genomes are of particular interest because cytosine forms part of the basis for epigenetic regulation, and epigenetic modification of cytosine has been found to alter the stability of iM (Bhavsar-Jog et al., 2014; Xu et al., 2015; Wright et al., 2017). Moreover, ligands which bind and stabilize iM have been shown to modulate biological functions (Amato et al., 2014b). For example, stabilization of the iM forming sequence in the human telomere was found to inhibit telomerase activity and interfere with telomere biology (Li et al., 2006; Chen et al., 2012); stabilization of iM forming sequence in the promoter region of BCL2 was found to cause an increase in gene expression (Kang et al., 2014; Kendrick et al., 2014) and an iM interacting compound was found to downregulate PDGFR- $\beta$  promoter activity (Brown et al., 2017).

As it appears that formation of iM and/or G4 structures could incite different biological outcomes, it is important to understand the potential structures a compound is able to interact with. In contrast to the hundreds of G4 binding ligands (Pagano et al., 2007, 2010, 2015; Di Leva et al., 2013; Li et al., 2013; Amato et al., 2014a, 2018), there are comparatively very few iM binding compounds reported in the literature (Day et al., 2014). Some ligands which were described to bind G4 have also been found to bind iM (Fedoroff et al., 2000; Wright et al., 2016a; Xu et al., 2016), so we decided to assess and compare the capability to interact with iM-forming DNA of several known bioactive G4 binding agents: Berberine (1) (Franceschin et al., 2006), BRACO-19 (2) (Gowan et al., 2002), Mitoxantrone (3) (Huang et al., 2007), Phen-DC3 (4) (De Cian et al., 2007a), Pyridostatin (5) (Rodriguez et al., 2008), and RHPS4 (6) (Izbicka et al., 1999) (Figure 1). The interaction of these compounds with G4- and iM-forming sequences were investigated *in vitro* in different experimental conditions employing several biophysical methodologies (Pagano et al., 2012) (Figure 2). The data unequivocally demonstrate that, even if in different ways, actually these molecules interact with both DNA, thus acting *de facto* as multi-target-directed compounds.

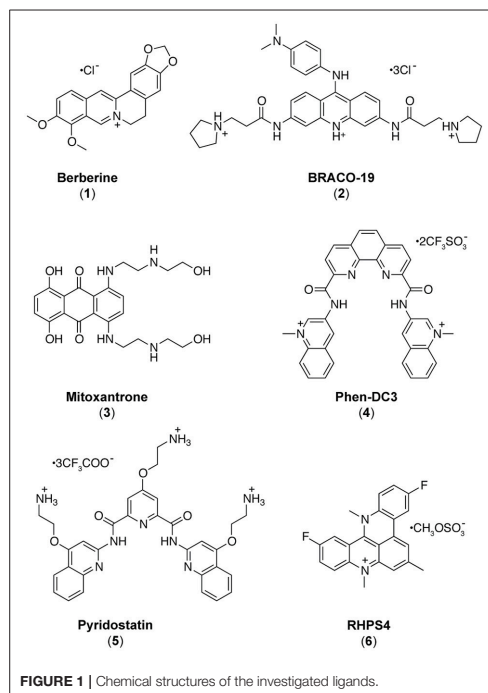
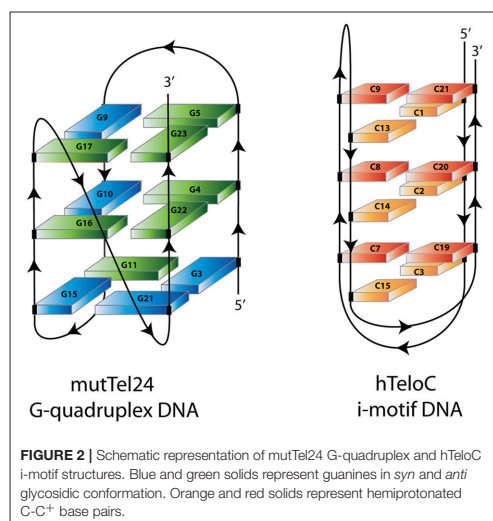


FIGURE 1 | Chemical structures of the investigated ligands.

## MATERIALS AND METHODS

### Oligonucleotide Synthesis and Sample Preparation

An ABI 394 DNA/RNA synthesizer (Applied Biosystem) was employed to prepare DNA sequences at 5- $\mu$ mol scale by using standard  $\beta$ -cyanoethylphosphoramidite solid phase chemistry. The subsequent detachment of DNA from support and its deprotection were carried out by means of an aqueous solution of concentrated ammonia at 55°C for 12 h. The filtrates and the washings were combined and concentrated under reduced pressure, solubilized in water, and then purified by high-performance liquid chromatography (HPLC) equipped with a Nucleogel SAX column (Macherey-Nagel, 1000-8/46). Two buffers were employed for the purification: buffer A, consisting of a 20 mM  $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$  aqueous solution (pH 7.0) and containing 20% (v/v)  $\text{CH}_3\text{CN}$ , and buffer B, consisting of 1 M KCl, 20 mM  $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$  aqueous solution (pH 7.0), containing 20% (v/v)  $\text{CH}_3\text{CN}$ , combined in a 30 min linear gradient going from 0 to 100% B with a flow rate of 1 mL/min. The purified fractions of the oligomers were then desalted by using C-18 cartridges (Sep-pak). The purity of the isolated oligomer was evaluated by NMR and it turned out to be higher than 98%. In particular, the



following oligonucleotides were employed for the experiments: d[CCCT(AACCCT)<sub>3</sub>] (hTeloC) and d[(TT(GGGTTA)<sub>3</sub>GGGA)] (mutTel24). The oligonucleotide concentrations were established by measuring the UV absorption at 90°C taking into account the molar extinction coefficient values  $\epsilon$  ( $\lambda = 260$  nm) determined by the nearest neighbor model (Cantor et al., 1970). hTeloC and mutTel24 were dissolved in 10 mM sodium phosphate buffer and 10 mM potassium phosphate buffer, respectively, at different pH values before establishing the experimental conditions to be used (pH 4.3 and pH 5.7). DNA samples were heated at 90°C for 5 min, and then gradually cooled to room temperature overnight.

### Circular Dichroism Spectroscopy

Circular dichroism (CD) experiments were recorded on a Jasco J-815 spectropolarimeter equipped with a PTC-423S/15 Peltier temperature controller. Each spectrum was recorded in the 220–360 nm wavelength range, averaged over three scans and subtracted of the buffer baseline. The scan rate was set to 100 nm/min, with a 1 s response time, and 1 nm bandwidth. Spectra were analyzed using Origin 7.0 software. CD experiments (spectra and melting) were performed using 10–15  $\mu$ M oligonucleotide concentration, in the absence and presence of 5 molar equivalents of ligands (10 mM in DMSO). CD melting was performed at 1°C/min heating rate in the 5–90 and 20–100°C temperature range for hTeloC and mutTel24, respectively. Changes of CD signal were followed at the wavelengths of the maximum CD intensity, 288 and 290 nm for hTeloC and mutTel24, respectively. The melting temperatures ( $T_{1/2}$ ) were mathematically calculated by using the curve fitting function in Origin 7.0 software.  $\Delta T_{1/2}$  values represent the difference between the melting temperature of the DNA with and without ligands.

### UV-Melting

A JASCO V-730 UV-visible spectrophotometer equipped with a Peltier temperature controller was employed to perform the UV thermal denaturation experiments. The oligonucleotide concentrations were 10  $\mu$ M for both hTeloC and mutTel24 DNA in the appropriate buffer, as indicated above. Experiments were performed by following changes of UV signal at 295 nm, at a heating rate of 1°C/min, in the temperature ranges of 5–100 and 20–100°C for hTeloC and mutTel24, respectively. The melting temperatures ( $T_{1/2}$ ) were mathematically calculated by using the curve fitting function in Origin 7.0 software.  $\Delta T_{1/2}$  values represent the difference between the melting temperature of the DNA with and without ligands.

### Nuclear Magnetic Resonance Experiments

A 700 MHz Varian Unity INOVA spectrometer was employed to perform the NMR experiments. One-dimensional proton spectra were recorded at 7°C using pulsed-field gradient DPGFSE for water suppression. All DNA samples were prepared at 0.2 mM strand concentration in 0.22 mL (H<sub>2</sub>O/D<sub>2</sub>O 9:1) buffer solution. DNA/ligand mixtures were obtained by adding aliquots of a stock solution of the six ligands in DMSO-d<sub>6</sub> directly to the DNA solution inside the NMR tube (Randazzo et al., 2002; Amato et al., 2017). NMR data were processed using the iNMR software (www.inmr.net).

### FRET and FRET-Melting

A FP-8300 spectrofluorometer (Jasco) equipped with a Peltier temperature controller accessory (Jasco PCT-818) was employed to carry out FRET experiments. The dual-labeled oligonucleotides corresponding to the G4 forming sequence 5'-FAM-d(GGG[TTAGGG]<sub>3</sub>)-TAMRA-3' (G4-F21T) (Amato et al., 2016; Salvati et al., 2017) and the iM forming sequence 5'-FAM-d(TAACC)<sub>4</sub>-TAMRA-3' (iM-F24T) were used. Such sequences are characterized by the presence of the donor fluorophore FAM (6-carboxyfluorescein) and the acceptor fluorophore TAMRA (6-carboxytetramethylrhodamine) that are covalently bound at 5'- and 3'-ends, respectively. Labeled oligonucleotides were purchased from Biomers (Germany). G4-F21T and iM-F24T were prepared at 1  $\mu$ M concentration in 10 mM potassium phosphate buffer and 10 mM sodium phosphate buffer, respectively. Samples were annealed in a hot water bath at 90°C for 2 min, and then cooled to room temperature overnight. FRET measurements were performed both in the absence and presence of 5 molar equivalents of compounds 1–6. The final concentration of G4-F21T and iM-F24T was 0.1  $\mu$ M. A sealed quartz cuvette with a path length of 1 cm was used. FRET spectra were acquired before (at 5 and 20°C for iM and G4, respectively) and after (at 90°C) melting assay. The dual-labeled oligonucleotides were excited at 492 nm, and emission spectra were recorded between 500 and 650 nm using 100 nm/s scan speed. Excitation and emission slit widths were both set at 5 nm. FRET-melting experiments were performed by setting the excitation wavelength at 492 nm and the detection wavelength at 522 nm. The emission intensity of FAM was then normalized between 0 and 1. Data analysis was carried out using Origin 7.0 software.

## Fluorescent Intercalator Displacement (FID) Assay

For the FID experiments, oligonucleotides d[(TAACCC)<sub>4</sub>] (hTeloC<sub>FID</sub>) and d[(TTGGGG(TTAGGG)<sub>3</sub>A)] (mutTel24<sub>FID</sub>) were purchased from Eurogentec and then purified via HPLC. Solid DNA samples were initially dissolved as a 1 mM stock solution in MilliQ water. 10 mM stock solutions of the candidate ligands were prepared in DMSO. Further dilutions were carried out in buffer: 10 mM NaH<sub>2</sub>PO<sub>4</sub> for hTeloC<sub>FID</sub> and 10 mM KH<sub>2</sub>PO<sub>4</sub> for mutTel24<sub>FID</sub>. DNA samples were thermally annealed at 90 μM in the respective buffers in an Applied Biosystems Veriti 96 well thermal cycler by holding at 95°C for 5 min and cooling at a rate of 1°C/min to 20°C. FID experiments were performed on a BMG CLARIOstar plate reader using Corning 96-Well Solid Black Flat Bottom plates. A 10 mM stock solution of thiazole orange (TO) was prepared in DMSO and diluted to 2 μM in the appropriate buffer for either hTeloC<sub>FID</sub> or mutTel24<sub>FID</sub>. Ninety microliters of the 2 μM TO solution were added to each well and fluorescence emission at 450 nm measured with excitation at 430 nm; this was normalized to 0% representing background fluorescence. One microliter of DNA was then added, shaken using double orbital shaking at 700 rpm in the plate reader for 15 s, and allowed to equilibrate for 15 min. After equilibration, fluorescence emission was measured as before, and normalized to 100% representing maximal fluorescence enhancement from the TO probe binding to the DNA secondary structure. 0.9 μL aliquots of ligand were titrated into each well (in triplicate) and measured as before. Fluorescence measurements after ligand addition were normalized between the 0 and 100% levels determined per the respective well. Percentage TO displacement was calculated as the difference between the normalized 100% fluorescence level and the normalized fluorescence measured after each ligand addition. The concentration for each ligand at which 50% of the TO was displaced (DC<sub>50</sub>) was calculated by using Origin data analysis software to plot percentage TO displacement against ligand concentration. These data were fitted with dose-response curves and the equations of the curves were solved for  $y = 50$  to give the DC<sub>50</sub> values.

## RESULTS

Most of the investigations reported in the literature dealing with the determination of iM structures in solution have been accomplished in sodium buffer and under acidic conditions, generally at pH values down to 4.3 (Gehring et al., 1993; Gallego et al., 1997; Malliavin et al., 2003). This is because the cytidine pK<sub>a</sub> is about 4.2, and the use of low pH values guarantees to obtain stable hemi-protonated C-C<sup>+</sup> pairs. However, these conditions are far away from physiological. This may also have consequences on the study of the interaction between these target molecules and potential ligands, which may be differently protonated with respect to their state under physiological pH. Therefore, in order to find experimental conditions as close to physiological as possible, the behavior of the telomeric iM-forming sequence (hTeloC) in 10 mM sodium buffer at different pH values was

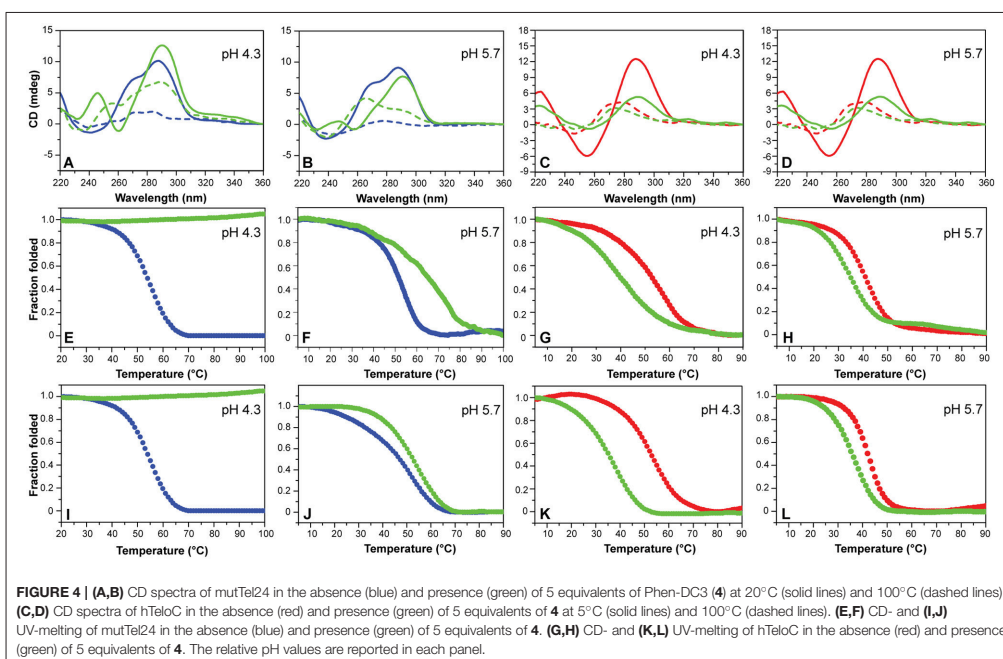
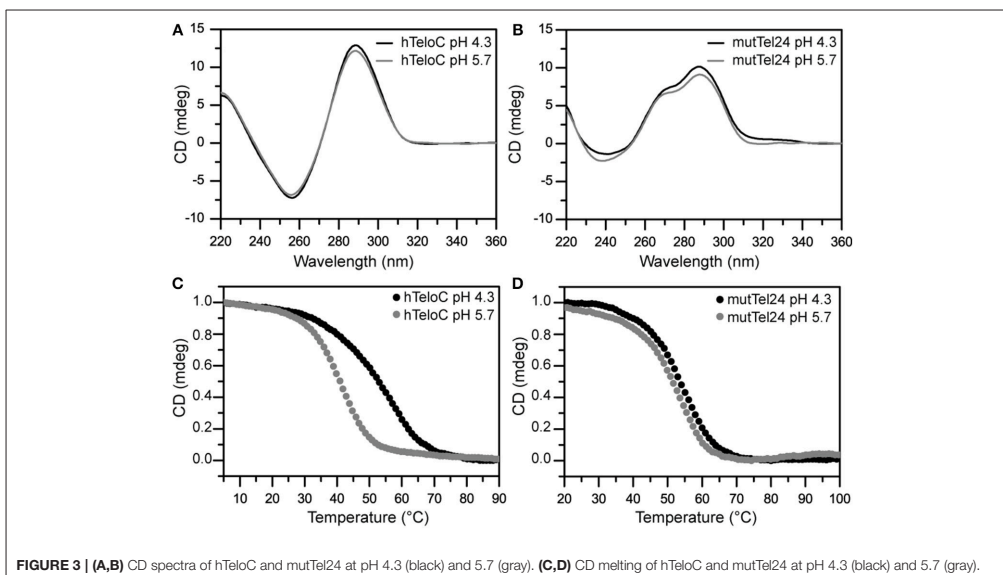
investigated by 1D <sup>1</sup>H-NMR and CD spectroscopies (Figures S1, S2, Supplementary Material). In particular, the pH range 4.3–7.0 was explored. NMR and CD spectra clearly indicated that the hTeloC sequence turned out to be folded into an iM structure only at pH values lower than 5.7, while above this value the iM structure is in equilibrium with the unfolded species. Therefore, we decided to perform our studies at the two boundaries pH values, namely 4.3 and 5.7. Since the main aim of this investigation was to evaluate the interaction of some known G4 ligands with an iM-forming DNA, experiments on the human telomeric G4 (mutTel24) were performed in parallel for comparison.

## Circular Dichroism Studies

The structures adopted by hTeloC and mutTel24 were first investigated by circular dichroism (CD) spectroscopy in the absence of ligands. At both pH 4.3 and 5.7, hTeloC showed almost superimposable CD spectra having a positive band at 288 nm and a negative one at around 260 nm (Figure 3A). These bands are characteristic of an iM folding topology (Guo et al., 2008). mutTel24 also displayed almost superimposable CD spectra at both pH values. These spectra were characterized by two positive bands at around 290 and 270 nm, and a negative one at around 240 nm (Figure 3B). These bands are perfectly superimposable to those observed for the same molecule at pH 7.0, thus indicating the presence of the expected hybrid [3+1] G4 structure (hybrid-1) as the major conformation under acidic conditions (Karsisiotis et al., 2011; Gray et al., 2014).

CD experiments were also performed to examine the potential of compounds 1–6 to alter the native folding topology of the two investigated DNA structures both at pH 4.3 and 5.7. DNA/ligand mixtures were obtained by adding 5 molar equivalents of compound to the folded G4 and iM structures so as to have an excess with respect to potential binding sites. In the case of mutTel24, regardless of the pH, Berberine (1), BRACO-19 (2), Phen-DC3 (4) and RHPS4 (6) induced a significant change in the CD spectrum of the G4 structure (Figures 4A,B and Figures S3, S4, Supplementary Material). In particular, the loss of the band at 270 nm followed by an intensity's increase of the band at 290 nm suggested a conformational change of the G4 topology from the hybrid to the antiparallel conformation (Masiero et al., 2010; Randazzo et al., 2013). Conversely, Mitoxantrone (3) and Pyridostatin (5) did not produce any measurable conformational change of the G4 structure, even if a decrease of the band at 290 nm is observed upon their addition.

As for the iM structure, at pH 4.3, Mitoxantrone (3), Phen-DC3 (4), and RHPS4 (6) induced a significant hypochromic shift of the positive band at 288 nm, with more marked effects observed for 3 and 4. Conversely, Berberine (1), BRACO-19 (2), and Pyridostatin (5) caused a hyperchromic shift of the band at 288 nm at this pH (Figure 4C and Figure S5, Supplementary Material). On the other hand, at pH 5.7 all the compounds have been shown to induce a hypochromic effect of the band at 288 nm that turned out to be particularly marked in the case of 2, 3, and 4 (Figure 4D and Figure S6, Supplementary Material). These data suggest that some interaction takes place and that, in some cases, the molecules seem to induce the unfolding of the structure.



## CD-Melting Analysis

CD-melting experiments were employed to evaluate the thermal stability of iM and G4 structures adopted by hTeloC and mutTel24 sequences, respectively, under the two experimental conditions used (pH 4.3 and 5.7). The melting temperatures ( $T_{1/2}$ ) of the iM structure were found to be 52.8 and 40.6°C at pH 4.3 and 5.7, respectively (Figure 3C). The lower thermal stability of the iM structure observed at pH 5.7 can be ascribed to the lower extent of protonation of the cytosines. On the other hand, only a very small variation in  $T_{1/2}$  values was observed for the G4 structure between pH 4.3 and 5.7 (53.5 and 51.1°C, respectively, Figure 3D).

Then, the effect of the investigated compounds on the stability of the DNA secondary structures was evaluated by measuring the ligand-induced change in the melting temperature ( $\Delta T_{1/2}$ ) of G4 and iM at pH 4.3 and 5.7. CD-melting curves of DNA in the absence and presence of each compound were obtained by following the variations of the intensity of CD signal at the wavelengths of 290 and 288 nm for mutTel24 and hTeloC, respectively (Figure 4 and Figures S7–S10, Supplementary Material). Very intriguingly, results of these experiments (Table 1) clearly indicated that 1–6 exert a different effect on iM compared to G4. As expected, all the tested compounds were able to thermally stabilize the G4 structure at both pH values, even if to a different extent (Figures 4E,F, and Figures S7, S8, Supplementary Material). On the contrary, regardless of pH, Berberine (1), Pyridostatin (5), and RHPS4 (6) did not show a remarkable influence on the iM thermal stability, while BRACO-19 (2), Mitoxantrone (3), and Phen-DC3 (4) significantly decreased it (Figures 4G,H and Figures S9, S10, Supplementary Material). These results are in agreement with those obtained from the CD spectra analysis.

## UV-Melting Analysis

The effect of compounds 1–6 on the stability of the G4 and iM structures was also investigated by UV-melting experiments. As for CD-melting studies, the ligand-induced changes in the melting temperature ( $\Delta T_{1/2}$ ) of the two DNA structures were obtained by recording UV-melting experiments in the absence

and presence of each compound at both pH 4.3 and 5.7. UV-melting curves were acquired by following the change in UV signal intensity at 295 nm for both mutTel24 and hTeloC (Figures 4I–L, and Figures S11–S14, Supplementary Material). Results of these experiments (Table 1) are consistent with CD-melting ones, and denote, once again, a different behavior for the investigated compounds toward the iM and G4 DNA structures.

## Nuclear Magnetic Resonance Studies

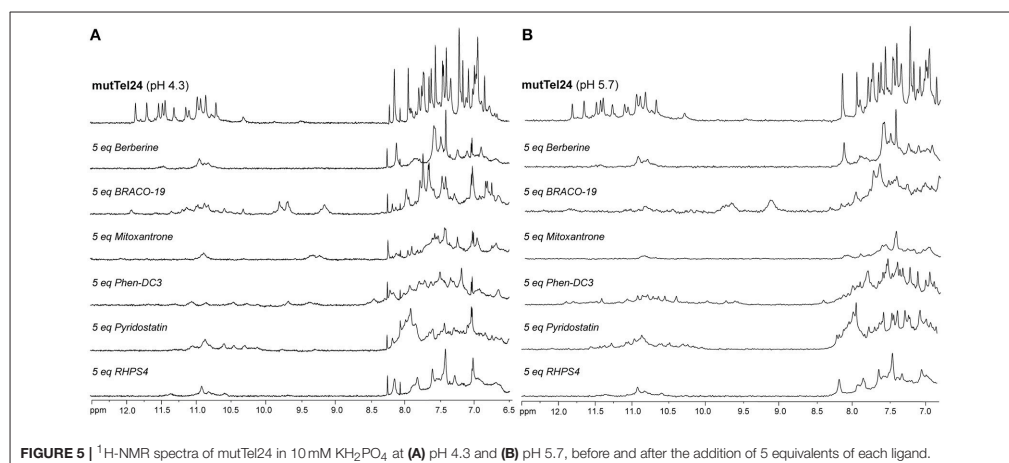
NMR spectroscopy was employed in order to obtain structural information about the DNA interaction of the six compounds. Also in this case, 1D <sup>1</sup>H-NMR spectra were recorded at pH 4.3 and 5.7. Under the experimental conditions used, the mutTel24 sequence forms a single G4 conformation characterized by 12 well-resolved imino proton peaks, corresponding to the 12 guanines involved in the three G-tetrad planes (Luu et al., 2006). On the other hand, hTeloC folds in an i-motif structure characterized by 3 well-resolved imino proton peaks that correspond to the 6 intercalated C-C<sup>+</sup> pairs (Phan et al., 2000).

The imino and aromatic proton regions of mutTel24 and hTeloC in the absence and presence of 5 equivalents of each compound are shown in Figures 5, 6, respectively. Regardless of the pH, both imino and aromatic proton signals of mutTel24 turned out to be significantly affected by the addition of the ligands (Figure 5). On the other hand, addition of compounds to the iM structure led to different results. At pH 4.3, the main changes were observed for Berberine (1) and RHPS4 (6) in both aromatic and imino regions, while little changes could be observed for BRACO-19 (2) and Mitoxantrone (3). Instead, Phen-DC3 (4) caused a general decrease of signal intensities (Figure 6A). Very little changes could be observed in the spectrum of iM upon addition of Pyridostatin (5). At pH 5.7, changes were observed for Berberine (1) BRACO-19 (2), Mitoxantrone (3), and RHPS4 (6) (Figure 6B), with BRACO-19 affecting the most the NMR spectrum of the iM structure. Interestingly, as for the experiment at pH 4.3, PhenDC3 (4) and Pyridostatin (5) caused a general decrease of the signal intensities.

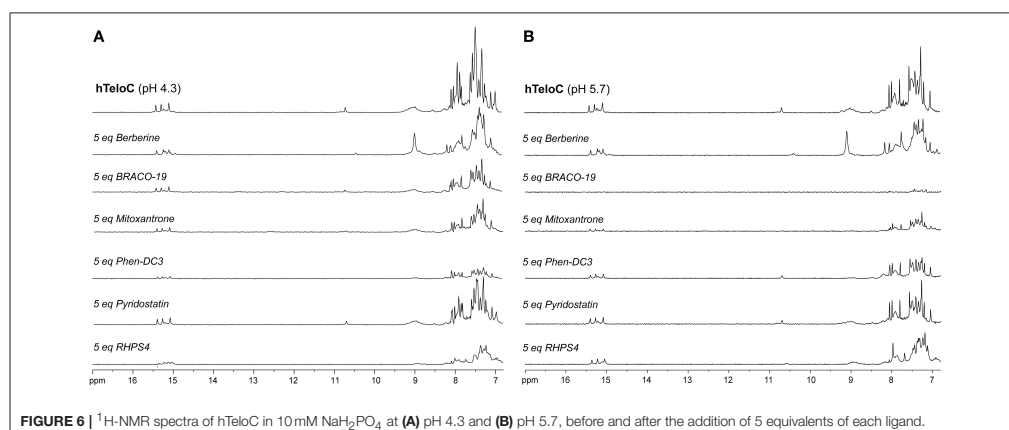
**TABLE 1** | Ligand-induced thermal stabilization of hTeloC and mutTel24 DNA measured by CD and UV melting experiments.

	$\Delta T_{1/2}$ (°C)* CD melting				$\Delta T_{1/2}$ (°C)* UV melting			
	hTeloC		mutTel24		hTeloC		mutTel24	
	pH 4.3	pH 5.7	pH 4.3	pH 5.7	pH 4.3	pH 5.7	pH 4.3	pH 5.7
Berberine	-2.5	-0.8	+13.4	+12.4	+1.2	+0.9	+11.7	+15.3
BRACO-19	-13.4	-9.2	+12.4	+8.9	-17.6	-6.5	ND	ND
Mitoxantrone	-4.8	-9.9	+7.7	+4.5	-16.6	-1.0	+4.1	+0.6
Phen-DC3	-13.4	-6.8	ND	+14.4	-17.1	-6.3	ND	+5.3
Pyridostatin	-2.8	+0.8	+12.9	+8.8	+1.1	+0.2	+9.3	+5.5
RHPS4	-1.0	-0.3	+22.0	+20.6	+0.7	+0.1	ND	ND

\* $\Delta T_{1/2} = T_{1/2}(\text{DNA}+\text{ligand}) - T_{1/2}(\text{DNA})$ . All experiments were performed in duplicate, and  $\Delta T_{1/2}$  values are reported as the mean. Errors were  $\pm 0.5^\circ\text{C}$ . ND, not determined.



**FIGURE 5** |  $^1\text{H-NMR}$  spectra of mutTel24 in 10 mM  $\text{KH}_2\text{PO}_4$  at (A) pH 4.3 and (B) pH 5.7, before and after the addition of 5 equivalents of each ligand.



**FIGURE 6** |  $^1\text{H-NMR}$  spectra of hTeloC in 10 mM  $\text{NaH}_2\text{PO}_4$  at (A) pH 4.3 and (B) pH 5.7, before and after the addition of 5 equivalents of each ligand.

### FRET and FRET-Melting Studies

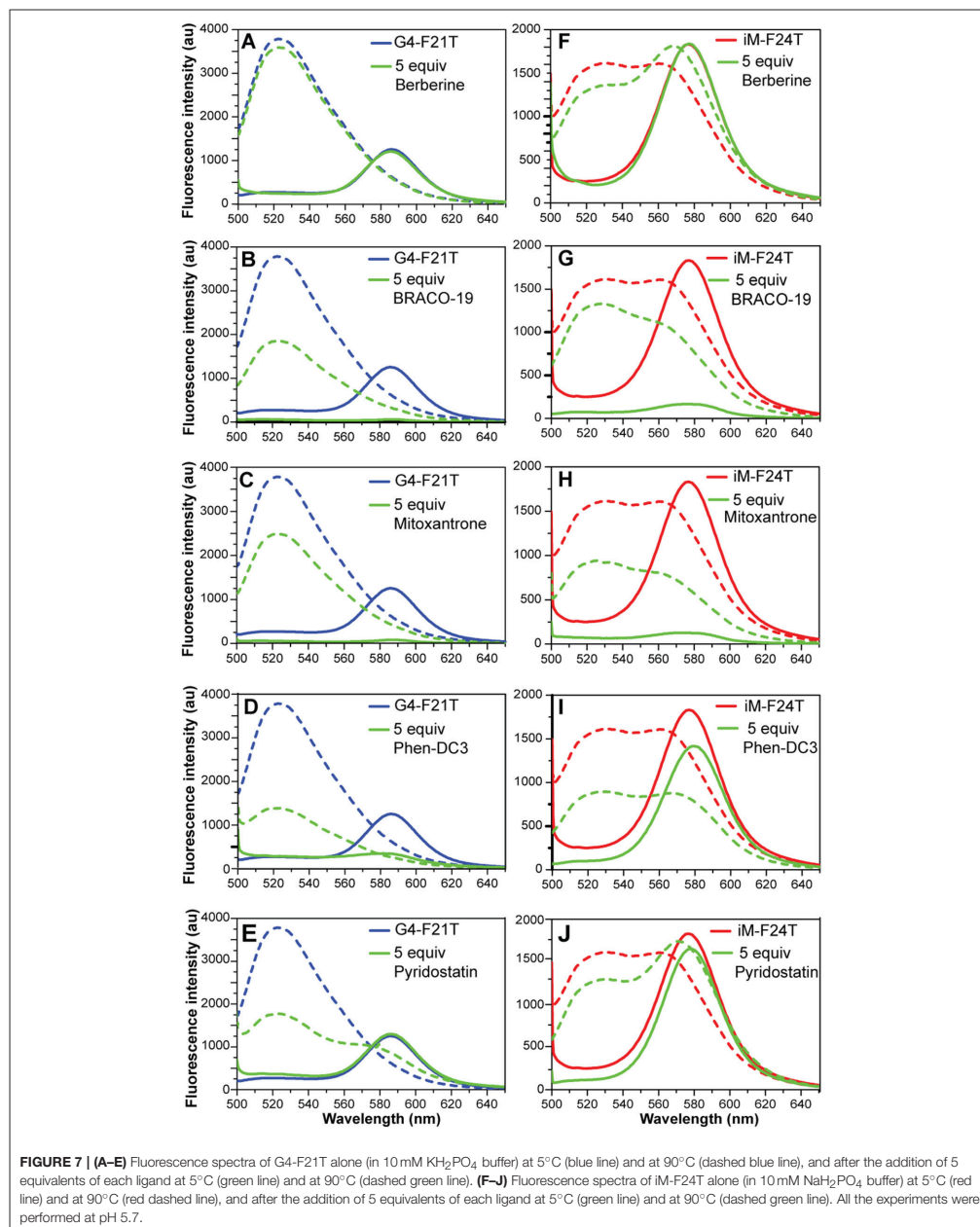
Dual labeled (FAM/TAMRA) human telomeric sequences G4-F21T and iM-F24T, which are able to form G4 and iM structures, respectively, were used (see Experimental section). Experiments were performed only at pH 5.7, because the fluorescence of the FAM is not stable at pH 4.3. In order to verify that 1-6 did not interfere in the FAM emission spectrum, the fluorescence spectrum of each compound was recorded by exciting at 492 nm and collecting its emission spectrum between 500 and 650 nm. Unfortunately, the emission spectrum of RHPS4 (6) was found to overlap with FAM, and, therefore, it was not used in these experiments.

FRET-melting assays were then performed to further investigate the compound-induced effects on the iM and G4

thermal stabilities. In agreement with CD- and UV-melting experiments, all the tested compounds induced a thermal stabilization of G4-F21T (Figure S15, Supplementary Material). Conversely, in the case of iM-F24T, FRET melting data did not agree at all with CD and UV melting results. In fact, all ligands, except Berberine, showed a significant thermal stabilization of the iM structure (Figure S16, Supplementary Material).

FRET spectra of FAM/TAMRA-modified oligonucleotides (G4-F21T and iM-F24T) in the absence and presence of each compound (at 1:5 DNA/ligand ratio) were also analyzed (Figure 7). The results show that, in both cases, BRACO-19 (2), Mitoxantrone (3), Phen-DC3 (4), and Pyridostatin (5) caused a significant decrease of the band intensity at 580 nm, thus suggesting that actually they could interact with the probes.





## Fluorescent Intercalator Displacement (FID) Assay

FID assays are long established and have been used extensively to determine relative binding affinities for ligands with duplex DNA (Boger et al., 2000). They have recently been adapted and validated for use with both G4 and iM structures (Monchaud et al., 2006; Sheng et al., 2017). The assay relies on a light-up fluorescent probe, in this case thiazole orange (TO), which binds to the structure of interest and can be competitively displaced by candidate ligands thus enabling the determination of their relative binding affinity to the structure under examination. Here, aliquots of ligands 1-5 were titrated in triplicate against both of the G4 and iM structures and the concentrations at which 50% displacement ( $DC_{50}$ ) was achieved were calculated from dose-response curves fitted to this data (Figures S17, S18, Supplementary Material). Unfortunately, as with the FRET, RHPS4 (6) was excluded from analysis due to the fluorescence profile of the ligand overlapping with the assay parameters. The results of the FID assay showed that all ligands (1-5) bound to both the G4 and the iM DNA. Unsurprisingly, they also all showed a slightly higher affinity for the G4 formed by mutTel24<sub>FID</sub> over the iM formed by hTeloC<sub>FID</sub> (Table 2). Nevertheless, all ligands tested were found to displace TO at both pH 4.3 and 5.7, indicative of an interaction with the iM-forming DNA.

## DISCUSSION

DNA has a well-known propensity to adopt various alternative non-canonical conformations *in vitro*, including G4 and iM. Recent investigations have demonstrated the formation of such structures in regulatory regions of the human genome, including gene promoters and telomeres, also providing evidences for the key role that G4s and iMs can play in several biological pathways (Kang et al., 2014; Salvati et al., 2014; Maizels, 2015).

Human telomeric DNA consists of a 2–20 kb double-stranded region composed by (TTAGGG)<sub>n</sub>(CCCTAA)<sub>n</sub> repeats, and of a single-stranded 3'-end G-rich sequence. The G-rich strand can adopt G4 conformations, while the opposite C-rich strand can fold into the iM structure. Of the two structures, telomeric G4 is by far the most studied. This disparity is mainly due to acidic pH required for the protonation of cytosine, since the parallel duplexes, the basic component of iM, are stabilized by hemiprotonated C:C<sup>+</sup> base pairs. In this investigation, the behavior of the human telomeric iM-forming sequence (hTeloC) was studied in the pH range of 4.3–7.0 by 1D <sup>1</sup>H-NMR and CD spectroscopies. Results clearly indicate that under the experimental conditions used, the iM structure is still well-preserved at pH 5.7, while, just above this pH value, it turns out to be in equilibrium with the random coil (Figures S1, S2, Supplementary Material). Moreover, the lower stability of the iM structure observed at pH 5.7 well reflects the lower extent of cytosine protonation compared to pH 4.3, thus confirming, once again, that iM structures are very sensitive to pH (Kovanda et al., 2015). On the other hand, results of NMR and CD experiments on the G-rich telomeric sequence clearly showed that it retains

the hybrid [3+1] G4 structure as major conformation at both pH 4.3 and 5.7, and that its thermal stability is basically not affected by the pH.

Once the working conditions were established, the interaction of the well-known G4 ligands Berberine (1), BRACO-19 (2), Mitoxantrone (3), Phen-DC3 (4), Pyridostatin (5), and RHPS4 (6), with the iM-forming sequence was explored in comparison with the G4, by using a combination of spectroscopic techniques. Experiments were performed at the two boundary pH values. Results of the different experiments undoubtedly showed that all the investigated compounds actually are able to interact, even if in a different way, with both G4- and iM-forming DNA.

Fluorescent intercalator displacement (FID) assay clearly showed that the thiazole orange probe, which binds to the investigated structures, is competitively displaced by the compounds 1-5 (compound 6 could not be used). The consequent determination of their relative affinity for the DNA under examination reveals that, regardless of the pH (4.3 or 5.7), the compounds exhibit only a slightly higher affinity for mutTel24 over hTeloC.

Ligand-induced effects on both G4 and iM structures were examined by means of UV, CD, and NMR. As far as mutTel24 is concerned, results of CD experiments show that four (Berberine, BRACO-19, Phen-DC3, and RHPS4) out of the six ligands induce a conformational change from the hybrid [3+1] to an antiparallel structure (Figures S3, S4). All the ligands also indiscriminately affect the NMR spectrum of mutTel24 (Figure 5). Results of CD- and UV-melting experiments agree in indicating that the six compounds are able to stabilize the G4 at both pH values (Figures S7, S8, S11, S12, Supplementary Material). Interestingly, the two ligands not inducing the G4 conformational change (Mitoxantrone and Pyridostatin) seem also to be the less effective in terms of thermal stabilization under the experimental conditions used. Moreover, some of the investigated ligands induced a higher increase of G4 melting temperature ( $\Delta T_{1/2}$ ) at pH 4.3 rather than 5.7 (Table 1), that may be ascribed in part to the different protonation states of such molecules at the different pHs. This was, for example, the case of BRACO-19 (2) and Pyridostatin (5). For these molecules the protonation state was theoretically determined in the 3.5–7.5 pH range, by using the pK<sub>a</sub> prediction program (ChemAxon, www.chemaxon.com) based on the calculation of partial charge of atoms in the molecule. The computed distribution of the microspecies for BRACO-19 and Pyridostatin, reported in Figure S19, shows that actually there is a variation between pH 4.3 and 5.7 for these molecules.

Concerning the iM, some differences were observed between pH 4.3 and 5.7. In particular, the addition of 1-6 to the iM caused major effects at pH 5.7, in which the iM structure is less stable *per se*.

Interestingly, CD- and UV-melting results clearly indicate that the molecules which were found to mainly decrease the CD and NMR signals (namely BRACO-19, Mitoxantrone, and Phen-DC3) were also those that significantly decreased iM thermal stability at both pH values (Figures S9, S10, S13, S14, Supplementary Material).



**TABLE 2** | Ligand DC<sub>50</sub> values for hTeloC<sub>FID</sub> and mutTel24<sub>FID</sub> determined using the FID assay\*.

	hTeloC <sub>FID</sub>				mutTel24 <sub>FID</sub>			
	pH 4.3		pH 5.7		pH 4.3		pH 5.7	
	DC <sub>50</sub> ( $\mu$ M)	SE ( $\mu$ M)	DC <sub>50</sub> ( $\mu$ M)	SE ( $\mu$ M)	DC <sub>50</sub> ( $\mu$ M)	SE ( $\mu$ M)	DC <sub>50</sub> ( $\mu$ M)	SE ( $\mu$ M)
Berberine	30.38	1.58E-02	1.46	5.88E-03	3.32	2.63E-02	1.26	7.22E-03
BRACO-19	0.66	1.61E-03	0.87	8.08E-04	0.26	5.14E-03	0.50	1.95E-03
Mitoxantrone	0.70	2.87E-03	1.34	6.99E-03	0.54	3.03E-04	0.95	6.07E-03
Phen-DC3	0.97	1.32E-02	0.95	3.79E-03	0.26	1.58E-03	0.39	1.42E-03
Pyridostatin	9.09	6.18E-03	18.02	5.89E-02	3.15	2.33E-03	9.42	3.90E-02

\*All experiments were performed in triplicate and DC<sub>50</sub> values are reported as the 50% displacement value calculated from fitted dose response curves. Standard errors are calculated using R-square values from the statistics on the data fit.

Overall, these results show that compounds 1-6 are able to interact with the telomeric iM-forming DNA. However, three of them (Berberine, Pyridostatin, and RHPS4) do not have relevant effects on the thermal stability of iM, while the others (BRACO-19, Mitoxantrone, and Phen-DC3) are able to destabilize it. We speculate that these compounds could reasonably make some non-specific interactions with the single-stranded (unfolded) C-rich DNA, resulting in a shift of folded-unfolded equilibrium toward the unfolded form, especially during the melting experiment, which in turn results in a decrease of  $T_{1/2}$ .

The FRET methodology, used here to further characterize the interaction of G4 and iM DNA with the investigated ligands, deserves a separate discussion. With respect of other spectroscopic techniques (such as UV, CD, and NMR), FRET has a higher sensitivity and it can explore a large range of ligand concentrations (Monchaud et al., 2006; Sheng et al., 2017). Additionally, it turns out to be the main methodology when the UV absorbance of a ligand overlaps with that of the DNA (Guédin et al., 2010). However, some artifacts may occur when compounds are inherently fluorescent and/or interact with the fluorescent probes rather than the DNA itself (De Cian et al., 2007b), and this is what probably happened in this case. Indeed, FRET melting results did not agree at all with both CD and UV melting data, especially in the case of iM-F24T (Figure S16, Supplementary Material). To understand the reasons for this different behavior, FRET spectra of labeled DNA (G4-F21T and iM-F24T) in the absence and presence of compounds were analyzed. Typically, when DNA is folded, the two dyes are in close proximity so FAM fluorescence peak at 522 nm (upon excitation at 492 nm) is quenched and its energy is transferred to TAMRA, which then emits light at 580 nm. On the other hand, FAM's fluorescence is no longer quenched when sufficient spatial separation of the two dyes occurs (for example upon unfolding of the DNA structure), therefore its fluorescence signal at 522 nm is observable. In principle, compounds that are able to interact with the fluorophores may affect the emission properties of the probes and decrease the intensity of the bands at 580 nm (if the DNA is structured) or at 522 nm (if the DNA is unstructured). Interestingly, four compounds (BRACO-19,

Mitoxantrone, Phen-DC3, and Pyridostatin) caused a significant decrease of the band intensity at 580 nm (for both G4-F21T and iM-F24T), clearly suggesting that they actually interact with the fluorophores. Therefore, the different stabilizing effects observed for the iM structure across the different spectroscopic techniques could be ascribed to ligand interaction with the FRET probes. This hypothesis is further corroborated by the fact that Berberine (1), which did not cause any observable change in the FRET spectrum, showed no variation in the DNA's thermal stability by FRET, in agreement with CD and UV experiments. For the same reasons, this suggests that the thermal stabilizations of G4-F21T measured by FRET are not accurate, being also potentially affected by the ligands' interaction with the probes. Therefore, a clear message came out from a careful examination of FRET data, meaning that false-positive responses can be obtained due to ligands ability to bind end-labeling DNA probes. This could occur particularly when the investigated compounds have an extended aromatic core for which  $\pi$ - $\pi$  stacking interactions with the large aromatic surface of the probes could be favored. This is especially the case of G4-interacting molecules. Overall, this study emphasizes the need of using a combination of techniques when examining DNA targeting ligands, in order to avoid an inaccurate evaluation of their binding/stabilizing properties.

## CONCLUSIONS

Herein, a combination of spectroscopic techniques was employed to determine whether well-known bioactive G4 ligands, namely Berberine (1), BRACO-19 (2), Mitoxantrone (3), Phen-DC3 (4), Pyridostatin (5), and RHPS4 (6) are able to interact with an iM-forming DNA. Two human telomeric DNA sequences able to form iM and G4 structures were studied and the experiments performed at two different pH values. The experimental results showed that all the investigated G4 ligands were also able to interact with the telomeric iM-forming DNA. Very interestingly, BRACO-19, Mitoxantrone, and Phen-DC3 have been shown to destabilize the iM structure.

The majority of iM forming sequences are generally less stable than G4s under physiological conditions. The

delicate equilibrium between the folded and unfolded DNA forms is highly sensitive to the environmental conditions (such as pH and ionic strength), and ligands may affect this fine equilibrium, shifting it toward different/less stable forms.

The here reported results are even more interesting if viewed in the context of regulation of gene expression. Indeed, recent investigations have suggested that G4 and iM structures may have opposing functions in the control of oncogene transcription: while G4 formation and its ligand-induced stabilization generally inhibits gene expression, stabilization of iM seems to have transcription activating capabilities (Kang et al., 2014; Kendrick et al., 2014). Therefore, a molecule that is able to stabilize a G4 structure and to destabilize an iM structure may exert a synergistic effect on the inhibition of transcription. These are the cases of BRACO-19 (2), Mitoxantrone (3), and Phen-DC3 (4), whose biological activity may be ascribed to both mechanisms.

Overall, the present study highlights the necessity, when studying G4-targeting compounds, of evaluating also their effects on the i-motif counterparts, especially if one is looking for a “specific” drug.

## REFERENCES

- Amato, J., Iaccarino, N., Pagano, B., Morigi, R., Locatelli, A., Leoni, A., et al. (2014a). Bis-indole derivatives with antitumor activity turn out to be specific ligands of human telomeric G-quadruplex. *Front. Chem.* 2:54. doi: 10.3389/fchem.2014.00054
- Amato, J., Iaccarino, N., Randazzo, A., Novellino, E., and Pagano, B. (2014b). Noncanonical DNA secondary structures as drug targets: the prospect of the i-motif. *ChemMedChem* 9, 2026–2030. doi: 10.1002/cmdc.201402153
- Amato, J., Morigi, R., Pagano, B., Pagano, A., Ohnmacht, S., De Magis, A., et al. (2016). Toward the development of specific g-quadruplex binders: synthesis, biophysical, and biological studies of new hydrazone derivatives. *J. Med. Chem.* 59, 5706–5720. doi: 10.1021/acs.jmedchem.6b00129
- Amato, J., Pagano, A., Capasso, D., Di Gaetano, S., Giustiniano, M., Novellino, E., et al. (2018). Targeting the BCL2 gene promoter G-Quadruplex with a new class of furopridazinone-based molecules. *ChemMedChem* 13, 406–410. doi: 10.1002/cmdc.201700749
- Amato, J., Pagano, A., Cosconati, S., Amendola, G., Fotticchia, I., Iaccarino, N., et al. (2017). Discovery of the first dual G-triplex/G-quadruplex stabilizing compound: a new opportunity in the targeting of G-rich DNA structures? *Biochim. Biophys. Acta* 1861, 1271–1280. doi: 10.1016/j.bbagen.2016.11.008
- Bedrat, A., Lacroix, L., and Mergny, J. L. (2016). Re-evaluation of G-quadruplex propensity with G4Hunter. *Nucleic Acids Res.* 44, 1746–1759. doi: 10.1093/nar/gkw006
- Bhavsar-Jog, Y. P., Van Dornshuld, E., Brooks, T. A., Tschumper, G. S., and Wadkins, R. M. (2014). Epigenetic modification, dehydration, and molecular crowding effects on the thermodynamics of i-motif structure formation from C-Rich DNA. *Biochemistry* 53, 1586–1594. doi: 10.1021/bi401523b
- Biffi, G., Tannahill, D., McCafferty, J., and Balasubramanian, S. (2013). Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* 5, 182–186. doi: 10.1038/nchem.1548
- Boger, D. L., Fink, B. E., and Hedrick, M. P. (2000). Total synthesis of distamycin, A, and 2640 analogues: a solution-phase combinatorial approach to the discovery of new, bioactive dna binding agents and development of a rapid, high-throughput screen for determining relative DNA binding affinity or DNA bind. *J. Am. Chem. Soc.* 122, 6382–6394. doi: 10.1021/ja994192d
- Brazier, J. A., Shah, A., and Brown, G. D. (2012). I-Motif formation in gene promoters: unusually stable formation in sequences complementary to known G-quadruplexes. *Chem. Commun.* 48:10739. doi: 10.1039/c2cc30863k
- Brown, R. V., Wang, T., Chappeta, V. R., Wu, G., Onel, B., Chawla, R., et al. (2017). The consequences of overlapping G-Quadruplexes and i-Motifs in the platelet-derived growth factor receptor  $\beta$  core promoter nucleic hypersensitive element can explain the unexpected effects of mutations and provide opportunities for selective targeting of both structures by small molecules to downregulate gene expression. *J. Am. Chem. Soc.* 139, 7456–7475. doi: 10.1021/jacs.6b10028
- Burge, S., Parkinson, G. N., Hazel, P., Todd, A. K., and Neidle, S. (2006). Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.* 34, 5402–5415. doi: 10.1093/nar/gkl655
- Cantor, C. R., Warshaw, M. M., and Shapiro, H. (1970). Oligonucleotide interactions. III. Circular dichroism studies of the conformation of deoxyoligonucleotides. *Biopolymers* 9, 1059–1077. doi: 10.1002/bip.1970.360090909
- Cerofolini, L., Amato, J., Giachetti, A., Limongelli, V., Novellino, E., Parrinello, M., et al. (2014). G-triplex structure and formation propensity. *Nucleic Acids Res.* 42, 13393–13404. doi: 10.1093/nar/gku1084
- Chambers, V. S., Marsico, G., Boutell, J. M., Di Antonio, M., Smith, G. P., and Balasubramanian, S. (2015). High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat. Biotechnol.* 33, 877–881. doi: 10.1038/nbt.3295
- Chen, Y., Qu, K., Zhao, C., Wu, L., Ren, J., Wang, J., et al. (2012). Insights into the biomedical effects of carboxylated single-wall carbon nanotubes on telomerase and telomeres. *Nat. Commun.* 3:1074. doi: 10.1038/ncomms2091
- Day, H. A., Pavlou, P., and Waller, Z. A. E. (2014). I-Motif DNA: structure, stability and targeting with ligands. *Bioorg. Med. Chem.* 22, 4407–4418. doi: 10.1016/j.bmc.2014.05.047
- De Cian, A., DeLemos, E., Mergny, J. L., Teulade-Fichou, M. P., and Monchaud, D. (2007a). Highly efficient G-quadruplex recognition by bisquinolinium compounds. *J. Am. Chem. Soc.* 129, 1856–1857. doi: 10.1021/ja067352b
- De Cian, A., Guittat, L., Kaiser, M., Saccà, B., Amrane, S., Bourdoncle, A., et al. (2007b). Fluorescence-based melting assays for studying quadruplex ligands. *Methods* 42, 183–195. doi: 10.1016/j.ymeth.2006.10.004

## AUTHOR CONTRIBUTIONS

AR, JA, BP, and ZW conceived and designed the experiments. AP, NI, EG, and AD performed the CD, UV, and FRET experiments. DB performed the NMR experiments. MA performed the FID experiments. AP and JA carried out the synthesis of oligonucleotides. AP, NI, MA, EN, ZW, BP, JA, and AR analyzed the results. AR, JA, and ZW wrote the paper. All authors verified the data, contributed to the manuscript, and approved the final version.

## ACKNOWLEDGMENTS

This work was supported by the Italian Association for Cancer Research (AIRC) (IG-18695 to AR, IG-16730 to BP). MA was supported by a Ph.D. studentship from the Eastern Academic Research Consortium.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fchem.2018.00281/full#supplementary-material>

- Di Leva, F. S., Zizza, P., Cingolani, C., D'Angelo, C., Pagano, B., Amato, J., et al. (2013). Exploring the chemical space of G-quadruplex binders: discovery of a novel chemotype targeting the human telomeric sequence. *J. Med. Chem.* 56, 9646–9654. doi: 10.1021/jm401185b
- Dzatzko, S., Krafčiková, M., Hänsel-Hertsch, R., Fessl, T., Fiala, R., Loja, T., et al. (2018). Evaluation of the stability of DNA i-Motifs in the Nuclei of living mammalian cells. *Angew. Chem. Int. Ed. Engl.* 57, 2165–2169. doi: 10.1002/anie.201712284
- Fedoroff, O. Y., Rangan, A., Chemeris, V. V., and Hurley, L. H. (2000). Cationic porphyrins promote the formation of i-motif DNA and bind peripherally by a nonintercalative mechanism. *Biochemistry* 39, 15083–15090. doi: 10.1021/bi001528j
- Fleming, A. M., Ding, Y., Rogers, R. A., Zhu, J., Zhu, J., Burton, A. D., et al. (2017). 4n-1 Is a “Sweet Spot” in DNA i-Motif folding of 2'-Deoxycytidine homopolymers. *J. Am. Chem. Soc.* 139, 4682–4689. doi: 10.1021/jacs.6b10117
- Franceschin, M., Rossetti, L., D'Ambrosio, A., Schirripa, S., Bianco, A., Ortaggi, G., et al. (2006). Natural and synthetic G-quadruplex interactive berberine derivatives. *Bioorg. Med. Chem. Lett.* 16, 1707–1711. doi: 10.1016/j.bmcl.2005.12.001
- Gallego, J., Chou, S. H., and Reid, B. R. (1997). Centromeric pyrimidine strands fold into an intercalated motif by forming a double hairpin with a novel T:G:G:T tetrad: solution structure of the d(TCCCGTTTCCA) dimer. *J. Mol. Biol.* 273, 840–856. doi: 10.1006/jmbi.1997.1361
- Gehring, K., Leroy, J.-L., and Guéron, M. (1993). A tetrameric DNA structure with protonated cytosine-cytosine base pairs. *Nature* 363, 561–565. doi: 10.1038/363561a0
- Gowan, S. M., Harrison, J. R., Patterson, L., Valenti, M., Read, M., a, Neidle, S., et al. (2002). A G-quadruplex-interactive potent small-molecule inhibitor of telomerase exhibiting *in vitro* and *in vivo* antitumor activity. *Mol. Pharmacol.* 61, 1154–1162. doi: 10.1124/mol.61.5.1154
- Gray, R. D., Trent, J. O., and Chaires, J. B. (2014). Folding and unfolding pathways of the human telomeric G-quadruplex. *J. Mol. Biol.* 426, 1629–1650. doi: 10.1016/j.jmb.2014.01.009
- Guédin, A., Lacroix, L., and Mergny, J.-L. (2010). Thermal melting studies of ligand DNA interactions. *Methods Mol. Biol.* 613, 25–35. doi: 10.1007/978-1-60327-418-0\_2
- Guo, K., Gokhale, V., Hurley, L. H., and Sun, D. (2008). Intramolecularly folded G-quadruplex and i-motif structures in the proximal promoter of the vascular endothelial growth factor gene. *Nucleic Acids Res.* 36, 4598–4608. doi: 10.1093/nar/gkn380
- Haeusler, A. R., Donnelly, C. J., Periz, G., Simko, E. A. J., Shaw, P. G., Kim, M. S., et al. (2014). C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* 507, 195–200. doi: 10.1038/nature13124
- Huang, H.-S., Chen, L.-B., Huang, K.-F., Lu, W.-C., Shieh, F.-Y., Huang, Y.-Y., et al. (2007). Synthesis and human telomerase inhibition of a series of regioisomeric disubstituted aminoanthraquinones. *Chem. Pharm. Bull.* 55, 284–292. doi: 10.1248/cpb.55.284
- Huppert, J. L., and Balasubramanian, S. (2007). G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res.* 35, 406–413. doi: 10.1093/nar/gkl1057
- Izbicka, E., Wheelhouse, R. T., Raymond, E., Davidson, K. K., Lawrence, R. A., Sun, D., et al. (1999). Effects of cationic porphyrins as G-quadruplex interactive agents in human tumor cells. *Cancer Res.* 59, 639–644.
- Kang, H. J., Kendrick, S., Hecht, S. M., and Hurley, L. H. (2014). The transcriptional complex between the BCL2 i-motif and hnRNP LL is a molecular switch for control of gene expression that can be modulated by small molecules. *J. Am. Chem. Soc.* 136, 4172–4185. doi: 10.1021/ja4109352
- Karsisiotis, A. I., Hessari, N. M. A., Novellino, E., Spada, G. P., Randazzo, A., and Webba Da Silva, M. (2011). Topological characterization of nucleic acid G-quadruplexes by UV absorption and circular dichroism. *Angew. Chem. Int. Ed. Engl.* 50, 10645–10648. doi: 10.1002/anie.201105193
- Kendrick, S., Kang, H. J., Alam, M. P., Madathil, M. M., Agrawal, P., Gokhale, V., et al. (2014). The dynamic character of the BCL2 promoter i-motif provides a mechanism for modulation of gene expression by compounds that bind selectively to the alternative DNA hairpin structure. *J. Am. Chem. Soc.* 136, 4161–4171. doi: 10.1021/ja410934b
- Kovanda, A., Zalar, M., Šket, P., Plavec, J., and Rogelj, B. (2015). Anti-sense DNA d(GGCCCC)nexpansions in C9ORF72 form i-motifs and protonated hairpins. *Sci. Rep.* 5:17944. doi: 10.1038/srep17944
- Lam, E. Y. N., Beraldi, D., Tannahill, D., and Balasubramanian, S. (2013). G-quadruplex structures are stable and detectable in human genomic DNA. *Nat. Commun.* 4, 1796. doi: 10.1038/ncomms2792
- Li, Q., Xiang, J. F., Yang, Q. F., Sun, H. X., Guan, A. J., and Tang, Y. L. (2013). G4LDB: A database for discovering and studying G-quadruplex ligands. *Nucleic Acids Res.* 41, D1115–D1123. doi: 10.1093/nar/gks1101
- Li, X., Peng, Y., Ren, J., and Qu, X. (2006). Carboxyl-modified single-walled carbon nanotubes selectively induce human telomeric i-motif formation. *Proc. Natl. Acad. Sci. U.S.A.* 103, 19658–19663. doi: 10.1073/pnas.0607245103
- Luu, K. N., Phan, A. T., Kuryavyi, V., Lacroix, L., and Patel, D. J. (2006). Structure of the human telomere in K+ solution: an intramolecular (3 + 1) G-quadruplex scaffold. *J. Am. Chem. Soc.* 128, 9963–9970. doi: 10.1021/ja062791w
- Maizels, N. (2015). G4-associated human diseases. *EMBO Rep.* 16, 910–922. doi: 10.15252/embr.201540607
- Malliavin, T. E., Snoussi, K., and Leroy, J. L. (2003). The NMR structure of [Xd(C2)]<sub>4</sub> investigated by molecular dynamics simulations. *Magn. Reson. Chem.* 41, 18–25. doi: 10.1002/mrc.1109
- Masiero, S., Trotta, R., Pieraccini, S., De Tito, S., Perone, R., Randazzo, A., et al. (2010). A non-empirical chromophoric interpretation of CD spectra of DNA G-quadruplex structures. *Org. Biomol. Chem.* 8, 2683–2692. doi: 10.1039/c003428b
- Mergny, J. L., Lacroix, L., Hélène, C., Han, X., and Leroy, J. L. (1995). Intramolecular Folding of Pyrimidine Oligodeoxynucleotides into an i-DNA Motif. *J. Am. Chem. Soc.* 117, 8887–8898. doi: 10.1021/ja00140a001
- Mir, B., Serrano, I., Buitrago, D., Orozco, M., Escaja, N., and González, C. (2017). Prevalent sequences in the human genome can form mini i-Motif structures at physiological pH. *J. Am. Chem. Soc.* 139, 13985–13988. doi: 10.1021/jacs.7b07383
- Monchaud, D., Allain, C., and Teulade-Fichou, M.-P. (2006). Development of a fluorescent intercalator displacement assay (G4-FID) for establishing quadruplex-DNA affinity and selectivity of putative ligands. *Bioorg. Med. Chem. Lett.* 16, 4842–4845. doi: 10.1016/j.bmcl.2006.06.067
- Pagano, B., Amato, J., Iaccarino, N., Cingolani, C., Zizza, P., Biroccio, A., et al. (2015). Looking for efficient G-quadruplex ligands: evidence for selective stabilizing properties and telomere damage by drug-like molecules. *ChemMedChem* 10, 640–649. doi: 10.1002/cmdc.201402552
- Pagano, B., Cosconati, S., Gabelica, V., Petraccone, L., De Tito, S., Marinelli, L., et al. (2012). State-of-the-art methodologies for the discovery and characterization of DNA G-quadruplex binders. *Curr. Pharm. Des.* 18, 1880–1899. doi: 10.2174/138161212799958332
- Pagano, B., Fotticchia, I., De Tito, S., Mattia, C., a, Mayol, L., Novellino, E., et al. (2010). Selective Binding of Distamycin A Derivative to G-Quadruplex Structure [d(TGGGGT)]<sub>4</sub>. *J. Nucleic Acids* 2010, 1–7. doi: 10.4061/2010/247137
- Pagano, B., Mattia, C. A., Virno, A., Randazzo, A., Mayol, L., and Giancola, C. (2007). Thermodynamic analysis of quadruplex DNA-drug interaction. *Nucleosides Nucleotides Nucleic Acids* 26, 761–765. doi: 10.1080/15257770701499069
- Phan, A. T., Guéron, M., and Leroy, J. L. (2000). The solution structure and internal motions of a fragment of the cytidine-rich strand of the human telomere. *J. Mol. Biol.* 299, 123–144. doi: 10.1006/jmbi.2000.3613
- Rajendran, A., Nakano, S., and Sugimoto, N. (2010). Molecular crowding of the cosolutes induces an intramolecular i-motif structure of triplet repeat DNA oligomers at neutral pH. *Chem. Commun.* 46, 1299–1301. doi: 10.1039/b922050j
- Randazzo, A., Galeone, A., Esposito, V., Varra, M., and Mayol, L. (2002). Interaction of distamycin a and netropsin with quadruplex and duplex structures: A comparative 1H-NMR study. *Nucleosides, Nucleotides Nucleic Acids* 21, 535–545. doi: 10.1081/NCN-120015067
- Randazzo, A., Spada, G. P., and Da Silva, M. W. (2013). Circular dichroism of quadruplex structures. *Top. Curr. Chem.* 330, 67–86. doi: 10.1007/128\_2012\_331
- Rodriguez, R., Müller, S., Yeoman, J. A., Trentesaux, C., Riou, J. F., and Balasubramanian, S. (2008). A novel small molecule that alters shelterin integrity and triggers a DNA-damage response at telomeres. *J. Am. Chem. Soc.* 130, 15758–15759. doi: 10.1021/ja805615w

- Salvati, E., Botta, L., Amato, J., Di Leva, F. S., Zizza, P., Gioiello, A., et al. (2017). Lead Discovery of Dual G-Quadruplex Stabilizers and Poly(ADP-ribose) Polymerases (PARPs) Inhibitors: a new avenue in anticancer treatment. *J. Med. Chem.* 60, 3626–3635. doi: 10.1021/acs.jmedchem.6b01563
- Salvati, E., Zizza, P., Rizzo, A., Iachettini, S., Cingolani, C., D'angelo, C., et al. (2014). Evidence for G-quadruplex in the promoter of vegfr-2 and its targeting to inhibit tumor angiogenesis. *Nucleic Acids Res.* 42, 2945–2957. doi: 10.1093/nar/gkt1289
- Sheng, Q., Neaverson, J. C., Mahmoud, T., Stevenson, C. E. M., Matthews, S. E., and Waller, Z. A. E. (2017). Identification of new DNA i-motif binding ligands through a fluorescent intercalator displacement assay. *Org. Biomol. Chem.* 15, 5669–5673. doi: 10.1039/C7OB00710H
- Siddiqui-Jain, A., Grand, C. L., Bearss, D. J., and Hurley, L. H. (2002). Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. U.S.A.* 99, 11593–11598. doi: 10.1073/pnas.182256799
- Sun, D., and Hurley, L. H. (2009). The importance of negative superhelicity in inducing the formation of G-quadruplex and i-motif structures in the c-Myc promoter: Implications for drug targeting and control of gene expression. *J. Med. Chem.* 52, 2863–2874. doi: 10.1021/jm900055s
- Wright, E., Lamparska, K., Smith, S. S., and Waller, Z. A. E. (2017). Substitution of cytosine with guanylyurea decreases the stability of i-Motif DNA. *Biochemistry* 56, 4879–4883. doi: 10.1021/acs.biochem.7b00628
- Wright, E. P., Day, H. A., Ibrahim, A. M., Kumar, J., Boswell, L. J. E., Huguin, C., et al. (2016a). Mitoxantrone and analogues bind and stabilize i-motif forming DNA sequences. *Sci. Rep.* 6:39456. doi: 10.1038/srep39456
- Wright, E. P., Huppert, J. L., and Waller, Z. A. E. (2016b). Identification of multiple genomic DNA sequences which form i-motif structures at neutral pH. *Nucleic Acids Res.* 45, 2951–2959. doi: 10.1093/nar/gkx090
- Xu, B., Devi, G., and Shao, F. (2015). Regulation of telomeric i-motif stability by 5-methylcytosine and 5-hydroxymethylcytosine modification. *Org. Biomol. Chem.* 13, 5646–5651. doi: 10.1039/C4OB02646B
- Xu, L., Hong, S., Sun, N., Wang, K., Zhou, L., Ji, L., et al. (2016). Berberine as a novel light-up i-motif fluorescence ligand and its application in designing molecular logic systems. *Chem. Commun.* 52, 179–182. doi: 10.1039/C5CC08242K
- Zeraati, M., Langley, D. B., Schofield, P., Moye, A. L., Rouet, R., Hughes, W. E., et al. (2018). I-motif DNA structures are formed in the nuclei of human cells. *Nat. Chem.* 10, 631–637. doi: 10.1038/s41557-018-0046-3
- Zhao, J., Bacolla, A., Wang, G., and Vasquez, K. M. (2010). Non-B DNA structure-induced genetic instability and evolution. *Cell. Mol. Life Sci.* 67, 43–62. doi: 10.1007/s00018-009-0131-2
- Zhou, J., Wei, C., Jia, G., Wang, X., Feng, Z., and Li, C. (2010). Formation of i-motif structure at neutral and slightly alkaline pH. *Mol. Biosyst.* 6, 580–586. doi: 10.1039/B919600E
- Zizza, P., Cingolani, C., Artuso, S., Salvati, E., Rizzo, A., D'Angelo, C., et al. (2016). Intragenic G-quadruplex structure formed in the human CD133 and its biological and translational relevance. *Nucleic Acids Res.* 44, 1579–1590. doi: 10.1093/nar/gkv1122

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Pagano, Iaccarino, Abdelhamid, Brancaccio, Garzarella, Di Porzio, Novellino, Waller, Pagano, Amato and Randazzo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

## Destabilization of i-Motif DNA at Neutral pH by G-Quadruplex Ligands

Mahmoud A. S. Abdelhamid,<sup>†,‡</sup> Andrew J. Gates,<sup>\*,§,‡</sup> and Zoë A. E. Waller<sup>\*,†,‡</sup>

<sup>†</sup>School of Pharmacy, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, United Kingdom

<sup>‡</sup>Centre for Molecular and Structural Biochemistry, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, United Kingdom

<sup>§</sup>School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, United Kingdom

### Supporting Information

**ABSTRACT:** Numerous studies have been published stressing the importance of finding ligands that can bind specifically to DNA secondary structures. Several have identified ligands that are presented as having specific binding to the G-quadruplex; however, these were not originally tested on the complementary i-motif structure. The i-motif was overlooked and presumed to be irrelevant due to the belief that the hemiprotonated (cytosine<sup>+</sup>–cytosine) base pair at the core of the structure required acidic pH. The pathophysiological relevance of i-motifs has since been documented, as well as the discovery of several genomic sequences, which can form i-motif at neutral pH. Using different biophysical methodologies, we provide experimental evidence to show that widely used G-quadruplex ligands interact with i-motif structures at neutral pH, generally leading to their destabilization. Crucially, this has implications both for the search for quadruplex binding compounds as well as for the effects of compounds reported to have G-quadruplex specificity without examining their effects on i-motif.

Contrary to it once being regarded as a passive code of instructions, the sequences, structures, and functions of DNA are inextricable from both homeostasis and pathogenesis.<sup>1</sup> A diverse array of noncanonical nucleic acid structures exists conferring on DNA these properties, and two types of four-stranded structures have emerged as potential key players in this regard: the G-quadruplex and the i-motif.<sup>2,3</sup> The G-quadruplex is formed by sequences containing stretches of consecutive guanine nucleotides,<sup>4</sup> and the i-motif is dependent on a sequence with cytosine base repeats;<sup>5</sup> both quadruplex-forming sequences have interspersed loop-forming regions.

The more prominent quadruplex of the two in terms of research output is undoubtedly the G-quadruplex. An already substantial body of evidence for the significance of the G-quadruplex is growing and garnering further interest due to the implications raised by their location in the genome and their biological activity.<sup>6,7</sup> Additionally, recent work estimates a 2–10-fold higher number of G-quadruplexes in the human genome than previously proposed.<sup>8</sup> A similar body of evidence has been accumulating for the “other” quadruplex structure known as i-motif. As far as has been revealed to date, the

genomic locations and biological effects of the i-motif are apparently just as interesting as those of the G-quadruplex.<sup>9,10</sup> However, the number of sequences in the genome with the propensity to form i-motif is not clearly known. The primary cause for the discrepancy in research focus was the i-motif's requirement for hemiprotonated cytosine base pairing, which gave rise to the assumption that these structures invariably require acidic pH to fold and so would not be physiologically relevant.<sup>3</sup> This was later shown to not be the case, and several examples of genomic DNA sequences that form i-motif structures at neutral pH were presented.<sup>10,11</sup> Furthermore, conditions of molecular crowding,<sup>12</sup> low temperature,<sup>13</sup> or negative superhelicity<sup>14</sup> have all been shown to be favorable for i-motif formation at neutral pH.

Historically, DNA was the first target for anticancer drugs and remains the mainstay of most treatment regimes.<sup>15</sup> Given their positions in the genome, the potential therapeutic impact of targeting quadruplex DNA structures to control homeostatic or pathogenic processes is vast.<sup>16,17</sup> Ground-breaking recent work allowed the *in vivo* visualization of the two structures in the nuclei of human cells. The G-quadruplex was shown first in 2013 by Balasubramanian and co-workers,<sup>18</sup> followed by the i-motif in 2018 by Christ and co-workers.<sup>19</sup> Christ's work did, however, reveal an issue with disregarding the i-motif structure: finding that the antibody used to confirm the *in vivo* existence of G-quadruplex in cells was also able to interact with the i-motif. More recent work has shown that at acidic pH, 5.3 and below, common agents assumed in the literature to be specifically binding G-quadruplex also interact with i-motif.<sup>20</sup> Herein, we describe the effects of known G-quadruplex ligands on two genomic i-motif-forming sequences under neutral pH conditions.

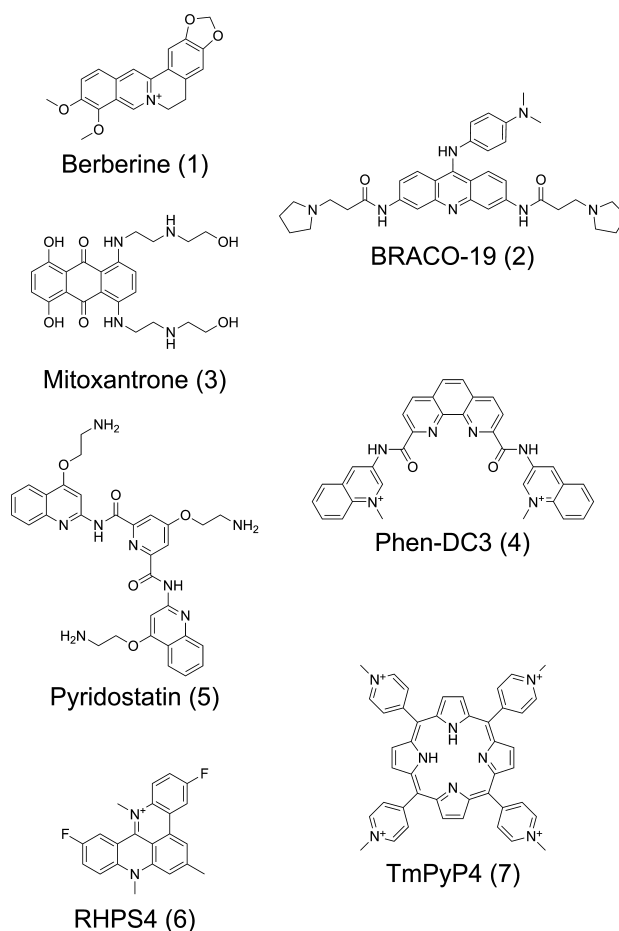
Almost all previous efforts in the literature to examine the interaction between small molecules and the i-motif have been performed at acidic pH.<sup>20</sup> While the effect of pH on the i-motif sequence under investigation is discussed at length, rarely is consideration given to the effect of pH on the ligand itself. Each small molecule will be affected by the pH in a different manner, depending on the functional groups that it contains.

**Special Issue:** Future of Biochemistry: The International Issue

**Received:** September 10, 2018

**Revised:** October 15, 2018

**Published:** October 17, 2018



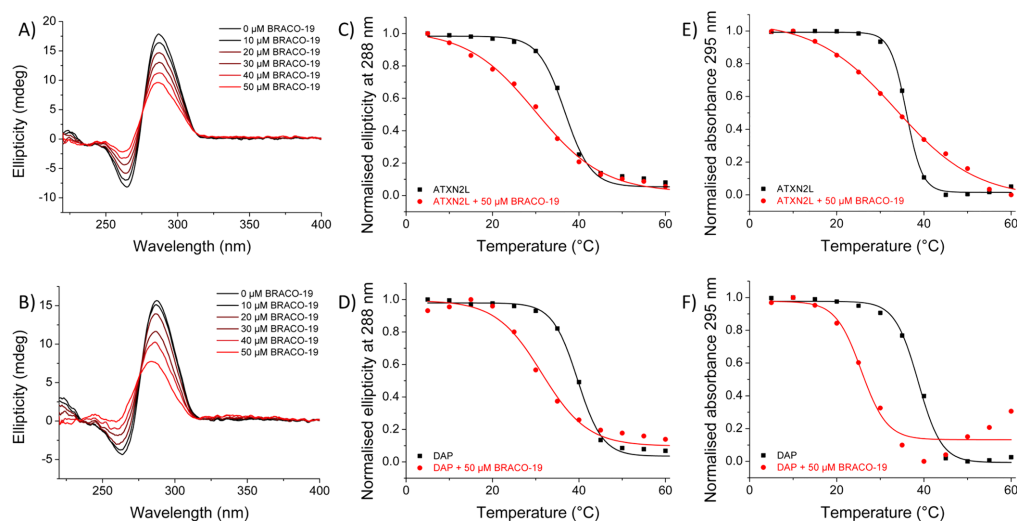
**Figure 1.** Chemical structures of the investigated ligands.

Aliphatic amino groups will be protonated at physiological pH, but aromatic amines would typically not be to the same extent. However, at acidic pH, the protonation of aromatic amines would be more likely. This will consequently have an effect on ligand interactions with different DNA structures. As such, we decided to examine how several established G-quadruplex ligands interact with two candidate i-motif-forming sequences at neutral pH. The two sequences used, ATXN2L 5'-(CCC-CCC)<sub>4</sub>-3' and DAP 5'-(CCC-CCG)<sub>4</sub>CCCC3', are from promoter regions of their respective genes in the human genome and have been shown to fold at neutral pH.<sup>11,21</sup> The G-quadruplex binding compounds used (Figure 1) were berberine (1),<sup>22</sup> BRACO-19 (2),<sup>23</sup> mitoxantrone (3),<sup>24</sup> Phen-DC3 (4),<sup>25</sup> pyridostatin (5),<sup>26</sup> RHPS4 (6),<sup>27</sup> and TmPyP4 (7).<sup>28</sup> We employed three techniques to monitor the interactions of these ligands with the neutral i-motifs. Circular dichroism (CD) titrations were used to determine the effects of the compounds on the i-motif structure. Both CD

and UV melting were used to observe changes in thermal stability upon addition of the ligand. Finally, a fluorescent intercalator displacement (FID) assay was used to compare the affinities of the ligands. Only DAP was used in the FID experiments as ATXN2L is a poor candidate for this technique due to the low fluorescence enhancement observed after equilibration of the fluorescent probe. While using a lower pH would have resolved this issue, to allow meaningful comparisons, we consistently use 10 mM sodium cacodylate buffer at pH 7.0 to better address the scepticism surrounding the physiological relevance of the i-motif.

Before any experiments, we used CD to confirm that DAP and ATXN2L formed i-motif structures under our experimental conditions at pH 7.0. Both sequences showed the characteristic positive peak at 288 nm and a negative one at 260 nm (Figure 2). This was followed by titration of each of the ligands (1–7) in 1 equiv (10 μM) increments to 5 equiv (50 μM). This upper concentration was chosen so as to give 5





**Figure 2.** CD spectra of 10  $\mu\text{M}$  ATXN2L (A) and DAP (B) with titration up to 5 equiv (50  $\mu\text{M}$ ) of BRACO-19. Normalized ellipticity at 288 nm of 10  $\mu\text{M}$  ATXN2L (C) and DAP (D) without ligand (black) and with 5 equiv (50  $\mu\text{M}$ ) of BRACO-19 (red). Normalized absorbance at 295 nm of 10  $\mu\text{M}$  ATXN2L (E) and DAP (F) without ligand (black) and with 5 equiv (50  $\mu\text{M}$ ) of BRACO-19 (red). Experiments performed at pH 7.0 in 10 mM sodium cacodylate buffer.

**Table 1.** Change in Melting Temperature ( $\Delta T_m$ ) of ATXN2L and DAP i-Motifs with G-Quadruplex Ligands Measured by CD and UV Melting Experiments<sup>a</sup>

ligand	ATXN2L		DAP	
	$\Delta T_m$ ( $^{\circ}\text{C}$ )		$\Delta T_m$ ( $^{\circ}\text{C}$ )	
	CD melting	UV melting	CD melting	UV melting
(1) berberine	$-1.5 \pm 0.3$	$-1.17 \pm 0.1$	$-3.9 \pm 0.3$	$-3.4 \pm 0.3$
(2) BRACO-19	$-6.4 \pm 0.3$	$-1.50 \pm 0.1$	$-7.3 \pm 0.7$	$-11.6 \pm 0.3$
(3) mitoxantrone	$-1.0 \pm 0.4$	$3.84 \pm 0.2$	$-7.1 \pm 0.3$	$-2.3 \pm 0.4$
(4) Phen-DC3	$-1.4 \pm 0.3$	$-0.61 \pm 0.1$	$-3.0 \pm 0.4$	$-4.4 \pm 0.8$
(5) pyridostatin	$-0.8 \pm 0.3$	$0.87 \pm 0.3$	$-1.8 \pm 0.3$	$-1.5 \pm 0.3$
(6) RHPS4	$-1.4 \pm 0.3$	$-7.48 \pm 0.1$	$-6.8 \pm 0.4$	$-8.7 \pm 0.3$
(7) TmPyP4	ND	$4.48 \pm 0.1$	ND	$-16.0 \pm 0.3$

<sup>a</sup> $T_m$  is the midpoint of the transition from each melting experiment, and SE is calculated using R-square values from the statistics on the data fit.  $\Delta T_m$  is the difference between the  $T_m$  of the DNA in the presence of 5 equiv of each ligand and the DNA on its own. ND, not determined.

M equiv of ligand with respect to the DNA, thereby potentially representing an excess of ligand to any potential binding sites on the structure (representative data is shown for ligand 2 in Figure 2A,B). At these concentrations, 1 and 5 did not have significant effects on the spectra of either i-motif, with 5 resulting in only a minimal hypochromic shift of the peak at 288 nm. By contrast, 2–4 and 6 resulted, to varying extents, in considerably greater hypochromic shifts of both the positive band at 288 nm and the negative band at 260 nm, as well as a small hypochromic shift of both the positive band at 288 nm, consistent with an apparent unfolding effect.<sup>29</sup> The DAP i-motif appears to be affected by these ligands to a greater extent compared with ATXN2L; this is most apparent when observing the negative band at 260 nm, which comes very close to zero and almost loses its negative character in the cases of ligands 2 and 4. The effects of 7 were most apparent, as it resulted in the most significant hypochromic shifts of the bands at 288 and 260 nm for both sequences. Interestingly, while the

bands at 288 nm initially exhibited a hypochromic shift similar to that seen with 2–4 and 6 as the ligand concentration increased, with 7 at 40 and 50  $\mu\text{M}$  a bathochromic shift is then observed. These data suggest that these ligands are interacting with the i-motif structures formed by ATXN2L and DAP (see SI). The hypochromicity observed with 2–4 and 6 suggests that they are inducing the unfolding of the i-motif structure, and while 7 appears to do the same at the lower concentrations, the subsequent bathochromic shift suggests that perhaps another structure is then adopted.

To assess the effects of the ligands on the stability of the i-motif structures, the melting temperature ( $T_m$ ) for ATXN2L and DAP was determined in the absence of ligands using both CD and UV melting experiments. The results from both techniques were in very good agreement with one another: using CD to monitor ellipticity at 288 nm, the  $T_m$  for ATXN2L was determined to be  $36.9 \pm 0.2$   $^{\circ}\text{C}$ , and for DAP, it was  $39.7 \pm 0.1$   $^{\circ}\text{C}$ ; using UV and monitoring absorbance at 295 nm, the

$T_m$  for ATNX2L was determined to be  $36.0 \pm 0.02$  °C, and for DAP, it was  $38.6 \pm 0.2$  °C. The extent of destabilization of the i-motif observed using CD melting (example data shown for ligand 2 in Figure 2C,D) in the presence of the ligands correlated very well with the intensity of the hypochromicity observed in the titrations ( $r = -0.7691$ ,  $p = 0.0035$ , Table S1). Ligands (1–6) all showed a destabilization of the i-motif structure; the destabilization of DAP was more pronounced than ATNX2L (Table 1). Due to the significant changes in the CD spectra of the i-motifs in the presence of 7, a  $T_m$  could not be determined using this method.

In contrast to the CD melting experiments, the results from UV melting were less straightforward to interpret (example data is shown for ligand 2 in Figure 2E,F). Monitoring the ellipticity at 288 nm versus temperature using CD is a direct reporter of the melting of the i-motif structure.<sup>30</sup> On the other hand, changes in absorbance at 295 nm are not unique to the i-motif.<sup>31</sup> In the cases of 3, 5, and 7 with the ATNX2L sequence, an increase of the  $T_m$  is observed possibly due to the stabilization of a proportion of the DNA that has formed an alternative secondary structure, or an otherwise altered unfolding process reflected in the changes in the slopes of the melting curves. This result is not completely surprising as reports in the literature have shown using some techniques that 3 can stabilize i-motif-forming sequences,<sup>32</sup> while others show destabilization.<sup>20</sup> Otherwise, a similar destabilization pattern is observed for all ligands 1–7 with DAP and for the remaining ligands (1–2, 4, and 6) with ATNX2L. Again, the stability of the DAP i-motif was more significantly affected than ATNX2L (Table 1).

To be able to compare the relative affinities of ligands 1–7 to DAP, an FID assay was used to determine the concentration at which each ligand could displace 50% of the fluorescent thiazole orange probe (DC<sub>50</sub>) (Table 2). The fluorescence

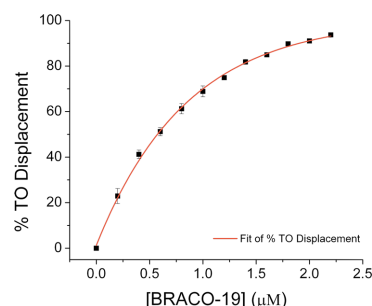
**Table 2. Ligand DC<sub>50</sub> Values for DAP Determined Using the FID Assay<sup>a</sup>**

Ligand	DC <sub>50</sub> (μM)
(1) berberine	27.86 ± 0.23
(2) BRACO-19	0.57 ± 0.00
(3) mitoxantrone	0.62 ± 0.00
(4) Phen-DC3	1.29 ± 0.02
(5) pyridostatin	14.61 ± 0.17
(6) RHPS4	ND
(7) TmPyP4	0.16 ± 0.00

<sup>a</sup>All experiments were performed at least in triplicate, and DC<sub>50</sub> values are reported as the 50% displacement value calculated from fitted dose–response curves. Standard errors are calculated using R-square values from the statistics on the data fit. ND, not determined.

profile of RHPS4 (6) overlapped with the parameters used in this assay and as such was excluded from analysis. As this assay relies on displacement, rather than occupation of all potential binding sites, the concentration range used for each ligand was selected to provide sufficient data points above and below the DC<sub>50</sub> to allow for reliable fitting of the data (example data is shown for ligand 2 in Figure 3).

Here we have shown that the G-quadruplex binding compounds 1–7 all interact with two i-motif sequences at neutral pH. This lends further credence to the issue highlighted by Christ's work on the i-motif-interacting antibody iMab: it is essential to examine both the G-



**Figure 3.** Dose–response curve from the FID assay of DAP (1 μM) at pH 7.0 in 10 mM sodium cacodylate buffer with BRACO-19. Error bars show the standard deviation across three repeats.

quadruplex and the i-motif when reporting on a molecule that appears to bind specifically to a given DNA secondary structure. This further demonstrates that the purported necessity for acidic pH when working with i-motif is false and that i-motifs that fold at neutral pH can be considered as potentially physiologically relevant structures with the same confidence as G-quadruplexes.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.8b00968.

Materials and methods, circular dichroism and UV spectroscopy, fluorescent intercalator displacement assay, Pearson correlation, and dose–response curves (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Authors

\*E-mail: a.gates@uea.ac.uk.

\*E-mail: z.waller@uea.ac.uk.

### ORCID

Mahmoud A. S. Abdelhamid: 0000-0002-1247-9915

Andrew J. Gates: 0000-0002-4594-5038

Zoë A. E. Waller: 0000-0001-8538-0484

### Author Contributions

The manuscript was written through contributions of all authors.

### Funding

This work was supported by the Eastern Academic Research Consortium Studentship (to M.A.S.A.) and the Biotechnology and Biological Sciences Research Council [BB/M00256X/1 to A.J.G. and BB/L02229X/1 to Z.A.E.W.].

### Notes

The authors declare no competing financial interest.

## ■ REFERENCES

- (1) Sinden, R. R. (1999) Biological implications of the DNA structures associated with disease-causing triplet repeats. *Am. J. Hum. Genet.* 64, 346–353.



- (2) Brooks, T. A., Kendrick, S., and Hurley, L. (2010) Making sense of G-quadruplex and i-motif functions in oncogene promoters. *FEBS J.* 277, 3459–3469.
- (3) Day, H. A., Pavlou, P., and Waller, Z. A. E. (2014) i-Motif DNA: structure, stability and targeting with ligands. *Bioorg. Med. Chem.* 22, 4407–4418.
- (4) Sen, D., and Gilbert, W. (1988) Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature* 334, 364–366.
- (5) Gehring, K., Leroy, J.-L., and Guéron, M. (1993) A tetrameric DNA structure with protonated cytosine-cytosine base pairs. *Nature* 363, 561–565.
- (6) Murat, P., and Balasubramanian, S. (2014) Existence and consequences of G-quadruplex structures in DNA. *Curr. Opin. Genet. Dev.* 25, 22–29.
- (7) Balasubramanian, S., and Neidle, S. (2009) G-quadruplex nucleic acids as therapeutic targets. *Curr. Opin. Chem. Biol.* 13, 345–353.
- (8) Bedrat, A., Lacroix, L., and Mergny, J. L. (2016) Re-evaluation of G-quadruplex propensity with G4Hunter. *Nucleic Acids Res.* 44, 1746–1759.
- (9) Abou Assi, H., Garavis, M., Gonzalez, C., and Damha, M. J. (2018) i-Motif DNA: structural features and significance to cell biology. *Nucleic Acids Res.* 46, 8038–8056.
- (10) Brazier, J. A., Shah, A., and Brown, G. D. (2012) I-motif formation in gene promoters: unusually stable formation in sequences complementary to known G-quadruplexes. *Chem. Commun.* 48, 10739–10741.
- (11) Wright, E. P., Huppert, J. L., and Waller, Z. A. E. (2017) Identification of multiple genomic DNA sequences which form i-motif structures at neutral pH. *Nucleic Acids Res.* 45, 2951–2959.
- (12) Bhavsar-Jog, Y. P., Van Dornshuld, E., Brooks, T. A., Tschumper, G. S., and Wadkins, R. M. (2014) Epigenetic modification, dehydration, and molecular crowding effects on the thermodynamics of i-motif structure formation from C-rich DNA. *Biochemistry* 53, 1586–1594.
- (13) Zhou, J., Wei, C., Jia, G., Wang, X., Feng, Z., and Li, C. (2010) Formation of i-motif structure at neutral and slightly alkaline pH. *Mol. Biosyst.* 6, 580–586.
- (14) Sun, D., and Hurley, L. H. (2009) The importance of negative superhelicity in inducing the formation of G-quadruplex and i-motif structures in the c-Myc promoter: implications for drug targeting and control of gene expression. *J. Med. Chem.* 52, 2863–2874.
- (15) Kohn, K. W. (1996) Beyond DNA cross-linking: history and prospects of DNA-targeted cancer treatment. *Cancer Res.* 56, 5533–5546.
- (16) Balasubramanian, S., Hurley, L. H., and Neidle, S. (2011) Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat. Rev. Drug Discovery* 10, 261–275.
- (17) Hansel-Hertsch, R., Di Antonio, M., and Balasubramanian, S. (2017) DNA G-quadruplexes in the human genome: detection, functions and therapeutic potential. *Nat. Rev. Mol. Cell Biol.* 18, 279–284.
- (18) Biffi, G., Tannahill, D., McCafferty, J., and Balasubramanian, S. (2013) Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* 5, 182–186.
- (19) Zeraati, M., Langley, D. B., Schofield, P., Moye, A. L., Rouet, R., Hughes, W. E., Bryan, T. M., Dinger, M. E., and Christ, D. (2018) I-motif DNA structures are formed in the nuclei of human cells. *Nat. Chem.* 10, 631–637.
- (20) Pagano, A., Iaccarino, N., Abdelhamid, M. A. S., Brancaccio, D., Garzarella, E. U., Di Porzio, A., Novellino, E., Waller, Z. A. E., Pagano, B., Amato, J., and Randazzo, A. (2018) Common G-Quadruplex Binding Agents Found to Interact With i-Motif-Forming DNA: Unexpected Multi-Target-Directed Compounds. *Front. Chem.* 6, 1–13.
- (21) Dzatko, S., Krafcikova, M., Hansel-Hertsch, R., Fessel, T., Fiala, R., Loja, T., Krafcik, D., Mergny, J. L., Foldynova-Trantirkova, S., and Trantirek, L. (2018) Evaluation of the Stability of DNA i-Motifs in the Nuclei of Living Mammalian Cells. *Angew. Chem., Int. Ed.* 57, 2165–2169.
- (22) Franceschin, M., Rossetti, L., D'Ambrosio, A., Schirripa, S., Bianco, A., Ortaggi, G., Savino, M., Schultes, C., and Neidle, S. (2006) Natural and synthetic G-quadruplex interactive berberine derivatives. *Bioorg. Med. Chem. Lett.* 16, 1707–1711.
- (23) Gowan, S. M., Harrison, J. R., Patterson, L., Valenti, M., Read, M. A., Neidle, S., and Kelland, L. R. (2002) A G-quadruplex-interactive potent small-molecule inhibitor of telomerase exhibiting in vitro and in vivo antitumor activity. *Mol. Pharmacol.* 61, 1154–1162.
- (24) Huang, H. S., Chen, I. B., Huang, K. F., Lu, W. C., Shieh, F. Y., Huang, Y. Y., Huang, F. C., and Lin, J. J. (2007) Synthesis and human telomerase inhibition of a series of regioisomeric disubstituted amidoanthraquinones. *Chem. Pharm. Bull.* 55, 284–292.
- (25) De Cian, A., Delemos, E., Mergny, J. L., Teulade-Fichou, M. P., and Monchaud, D. (2007) Highly efficient G-quadruplex recognition by bisquinolinium compounds. *J. Am. Chem. Soc.* 129, 1856–1857.
- (26) Rodriguez, R., Muller, S., Yeoman, J. A., Trentesaux, C., Riou, J. F., and Balasubramanian, S. (2008) A novel small molecule that alters shelterin integrity and triggers a DNA-damage response at telomeres. *J. Am. Chem. Soc.* 130, 15758–15759.
- (27) Izbiccka, E., Wheelhouse, R. T., Raymond, E., Davidson, K. K., Lawrence, R. A., Sun, D., Windle, B. E., Hurley, L. H., and Von Hoff, D. D. (1999) Effects of cationic porphyrins as G-quadruplex interactive agents in human tumor cells. *Cancer Res.* 59, 639–644.
- (28) Siddiqui-Jain, A., Grand, C. L., Bearss, D. J., and Hurley, L. H. (2002) Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. U. S. A.* 99, 11593–11598.
- (29) Chen, C., Li, M., Xing, Y., Li, Y., Joedecke, C.-C., Jin, J., Yang, Z., and Liu, D. (2012) Study of pH-induced folding and unfolding kinetics of the DNA i-motif by stopped-flow circular dichroism. *Langmuir* 28, 17743–17748.
- (30) Li, W., Miyoshi, D., Nakano, S., and Sugimoto, N. (2003) Structural competition involving G-quadruplex DNA and its complement. *Biochemistry* 42, 11736–11744.
- (31) Mergny, J.-L., Li, J., Lacroix, L., Amrane, S., and Chaires, J. B. (2005) Thermal difference spectra: a specific signature for nucleic acid structures. *Nucleic Acids Res.* 33, e138–e138.
- (32) Wright, E. P., Day, H. A., Ibrahim, A. M., Kumar, J., Boswell, L. J., Huguin, C., Stevenson, C. E., Pors, K., and Waller, Z. A. E. (2016) Mitoxantrone and Analogues Bind and Stabilize i-Motif Forming DNA Sequences. *Sci. Rep.* 6, 1–7.